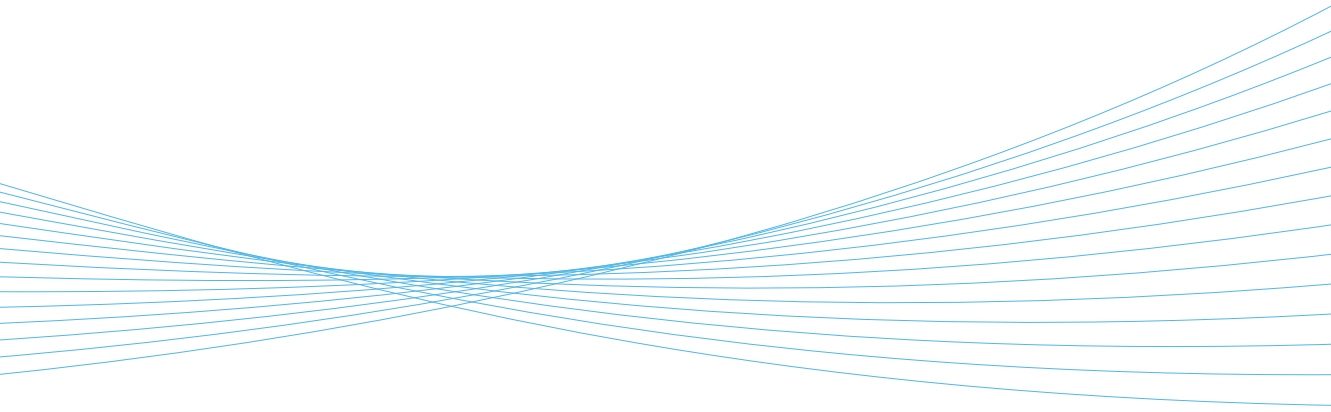
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

142
CONTRIBUTIONS

NEW METHODS USING IN-SITU AND REMOTE-SENSING OBSERVATIONS FOR IMPROVED METEOROLOGICAL ANALYSIS

ERIK GREGOW



FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS

No. 142

NEW METHODS USING IN-SITU AND REMOTE-SENSING OBSERVATIONS FOR IMPROVED METEOROLOGICAL ANALYSIS

Erik Gregow

Department of Physics
Faculty of Science
University of Helsinki
Helsinki, Finland

ACADEMIC DISSERTATION in meteorology

To be presented, with the permission of the Faculty of Science of the University of Helsinki, for public criticism in E204 auditorium at Physicum (Gustaf Hällströmin katu 2 A, Helsinki) on February 9th, 2018, at 12 noon.

Finnish Meteorological Institute
Helsinki, 2018

Supervisors Dr. Elena Saltikoff
Meteorological Research, Satellite and Radar Development
Finnish Meteorological Institute, Finland

Dr. Carl Fortelius
Meteorological Research, Numerical Weather Prediction
Finnish Meteorological Institute, Finland

Reviewers Associate Professor Piia Post
Institute of Physics
University of Tartu, Estonia

Dr. Niels Woetman Nielsen
Centre for Meteorological Models
Danish Meteorological Institute, Denmark

Custos Professor Heikki Järvinen
Department of Physics
University of Helsinki, Finland

Opponent Associate Professor Joan Bech
Department of Applied Physics - Meteorology
University of Barcelona, Spain

ISBN 978-952-336-042-6 (paperback)
ISBN 978-952-336-043-3 (pdf)
ISSN 0782-6117

Erweko
Helsinki 2018

FINNISH METEOROLOGICAL INSTITUTE

Published by Finnish Meteorological Institute
(Erik Palménin aukio 1), PL 503
00101 Helsinki

Series title, number and report code of publication
Finnish Meteorological Institute
Contributions 142, FMI-CONT-142
Date
February 2018

Author
Erik Gregow

Title
New methods using in-situ and remote-sensing observations for improved meteorological analysis

Abstract

Observations have been and are an important part of today's meteorological developments. Surface observations are very useful as they are, providing weather information for a point location. Though they do not give much information, if any, on what happens between the stations across a larger area. With models one can create an analysis of the meteorological situation, i.e. calculate and estimate what happens between these fixed observation points. Remote-sensing data, such as radar and satellite, are being processed and the output is given over a domain as an analysed product of their measurements. For example, radar gives a plot of where the rain is located, i.e. an analysis of the current precipitation.

With a series of radar images, a human (subjectively) or a computer (objectively) can process this information to estimate where the rain will move and be located within the next few minutes (even hours), i.e. a short forecast also called "nowcast". This applies to some extent also for other observations, such as satellite data (cloud propagation). But for most quantities (such as temperature, wind, etc) it is significantly harder to make such a nowcast, since these are influenced by many other factors and there is no linear development of them. Therefore, there are forecast models that solve physical and dynamic equations, so that one can estimate the future weather for the coming hours and days. A prerequisite for generating a forecast of high quality is to capture the initial weather conditions as best as possible. This is done using observations and they are introduced into the forecast model through different techniques, where the model creates its own analysis as the initial step. There remain problems since forecast models often are affected by physical disagreements, as the dynamic conditions are not in balance. This results in the model having a spin-up effect, where the meteorological quantities are not yet in balance with each other and the resulting weather conditions are not always reliable during the first hours. Hence, a lot of research is spent on how to reduce this spin-up effect and on the use of nowcast models, in order to deliver the best model results for the first few hours of the forecast period.

In this dissertation, the research work has been to improve the meteorological analysis, algorithms and functionality, using the Local Analysis and Prediction System (LAPS) model. Different kinds of observations were used and their interdependencies have been studied, in order to combine and merge information from various instruments. Primarily focus has been to improve the estimation of precipitation accumulation and meteorological quantities that affect wind energy. The LAPS developments have been used for several end-users and nowcasting applications, and experimentally as initial conditions for forecast modelling. The studies have been concentrated on Finland and nearby sea areas, with the available datasets for this domain.

By combining surface-station measurements, radar and lightning information, one can improve the precipitation-amount estimations. The use of lightning data further improves the estimates and gives the advantage of having additional data outside radar coverage, which can potentially be very useful for example over sea areas. In addition, the improved LAPS analyses (cloud-related quantities) and a newly developed model (LOWICE), calculating the electricity production during wintertime (taking into account the icing of wind turbine rotor blades which reduces efficiency), have shown good results.

Publishing unit
Finnish Meteorological Institute

Classification (UDC)	Keywords	
556.12, 551.508.85 621.548, 624.143	Precipitation, observations, radar, lightning wind energy, wind-power, wind-turbine icing, power-loss	
ISSN and series title		
0782-6117 Finnish Meteorological Institute Contributions		
ISBN	Language	Pages
ISBN 978-952-336-042-6 (paperback), 978-952-336-043-3 (pdf)	English	106

FINNISH METEOROLOGICAL INSTITUTE

Publicerat av Finska Meteorologiska Institutet
(Erik Palménin aukio 1), PL 503
00101 Helsinki

Serie titel, nummer och rapportkod för publicering
Finnish Meteorological Institute
Contributions 142, FMI-CONT-142
Publicerat
Februari 2018

Författare
Erik Gregow

Titel
Nya metoder för användande av observationer och förbättrade meteorologiska analyser

Abstrakt

Observationer har varit och är en betydelsefull del i den meteorologiska utvecklingen. Markobservationer är mycket användbara som de är, de tillför väderdata för en specifik punkt. Men de ger ingen information om vad som händer mellan dessa mätpunkter. Med modeller kan man skapa en analys, dvs beräkna och estimera vad som händer mellan dessa observationstationer. Radar och satellit ger data över områden och är en produkt där dess mätningar är analyserade. Till exempel, radar ger en bild av var regnet befinner sig, dvs en analys av nuläget.

Med en serie av radar bilder, kan en människa (subjektivt) eller en dator (objektivt) bearbeta denna information så att man får en uppfattning om var regnet kommer att befinna sig inom de närmaste minuterna (även timmarna), dvs en kort progonos även kallat "nowcast". Detta gäller även i stor utsträckning för övriga observationer, såsom satellit data (molnutbredning) etc. För meteorologiska parametrar såsom temperatur eller vind, är det dock betydligt svårare att göra en sådan nowcast, då dessa påverkas av många andra faktorer och det finns inte en linjär utveckling av dem. För att lösa detta problem finns det prognos-modeller, som löser de fysikaliska och dynamiska ekvationerna så att man kan få en bild av kommande väderparametrar för de kommande timmarna och dygnen. En förutsättning för en bra prognos är att man fångar det initiala väderläget så bra som möjligt. Detta görs med observationer och de introduceras i prognosmodellen via olika tekniker. Här kvarstår ett problem då modeller påverkas av fysikaliska oenigheter då de dynamiska förhållandena är i obalans. Detta resulterar ofta i att modellen under de första timmarna har en "spin-up" effekt där de meteorologiska parametrarna ännu inte är i balans med varandra och de utvecklade väderförhållandena ännu inte är helt tillförlitliga. Därav spenderas mycket forskning om hur man kan reducera denna spin-up effekt och användandet av nowcast-modeller för att tillföra bästa modell resultat för de närmaste timmarna.

I denna avhandling har fokus varit att förbättra den meteorologiska analysen (algoritmer och funktionalitet), genom att använda modellen Local Analysis and Prediction System (LAPS). Ett flertal observationer har använts och deras inbördes påverkan studerats, för att I bästa möjliga mån kombinera information från dessa olika instrument. Fokus har främst varit med avseende på nederbördsmängd och beräkning av meteorologiska parametrar som påverkar vindkraftsenergi. LAPS har även använts experimentellt i nowcasting syfte och som analys för prognos-model, för att förbättra prognoserna i närtid. Studierna har i första hand fokuserat på Finland, med närliggande havsområden och tillhörande observations nätverk och instrumentering.

Vi har funnit att genom användandet av mark-stationer, radar och blixtnedslags information så kan man förbättra bestämningen av nederbördsmängden. Användandet av blixtdata ger möjligheten att bestämma nederbörd över områden där det inte finns radar, till exempel över havsområden, vilket förr inte varit möjligt. Därtill har vi med förbättrade LAPS analyser (främst moln relaterade parametrar) och en nyutvecklad modell (LOWICE) påvisat positiva resultat vid beräkning av elproduktionen under vintertid, där man tar i beaktning nedisning av vindkraftverkens rotorblad, vilket sänker effektiviteten.

Publicerings instans		
Finska Meteorologiska Institutet		
Klassficering (UDK)	Sökord	
556.12, 551.508.85	Observationer, nederbördsmängd, radar, blixtar	
621.548, 624.143	Vindkraft, vindenergi, nedisning	
ISSN och serie titel		
0782-6117 Finnish Meteorological Institute Contributions		
ISBN	Språk	Antal sidor
ISBN 978-952-336-042-6 (paperback), 978-952-336-043-3 (pdf)	Svenska	106

Preface

The work presented in this thesis has been carried out in the Meteorological Research Unit of the Finnish Meteorological Institute (FMI), while participating in several different projects and through involvement in many operational development works.

I thank all the co-authors and people who have helped me during this work.

I want to thank FMI for the opportunity to work among experts within the NWP field and including me in many interesting projects. I have always felt warmly welcome and there has been a positive spirit surrounding the cooperative work within the unit and whole of FMI.

I want to thank Professors Heikki Järvinen and Hannu Savijärvi for the support and help during my PhD work.

I want to express my deepest gratitude to Dr. Elena Saltikoff who has always been there to help, supervise and guide me, whenever in scientific troubles. Many thanks to Drs Carl Fortelius and Sami Niemelä who have been excellent supervisors during the years of my work at FMI and Docent Curtis Wood for helping in the final polishing of the thesis.

Thanks to specialist Ben Bernstein for helping with scientific questions and becoming a dear friend during the years of work together, thank you.

A special thank you to Professor Wahé Balekjian, who I have always looked up to and who encouraged me to study for this Doctorate. My thoughts go also to my mother Margareta, my Aunt Birgitta, my siblings and Annika who are, and always will be, the foundation in my life. Last, but not least, my children Anni and Nelli: You are my everything!

Erik Gregow
Helsinki, February 2018

Contents

Acronyms - Abbreviations 9
List of the original publications 10
Summaries of the original publications 11
1. Introduction 14
2. Research motivations, questions and goals 17
3. Methods and materials 18
    3.1. The Local Analysis and Prediction System (LAPS) 19
        3.1.1 LAPS – The RandB method 20
        3.1.2 LAPS – Lightning Data Assimilation (LDA) method 21
        3.1.3 LAPS – LOWICE method 21
    3.2. Observational datasets 24
        3.2.1 Surface gauges 24
        3.2.2 Radars 25
        3.2.3 Lightning-detection system 26
        3.2.4 Icing detection 26
        3.2.5 Wind-power production measurements 27
4. Summary of the results 28
    4.1. Precipitation accumulation estimates 28
    4.2. Wind-power production and icing 31
5. Discussion 33
6. Conclusions 36

References 37

Acronyms - Abbreviations

3D – Three-Dimensional
AWS – Automatic Weather Station
CBZ – Cloud-Base Height
CORR – Correlation
CRR – Convective Rainfall Rate
DA – Data Assimilation
ECMWF – European Centre for Medium-Range Weather Forecasts
FMI – Finnish Meteorological Institute
FTA – Finnish Transport Agency
HIRLAM – High Resolution Limited Area Model
IFS – Integrated Forecasting System
INCA – Integrated Nowcasting through Comprehensive Analysis
LAPS – Local Analysis and Prediction System
LDA – Lightning Data Assimilation
LEA – Leading Edge Atmospherics
LF – Low Frequency
LLS – Lightning-Location System
LWC – Liquid Water Content
MAE – Mean Absolute Error
MESAN – Mesoscale Analysis System
MOS – Model Output Statistics
NORDLIS – Nordic Lightning-Information System
NWP – Numerical Weather Prediction
PPI – Plan-Position Indicator
PC – Precipitating Clouds
QC – Quality Control
QPE – Quantitative Precipitation Estimation
RandB – Regression and Barnes
RMSE – Root Mean Square Error
SLWC – Supercooled Liquid Water Content
STDEV – Standard Deviation
STEPS – Short-Term Ensemble Prediction System
VERA – Vienna Enhanced Resolution Analysis
VHF – Very-High Frequency
VPR – Vertical Profile of Reflectivity

List of the original publications

I Gregow, E., E. Saltikoff, S. Albers, and H. Hohti, 2013: Precipitation accumulation analysis – assimilation of radar–gauge measurements and validation of different methods, *Hydrol. Earth Syst. Sci.*, **17**, 4109–4120, doi:10.5194/hess-17-4109-2013

II Gregow, E., B. Bernstein, I. Wittmeyer, and J. Hirvonen, 2015: LAPS-LOWICE: A Real-Time System for the Assessment of Low-Level Icing Conditions and Their Effect on Wind Power, *J. Atmos. Oceanic Technol.*, **32**(8), 1447–1463, doi: http://dx.doi.org/10.1175/JTECH-D-14-00151.1

III Gregow, E., A. Pessi, A. Mäkelä, and E. Saltikoff, 2017: Improving the precipitation accumulation analysis using lightning measurements and different integration periods, *Hydrol. Earth Syst. Sci.*, **21**, 267–279, doi:10.5194/hess-21-267-2017

IV Mäkelä, A., E. Saltikoff, J. Julkunen, I. Juga, E. Gregow, and S. Niemelä, 2013: Cold-season thunderstorms in Finland and their effect on aviation safety, *Bull. Amer. Meteor. Soc.*, **94**, 847–858, doi:10.1175/BAMS-D-12-00039.1

Summaries of the original publications

I. Gregow, E., E. Saltikoff, S. Albers, and H. Hohti, 2013: Precipitation accumulation analysis – assimilation of radar–gauge measurements and validation of different methods, *Hydrol. Earth Syst. Sci.*, **17**, 4109–4120, doi:10.5194/hess-17-4109-2013

In this article, we investigate four different methods to produce precipitation accumulation fields, using radar data combined with precipitation-gauge observations. The Local Analysis and Prediction System (LAPS) is used as a platform to calculate four different hourly accumulation products over a 6-month verification period, including summer 2011. The study uses radar reflectivity, as well as three assimilation methods that blend together radar and surface data; linear analysis regression, Barnes objective analysis and a new method based on a combination of the regression and Barnes techniques (RandB). The performance of each method is verified against both dependent and independent observations (i.e. observations that are or are not included, respectively, into the precipitation-accumulation analysis) across Finland. Results showed that the newly developed RandB method performed the best. Although not as good as the RandB method, individual application of the regression or Barnes assimilation analysis also yielded improvements to results for the accumulation products, compared with precipitation accumulation derived from radar data alone. The lead author was responsible for all the analyses and for the major part of the calculations and writing.

II. Gregow, E., B. Bernstein, I. Wittmeyer, and J. Hirvonen, 2015: LAPS-LOWICE: A Real-Time System for the Assessment of Low-Level Icing Conditions and Their Effect on Wind Power, *J. Atmos. Oceanic Technol.*, **32**(8), 1447–1463, doi: http://dx.doi.org/10.1175/JTECH-D-14-00151.1

The wind-power industry is highly sensitive to weather; and atmospheric icing has a clear impact on turbine efficiency, sometimes causing rapid and substantial power losses and even total shutdown of wind farms. Therefore, accurate analyses and forecasts of wind- and icing-related meteorological variables are of great importance. The Local Analysis and Prediction System (LAPS) - LOWICE system has been developed to produce real-time hourly estimates of the presence, intensity, and impacts of icing on wind power production. Analysis of LAPS-LOWICE output and observations from wind farms indicated that wind-power losses were not well-correlated with measured ice loads. Instead, wind power losses were better

correlated with icing rate and its time history, in combination with the loss of ice due to melting, sublimation, and shedding. The lead author was responsible for the LAPS developments, partly involved in the development of the LOWICE model, and for a major part analysing the results and writing the article.

III. Gregow, E., A. Pessi, A. Mäkelä, and E. Saltikoff, 2017: Improving the precipitation accumulation analysis using lightning measurements and different integration periods, *Hydrol. Earth Syst. Sci.*, **21**, 267–279, doi:10.5194/hess-21-267-2017

The article introduces and compares new methods of precipitation-accumulation analysis, with special focus on heavy-rainfall events. The method assimilates lightning observations, in combination with radar and raingauge measurements, to give an estimate of precipitation accumulation. A new Lightning Data Assimilation (LDA) method has been implemented and validated within the Finnish Meteorological Institute (FMI) Local Analysis and Prediction System (LAPS). Precipitation accumulation analyses indicated the usefulness of lightning assimilation, together with radar information. Additionally, the impact of different integration times on the radar–gauge correction method was investigated in this article. The radar–gauge assimilation method was dependent on statistical relationships between radar and gauges, when performing the correction to precipitation accumulation field. Here we investigated the usage of different integration intervals; 1, 6, 12, 24 hours and 7 days. Such differences changed the amount of data used and affected the statistical calculation of the radar–gauge relations. Verification showed that the real-time analysis using the 1-hour integration time gave the best result. The work presented in this article was a continuation of previous work in the same research field, by Gregow et al. (2011). The lead author was responsible for all the analyses, implementing and utilizing the LDA method within FMI-LAPS, and for a major part of the calculations and writing.

IV. Mäkelä, A., E. Saltikoff, J. Julkunen, I. Juga, E. Gregow, and S. Niemelä, 2013: Cold-season thunderstorms in Finland and their effect on aviation safety, *Bull. Amer. Meteor. Soc.*, **94**, 847–858, doi:10.1175/BAMS-D-12-00039.1

A total of 13 commercial aeroplanes were struck by lightning in October (ten in one day) and December (three on separate days) of 2011 in the main Finnish

Helsinki–Vantaa airport corridor. The number of lightning-struck airplanes was extremely large, considering the time of year and the small number of strikes by the storms. The analysis suggested that a major cause for the large number of struck airplanes is that the planes took off directly into the convective core of the storm and the planes initialized the flashes themselves. The interview of the pilots of those aeroplanes struck by lightning showed that the pilots did not receive detailed information to allow them to avoid the situation. The lightning strikes did affect the pilots, causing temporary loss of sight and hearing, but luckily no fatalities or severe damage occurred. This paper gives an overview of the synoptic weather situation, as well as the forecasts, for these events. There were remarkable differences in the operational forecast models and the high-resolution non-hydrostatic model was superior in predicting the convective nature of the event, compared to the coarser-resolution hydrostatic model. LAPS analysis was used to determine and compare the vertical temperature and wind profiles for these cases. Additionally, the LAPS-calculated stability indexes (such as the K-index, Lifted index, and Total Totals index) provided useful information for estimating the risks of thunderstorms. The lead author's contribution was related to the LAPS analysis and results of the study.

