Evaluating atmospheric icing forecasts with ground-based ceilometer profiles

Karoliina Hämäläinen1 | Anne Hirsikko1 | Ari Leskinen2,3 | Mika Komppula2 | Ewan J. O’Connor1,4 | Sami Niemelä1

1Finnish Meteorological Institute (FMI), Helsinki, Finland
2Finnish Meteorological Institute (FMI), Kuopio, Finland
3University of Eastern Finland, Kuopio, Finland
4University of Reading, Reading, UK

Correspondence
Karoliina Hämäläinen, Finnish Meteorological Institute (FMI), Helsinki, Finland.
Email: karoliina.hamalainen@fmi.fi

Funding information
The study was mainly funded by a grant awarded by the Vilho, Yrjö and Kalle Väisälä Foundation supporting the dissertation work of Karoliina Hämäläinen. Additionally, Anne Hirsikko and Karoliina Hämäläinen received financial support from the Finnish Scientific Advisory Board for Defence for the project “Observations of Atmospheric Icing Conditions with Ground-Based Remote Sensing Techniques”.

Abstract
Atmospheric icing can cause severe damage and danger for many different sectors of society, and accurate forecasts of icing can help mitigate this. The aim of the study is to demonstrate how new ceilometer-based icing observations can be used in icing model verification. The study focuses on icing in the lowest 2 km of the atmosphere. The lowest 400 m of the atmosphere is essential for wind energy production and electricity power lines, while the higher elevations are essential from the aviation point of view. For aviation, the most risky situations are take-offs and landings. The ceilometer observations provide a new way to investigate atmospheric icing. Ceilometers provide vertical profiles of clouds which can be further processed in order to identify in-cloud icing conditions. The new type of ceilometer-based icing observations enables more detailed verification of icing forecasts both in the vertical and in terms of geographical location. Overall, the model shows good skill in predicting clear-sky situations. The verification results show that the icing model performs better for inland areas than for coastal areas, which is attributed to the tendency of the model to forecast clouds, and in-cloud icing, at higher altitudes than observed.

KEYWORDS
atmospheric icing, ceilometer, icing forecast, rime ice

1 | INTRODUCTION

Atmospheric icing is a severe risk for aviation safety and can cause financial losses for wind energy producers and electricity network operators. Aircraft dynamics degrade as ice accumulates on the wings, reducing lift, and is especially dangerous during landing and take-off (Bernabò et al., 2015). Wind turbine blades suffer similar aerodynamic degradation (Kraj and Bibeau, 2010), causing power production loss. Hoar frost icing on power lines increases the corona discharge on power lines, reducing the efficiency of the electricity network (Yin et al., 2016). Furthermore, heavy ice loads on cables can exceed the structural limits of the transmission towers (Nygaard et al., 2016).

To support these vulnerable infrastructures, many institutes and companies have developed forecasts to issue warnings of icing events, traditionally for aviation (FMI, 2020; NOAA-NCEP, 2020). However, during the past 10 years, wind energy production has been growing...
in cold climate regions, and the demand from the energy sector for similar products has greatly increased. Many institutes have answered the demand for accurate icing forecasts tailored to the energy sector (Kjeller Vindteknikk, 2020; Meteotest, 2020; SMHI, 2020). Typically, such forecasts are based on weather forecasts provided by the local operational numerical weather prediction (NWP) model with the actual icing forecast generated by post-processing the NWP output with dedicated icing process models (Makkonen, 2000; Molinder et al., 2018; Scher and Molinder, 2019). The choice of icing process model depends on the demand of the particular forecast problem. When forecasting for aviation or wind energy production purposes, on-off or instantaneous icing intensity information are typically sufficient. However, in estimating products such as wind energy potential, more detailed information on ice accumulation over a longer period of time is needed (Hämäläinen and Niemelä, 2017).

The present study focuses on rime ice accumulation. Rime ice is formed when liquid water particles, such as cloud droplets, collide with a cold surface (Makkonen, 2000). The liquid particles freeze immediately on contact with the cold surface, forming rime ice on the wind side of the structure. For supercooled liquid water droplets, the risk for ice accumulation is even larger and clouds containing supercooled liquid water droplets are very common in the mid to high latitudes. Furthermore, Westbrook and Illingworth (2011) have shown that super-cooled liquid water droplets can exist at temperatures as low as −37°C.

Measuring icing is very challenging. There are several instruments (e.g. Goodrich, 2020; Labkotec, 2020) developed for detecting icing, but they have limitations in severe icing events. Many of these advanced instruments are based on a heating cycle where the instruments measure icing and then melt the accumulated ice by heating itself in order to start a new measurement cycle. However, in severe icing events, the instrument supporting structures start to accumulate ice, too, which eventually covers the whole instrument and the heating is not sufficient to melt the accumulated ice to commence the next measurement cycle. Nevertheless, such instruments are widely used in wind parks and at airports. Another disadvantage, from the model verification point of view, is that these instruments provide point measurements representing a limited volume, both spatially and in the vertical. Icing, as a phenomenon, is very heterogeneous spatially and also depends on the vertical distribution of clouds and temperature. Therefore, to enable verification of the icing model in many applications, more extensive information on the vertical structure of the icing process and more extensive geographical coverage of icing observations is required.

Icing can also be detected by satellites (Minnis et al., 2004). Geostationary and polar-orbiting meteorological satellites provide good spatial and temporal coverage, but limited vertical resolution. If there are multiple cloud layers, most satellites will not see the lowest levels of the atmosphere. Hence, satellites provide important information, such as cloud top and liquid water layers, for in-flight icing conditions in the upper troposphere (Ellrod and Bailey, 2007), but are not as reliable for activities taking place closer to the surface, such as energy sector operators or aircraft take-off and landing.

