



Supplement of

Cloud fraction estimation using random forest classifier on sky images

Sougat Kumar Sarangi et al.

Correspondence to: Chandan Sarangi (chandansarangi@civil.iitm.ac.in)

The copyright of individual parts of the supplement might differ from the article licence.

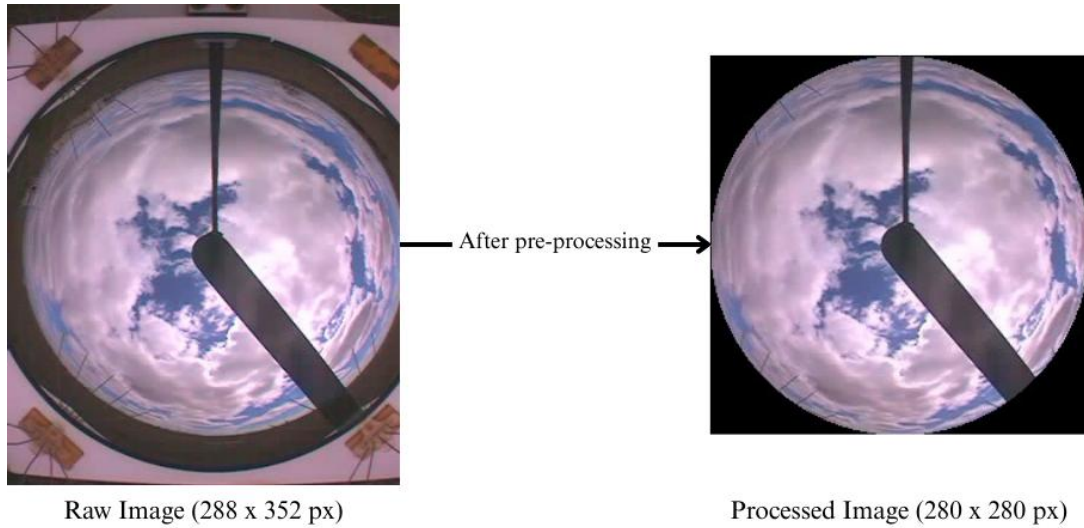


Figure S1. Left side - Raw image (from Lamont Canada) captured at 5:07PM on 07-04-2010 showing image of sky along with surrounding area that is not essential for cloud fraction estimation. Right side – Processed image after removing dead zones by cropping and resizing the raw image and then applying a circular mask to highlight only the usable area of the raw image.

