

Review of: A novel machine learning retrieval for the detection of ice crystal icing conditions based on geostationary satellite imagery

Summary statement

In this paper, Aricò et al. present a method for detecting high ice water content (HIWC) conditions associated with aviation ice crystal icing (ICI) events using geostationary satellite data. They apply a machine learning approach to derive a set of input variables from Meteosat Second Generation (MSG) SEVIRI derived products. Training of a random forest is accomplished using CloudSat radar and lidar (DARDAR) products. Finally, the authors obtained a database comprised of in-service ICI events which they have used to assess performance of their detection method.

The paper is clearly written and well-organized. It begins with a comprehensive description of the ICI threat to aviation and a review of previously published HIWC detection methods. It provides thorough descriptions of the various data sets used. Explanations of the random forest approach, why it was chosen, and how it was applied to this problem are detailed and clear. Results presented are meaningful, and conclusions are well-justified. Significantly, the authors have demonstrated the need for a dedicated HIWC detection product by showing displacement between their HIWC Mask and detected convective cells. Overall, this is a very good paper that needs only minor revision.

Specific points:

- 1.If I understand correctly, the method is limited to cruise altitudes, but it's not clear why. ICI engine events have occurred during ascents and descents, so the hazard is not limited to cruise altitudes.
- 2.While I understand why the method is limited to daytime, given its reliance on products derived from visible channels, I think it would be appropriate to at least discuss how you might develop a corresponding nighttime method.
- 3.It's difficult to see the symbols in certain figures (e.g., Figs 1, 7, 9, 10). The images and symbols could be enlarged and/or the symbol color could have better contrast with the background image.
- 4.Verification against in-service ICI events is very important, but the Lufthansa database apparently only includes air data system (ADS) events (e.g., TAT anomalies), not engine events. I assume the authors did not have access to the latter. Some discussion of the relationship between ADS events and engine events would bolster the significance of your results.
- 5.In the box and whisker plots shown in Fig. 6, some of the variables on the vertical axis only have outlier points, i.e., no box and whiskers. Could you explain how this should be interpreted?

Comment 1: If I understand correctly, the method is limited to cruise altitudes, but it's not clear why. ICI engine events have occurred during ascents and descents, so the hazard is not limited to cruise altitudes.

Author reply:

Although ICI events were observed during other flight phases than cruise, we tried to convey the message that geostationary satellites are not able to detect HIWC conditions in-clouds, thus making this scenario challenging to detect. This can be found in:

Page 9, Line 197: *Most of the events collected by Lufthansa fall within the specified boundaries, except*

for three cases. 88% events occur between 9000 m (29527 ft) and 13000 m (42650 ft). Continues in Line 196: Furthermore, while testing for multiple cruise level limits (not shown), we observed that the correct detection of HIWC conditions was more likely when these conditions occur at higher altitudes, as observed also by de Laat et al. (2017). We speculate that this is due to passive sensors mainly measuring emitted and reflected radiation in proximity to cloud tops.

Page 20, Line 347: However, in this instance, cruise levels are chosen according to the altitude at which ICI events occur in the Lufthansa ICI database. Nonetheless, the best trade-off to retrieve cloud properties remains challenging to find. Cloud properties vary within the cloud structure, while passive sensors can only detect cloud top characteristics or column-integrated quantities.

Manuscript changes:

We tried to emphasize further this concept:

Page 9, Line 201: We speculate that this is due to passive sensors measuring emitted and reflected radiation in proximity to cloud tops, thus inherently limiting the in-cloud HIWC detection.

Page 25, Line 424: Cruise levels are considered because, even if ICI events are possible during the ascent and descent of an aircraft, passive remote sensing platforms are more sensitive to cloud tops and column-integrated quantities.

Comment 2: While I understand why the method is limited to daytime, given its reliance on products derived from visible channels, I think it would be appropriate to at least discuss how you might develop a corresponding nighttime method.

Author reply:

It is a very good point, and we have added a night-time product demonstration for few Lufthansa ICI cases. Although a dedicated night-time product was not developed, it gave us the chance to test the current algorithm version during night-time.

Manuscript changes:

Section 4.3, and additional Lufthansa ICI case in the Appendix.

