



Research papers

Improving the accuracy of streamflow data acquired from the acoustic tomography technique using data despiking algorithms



Amirhosein Hasanabadi ^a, Masoud Bahreiniotlagh ^{b,*}, Ebrahim Jabbari ^a, Kiyosi Kawanisi ^c, Hosein Alizadeh ^a, Yousef Olfatmiri ^a

^a Department of Civil Engineering, Iran University of Science & Technology, Tehran 13114-16846, Iran

^b Department of Water Resources Studies and Research, Water Research Institute, Tehran 1658954381, Iran

^c Graduate School of Advanced Science and Engineering, Hiroshima University, 1-4-1, Kagamiyama, Higashi Hiroshima 739-8527, Japan

ARTICLE INFO

Keywords:

Fluvial acoustic tomography
Outlier detection
Time series preprocessing
Three-dimensional Rousseeuw phase-space thresholding method
Improve streamflow accuracy

ABSTRACT

Streamflow is an important hydrological factor in water resources management. The Fluvial Acoustic Tomography (FAT) as a cutting-edge river monitoring technology measures the flow velocity using the travel time of the acoustic signals. The velocity is then multiplied by the river cross-section to obtain accurate and continuous streamflow data. Similar to all acoustic instruments, the FAT output velocity contaminated by spike data. Standard deviation (STD) as a simple method of spike removing, was used in previous FAT studies. The disadvantages of the STD method are that it not only does not replace the identified spikes with the appropriate values and therefore results in measurement discontinuity, but it also requires engineering judgment that may increase the measurement error. In this study, two despiking methods, i.e., Phase-Space Thresholding (PST) and three-dimensional Rousseeuw Phase-Space thresholding (3D-RPS) are applied to detect/modify outliers of a 6-day FAT velocity signals. The results showed that the 3D-RPS method has the best performance with a Nash-Sutcliffe efficiency coefficient (NSC) of 0.731 compared to 0.572 and 0.44 NSC values of STD and PST methods, respectively. The FAT-based streamflow data were also compared with the reference streamflow obtained from the Rating-Curve method. The results reveal that the streamflow measurement accuracy improves significantly, where the relative errors decrease from $\pm 40\%$ for the STD method to $\pm 15\%$ for the 3D-RPS approach. As a result, the 3D-RPS approach appears to be an effective method for improving the accuracy of the FAT-acquired streamflow data.

1. Introduction

1.1. The Fluvial acoustic tomography overview

Prediction and awareness of hydrological processes, particularly river behavior, are essential for successful water resource management (Abbasi et al., 2021). One of the most critical and challenging issues of river monitoring is the development of an instrument that provides real-time and continuous streamflow data (Al Sawaf et al., 2021).

FAT is an innovative technology for continuous monitoring of rivers and estuaries (Kawanisi et al., 2010) even in the presence of high suspended sediment concentration during passing floods (Al Sawaf and Kawanisi, 2020; Bahreiniotlagh et al., 2020). The backbone of the FAT concept is the “time-of-travel” principle, which is analogous to the acoustic velocity meter (AVM). However, the fundamental benefit of the

FAT is that the average cross-sectional velocity can be determined without any complicated post-processing operations (Kawanisi et al., 2012).

Fig. 1 shows the conceptual model of a pair of FAT stations deployed in a river. FAT controllers are synchronized with nanosecond accuracy by connecting to the Global Navigation Satellite System (GNSS) and transmitting sound waves at regular short intervals (e.g., 30 s). The system consists of two processors connected by a cable to the transducers immersed in water transmitting and receiving sound signals. The range-average flow velocity along the ray path is measured by determining the horizontal distance between the two transducers and recording the arrival time of the sound waves in the upstream and downstream stations. The streamflow can then be obtained by multiplying the cross-sectional river area by the FAT-measured flow velocity along the flow path.

Considering two acoustic stations distanced by a length L in a fluid

* Corresponding author.

E-mail address: m.bahreini@wri.ac.ir (M. Bahreiniotlagh).

Nomenclature

Symbol	Meaning
FAT	Fluvial Acoustic Tomography
STD	standard deviation method
PST	phase-space thresholding method
3D-RPS	three dimensional Rousseeuw phase space thresholding method
3D-PST	three dimensional phase-space thresholding method
mPST	modified phase-space thresholding method
Symbol	Meaning
RE	robust estimation method
R	Pearson's correlation Coefficient
RMSE	root mean squared error parameter
MAPE	mean absolute percentage error parameter
NSC	Nash-Sutcliffe model efficiency coefficient
MAD	median absolute deviation parameter

medium flowing at a velocity u_m , the forward and reverse travel times, t_1 and t_2 , are respectively expressed by (Kawanisi et al., 2010):

$$t_1 = \frac{L}{C_m + u_m} \quad (1)$$

$$t_2 = \frac{L}{C_m - u_m} \quad (2)$$

where C_m is the sound speed in water and u_m represents averaged water velocity along the sound path that are respectively given by:

$$c_m = \frac{L}{2} \left(\frac{1}{t_1} + \frac{1}{t_2} \right) \approx \frac{L}{t_m} \quad (3)$$

$$u_m = \frac{L}{2} \left(\frac{1}{t_1} - \frac{1}{t_2} \right) \approx \frac{L}{2} \frac{\Delta t}{t_m^2} = \frac{c_m^2}{2L} \Delta t \quad (4)$$

where $t_m = \frac{t_1+t_2}{2} \approx t_1 \approx t_2$, and $\Delta t = t_1 - t_2$.

Finally, the FAT estimates the streamflow using Eq. (5):

$$Q_{FAT} = u_m \times A \times \tan\theta \quad (5)$$

where A is the oblique cross-sectional area along the transmission line, and θ is the angle between the ray path and river flow.

It should be noted that other acoustic devices applied for velocity measurement (e.g., ADV¹, ADCP², and AVM³s) propagate sound waves in a straight line; therefore, those devices can not provide information of the average velocity of the entire cross-section of the river. Whereas, the sound waves of the FAT propagate the entire cross-section of the river, allowing the average cross-sectional flow velocity to be measured in a fraction of a second (Al Sawaf et al., 2017). Furthermore, since the propagated sound waves are reflected from the riverbed and the water surface, the geometry of the bed does not affect the measurements. Therefore, acoustic tomography measurements are independent of the riverbed geometry, and the average cross-sectional velocity can be measured directly without the need for postprocessing (Bahreinimotlagh et al., 2016).

1.2. Spikes in FAT velocity data

The presence of spike data is a problem in all acoustic measurement

systems, decreasing the quality and precision of their datasets. A frequently used definition for the concept spike is expressed as: "An observation which deviates significantly from other observations that it raises the possibility that it was created by a different mechanism" (Hawkins, 1980). Therefore, the data must be cleaned before performing data mining processes (i.e., raw data preprocessing and preparation).

The primary goal of the despiking process is to identify spike data as observations that do not fall within the common behavior region defined by the despiking algorithm. However, it is challenging to pick a zone encompassing all normal data behaviors. One of the primary challenges of such a process is determining the exact boundaries of this area. Moreover, the precise meaning of spike data can differ across fields and disciplines. Minor changes in values are observed in medical science datasets, for example, maybe significant, whereas, the same amount of change in economic or engineering observations may be negligible. As a result, the varying concepts of spike data in different disciplines make it complicated to apply methodologies created for one field to another.

Flow velocity measured by FAT, similar to other acoustic measurement systems such as ADV, ADCP, and AVM, contains spike data. STD was the simply method to despike FAT data. The STD method considers a normal behavior region that is located around the primary trend of the signal by a certain distance. This distance is determined as a specific coefficient of the signal standard deviation value. Data points outside the normal region are detected as spikes (Bahreinimotlagh et al., 2016).

The most crucial drawback of the STD method is the ignorance of the replacement of the detected spikes data with appropriate values, which results in a discontinuous measurement. Moreover, the STD is not an intelligent method, and its parameters are determined by user judgment.

Goring and Nikora (2002) developed the Phase-Space Thresholding (PST) despiking method to show its superiority with compared to previous despiking techniques (e.g., RC filters and Tukey 53H methods (Otnes and Enochson, 1978), acceleration thresholding method (Nikora and Goring, 2000), and wavelet thresholding method (Donoho and Johnstone, 1994)).

Wahl (2003) proposed a modified version of PST called the three-dimensional phase-space thresholding method (3D-PST) to improve the performance of the previous model using a three-dimensional space. Mori et al (2007) implemented Wahl's modified 3D-PST method on an ADV contaminated dataset and compared it to the original PST method. Their findings revealed that the 3D-PST method outperformed the PST method.

Zhong et al (2020) developed and implemented the three-dimensional Rousseeuw phase-space thresholding method (3D-RPS) by optimizing the PST method and combining it with the RE method through a two-step filtering process, improving the despiking procedure and reducing the iterations needed to implement the PST method. The results showed that 3D-RPS is a promising method that effectively improves the accuracy of outlier detection and considerably reduces the over-processing phenomenon.

