Quantification of Rain to Improve Scatterometer Wind Speed by a Support Vector Machine Method

Xingou Xu [1], Ad Stoffelen [2*]

1 National Space Science Center, Chinese Academy of Sciences (NSSC, CAS), China. E-mail: xuxingou@mirslab.cn

2 Royal Netherlands Metrological Institute (KNMI). E-mail: ad.stoffelen@knmi.nl*

Ku-band backscatter (NRCS) is more affected by rain than C-band for scatterometers, with 10 times more rejections in wind Quality Control (QC). Their normalized residual values from wind retrieval (Rn) is correlated with rain rate, while the spatially informed QC indicator (Joss) demonstrates improved rain screening. Hence, both geophysical and spatial retrieval properties can inform rain QC. In this research, SVM is applied to collocations from spatiotemporal registered C- and Ku-band wind vector cells (WVCs) from ASCAT and OSCAT-2, with rain references from Global Precipitation Mission (GPM) products. A parameter relevant to rain fraction in a Wind Vector Cell (WVC) is also proposed to describe the underlying physics of the problem.

A recognition-regression procedure for rain rate and wind regression approach have been conducted. Results for rain identification on Ku-band scatterometer QC rejections achieves 70.2% reinstated WVCs with associated corrected wind speed of good accuracy. The correlation coefficient with GPM and estimated rain rate is 0.42 for rain rates ranging from 0 to 10 mm/h. Results can provide both wind and rain information for applications, and as a good reference for wind-rain uncertainty for data assimilation.

Further research would be to investigate the kernel functions for the uncertainty determination of the established SVM.

1 Background & Introduction

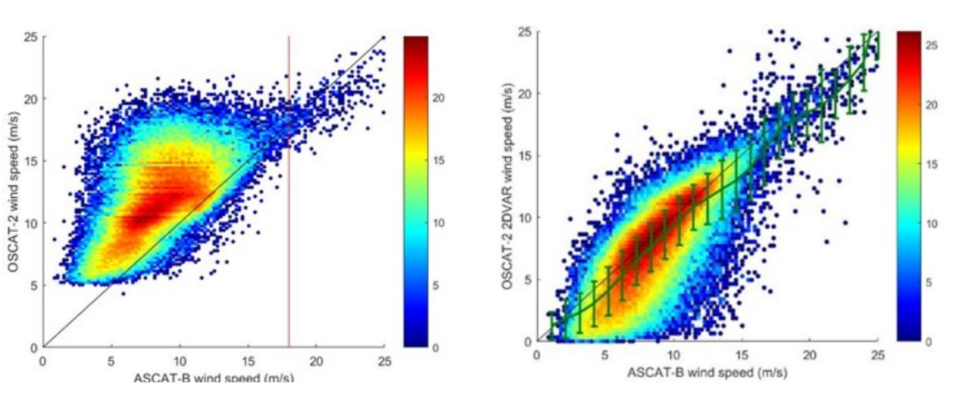
Parameters	Descriptions
Geophysical Model Function	GMF, an empirical model function that maps a wind vector to NRCS. It is employed in wind inversion by Maximum Likelihood Estimation (MLE) to obtain wind from NRCSs.
The analysis wind speed	Also referred to as 2DVAR wind speed. It is obtained in the 2DVAR ambiguity removal procedure after wind retrieval balancing the differences of observed and background winds. Referred to as *f* hereafter.
Wind scatterometry Quality Control	Referred to as QC here. In tropical regions, QC rejections are generally associated with rain, and labelled by QC indicators. Ku-band observations are 10 times more frequently (~5%) rejected than C-band (~0.5%) due to the shorter wavelength.
R_n	•A widely applied QC indicator obtained during wind inversion, as normalized distances between NRCS and GMF calculated in the MLE procedure.
	•Featured for correlations with rain features, and also features for different rain developing phases.
J_{oss}	•A spatially-informed QC indicator representing difference of observational and analysis wind speeds.
	•Featured for its rain screening ability and False Alarm Rate (FAR) reduction ability.
α	•Proposed in this research as a rain fraction parameter in a WVC: $\alpha = (J_{oss})/(f-18)$