The study investigates the usability of ground-based profiling ceilometer observations for icing model verification. Such novel observations potentially provide the urgently required information on the vertical structure of the icing process. Furthermore, ceilometer networks are widespread and relatively dense, with more stations than the sparse in-situ point measurements on towers or masts. Previously, similar active remote-sensing methods have been tested by NASA (Reehorst et al., 2005, 2006; Johnston et al., 2011), using a combination of ceilometer, microwave radiometer and radar, to detect icing layers. Dual-polarization radars have also been used to identify supercooled liquid water droplets in clouds (Bernabò et al., 2016). Furthermore, Harstveit (2009) developed a method to estimate icing from cloud observations using air temperature and wind speed measurements from a mountain site combined with Metar-data from a nearby airport. Here we use ceilometer-only observations from the ground-based remote-sensing network operated by the Finnish Meteorological Institute (FMI) (Hirsikko et al., 2014) together with novel cloud detection and classification algorithms (Hirsikko et al., 2019) to identify liquid water clouds potentially causing icing, and compare these results against an operational atmospheric icing model forecast. This will pave the way for the wider use of ceilometer observations, taking advantage of the full capacity of such networks.

The remainder of the paper is structured as follows. Section 2 describes the modelling methods and observations that were used, together with the verification scores used to compare the performance of the new observations and the operational icing model. Section 3 presents the main results, which are then discussed in Section 4. Finally, the conclusions are presented in Section 5.

2 METHODS

2.1 The NWP model Harmonie-AROME

The Harmonie-AROME (Harmonie = HIRLAM–ALADIN Research on Mesoscale Operational NWP in
Euromed, using AROME physics) is the operational weather forecasting model used in 10 different European meteorological services (Bengtsson et al., 2017), including the FMI. The model is continually developed as part of the shared ALADIN-HIRLAM system, which is a joint collaboration between 26 countries from Europe and North Africa (Termonia et al., 2018). The Harmonie-AROME is a convection-permitting model using non-hydrostatic dynamics (Bubnová et al., 1995; Benard et al., 2010) with the fully compressible Euler equation system (Simmons and Burridge, 1981; Laprise, 1992). This equation set is solved in space and time with a semi-Lagrangian advection scheme and a semi-implicit two-time-level scheme, originating from the European Centre for Medium-Range Weather Forecasts’ (ECMWF) global model (ECMWF, 2015). For a detailed description of the physical parametrization schemes used in the Harmonie-AROME, see Seity et al. (2011) and Bengtsson et al. (2017). An outline of the most essential parameterization schemes from the point of view of the study, the cloud microphysics and turbulence schemes, is presented herein.

The Harmonie-AROME employs a bulk cloud microphysics scheme to handle water vapour, liquid water content and ice water content. The ICE3 scheme (Pinty and Jabouille, 1998; Lascaux et al., 2006) treats five different types of hydrometeors as prognostic values: three classes of solid hydrometeors separated into cloud ice, dry snow and a combination of graupel and hail; and two liquid hydrometeor classes, cloud liquid water and rain. The ICE3 scheme accounts for hydrometeor horizontal advection and vertical sedimentation (Bouteloup et al., 2011), and the evaporation, sublimation and other interactions between the water phases.

The turbulence scheme is called HARATU, Harmonie with RACMO Turbulence (van Meijgaard et al., 2012). The scheme has a prognostic equation for turbulent kinetic energy (TKE), which includes source and sink terms: wind shear (+), buoyancy (±), transport (±) and dissipation of the TKE (−). The turbulent mixing length is described by Lenderink and Holtslag (2004), and is applied in two parts: in stable conditions and in a near-neutral convective condition. The stability coefficients (also described by Lenderink and Holtslag, 2004) take into account moist processes, such as condensation and the evaporation of cloud droplets, and their effect on latent heat. Thus, the turbulence processes affect on cloud formation and wind speed.

The present study uses version cy40h1.1 of the Harmonie-AROME model configuration (Bengtsson et al., 2017) with a horizontal grid spacing of 2.5 km and 65 levels in the vertical. The lowest model level is at 12 m above the surface, and the model top at 10 hPa. There are 27 model levels within the lowest 2 km, which is the region of interest for the study. The internal model time step is 75 s. The lateral boundary conditions are taken from the ECMWF’s global forecasts. The model is run in data-assimilation cycles, where new observations are included every 3 hr to improve the initial conditions, to simulate the operational usage of the model. The data-assimilation process uses observations from radio-soundings, aircraft and surface weather stations, both the 3D-Var and Optimal interpolation used in Harmonie-AROME are described in (Gustafsson et al., 2018). Once a day, a longer forecast is run from 0000 UTC to +24 hr, and hourly output from this forecast run is used as input into the separate icing model in order to produce the icing forecasts.

### 2.2 Icing model

The FMI operational icing model is based on the algorithm of Makkonen (2000) and calculates the rime ice accumulation in terms of a 1 m-long freely rotating standard cylinder (ISO, 2017), which has a diameter of 3 cm and is vertically oriented. Originally this model was designed for high-voltage power lines, hence the choice of diameter, and the results can be used directly for electricity cables. Our previous studies on wind turbine icing (Turkia et al., 2013; Hämäläinen and Niemelä, 2017) also show that icing intensities > 10 g·hr⁻¹ begin to affect wind turbine blade aerodynamics and cause power loss.

In this model, the diameter of the ice-covered cylinder can increase with ice accumulation and decrease by melting (melt). For a cylinder presenting a surface area, A (m²), into the wind, the modelled icing rate (g·s⁻¹) is the rate of change of ice mass, M, over time, t, given by:

\[ \frac{dM}{dt} = \alpha_1 \cdot \alpha_2 \cdot \alpha_3 \cdot \text{LWC} \cdot A \cdot V - \text{melt}, \]  

(1)

where \( \alpha_1 \), \( \alpha_2 \) and \( \alpha_3 \) are unitless collision, sticking and accretion co-efficients, respectively; LWC is the liquid water content (g·m⁻³) of the water droplets impacting the cylinder; and \( V \) is the wind speed (m·s⁻¹). The LWC and \( V \) are provided by the NWP model, together with temperature, air pressure and relative humidity. In the Harmonie-AROME, the LWC refers to the total water mass, including both cloud liquid water and liquid precipitation. Furthermore, the icing model includes important internal parameters related to the assumed droplet size distribution. The cloud droplet particle concentration (Nd) was set to 100 (cm⁻³) (Hämäläinen and Niemelä, 2017) and the droplet median volume diameter...
(MVD) is defined in terms of a Gaussian distribution of the LWC as a function of the Nd.

The unitless collision, sticking and accretion co-efficients characterize the properties of the impacting droplets. The collision co-efficient \( \alpha_1 \) is sensitive to the particle MVD and varies between 0 and 1. Large particles have a value of 1, meaning that the particles hit the object due to their inertia, whereas small light particles follow the wind streamlines more readily, and are therefore more likely to pass around the object (cylinder).