In this study, PST and 3D-RPS despiking methods, which were previously applied to ADV devices, are applied to the FAT data. The results are compared with the output of the STD method, which was used in previous FAT studies. Finally, the modified FAT-based streamflow data using these three approaches are compared to the reference streamflow obtained from the Rating-Curve method, and the accuracy improvement in the FAT measurement is assessed.

2. Materials and methods

2.1. Standard deviation method

The standard deviation method is one of the most straightforward procedures for removing spikes. This method was used to detect and eliminate apparent spikes in the FAT (Bahreinimotlagh et al., 2016). In this method, data points within a certain distance from the signal's

¹ Acoustic Doppler Velocimeter

² Acoustic Doppler Current Profiler

³ Acoustic Velocity Meter

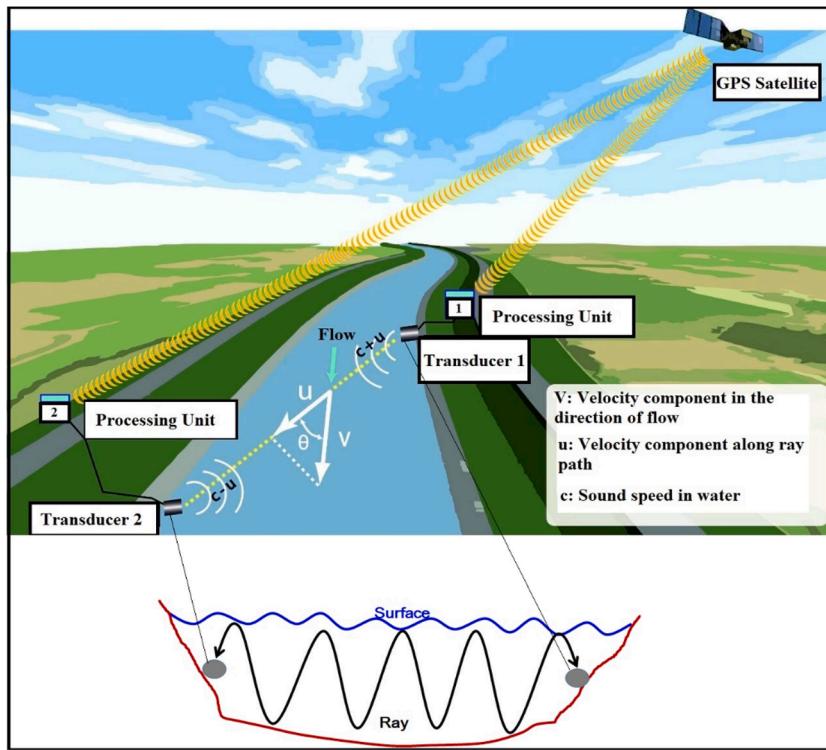


Fig. 1. Schematic representation of FATS and its operation in a river (Bahreinimotagh et al., 2019).

primary trend are identified as usual and reliable data, whereas, points outside this range are identified as spikes and deleted. This distance is arbitrarily chosen by the operator based on the type of signal (engineering judgment), which can be considered as a disadvantage.

2.2. Phase-Space thresholding (PST) method

The PST method consists of three concepts:

- Differentiating a signal enhances its high-frequency portion and displays some additional features of the signal that are not otherwise visible, and helps detect the spikes better.
- The expected absolute maximum value of a random time-series is calculated using the Universal threshold value.
- Valid good data points form a cluster in a dense cloud of phase space.

These concepts are used to construct three threshold ellipsoids in three phase-spaces of $u-\Delta u$, $u-\Delta^2 u$, and $\Delta u-\Delta^2 u$; afterwards, the spikes are detected based on recognition of data points that lie outside of these ellipsoids. The advantage of this method over many other developed despiking methods such as RC filters, the Tukey 53H method (Otnes and Enochson, 1978) and the Acceleration thresholding method (Nikora and Goring, 2000) is that it does not use any parameter that needs to be specified by the operator experimentally and approximately.

The PST technique is an iterative approach in which spikes are iteratively recognized and replaced with more reliable and reasonable values. This process continues until the number of valid points becomes constant inside the threshold ellipses in all three phase spaces between consecutive iterations.

The despiking process with the PST method is applied in such a manner that after removing the mean of the signal, the first and second derivatives of the signal are calculated. The three-phase spaces of $u-\Delta u$, $u-\Delta^2 u$, and $\Delta u-\Delta^2 u$ are drawn, and the parameters required to determine the dimensions of the three threshold ellipses, which are the universal threshold and the standard deviation of the original signal and its derivatives, are calculated. Finally, data points outside these ellipses in

each of the three-phase spaces are detected as spikes.

PST method can be stepped up in the following way:

1. Calculation of the first and second derivatives of the input velocity signal:

$$\Delta u_i = \frac{(u_{i+1} - u_{i-1})}{2} \quad (6)$$

$$\Delta^2 u_i = \frac{(\Delta u_{i+1} - \Delta u_{i-1})}{2} \quad (7)$$

2. Calculation of standard deviation and universal threshold parameters: for each u , Δu , $\Delta^2 u$ variables, the standard deviation is calculated, and the expected maximum value for a random and normal variable, whose standard deviation is estimated by σ and whose mean is zero, is computed using the universal threshold parameter:

$$\lambda_U = \sqrt{2 \ln(n)} \quad (8)$$

where n is the total number of signal data points.

3. Calculation of the rotation angle of the threshold ellipsoid principal axis in the $u-\Delta^2 u$ space:

$$\theta = \tan^{-1} \left(\frac{\sum (u_i \Delta^2 u_i)}{\sum (u_i)^2} \right) \quad (9)$$

4. Dimension calculation of large and small principal axes of threshold ellipsoids: for the phase space of $u-\Delta u$, the sizes of the large and small axes of the ellipsoid are $\lambda_U \sigma_U$ and $\lambda_U \sigma_{\Delta u}$, respectively; for the phase space of $\Delta u-\Delta^2 u$, the sizes of the large and small axes of the ellipsoid are $\lambda_U \sigma_{\Delta u}$ and $\lambda_U \sigma_{\Delta^2 u}$, respectively; moreover, for the phase space of $u-\Delta^2 u$, the sizes of the large and small axes of the ellipsoid are

obtained by solving the following equations simultaneously, in which a is the large axis and b is the small axis of ellipsoid:

$$(\lambda_u \sigma_u)^2 = a^2 \cos^2(\theta) + b^2 \sin^2(\theta) \quad (10)$$

$$(\lambda_u \sigma_{\Delta^2 u})^2 = a^2 \sin^2(\theta) + b^2 \cos^2(\theta) \quad (11)$$

2.3. Three dimensional Rousseeuw phase-space (3D-RPS) method

The 3D-RPS method, proposed by Zhong et al. (2020), combines two different algorithms: 3D-PST of Wahl (2003), and a combination of the RE method of Rousseeuw and Leroy (1998), and the mobile window method of Zou et al. (2008). These methods are applied separately to the input signal and sum of the points identified as outliers in each of these two steps are considered as the final result of the despiking process. The detected spikes are then replaced with a more appropriate value.

This algorithm attempts to use a three-dimensional threshold ellipsoid instead of using the projection of three-dimensional threshold ellipsoids on three planes: $u \cdot \Delta u$, $u \cdot \Delta^2 u$, and $\Delta u \cdot \Delta^2 u$ to identify spikes.

The other primary part of the 3D-RPS method includes using robust statistical parameters to develop this algorithm, in a way that instead of using the mean parameter, the median is used as a robust location estimator. As well, instead of using the standard deviation parameter, the median absolute deviation (MAD) is applied as a scale estimator. The use of these robust statistical parameters makes this method more efficient than the PST method. This means that the spike values do not affect the estimators, while these effects on the mean and standard deviation parameters are high. In addition, instead of using the universal threshold value as the boundary separating the normal and outlier regions, the Chauvenet's criterion is applied (Coleman & Steele 1999).

The three-dimensional threshold is used because if the two-dimensional threshold in each principal plane of $u \cdot \Delta u$, $u \cdot \Delta^2 u$, and $\Delta u \cdot \Delta^2 u$, were to be used, the process of despiking would require three comparisons to be made in every plane, and spikes would ultimately be detected if they were located outside of a complex 3D shape analogous to an ellipsoid, but with knobby projections in each of the eight phase-space quadrants. The knobs allow some data points outside the 3D ellipsoid to be retained as normal data. To prevent this error, a single-comparison test using the 3D ellipsoid shape, after removing the median of the signal and calculating the first and second derivatives of the signal similar to the PST method, is implemented. And the ellipsoid is defined in spherical coordinates of ρ , ϕ and θ as:

$$\frac{1}{\rho^2} = \frac{(\sin\varphi \times \cos\theta \times \cos\alpha + \cos\varphi \times \sin\alpha)^2}{a^2} + \frac{(\sin\varphi \times \cos\theta \times \sin\alpha - \cos\varphi \times \cos\alpha)^2}{b^2} + \frac{(\sin\varphi \times \sin\theta)^2}{c^2} \quad (12)$$

where, a and b are the large and small principal axes of the threshold ellipsoid, which are projected on the $u \cdot \Delta^2 u$ plane, respectively, and c is the large principal axis of the threshold ellipsoid that is projected on the $\Delta u \cdot \Delta^2 u$ plane. α shows the rotation angle of the ellipsoid in the $u \cdot \Delta^2 u$ plane, which is calculated using Eq. (4).