1. Introduction

The atmosphere is in constant motion as it tries to adjust for imbalances — for example developed by differences in air masses, topographical and land–sea effects, etc — and it does so by developing weather phenomena (e.g. frontal activities, thunderstorms, precipitation, formation of clouds, etc). An analysis of the weather describes the atmospheric state (e.g. meteorological quantities and fields), by the use of meteorological observations from both surface and upper-air (Daley, 1991). The analysis can give a solution which might not be exactly physically consistent or in perfect balance between all meteorological quantities (i.e. a cloud pattern might not be in balance with the wind vectors/motions at the same time and placement), but still the analysis describes individual fields with reasonable accuracy. A numerical weather prediction (NWP) model produces forecasts of the weather for the upcoming days, normally out to 3 days but some forecasts even up to 10 days. NWP models use an analysis as their starting point, where the observations usually are ingested via data-assimilation (DA) techniques (Kalnay, 2003). A nowcast model is based upon the ability to describe existing meteorological conditions at very-high resolution (i.e. an analysis) and extrapolate this information forward in time (Vivoni et al., 2006). It operates on time-scales of a few hours ahead: exactly those time-ranges where NWP models suffer from problems such as assimilating observations at the convective scale, accurately representing physical processes, problems of model spin-up and rapid error growth at the convective scale, etc (Sun and Wang, 2013).

It is important to distinguish between whether the goal is to achieve an as accurate analysis as possible, or to create an analysis with the intention to be used as initial condition for starting NWP forecast models. Because, in order to use the analysis within NWP models, there are limitations in its use of observations. This is because the NWP model state and analysed variables need to be consistent, and in balance with each other, and this limits the usage of observations within NWP (Talagrand, 1997). For example satellite observations (microwave spectrums), which provide much information on clouds, are difficult to use and therefore much of the data is discarded in the NWP DA systems. It is also noticed that there is a quick loss of information from observations during the early forecast steps in NWP models (Bauer et al., 2011). Methods used for an as accurate analysis as possible, and to some extent also in nowcast models, combine and merge observations where the observed variables are retained as much as possible. In this research work, focus has been to create an as accurate analysis as possible and to use all available high-quality observations.

There is an extensive amount of meteorological observations available, especially from meteorological institutes but also commercial actors, crowdsourcing (e.g. citizen observations) and social media (Hyvärinen and Saltikoff, 2010); and they have

been substantially increasing during recent decades. At the surface, automatic weather stations (AWS) are growing in number as well as the instrumentation at them. Remote-sensing measurements, such as radar and satellites, produce large quantities of fine-scale observations, both at surface and in upper air (Kelly and Thépaut, 2007). The information from each measurement can be useful as an independent observation, but in many applications a gridded product is needed (e.g. for nowcasting or as input to end-user applications). The gridding process (i.e. the weather quantities are calculated at distinct spatially equidistant points for the area of interest), can be performed with a minimum of human intervention needed at different degrees of advanced levels and for different purposes. An analysis model provides meteorological quantities interpolated onto a grid for a certain valid time, whereas a nowcasting model produce both an analysis and short-term forecasts. With the increased amount of observations, these models can potentially reach an improved analysis state through a more-complete data coverage in time and space. This has also the potential to affect and bring positive impact to the weather forecasts and end-user specific applications, for example model initialisation for hydrological, fire-weather and wind-power uses.

Several analysis and nowcasting systems are made available around the world, which produce analysis and/or nowcast fields of meteorological quantities. They all have different features, making them attractive for different end-users in weather prediction, and there is a range of new methods being developed in the modelling community. The systems use different techniques to combine the available observations, and usually the process involves an NWP model used as initial background field. There are three commonly used techniques; i) objective interpolation (Barnes, 1973), ii) optimal interpolation (Lorenc, 1981) and iii) variational assimilation (Le Dimet and Talagrand, 1986). The target is to obtain a description of the atmosphere in terms of meteorological variables (Lahoz et al., 2010). Most meteorological institutes run their own preferred system with their own specifications, i.e. predefined domain, resolution, ingest of observations, etc. Below are the principles for some of the existing and operationally running analysis systems.

The Local Analysis and Prediction System (LAPS) is capable of producing three-dimensional (3D) analyses of the atmosphere, for several meteorological variables (Albers et al., 1996). Whereas, other analysis system only produce two-dimensional (2D) analysis and only for certain variables. LAPS is able to use a wide range of different observations as input and there are several NWP model options to choose from, to be used as background field. The software is open-source and there is a large user-group worldwide. Altogether, this was a favourable system to adopt at FMI.

The Mesoscale Analysis System (MESAN) create a 2D mesoscale analysis of

selected meteorological variables and use the high-resolution limited-area model (HIRLAM) as a background. The analysed quantities are of general interest in operational weather forecasting and to produce initial information to be used for nowcasting tools (Häggmark et al., 2000). MESAN is not an open-source program.

The Integrated Nowcasting through Comprehensive Analysis (INCA) system provides analysis and nowcasting for a selection of products, including both 3D quantities (temperature, humidity and wind) and 2D quantities (precipitation amount, precipitation type, cloudiness and global radiation). The nowcast is merged into an NWP forecast provided by a limited-area model. INCA has been especially developed for use in mountainous terrain (Haiden et al., 2011). The source-code is not freely available.

The Vienna Enhanced Resolution Analysis system (VERA) has similarities with INCA and the analysis is used in complex terrain and, also, not freely available. It has the advantage of not needing background fields to create the analysis and the system includes a sophisticated data quality-control (QC) tool (Schneider et al., 2008).

The Short-Term Ensemble Prediction System (STEPS) is primarily focused on precipitation fields and is not an open-source program. It is a probabilistic nowcasting system, using both the extrapolation of radar images and the downscaled precipitation output of NWP models to produce seamless 2D forecasts (Seed et al., 2013).

2. Research motivations, questions and goals

Through involvement in several projects, both domestic and international, two target areas were identified as important research topics: precipitation accumulation and wind power production. For example, the uncertainty in hydrological predictions is mainly due to the quality of the estimated rainfall, used as input to hydrological models. This affects the catchment hydrology (i.e. for hydropower) and the control of urban drainage and sewer systems (assessment of flooding risk). There is a need to know the amount of precipitation at every point in Finland, even in places where there is no surface gauge data, with as good quality as possible (Jasper et al., 2002). Wind-power production is a growing source of energy but also power production is sensitive to icing, especially in Nordic countries and areas with high elevation. It has been shown that energy production decreases rapidly when ice forms on turbine blades, and there is also a safety risk due to shedding ice blocks (Cattin et al., 2007). Modelling the icing process has been a topic of great interest in recent years and the outcome may have large economic impact both on the turbine and wind-park owners, as well as for the electricity market, i.e. the buying and selling of expected generated electricity.

The scientific research goal of this work was to improve the existing data-fusion methods and to develop new ones, to use and combine both the standard and new high-temporal resolution observations in the best way. In order to do this, we needed to solve the problem of combining observation types from different instrumentation with different measuring scales and error characteristics in a meaningful way. The final goal was to produce a high-quality gridded analysis for i) precipitation-accumulation estimates and of full atmosphere for ii) icing-related quantities (clouds, temperature, humidity, etc.), to be used in operational products and as input for end-user-specific models.

As part of the FMI institutional duties, the obligations are to produce and deliver high-quality meteorological datasets to public and other end-users, free of charge. The meteorological services, not only in Finland but all over the world, are aiming at a higher degree of automation, i.e. to replace the majority of manual observations by automatic stations and to increase the number of stations. These automatic observations, together with the utilization of remote-sensing data from radars and satellites, provided the possibility to develop and improve the mesoscale FMI-LAPS analysis.

The work to combine surface gauges, radar and lightning data in order to perform a better precipitation accumulation analysis, are published in articles I and III. The assessment of the cloud analysis and related icing parameters to determine the effects of icing on wind power production is described in article II. The usage of LAPS products in end-user applications (e.g. risk of lightning) is described in article IV.

3. Methods and materials

LAPS was adopted as the operational regional mesoscale analysis system at FMI in 2009. The system offers the advantages of assimilating a wide variety of observations into an analysis of the entire atmosphere, hereafter referred to as 3D analysis. There are only a few other analysis models with this capability (several systems produce only surface analyses) and it is an excellent platform to introduce and test newly developed routines and concepts, for either ingestion or output of new meteorological quantities. LAPS has been used in research and operational duties all over the world for almost three decades. Besides being a reliable analysis model, there are also other factors that made it a successful tool; open-source code, a large user-group forum for questions and discussions, and it has very good support by the developers at NOAA. The LAPS products were found to be useful in many of the meteorological applications, both within FMI and other Finnish companies.

One of the main developments reported in this thesis is related to the precipitation-accumulation process. With the use of more observations (both surface and remote-sensing instruments) and newly developed assimilation routines, accumulation estimates became better and more useful for the end-users. The high-quality observations from SYNOP stations were complemented by the Road-Weather stations network and with this, good surface-observational coverage of the Finland domain was achieved. The first improvement was incorporated by combining surface observations with radar measurements, in order to make corrections to the radar-accumulation field. Here, the surface observations usually measure the precipitation accumulation correctly, while the radar network has much better areal coverage and the ability to resolve what happens between the surface stations. Lightning information was investigated as a complementary observation to further improve the accumulation analysis. The goal was to improve the result in heavy rainfall situations (i.e. during thunderstorms). The outcome of the LAPS precipitation accumulation process is described in articles I and III. The LAPS 3D analysis is useful as a starting point for other models. In article II an icing model was developed which estimates the icing effects on wind-power production. Here LAPS is important with its high resolution meteorological fields, such as the temperature, humidity and especially clouds within the atmospheric boundary layer and at wind turbine levels. The LAPS analysis also includes derived atmospheric-stability indexes for upcoming convective weather situations. These data were used in article IV to detect winter thunderstorms, including lightning, and for the risks in aviation safety. The LAPS mesoscale analysis and new development methods are described in Section 3.1. The observational datasets that have been used within articles I–IV are described in Section 3.2.

3.1. The Local Analysis and Prediction System (LAPS)

LAPS is used for the production of 3D analysis fields of many weather quantities (Albers et al., 1996). The system uses a data-fusion method, in which a high-resolution spatial analysis is performed on top of the coarser resolution background fields. Observations are fitted, mainly by using an objective analysis, with a successive-correction method (Barnes, 1994) while high-resolution topographical datasets are taken into account when creating the final high-resolution analysis fields. In the successive-corrections method the field variables are modified by the observations in an iterative manner. Successive iterations are made at every grid point, updating the variable at each grid point based on a first-guess field and the observations surrounding that grid point.

A field at grid point i is updated according to the following formula

$$f_i^{m+1} = f_i^m + \frac{\sum_{K=1}^{K_i} w_{iK}^m \left(O_K - f_K^m\right)}{\sum_{K=1}^{K_i} w_{iK}^m + \epsilon^2}, \tag{1}$$

where f_i^m is the value of the variable (e.g. T, q, u, etc) at the i'th grid point at the m'th iteration, O_K is the K'th observation surrounding the grid point, $w_i k^m$ is a weighting function which depends on how far the observation is from the grid point (Eq. 2), and ϵ^2 is an estimate of the ratio of the observation error to the first-guess field error (if the observations were perfect then $\epsilon^2 = 0$). The weighting function is

$$w_{iK}^m = e^{-\frac{r_{iK}^2}{2R_m^2}}, \tag{2}$$

where r is the radius between the observation station and the grid-point and R is the radius of influence (set by the user).

In the FMI version of LAPS (Fig. 1), hereafter FMI-LAPS (Koskinen et al., 2011), the coarser background fields are taken from the latest available forecast from the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS). The fine-scale structures in the resulting 3D-analysis are extracted from the observations. Therefore, LAPS relies heavily on the existence of a high-resolution spatial and temporal data from observational networks and remote sensors. The horizontal grid-spacing is 3 kilometres and the domain covers all of Finland and parts of neighbouring countries. The setup uses a pressure coordinate system including 44 vertical levels, distributed with a finer resolution (e.g. 10 hPa) at lower altitudes and decreasing with height. At present, FMI-LAPS is able to process several types of in-situ and remotely sensed observations such as: radar reflectivity and radial winds, weighing gauges, road-weather observations, atmospheric soundings, SYNOPs, METARs, air-traffic observations, lidars and Meteosat-9 satellite data. There are QC's of the input data, which are important in an automated analysis system where

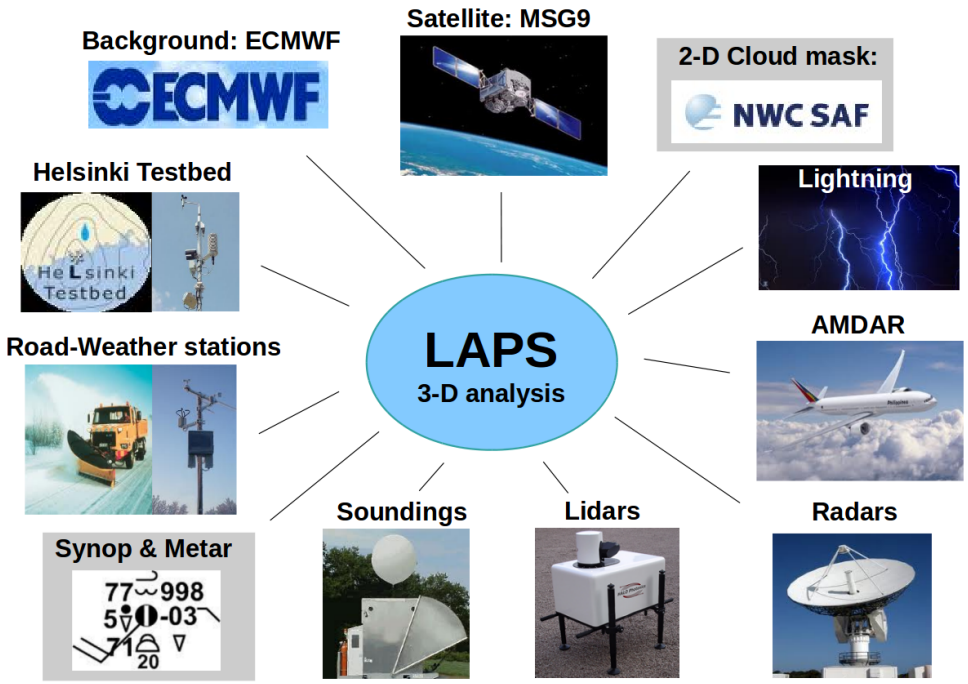


Figure 1: Example of the observational datasets and background model used in LAPS at FMI.

there is a minimum of human intervention. FMI-LAPS has been especially developed to better analyse the precipitation and cloud-related fields, this in order to perform better in accumulation calculations and in wind-power estimates (as in articles I, III and II, respectively).

3.1.1. LAPS – The RandB method

The newly developed LAPS- Regression and Barnes (RandB) method consists of two combined precipitation-correction calculations, which are run in sequence after each other; Regression- and Barnes methods, within the LAPS routines. The linear-regression-analysis method, used in the first step, calculates the quotient between the gauge–radar pairs from all given station-points within the LAPS area. The pairs undergo QC, based on thresholds, to prohibit dubious differences between gauge and radar values (i.e. to avoid including uncertain radar measurements and spurious surface observations). Once the QC's criteria are enforced, the remaining data form a dataset of representative gauge–radar pairs from which a linear regression

can be established, calculated with the least-square method (i.e. minimize the errors between the measurement pairs). The Regression method is used to correct the overall radar estimate at all grid-points, i.e. a constant correction over the domain for a given time-step.

In the second step, the Barnes method forces the radar field to converge towards accumulation values measured by the gauges, using an objective multi-pass telescoping strategy (Barnes, 1964; Hiemstra et al., 2006). Also here the gauge–radar quotients are used (including QC) and in order to optimize the result, several iteration steps are performed within the Barnes analysis, at successively finer scales. The corrections are weighted with distance (i.e. less impact from gauge observation further away from surface station) and rectifies the radar field in the surroundings to gauge stations.

3.1.2. LAPS – Lightning Data Assimilation (LDA) method

The Lightning Data Assimilation (LDA) method is constructed to build up statistical relationships between radar–lightning measurements and the new dataset is further used to improve the precipitation-accumulation estimates (article III). LDA counts the number of cloud-to-ground lightning strikes and converts lightning intensity into vertical radar-reflectivity profiles. The FMI-LAPS LDA method uses a 5 min interval of lightning and radar data, within a LAPS grid-box of resolution 3*3 km (Fig 2).

The collected strikes are divided into binned categories using an exponential division (i.e. $2^n ... 2^{n+1)}$, following the same method used in (Pessi, 2013). This results in 6 different lightning categories (e.g. with 1, 2–3, 4–7, 8–15, 16–31 and 32–63 bins) for the Finnish lightning detection dataset. For each of these 6 categories the average reflectivity is calculated at each grid-point, for each level, and results in the final radar–lightning reflectivity profiles. For the Finland domain, the climatological radar–lightning relationship profiles were estimated using lightning information and operational radar-volume data from summer 2014. Approximately 220,000 lightning strikes were used for this calibration.

3.1.3. LAPS – LOWICE method

A system to detect icing, calculating the ice load and the wind-power losses (LOWICE) has been developed by Leading Edge Atmospherics (LEA) and FMI (Gregow et al., 2015). The LOWICE model was run over the Nordic area using LAPS as its primary data source to produce gridded fields of relevant quantities for near-surface icing: temperature, humidity, winds, cloud microphysics, cloud fraction and precipitation type. LOWICE produces estimates of the following at the height of the wind turbine: wind speed, supercooled liquid water content (SLWC), drop size, icing intensity (i.e. accretion rate) and ice load.

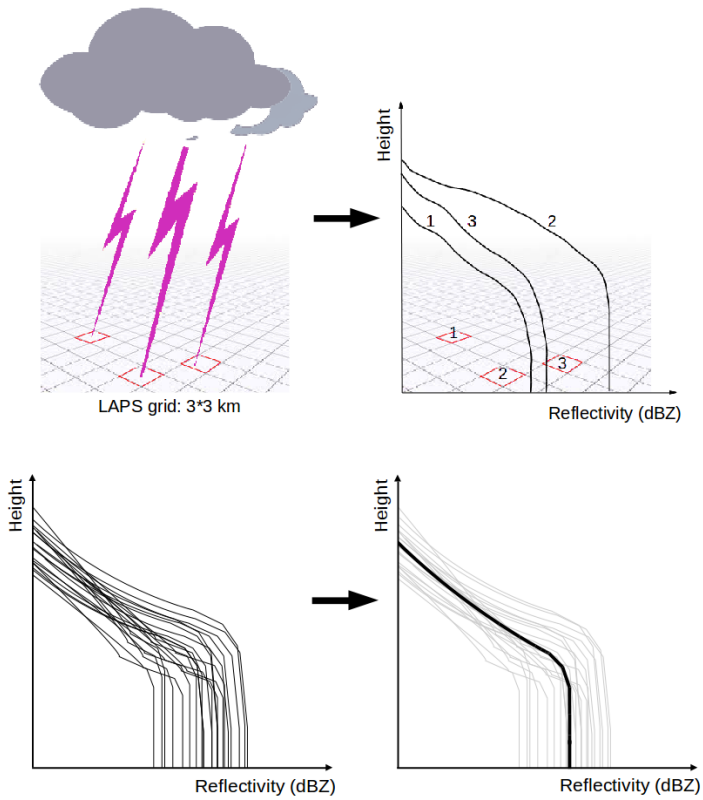


Figure 2: Scheme of the LDA method. Lightning strikes remapped onto LAPS grid (upper-left panel), where radar-reflectivity profiles are collected for the same grid-boxes (upper-right panel). The assembled radar profiles are handled for a time-period of 5 minutes and binned into classes (lower-left panel; example of one bin) and finally, the average profile is calculated for each bin (lower-right panel; thick line).