The sticking co-efficient \( \alpha_2 \) describes the likelihood that particles stick to the icy surface on impact. This is set to 1 when the air temperature is < 0°C to indicate that all liquid droplets are super cooled and will freeze immediately on hitting a cold object.

The accretion co-efficient \( \alpha_3 \) takes into account the cylinder surface energy balance. In dry growth (all particles freeze) \( \alpha_3 \) is set to 1, but for wet growth it is < 1. During wet growth there is a liquid water layer on top of the cylinder and the icing rate is controlled by latent heat release from the surface of the object. The heat balance of the icing surface in the model is affected by sensible heat loss, heat loss due to evaporation and/or heat loss due to long wave radiation. These mechanisms are described in more detail in ISO Standard 2000 (ISO, 2017).

Finally, the melt term takes into account the melting and shedding of ice and is expressed as:

\[
melt = (Q_h + Q_e + Q_l)\cdot S_f,
\]

where \( Q_h \) is the sensible heat flux; \( Q_e \) is the latent heat flux; \( Q_n \) is the net radiation; and \( S_f \) is an empirical ice-shedding factor. The energy balance at the cylinder surface takes these three terms into account, \( Q_h, Q_e \) and \( Q_n \), and melting starts when the temperature is > 0°C. The net radiation term is obtained through the combination of long wave \( (Q_l) \) and short wave radiation \( (Q_n) \), and, for the study, \( Q_n \) was set to 0 due to the challenge in determining the surface albedo during cloudy and dim conditions in winter. The empirical ice-shedding factor, \( S_f \), was set to 3.5 based on our previous studies and it speeds up ice loss during melting, mimicking how ice falls off structures in the real environment.

2.3 Ceilometer observations

The FMI operates a network of ceilometers distributed throughout Finland. Ceilometers are low-power, single-channel lidar systems traditionally used for monitoring the cloud ceiling in order to track cloud layer heights and provide cloudiness. Over 20 ceilometers in the FMI network now also routinely collect vertical profiles of the attenuated backscatter co-efficient, from which it is possible to discern the presence of aerosol, liquid cloud, ice cloud or precipitation in the form of rain or snow. The information content of these high vertical and temporal resolution vertical profiles shows great potential for developing new innovative data products such as hydrometeor classification.

Ceilometers have already shown their capability for detecting liquid clouds, including supercooled liquid layers (Hogan et al., 2003; Van Tricht et al., 2014), with mature algorithms for liquid cloud detection being incorporated into operational processing schemes (e.g. Cloudnet; Illingworth et al., 2007). The identification of liquid-only layers also forms part of the ceilometer-attenuated backscatter co-efficient calibration methods (O’Connor et al., 2004; Hopkin et al., 2019). Once a cloud layer containing liquid water droplets has been identified, the layer can then be classified as warm or supercooled based on the wet-bulb temperature at the cloud base. The wet-bulb temperature can be calculated using the meteorological profile information available from the NWP models; here we used output from the Global Data Assimilation System (GDAS) as an independent data set. The expected uncertainty in the height of the freezing isotherm is about ±150 m (Mittermaier and Illingworth, 2003). Recent ceilometer algorithm development has extended the capability to identify and distinguish between fog, precipitation and ice clouds (Tuononen et al., 2019), but currently classification of the phase of the precipitation from ceilometer profiles alone is not yet reliable.

Hence, by combining the Tuononen et al. (2019) algorithm with meteorological profiles, we have a hydrometeor classification scheme suitable for identifying hydrometeors liable to cause icing. For the study, we used attenuated backscatter profiles from six stations (Table 1). The station locations are shown in Figure 1.

2.4 Comparison between icing model forecast and observations

The observations used in the comparison were measured at Utö, Hyttylä, Kumpula, Sodankylä, Savilahd and Vehmashäki during one icing season (December 4, 2016–March 31, 2017). The ceilometer measurement frequency is < 30 s, whereas the icing model depends on the output frequency of the NWP model, and thus is typically 1 hr. For our model-observation comparison we used the same 1 hr time window for both data sets, with ceilometer observations collected within a time window of ±30 min centred on the hour. A time window was marked as an icing event if > 30% of the observed
ceilometer profiles within this time window indicated icing. The length of time window and required fraction of icing observations were chosen based on sensitivity tests performed on one month of observations from Vehmasmäki. These tests (data not shown) indicated that the comparison results were not particularly sensitive to the choice of time window and required fraction.

The model-based icing events were defined by the icing intensity produced by the icing model. The sensitivity tests showed that the choice of icing intensity threshold was the dominant factor in defining icing events from the model output. Previous studies (Hämäläinen and Niemelä, 2017) have shown that if the icing intensity is > 10 g·hr⁻¹, it impacts wind power production and aviation safety (RAYTHEON, 1999). Since the ceilometer is very sensitive to the presence of any liquid water in the atmosphere (Vaisala, 2020) and detects clouds with a wide range of liquid water content (e.g. Gierens et al., 2020), the ceilometer can detect potential icing events with much lower icing intensities than the threshold suggested above. Thus, in order to generate a representative comparison between the model and observations, an icing intensity threshold of 1 g·hr⁻¹ was selected, that is, a case was denoted an icing event if the model icing intensity was > 1 g·hr⁻¹.

This process creates a binary time series (icing present, no icing present) from the model and ceilometer, which can then be directly compared. As the ceilometer provides profiles with 10–30 m vertical resolution (Table 1), the observations closest in altitude to each model level were selected and the results for the lowest 2 km in altitude are presented in the paper. Levels above this were also investigated, but the comparison begins to suffer from the lack of observations; liquid water layers rapidly attenuate the ceilometer laser beam so that, typically, only the first cloud layer can be detected with no information available from above the first layer. However, the majority of applications are interested in the region < 2 km, and this is the first time we have been able to study the vertical profile of icing.