The process of identifying data points outside this 3-D threshold ellipsoid is performed in such a way that at first, the position of the data points is calculated in three-dimensional space of $u \cdot \Delta u \cdot \Delta^2 u$ in spherical coordinates as distance from the center of coordinates (ρ), and two angles of ϕ and θ , which are the angles between the line connecting the data point to the center of coordinates and the principal axes of $\Delta^2 u$ and u , respectively and then the value of ρ is compared to value of ρ_e , that ρ_e is the distance of a point to center of coordinates, which is at the surface of 3D ellipsoid and has the same angles of φ and θ with the data point that intends to specify its type. Finally, if ρ is greater than ρ_e , it indicates

the presence of a data point inside the threshold, and if ρ is smaller than ρ_e , it shows that the data point is outside the threshold and is an outlier.

To implement the second step of this method, the RE algorithm combines with the mobile window method in such a manner that at first, all data points are separated as different groups with a fixed-width mobile window moving across the signal, then with the help of Median parameter, location is estimated using Eq. (13), and with the help of MAD parameter, the scale estimated using Eq. (14), and with the use of these two parameters data points in each window are standardized using Eq. (15):

$$M = \text{Median}_{i=1,\dots,n}(u_i) \quad (13)$$

$$S = e \times (\text{Median}_{i=1,\dots,n}|u_i - M|) = e^* \text{MAD} \quad (14)$$

$$Z_i = \frac{(u_i - M)}{S} \quad (15)$$

$$O_i = |Z_i| - c \quad (16)$$

where, n is the number of signal data points, u_i is the velocity value of each data point, e is an estimated coefficient with the value of 1.483, Z_i is the standardized value of data points, O_i in Eq. (16) is the parameter that determines whether the data point is an outlier or normal and c is the threshold value between normal and outlier region. The data points that have an O_i greater than zero are deemed as spikes.

To detect spikes, a threshold value is set on standardized data points, and data points that lie outside this threshold are deemed as outlier data. In this study, Chauvenet's criterion is used as the threshold. Chauvenet's criterion value depends on the window width and the number of data points located in every window. To find the optimal value for this parameter, different values for the width of windows and the best option that detects most of the spikes and flattens the input signal optimally need to be assumed and chosen.

The combination of these two steps in the 3D-RPS method increases the accuracy of the despiking process. Consequently, due to the non-iterative process of this method, the occurrence of an overprocessing phenomenon is prevented. Flowcharts of the PST and 3DRPS algorithms are shown in Fig. 2 to demonstrate the despiking-replacement steps.

2.4. Performance evaluation of despiking methods

To evaluate the performance of the applied methods, appropriate statistical parameters such as Pearson Correlation Coefficient (R), Root

Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and Nash-Sutcliffe model Efficiency coefficient (NSC) are been used. The streamflow data are compared with the ones obtained from the Rating-Curve method as the reference.

In addition, the parameter relative discharge differences have been used to compare the output of the despiking approaches and the Rating-Curve method, using the following equation:

$$\text{RelativeDischargeDifference}_{(D,D_{\text{reference}})} = \frac{D - D_{\text{reference}}}{D_{\text{reference}}} \quad (17)$$

2.5. Study area and FAT data

The FAT observation is carried out in the Gono River located in Miyoshi City, Hiroshima Prefecture, Japan. The length of the Gono River

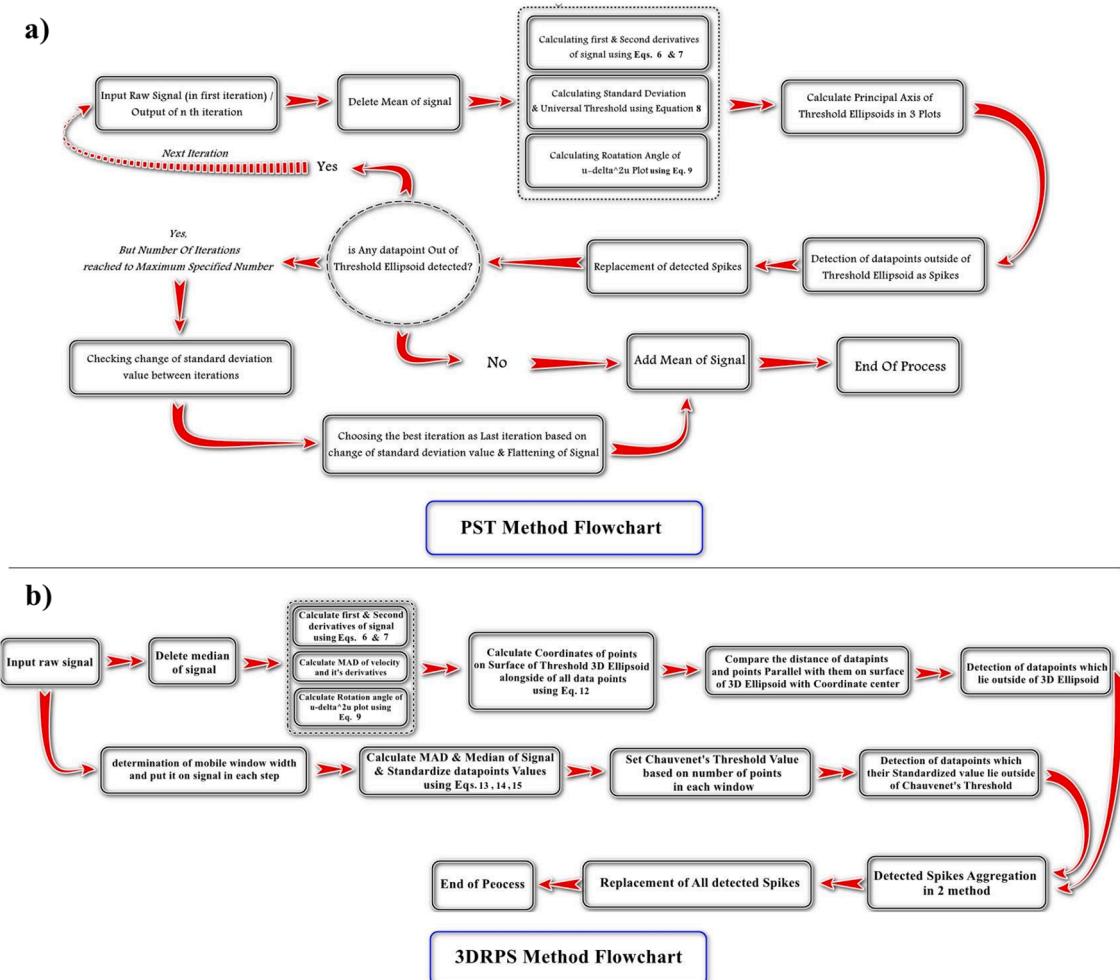


Fig. 2. Flowcharts of a) PST and b) 3DRPS methods.



Fig. 3. location of the site and the position of the downstream (T2) and upstream (T1) stations.

is approximately 194 km. The Saiju and Basen rivers are the two main tributaries of the Gono River, which meet about 3 km from the data collection site. The slope of the river bed in the selected reach region was measured 0.11%, and the river bed is composed of sand and boulder. The average annual flow rate is approximately 73 cubic meters per second (Kawanisi et al., 2016).

A 115-meter-wide reach was selected for the location of the transducers. To collect data at this site, two 30-kHz FATS were diagonally deployed on both sides of the river, where the horizontal distance between them was set 294.6 m. The data was collected every 30 s for six days from February 1, to February 6, 2015. The total amount of data is 10,600. Fig. 3 shows the data collection area and location of upstream and downstream stations.

Fig. 4 shows the raw flow velocity data measured by the FATS. There are numerous spikes that cause significant error in velocity and streamflow calculations.

3. Results and discussion

3.1. Standard deviation method results

In this study, data in the range of 0.05 to 1.8 m/s were considered as normal data, and points outside this range were assumed to be a spike. Fig. 5-a shows the data points identified as outliers by the STD method. Fig. 5-b also shows a comparison of the unspiked signal with the STD method output. Out of 10,600 data, 156 data (1.47% of the total) were identified as outliers. Although most of the obvious spikes are detected by this method, it should be noted that using user judgment in flood events, when the flow velocity increases significantly or even in estuaries when the tides occur (tidal velocities are shown by negative values), increases the error. Furthermore, the existence of significant fluctuations in the processed signal implies that the STD does not perform well in despiking acoustic tomographic output.

