**Soil Moisture Active Passive (SMAP) Project:**

**Quality-Controlled SMAP Polarimetric L-Band Synthetic Aperture Radar Data (Global April 1 to July 7, 2015) Research Data**

**Citation:**

M. Mousavi1, A. Colliander1, S. Dunbar1, S. Yueh1, and D. Entekhabi2, January 20, 2022. *Quality-Controlled SMAP Polarimetric L-Band Synthetic Aperture Radar Data (Global April 1 to July 7, 2015) Research Data*, SMAP Project, JPL D-<#####>, Jet Propulsion Laboratory, Pasadena, CA.

Paper copies of this document may not be current and should not be relied on for official purposes. The current version is in the Product Data Management System (PDMS): <https://pdms.jpl.nasa.gov/>

**January 21, 2022**

**JPL D-<#####>**

National Aeronautics and Space Administration



Jet Propulsion Laboratory

4800 Oak Grove Drive

Pasadena, California 91109-8099

California Institute of Technology

Copyright 2021 California Institute of Technology.  U.S. Government sponsorship acknowledged.

**Contributors to this report:**

M. Mousavi1, A. Colliander1, S. Dunbar1, S. Yueh1, and D. Entekhabi2.

1Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA

2Massachusetts Institute of Technology, Cambridge, MA 02139 USA

This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. © 2022. All rights reserved.**TABLE OF CONTENTS**

1 EXECUTIVE SUMMARY 4

2 Outlier Detection Algorithm 5

2.1 Background 5

2.2 Method 6

3 Results 8

4 Summary 9

5 ACKNOWLEDGEMENTS 10

6 REFERENCES 11

# EXECUTIVE SUMMARY

The Soil Moisture Active Passive (SMAP) satellite has been developed by NASA in response to the National Research Council’s (NRC) which was released in 2007 [1, 2]. The SMAP data has been provided global coverage from low (40km) to moderate (enhanced to 9km) and high resolution (1km SAR) with high accuracy, which has been used across many science and applications discipline including hydrology, climate, carbon cycle, and the meteorological, environmental and ecology applications communities [3, 4, 5]. The L-band radiometer on the SMAP satellite was launched in January 2015 and continues nominal operations since March 2015, while the its radar failed in July 2015. This leaves us with about two months of SMAP SAR data, which can be used not only to construct a global high-resolution soil moisture retrieval over those time period but also as a benchmark to evaluate other active-passive soil moisture retrieval algorithms, such as SMAP-Sentinel L-C product [6]. It can also quantify advantages of transition to NASA-ISRO SAR (NISAR) [7].

All the aforementioned benefits and advantages of SMAP SAR dataset are based on a cleaned SAR data. However, the SMAP SAR data have not been fully and carefully investigated since the radar failure in 2015. In this report, the approximately two months of 1km SMAP synthetic aperture radar (SAR) data has been revisited and scrubbed. The SAR bad data (aka outlier) are detected and removed by statistically investigating the time series difference between scatterometer and linearly averaged SAR measurements (1 km) within the SMAP antenna footprint (~38 km). It is performed orbit by orbit. The outlier or bad orbits were identified when a data point is more than three scaled median absolute deviations (MAD) away from the median. On average only about 10% of all SAR orbits (more than 700), in each polarization, are classified as outliers.

# Outlier Detection Algorithm

## Background

In order to detect bad and erroneous SMAP SAR orbits, the scatterometer data were used as a baseline since they do not include the SAR processing, and their spatial resolution is limited by the antenna footprint size on the ground (~40km). This approach assumes that there are negligible to no errors due to the antenna itself and the radar instrument at the time of operation, which is a true assumption as the scatterometer data has been quality checked before, and the antenna is being shared by the radiometer, which continues to have nominal operation. In fact, the radar will employ pulse compression in range and Doppler discrimination in azimuth to sub-divide the antenna footprint, which is analogous to SAR processing techniques utilization for the SMAP conically scanning radar. Our hypothesis is that this processing can be the main source of bad data production in some orbits. The SMAP conically scanning measurement geometry is shown in Fig. 2.1 [9], where the spacecraft ground velocity vector is shown by . Using this geometry, the SAR 1km pixels that fall into the egg-shaped footprint are linearly averaged in order to be compared with the scatterometer measurements. In order to find the SAR pixels inside the scatterometer footprint, first the equation of the ellipse at each scan angle is defined, then the SAR pixels that their center fall inside the ellipse are being picked.