Clouds have a key role when calculating icing-related quantities, affecting wind-power turbines; and one of the main features within the FMI-developed LAPS version is the cloud-resolving process (Fig. 3). The cloud analysis is dependent on the use of satellite input, visible and infrared channels, which are used to detect the cloud-mask and estimate the cloud-top temperature/height. Additionally, information from SYNOP, METAR and radar is used to fill in the vertical cloud structure and the height of the cloud-base. New FMI-LAPS developments include the use of cloud-mask information from a NWCSAF product (Dybbroe et al., 2005) and the detection of clouds captured within temperature inversions (e.g. low clouds).

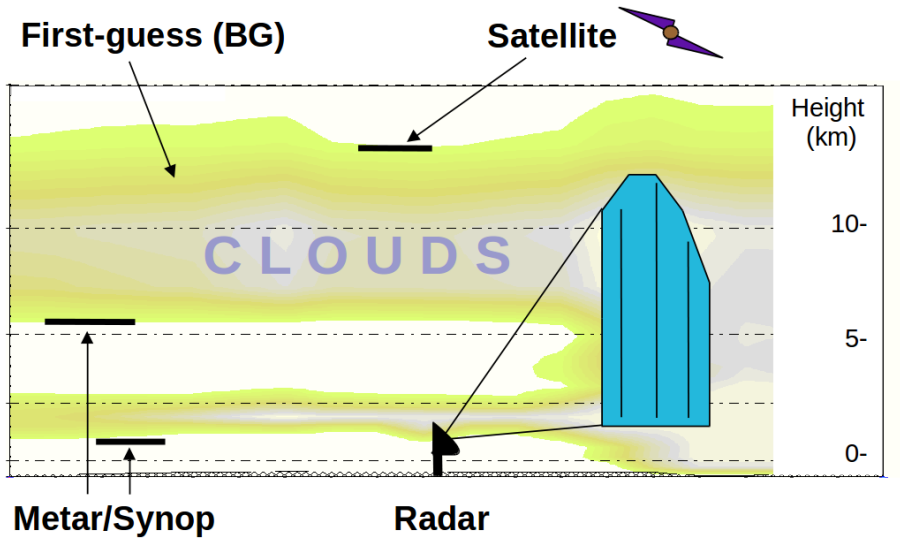


Figure 3: Sketch of how observations and NWP model are used within the FMI-LAPS cloud process. The satellite detects the cloud-top; surface stations (e.g. Metar and Synop) measures the cloud-base heights and the vertical structures of the clouds are filled together with radar measurements and first-guess fields from background (BG; from NWP forecast model).

The icing intensity is estimated using liquid water content (LWC; g m^{-3}), the temperature and the wind speed at the height of interest (e.g 100 metres above ground level). Temperature and wind speed are interpolated from LAPS vertical levels. LOWICE estimates the LWC by calculating the change of saturation-mixing ratio between two levels, in a similar manner as (Betts, 1987). Therefore, when the level of interest is located at or above the estimated cloud-base height (CBZ), the LWC is approximated by; 1) estimating the saturated-mixing ratio at CBZ, 2) taking the difference between the mixing ratios at cloud-base and the height of interest (assuming

a moist adiabatic temperature lapse rate) and finally, 3) compensating for density. Also, LOWICE is taking into account the depletion of LWC due to precipitation (especially snow), by analysing the observation reports from nearby stations. The depletion factor is depending on distance to observation and the LWC can be reduced with maximum 50% of its original value. To calculate the icing rate, with input of the atmospheric icing conditions from above, a standard icing-rate equation is used (Makkonen, 2000)

$$\frac{dm}{dt} = SLWCAv\alpha_1\alpha_2\alpha_3, \tag{3}$$

where $SLWC$ is the supercooled LWC, A is the cross-sectional area of the object, v is the wind velocity; and here the collision (α_1), sticking (α_2) and accretion (α_3) efficiencies are set equal to 1.0. The cross-sectional area (A) is set equal to 0.015 m^2, based on the ISO 12494 standard of a 0.5 m-long, 30 mm-diameter cylinder, but it could be changed to accommodate other objects (ISO, 2001). The wind speed (v; m s^{-1}) is taken directly from LAPS. By summing the hourly icing rate and thereby accumulating ice, when temperatures are in a suitable range, the ice load is estimated for the reference cylinder. This load will build during periods of active icing and can be depleted by melting and sublimation during periods when icing is not active. The melting scheme is currently based on temperature and the sublimation scheme use both wind speed and relative humidity.

There are three power-loss schemes tied to the icing properties, driven by ice load and icing rate, as well as time, depending on LOWICE version. V0 uses a simple method where the build-up of ice load determines the degree of power loss (e.g. when ice load increases from 0.0 to 10.0 kg m^{-1}, the power-loss increases linearly from 0% to 100%). LOWICE's alternative power loss schemes (V1 and V2) combined the "building" and "clearing" effects, where ice rate and time are key factors. LOWICE versions V1 and V2 outperformed V0.

3.2. Observational datasets

3.2.1. Surface gauges

Until the year 2013, FMI managed 77 stations instrumented with the weighing gauge Vaisala model VRG10 (used in article I). These instruments were replaced and since 2013 FMI now operates 102 stations instrumented with the weighing gauge OTT Messtechnik Pluvio2 (used in article II). The Finnish Transport Agency (FTA) operates 370 road-weather stations with optical sensor measurements (Vaisala Present Weather Detectors models PWD11 and PWD22; used in article I and II). The FTA observation sites are not selected according to meteorological standards. Hence, their location in

the immediate vicinity of roads with heavy traffic, where "splash-effects" and wind eddies (generated by big vehicles) occasionally affect to the measurement quality and representativeness, compared to FMI stations.

3.2.2. Radars

As of summer 2016, FMI operated ten C-band Doppler radars where all but one of the stations have dual-polarization. At the moment, the quantitative-precipitation estimation based on dual-polarization is not used operationally but the polarimetric properties contribute to the improved clutter cancellation (i.e. removal of non-meteorological echoes; especially sea clutter, birds and insects). In FMI's general radar processing, clutter is removed with Doppler-filtering and any residual clutter with a post-processing procedure based on fuzzy logic (Peura, 2002). Additionally, the vertical profile of reflectivity (VPR) method is correcting the range-dependent errors (Zawadzki, 1984) and also compensates for overestimation in a melting layer (Koistinen et al., 2004).

In southern Finland the distance between radars is 140–200 km, but in the north it can be as much as 260 km. The basic radar volume scan consists of thirteen PPI (so-called plan-position indicator) sweeps, which are scanned out to 250 km and repeated every 5 minutes. The location of the radars and the coverage is shown in Fig. 1a of article I. Because Finland has no high mountains, the horizon of all the radars is near zero elevation with no major beam blockage and, in general, the radar coverage is very good except in the most northern part of the country. The Finnish radar network has a very high system reliability (no interruption of data). During year 2014 and 2015 the reliability was > 99%. Further details of the FMI radar network and processing routines are described in (Saltikoff et al., 2010).

The effective radar-reflectivity factor Z_e (usually called reflectivity) is derived from the expression

$$Z_e = \frac{P_e r^2}{LC|K|^2}, \tag{4}$$

where P_e is the average received microwave power, r is the measurement range, L is the two-way attenuation in the propagation path (antenna–scatterers–antenna), C is a radar constant (including parameters of the radar hardware) and $|K|$ is the dielectric factor (depending on the relative fraction of ice and water in the hydrometeors).

The reflectivity uses dBZ as unit, which is expressed as

$$dBZ = 10log_{10}\left(Z_e\right). \tag{5}$$

The output of weather radars is the measured reflectivity (dBZ; Equation 4 and 5), which is further used to calculate the precipitation intensity.

3.2.3. Lightning-detection system

The Lightning Location System (LLS) of the FMI is part of the Nordic Lightning Information System (NORDLIS). NORDLIS started around the years 2001–2002 as a cooperation between Finland, Norway and Sweden, and gives a sufficient coverage of lightning detection for these countries. The system detects primarily cloud-to-ground strikes in the low-frequency (LF) domain, using IMPACT ES-type sensors and a central processor; LP2000 (Cummins et al., 1998). As of 2001, another system was installed in southwest Finland, the SAFIR total lightning network with three sensors (SAFIR 3000) and a central processor. The SAFIR system is using very high frequency (VHF) interferometry to locate all lightning discharges within its coverage area and also has a separate LF, E-field measurement for identifying ground strokes. SAFIR sensors have improved the detection efficiency and location accuracy of ground flashes in southern Finland, providing three extra time measuring stations in the network. In August 2004, a CP8000 central processor was installed which enabled the raw data from both IMPACT and SAFIR sensors to be processed at the same time. One compatible sensor (LS7000) was installed in Estonia in 2005 and was connected to FMI's central processor, which improved the detection efficiency and further widened the coverage area toward the south. At present, the FMI LLS uses this configuration: ground flash (LF) data available from Nordic countries and neighbouring areas, and cloud lightning (VHF) data that is available from SW Finland and the adjacent maritime areas.

3.2.4. Icing detection

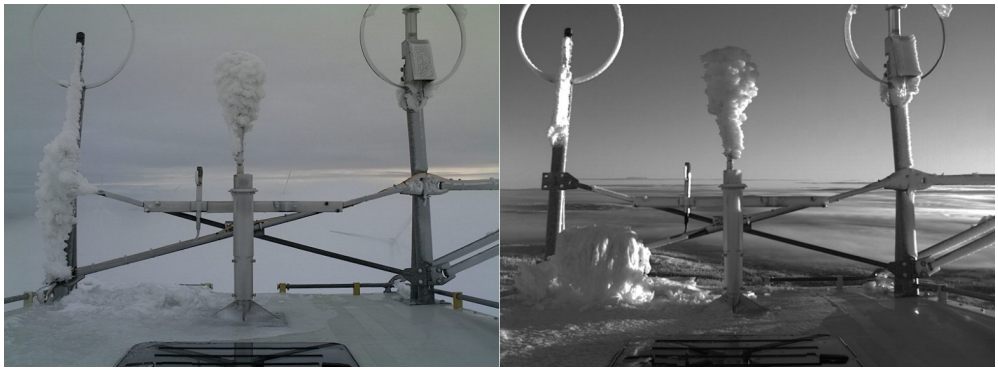


Figure 4: Web-camera images from turbine-hub roof, on the 7th and 8th of January 2013, illustrating the visual assessment of icing. The ice-load instrument is in the centre of each image.

In article II, the icing detection was mainly observed through manual assessments,

i.e. determined by visual inspection of web-camera images (Fig. 4). These observations, of both the presence of ice on the instruments and its growth (active icing), are generally quite reliable. Though, in some cases they are more difficult to determine, such as when thin glazes of icing might be present or when lighting is relatively poor.

3.2.5. Wind-power production measurements

Observations of power production were examined for several wind farms across Sweden, at locations around 800 m above sea level, over five icing seasons: 2009–2014. Because of the confidentiality agreements with wind power companies, the actual production, wind farm names, and certain other data could not be included in this thesis.

4. Summary of the results

4.1. Precipitation accumulation estimates

One main focus in the development of FMI-LAPS was to improve the precipitation accumulation estimates. The research results have been communicated to the scientific community in a series of two articles; Gregow et al. (2013) and Gregow et al. (2017).

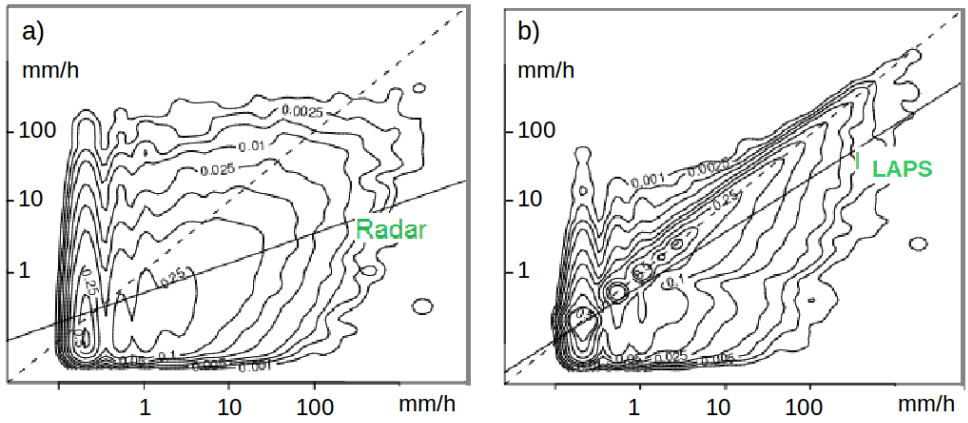


Figure 5: Verification of precipitation accumulation estimates using a) radar alone and b) LAPS-RandB method for summer 2011. Density plots of RandB analysed precipitation accumulation (y-axis: log-scale) against dependent rain-gauge observations (x-axis: log-scale). The solid line is a linear fit to the datasets and the dashed line represents the perfect 1:1 fit in the plots.

The method developed by the author (LAPS-RandB; Section 3.1.1) gives the best results when compared with already-existing methods, e.g. Regression and Barnes, and radar precipitation estimates (Fig. 5). To further investigate and verify the new RandB method, different weather situations were divided into two categories describing their airmass stability: strong convection (i.e. thunderstorms; hereafter "convective") and light-to-moderate convection (i.e. warm-fronts; hereafter "non-convective"). Also with this categorisation, the RandB method performs best, with lower mean absolute errors (MAE) and root-mean-square errors (RMSE), compared with radar data alone (Fig. 6). Verification was done using both dependent and independent observations (i.e. observations that are or are not included, respectively, into the precipitation accumulation analysis). The performance using either the regression or Barnes assimilation analysis separately still yields better results for the accumulation products, compared to precipitation accumulation derived from radar data alone (Gregow et al., 2013).

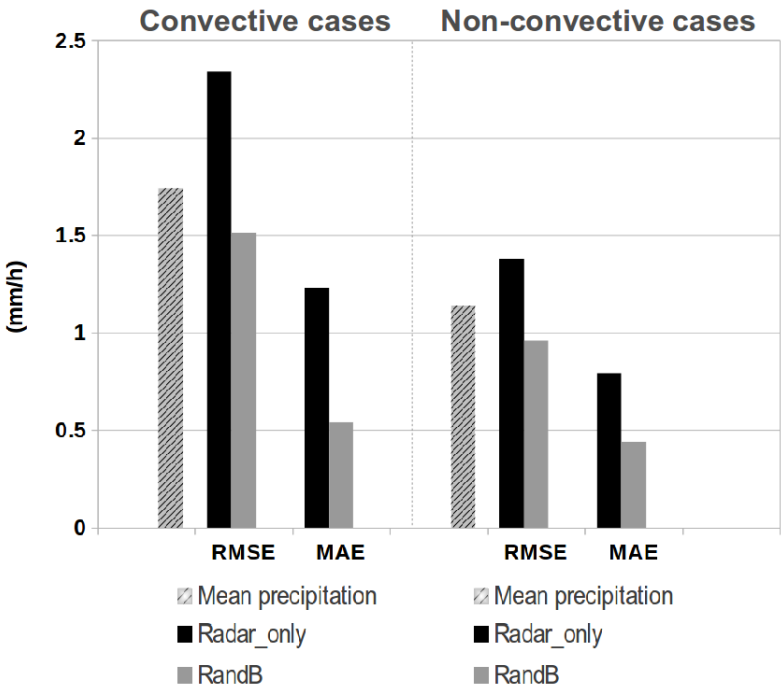


Figure 6: Verification results for RandB method, root-mean-square errors (RMSE) and mean absolute errors (MAE), using dependent raingauge observations for two different airmass stability situations over the Finland area. Left panel: convective cases (i.e. thunderstorms) and right panel: non-convective cases (i.e. warm-fronts). The mean precipitation values, calculating the average of the raingauge data, for all the convective and non-convective cases are included as a hatched bar.

Furthermore, in a case study, the advantage of using the RandB method could be seen during a heavy-precipitation event on the 22 August 2011 at Kaisaniemi station (dependent station), located in the centre of Helsinki. For this event, the radar suffered from attenuation and gave low accumulation values (blue bar), while the LAPS-RandB method improved the analysis (red bar) and the estimated accumulation values were closer to the observed amounts (green bar; Fig. 7).

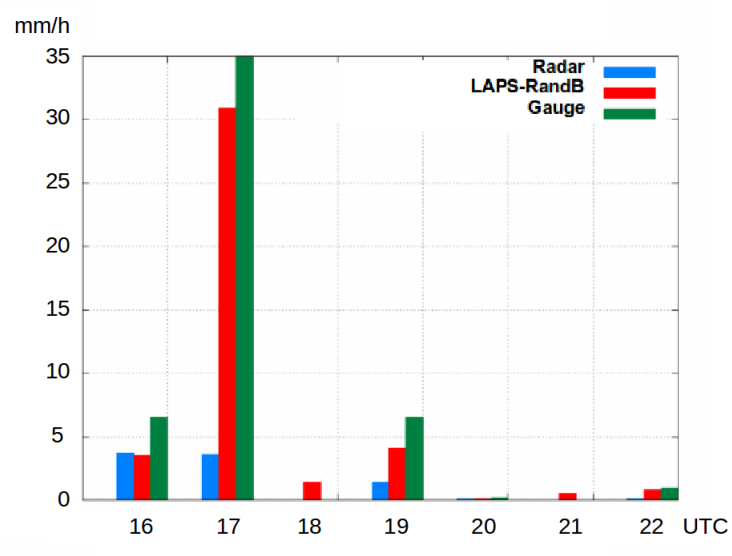


Figure 7: Case study during a heavy rainfall event, at Kaiseniemi Helsinki; 22 August 2011, 16–22 UTC. Radar-estimated precipitation accumulation (blue colour), LAPS-RandB model (red colour) and raingauge observations (green colour).

The integration time of one hour show the best result for LAPS-RandB method, when tested against a selection of longer periods (e.g. the previous 6, 12, 24 hours and 7 days of data; Fig. 8). The verification results from summer 2015 clearly indicate that the RandB method, using radar and raingauge data from the latest hour, gives the best scores for the RMSE and standard deviation (STDEV) verification values (Gregow et al., 2017).

With the use of lightning observations and the newly developed LAPS-LDA method (Section 3.1.2), heavy precipitation cases can be better resolved and the precipitation estimates improved. The studies were performed for summer periods during years 2014–2016 and the verification of the dataset also considered the distance dependencies to radar stations (i.e. gauges situated further away than 100- and 150 km), which showed the same improvement in the results. The strength of the LDA method is that the radar and lightning information can be merged and complement

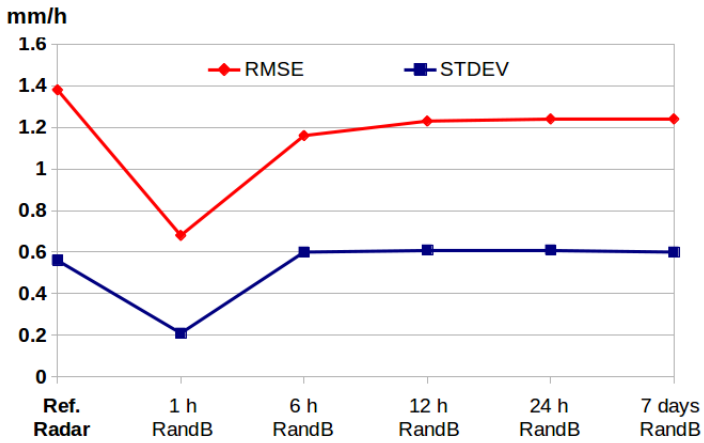


Figure 8: Impact of the integration time on RandB-method for datasets during summer 2015. The scores are shown as RMSE (red colour; mm/h) and STDEV (blue; mm/h) for the different integration times, verified using independent raingauge data. The radar-estimated accumulation (Ref. Radar) scores are included as reference values.

each other. This is especially important in areas of poor, or even no, radar coverage, where the lightning information will improve the hourly precipitation accumulation analysis (Gregow et al., 2017).