Page 23, Line 389: The retrieval is here tested during nighttime. In this scenario, the random forest model does not have access to visible channel information and cloud optical thickness. Furthermore, it has been trained exclusively with day-time samples. Nevertheless, it can access infrared channels and convection related variables. In night-time mode, we decided to use instrumental values to fill the missing optical information required by the random forest approach.

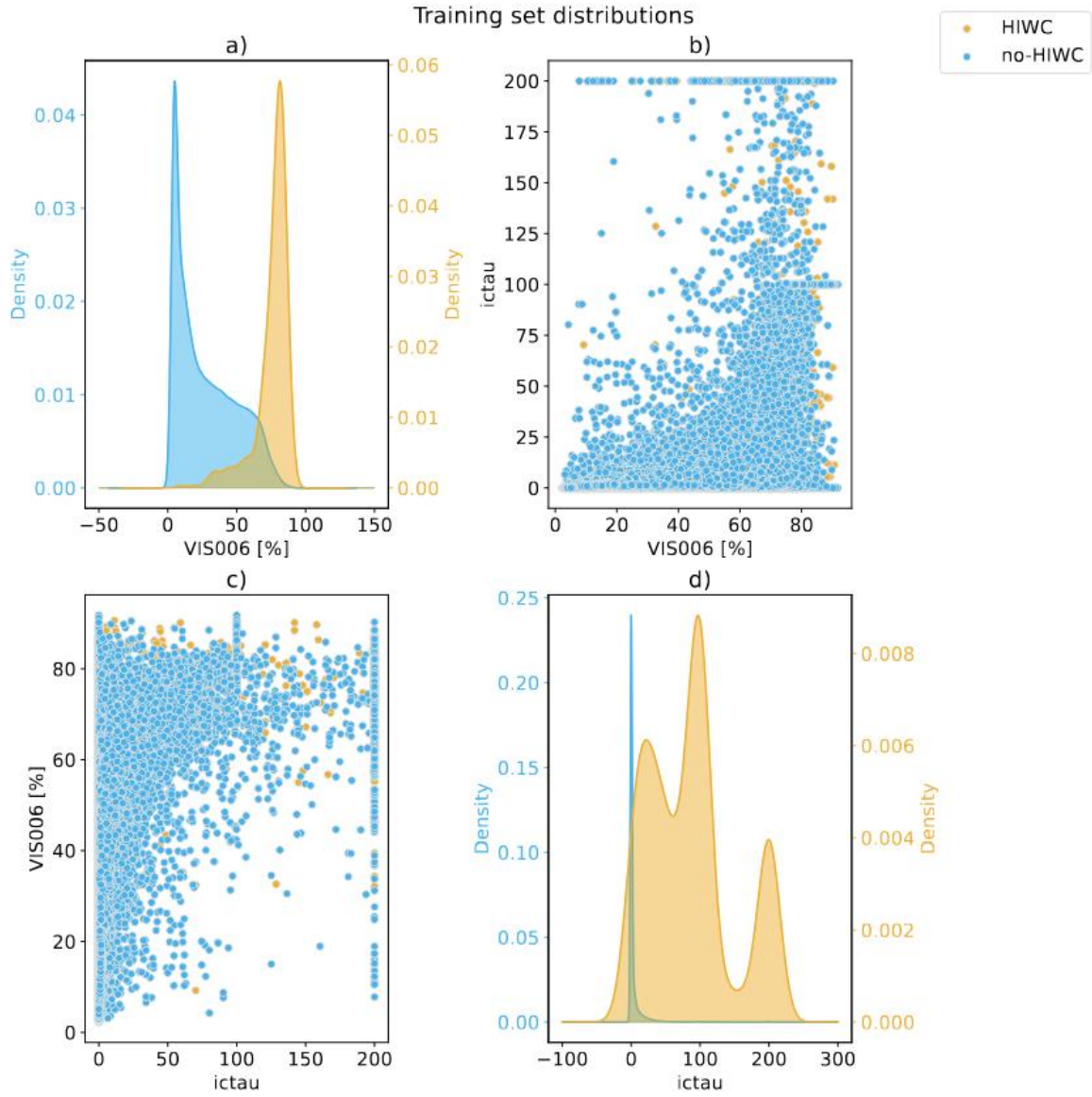


Figure 11. a) $VIS006$ and d) $ictau$ samples distributions in the training dataset, for HIWC (orange) and no-HIWC (blue). Negative values in both distributions arise because of the curve smoothing for plot purposes. The real sample locations are visible in panels b), and d).

