Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

Seongmun Sim¹ · Seonghyeon Ha¹ · Junghee Lee¹ · Jungho Im¹

¹ School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology (UNIST)
Contents

01. Introduction

02. Previous studies

03. Research content

1) Icing masking model based on COMS

2) Icing altitude estimation based on COMS

3) Icing masking model based on Himawari-8

04. Conclusion
Introduction

- Aviation accidents caused by icing

Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

- 12% of the total weather accidents are caused by icing
- Fatal accidents occur every year

⇒ Critical issues for aviation safety
INTRODUCTION

- What is icing?

Super-Cooled Droplet (SCD) clouds

Super-Cooled Droplet (SCD) occurs under 0°C and stable condition. When SCD collides on an object, SCD turns into ice form, which is ‘Icing’. Natural phenomenon, but too dangerous.

→ **Accurate observation and monitoring are required**
INTRODUCTION
- GEO-KOMPSAT-2 (GK-2A)

COMS

GK-2A

It can be good proxy data for GK-2A icing product!!

Himawari - 8

Icing!!
PREVIOUS STUDIES

- Current Icing Product/Forecast Icing Product (CIP/FIP)

- Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

Geostationary Operational Environmental Satellite (GOES)

CIP/FIP operating map

CIP/FIP algorithm
PREVIOUS STUDIES

Communication, Ocean and Meteorological Satellite (COMS)

<table>
<thead>
<tr>
<th>GOES - Imager</th>
<th>COMS – MI</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Band</strong></td>
<td><strong>Bandwidth, μm</strong></td>
</tr>
<tr>
<td>Vis</td>
<td>0.55–0.75</td>
</tr>
<tr>
<td>ShortWave</td>
<td>3.80–4.00</td>
</tr>
<tr>
<td>Moisture</td>
<td>6.50–7.00</td>
</tr>
<tr>
<td>IR-1</td>
<td>10.20–11.20</td>
</tr>
<tr>
<td>IR-2</td>
<td>11.50–12.50</td>
</tr>
</tbody>
</table>

**Computation Sources**

<table>
<thead>
<tr>
<th>Contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper tropospheric humidity</td>
</tr>
<tr>
<td>Cloud analysis</td>
</tr>
<tr>
<td>Cloud top temperatures &amp; heights</td>
</tr>
</tbody>
</table>

- Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

- GOES – Imager
- COMS – MI
PREVIOUS STUDIES
- Communication, Ocean and Meteorological Satellite (COMS)

- Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

<table>
<thead>
<tr>
<th></th>
<th>COMS KMA algorithm</th>
<th>GOES algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>POD (%)</td>
<td>57.89</td>
<td>59.90</td>
</tr>
<tr>
<td>FAR (%)</td>
<td>99.02</td>
<td>5.51</td>
</tr>
</tbody>
</table>

POD: Probability Of Detection  
FAR: False Alarm Rate
1. GOAL OF RESEARCH
   - To develop icing detection models using COMS and Himawari-8 based on machine learning approaches

2. RESEARCH PROCESS
   - Icing reference data
     - Relied only on the PIREPs data as reference
   - Cloud-related variables are determined
     - L1B data, Cloud analysis data, and Upper atmospheric variables
   - Machine learning approaches
     - Decision Trees(DT), Random Forest(RF), Support Vector Regression(SVR)

3. THREE ICING MODELS
   1) Icing masking model using COMS data
   2) Icing altitude estimation using COMS data
   3) Icing masking model using Himawari-8 data
Machine Learning is a sort of the artificial intelligence (AI). Machine learning develops a model that learns from and makes prediction of data.

### 1. Decision Trees (DT)

- Leaf nodes
- Split nodes

### 2. Random Forest (RF)

Average prediction

\[(0.23 + 0.19 + 0.34 + 0.22 + 0.26 + ... + 0.31) / \# \text{Trees} = 24\]

### 3. Support Vector Regression (SVR)
1) Icing masking model based on COMS
   - Data and Methodology

- Reference dataset was prepared based on the PIREPs
  - Consisted of 22 icing sites and 169 non-icing sites acquired from PIREPs between 1 Apr 2011 and 5 Sep 2015
- Input variables from the Level-1b and Level-2 data

<table>
<thead>
<tr>
<th>Level-1b</th>
<th>Level-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visible</td>
<td>Cloud Optical Thickness; COT</td>
</tr>
<tr>
<td>Shortwave Infrared (SWIR)</td>
<td>Cloud Top Temperature; CTT</td>
</tr>
<tr>
<td>Water Vapor (WV)</td>
<td>Cloud Top Height; CTH</td>
</tr>
<tr>
<td>Infrared1 (IR1)</td>
<td>Upper Tropospheric Humidity; UTH</td>
</tr>
<tr>
<td>Infrared2 (IR2)</td>
<td>Cloud Effective Radius; CER</td>
</tr>
<tr>
<td>BTD1 (SWIR - IR1)</td>
<td>Cloud Phase; CP</td>
</tr>
<tr>
<td>BTD2 (IR1-IR2)</td>
<td></td>
</tr>
</tbody>
</table>

Excluded for training data through the result of tests
1) Icing masking model based on COMS

- Results

- Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches

Random Forest

Decision Trees

KMA algorithm

Non-icing
Icing
2) Icing altitude estimation based on COMS

- Data and Methodology

- Cloud Top Temperature (CTT), Cloud Top Pressure (CTP), Cloud Top Height (CTH), and Cloud Optical Thickness (COT) are related to the internal properties of clouds such as temperature and particles.