We investigate by its QC rejected Ku-band WVCs collocated with by its QC accepted C-band WVCs, information provided by α, Ku-band observational speed, ECMWF wind speed, complemented by reference rain rates from the Global Precipitation Mission (GPM) and wind speed from the C-band products. A support vector machine (SVM) is employed to model rain features and rain effect corrections to wind speeds from Ku-band observations.

2 Data Set & Collocation

The QC accepted C-band winds are not much affected by rain and used as reference in this research (Tropical region). The spatial distance is less than 25km and observing time lag less than 25 min (minutes) between scatterometers and 2.4 min of OSCAT-2 and GPM rain rates. 11,274 WVCs are obtained.

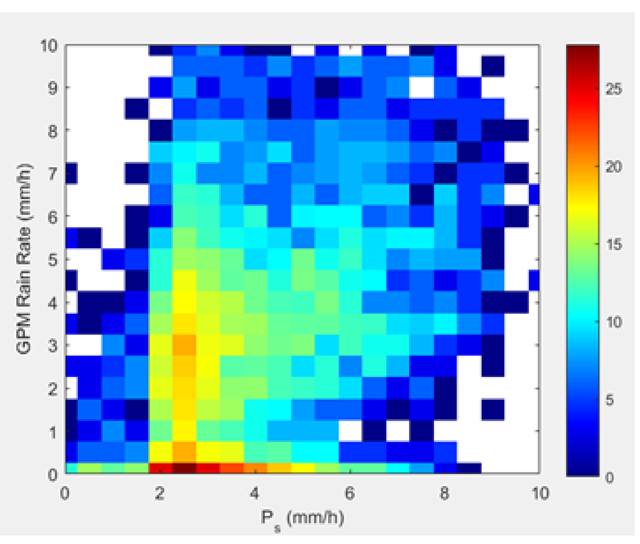
Data	Source
Ku-Scatterometer	OSCAT-2 onboard SCATSAT-1
C-Scatterometer	Advanced Scatterometers (ASCAT-A and ASCAT-B) onboard MetOP-A and MetOP-B
Precipitation	NASA GPM L (Level 3 Integrated Multi-satellitE Retrievals) in version 5.
Model winds	ECMWF winds.



Winds from C-band accepted and Ku-band rejected set. Horizontal axis are ASCAT-B winds, vertical axis are OSCAT-2 observational winds (a) and OSCAT-2 2DVAR winds (b). Color bar indicates WVC density in dB.

From the left panel, rain effect would saturate and observations demonstrate a confusion with wind speed and rain cloud effect in heavy rains. Since analysis winds are in good accordance with C-band good quality winds, considering also sample numbers, WVCs with analysis speed under 18m/s of QC accepted by C-band and rejected by Ku-band is extracted for discriminated rain features .

3 α and Rain Rate: a Least Square Regression Approach



Rain rates from GPM are related to α in the regressed equation (Ps for rain rate in mm/h):

$$P_s = 9.7767*\alpha - 0.397$$

The left panel shows the density plot of the regressed rain Ps and GPM rain rates, color indicates density in dB. The correlation coefficients for all rain rates as well as rain rates in the range of [0, 10] are the best in all sets, as **0.4132** and **0.4147** respectively.

The LSR approach regresses rain rate and α, which latter depends on Joss. Independently of Joss, rain rates are associated with high Rn. Moreover, differences of observational wind speed and background speed from ECMWF increase with rain. We use this in SVM.

4 Support Vector Machines (SVMs):

A SVM transforms several inputs into a space of linearized outputs using kernel functions. It is used to solve problems that are non-convex and difficult to solve in the origin input space. The kernel function applied is the (Gaussian) radial basis function kernel, where L2 distances are minimized in the procedure.