3.2. PST method results

Fig. 6 shows the three-phase spaces for the first and fourth iterations of the PST method. The values of the data points and their position relative to the threshold ellipses are specified as well. The data points with scattered values and farther away from other points, which have formed a dense cluster, are removed during these iterations. The reduction of the dimensions of the ellipses in these four iterations shows the proper identification of outlier points with this method.

It should be pointed out that the completion criterion of the PST iterations is that no outliers beyond these ellipses must be detected in the last iteration. However, in this study, this criterion was not met for the FATS signal. In addition to the weaknesses and limitations of the PST method in despiking signals with a high percentage of outliers, as previously stated in (Islam and Zhu, 2013), due to the dense spikes in the signal, causing improper performance of the replacement method.

Finally, to solve this problem, by examining the changes in the standard deviation value of the signal fluctuations around its average during different iterations of this method, in the fourth to fifth iteration, this change was tiny (on the scale of less than one-thousandth). Thus, the

overall dispersion of signal points relative to the mean value, in these consecutive iterations did not change much due to the proper flattening of the signal after the fourth iteration. This iteration was considered the last repetition. It should be noted that the standard deviation value of the data fluctuations around the average signal value before the despiking process was 0.201 and after the fourth iteration reached 0.0866.

Fig. 7a depicts the total number of points identified as outliers in four iterations of the PST approach. Fig. 7b shows a comparison of the input signal to the PST algorithm output. Out of 10,600 data, 3475 data (32.78% of the total data) were identified and replaced as outliers.

3.3. 3DRPS method results

This method involves two steps: In the first step, which is the 3DPST algorithm, data points were shown in three-dimensional space with a threshold ellipse after eliminating the median value from the signal and calculating the required parameters of Eq. (7). The data points outside the three-dimensional threshold ellipsoid were identified by comparing the values of the ρ and ρ' Parameters. Out of 10,600 data, 2220 (20.94% of the total data) were identified as outliers. Fig. 8A shows the position of the data points outside the three-dimensional threshold ellipsoid.

To apply the RE method in the second step, the data points must first be standardized using Eq. (10). The mobile window method is then used to consider a window with a specific width, which depends on the optimum Chauvenet's criterion for the input signal. It depends on the number of data points in each window and is used as the threshold value. The spike data points would be identified if their standardized value be outside this threshold.

The signals are divided into 21 groups of 500 data, and the threshold value, which depends on the selected window width, is applied. Chauvenet's criterion used to identify outliers, is 3.29 as a threshold value for 500 data points. The state of the data after the standardization and application of Chauvenet's criterion is shown in Fig. 8B. As shown, this parameter does an excellent job of distinguishing between good and bad data.

The despiked output of this method was obtained after aggregating the spikes identified in both steps of the method and replacing these points with the average value of two data points on each side of them. The signal standard deviation before despiking was 0.201 and decreased to 0.094 after the process, indicating a reduction in data dispersion around the signal average.

Fig. 9a shows the total points identified as outliers in the 3DRPS method. Fig. 9b represents the output of the 3DRPS method and the raw signal before despiking the output signal of the 3DPST method and the output signal of the RE method. Out of 10,600 data, 1726 data (16.28% of the total data) were identified and replaced as outliers by the 3DRPS method. The RE method alone outperformed the 3DPST method in most areas, and their combination and aggregation produced a better result than each of them individually did.

3.4. Streamflow results

In this section, the accuracy of different despiking methods is evaluated. These results are compared with the streamflow of the Rating

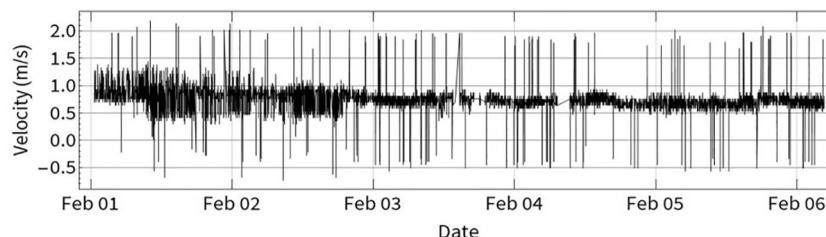


Fig. 4. Raw flow velocity signal in the direction of transmitted FATS Waves.

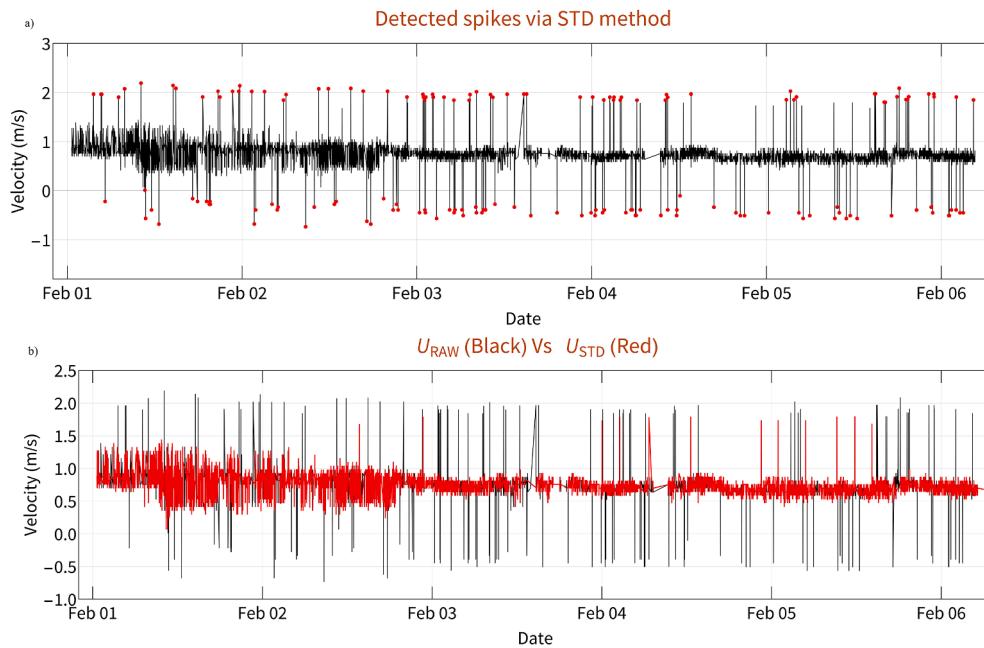


Fig. 5. a) Detected Spikes in STD method (Red Points) and their positions in the input signal (Black)- b) Comparison between Raw signal before and after Despiking with STD method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

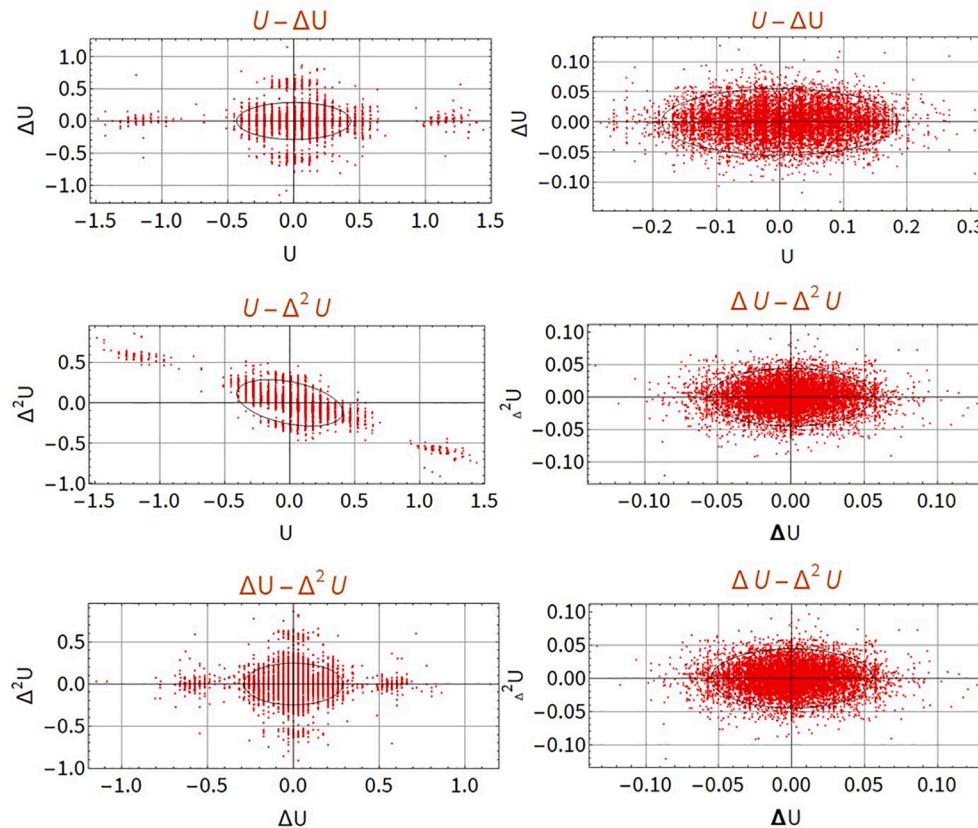


Fig. 6. Comparison of data points position on 2D PST Plots relative to threshold ellipses in 1st (left side) and 4th (right side) iteration of PST method.