### 2.5 Verification scores

The performance of the model was evaluated using a set of verification skill scores. First, a 2 × 2 contingency table based on the observed and forecast icing events was generated (Table 2). For each event, a binary score, “Yes” or “No” is given for both the observation and the forecast, using the selected icing intensity threshold of 1 g·hr⁻¹. This gives four options for the contingency table: the variable (1) means a hit, as both observations and forecast

![Ceilometer locations in Finland](image-url)

**TABLE 1** Ceilometer sites and specifications

| Site   | Latitude  | Longitude | Location type | Instrument type | Vertical resolution (m) | Base elevation (m) |
|--------|-----------|-----------|---------------|-----------------|------------------------|-------------------|
| Utö    | 59.78     | 21.37     | Coastal       | Vaisala CL31    | 10                     | 8                 |
| Kumpula| 60.33     | 25.60     | Coastal       | Vaisala CL51    | 10                     | 26                |
| Hyytiäla| 61.85     | 24.29     | Inland        | Vaisala CL31    | 30                     | 179               |
| Vehmasmäki| 62.74    | 27.54     | Inland        | Vaisala CL51    | 10                     | 190               |
| Savilahti| 62.89     | 27.63     | Inland        | Vaisala CT25K   | 30                     | 90                |
| Sodankylä| 67.37     | 26.63     | Lapland       | Vaisala CL31    | 30                     | 181               |

**FIGURE 1** Ceilometer locations in Finland

[Correction added on 4 December 2020, after first online publication: ‘Sweden’ amended to ‘Finland’ in the Figure 1 caption, correcting a copy-editing error in the originally published article.]
TABLE 2  Contingency table

|       | Observations |
|-------|--------------|
|       | Yes | No |
| Forecast | Yes | a   | b  |
|         | No  | c   | d  |

indicate “Yes” at the same time, (2) means a false alarm, where icing was forecast but not observed, (3) is a miss, where icing was observed but not forecast, and (4) means a correct rejection, where icing was not observed or forecast. A set of verification skill scores can then be derived from the contingency table (Wilks, 2006).

Frequency bias (Bias) calculates the ratio between correctly predicted cases and observed cases:

$$\text{Bias} = \frac{a + b}{a + c}.$$  \hspace{1cm} (3)

which varies between 0 and $\infty$, with a value of 1 describing a perfect forecast. Values < 1 indicate that the system is under-forecasting the number of icing events, and > 1 indicate a system that is over-forecasting.

Proportion correct (PC), given by:

$$\text{PC} = \frac{a + d}{a + b + c + d}.$$  \hspace{1cm} (4)

is influenced by the number of correctly predicted and correctly rejected forecasts and provides the fraction of correctly forecast cases from the total number of cases. A value of 1 indicates perfect skill and 0 indicates no skill.

Similar to PC, is the threat score (TS),

$$\text{TS} = \frac{a}{a + b + c},$$  \hspace{1cm} (5)

which is also known as the critical success index, ranging from 0 to 1, with 1 meaning a perfect forecast and 0 indicating no skill. The TS identifies the fraction of correctly forecast cases (hits) relative to all forecast and/or observed cases, that is, the PC after removing correctly rejected cases.

The probability of detection (POD):

$$\text{POD} = \frac{a}{a + c},$$  \hspace{1cm} (6)

also termed hit rate, describes the proportion of correctly forecast cases out of all observed cases, with values from 0 to 1, with 1 being a perfect score. This score is sensitive to misses, but not to false alarms, and means that the POD should not be used in isolation, but together with other skill scores that investigate false predictions. One such score is the false alarm ratio (FARatio):

$$\text{FARatio} = \frac{b}{a + b}.$$  \hspace{1cm} (7)

providing the fraction of “false positive” events versus all forecast events. This score is sensitive to false alarms, but neglects misses. The FARatio varies between 0 and 1, with 0 indicating a perfect score.

From the ceilometer output we can identify the layers from which we obtain a proper signal. This information was used to mask out the situations with no observation data available. This quality control took place before model verification to exclude the cases with no observation.

3 | RESULTS

3.1 | Icing as a function of height and geographical location

The ceilometer observations used to verify the icing model represent different geographical and climatic environments in Finland, providing the potential to investigate possible regional differences in the model predictive skill. Furthermore, vertical profile observations allow the investigation of the model performance in predicting the vertical structure of icing events.

Figure 2 shows (1) the total number of investigated cases (observed-forecast pairs of icing/no-icing events), (2) the number of observed icing cases and (3) the number of icing events forecast in the lowest 2 km for each ceilometer location. Figure 2a gives the data availability at each station, that is, where the ceilometer signal is deemed to be above the threshold suitable for the specific instrument at each site. This data availability includes all cases where aerosol is observed, hence data availability is expected to be greater for the more sensitive instrument (CL51) and in locations where more aerosol is expected. The total number of cases for the three inland stations (Savilahti, Hyytiälä and Sodankylä) decreases rapidly as a function of height (Figure 2a). In Savilahti, this decrease in the number of cases stabilizes at about 500 m, whereas in Hyytiälä and Sodankylä, the decrease continues and stabilises > 1 km. This decrease is attributed to the detection capability for the CL31 and CT25K ceilometers in the low aerosol concentration conditions in Finnish inland areas. The common quality control algorithm applied by the FMI to all ceilometers to discriminate between signal and noise has not performed as well for
the CT25K instrument as for the CL31 instruments and is likely responsible for the difference between the CL31 and CT25K data availability profiles. This does not impact the detection of icing events, which use signals much further above the detection threshold. In Utö, Kumpula and Vehmasmäki, the profiles behave differently. The total number of cases remains constant from the surface upwards. At Utö, a location with marine aerosol present, the number of cases starts to decrease between 500 and 1,000 m, whereas in Vehmasmäki and Kumpula with the more sensitive CL51 ceilometer, the decrease in data availability does not start until > 1,000 m in altitude.

Figure 2b shows that, in the lowest 500 m, the ceilometer at the coastal stations (Utö and Kumpula) detects less icing cases than inland and Lapland stations. The proportion of icing events from all cases in coastal areas (Figure 2) varies between 12% and 30%, depending on height. Furthermore, at coastal stations, the number of observed icing events peaks at altitudes close to 400 m. However, for inland stations, the proportion of icing events may reach even 45% at < 500 m, peaking at 200–300 m. For all stations, the number of observed icing events rapidly decreases above the peak height (200–400 m). The rapid decrease in cloud observations as a function of height is expected since any cloud layer above the lowest liquid water layer will not be observed due to full attenuation of the ceilometer beam.