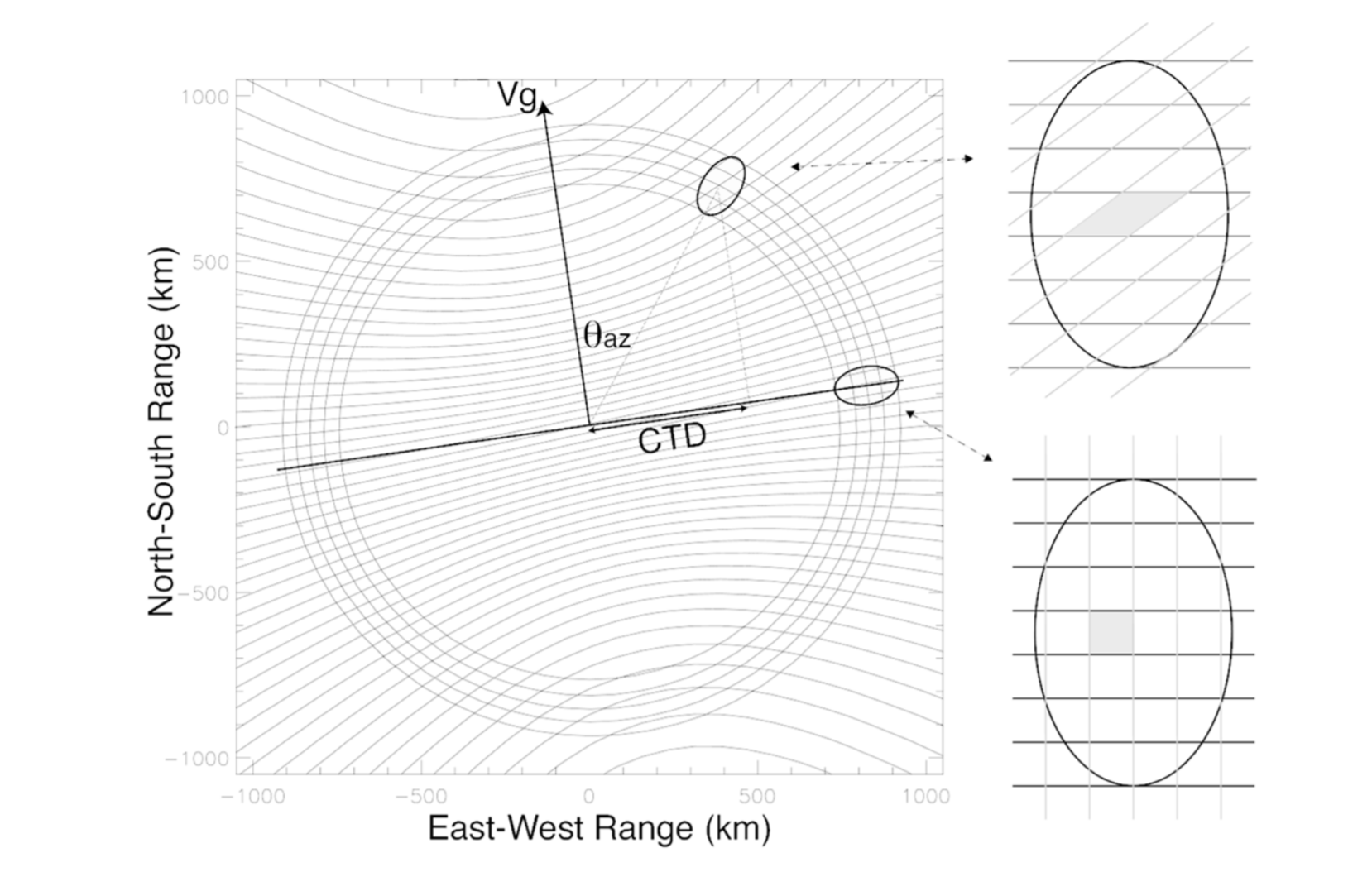


Figure 2.1. SMAP radar measurement geometry as a function of scan angle.

## Method

Any real-world data may suffer from outliers, which can drastically distort the mean, variance, and other data statistics. It is obvious that the quantity and value of these outliers are important. For example, if there are two large value outliers in a small dataset, they may become unnoticed. So, one has to choose the right outlier detection method based on the dataset size, distribution, and etc.