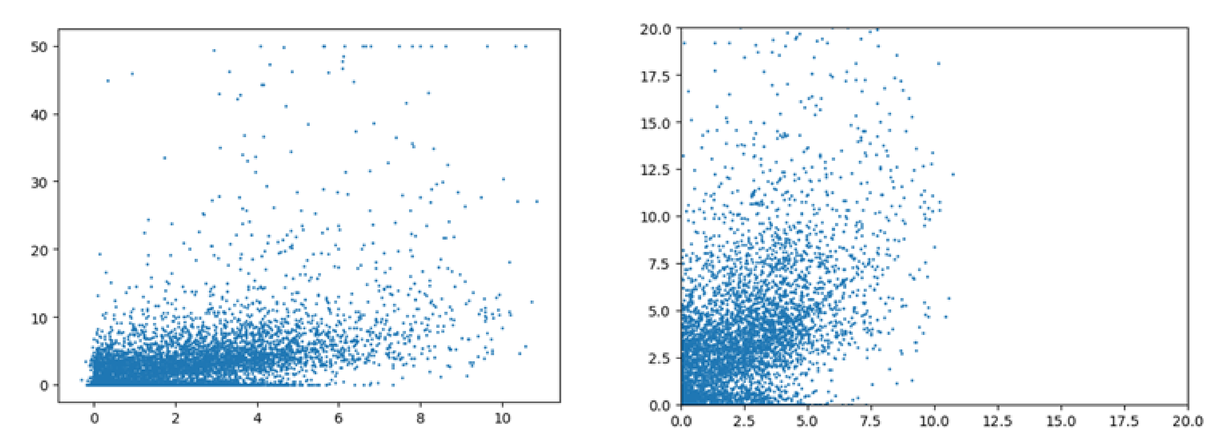
In this research, using the same inputs, SVMs for 1) rain/no-rain labelling, for 2) rain rate regression and for 3) wind speed correction have been established respectively. The former two SVMs investigate the ability of this method to detect rain, given the set inputs.

6,764 WVCs used for training (60%), and 4,510 WVCs for testing.

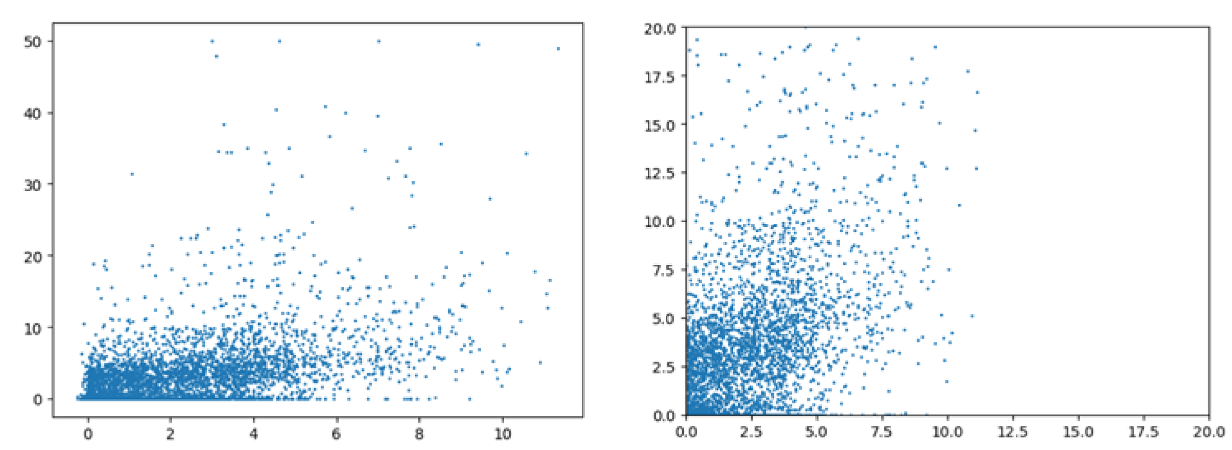
Inputs for SVMs	Output of SVMs	Values of output for training
• Rn in dB space, • α,	1) Rain / 0 surface rain rate label	1) GPM rain rates, 0 and non-zero cases
• Analysis speed,	2) Rain rates (mm/h)	2) GPM rain rates in mm/h
• Observational speed, • ECMWF speed.	3) Corrected wind speed (m/s)	3) wind speed form ASCAT-A or ASCAT-B