4.2. Wind-power production and icing

The LAPS-LOWICE system was developed to produce real-time, hourly estimates of the presence, intensity, and impacts of icing on wind power production. The validation results from one wind farm in January 2013 showed observed average power losses of 19.5%, while the LOWICE estimated power losses were 18.3%. Other validation results can be seen as bars in lower part of figure 9 and the graphs show that the LOWICE (versions V0, V1 and V2, respectively) compares quite well with the time-trends of turbine observed power production. It was demonstrated that wind-power energy losses due to icing were poorly correlated with ice loads. In fact, power losses were most evident when icing was active (e.g. SLWC is present and ice was actively growing at the site) and the intensity of the power losses were more strongly related to the icing rate, rather than ice load (Gregow et al., 2015).

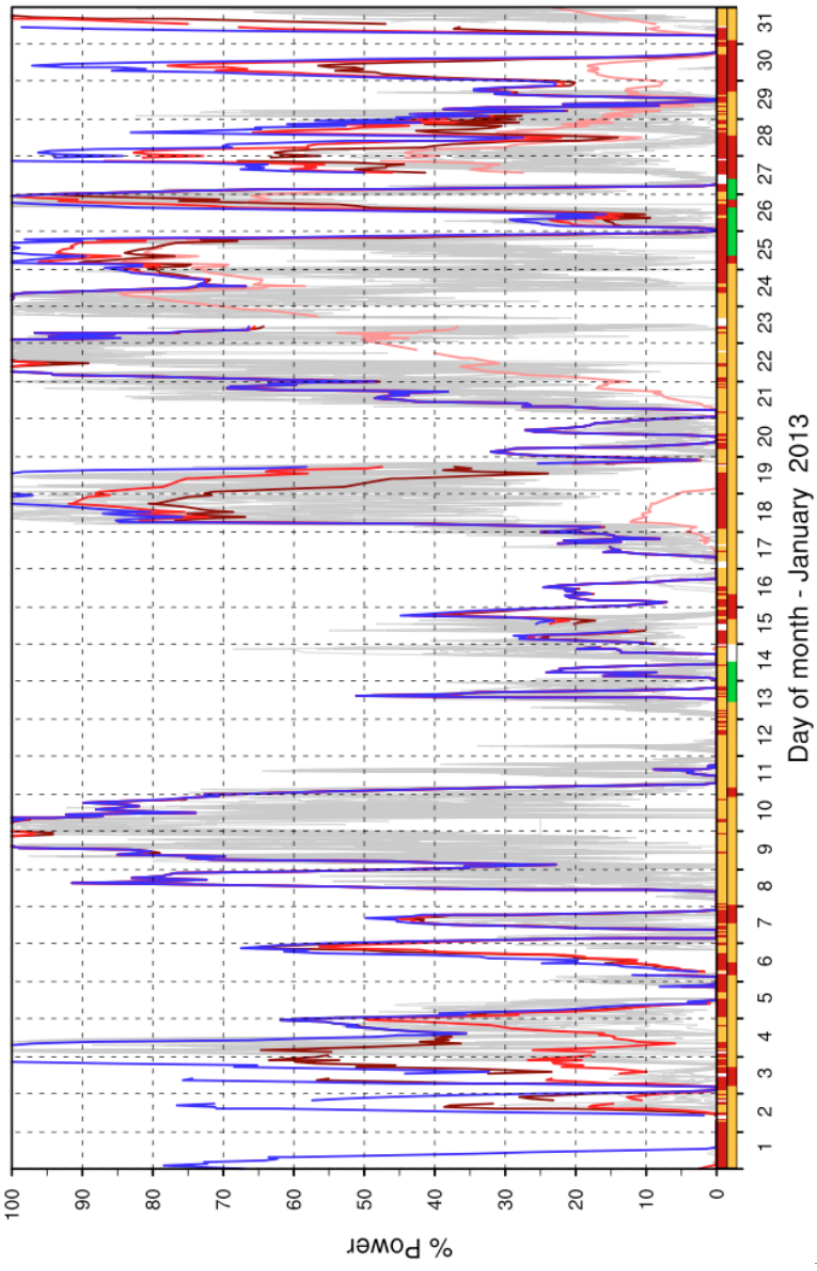


Figure 9: Time-series for January 2013 of the percentage of maximum power from wind turbines (grey), LOWICE clean power (blue) and iced power (red, brown and tan lines for LOWICE versions V2, V1 and V0). Colourbars at the bottom provide an hour-by-hour assessment of the presence of icing as follows: Upper bar; LOWICE-expected icing (red=active icing [rate >10g/h], yellow=inactive icing [rate <10 g/h], green=no icing, white: no information). Lower bar; manually assessed icing through web-camera (red=active icing, yellow=active icing possible, but not confirmed, green=no active icing, white: no information).

5. Discussion

LAPS is capable of using many different observations as input to its analysis and the system is being developed to use new measurements (Fig. 1). With the increase of AWS (by FMI and FTA) and the potential inclusion of data from crowdsourcing and social media (where the stations are unattended and many times the placement of a measurement is not following meteorological standards), there is a risk that data will vary in quality. Datasets may contain erroneous observations caused by for example technical problems (e.g. calibration of instruments) and physical errors (e.g. freezing of wind-anemometers or poor instrument placement), errors which might not be detected for hours or even days, depending on the provider of data. Therefore, the capability to QC the observational datasets is becoming more and more important. FMI's database contains QC'ed observational datasets, data which are used within FMI-LAPS analysis. Also, FMI-LAPS is doing its own QC on several ingest parameters (outlier checks and comparison with fields from forecast model). This becomes especially important when introducing new observations (for example lightning) and with new datasets, which are not necessarily from the FMI database.

In the process of estimating precipitation accumulation, LAPS combines several different observations from raingauges, radar and lightning data. Special focus has been to improve the heavy rainfall events and weather situations which are generally characterized by relatively small spatial scales and with strong vertical motions, such as convection. These precipitation events are known to be difficult in terms of quantitative precipitation estimation (QPE) because: (i) their small spatial scales may not be adequately sampled by gauges, and (ii) radar systems can experience problems due to attenuation of the signal, hail contamination etc. Therefore, problems can occur when strong horizontal reflectivity gradients cause disagreement between radar and raingauge values, since the raingauge is a point observation while radar measures over a volume of the order of 1 km^3. Vertical wind shear, along with strong winds, and hail at gauge station could also contribute to a severe mismatch between radar and gauge measurements. Despite these limitations, examples based upon disdrometer data suggest that generalized relations between two variables are useful over a wide range of remote-sensing problems and a wide range of scales (Jameson and Kostinski, 2001). Data collected over disparate sampling-volumes and sampling-frequencies can be combined to yield meaningful estimates. Although additional testing is required, this allows us to use methods which combine estimates using remote-sensing techniques with sparse but direct rainfall observations. Therefore, in the FMI-LAPS radar–gauge correction method, we assume that: i) raingauge measurements are accurate for the raingauge's location, ii) radar successfully measures relative spatial and temporal variabilities of precipitation, iii) raingauge and radar measurements are valid for the same locations in time and space, and iv) relationships based on comparisons between

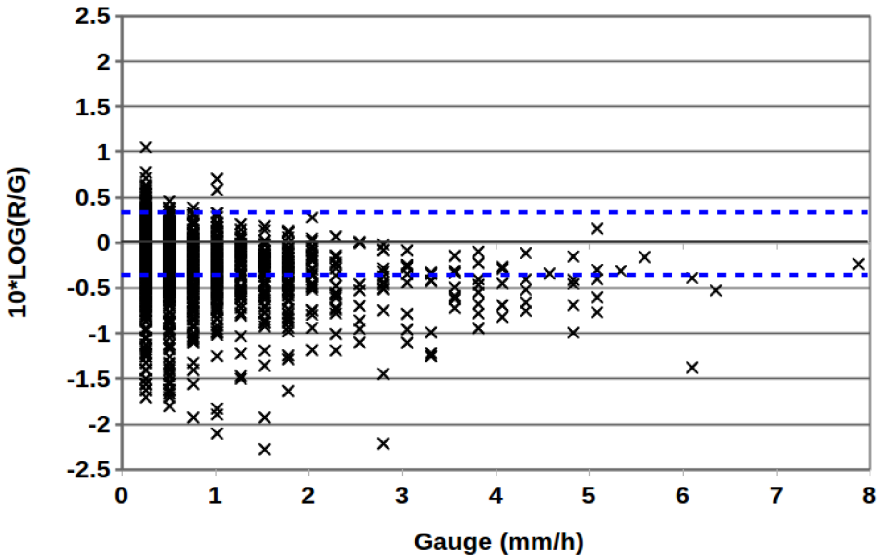


Figure 10: Example of the R/G quotient (y-axis in 10*log-scale) compared with raingauge observations (x-axis) from July 2015. The two blue lines indicate the 2 and 0.5 quotient, respectively.

raingauges and radars are valid for other locations in space and time (e.g. Finland and 1 hour, respectively). Though, in case the measured precipitation amount differs too much between radar and raingauge, which leads to large discrepancies in the radar–gauge quotient (as illustrated in figure 10), we limit the use of datasets. Hence, if $R/G > 2$ or $R/G < 0.5$ it is likely that the radar and raingauge do not measure the same phenomenon, at least they are uncertain values, and the raingauge data is therefore not used in field adjustment (see article I and III).

The use of the LDA represents an important source of information to improve the QPE for convective weather situations. In the LDA method, the correlation between lightning and radar is built upon statistical relationships, such as a period of approximately 1-year for the area of interest. There are concerns regarding the spatial and timeliness representation between these two observation types. Lightning is not always collocated with the reflectivity core and the area of highest lightning frequency varies its position with respect to maximum reflectivity core over the lifetime of the storm (Lopez et al., 1990; Watson et al., 1995). Tapia et al. (1998) show that calculating the precipitation estimates using their lightning–radar method gives geographical variability and potentially exhibit substantial dependence on storm type. The variability of the spatial correlation between lightning and rainfall within the storm area suggests the use of a uniform distribution of rainfall around lightning flashes. They

show that there is a high correlation between the temporal evolution of rainfall rate and lightning frequency, using rainfall distributed uniformly within a 10-km diameter circle around the location of the lightning strike and within 5 minutes. These methods and results correlate with the FMI-LAPS LDA method (see article III).

Snowfall is more difficult to quantify than rain, because of the large variability in density and size of particles, which must be accounted for (Matrosov, 1998). Weather-radar systems are usually calibrated to measure the water equivalent of the reflectivity (Z_e). The dielectric factor (K) included in the Z_e transformation equation is incorporated as a calibration constant, which is not altered when the precipitation form changes from liquid to solid (i.e. snow/hail/graupel). Therefore, the reflectivity-rain rate (e.g. Z–R) relationship has to be changed in order to obtain a phase change of particles (Smith, 1984). Different radar-reflectivity–snow-rate relationships have been investigated throughout the years (von Lerber et al., 2017; Huang et al., 2010; Löffler-Mang and Blahak, 2001), including the use of dual-polarization radars in wintertime conditions (Thompson et al., 2014). Also surface observations (i.e. raingauges) can suffer from incorrect measurements during winter-time (Martinaitis et al., 2014). Therefore, the studies in this compendium focus on the liquid precipitation phase to avoid additional uncertainties of the solid phase explained above.

Satellite-precipitation products — e.g. the NWCSAF geostationary Convective Rainfall Rate (CRR) or polar orbiting Precipitating Clouds (PC) products — are planned to be included in the precipiation estimation process. The challenges with satellite based precipitation products are that the error characteristics are very different from the other measurements (Yilmaz et al., 2005) and new algorithms will need to take this into account when being developed.

Other new observations from crowdsourcing, and possibly even from social media, are becoming interesting to use within weather models. During the summer of 2017, FMI started a pilot-project to collect "citizen observations" and depending on the outcome of the project, there are opportunities to include these new datasets into for example FMI-LAPS. In the future, the variational version of LAPS is of interest because of the use of more-advanced and updated assimilation methods.

FMI-LAPS has been used as the operational analysis system since 2009 and in several other applications at the institute; calculations of gridded forecast products from Model Output Statistics; MOS (Carter et al., 1989; Glahn and Lowry, 1972) predicted station values, post-processing of radiation quantities, and hot-start of HARMONIE and other forecast models to improve the initial boundary conditions, especially hydro-meteors, and thereby improve the short-term forecast (Tiesi et al., 2016).

6. Conclusions

In order to support the FMI operational production needs, there was a clear interest to perform research with the scientific goal of improving the meteorological analysis methods underlying the identified end-user products for i) precipitation accumulation estimates and ii) icing-related quantities.

To more accurately calculate the precipitation accumulation estimate, new methods were developed within the FMI-LAPS framework. This was done by combining both the standard and new high-temporal resolution observations in new ways, taking into account the different measuring scales and error characteristics. The RandB method produces high quality gridded analysis, with improved precipitation estimates for the 1-hour integration time. Applying the LDA method lightning data does add information, especially to areas not covered by radar, and improve the results. With these developments the final precipitation accumulation product became better as of quality and temporal and spatial resolution.

The outcome of a 3-year wind-power pilot-project clearly showed the importance of high-quality meteorological fields, especially within the atmospheric boundary layer, to estimate the icing on wind-turbine blades and their effect on wind-power production. Developments of FMI-LAPS, especially cloud processes, improved the analysis of icing-related quantities (clouds, temperature, humidity etc). The newly developed icing-model LOWICE calculates wind-power production and takes into account the reduction due to icing. Using FMI-LAPS analyses as input to the newly developed icing-model, LOWICE showed good results and it was concluded that it provides useful information for the wind-power industry.

The high-quality gridded FMI-LAPS analysis products are presented at both public web-pages and within FMI in-house visualization tool (SMARTMET), and it has been found to be useful for several nowcasting purposes. Additionally, LAPS products add value when used as input to specific end-user products, such as hydrological models (used at SYKE), the FMI road-weather model and as input for the FMI fire-weather index warning delivered to the public. Thus far, it is clear that the use of FMI-LAPS analysis can be very useful for several applications, models, and meteorological services.

References

Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Weather and Forecasting*, **11**(3), 273–287.

Barnes, S. L., 1964: A technique for maximizing details in numerical weather map analysis. *Journal of Applied Meteorology*, **3**(4), 396–409.

Barnes, S. L., 1973: Mesoscale objective map analysis using weighted time-series observations. *NOAA Tech. Memo*. ERL NSSL-62, National Severe Storms Laboratory, Norman, Oklahoma, 60 pp. [NTIS COM-73-10781].

Barnes, S. L., 1994: Applications of the Barnes objective analysis scheme. Part I: Effects of undersampling, wave position, and station randomness. *Journal of Atmospheric and Oceanic Technology*, **11**(6), 1433–1448.

Bauer, P., G. Ohring, C. Kummerow, and T. Auligne, 2011: Assimilating satellite observations of clouds and precipitation into NWP models. *Bulletin of the American Meteorological Society*, **92**(6), ES25–ES28.

Betts, A. K., 1987: Thermodynamic constraint on the cloud liquid water feedback in climate models. *Journal of Geophysical Research: Atmospheres*, **92**(D7), 8483–8485.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Weather and Forecasting*, **4**(3), 401–412.

Cattin, R., S. Kunz, A. Heimo, G. Russi, M. Russi, and M. Tiefgraber, 2007: Wind turbine ice throw studies in the Swiss Alps. *European Wind Energy Conference Milan*.

Cummins, K. L., M. J. Murphy, E. A. Bardo, W. L. Hiscox, R. B. Pyle, and A. E. Pifer, 1998: A combined TOA/MDF technology upgrade of the US National Lightning Detection network. *Journal of Geophysical Research: Atmospheres*, **103**(D8), 9035–9044.

Daley, R., 1991: Atmospheric data analysis, Cambridge atmospheric and space science series. *Cambridge University Press*, **6966**, 25.

Dybbroe, A., K.-G. Karlsson, and A. Thoss, 2005: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part I: Algorithm description. *Journal of Applied Meteorology*, **44**(1), 39–54.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, **11**(8), 1203–1211.

Gregow, E., B. Bernstein, I. Wittmeyer, and J. Hirvonen, 2015: LAPS–LOWICE: A Real-Time System for the Assessment of Low-Level Icing Conditions and Their Effect on Wind Power. *Journal of Atmospheric and Oceanic Technology*, **32**(8), 1447–1463.

Gregow, E., A. Pessi, A. Mäkelä, and E. Saltikoff, 2017: Improving the precipitation accumulation analysis using lightning measurements and different integration periods. *Hydrology and Earth System Sciences*, **21**(1), 267.

Gregow, E., E. Saltikoff, S. Albers, and H. Hohti, 2013: Precipitation accumulation analysis–assimilation of radar-gauge measurements and validation of different methods. *Hydrology and Earth System Sciences*, **17**(10), 4109–4120.

Häggmark, L., K.-I. Ivarsson, S. Gollvik, and P.-O. Olofsson, 2000: MESAN, an operational mesoscale analysis system. *Tellus A*, **52**(1), 2–20.

Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber, 2011: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the Eastern Alpine region. *Weather and Forecasting*, **26**(2), 166–183.

Hiemstra, C. A., G. E. Liston, R. A. Pielke Sr, D. L. Birkenheuer, and S. C. Albers, 2006: Comparing Local Analysis and Prediction System (LAPS) assimilations with independent observations. *Weather and forecasting*, **21**(6), 1024–1040.

Huang, G.-J., V. Bringi, R. Cifelli, D. Hudak, and W. Petersen, 2010: A methodology to derive radar reflectivity–liquid equivalent snow rate relations using C-band radar and a 2D video disdrometer. *Journal of Atmospheric and Oceanic Technology*, **27**(4), 637–651.

Hyvärinen, O., and E. Saltikoff, 2010: Social media as a source of meteorological observations. *Monthly Weather Review*, **138**(8), 3175–3184.

ISO, E., 2001: Atmospheric icing of structures. *INTERNATIONAL STANDARD. ISO*, **12494**.

Jameson, A., and A. Kostinski, 2001: Reconsideration of the physical and empirical origins of Z–R relations in radar meteorology. *Quarterly Journal of the Royal Meteorological Society*, **127**(572), 517–538.

Jasper, K., J. Gurtz, and H. Lang, 2002: Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. *Journal of Hydrology*, **267**(1), 40–52.

Kalnay, E., 2003: *Atmospheric modeling, data assimilation and predictability*. Cambridge University Press, 342 pp.

Kelly, G., and J.-N. Thépaut, 2007: Evaluation of the impact of the space component of the Global Observing System through Observing System Experiments. *ECMWF Newsletter*, **113**, 16–28.

Koistinen, J., D. B. Michelson, H. Hohti, and M. Peura, 2004: Operational measurement of precipitation in cold climates. *Weather Radar*. Springer, 78–114.

Koskinen, J. T., J. Poutiainen, D. M. Schultz, S. Joffre, J. Koistinen, E. Saltikoff, E. Gregow, H. Turtiainen, W. F. Dabberdt, J. Damski, et al., 2011: The Helsinki Testbed: a mesoscale measurement, research, and service platform. *Bulletin of the American Meteorological Society*, **92**(3), 325–342.

Lahoz, W., B. Khattatov, and R. Menard, 2010: *Data assimilation: making sense of observations*. Springer Science & Business Media, 718 pp.

Le Dimet, F.-X., and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A: Dynamic Meteorology and Oceanography*, **38**(2), 97–110.

Löffler-Mang, M., and U. Blahak, 2001: Estimation of the equivalent radar reflectivity factor from measured snow size spectra. *Journal of Applied Meteorology*, **40**(4), 843–849.

Lopez, R., W. Otto, R. Ortiz, and R. Holle, 1990: The lightning activity of convective cloud systems in northeastern Colorado. *Preprints, 18th Conf. on Interactive Information and Processing Systems for Meteorology*. Oceanography and Hydrology, Orlando, FL, AMS, 727–731.