In Fig. 11, the distribution in the training dataset of $VIS006$ and $ictau$ for HIWC and no-HIWC is shown. These distributions allowed us to select a bias-free value with which we filled the missing information in nighttime mode. In particular, this bias-free value is selected such that it favors neither HIWC prediction, nor non-HIWC, i.e. the instrumental value should be in a range where HIWC and no-HIWC training samples distributions overlap. The values are set to $VIS006=80\%$ and $ictau=50$. The significance of this choice is shown in Fig. 12.

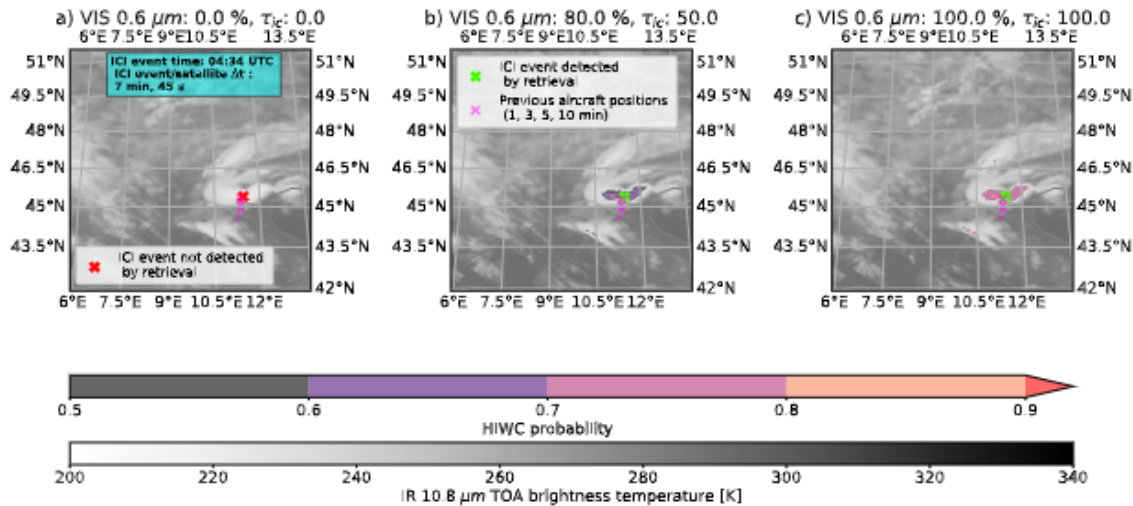


Figure 12. ICI retrieval nighttime mode demonstration example for a Lufthansa ICI event. Panel a) shows the HIWC mask setting instrumental values that favor no-HIWC prediction (see Fig. 11). Panel b) depicts the HIWC mask with bias free instrumental values. Panel c) shows the HIWC mask with instrumental values favoring HIWC predictions.

The mask in panel a), where we set $VIS006 = 0\%$ and $ictau = 0$, is absent because the HIWC probability never exceeds 0.5. In panel b), the bias-free choice of $VIS006 = 80\%$ and $ictau = 50$ leads to a smooth transition of HIWC probability between areas without detected HIWC and areas where HIWC is detected. Panel c) displays instead a sharp transition to high HIWC probabilities, as soon as this is detected by overcoming the 0.5 probability threshold. We observe that the constant instrumental values with which we fill missing information modulate the HIWC probability mask significantly. The choice made for panel b) is the best to achieve realistic results even with missing solar information. This demonstrates the good performance of the model even during nighttime.

Comment 3: It's difficult to see the symbols in certain figures (e.g., Figs 1, 7, 9, 10). The images and symbols could be enlarged and/or the symbol color could have better contrast with the background image.

Author reply:

Thanks for pointing that out and for providing a suggestion for a new layout. They have been modified in the manuscript accordingly.

Manuscript changes:

Figure 1, 7, 9, 10 were adjusted as suggested.