Curve method as the reference. As shown in Fig. 10a, the inadequate performance of the STD method in these comparisons and the existence of sudden severe fluctuations at several points of the output signal of this method are evident.

Interestingly, the result of the STD method is more consistent with

the output of the Rating-curve than the PST method in some regions (e.g., February 5th). The reason for this error indicates the improper performance of the PST method in signals with a high percentage of the presence of outliers. As a result, the PST method is unsuitable for the despiking of tomographic data.

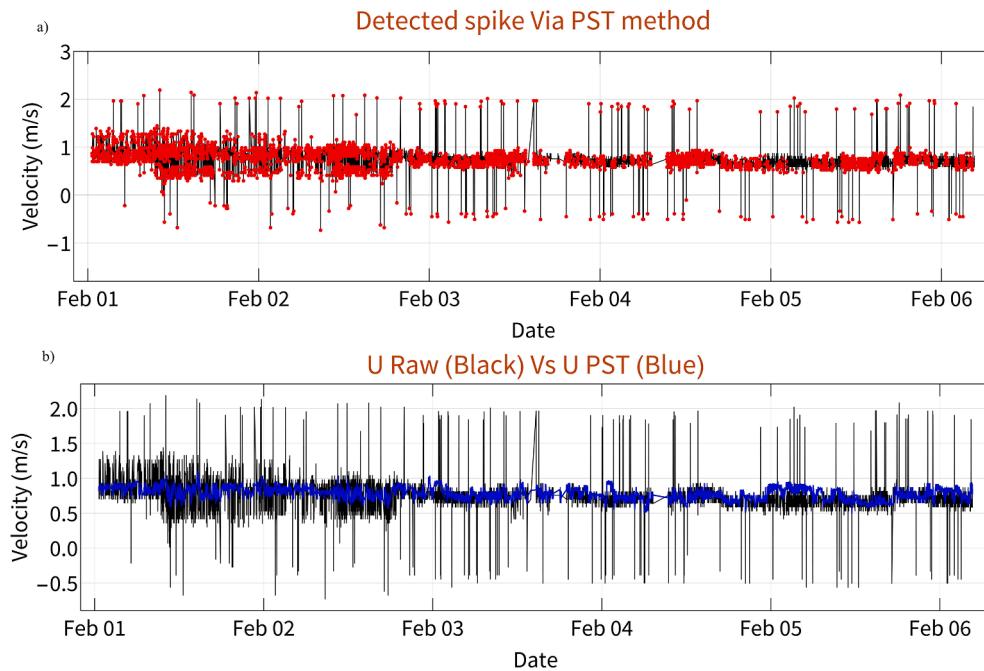


Fig. 7. a) Detected spikes by PST method (red points) and their positions in the input signal (black), b) Comparison between raw signal before and after despiking with PST method. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

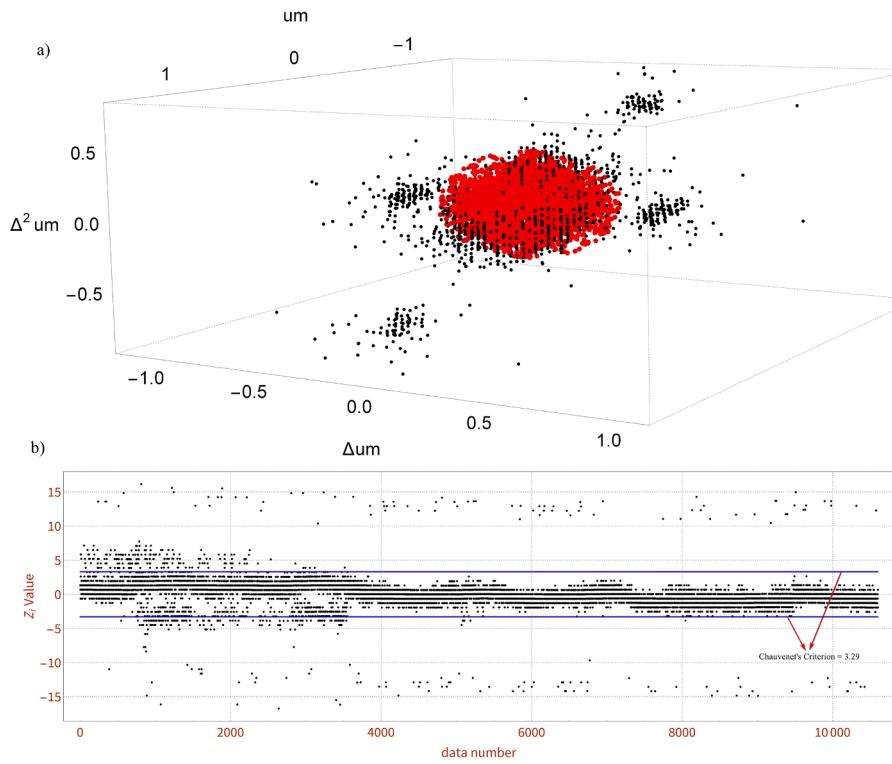


Fig. 8. a) Detected spikes in 3D Space and their position relative to the 3D Ellipsoid: black points are spikes out of threshold and red points are points on the surface of the 3D Ellipsoid, b) Data points standardized Z_i values and implemented Chauvenet's Criterion as cut-off threshold. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The comparison of the 3DRPS method with PST and STD methods reveals a better performance of 3DRPS. Fig. 10b illustrates the output fluctuations of the 3DRPS method around the RC reference method which are significantly reduced compared to the PST and STD methods. It is especially visible between 3rd and 6th February. Consequently, the output of this method is flatter than the ones of the other two methods

and involves fewer sudden fluctuations.

3.5. Uncertainty analysis

Fig. 11 shows the relative difference of streamflow between the three proposed methods and the Rating-Curve data as the reference. Fig. 11a

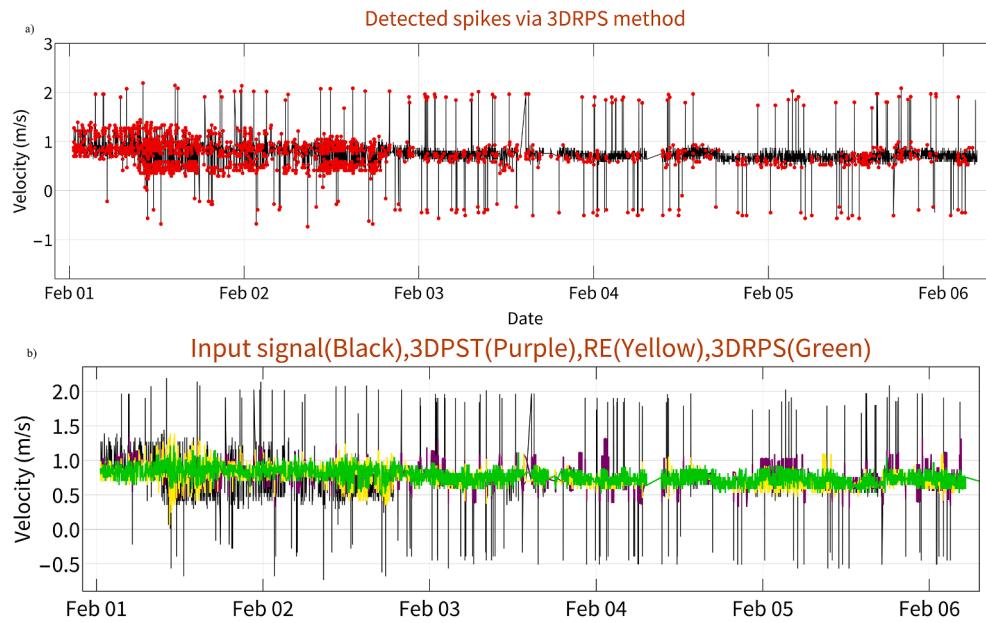


Fig. 9. a) Detected spikes in 3DRPS method (red points) and their positions in the input signal (black), b) Comparison of raw signal vs results of: 3DPST method (purple) vs RE method (yellow) vs 3DRPS method (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