Lorenc, A., 1981: A global three-dimensional multivariate statistical interpolation scheme. *Monthly Weather Review*, **109**(4), 701–721.

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

Martinaitis, S., S. Cocks, Y. Qi, B. Kaney, J. Zhang, and K. Howard, 2014: An examination of the impacts of frozen precipitation on gauge networks during winter precipitation events. *Extended Abstract, 39th Natl. Wea. Assoc, Annual Meeting*.

Matrosov, S. Y., 1998: A dual-wavelength radar method to measure snowfall rate. *Journal of Applied Meteorology*, **37**(11), 1510–1521.

Pessi, A., 2013: Characteristics of lightning and radar reflectivity in continental and oceanic thunderstorms. *93th Annual American Meteorological Society Meeting, Austin, Texas, USA*, volume 610.

Peura, M., 2002: Computer vision methods for anomaly removal. *Proc. ERAD*, 312–317.

Saltikoff, E., A. Huuskonen, H. Hohti, J. Koistinen, and H. Järvinen, 2010: Quality assurance in the FMI Doppler weather radar network. *Boreal environment research*, **15**(6), 579–594.

Schneider, S., B. Chimani, H. Kaufmann, B. Bica, C. Lotteraner, S. Tschannett, and R. Steinacker, 2008: Nowcasting of a supercell storm with VERA. *Meteorology and atmospheric physics*, **102**(1-2), 23–36.

Seed, A. W., C. E. Pierce, and K. Norman, 2013: Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme. *Water Resources Research*, **49**(10), 6624–6641.

Smith, P. L., 1984: Equivalent radar reflectivity factors for snow and ice particles. *Journal of Climate and Applied Meteorology*, **23**(8), 1258–1260.

Sun, J., and H. Wang, 2013: Radar data assimilation with WRF 4D-Var. Part II: Comparison with 3D-Var for a squall line over the US Great Plains. *Monthly Weather Review*, **141**(7), 2245–2264.

Talagrand, O., 1997: Assimilation of Observations, an Introduction. *Journal of the Meteorological Society of Japan. Ser. II*, **75**(1B), 191–209.

Tapia, A., J. A. Smith, and M. Dixon, 1998: Estimation of convective rainfall from lightning observations. *Journal of Applied Meteorology*, **37**(11), 1497–1509.

Thompson, E. J., S. A. Rutledge, B. Dolan, V. Chandrasekar, and B. L. Cheong, 2014: A dual-polarization radar hydrometeor classification algorithm for winter precipitation. *Journal of Atmospheric and Oceanic Technology*, **31**(7), 1457–1481.

Tiesi, A., M. M. Miglietta, D. Conte, O. Drofa, S. Davolio, P. Malguzzi, and A. Buzzi, 2016: Heavy Rain Forecasting by Model Initialization With LAPS: A Case Study. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **9**(6), 2619–2627.

Vivoni, E. R., D. Entekhabi, R. L. Bras, V. Y. Ivanov, M. P. Van Horne, C. Grassotti, and R. N. Hoffman, 2006: Extending the predictability of hydrometeorological flood events using radar rainfall nowcasting. *Journal of Hydrometeorology*, **7**(4), 660–677.

von Lerber, A., D. Moisseev, L. F. Bliven, W. Petersen, A.-M. Harri, and V. Chandrasekar, 2017: Microphysical Properties of Snow and Their Link to Z_e–S Relations during BAECC 2014. *Journal of Applied Meteorology and Climatology*, **56**(6), 1561–1582.

Watson, A. I., R. L. Holle, and R. E. López, 1995: Lightning from two national detection networks related to vertically integrated liquid and echo-top information from WSR-88D radar. *Weather and forecasting*, **10**(3), 592–605.

Yilmaz, K. K., T. S. Hogue, K.-l. Hsu, S. Sorooshian, H. V. Gupta, and T. Wagener, 2005: Intercomparison of rain gauge, radar, and satellite-based precipitation estimates with emphasis on hydrologic forecasting. *Journal of Hydrometeorology*, **6**(4), 497–517.

Zawadzki, I., 1984: Factors affecting the precision of radar measurements of rain. *22nd Conference on Radar Meteorology, Zurich, Switzerland*, 251–256.

I

Reprinted, with permission, from
Hydrol. Earth Syst. Sci., 17, 4109–4120,
doi:10.5194/hess-17-4109-2013

Hydrol. Earth Syst. Sci., 17, 4109–4120, 2013
www.hydrol-earth-syst-sci.net/17/4109/2013/
doi:10.5194/hess-17-4109-2013

Hydrology and Earth System Sciences
Open Access

Precipitation accumulation analysis – assimilation of radar-gauge measurements and validation of different methods

E. Gregow[1], **E. Saltikoff**[1], **S. Albers**[2,3], **and H. Hohti**[1]

[1]Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
[2]NOAA/ESRL/Global Systems Division, Boulder, Colorado, USA
[3]Cooperative Institute for Research in the Atmosphere, Fort Collins, Colorado, USA

Correspondence to: E. Gregow (erik.gregow@fmi.fi)

Received: 25 January 2013 – Published in Hydrol. Earth Syst. Sci. Discuss.: 28 February 2013
Revised: 31 July 2013 – Accepted: 2 September 2013 – Published: 23 October 2013

Abstract. We investigate the appropriateness of four different methods to produce precipitation accumulation fields using radar data alone or combined with precipitation gauge data. These methods were validated for high-latitude weather conditions of Finland. The reference method uses radar reflectivity only, while three assimilation methods are used to blend radar and surface observations together, namely the linear analysis regression, the Barnes objective analysis and a new method based on a combination of the regression and Barnes techniques (RandB). The Local Analysis and Prediction System (LAPS) is used as a platform to calculate the four different hourly accumulation products over a 6-month period covering summer 2011. The performance of each method is verified against both dependent and independent observations (i.e. observations that are or are not included, respectively, into the precipitation accumulation analysis). The newly developed RandB method performs best according to our results. Applying the regression or Barnes assimilation analysis separately still yields better results for the accumulation products compared to precipitation accumulation derived from radar data alone.

1 Introduction

The concept of precipitation accumulation is of great importance for various applications in meteorology and hydrology. Climate projections under possible climate change scenarios point to likely higher frequency of storms, with intensified precipitation over Europe. This will most probably have a significant effect on the surface water balance, therefore having a large impact on society and its economical aspects. Hydrological models, which are based on analyzed precipitation accumulation, do need a very high accuracy of the precipitated water amount in order to issue warnings, e.g. for sudden flooding. Fire and weather warnings are another example of products where end-users require high-quality data of precipitation accumulation during the summer period.

Radar-derived precipitation products are generated at high spatial resolution but embed measurement uncertainties. On the other hand, surface precipitation observations, such as standard gauge observations and road-weather measurements, have usually higher accuracy and are essential when used for correcting radar-based precipitation accumulation fields, but have limited spatial representativeness. The literature provides many studies on the benefits one can gain from the combination of radar measurements and surface observations to derive the final accumulated precipitation product (Goudenhoofdt and Delobbe, 2009). Radar reflectivity generates a good first guess for the accumulated precipitation, with the advantage of high spatial resolution, though there are certain inherent inaccuracies when deriving this product from radars (Koistinen and Michelson, 2002). Measurements of precipitation at ground level are performed at point location and the errors associated with the observations are well characterized (Steiner et al., 1999). Different, more or less sophisticated assimilation methods exist, whereby surface point observations are blended together with radar data in order to establish a corrected precipitation accumulation, e.g.: co-kriging (Sun et al., 2000), the statistical objective analysis method (Pereira et al., 1998), combined bias-adjustments method (Overeem et al., 2009) and bias

adjustments using the Kalman filter (Chumchean et al., 2006; Anagnostou and Krajewski, 1999). A summary of the methods and operational usage in different countries is compiled in the COST-717 report (Gjertsen et al., 2003). Problems linked to radar-gauge bias correction methods have been discussed in, e.g. Seo and Breidenbach (2002).

In this study, we use the Local Area and Prediction System – LAPS (McGinley et al., 1991, 1992) as a platform for testing and validating 4 different precipitation accumulation analyses: the radar only (hereafter LAPS_radar) and 3 assimilation methods, namely the linear analysis regression, the Barnes objective analysis and a combination of those two methods (hereafter Regression, Barnes and RandB, respectively). Here the RandB is a new method, while the three others are more widely used. Geostatistical methods have shown good results in other studies for daily accumulation sums (e.g. Goudenhoofdt and Delobbe, 2009). However, they are sensitive to networks density, and the density of stations measuring hourly precipitation in Finland is very low. Therefore, in this paper we concentrate on further development of methods already used in LAPS, such as Regression and Barnes. LAPS is applicable for operational usage (Albers et al., 1996; Amy 2003), which is of critical interest for end-users who demand as close to real-time products as possible.

According to the classic Köppen classification, the climate of southern coastal Finland belongs to class Dfb and the rest of the country to Dfc, i.e. a cool and moist continental, subarctic climate of cold and snowy winters and precipitation throughout the year. Summer is warm, not hot, and in the north it is also short (Jylhä et al., 2010). The only mountains are in northern Finland but do not exceed 1350 m, while Finland is embraced by two Gulfs of the Baltic Sea (Gulf of Finland and Bothnian Bay) from two sides.

The aim of this article is to test and validate our new RandB method against three conventional methods, for typical high latitudes summer weather conditions encountered in Finland (extending between 60 and 70° N) and to provide some guidance in the use of these methods. Section 2 introduces the LAPS model (Sect. 2.1), the radar data (Sect. 2.2) and the gauge network data (Sect. 2.3). The different analysis methods for estimating precipitation accumulation are introduced in Sect. 3. The results are presented and analysed in Sect. 4, while Sect. 5 provides some conclusions and outlooks.

2 Methods and material

We describe here the model and data used to determine the gridded background fields involved in the estimation of the precipitation accumulation.

2.1 The Local Analysis and Prediction System (LAPS)

The Finnish Meteorological Institute (FMI) operates the Local Analysis and Prediction System (LAPS) for production of 3D analysis fields of different weather parameters (Albers et al., 1996). LAPS uses a data fusion method, in which a high-resolution spatial analysis, using statistical methods, is performed on top of the coarser resolution background fields. Observations are fitted to the coarser first-guess analysis mainly by successive correction method, while high-resolution topographical data sets are taken into account when creating the final high-resolution analysis fields. Those analysis products are mainly used for now-casting purposes; i.e. what is currently happening and what will happen in the next few hours.

The coarser background first-guess field is the latest available forecast from the European Centre for Medium-Range Weather Forecasts (ECMWF) model, with a current horizontal grid spacing of approximately 16 km (ECMWF, 2011). The following ECMWF parameters are used at 16 vertical pressure levels: vertical velocity, specific humidity, temperature, geopotential, vectorized winds, surface geopotential, surface pressure, pressure at mean sea level, 2 m temperature and dew-point temperature, vectorized wind at 10 m, sea surface temperature, skin temperature and land-sea mask.

The FMI LAPS setup uses a pressure coordinate system, including 44 vertical levels distributed with a higher resolution (e.g. 10 hPa) at lower altitudes and decreasing with height. The horizontal resolution is 3 km and the domain used in this article covers the entirety of Finland and some parts of the neighbouring countries (see Fig. 1a).

The fine-scale structures in the resulting 3-D analysis are extracted from the observations. Therefore, LAPS highly relies on the existence of high-resolution, both spatial and temporal, observational network and especially on remote sensing data. At present, the LAPS suite implemented at FMI is able to process several types of in situ and remotely sensed observations such as radar reflectivity, weighting gauges, road-weather observations, radar radial winds, soundings, Synop, Metar, air traffic observations, lidars and Meteosat9 satellite data. The first three of these listed measurements are used for calculating the precipitation accumulation within LAPS. The Finnish radar volume scans are read into LAPS as NetCDF format files, thereafter the data is remapped to LAPS internal Cartesian grid and the mosaic process combines data of the different radar stations (Albers et al., 1996). In LAPS the rain rates are calculated from the lowest levels of the LAPS 3-D radar mosaic data, via the standard $Z - R$ equation formula (Marshall and Palmer, 1948), which is then used for precipitation accumulation calculations, either as radar only accumulation, see Sect. 3.1, or merged with gauge observations, see Sects. 3.2, 3.3 and 3.4.

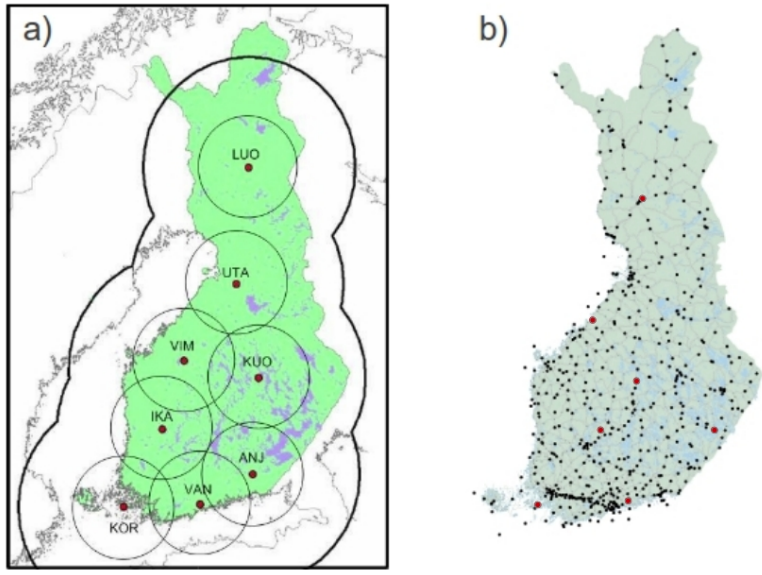


Fig. 1. (a) The rectangular frame of the map depicts the LAPS analysis domain. The red dots represent the 8 Finnish radar stations and the thick, black curved lines display their coverage. The thin circles surrounding each radars represent the areas where measurements are performed below 2 km height. **(b)** The Finnish surface gauge network (dots on the map) used to measure precipitation accumulation. The red dots indicate the position of the seven "independent" stations used for the verification.

2.2 The radar network

FMI operates eight C-band Doppler radars, which nearly cover the whole country. In southern Finland, the distance between radars is 140–200 km and measurements are made in bins that are 500 m long and 1° wide, up to 250 km in range. Thus, data from two or three radars are available over most of the study area. The location of the radars and their coverage is shown in Fig. 1a. As Finland has no high mountains, the horizon of all the radars is near 0° elevation with no major beam blockage, and, in general, the radar coverage is excellent up to 68° N latitude.

The effective radar reflectivity factor Z_e (usually called reflectivity) is derived from the expression

$$Z_e = \frac{P_r \cdot r^2}{L \cdot C \cdot K^2}, \quad (1)$$

where P_r is the average received microwave power, r is the measurement range, L is the two-way attenuation in the propagation path (antenna – scatterers – antenna), C is a radar constant (including parameters of the radar hardware) and K is the dielectric factor (depending on the relative fraction of ice and water in the hydrometeors).

The reflectivity uses dBZ as a unit, which is expressed as

$$\mathrm{dBZ} = 10 \cdot \log_{10} Z_e. \quad (2)$$

The uncertainty factors affecting radar reflectivity are the electronic miscalibration, beam blocking, and attenuation due to both precipitation (Battan, 1973) and wet radome (Germann, 1999). At mid-latitudes, the main source of uncertainty of radar-based rainfall estimates is the vertical profile of reflectivity (VPR), which causes a range-dependent error (Zawadski, 1984). At large distances, the radar probes the upper parts of the cloud, where reflectivity is weaker. In FMI's general radar processing chain, this is compensated with the VPR correction, which also compensates for overestimation in a melting layer when appropriate (Koistinen et al., 2003). The radar ingest to LAPS system, used in this study, processes original 3D volumes and therefore no VPR correction is needed. Before the radar volume data is ingested into LAPS, clutter is removed with Doppler-filtering and any residual clutter with a post-processing procedure based on fuzzy logics (Peura, 2002).

The output of a weather radars is reflectivity, Z, which depends on sum of sixth power of drop diameter. When converting reflectivity to precipitation intensity, one has to assume the size of measured drops. The real drop size distribution is highly variable depending on the type of precipitation, but because it is usually unknown, a default drop size distribution is used (Battan, 1973). This leads to errors when the drop sizes differ from average values. It has been noted, both in literature and in our experiments, that during small-droplet precipitation (drizzle), the gauges usually give larger values

compared to radar, with a factor often exceeding values of 30. On the contrary, in large drop situations, typically related to heavy precipitation cases (rain showers with embedded cumulonimbus clouds), the observed gauge-to-radar ratio often gets less than 0.25. This discrepancy is related to the use of the standard $Z - R$ equation formula (Marshall and Palmer, 1948) for all liquid precipitation cases, even though we know that drop size distributions vary from one precipitation case to another. Another well-known factor causing differences between measurements with gauge and radar is the radar beam overshooting in shallow drizzle events. These circumstances could breed a substantial impact on the analysis and therefore the gauge-to-radar ratio has to be controlled carefully (see Sects. 3.2 and 3.3).

Comparing radars and gauges, an additional challenge arises from the different sampling sizes of the instruments. Radar measurement volume can be several kilometres wide and thick (one degree beam is ca. 4 km wide at a 250 km distance from antenna), while the measurement area of a gauge is 400 cm^2 (weighting gauges) or 100 cm^3 (optical instruments). The measurements in the FMI network have been designed to use the radar composite in Cartesian grid of 1 km × 1 km. Details of the FMI radar network and processing routines are described in Saltikoff et al. (2010).

In this study, the radar data were used as volume measurements, repeated every 5 min and consisting of 5 elevation angles, typically between 0.4 and 45°. LAPS processes the radar data directly onto its own gridded coordinate system, which has a resolution of 3 km × 3 km.

2.3 Surface observations

For this study, a total of 447 rain gauges, both weighting gauges and optical sensors, provide detailed point information, which is used to correct the radar first-guess field (introduced in Sect. 2.2). The verification period ranges from 11 April and 14 October 2011, i.e. by and large the non-winter season (no-snow-phase precipitation).

The surface precipitation observations are from standard weighting gauges and optical sensors mounted on road-weather masts. Weighting gauges are subject to different sources of random errors such as mechanical malfunction, wind drift (Hanna, 1995) and icing, which all affect the accuracy of measurements. FMI manages 77 stations instrumented with the weighting gauge Vaisala model VRG101. Measurements with this instrument have high cumulative accuracy (0.2 mm) provided that the precipitation event exceeds 0.5 mm. Depending on the station, the gauges measure the accumulated precipitation in intervals of 10 to 60 min. Summing these measurements over a 60 min period yields 1 h accumulation data.

The Finnish Transport Agency (FTA) runs 370 road-weather stations with optical sensor measurements (Vaisala Present Weather Detectors models PWD11 and PWD22), which have a precipitation detection sensitivity of 0.05 mm h^{-1} or less, within 10 min. The precipitation intensity is measured in intervals ranging between 10 s and 5 min and finally summed up to 1 h precipitation accumulation information. A performance study between PWD22 sensor and VRG weighting gauges against Geonor weighting gauges has been done by Wong (2012). The study shows that the PWD22 has a larger negative mean error (underestimation) and a more than four times larger standard error than the VRG. The Finnish road-weather station sites have not been selected for best meteorological quality or representativeness. Hence they may have additional uncertainties connected to their location in the immediate vicinity of roads with heavy traffic, where splash effects and wind eddies, generated by big vehicles, occasionally affect the resulting accumulation. Such effects would be hard to quantify, and as the FTA mainly need qualitative information of precipitation, they have not published accuracy estimates of these measurements.

Another source of uncertainty in surface accumulation observations results from the limited spatial representativeness of many stations with respect to their surroundings, due to the insufficient density of measuring stations for certain areas (Cherubini et al., 2002). Note that if measurements consistently indicate poor data quality, those stations are blacklisted within LAPS and do not contribute to the precipitation accumulation analysis. Hereafter in this article, the weighting gauges and road-weather measurements are indistinctly called gauges and their distribution in Finland is shown in Fig. 1b.