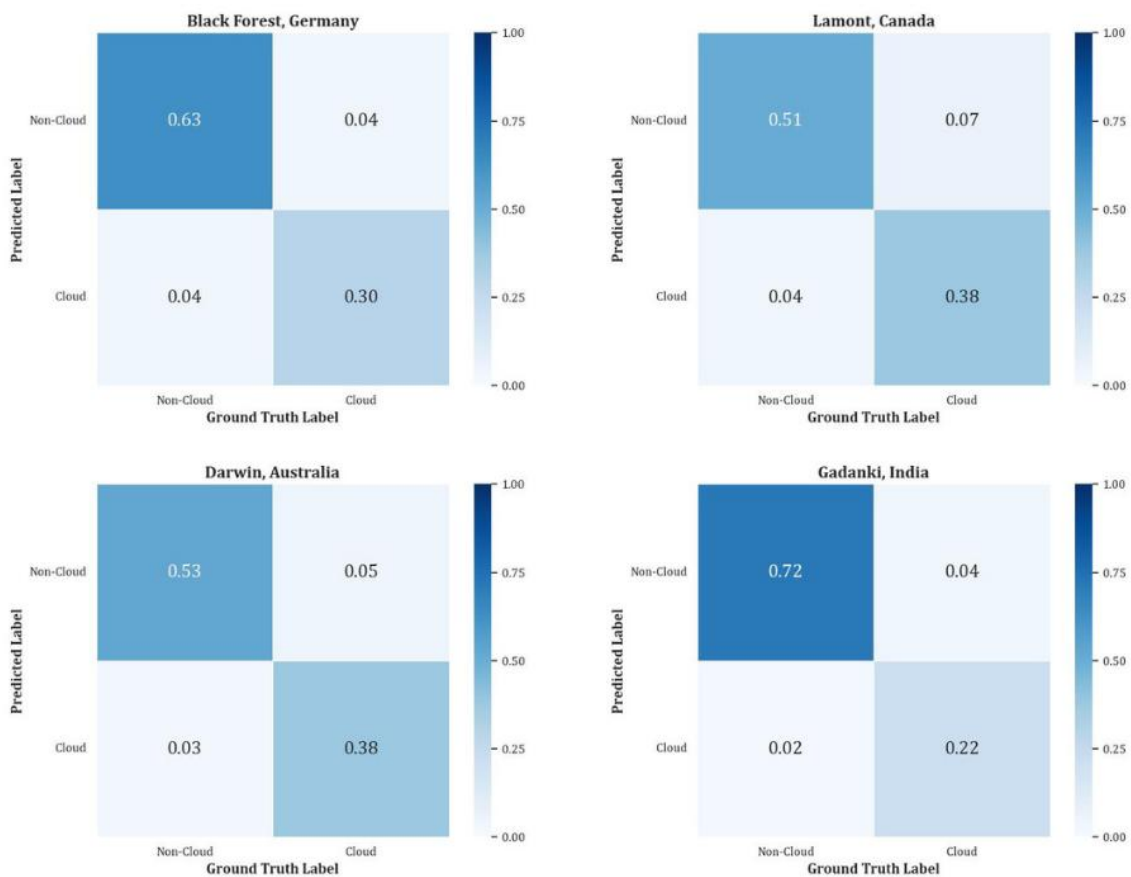


Figure S2. Confusion matrix obtained from the test set for each location. The standard layout of the confusion matrix is as follows:

	Actual: Non-Cloud	Actual: Cloud
--	-------------------	---------------

Predicted: Non-Cloud	$T_{nonCloud}$	$F_{nonCloud}$
Predicted: Cloud	F_{cloud}	T_{cloud}

Thus, the performance metrics are defined as:

$$Accuracy = \frac{T_{cloud} + T_{nonCloud}}{T_{cloud} + T_{nonCloud} + F_{cloud} + F_{nonCloud}}$$