Our dataset is time series of the difference between scatterometer measurements and linearly averaged SAR measurements (1 km) within the antenna footprint, as discussed in Section 2.1, for each orbit (~700 for each SAR polarization). As an example, the time series of orbital mean of the SAR and scatterometer measurement difference (H2L-Lo) for HH channel in each day from May 15, 2015 till July 05, 2015 is shown in Fig. 2.2. The “H2L” stands for “High to Low Resolution”, which is the linearly averaged SAR measurements in the antenna footprint, and “Lo” stands for “Low Resolution”, which is the scatterometer measurements. It can be observed that the June 15-17, 2015 days are outliers where the expected value is 0 dB (ratio of one). This discrepancy could be due to either one, couple, or all outlier orbits for those dates. To better illustrate this, Fig. 2.3 shows the global map of the H2L-Lo for descending (Desc) orbits of HH channel on June 15, 2015. It can be observed that there are at least three bad SAR orbits.

Since there is no presumption about our data distribution, an outlier labeling method, such as Tukey method based on interquartile range (IQR) [10], Z-score method [11], and MADE method based on median and median absolute deviation (MAD) [12] will be pursued. Among all these methods, the MADE method is the most robust for our small dataset with the presence of extreme values. This method identifies an outlier if a data point is more than three times scaled MAD (MADE) away from the median (), as given by

|  |  |
| --- | --- |
|  | (1) |

where is a binary state for each data point indicating outlier (1) or clean (0) data, and is the scaled version of MAD as given by

|  |  |
| --- | --- |
|  | (2a) |
|  | (2b) |

where is the total number of data points. This scaling factor assures that the MADE approximates the actual standard deviation of the data. Instead of three times MADE in (1) one can use two times MADE, but we chose three to be more conservative and confident in the final cleaned SMAP SAR dataset.



Figure 2.2. The time series of orbital mean of the H2L-Lo scattering coefficient for HH channel.