Figure 2c shows the number of forecast icing events. In coastal areas (Utö and Kumpula), the forecast profiles clearly differ from observations. The peak in icing events is generally forecast to be at higher levels (550–750 m) than the observations indicate. Consequently, there are much less predicted icing events in the lowest 500 m than observed. Furthermore, the model predicts almost three times more icing cases than observed > 1 km at all stations. However, the results presented in Figure 2 are the total number of cases and may indicate how many observations we still might be missing even with this new measurement technique.

3.2 | Verification results

Figure 3 displays the profiles of the Bias, POD, FARatio and PC in different geographical areas (coastal, inland, Lapland). For inland locations, the Bias is close to unity at < 500 m altitude, indicating good model performance in predicting the number of icing events. However, the model is under-forecasting in coastal locations and the Bias decreases towards the surface. Furthermore, as already seen in Figure 2, the Bias indicates that the model tends to over-forecast the number of icing events at all stations, especially at > 1 km.

Profiles of the POD and FARatio are presented in Figures 3b, c, respectively. The POD is best in inland areas and within the lowest 500 m, with the highest scores obtained at 300 m altitude. Moreover, the FARatio is also slightly better in inland areas than in coastal areas, although differences in this score between regions and heights are rather small. Coastal stations demonstrate much poorer POD scores and the lowest value (0.18) is obtained close to the surface, indicating poor model capability in predicting very-low-altitude icing events accurately.

Figure 3d presents the PC profiles. The overall skill to forecast correctly both icing events and non-icing events is almost equally good for all regions. Furthermore, the PC score does not vary much (0.65–0.80) as a function of height. Coastal areas show slightly better PC scores near the surface, which is mainly due to the correct forecasting of non-icing events, which are more
frequent in coastal regions than inland and Lapland regions.

The performance diagram (Roebber, 2009) combines four scores (Figure 4): the POD, success ratio \((1 - \text{FARatio})\), Bias and TS. The latter, also known as the critical success index, describes the balance between the POD and success ratio. In an ideal forecast, the dots would congregate in the top-right corner of the performance diagram, having POD = 1 and SR = 1 with Bias and TS = 1. The results (Figure 4) indicate that inland and Lapland areas obtain the best forecast skill in a layer between 100 and 400 m (circled with a black line). The overall skill of forecasting icing events at coastal areas is much worse with a lower POD and TS than inland and Lapland regions. All stations have almost equally good/poor skill at higher levels (1–2 km).

4 DISCUSSION

The results indicate that the ability of the icing model varies between different regions and also as a function of height. This section highlights the main findings of the results.

The icing process in the model is driven by two main factors: temperature and the amount of liquid water. Therefore, any error in these parameters will directly influence the model’s capability to predict icing events. In general, clouds are more difficult to predict spatially and temporally than temperature (Haiden et al., 2015). Consequently, the icing verification presented in the present study will mostly give information on the NWP model performance in forecasting cloud. However, according to the operational NWP verification performed by the FMI, the Harmonie-AROME NWP model tends to have a
minor near-surface cold bias (< 1°C). Therefore, part of the wrongly predicted missing icing cases might be due to errors in incorrect temperature predictions, and conditions where temperatures are close to 0°C are especially prone to temperature-related errors in predicting icing events.

The results show that there appears to be two different profile shapes regarding the profile of the total number of cases. One is where the number of cases decreases rapidly as a function of height, stabilizing around 1 km; and the second is where stations display a constant number of cases up to 500–1,000 m and then there is a suddenly a rapid decrease. This behaviour can be explained partly by the station location, with coastal stations exhibiting the second type of behaviour. However, Vehmasmäki station makes an exception.

The FMI operates three different ceilometer models, which have differing levels of performance with respect to aerosol observations. However, all models have more than enough sensitivity to liquid water clouds, especially when < 2 km in altitude, and no model will see through the first completely attenuating layer.

The results indicate that the model behaves better for inland stations. This behaviour can partly be explained by the number of cases, where, for inland stations, there is almost an equal number of modelled and observed cases. The number of cases forecast in coastal areas at < 500 m is too low, resulting in poor skill as indicated by the POD and Bias, but this does not have a clear impact on the FARatio, which remains quite similar at locations. This clearly indicates that the model tends to miss actual icing events more often in coastal areas. Overall, the skill in predicting icing cases decreases rapidly at > 500 m due to over-forecasting compared with available observations.

In the lowest 300 m in coastal areas, the model predicts only half the observed icing cases. The maximum number of modelled and observed icing cases is very similar, but the altitudes are different with the maximum in the model around 600 m and the observations peaking at 400 m. This altitude shift in the model-produced icing clouds has a direct effect on verification scores and is probably due to the poor ability of the NWP model to predict low-level clouds in coastal regions.

Although coastal areas seem to exhibit the worst scores in terms of the POD and Bias, non-icing cases are forecast equally well or even slightly better (Figure 3d) than elsewhere, which is because the Bias, POD or FARatio do not take into account correctly rejected cases. This indicates that the model has skill in forecasting clear-sky cases and cloudy cases with temperatures above freezing. This finding is also in line with the low number of observed icing events in coastal regions. Overall, the PC varies between 70% and 80% for all stations, indicating that the model has good skill in distinguishing between cloudy and non-cloudy cases in general.

The Sodankylä station in Lapland was studied separately from the rest of the inland stations due to its northern location and colder climate. In Sodankylä, the number of cases predicted by the model is quite similar compared with observations in the lowest 400 m. However, the model over-predicts the number of icing events at higher levels. Furthermore, the number of observed icing events in Sodankylä is lower compared with other inland stations, most probably a consequence of the drier air in the cold climate.

In all stations, scores were generally poorer at > 1 km. The model clearly over-predicts the number of cases, having almost twice as many icing cases as observed. However, the over-prediction above the peak in ceilometer-observed icing is to some extent related to the lack of observations in multilayered cloud situations. The ceilometer is not capable of detection in these situations, reducing the number of cases for verification and resulting in deteriorating verification scores. The addition of other active remote-sensing instruments such as cloud radars would likely improve the representativeness of icing at higher altitudes (Illingworth et al., 2007), but a dense network of such instruments is not yet practical because of cost. However, we also expect icing to decrease as a function of height as the air gets colder and drier, especially since the fraction of clouds containing supercooled liquid begins to decrease once the temperature at cloud top falls to < −20°C (Westbrook and Illingworth, 2011). Therefore, over-forecasting may also be due to a combination of over-forecasting of mixed-phase cloud in general and an under-forecasting of the fraction of ice-phase only clouds in the Harmonie-AROME at colder temperatures. Previous issues with the Harmonie-AROME predicting too much liquid in mixed-phase clouds have been noted by Müller et al. (2017) and Bengtsson et al. (2017) and corrections introduced; however, further evaluation of the liquid to ice-phase ratio in mixed-phase clouds produced by the model is likely required. In general, the model has only a moderate skill in predicting icing at higher levels.