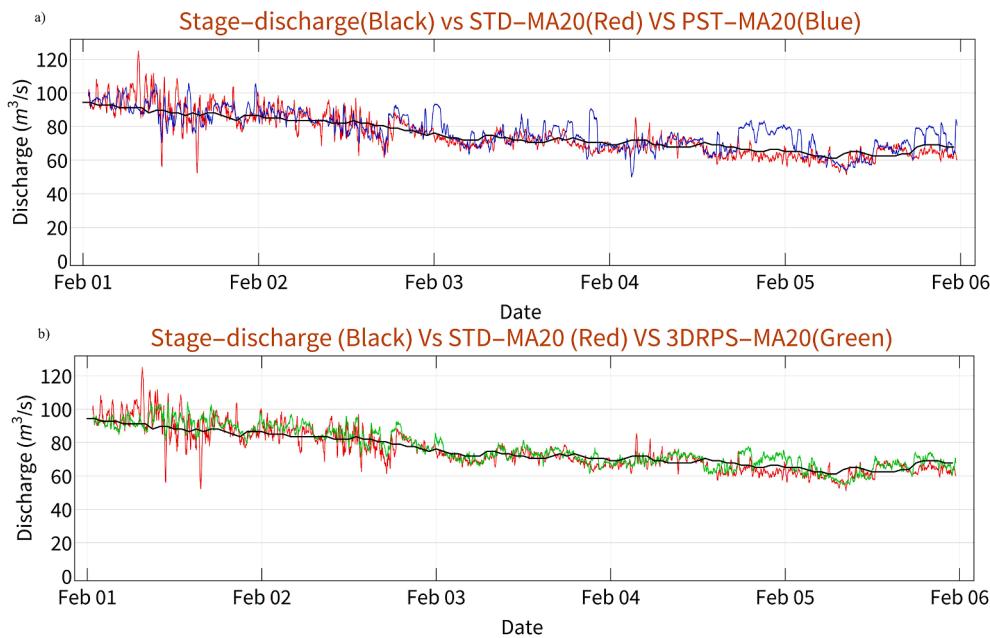


Fig. 10. Comparison between rating-curve reference streamflow (black), STD method (red), PST method (blue) and b) 3DRPS method (green). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

shows the performance of the STD method. The relative discharge difference of most data points is in the range of $\pm 15\%$. However, for some outliers that STD is unable to identify, the relative discharge difference exceeds up to 40%.

Fig. 11b illustrates the relative discharge difference of the PST method. It is clear that the majority of the points have a relative discharge difference in the range of -10 to $+20\%$, with some points exceeding this range by even more than 20%. Furthermore, when comparing this figure to the previous one, it is observed that, although the output of the STD method is more relative in most points than the one in PST method, the PST performed better in signal flattening.

Fig. 11c represents the relative discharge difference of the 3DRPS

method. Most of the points are in the range of $\pm 10\%$, and a few points are also seen outside this range and in the range of 15 to 18%. By comparing these results, the functional superiority of the 3DRPS method compared to other methods in comparison with the RC data can be concluded.

The correlation diagram of the despiking methods versus the Rating-Curve data is shown in Fig. 12. The reduction in data correlation with the RC method in the case of the PST method compared to undespike data is an exciting point that acknowledges the limitation of the PST method in the case of highly contaminated signals. Finally, this parameter demonstrates the superiority of the 3DRPS method in comparison with other methods.

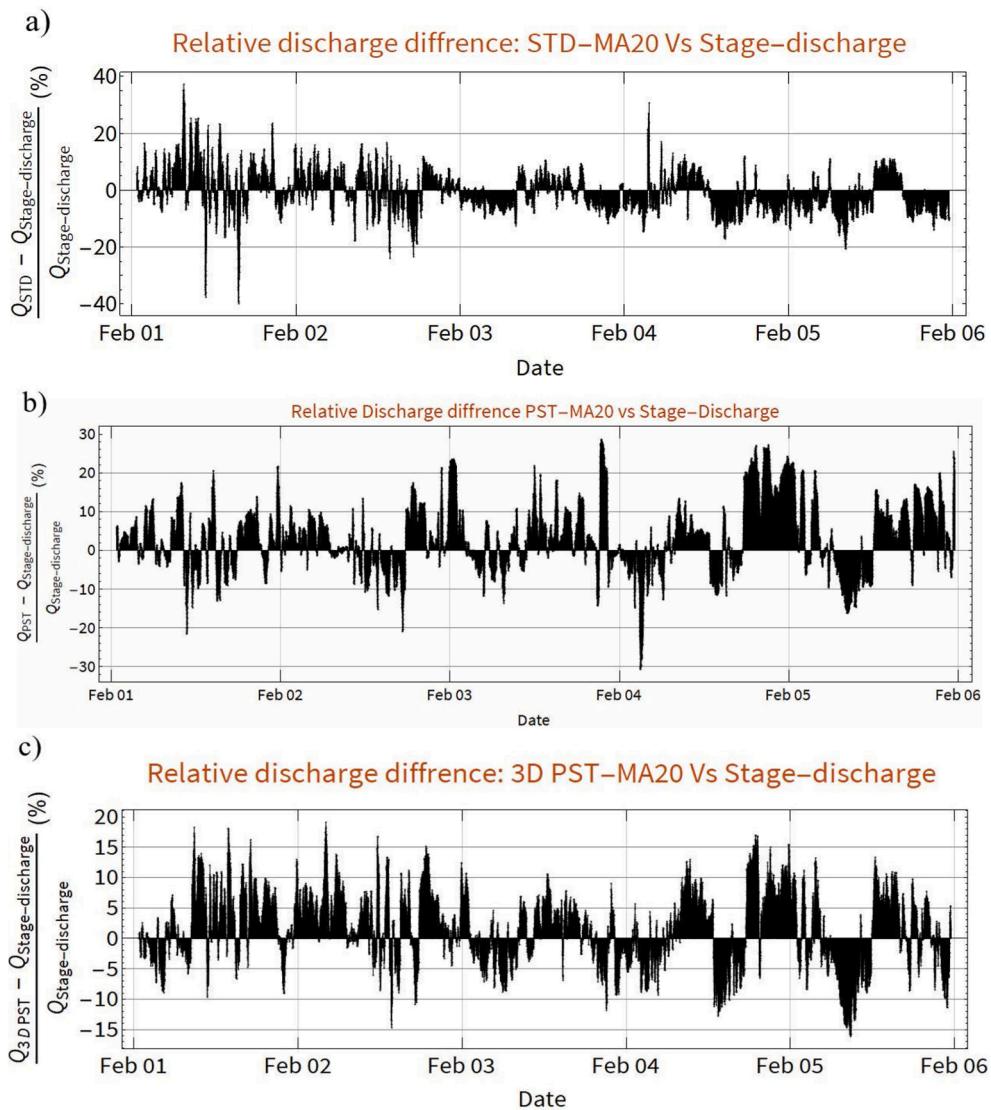


Fig. 11. Relative discharge difference between Rating-Curve method data (as reference), and: a) STD, b) PST, and 3) 3DRPS methods.

Table 2 summarizes the performance of various despiking methods that are statistically compared with the Rating-Curve data. The R^2 coefficients were 0.883, 0.798, and 0.913 for the STD, PST, and 3DRPS methods, respectively. The RMSE demonstrates the superiority of the 3DRPS method over the other existing methods. The STD method has a value of $5.899 \text{ m}^3/\text{s}$, which reduces to $4.676 \text{ m}^3/\text{s}$ for the 3DRPS technique. Furthermore, the PST technique with the lowest performance, indicates that the PST technique has a functional constraint for severely contaminated signals.

The functional superiority of the 3DRPS method was also confirmed when the MAPE decreased from 5.838 % for the STD method to 5.123 % for 3DRPS. This parameter also indicates the poor performance of the PST method, as the MAPE of PST is even higher than the MAPE of the STD method, and it comes to mind that considering the level of the signal, flattening with the help of the despiking techniques alone cannot help to evaluate their accuracy and performance.

Furthermore, by examining the NSC coefficient, which is used to evaluate the predictive ability of hydrological models, and the closer the value of this parameter to 1, the better the performance of that method. The 3DRPS method has the best performance among the applied methods, and the NSC value has increased from 0.572 for the STD to 0.731 for 3DRPS. For this parameter, the PST method is still recognized as the worst method in terms of performance.

The computational time effort of a computer system to process the algorithms and assess the computational cost of each of them is another metric that may be used to compare the performance of algorithms. The computational times of the PST and 3DRPS algorithms were examined using the Wolfram MATHEMATICA version 12 software in a computer system with the following specifications: Processor: Intel(R) Core(TM) i7-4510U CPU @ 2.00 GHz 2.60 GHz, RAM: 8.00 GB DDR3 (7.90 GB useable), system Type: 64-bit operating system, x64-based processor. The computational time effort was 4 min and 2 s and 1 min and 54 s for the PST and 3DRPS methods, respectively.

As a result, the 3DRPS approach is more efficient than the PST algorithm, as evidenced by the results. This may have been inferred even before the despiking process, because the PST method is iterative. The computational cost of this algorithm appears to be higher than the 3DRPS method due to the repetition of calculations in successive iterations.