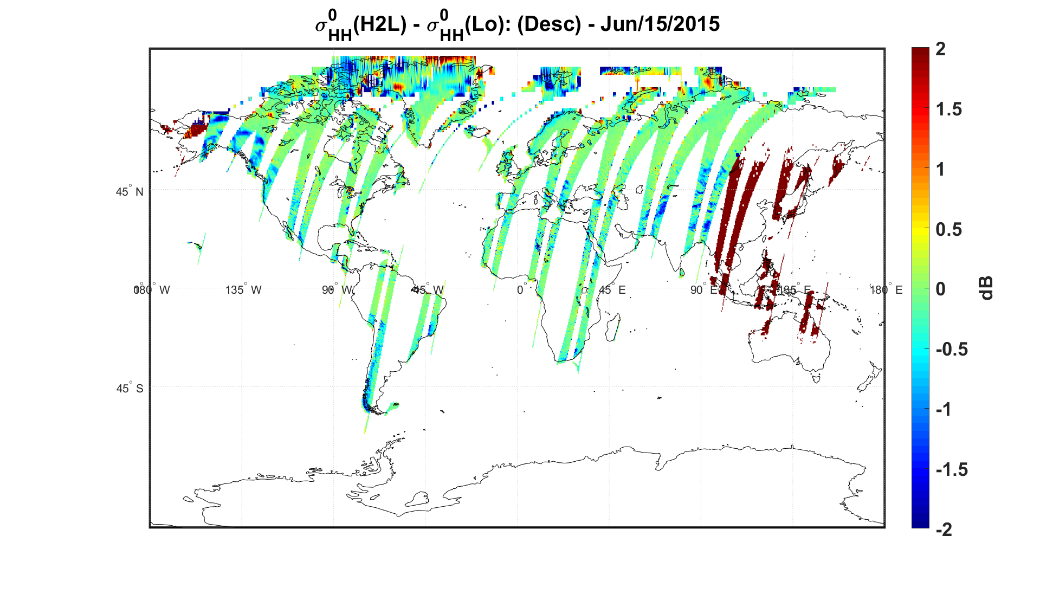


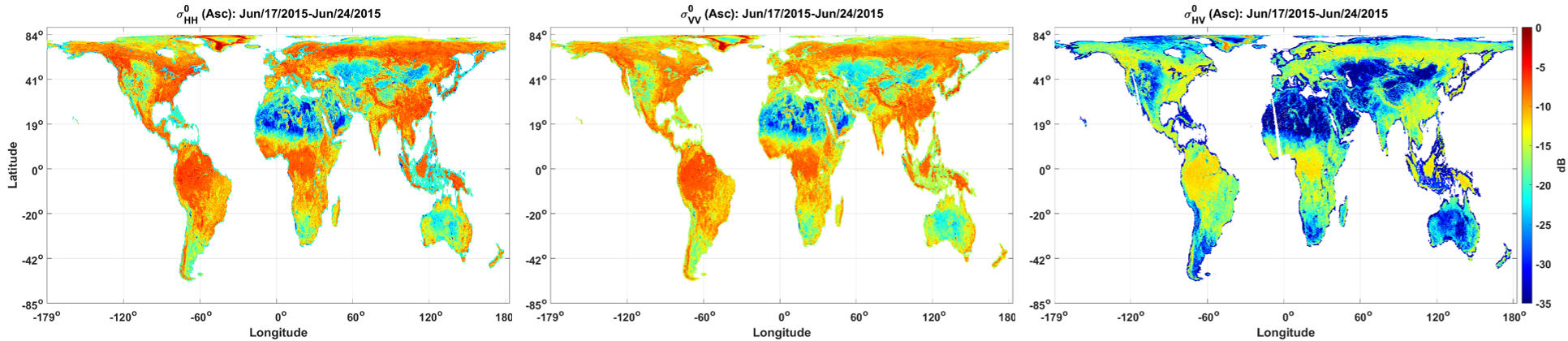
Fig. 2.3. Global map of the H2L-Lo scattering coefficient for descending (Desc) orbits.

# Results

After applying the proposed MADE method, as explained in Section 2.2, to all orbits for each channel (HH, VV, and HV) and separately for descending (Desc) and ascending (Asc) orbits, on average, less than 10% of the total number of orbits (~700) were identified and removed as outlier for each channel and each orbit (Desc and Asc). This proportion of outliers is a tolerable amount of lost data to still have a clean global coverage. The detail of the number of outlier orbits for each channel and orbit is shown in Table 3.1. As an illustrative example, after removing the outlier orbits, an 8 days (June 17-24, 2015) average map of each channel scattering coefficient () for Asc and Desc orbits are shown in Figures 3.1 and 3.2, respectively. It can be observed that there are few missing orbits in HV channel as it was expected due to their largest number of outlier orbits. That said, there is still a very good global coverage of SMAP SAR clean and quality checked dataset, which can be used for further analysis in applications, such as soil moisture retrieval, and land freeze/thaw classification.

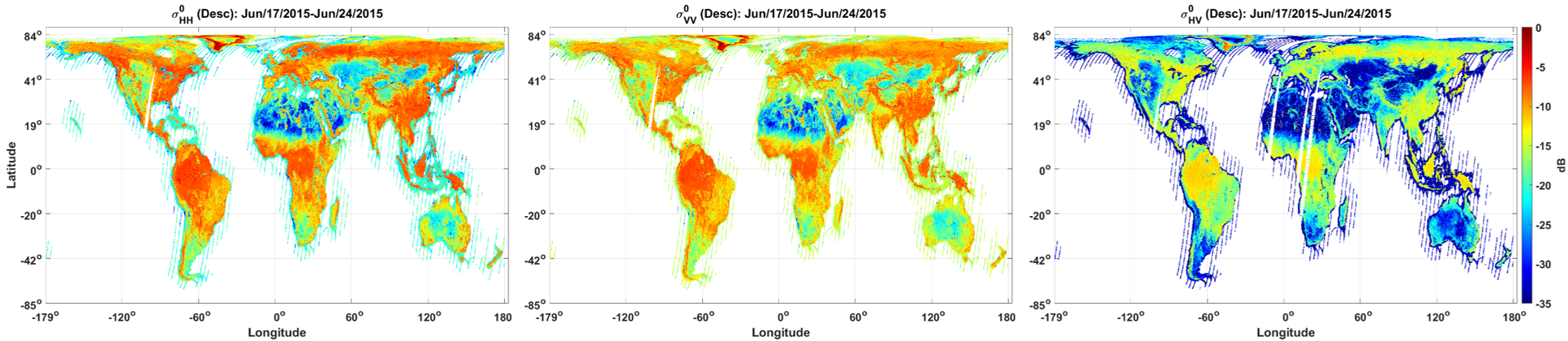
Table 3.1. The number of outlier orbits for each channel.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Ascending Orbits** | | **Descending Orbits** | |
| **Channel** | **Outlier** | **Total** | **Outlier** | **Total** |
| *HH* | 33 (4.47%) | 739 | 30 (4.06%) | 738 |
| *VV* | 43 (5.82%) | 739 | 27 (3.66%) | 738 |
| *HV* | 29 (3.93%) | 739 | 72 (9.75%) | 738 |



|  |  |  |
| --- | --- | --- |
| (a) | (b) | (c) |

Fig. 3.1. The 8 days average global maps of (a) HH, (b) VV, (c) HV scattering coefficients () for ascending (Asc) orbits.



|  |  |  |
| --- | --- | --- |
| (a) | (b) | (c) |

Fig. 3.2. The 8 days average global maps of (a) HH, (b) VV, (c) HV scattering coefficients () for descending (Desc) orbits.