One potential next step is to start generating icing forecasts based on ensemble predictions. Ensemble predictions provide information on forecast uncertainty, which is related to the particular weather situation. Since predicted icing is highly dependent on the uncertainty of temperature, wind speed and liquid water available in the atmosphere, ensembles would aid the understanding of the reliability of the forecast both temporally and spatially. The FMI has been running operationally a meso-scale ensemble prediction system (MEPS) in joint collaboration with the Swedish Meteorological and
Hydrological Institute, the Meteorological Institute of Norway, and the Estonian Environmental Agency (MetCoOp) (Müller et al., 2017). The MEPS is a 30-member lagged ensemble covering Northern Europe with a 2.5 km horizontal grid size and +61 hr lead time (Andrae et al., 2020). The comprehensive ceilometer observation network will be a great benefit in the verification of probabilistic cloud and icing forecasts in future.

5 | CONCLUSIONS

The aim of the study was to demonstrate how new ground-based ceilometer data product could be used in icing model verification. The vertical profiles of icing model output and a ceilometer icing product based on attenuated backscatter co-efficient profiles were compared. This enabled the investigation of the model's capability in simulating and forecasting the icing process in both the vertical dimension and different geographical regions. In general, the results show good similarity between the observations and the model. However, in coastal areas, the model tends to predict icing cases higher above the surface and in deeper layers than observed. Consequently, in these areas, the model predicts too few icing cases at < 500 m and overpredicts the number of icing cases at > 500 m. The vertical structure of icing is better captured at inland stations, with the model showing the best skill at < 500 m, which is important for wind energy and aviation applications. Clear sky conditions are in general well predicted. Improving the icing model forecast will likely require improvements in the cloud microphysics of the numerical weather prediction (NWP) model used as the input.

The study focused on icing profiles in six different locations in Finland and showed that ceilometer observations can be used to evaluate forecasts of the icing profiles. These profile observations bring added value to the verification process of the icing model compared with mast-based point observations. Since the study has taken place, the number of ceilometer-equipped weather stations in Finland providing profiles has grown to 24 stations, which will enable a more thorough study of the spatial behaviour of the icing process. There are relatively dense ceilometer networks to exploit (Illingworth et al., 2015), and the vertical profiling capability opens new ways to explore, understand and correct errors in the NWP and icing models.

ACKNOWLEDGMENTS

The authors thank the project team for end-user feedback and fruitful conversations during the project.

ORCID

Karoliina Hämäläinen 🆐 https://orcid.org/0000-0002-3449-7559

REFERENCES

Andrae, U., Frogner, I.-L., Vignes, O., Singleton, A., Azad, R., Partio, M. and Sokka, N. (2020) A continuous EDA-based ensemble in MetCoOp. ALADIN-HIRLAM Newsletter, 2020, 189–199.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., Gleeson, E., Hansen-Sass, B., Homleid, M., Hortal, M., Ivarsson, K.-I., Lenderink, G., Niemelä, S., Nielsen, K.P., Onvlee, J., Rontu, L., Samuelsson, P., Muñoz, D. S., Subias, A., Tijm, S., Toll, V., Yang, X. and Koltzow, M.O. (2017) The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system. Monthly Weather Review, 145 (5), 1919–1935.

Bernabò, P., Cuccioli, F. and Baldini, L. (2015) Icing hazard for civil aviation. In: 2015 IEEE Metrology for Aerospace (MetroAeroSpace). Benevento, Italy: IEEE, pp. 295–300.

Bernabò, P., Cuccioli, F., Baldini, L. and Chandrasekar, V. (2016) Super-cooled liquid water droplets detection using dual-polarization radar. In: 2016 IEEE Metrology for Aerospace (MetroAeroSpace). Florence, Italy: IEEE, pp. 118–123.

Bénard, P., Vivoda, J., Mašek, J., Smolíková, P., Yessad, K., Smith, C., Brožková, R. and Geleyn, J.-F. (2010) Dynamical kernel of the Aladin–NH spectral limited-area model: revised formulation and sensitivity experiments. Quarterly Journal of the Royal Meteorological Society, 136(646), 155–169.

Bouteloup, Y., Seity, Y. and Bazile, E. (2011) Description of the sedimentation scheme used operationally in all Météo-France NWP models. Tellus A, 63(2), 300–311.

Bubnová, R., Hello, G., Bénard, P. and Geleyn, J.-F. (1995) Integration of the fully elastic equations cast in the hydrostatic pressure terrain-following coordinate in the framework of the ARPEGE/Aladin NWP system. Monthly Weather Review, 123 (2), 515–535.

ECMWF (2015). Operational implementation 12 May 2015. Part III: dynamics and numerical procedures. Number 3 in IFS Documentation. ECMWF. Available at: https://www.ecmwf.int/node/9210 [Accessed 26th February, 2020].

Ellrod, G.P. and Bailey, A.A. (2007) Assessment of aircraft icing potential and maximum icing altitude from geostationary meteorological satellite data. Weather and Forecasting, 22(1), 160–174.

FMI (2020). Ilmailusää. Available at: https://www.ilmailusaa.fi/ [Accessed 6th May 2020].

Gierens, R., Kneifel, S., Shupe, M.D., Ebell, K., Maturilli, M. and Löhnert, U. (2020) Low-level mixed-phase clouds in a complex Arctic environment. Atmospheric Chemistry and Physics, 20, 3459–3481.

Goodrich (2020). 0871LH1 freezing-rain sensor. Available at: https://www.campbellsce.eu/0871lh1 [Accessed 6th May 2020].