5 Results

- **The classification accuracy for training set is 73.4%, for testing set is 71.4%.**
- **Rain regression: In the following figures, training set results are displayed first, followed by those from the testing set.**



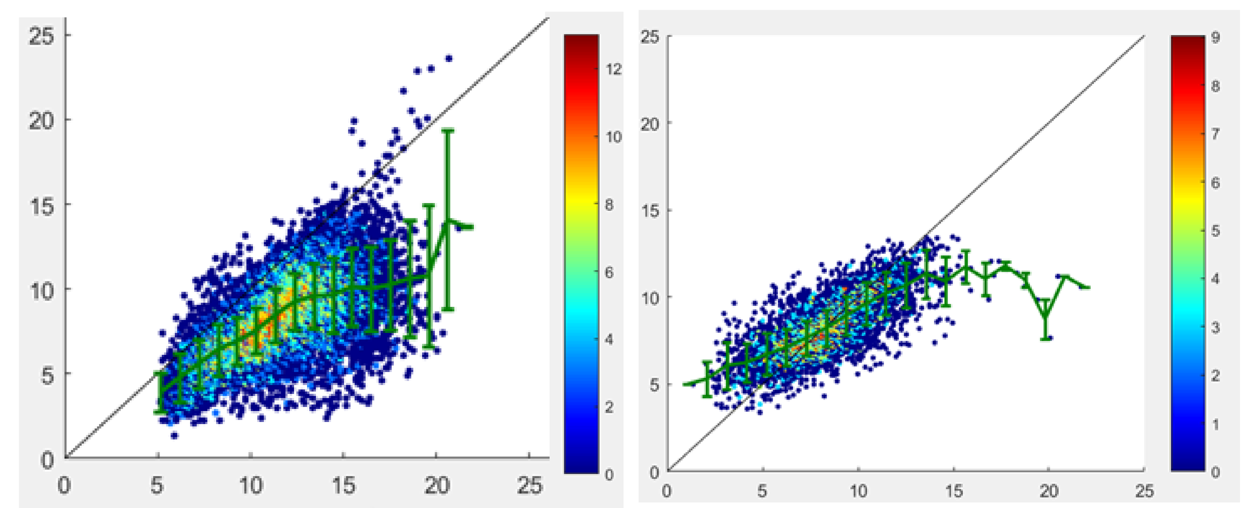
Correlation coefficient of regressed rain to GPM is 0.504 for the training set .



Correlation coefficient of regressed rain to GPM is 0.481 for the testing set .

The horizontal axis for the left four figures are the regressed rain rates, and the vertical axis are rain rates from GPM. They are in mm/h.
The right two panels are the zoomed figures for the left two.

- **Wind speed correction:**



Observational (left) and SVM corrected wind speed (right) with references to ASCAT speed (horizontal axis). Color indicate number density of WVCs in dB.

The horizontal axis for the two figures are the wind speed from ASCAT products, and the vertical axis are those from OSCAT-2 observational speed (left) and regressed wind speed(right) for testing sets, in m/s.

6 Further Research

- **To investigate the kernel functions for the uncertainty determination of the established SVM**
- **To establish rain and wind speed outputs in one SVM for modelling different rain effects for correction of wind speeds.**