Comment 4: Verification against in-service ICI events is very important, but the Lufthansa database apparently only includes air data system (ADS) events (e.g., TAT anomalies), not engine events. I assume the authors did not have access to the latter. Some discussion of the relationship between ADS events and engine events would bolster the significance of your results.

Author reply:

Yes, that is correct, we do not have access to the engine events database. However, both TAT anomaly and engine failures can happen during the flight in HIWC regions. In conditions where TAT anomalies occur, the ice concentration is high, hence increasing the probability to have an engine failure. For this reason, TAT anomalies can be used as engine failure precursors. I do not have any information about the extent to which they correlate. We used TAT anomalies because those are the data available to us, and TAT anomalies are well correlated with HAIC regions.

Eventually the goal is to provide a prototype detection for those regions.

Explained in:

Page 2, Line 25: *this is where ICI events can occur because ice particles can build up inside the engine and lead to performance loss and damage (Grzych, 2010, 2015; Bravin et al., 2015; Haggerty et al., 2019), or they can clog the pitot tube which in turns result into a wrongful transmission of information to the autoflight system; this latter occurrence has caused two fatal accidents in recent years (S. Ayra et al., 2020).*

Manuscript changes:

Added to emphasize the concept:

Page 2, Line 28: *Because those failures can happen in high ice concentration regions, on-board sensor anomalies, as for example the total air temperature (TAT) anomalies, are often used as precursors for engine failures (Haggerty, 2016; Rodríguez-Sanz et al., 2018).*

References:

Bravin, M., Strapp, J. W., and Mason, J.: An Investigation into Location and Convective Lifecycle Trends in an Ice Crystal Icing Engine Event Database, in: SAE Technical Paper Series, SAE Technical Paper Series, SAE International400 Commonwealth Drive, Warrendale. PA, United States, <https://doi.org/10.4271/2015-01-2130>, 2015.

Grzych, M.: Avoiding convective weather linked to ice-crystal icing engine events, Boeing Aeromagazine, 2010.

Grzych, M., Tritz, T., Mason, J., Bravin, M., and Sharpsten, A.: Studies of Cloud Characteristics Related to Jet Engine Ice Crystal Icing Utilizing Infrared Satellite Imagery, in: SAE Technical Paper Series, SAE Technical Paper Series, SAE International400 Commonwealth Drive, Warrendale, PA, United States, <https://doi.org/10.4271/2015-01-2086>, 2015.

Haggerty, J. A.: High Ice Water Content and Airborne Temperature Measurement Anomalies in Tropical Convection, 32nd Conference on Environmental Information Processing Technologies, p. 1, 2016.

Haggerty, J. A., Rugg, A., Potts, R., Protat, A., Strapp, J. W., Ratvasky, T., Bedka, K., and Grandin, A.: Development of a Method to Detect High Ice Water Content Environments Using Machine Learning, Journal of Atmospheric and Oceanic Technology, 37, 641–663, <https://doi.org/10.1175/JTECH-D-19-0179.1>, 2020.

Rodríguez-Sanz, Á., Arnaldo, R. M., Sánchez Ayra, E., and Gómez Comendador, F.: Detecting HAIC icing events from TAT anomalies vs8, 31st Congress of the International Council of the Aeronautical Sciences, 2018.

S. Ayra, E., Rodríguez Sanz, Á., Arnaldo Valdés, R., Gómez Comendador, F., and Cano, J.: Detection and warning of ice crystals clogging pitot probes from total air temperature anomalies, Aerospace Science and Technology, 102, 105 874, <https://doi.org/10.1016/j.ast.2020.105874>, 2020.

Comment 5: In the box and whisker plots shown in Fig. 6, some of the variables on the vertical axis only have outlier points, i.e., no box and whiskers. Could you explain how this should be interpreted?

Author reply:

Thank you for bringing out this point, the explanation is now added.

Manuscript changes:

Page 16, Line 288: *This is denoted by the boxplot collapsing into a single line, which indicates that all the simulations carried out led to the same decrease in performance score, thus producing no distributions. Few outliers present for some variables, as for example D_A-1_2, D_A-1_3, IR_016, and ictau, indicates that only a minority of simulations led to a change in performance score.*