4. Discussion

The signal to noise ratio (SNR), which is defined as 10 times the logarithm of the peak amplitude of received signals, is one of the important factors in FAT and is used as a determinant component between clean and noisy data, where, data with SNRs less than 10 dB are

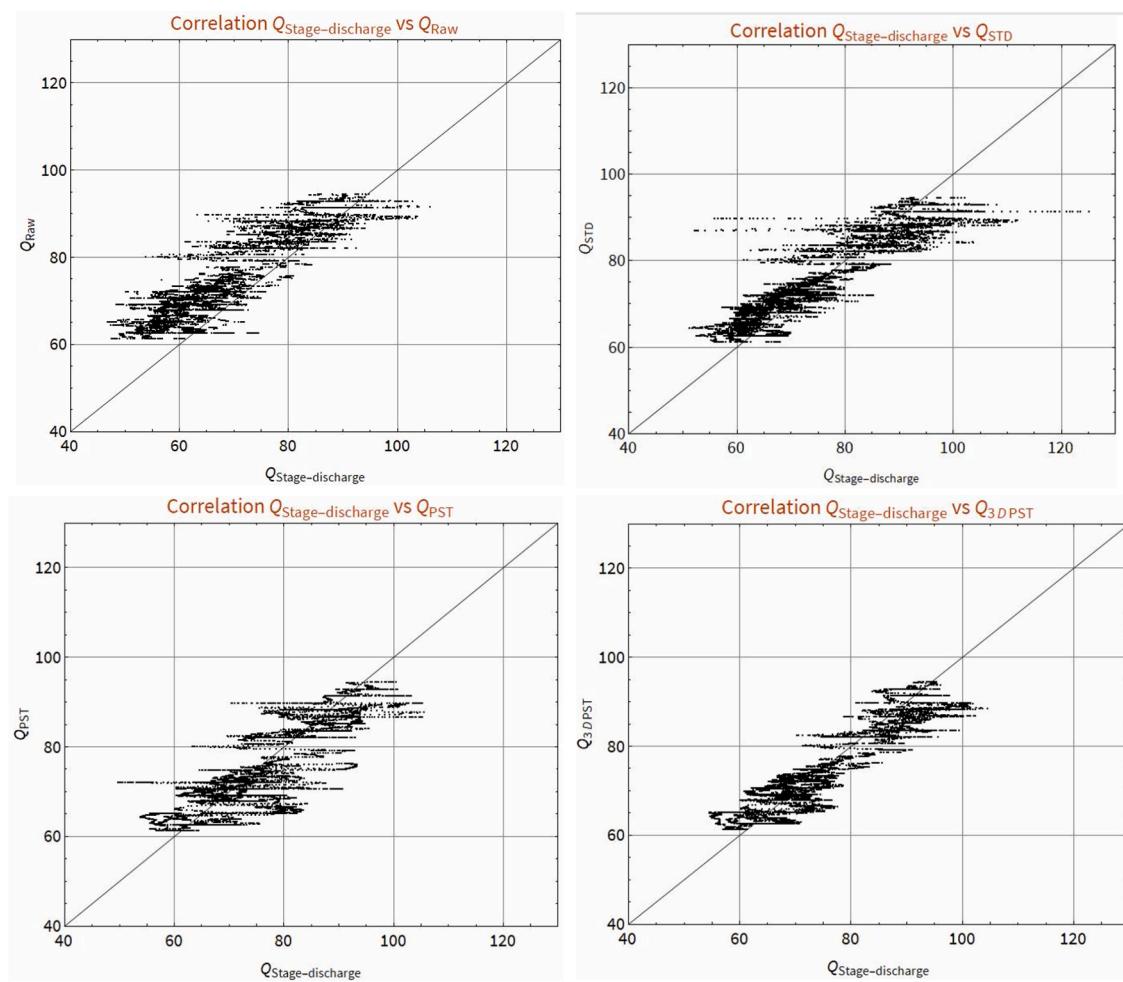


Fig. 12. Correlation Plots of stage-discharge method vs raw, STD, PST, 3DRPS methods.

Table 2
Statistical Performance of Despiking Methods (Rating-Curve Method as Reference).

	RMSE (m^3/s)	MAPE (%)	NSC	R^2
STD	5.899	5.838	0.572	0.883
PST	6.753	7.142	0.440	0.784
3DRPS	4.676	5.123	0.731	0.913

removed as noisy data. Figs. 13 and 14 show the SNR values for all detected spikes using the PST and 3DRPS methods for upstream (T1) and downstream (T2) stations (red points), respectively, and compared with ones for all received signals (black points). It was observed that the detected spikes in both PST and 3DRPS methods have acceptable and sufficient SNR levels to be regarded as valid in terms of the clean pulse. As a result, there is not any relationship between the SNR value of data and the outliers.

It is necessary to mention that in the PST method, no parameters must be quantified by the analyst, and the process of thresholding definition for all phase spaces is completed by combining the standard deviation of the velocity signal and its derivatives with universal threshold parameter that requires no tuning and only depends on the number of data points, which is the fundamental advantage of this algorithm over others.

For the 3DRPS method, however, the operator should decide on the value of the Chauvenet's parameter, which is the "c" parameter in Eq. (16), as the threshold for standardized values. Because the value of

Chauvenet's criterion depends on the number of data points in each window, this parameter was quantified using standardized data values that helped in selecting the optimal number of data points in each window which corresponded to an optimum Chauvenet value that appropriately separated two areas of normal and spike data. As shown in Fig. 8, the optimum value of the Chauvenet's threshold that can separate the normal and spike region is approximately 3. By selecting 500 data for each window, which is equivalent to the value of 3.29 for Chauvenet's criterion, the normal and spike regions are separated optimally and data points lie outside the normal region considered as spikes are detected by the RE method. Furthermore, another coefficient used in the 3DRPS method (i.e., "e" in Eq. (14)) is quantified as 1.483 due to the explanation of (Wahl, 2003), which stated that the 1.483 coefficient is a factor that makes the scale estimator (i.e., "S" in Eq. (14)) analogous to the standard deviation, the usual scale parameter of a normal distribution.

5. Conclusion

For hydrological studies as well as water resources management, it is vital to understand streamflow properties under normal and extreme climatic conditions over the short-term (few hours) and long-term (several days to several years). Acoustic Tomography technique provides such streamflow data. However, the presence of spikes is an issue that affects the quality and validity of streamflow data. A despiking process is needed to purify the FAT output signal. In this study, two despiking methods of phase-space thresholding (PST) and three-dimensional Rousseeuw Phase-Space thresholding (3D-RPS) are

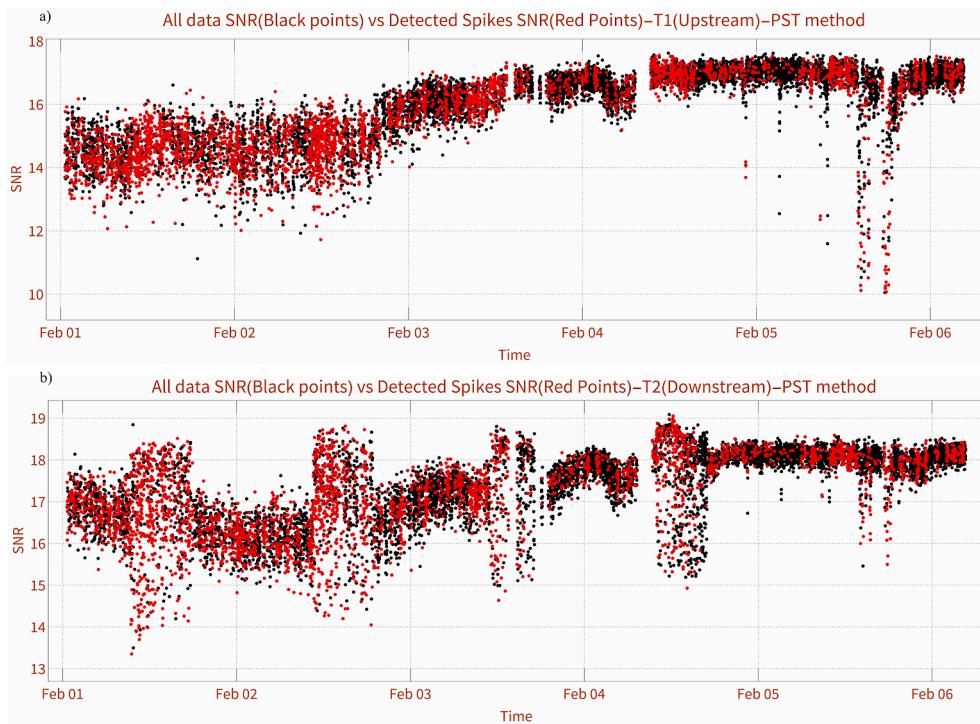


Fig. 13. SNR Plots of all data (black points) and detected spikes (red points) by PST method in: a) station T1 b) station T2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