$$Precision = \frac{T_{cloud}}{T_{cloud} + F_{cloud}}$$

$$Recall = \frac{T_{cloud}}{T_{cloud} + F_{nonCloud}}$$

$$F1\ Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

$$IoU = \frac{T_{cloud}}{T_{cloud} + F_{cloud} + F_{nonCloud}}$$

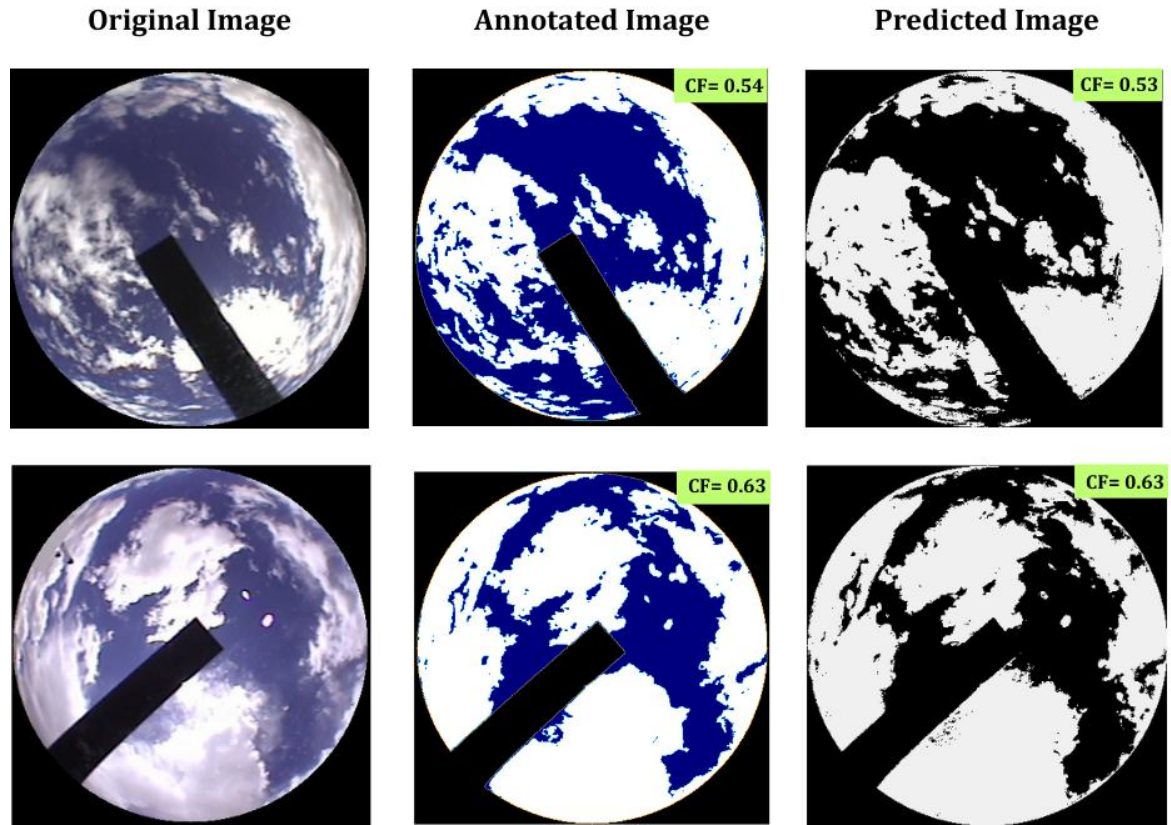


Figure S3. A few predicted outputs of the RF classifier for images taken at Merak India, that are accurate

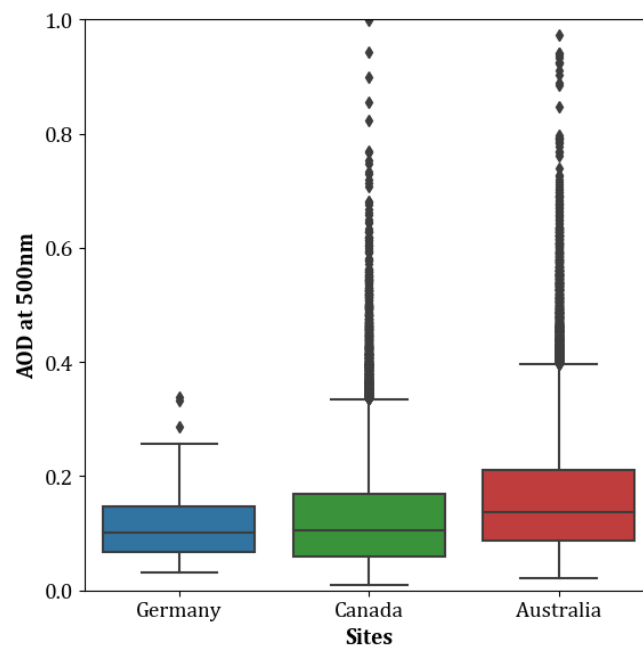


Figure S4. Box plot of one-year aerosol optical depth (AOD) at 500nm from three ARM sites: Black Forest, Germany (2007); Lamont, Canada (2010); and Darwin, Australia (2010). The plot shows the occurrence of higher AOD in Canada and Australia compared to Germany.