Gustafsson, N., Janjić, T., Schraff, C., Leuenberger, D., Weissmann, M., Reich, H., Brousseau, P., Montmerle, T., Wattrelot, E., Bučánek, A., Mile, M., Hamdi, R., LindsKog, M., Barkmeijer, J., Dahlbom, M., Macpherson, B., Ballard, S., Inverarity, G., Carley, J., Alexander, C., Dowell, D., Liu, S., Ikuta, Y. and Fujita, T. (2018) Survey of data assimilation
methods for convective-scale numerical weather prediction at operational centres. *Quarterly Journal of the Royal Meteorological Society*, 144(713), 1218–1256.

Haiden, T., Forbes, R., Ahlgrimm, M., and Bozzo, A. (2015). The skill of ECMWF cloudiness forecasts. *Newsletter No. 143*, *Spring* 2015, pp. 14–19. Available at: https://www.ecmwf.int/node/14587 [Accessed 13th May 2020].

Harsteveit, K., and Vindekleivik, K. (2009) Using meter-data to calculate in-cloud icing on a mountain site near by the airport. In: *13th International Workshop on Atmospheric Icing*. Andermatt, Switzerland.

Hirsikko, A., Hämäläinen, K., Ruuskanen, A., Leskinen, A., Komppula, M. and O’Connor, E. (2019) Advances in ceilometer data processing—detecting atmospheric icing in the vertical profile. In: *99th American Meteorological Society Annual Meeting*. Phoenix, AZ: AMS. Available at: https://ams.confex.com/ams/2019Annual/webprogram/Paper351187.html [Accessed 23rd September 2020].

Hirsikko, A., O’Connor, E., Komppula, M., Korhonen, K., Pfuller, A., Giannakaki, E., Wood, C., Bauer-Pfundstein, M., Poikonen, A., Karpipinen, T., et al. (2014) Observing wind, aerosol particles, clouds and precipitation: Finland’s new ground-based remote-sensing network. *Atmospheric Measurement Techniques*, 7, 1351–1375.

Hämäläinen, K. and Niemelä, S. (2017) Production of a numerical icing atlas for Finland. *Wind Energy*, 20(1), 171–189.

Hogan, R.J., Illingworth, A.J., O’Connor, E.J. and Poiares Baptista, J.P.V. (2003) Characteristics of mixed-phase clouds. II: a climatology from ground-based lidar. *Quarterly Journal of the Royal Meteorological Society*, 129(592), 2117–2134.

Hopkin, E., Illingworth, A.J., Charlton-Perez, C., Westbrook, C.D. and Ballard, S. (2019) A robust automated technique for operational calibration of ceilometers using the integrated backscatter from totally attenuated liquid clouds. *Atmospheric Measurement Techniques*, 12, 4131–4147.

Illingworth, A.J., Cimini, D., Gaffard, C., Haefelin, M., Lehmann, V., Löhner, U., O’Connor, E.J. and Ruffieux, D. (2015) Exploiting existing ground-based remote sensing networks to improve high-resolution weather forecasts. *Bulletin of the American Meteorological Society*, 96(12), 2107–2125.

Illingworth, A.J., Hogan, R.J., O’Connor, E.J., Bouniol, D., Brooks, M. E., Delanoë, J., Donovan, P.D., Eastment, J.D., Gaussiat, N., Goddard, J.W.F., Haefelin, M., Baltink, H.K., Krasnov, O.A., Pelon, J., Piriou, J.-M., Protat, A., Russchenberg, H.W.J., Seifert, A., Tompkins, A.M., van Zadelhoff, G.-J., Vinit, F., Willén, U., Wilson, D.R. and Wrench, C.L. (2007) Cloudnet: continuous evaluation of cloud profiles in seven operational models using ground-based observations. *Bulletin of the American Meteorological Society*, 88(6), 883–898.

ISO (2017). Atmospheric icing of structures. International Standard. ISO, 12494. Available at: https://www.iso.org/standard/72443.html [Accessed 25th February 2020].

Johnston, C.J., Serke, D.J., Adriaansen, D.R., Reehorst, A.L., Politovich, M.K., Wolff, C.A. and McDonough, F. (2011) Comparison of in-situ, model and ground based in-flight icing severity. *NASA Technical Reports*, 1–12. https://ntrs.nasa.gov/search.jsp?R=20120000908 [Accessed 25th February 2020].

Kjeller Vindekleivik (2020). *Weather modelling wind & icing*. https://www.vindekleivik.com/services/wind-energy/weather-modelling-wind-icing/ [Accessed 6th May 2020].

Kraj, A.G. and Bibeau, E.L. (2010) Phases of icing on wind turbine blades characterized by ice accumulation. *Renewable Energy*, 35 (5), 966–972.

Labkotec (2020). *LID-3300IP type 2 ice detector*. Available at: https://www.labkotec.fi/en/products/ice-detection-system-for-wind-turbine/lid-3300ip-type-2-ice-detector [Accessed 6th May 2020].

Laprise, R. (1992) The Euler equations of motion with hydrostatic pressure as an independent variable. *Monthly Weather Review*, 120(1), 197–207.

Lascaux, F., Richard, E. and Pinty, J.-P. (2006) Numerical simulations of three different MAP IOPs and the associated microphysical processes. *Quarterly Journal of the Royal Meteorological Society*, 132(619), 1907–1926.

Lenderink, G. and Holtslag, A.A.M. (2004) An updated length-scale formulation for turbulent mixing in clear and cloudy boundary layers. *Quarterly Journal of the Royal Meteorological Society*, 130 (604), 3405–3427.

Makkonen, L. (2000) Models for the growth of rime, glaze, icicles and wet snow on structures. *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 358(1776), 2913–2939.

Meteor (2020). *Icing Services*. Available at: https://metoeastest.ch/en/division/icing [Accessed 6th May 2020].

Minnis, P., Smith, W.L., Jr., Nguyen, L., Khaiyer, M.M., Spangenberg, D.A., Heck, P.W., Palikonda, R., Bernstein, B.C. and McDonough, F. (2004) A real-time satellite-based icing detection system. *NASA Technical Reports Server*, 1–5. https://ntrs.nasa.gov/search.jsp?R=20040087186 [Accessed 28th May 2020].

Mittermaier, M.P. and Illingworth, A.J. (2003) Comparison of model-derived and radar-observed freezing-level heights: implications for vertical reflectivity profile-correction schemes. *Quarterly Journal of the Royal Meteorological Society*, 129(587), 83–95.