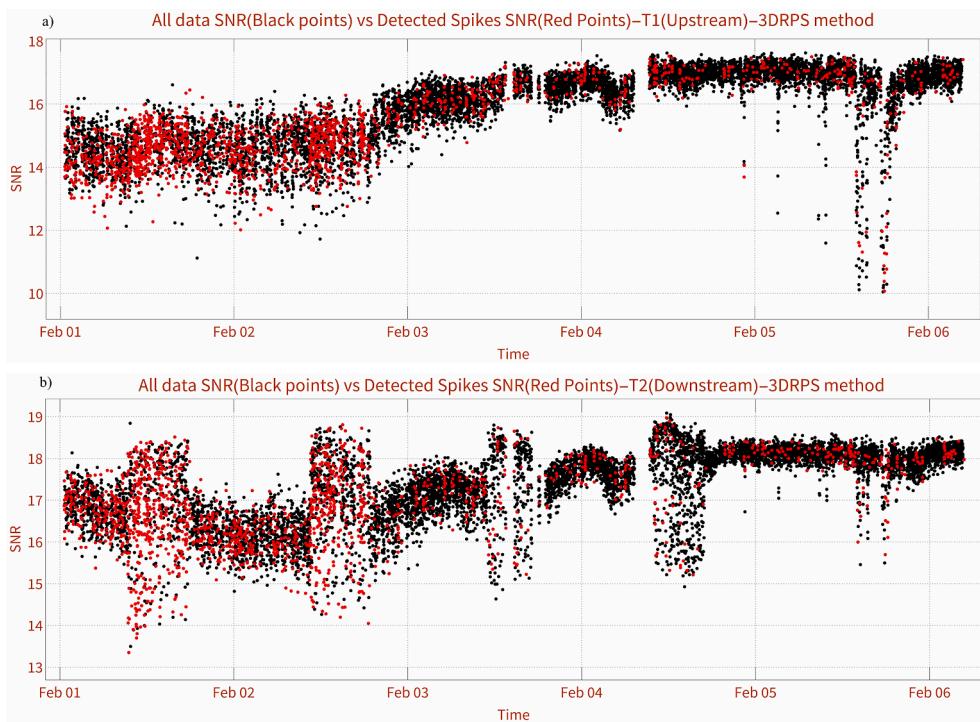


Fig. 14. SNR Plots of all data (black points) and detected spikes (red points) by 3DRPS method in: a) station T1 b) station T2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

applied to detect FAT outliers. In addition, the results of these methods were compared with the results of the simple standard deviation method, which was previously used to detect the spikes of FAT data.

In conclusion, the 3D-RPS method has the best performance, with a Nash-Sutcliffe efficiency coefficient of 0.731 compared to 0.572 and 0.44 for the STD and PST methods, respectively. The outcomes

demonstrate that streamflow measurement accuracy is also significantly improved after the despiking process, with the relative errors with Rating-Curve data as a reference, decreasing from $\pm 40\%$ for the STD approach to $\pm 15\%$ for the 3DRPS approach.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Abbasi, M., Farokhnia, A., Bahreiniotlagh, M., Roozbahani, R., 2021. A hybrid of Random Forest and Deep Auto-Encoder with support vector regression methods for accuracy improvement and uncertainty reduction of long-term streamflow prediction. *J. Hydrol.* 597, 125717. <https://doi.org/10.1016/j.jhydrol.2020.125717>.
- Al Sawaf, M.B., Kawanisi, K., Jilati, M.N., Xiao, C., Bahreiniotlagh, M., 2021. Extent of detection of hidden relationships among different hydrological variables during floods using data-driven models. *Environ. Monit. Assess.* 193 (11) <https://doi.org/10.1007/s10661-021-09499-9>.
- Al Sawaf, M.B., Kawanisi, K., 2020. Assessment of mountain river streamflow patterns and flood events using information and complexity measures. *J. Hydrol.* 590, 125508. <https://doi.org/10.1016/j.jhydrol.2020.125508>.
- Al Sawaf, M.B., Kawanisi, K., Kagami, J., Bahreiniotlagh, M., Danial, M.M., 2017. Scaling characteristics of mountainous river flow fluctuations determined using a shallow-water acoustic tomography system. *Phys. A Stat. Mech. its Appl.* 484, 11–20. <https://doi.org/10.1016/j.physa.2017.04.168>.
- Bahreiniotlagh, M., Kawanisi, K., Kavousi, A., Roozbahani, R., Abbasi, M., Al Sawaf, M., 2020. Influence of Suspended Sediment Concentration and Particle Sizes on the Sound Attenuation of the Fluvial Acoustic Tomography Technique. *J. Water Environ. Technol.* 18 (5), 338–348. <https://doi.org/10.2965/jwet.20-024>.
- Bahreiniotlagh, M., Kawanisi, K., Al Sawaf, M.B., Roozbahani, R., Eftekhari, M., Khoshuei, A.K., 2019. Continuous streamflow monitoring in shared watersheds using advanced underwater acoustic tomography system: a case study on Zayanderud River. *Environ. Monit. Assess.* 191 (11) <https://doi.org/10.1007/s10661-019-7830-4>.
- Bahreiniotlagh, M., Kawanisi, K., Danial, M.M., Al Sawaf, M.B., Kagami, J., 2016. Application of shallow-water acoustic tomography to measure flow direction and river discharge. *Flow Meas. Instrum.* 51, 30–39. <https://doi.org/10.1016/j.flowmeastinst.2016.08.010>.
- Coleman, H.W., Steele, W.G., 1999. *Experimentation and uncertainty analysis for engineers*, 2nd Ed. Wiley, New York, pp. 34–37.
- Goring, D.G., Nikora, V.I., 2002. Despiking acoustic doppler velocimeter data. *J. Hydraul. Eng.* 128 (1), 117–126. [https://doi.org/10.1061/\(ASCE\)0733-9429\(2002\)128:1\(117\)](https://doi.org/10.1061/(ASCE)0733-9429(2002)128:1(117)).
- Hawkins, D. M., 1980. Identification of Outliers. London, U.K.: Chapman and Hall, 10.1007/978-94-015-3994-4.
- Islam, M. R., Zhu, D. Z., 2013. Kernel Density-Based Algorithm for Despiking ADV Data, *J. Hydraul. Eng.*, vol. 139, no. 7, pp. 785–793, Jul. 2013, doi: 10.1061/(ASCE)HY.1943-7900.0000734.
- Kawanisi, K., Bahreiniotlagh, M., Al Sawaf, M.B., Razaz, M., 2016. High-frequency streamflow acquisition and bed level/flow angle estimates in a mountainous river using shallow-water acoustic tomography. *Hydrolog. Process.* 30 (13), 2247–2254. <https://doi.org/10.1002/hyp.v30.1310.1002/hyp.10796>.
- Kawanisi, K., Razaz, M., Ishikawa, K., Yano, J., Soltaniasl, M., 2012. Continuous measurements of flow rate in a shallow gravel-bed river by a new acoustic system. 48. 1–10. <https://doi.org/10.1029/2012WR012064>.
- Kawanisi, K., Razaz, M., Kaneko, A., Watanabe, S., 2010. Long-term measurement of stream flow and salinity in a tidal river by the use of the fluvial acoustic tomography system. *J. Hydrol.* 380 (1–2), 74–81. <https://doi.org/10.1016/j.jhydrol.2009.10.024>.
- Mori, N., Suzuki, T., Kakuno, S., 2007. Noise of acoustic Doppler velocimeter data in bubbly flows. *Journal of engineering mechanics* 133 (1), 122–125.
- Nikora, V., Goring, D., 2000. Flow Turbulence over Fixed and Weakly Mobile Gravel Beds. *J. Hydraul. Eng.* 126 (9), 679–690. [https://doi.org/10.1061/\(ASCE\)0733-9429\(2000\)126:9\(679\)](https://doi.org/10.1061/(ASCE)0733-9429(2000)126:9(679)).
- Otnes, R.K., Enochson, L., 1978. *Applied time series analysis*, Vol. 1. Wiley, New York.
- Rousseau, P.J., Leroy, A.M., 1998. *Robust regression and outlier detection*, Vol. 589. John wiley & sons.
- Wahl, T.L., 2003. Discussion of “Despiking acoustic doppler velocimeter data” by Derek G. Goring and Vladimir I. Nikora.“. *Journal of Hydraulic Engineering* 129 (6), 484–487.
- Zhong, C., Yin, F., Zhang, J., Zhang, S., Wan, R., Kitazawa, D., 2020. Optimized Algorithm for Processing Outlier of Water Current Data Measured by Acoustic Doppler Velocimeter. *Journal of Marine Science and Engineering* 8 (9), 655.
- Zou, H., Xiang, D.W., Jing, Y.G., 2008. Data smoothing methods for DVL. *Tech. Acoust.* 27, 507–510.