3 Description of the four analysis methods

Thanks to its high-resolution reflectivity pattern, weather radar data provide the best first-guess to calculate precipitation accumulation. The radar-based accumulation is calculated in the LAPS routine with the standard $Z - R$ equation formula (Marshall and Palmer, 1948). On the other hand, gauges usually measure the accumulation with higher quality and are consequently used to correct the radar field. In this study, three different assimilation methods have been tested in the LAPS routines as to their capacity to perform the best radar-gauge correction: the Regression, the Barnes and new RandB methods. These methods use the quotient between gauge and radar (hereafter G/R) for their corrections.

3.1 LAPS_radar-based accumulation

The reflectivity Z parameter measured by the radar is converted to precipitation intensity R (mm h^{-1}) within LAPS accumulation process (see Sect. 2.1), using a pre-selected $Z - R$ equation (Marshall and Palmer, 1948) as of the type

$$Z = A \cdot R^b, \quad (3)$$

where A and b are empirical factors describing the shape and size distribution of the hydro-meteors. In FMI's

implementation of LAPS we used $A = 315$ and $b = 1.5$ for liquid precipitation, which is relevant in this study carried out during the summer period. This is a gross simplification since the drop size and particle shapes vary according to weather situation (drizzle/convective, wet snow/snow grain), as described in Sect. 2.2. Problematic situations include both convective showers with heavy rainfall and the opposite case of drizzle with little precipitation. Although such situations contribute only a fraction of the annual precipitation amount, they might be important during, e.g. flooding events. On the other hand, the same factors have been used for many years in FMI's other operational radar products, and looking at long-term averages, the radar accumulation data match the gauge accumulation values within reasonable accuracy (Aaltonen et al., 2008). After correcting for vertical profile of reflectivity (Sect. 2.2), mainly due to major sampling differences between the two sensors, random errors remain at 2–3 dB, which is a typical, reasonably accurate figure in operational radar measurements (Koistinen et al., 2003; Collier, 1986).

In LAPS the intensity field (R in Eq. 3) is calculated every 5 min and the 1 h accumulation is thereafter obtained by summing up over the 5 min intervals.

The linear regression analysis method as described above, in addition to sampling differences, such as accumulation estimates based only on radar data, can differ from gauge observation values either due to radar errors (see Sect. 2.2) or problems with the gauges (Sect. 2.3). This is why various statistical methods have been used to address and reduce these differences; for example, a model using a regression method is described in Sokol (2003). In the linear regression analysis method (hereafter Regression method) used in this article, as a first step, the gauge-radar pairs from a given grid point undergo a quality check to prohibit dubious differences between gauge and radar values. The aim is to avoid comparisons involving uncertain radar measurements and spurious surface observations. The selection is performed by discarding gauge-radar pairs exceeding specific thresholds based on the G/R quotient. The thresholds are based on approximately 2 times standard deviation, STDEV (R/G), from LAPS_radar dependent data set (see Table 1). The thresholds used in the Regression method within the LAPS routine are as follows:

- if $G/R > 2.0$ then the gauge-radar pair is discarded;
- if $G/R < 0.5$ then the gauge-radar pair is discarded.

The first threshold handles surface observations that are suspected to be false. The second criteria attempt to avoid cases where the radar gives too high a reflectivity, for example in strong convective precipitation (including hail). Once these criteria are enforced, the remaining data form a data set of representative gauge-radar pairs from which a linear regression can be established, calculated with the least square method, which minimized the errors between the measurement pairs. The outcome are values for k and c in the linear regression formula

$$Y = k \cdot X + c. \quad (4)$$

The next step is to calculate the newly corrected radar estimate using Eq. (4). Here, Y is the corrected radar estimate, X is the first-guess accumulation from radar and the regression coefficients, with k (the slope) and c (the interception point with the y axis) derived from the regression analysis.

The Regression method has the limitation of requiring a large number of valid gauge-radar pairs in order to fulfil the least square calculations and thereby creating a sufficient linear curve fit between the gauge network and radar observations. If there are not enough valid pairs, or if the criteria for a linear dependency are not fulfilled, then the regression method will not be used and the analysis will fall back to the original LAPS_radar-based initial precipitation accumulation field. The behaviour of the linear curve has to be constrained since the shape of the curve is strongly influenced by the amount of gauge-radar pairs. Criteria for this have been set so to constrain k values between 0.2 and 5.0, and c values between -5 and $+5$ mm, in Eq. (4). These constraints were based on average vertical profile adjustments of reflectivity and relates to ranges of up to 200 km from radar station, during the summer period (Koistinen et al., 2003). The linear function is applied to the whole radar accumulation field, i.e. corresponds to a regional-scale correction.

Table 1. Statistical verification results of the different methods for the dependent stations data set.

	LAPS_radar	Regression	Barnes	RandB
Number of observations	111 821	102 016	111 821	111 821
STDEV (R/G)	1.11	1.23	0.53	0.55
STDEV ($\log(R/G)$)	0.51	0.51	0.37	0.38
RMSE	1.38	1.32	1.03	0.98
MAE	0.73	0.69	0.43	0.39
RMSE–MAE	0.85	0.63	0.60	0.59
CORR	0.51	0.56	0.69	0.72

3.2 Barnes objective analysis method

The Barnes interpolation forces the radar field to converge towards gauge accumulation measurements, using an objective multi-pass telescoping strategy (Barnes, 1964, Heimstra et al., 2006) in the LAPS routine. The G/R quotient is used to interpolate the first-guess radar field closer to the observation value and in order to optimize the result, several iteration steps are performed within the Barnes analysis at successively finer scales. For grid points far from any G/R observations, the G/R field tends smoothly towards a value of 1.

Depending on the precipitation pattern, this method can potentially result in a highly overestimated or underestimated reflectivity field being spread to the surroundings. For example, if there is one ground station situated at the border of a convective rain shower (cumulonimbus cloud), where only

light precipitation occurs, the G/R quotient would probably exceed the value of 30 in this case, as described in Sect. 3. For the station point itself, this quotient gives an adequate correction but spreading this large quotient to the surrounding precipitation pattern could potentially give very large overestimates of the accumulation within, for example in this case, the nearby core of a rain shower with heavy precipitation. Quality checks and thresholds have been set to avoid situations where such over- or underestimations of nearby precipitation areas are likely. If the G/R quotient gives very large (more than 30) or very small (less than 0.25) values, this might still give a signal of an adequate trend, even though the signal is overamplified. This trend has to be maintained and adapted but is given less weight in the resulting accumulation. Consequently, the chosen criteria must incorporate these aspects. The thresholds for the Barnes G/R quotient are based on approximately 2 times standard deviation, STDEV (R/G), from the LAPS_radar-dependent data set (see Table 1). The following thresholds were used:

- if $0.25 < G/R < 2.0$ then allow the derived quotient;
- if $0.25 < G/R$ and $G/R > 0.0$ then reset $G/R = 0.25$;
- if $2.0 < G/R$ and $G/R < 100.0$ then reset $G/R = 2.0$.

The modified Barnes scheme allows weighting (w_0) with distance (d) from the gauge station point with respect to the radius of influence (r), normalized by the instrument error (err_0), which is here set to be 1.5 in Eq. (5). The G/R increment gives the initial increment (p_0) at the first iteration step, and the background weight (w_b), set to 0.02, adjusts the output to be closer to radar value further away from the observation point in Eq. (6).

$$w_0 = \frac{e^{-\left(\frac{d}{r}\right)^2}}{\mathrm{err}_0^2}, \tag{5}$$

$$p_{ij} = \frac{\sum (p_0 \cdot w_0)}{\sum p_0 + w_\mathrm{b}} \tag{6}$$

After the first iteration step, the p_{ij} output becomes the new G/R increment (p_0) for the next iteration step in Eq. (6). The iterations continue with successively decreasing values of r, by a factor of 2 for each iteration, in Eq. (5) until the observation increments have been diminished to a preset value in LAPS, in this case RMSE = 0.13 mm, or alternatively after 10 iteration steps in order to minimize the calculation time.

3.3 New method, combination of Regression and Barnes methods

This new method combines the above described Regression and Barnes analyses. First, the Regression method is used to correct the overall radar estimate, i.e. a regional-scale correction. The resulting accumulation field is thereafter used as a new first guess, initializing the Barnes analysis, which rectifies the radar field on local scales. Assuming that the new first-guess field from the Regression analysis is closer to the real precipitation accumulation, the Barnes correction method will not need to be too aggressive in its correction, thus minimizing the risk of exaggerating the surrounding precipitation with too low, alternatively too high, G/R quotients.

4 Results and verification

The performance of the different methods has been verified against surface gauge observations of precipitation accumulation data. The verification period spans from 11 April to 14 October 2011, therefore assuming precipitation is in the form of liquid water, and the time sampling interval is one hour. The observations have been divided into two subsets: (i) one set including observations of all stations (but 7 of them) and (ii) a group of 7 Synop stations (excluded from the former set) used as an objective data set for verification (Figs. 2–3 and 4–5, respectively). Accordingly, in the calculation of the 1 h precipitation accumulation, the analysis depends on the station information from the first subset (i), hereafter called "dependent" stations, while the accumulation analysis is independent of the 7 stations in the second subset (ii), hereafter called "independent" stations. As the total number of gauge stations in Finland is low, compared to radar pixels, and the experiment was run using the operational system (i.e. results are used in end-users applications), we could not set more stations aside without risking the quality of the end product. The seven independent stations were selected subjectively from different physiographical areas such as coastline, inland, lake district, and proximity to each other. On average, within a radius of 50 km from the independent station point, there are 11 dependant stations and the average distance to the nearest dependant station is 9.8 km.

The statistical quantification of the validation of the different analysis methods are based on the root mean square error (RMSE Eq. 7), and the mean absolute error (MAE Eq. 8), calculated with these data sets:

$$\mathrm{RMSE} = \sqrt{\frac{\sum (\text{Analysis-Gauge})^2}{N}}, \tag{7}$$

$$\mathrm{MAE} = \frac{\sum (|\text{Analysis-Gauge}|)}{N}. \tag{8}$$

RMSE is a quadratic scoring rule, which measures the average magnitude of the error. Since the errors are squared before they are averaged, RMSE gives a relatively high weight to large errors. MAE measures the average magnitude of the errors in a set of analyses, without considering their direction. It measures the accuracy for continuous variables. MAE is a linear score, which means that all the individual differences are weighted equally in the average. MAE and RMSE can be used together to diagnose the variation in the errors

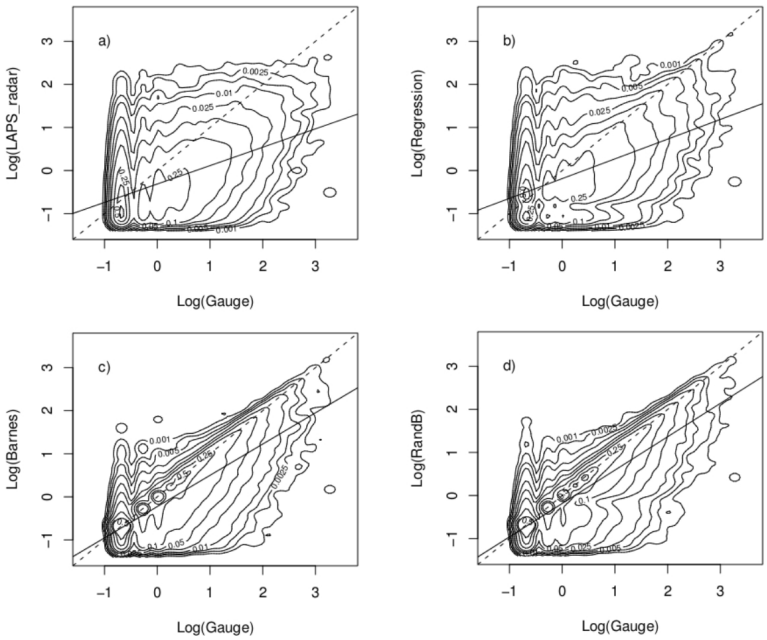


Fig. 2. Density plots of analyzed precipitation accumulation (y axis) against observed rain-gauge values (x axis) for the dependent stations: **(a)** LAPS_radar; **(b)** Regression; **(c)** Barnes, and **(d)** RandB. The continuous line is a linear fit to the data set and the dashed line represents the perfect 1 : 1 fit in the plots.

in a set of analyzes. RMSE will always be larger or equal to MAE. The greater the difference between them (RMSE–MAE), the greater the variance in the individual errors in the sample (see Tables 1 and 2). If RMSE = MAE, then all the errors are of the same magnitude.

Results are shown as density plots with logarithmic scales, where data points less than 0.3 mm h^{-1} are discarded in order to avoid artificial effects due to different detection sensitivities of the different instruments (criteria applied in Figs. 2–5). In Fig. 2 we show, separately for the four different methods, the relationship between the analyzed accumulation data at the LAPS grid point closest to a gauge station and the corresponding gauge observations for the dependent stations. The correlation calculated from the data sets and the statistics of the comparisons are compiled in Table 1. It appears from these comparisons that the new RandB method yields the best agreement for accumulation precipitation compared to gauge observations, though the Barnes method also provides reasonable results. On the other hand, the regression method alone is not very successful but still improves the accumulation analysis to some extent. The LAPS_radar method, which is based on radar information only, gives the poorest results in our study.

In order to investigate the error dependencies between radar and gauges, we use an indicator that describes the hydrological aspects of the errors (Szturc et al., 2011), namely, the absolute difference between observed and analyzed precipitation accumulation as a function of the magnitude of the observed value (i.e. gauge data). Figure 3 shows that the linear fit has a smaller angle coefficient as one passes from the LAPS_radar, to Regression, Barnes and RandB analysis methods. This shows that the departure between analyzed and observed values decreases and again the RandB analysis performs best of the different methods.

We next investigate the agreement between the analyzed precipitation accumulation values and observations (gauge values) for the independent stations (Table 2). Note that for independent stations, there is much less data available.

Table 2. Statistical verification results of the different methods for the independent stations data set.

	LAPS_radar	Regression	Barnes	RandB
Number of observations	2648	2436	2648	2648
STDEV (R/G)	1.67	1.47	1.41	1.19
STDEV ($\log(R/G)$)	0.47	0.47	0.37	0.37
RMSE	1.29	1.23	0.95	0.91
MAE	0.72	0.68	0.44	0.40
RMSE–MAE	0.57	0.55	0.51	0.51
CORR	0.60	0.65	0.80	0.81

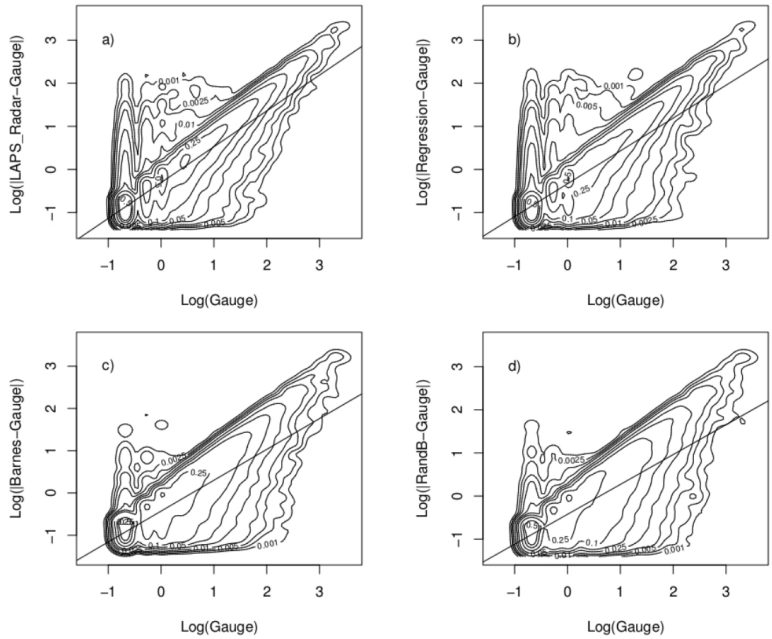


Fig. 3. Absolute value of the difference between observed and analyzed precipitation accumulation (y axis) plotted against observed rain-gauge values (x axis), for the dependent stations: **(a)** |LAPS_radar-Gauge|; **(b)** |Regression-Gauge|; **(c)** |Barnes-Gauge|, and **(d)** |RandB-Gauge|. The continuous line is a linear fit to the data set.

Through the independent stations we want to prove that the methods also work for areas where there are no observing stations available. Thus, verifying that there are no over- or underamplified accumulation patterns devolving especially from the Barnes method (see Sect. 3.3), but also from the Regression method. The density plots (Fig. 4) indicate less scatter and slightly better agreement, i.e. smaller RMSE, MAE and higher correlation coefficient, compared to the dependent stations analysis (Fig. 2). The linear fitted curves in Fig. 4 are strongly influenced by the small amount of observation points, because the data is not normally distributed, hence the distribution of high accumulation values (i.e. corresponding to over 10 mm h^{-1}) have a large impact on the fitted curve. The comparison between the linear fitted curves in Fig. 4a–d gives a clear indication of how the different methods compare to each other. We also plotted the absolute difference between analyzed precipitation accumulation and observation as a function of gauge observations for the independent stations (Fig. 5). The same trend is observed as with dependent station data: less dependence of Barnes and RandB methods, compared with LAPS_radar and Regression methods.

In Sects. 2.2 and 3.1 we gave an explanation for the errors that are attributed to radar measurements, such as the range-dependent error and $Z - R$ inaccuracies. These errors are related to the prevailing weather situation (e.g. thunderstorms or warm fronts) and, hence, the type of precipitating hydro-meteors occurring at that time. Such influence was further investigated by dividing the different weather situations into two categories describing their air-mass stability: strong convection (hereafter convective) and light-moderate convection (hereafter non-convective), which relates to thunderstorms and warm fronts, respectively. Each category includes 10 cases of a full 24 h day, also selected from the period 11 April to 14 October 2011. The convective cases were determined by using FMI's lightning location system (Tuomi and Mäkelä, 2008) together with FMI radar archive, while the non-convective (warm front) cases were selected from analyzed frontal passages over southern Finland as tagged by the duty forecaster at FMI.

The data set representing the convective weather situations have fewer data values, compared to warm front cases (see # values in Fig. 6). This is expected since convective precipitation is less likely to hit a gauge measuring device and generally last for shorter time, while large-scale precipitation events occurring during warm fronts, have a much higher probability to come across a gauge station and have a larger temporal and spatial dimension. The results (Fig. 6) clearly show that the convective cases give larger RMSE and MAE values, compared to non-convective cases. This is expected, as convective precipitation situations display more

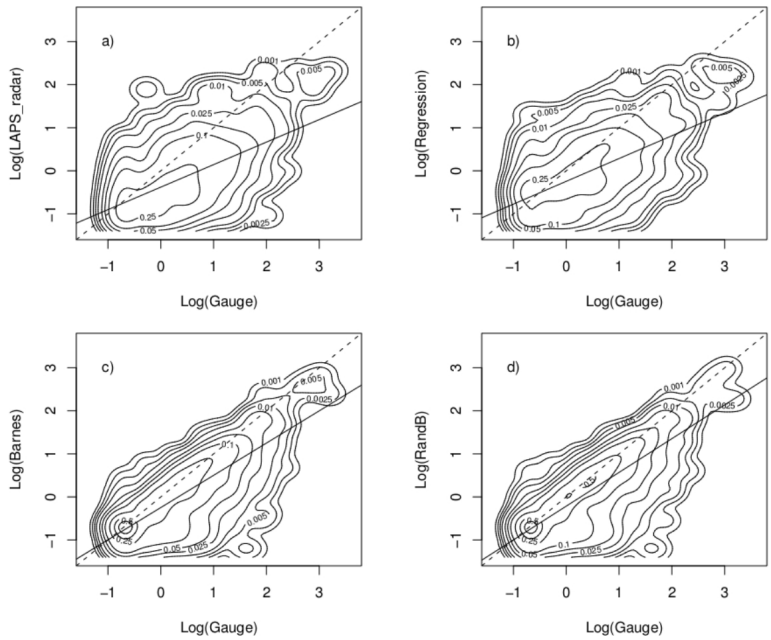


Fig. 4. Density plots of analyzed precipitation accumulation (y axis) plotted against rain-gauge values (x axis), for the seven independent stations: **(a)** LAPS_radar; **(b)** Regression; **(c)** Barnes, and **(d)** RandB. The continuous line is a linear fit to the data set and the dashed line represents the perfect 1 : 1 fit in the plots.

spatial heterogeneity and thus a stronger decoupling from the gauge observations. This categorisation also indicates that the RandB method performs best out of the four different methods, though only slightly better than the Barnes method.