- Distribution of cloud internal properties is related with the vertical icing potential.

- Altitude of icing from the PIREPs as a dependent variable.

- Input variables: L1B and CTT, CTP, CTH, and COT data.

- Modeling approach: Support Vector Regression (SVR).
2) Icing altitude estimation based on COMS

- Result

Calibration

R² = 0.9746
RMSE = 0.54 km

Validation

R² = 0.9739
RMSE = 0.55 km

Errors are generally within ±300m vertically, which is the significant level of icing from PIREPs.
3) Icing masking model based on Himawari-8

- Himawari-8

- Geostationary satellite of JMA, launched in October, 2014
- Provide data from July 2015
- Images of 16 channels are provided for weather observations and environmental monitoring
- Spatial resolution ranges from 0.5km to 2km
- Temporal resolution ranges from 0.5min to 10min
- Has similar channel characteristics with GEO-KOMPSAT-2 (GK-2A), so it is good proxy data for GK-2A

<table>
<thead>
<tr>
<th>Channel</th>
<th>Centerwavelength[μm]</th>
<th>Bandwidth[μm]</th>
<th>Resolution[km]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.4703</td>
<td>0.0407</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>0.5105</td>
<td>0.0308</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.6399</td>
<td>0.0817</td>
<td>0.5</td>
</tr>
<tr>
<td>4</td>
<td>0.8563</td>
<td>0.0345</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1.6098</td>
<td>0.0409</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2.257</td>
<td>0.0441</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>3.8848</td>
<td>0.2006</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>6.2383</td>
<td>0.8219</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td>6.9395</td>
<td>0.4019</td>
<td>2</td>
</tr>
<tr>
<td>10</td>
<td>7.3471</td>
<td>0.1871</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>8.5905</td>
<td>0.3727</td>
<td>2</td>
</tr>
<tr>
<td>12</td>
<td>9.6347</td>
<td>0.3779</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td>10.4029</td>
<td>0.4189</td>
<td>2</td>
</tr>
<tr>
<td>14</td>
<td>11.2432</td>
<td>0.6678</td>
<td>2</td>
</tr>
<tr>
<td>15</td>
<td>12.3828</td>
<td>0.9656</td>
<td>2</td>
</tr>
<tr>
<td>16</td>
<td>13.2844</td>
<td>0.5638</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Observations per timeline</th>
<th>Time cycle [min.]</th>
<th>Observations per day</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Disk</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Japan Area</td>
<td>4</td>
<td>2.5</td>
</tr>
<tr>
<td>Target Area</td>
<td>4</td>
<td>2.5</td>
</tr>
<tr>
<td>Landmark Area</td>
<td>20</td>
<td>0.5</td>
</tr>
<tr>
<td>Landmark Area</td>
<td>20</td>
<td>0.5</td>
</tr>
</tbody>
</table>
3) Icing masking model based on Himawari-8

- Data and Methodology

- Reference dataset
  - Consists of 2 icing sites and 7 non-icing sites acquired from PIREPs between 1 Jul 2015 and 31 Aug 2015

- Input variables: 16 channels from full disk images

- Very limited number of samples during 2 months
3) Icing masking model based on Himawari-8

- Result

Random Forest  
2015.08.23 09:00 (KST)

Himawari-8
2015.08.23 09:00 (KST)

Decision Tree
2015.08.23 09:00 (KST)

COMS
2015.08.23 09:00 (KST)

- Icing detection from geostationary satellite data over Korea and Japan using machine learning approaches
CONCLUSION

- Icing masking model based on COMS by DT & RF approaches
  - Similar patterns by two models
  - Decision trees estimated icing more than random forest.

- Icing altitude model based on COMS by SVR approach
  - Errors are generally within ±300m vertically, which is the significant level of icing from PIREPs.

- Icing masking model based on Himawari-8 by DT & RF approaches
  - Similar patterns by two models
  - Much more icing areas were produced from Himawari-8 than COMS.

- Very limited amount of data based solely on PIREPs as reference
  - More PIREPs will be available in the future, but might not be sufficient for modeling
  - Will investigate if the CloudSat Icing Potential (CLIP) algorithm based on cloud type and vertical profile of temperature can be further improved to provide more reliable icing masks.
Thank you

Intelligent Remote sensing and geospatial Information Systems (IRIS)

School of Urban and Environmental Engineering
Ulsan National Institute of Science and Technology, Ulsan, S. Korea

UNIST-gil 50, Ulsan 689-798, Republic of Korea
Tel : +82 52 217 2887
E-mail : iris-lab@unist.ac.kr