# Summary

NASA’s SMAP satellite, launched in January 2015, were planned to offer a valuable set of low frequency radiometer (1.41 GHz) and radar (1.26 GHz) measurements since April 2015 to monitor the soil moisture and land surface freeze/thaw at high resolution and accuracy. After the radar failure in July 2015, SMAP only offers radiometer measurements, and its radar left us with about two months of data. The SMAP community has utilized the Sentinel radar measurements at C-band to substitute for the SMAP L-band SAR in active/passive soil moisture retrieval or used an enhanced version of SMAP radiometer measurements at 9km.

However, the SMAP SAR data has not yet been carefully investigated since its failure. In this report, we revisited and cleaned the available SMAP SAR measurements from May 15 till July 05 of 2015. We compared the SMAP SAR with its scatterometer measurements, orbit by orbit, and employed an outlier detection algorithm based on the scaled median absolute deviation (MADE) to detect bad and erroneous SAR orbits. On average for all channels (HH, VV, and HV) and both ascending and descending orbits, about 10% of all orbits (~700) were identified as outlier and erroneous SAR orbits. Finally, with this minimal loss of SMAP SAR data, this cleaned dataset would provide the first global L-band SAR mapping to be a great benchmark to evaluate different active/passive soil moisture retrieval algorithms, such as SMAP-Sentinel L-C product, and quantify advantages of transition to utilizing NASA-ISRO SAR (NISAR) future mission.

# ACKNOWLEDGEMENTS

This document resulted from many hours of diligent analyses and constructive discussion among the SMAP project team members.

# REFERENCES

|  |  |
| --- | --- |
| [1] | D. Entekhabi et al., "The Soil Moisture Active Passive (SMAP) Mission," *Proceedings of the IEEE,* vol. 98, no. 5, pp. 704-716, 2010. |
| [2] | National Research Council, Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, Washington, DC: The National Academies Press, 2007. |
| [3] | M. Mousavi, A. Colliander, J. Z. Miller and J. Kimball, "Evaluation of surface melt on the Greenland ice sheet using SMAP L-band microwave radiometry," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* vol. 14, pp. 11439-11449, 2021. |
| [4] | S. H. Yueh et al., "SMAP L-Band Passive Microwave Observations of Ocean Surface Wind During Severe Storms," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 54, no. 12, pp. 7339-7350, 2016. |
| [5] | J. Du, J. S. Kimball, M. Azarderakhsh, R. S. Dunbar, M. Moghaddam and K. C. McDonald, "Classification of Alaska Spring Thaw Characteristics Using Satellite L-Band Radar Remote Sensing," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 53, no. 1, pp. 542-556, 2015. |
| [6] | N. Das and R. S. Dunbar, "Level 2 SMAP/Sentinel Active/Passive Soil Moisture Product Specification Document," 2018. |
| [7] | P. A. Rosen et al., "The NASA-ISRO SAR mission - An international space partnership for science and societal benefit," in *2015 IEEE Radar Conference (RadarCon)*, 2015. |
| [8] | M. E. Brown et al., "NASA’s Soil Moisture Active Passive (SMAP) mission and opportunities for applications users," *Bull. Amer. Meteor. Soc.,* vol. 94, pp. 1125-1127, 2013. |
| [9] | M. Spencer, S. Chan, L. Veilleux and K. Wheeler, "The Soil Moisture Active/Passive (SMAP) mission radar: A novel conically scanning SAR," in *2009 IEEE Radar Conference*, 2009. |
| [10] | J. W. Tukey, Exploratory Data Analysis, Addison-Wesley Pub, 1977. |
| [11] | R. E. Shiffler, "Maximum Z Scores and Outliers," *The American Statistician,* vol. 42, no. 1, pp. 79-80, 1988. |
| [12] | S. Burke, "Missing values, outliers, robust statistics & non-parametric methods," *Scientific Data Management,* vol. 1, pp. 32-38, 1998. |