5 Discussions and conclusions

In this article we compare the results from 4 different analysis methods on how to calculate the hourly precipitation accumulation: LAPS_radar, Regression, Barnes and a new developed method RandB (combination of Regression and Barnes). The LAPS_radar serves as the reference method and since it is based on the common $Z - R$ formula, this method is also similar to what is used at many meteorological services. The LAPS_radar is further used as the first-guess field when merging gauges' data into the analysis routine of the three other methods. As described in Sect. 3.2, the Regression method benefits from having many gauge-radar pairs, since it will then create a more robust statistical relationship between the measurements. In cases with no valid pairs, or if the criteria for a linear dependency are not fulfilled, the analysis will become the same as the original LAPS_radar-based accumulation field. The Barnes method will in the same way fall back to the original LAPS_radar-based accumulation field if there are no observations available, or if the radar-gauge pairs do not fulfil the thresholds stipulated for the G/R quotient. The new RandB method encounters the same restrictions as described above, since it is a combination of the Regression and Barnes methods. In order to be meaningful for operational purposes, the studied merging methods should therefore show at least as good a result as the LAPS_radar precipitation accumulation analysis. Figures 2, 4 and 6 confirm that applying an assimilation method improves the overall results. In Figs. 3b–d and 5b–d one can see that the density values congregate closer to the zero value along the x axis, indicating a better match between analyzed and observed value. The calculated statistics, including both the dependent, independent, convective and non-convective data sets, also state that agreement is improved by applying a merging method. The error values of RMSE and MAE are decreasing, compared to LAPS_radar values, and for the RandB method with the dependent data set, the corresponding reduction in RMSE and MAE are 29 and 47 %, respectively. The correlation, for RandB dependent data set, increases (41 %) accordingly and the variance (RMSE–MAE) decreases when applying the different assimilation methods. Similar results are seen in the independent, convective and non-convective data sets.

When studying the results from two different stability weather situations, i.e. convective and non-convective, the

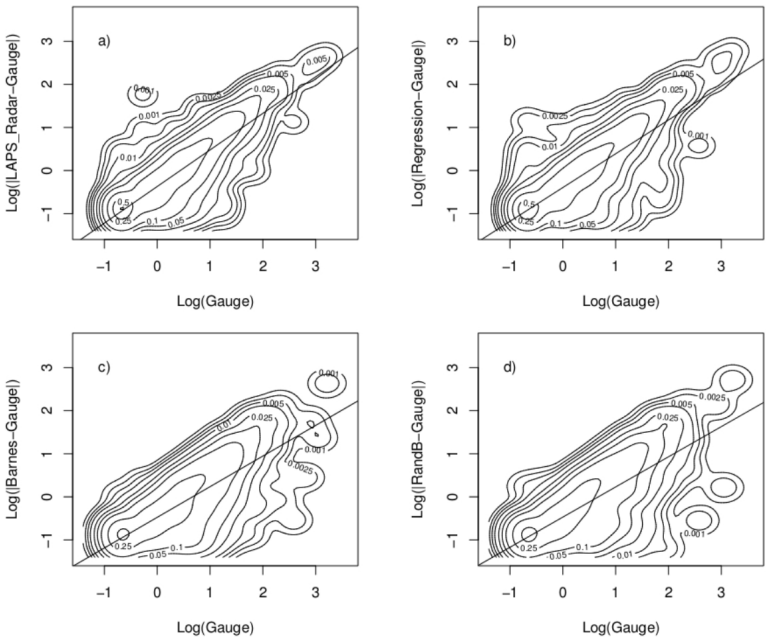


Fig. 5. Absolute value of the difference between observed and analyzed precipitation accumulation (y axis) plotted against rain-gauge values (x axis), for the seven independent stations: **(a)** |LAPS radar-Gauge|; **(b)** |Regression-Gauge|; **(c)** |Barnes-Gauge|, and **(d)** |RandB-Gauge|. The continuous line is a linear fit to the data set.

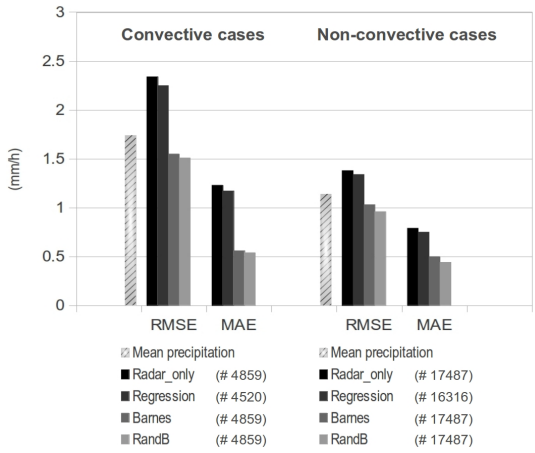


Fig. 6. Statistical verification results for the four different accumulation methods split into two different air-mass stability situations; left panel: convective cases (i.e. thunderstorms) and right panel: non-convective cases (i.e. warm fronts). The symbol # indicates the number of observations used in the calculations. The mean precipitation for each case, calculated from rain-gauge values, is included as a dashed stack.

main findings are that the RMSE and MAE are considerably higher in convective cases. This indicates that the four accumulation methods adopted in this study are more sensitive to convective situations. We interpret that this is related to the larger spatial variability of convective precipitation as well as different drop size distributions. In convective situations, the real intensity is variable within each radar measurement bin (typically representing several cubic kilometres), and it is a random process, which is only partly captured at a single gauge (orifice diameter of 22.6 cm). Also the $Z - R$ equation used in Finland has been optimized for total rainfall, which in areas of extra-tropical cyclones consists largely of frontal precipitation, e.g. warm fronts. As a consequence, when the discrepancy between radar and gauge observations (i.e. large G/R quotients) is significant for the convective cases, the thresholds (see Sects. 3.2 and 3.3) are more frequently exceeded within the Regression, Barnes and RandB analyses. This leads to fewer corrections being done from the gauge measurements and the resulting accumulation analysis is worse for convective weather situations, compared to non-convective cases.

On the other hand, optimising the $Z - R$ equation for some specific types of precipitation should lead to a more faithful merging, which should be reflected in the agreement between analysed and observed precipitation. When such

approach would be performed, using a much larger data set basis, the RMSE and MAE value of the agreement for specific precipitation types should naturally tend towards better performance than without any differentiation between precipitation types, and could be used thus as a test.

The conclusive results from this study are that the newly developed RandB method, i.e. the combination of Regression and Barnes analysis methods, generates the best estimate of 1 h precipitation accumulation. Also, applying either Barnes or Regression methods separately still yields a better result than solely using radar accumulation, i.e. LAPS_radar method.

Acknowledgements. We want to thank Victoria Sinclair for her help in finding the warm front cases, Sylvain Joffre for advice and encouragement, Otto Hyvärinen for the help with creating the density plots in R-tool and the reviewers for their helpful and constructive comments.

Edited by: R. Deidda

References

Aaltonen, J., Hohti, H., Jylhä, K., Karvonen, T., Kilpeläinen, T., Koistinen, J., Kotro, J., Kuitunen, T., Ollila, M., Parvio, A., Pulkkinen, S., Silander, J., Tiihonen, T., Tuomenvirta, H., and Vajda, A.: Strong precipitation and urban floods (Rankkasateet ja taajamatulvat RATU), Finnish Environment Institute (Suomen ympäristökeskus), Helsinki, Finland, 80 pp., 2008.

Albers, S. C., McGinley, J. A., Birkenheuer, D. L., and Smart, J. R.: The local analysis and prediction system (LAPS): Analyses of clouds, precipitation, and temperature, Weather Forecast., 11, 273–287, 1996.

Amy, H.: Precipitation Verification of the Local Analysis and Prediction System (LAPS) Storm Total Precipitation Estimates, St. Cloud State University and NOAA/NWS Weather Forecast Office, Minnesota, 2003.

Anagnostou, E. N. and Krajewski, W. F.: Real-time radar rainfall estimation Part I: Algorithm formulation, J. Atmos. Ocean. Tech., 16, 189–197, 1999.

Barnes, S. L.: A technique for maximizing details in numerical weather map analysis, J. Appl. Meteorol., 3, 396–409, 1964.

Battan, L. J.: Radar Observation of the Atmosphere, University of Chicago Press, Chicago, 1973.

Cherubini, T., Ghelli, A., and Lalaurette, F.: Verification of Precipitation Forecasts over the Alpine Regions Using a High-Density Observing Network, Weather Forecast., 17, 238–249, 2002.

Chumchean, S., Sharma, A., and Seed, A.: An Integrated Approach to Error Correction for Real-Time Radar-Rainfall Estimation, J. Atmos. Ocean. Tech., 23, 67–79, 2006.

Collier, C. G.: Accuracy of rainfall estimates by radar, Part I: Calibration by telemetering raingauges, J. Hydrol., 83, 207–223, 1986.

ECMWF: available at: http://www.ecwmf.int/research/ifsdocs/CY37r2/index.html, last access: 30 December 2011.

Germann, U.: Radome attenuation – a serious limiting factor for quantitative radar measurements?, Meteorol. Z., 8, 85–90, 1999.

Gjertsen, U., Salek, M., and Michelson, D. B.: Gauge-adjustment of radar-based precipitation estimates – a review, COST-717 working document No. WDD 02 200310 1, SMHI, Norrköping, Sweden, 27 pp., available at: available at: http://www.smhi.se/cost717/doc/WDD_02_200310_1.pdf, 2003.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.

Hanna, E.: How effective are tipping-bucket raingauges? A review, Weather, 50, 336–342, 1995.

Hiemstra, C. A., Liston, G. E., Pielke, R. A., Birkenheuer, D. L., and Albers, S. C.: Comparing Local Analysis and Prediction System (LAPS) Assimilations with Independent Observations, Weather Forecast., 21, 1024–1040, doi:10.1175/WAF961.1, 2006.

Jylhä, K., Tuomenvirta, H., Ruosteenoja, K., Niemi-Hugaerts, H., Keisu, K., and Karhu, J. A.: Observed and projected future shifts of climatic zones in Europe and their use to visualize climate change information, Weather Clim. Soc., 2, 148–167, 2010.

Koistinen, J. and Michelson, D. B.: BALTEX weather radar-based precipitation products and their accuracies, Boreal Environ. Res., 7, 253–263, 2002.

Koistinen, J., Michelson, D. B., Hohti, H., and Peura, M.: Operational Measurement of Precipitation in Cold Climates, in: Weather Radar Principles and Advanced Applications, edited by: Meischner, P., Springer, Germany, 78–114, 2003.

Marshall, J. S. and Palmer, W. M.: The Distribution of raindrops with size, J. Meteorol., 5, 165–166, 1948.

McGinley, J. A., Albers, S. C., and Stamus, P. A.: Validation of a composite convective index as defined by a real-time local analysis system, Weather Forecast., 6, 337–356, 1991.

McGinley, J. A., Albers, S. C., and Stamus, P. A.: Local Data Assimilation and Analysis for Nowcasting, Adv. Space Res., 12, 179–188, 1992.

Overeem, A., Holleman, I., and Buishand, A.: Derivation of a 10-year radar-based climatology of rainfall, J. Appl. Meteorol. Clim., 48, 1448–1463, 2009.

Pereira Fo, A. J., Crawford, K. C., and Hartzell, C. L.: Improving WSR-88D Hourly Rainfall Estimates, Weather Forecast., 13, 1016–1028, 1998.

Peura, M.: Computer vision methods for anomaly removal, in: Second European Conference on Radar Meteorology, ERAD, Delft, the Netherlands, 312–317, 2002.

Saltikoff, E., Huuskonen, A., Hohti, H., Koistinen, J., and Järvinen, H.: Quality assurance in the FMI Doppler Weather radar network, Boreal Environ. Res., 15, 579–594, 2010.

Seo, D. J. and Breidenbach, J. P.: Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements, J. Hydrometeorol., 3, 93–111, 2002.

Sokol, Z.: Utilization of regression models for rainfall estimates using radar-derived rainfall data and rain gauge data, J. Hydrol., 278, 144–152, 2003.

Steiner, M., Smith, J. A., Burges, S. J., Alonso, C. V., and Darden, R. W.: Effect of bias adjustment and rain gauge data quality control on radar rainfall estimation, Water Resour. Res., 35, 2487–2503, 1999.

Sun, X., Mein, R. G., Keenan, T. D., and Elliott, J. F.: Flood estimation using radar and raingauge data, J. Hydrol., 239, 4–18, 2000.

Szturc, J., Ośródka, K., and Jurczyk, A.: Quality index scheme for quantitative uncertainty characterization of radar-based precipitation, Meteorol. Appl., 18, 407–420, 2011.

Tuomi, T. J. and Mäkelä, A.: Thunderstorm climate of Finland 1998–2007, Geophysica, 44, 29–42, 2008.

Wong, K.: Performance of Several Present Weather Sensors as Precipitation Gauges, in: WMO TECO, 16–18 October 2012, Brussels, Belgium, 25 pp., 2012.

Zawadzki, I.: Factors affecting the precision of radar measurement of rain, in: 22nd Conference on radar meteorology, 10–13 September 1984, Zurich, Switzerland, 251–256, 1984.

Reprinted, with permission, from
Journal of Atmospheric and Oceanic Technology, 32, 1447–1463,
DOI:10.1175/JTECH-D-14-00151.1

II

LAPS–LOWICE: A Real-Time System for the Assessment of Low-Level Icing Conditions and Their Effect on Wind Power

E. GREGOW

Finnish Meteorological Institute, Helsinki, Finland

B. BERNSTEIN AND I. WITTMEYER

Leading Edge Atmospherics, Longmont, Colorado

J. HIRVONEN

Finnish Meteorological Institute, Helsinki, Finland

(Manuscript received 8 August 2014, in final form 16 April 2015)

ABSTRACT

The wind power industry is highly sensitive to weather, and there is a clear impact on turbine efficiency associated with icing, which can cause significant power losses and even result in the total shutdown of wind farms. Therefore, accurate analyses and forecasts of wind- and icing-related meteorological variables are of great importance. To this end, the Local Analysis and Prediction System (LAPS)–LOWICE system has been developed to produce real-time, hourly estimates of the presence, intensity, and impacts of icing on wind power production. As part of this development, it became clear that power losses did not correlate well with measured icing loads but correlated reasonably well with the time history of icing rate in combination with ice loss due to melting, sublimation, and shedding.

1. Introduction

The expanding use of renewable energy implies new developments within the wind power industry, and the global market for wind energy grew more than 56 billion Euro during 2011–12 (GWEC 2012). In Europe, wind power is now the fastest-growing source of energy (EWEA 2013). The energy available from wind is highly variable in time and space; therefore, better wind prediction may allow power companies to determine how, when, and where to transfer the energy generated into the electricity network. It is important to recognize the large economic impact such decisions can make. For example, the buying and selling of electricity is highly dependent on the forecasts of the potential for wind energy that is expected in the coming hours and days. Beyond the direct and obvious effects of wind speed, the accretion and persistence of ice can have a large impact on turbine efficiency and thus the amount of electricity generated. In some cases, the effects of icing from supercooled clouds, freezing precipitation, and wet snow can result in complete power loss and even shutdown of wind farms (Tammelin et al. 2000). Icing can also cause dynamic instability on a turbine due to irregular distribution of ice, resulting in fatigue and faster wear of mechanical parts. Furthermore, accumulated ice can fall from, or be thrown from, the turbine when it detaches from the blades and other surfaces, posing a potential risk of injuries to humans and nearby structures (Cattin et al. 2007). These elements of wind power are becoming increasingly important as more wind power capacity is being built in cold climates, where there is a large threat of icing.

Accurate analysis and forecasting of wind speed and icing can allow meteorologists to provide power companies and traders with valuable estimates of expected power and can allow those controlling the turbines to make informed decisions about when to activate deicing equipment and/or to shut down turbines to reduce the risks of damage to the turbines and their surroundings. Thus, there is a clear need for timely, high-quality analyses

Corresponding author address: Erik Gregow, Finnish Meteorological Institute, P.O. Box 503, FIN-00101 Helsinki, Finland.
E-mail: erik.gregow@fmi.fi

DOI: 10.1175/JTECH-D-14-00151.1

and forecasts of wind power and the effects of icing thereon.

It is with these concepts in mind that a wind power icing diagnostic system has been developed for Scandinavia, combining output from the Finnish Meteorological Institute's Local Analysis and Prediction System (LAPS: Albers et al. 1996) model and the LOWICE algorithm (Bernstein et al. 2011a). LAPS–LOWICE produces real-time, hourly estimates of wind speed and the presence, rate, and impacts of icing on wind power production. LOWICE employs multiple approaches to estimate the proportion of power expected to be lost due to icing. The first, simplest method [version 0 (V0)] ties power loss solely to the amount of ice "load," while newer versions [version 1 (V1) and version 2 (V2)] tie power loss to the recent history of icing rate, as well as the return of power due to melting, sublimation, and shedding of previously accreted ice. In this paper, LAPS–LOWICE is described and the output will be validated and verified against observations from a wind farm in Sweden.

2. Methods and material

Estimation of the likelihood and intensity of near-surface icing conditions is a challenging problem for meteorologists and those affected by this phenomenon, including the power industry. Beyond direct measurements of icing, wind, and power at a given location, there are numerous sources of independent meteorological data that can be used to estimate near-surface icing conditions indirectly. In particular, observations from satellites, radiosondes, surface stations, and radars (if available) provide a great deal of useful information, especially when paired with forecasts from numerical weather models. Each of these data sources has its strengths and weaknesses for the analysis and forecasting of icing conditions, and the information from each must be considered carefully in the context of the meteorological environment if information from them is to be employed effectively (as in Tafferner et al. 2003; Le Bot 2004). By blending 3D numerical model fields and observations from sources such as those described above in a manner that is consistent with the meteorology of icing, LAPS–LOWICE is able to produce high-resolution grids of icing probability, icing rate, and ice load, as well as expected "clean" (ice free) and "iced" power across Scandinavia.

a. LAPS

The Finnish Meteorological Institute (FMI) operates the LAPS to produce 3D analyses of a number of meteorological parameters (Albers et al. 1996; Koskinen et al. 2011). LAPS uses a data fusion method to generate high-resolution analyses using statistical methods to merge coarser-resolution background fields. Observations are fitted to the coarser first-guess analysis mainly by successive correction, while high-resolution topographical datasets are taken into account when creating the final analysis fields. Those analysis products are mainly used for nowcasting purposes, that is, what is currently happening and what will happen in the next few hours.

In FMI-LAPS, background first-guess fields are derived from the European Centre for Medium-Range Weather Forecasts (ECMWF) model grids, with horizontal grid spacing of ~16 km (ECMWF 2014). This includes 3D forecasts of geopotential height, temperature, specific humidity, and winds at 16 pressure levels. Single-level variables include surface geopotential height, surface pressure, 2-m temperature, and 10-m wind. FMI-LAPS output is produced at 3-km horizontal spacing across Scandinavia (Fig. 1) and on 44 vertical (pressure) levels, with 10-hPa spacing near the surface. LAPS relies heavily on high-resolution data from remote sensors and ground-based observational networks, including Meteosat satellite, surface synoptic observations (SYNOPs) and METARs, precipitation gauges, road weather sensors, soundings, AMDAR temperatures and winds, and radar data. LAPS output fields include 3D pressure, geopotential height, temperature, winds, relative humidity, model condensate, and both 2D and 3D cloud fields [e.g., cloud-top temperature, cloud-top and cloud-base height, fractional cloud cover, and cloud layering; described in more details in section 2a(1)]. Many of these LAPS fields are highly relevant for both wind power and icing, providing essential inputs to the LOWICE system.