Molinder, J., Körnich, H., Olsson, E., Bergström, H. and Sjöblom, A. (2018) Probabilistic forecasting of wind power production losses in cold climates: a case study. *Wind Energy Science*, 3, 667–680.

Müller, M., Homleid, M., Ivarsson, K.-I., Køltzow, M.A.O., Lindskog, M., Midtbø, K.H., Andrae, U., Aspelien, T., Berggren, L., Bjørg, D., Dahlgren, P., Kristiansen, J., Randriamampianina, R., Ridal, M. and Vignes, O. (2017) AROME-MetCoOp: a Nordic convective-scale operational weather prediction model. *Weather and Forecasting*, 32(2), 609–627.

NOAA-NCEP (2020). *Aviation Weather Center*. Available at: https://aviationweather.cpc.ncep.noaa.gov/icing [Accessed 6th May 2020].

Nygaard, C., Ågustsson, M., Moen, L., Welgaard, Ø., and Ellasson, Å. (2016). Monitoring and forecasting ice loads on a 420 kV transmission line in extreme climatic conditions. In *Proceeding 16th International Workshop on Atmospheric Icing of Structures*, 28 June to 3 July 2015, Uppsala, Sweden.

O’Connor, E.J., Illingworth, A.J. and Hogan, R.J. (2004) A technique for autocalibration of cloud lidar. *Journal of Atmospheric and Oceanic Technology*, 21, 777–786.

Pinty, J.-P. and Jabouille, P. (1998) A mixed-phased cloud parameterization for use in a mesoscale nonhydrostatic model:
simulations of a squall line and of orographic precipitation. In: *Proceedings of the Conference of Cloud Physics*. Everett, WA: American Meteorological Society, pp. 217–220.

RAYTHEON (1999). Ice Accretion Algorithm Development. Citeseer, Raytheon, Information Technology and Scientific Services (RTSS).

Reehorst, A., Politovich, M.K., Zednik, S., Isaac, G.A. and Cober, S. (2006) Progress in the development of practical remote detection of icing conditions. *NASA Technical Reports*, 1–17. https://ntrs.nasa.gov/search.jsp?R=20060013453 [Accessed 25th February 2020].

Reehorst, A.L., Brinker, D.J., Ratvasky, T.P., Ryerson, C.C. and Koenig, G.G. (2005) The NASA icing remote sensing system. *NASA Technical Reports*, 1–18. https://ntrs.nasa.gov/search.jsp?R=20050175843 [Accessed 25th February 2020].

Roebber, P.J. (2009) Visualizing multiple measures of forecast quality. *Weather and Forecasting*, 24(2), 601–608.

Scher, S. and Molinder, J. (2019) Machine learning-based prediction of icing-related wind power production loss. *IEEE Access*, 7, 129421–129429.

Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Boudt, F., Lac, C. and Masson, V. (2011) The AROME-France convective-scale operational model. *Monthly Weather Review*, 139(3), 976–991.

Simmons, A.J. and Burridge, D.M. (1981) An energy and angular-momentum conserving vertical finite-difference scheme and hybrid vertical coordinates. *Monthly Weather Review*, 109(4), 758–766.

SMHI (2020). *Icing forecasts for wind turbines now possible*. Available at: https://www.smhi.se/en/research/research-news/icing-forecasts-for-wind-turbines-now-possible-1.85372 [Accessed 6th May 2020].

Termonia, P., Fischer, C., Bazile, E., Bouyssel, F., Brožková, R., Bénard, P., Bochenek, B., Degrauwe, D., Derková, M., El Khatib, R., et al. (2018) The ALADIN system and its canonical model configurations AROME CY41T1 and ALARO CY40T1. *Geoscientific Model Development*, 11, 257–281.

Tuononen, M., O’Connor, E.J. and Sinclair, V.A. (2019) Evaluating solar radiation forecast uncertainty. *Atmospheric Chemistry and Physics*, 19, 1985–2000.

Turkia, V., Huttunen, S., and Tomas, W. (2013). *Method for estimating wind turbine production losses due to icing*. Number 114 in VTT Technology. VTT Technical Research Centre of Finland, Finland. Project code: 72921. Available at: https://www.vttresearch.com/sites/default/files/pdf/technology/2013/T114.pdf [Accessed: 13th May 2020].

Vaisala (2020). *Ceilometer CL31*. Available at: https://www.vaisala.com/sites/default/files/documents/CL31-Datasheet-B210415EN_0.pdf [Accessed 13 May 2020].

van Meijgaard, E., van Ulft, L.H., Lenderink, G., Roode, S.R.D., Wipfler, L., Boers, R. and van Timmermans, R.M.A. (2012) Refinement and application of a regional atmospheric model for climate scenario calculations of Western Europe. *KVR Research Report*. ISBN ISBN/EAN 978-90-8815-046-3, 1–46. http://library.wur.nl/WebQuery/wurpubs/fulltext/312258 [Accessed 27th February 2020].

van Tricht, K., Gorodetskaya, I.V., Lhermitte, S., Turner, D.D., Schween, J.H. and Van Lipzig, N.P.M. (2014) An improved algorithm for polar cloud-base detection by ceilometer over the ice sheets. *Atmospheric Measurement Techniques*, 7(5), 1153–1167.

Westbrook, C.D., and Illingworth, A.J. (2011) Evidence that ice forms primarily in supercooled liquid clouds at temperatures > −27 C. *Geophysical Research Letters*, 38(14), 1–4. http://dx.doi.org/10.1029/2011gl048021.

Wilks, D.S. (2006). In: Dmowska, R., Hartmann, D. and Rossby, H.T. (Eds.) *Statistical Methods in the Atmospheric Sciences*. Elsevier, 2nd edition. Cambridge, MA: Academic Press. ISBN 13: 978-0-12-751966-1.

Yin, F., Farzaneh, M. and Jiang, X. (2016) Laboratory investigation of AC corona loss and corona onset voltage on a conductor under icing conditions. *IEEE Transactions on Dielectrics and Electrical Insulation*, 23(3), 1862–1871.

How to cite this article: Hämäläinen K, Hirsiikko A, Leskinen A, Komppula M, O’Connor EJ, Niemelä S. Evaluating atmospheric icing forecasts with ground-based ceilometer profiles. *Meteorol Appl.* 2020;27:e1964. https://doi.org/10.1002/met.1964