LAPS CLOUD ANALYSIS

One of the dominant causes of icing on wind power plants is the freezing of supercooled liquid cloud drops onto turbine blades. To determine the presence or absence of clouds (the "cloud mask"), their heights, layering, etc., LAPS relies heavily on satellite data from Meteosat. During daytime, visible and multiple infrared (IR) channels can be used to determine the cloud mask, while only infrared channels are used at night. To estimate the cloud-top height, longwave IR temperatures are compared to the LAPS temperature and height profiles. LAPS then assimilates other existing measurements, such as surface observations (METARs and SYNOPs) to assess the cloud-base height (Fig. 2).

In northern latitudes, there is a high frequency of low clouds due to frontal systems, maritime influences, and temperature inversions. Especially during winter, these single-layered low clouds form due to strong temperature inversion, where the condensate is trapped and potentially induces a risk of icing. In these situations, it has been common for cloud-top heights to be overestimated (as in

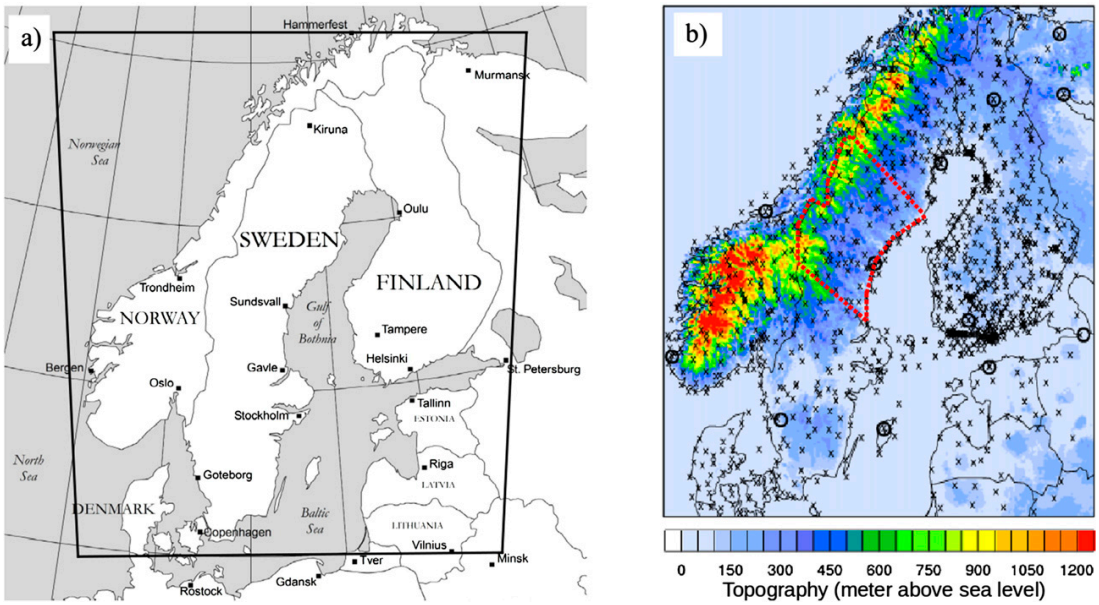


FIG. 1. (a) Areal coverage of LAPS analysis domain is shown with a black line. (b) Surface observations (black crosses) and sounding stations (open black circles) ingested into LAPS, together with the topography height (see color bar). The wind farm is located in the mountains of western Sweden, at an elevation of approximately 800 m above mean sea level. The location is within the enclosed area (red dotted line) in (b), which describes the region SE2 of the Swedish electrical power net.

Haggerty et al. 2008). LAPS searches the temperature profile downward, in order to match it with the measured cloud-top IR temperature, and the cloud top is put at the first vertical level where these temperatures are found. When inversions are present, multiple levels may meet this criterion. However, this was not considered in the original LAPS cloud-top identification method. Therefore, new algorithms have been added to locate subinversion clouds in weather situations where only a single-layer low cloud is present. When the new method finds a temperature inversion, the cloud height is set below the thermal inversion if the satellite infrared temperature is greater than the air temperature below the inversion by at least 10°C. As a result, the vertical cloud structure can be significantly changed in weather situations with strong low-level inversions. This cloud inversion method follows the same principles as those applied to Meteosat data by EUMETSAT (SAFNWC 2013).

In addition, the FMI-LAPS cloud analysis (Fig. 2) now uses cloud mask products from the Nowcasting Satellite Application Facilities at EUMETSAT (SAFNWC) to correct the cloud mask in the final stage of cloud analysis (Derrien and Le Gléau 2005). There are certain problems within the cloud analysis that are very difficult to solve, such as twilight effects (dawn/dusk), resulting in weak information from satellite channels, and how to distinguish between open and frozen water bodies, which sometimes cause overestimations of clouds over sea/lakes, especially during nighttime (Dybbroe et al. 2005).

b. LOWICE

The basic concepts behind LOWICE are similar to those used in the current icing product (CIP; Bernstein et al. 2005, 2006), which combines numerical model output with observations from satellite, surface stations, pilot reports, and a lightning detection network to produce 3D analyses of the probability and intensity of icing

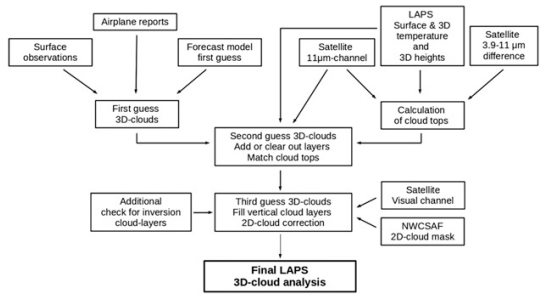


FIG. 2. FMI-LAPS cloud analysis flow diagram, describing the sequential processes of assimilating different input data, after Albers et al. (1996).

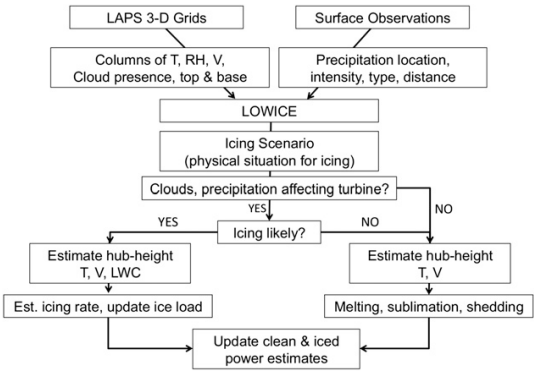


FIG. 3. LOWICE icing and power assessment flow diagram.

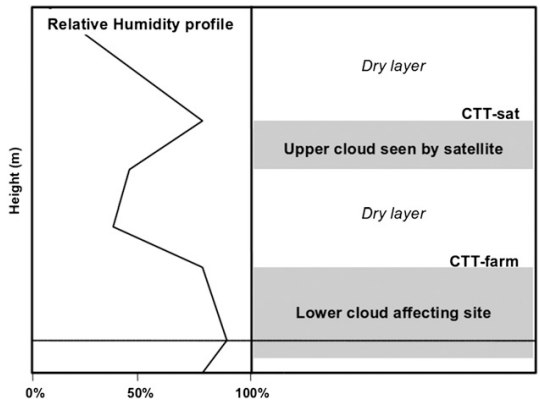


FIG. 4. Conceptual model of the LOWICE multilayer cloud scheme. (left) The RH profile is shown. (right) The CTT-sat and CTT-farm represents the CTT of the upper and lower cloud layers (gray), respectively. The horizontal solid line represents the elevation of the site (i.e., turbine hub height).

conditions aloft over the contiguous United States for the aviation community. LOWICE includes numerous unique concepts and code that are specific to the near-surface icing environment. In sections 2b(1)–2b(3) the LOWICE processes used to estimate wind power and power loss due to icing will be described.

1) ICING DETECTION, LIKELIHOOD, AND INTENSITY

Using the meteorological fields described above, LOWICE examines each vertical column of LAPS data. First, if clouds are present, then the vertical profiles of cloud and relative humidity (RH) are examined, along with observations of precipitation from surface stations, to determine whether the following environments appear to be present: single-layer clouds, multilayer clouds, classical freezing rain, nonclassical freezing drizzle (Huffman and Norman 1988), or a snow-dominated precipitation process. For each of these environments, the column of data must be examined uniquely to correctly assess the expected likelihood and intensity of icing near the surface, where wind turbines operate. Icing is considered possible at the turbine level if the temperature is subfreezing and the level is located between the highest cloud top and the lowest cloud base. Levels just below cloud base are also allowed to have icing due to uncertainty in ceiling height, especially in complex terrain. In addition, the possibility for icing is considered for all subfreezing levels between cloud base and the surface when liquid or freezing precipitation is reported, since these observations imply the potential for freezing precipitation above the ground when certain meteorological structures are present (e.g., the classical freezing rain structure). A flowchart showing the LOWICE analysis process is given in Fig. 3.

Cloud-top temperature (CTT) is a very useful parameter for assessing the likelihood of icing at a given location. Therefore, it is important to determine whether the clouds exposed to satellite view are those that are directly affecting the wind farm (Fig. 4). LOWICE performs its own assessment of the presence of multiple cloud layers, following the simple examination of the RH vertical profile described in Bernstein et al. (2005), searching downward from the highest cloud tops, looking for layers with RH $\leq$ 50%. If at least three model levels meet this criterion, then a dry layer is considered to be present. The downward progression continues and the system searches for a layer with RH $\geq$ 70%. If that criterion is met, then a new cloud layer is identified and the CTT of that layer is set to the temperature (T) where RH $\geq$ 70% was found. This process continues until the lowest cloud base height (CBZ) is reached, and then the CTT of the layer affecting the site (CTT-farm) is set to the CTT of the lowest cloud layer found above the site.

The icing likelihood [$\mathrm{ICE_{like}}$; Eq. (1)] is then initially estimated by multiplying the output from LOWICE's membership functions for T, RH, and CTT at the hub height (T_{map}, $\mathrm{RH_{map}}$, $\mathrm{CTT_{map}}$; Fig. 5; Bernstein et al. 2005). These functions are designed to account for some of the uncertainty inherent in the data being used, as well as the physical characteristics of supercooled liquid water in the atmosphere (e.g., icing is more likely to exist at relatively warm, subfreezing temperatures). $\mathrm{ICE_{like}}$ is then increased or decreased based upon reports of certain precipitation types (e.g., freezing drizzle, freezing fog, snow, snow grains) that have implications about icing. Nearby observations have a much greater influence than those that are more distant,

$$\mathrm{ICE_{like}} = T_{\mathrm{map}}\mathrm{CTT_{map}}\mathrm{RH_{map}}. \qquad (1)$$

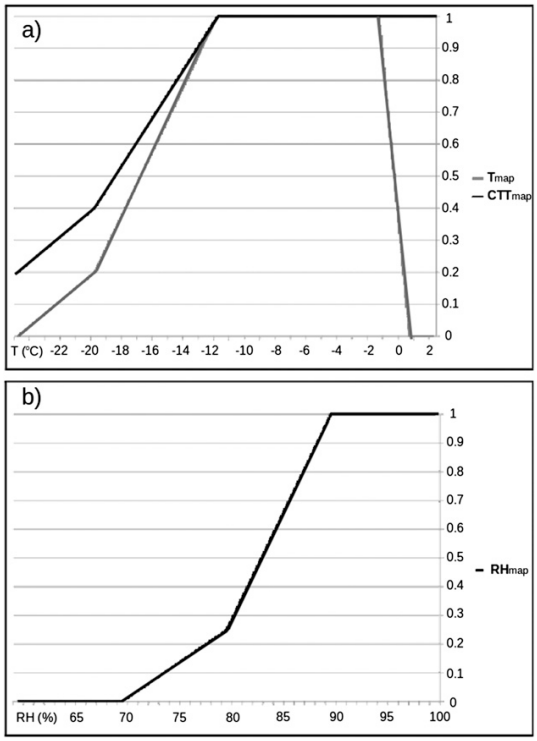


FIG. 5. Membership functions for (a) temperature (Tmap, gray line) and CTT (CTTmap, black line), and (b) RH (RHmap, %, black line). The *x* axes are temperature (°C) and RH with respect to water (%), while the *y* axes are unitless.

To estimate the icing rate, it is necessary to estimate T, liquid water content (LWC), and wind speed (v) at the site. Here T and v are taken directly from LAPS data, vertically interpolated to the site elevation. LOWICE then estimates LWC using the cloud and precipitation fields described above in combination with LAPS profiles of pressure (P), T, and RH. If the level of interest is located at or above the estimated CBZ, then a first-guess LWC is generated by assuming a moist adiabatic lapse rate from the level of interest down to CBZ. This is done by estimating the saturated mixing ratio at CBZ, via standard equations, and taking the difference between the mixing ratios at CBZ and the height of the wind farm (Z-farm), and then compensating for density. While using the adiabatic assumption is not ideal, it is not unusual for lapse rates to be nearly moist adiabatic over the shallow layer between Z-farm and CBZ. When lapse rates are more stable than moist adiabatic, initial LWC estimates are decreased to be more realistic. As noted earlier, because of the potential for local variability in CBZ between the site and the nearest station-reported ceiling height, LOWICE also allows for the possibility for icing at elevations slightly below CBZ. This allowable depth beneath the cloud base (dz) changes with the distance between the site and the reporting surface station. It is at its minimum value when the ceiling report is made close to the site and increases linearly when the ceiling report is made at the maximum allowable distance of 160 km. LWC estimates are set to nominal values between 0.0 and 0.1 g m^{-3} as dz increases from zero to the maximum allowable value.

One important aspect of the adiabatic assumption is that all condensate remains within the cloud. Many icing clouds produce precipitation, which depletes the supercooled LWC (SLWC) within them. This is particularly true when snow is expected to be present within the clouds, because the SLWC is expected to be at least partially depleted via riming (Rogers and Yau 1989). The closer the observation of precipitation is to the site, the greater the likelihood that depletion is occurring at the site. To address this process, LOWICE examines all surface observations within the radius of influence for the occurrence of precipitation. If precipitation is observed, then a depletion factor is calculated based on the distance between the site and the precipitation report. The initial estimate of LWC can be depleted by as much as 50%, depending on the distance to the reporting station. This approach is reasonably consistent with evidence from flights made in precipitating, well-mixed icing clouds sampled during natural icing flight programs (e.g., Bernstein et al. 2011b; Politovich and Bernstein 1995). Radar data can also be quite valuable for estimating the depletion of SLWC by riming, but radar data were not readily available over Sweden in LAPS through 2012 and therefore were not used in this version of LOWICE.

Once the T, LWC, and v have been estimated, the next step is to calculate the icing rate ($\mathrm{ICE_{rate}}$). This is done using a standard icing-rate equation for a cylinder (Finstad et al. 1988; Makkonen 2000; ISO 2001):

$$\frac{dm}{dt} = (\mathrm{SLWC}) A v \eta_1 \eta_2 \eta_3 , \qquad (2)$$

where SLWC is the supercooled LWC (g m^{-3}); A is the cross-sectional area (m^2) of the object; v is the wind velocity (m s^{-1}); and the unitless η_1, η_2, and η_3 are the correction factors for the collision, sticking, and accretion efficiencies, respectively, which are all set to 1.0 for simplicity. In this case, A is set equal to 0.015 m^2, based on the International Organization for Standardization (ISO) 12494 standard reference cylinder, which is 0.5 m long and has a 30-mm diameter. Because dm/dt is in grams per second and LOWICE values are produced once per hour, dm/dt is multiplied by 3600 (s h^{-1}) to

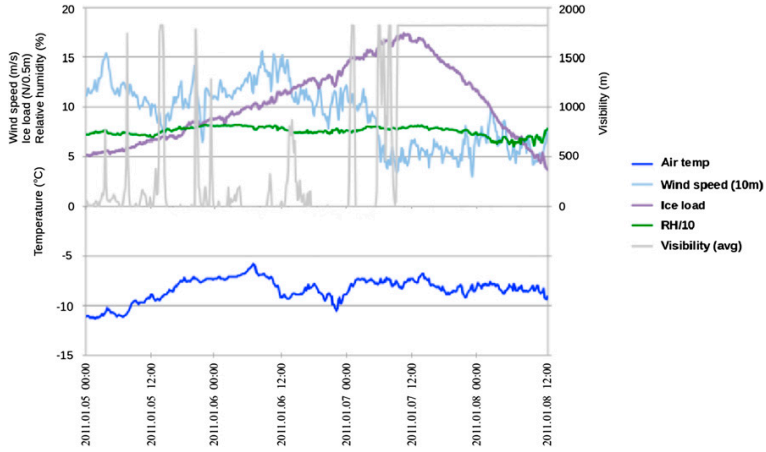


FIG. 6. Time series plot to illustrate the effects of sublimation. Observations from one wind farm in Sweden during periods of active icing (before 1100 UTC 7 Jan 2011) and sublimation (after 1100 UTC 7 Jan 2011). Observations are plotted against the same y axis, with different units; temperature (air temperature, °C), wind speed at 10 m (wind speed, m s^{-1}), ice load [ice load, N (0.5 m)$^{-1}$], RH (RH/10, %, divided by 10 for plotting purposes), and visibility (visibility, m).

generate the hourly icing rate in grams per hour. The final equation for the icing rate takes on the form

$$\mathrm{ICE_{rate}} = (\mathrm{SLWC})Av3600. \tag{3}$$

2) ICE-LOAD AND ICE-LOSS MECHANISMS

Applying the icing rate ($\mathrm{ICE_{rate}}$) for each hour, ice is accumulated and the ice mass ($\mathrm{ICE_{mass}}$) on the reference cylinder increases accordingly. $\mathrm{ICE_{mass}}$ builds when icing is active ($\mathrm{ICE_{rate}} > 0\,\mathrm{kg\,h^{-1}}$) and is depleted by melting and sublimation when icing is not active ($\mathrm{ICE_{rate}} = 0\,\mathrm{kg\,h^{-1}}$):

$$\text{If } \mathrm{ICE_{rate_1}} > 0, \quad \text{then}$$

$$\mathrm{ICE_{mass_1}} = \mathrm{ICE_{mass_0}} + [\mathrm{ICE_{rate_1}} \times \text{time (h)}] \tag{4a}$$

$$\text{If } \mathrm{ICE_{rate_1}} = 0, \quad \text{then}$$

$$\mathrm{ICE_{mass_1}} = \mathrm{ICE_{mass_0}} - [(\mathrm{MELT_{rate}} + \mathrm{SUB_{rate}})\text{time (h)}], \tag{4b}$$

where the subscripts "0" and "1" indicate the values from the previous and current times, respectively. Note that ice load (N) is calculated by multiplying ice mass (kg) by gravitational acceleration (~9.81 m s^{-2}). Melting of ice loads is obviously important, since even large ice loads can dissipate fairly quickly once the temperature becomes adequately warm. Rather than simply melting away the entire ice load as soon as the model temperature (T) exceeds 0°C, it is logical that a melting rate should be applied, based on how much temperature exceeds 0°C. Thus, LOWICE estimates a simple melting rate ($\mathrm{MELT_{rate}}$), which increases linearly from 0.0 kg h^{-1} (when $T \le 0$°C) to a maximum value of 10.0 kg h^{-1} (when $T \ge +5$°C). This estimation of melting rate is somewhat arbitrary, but sensitivity tests performed on several years of historical icing events at wind farms indicated that changes in the melting rate parameter had little effect on the overall estimates of ice load. Future versions of LOWICE may also include insolation as a factor, since sunlight on turbine blades could increase their temperature above 0°C, even when the air temperature is below 0°C.

Sublimation also plays an important role in the depletion of ice loads, especially in the Scandinavian winter, when temperatures rarely exceed 0°C. Evidence from the examination of meteorological measurements and webcam images during subfreezing, dry, windy weather indicated that the visible ice (on webcam images) and the measured loads decreased gradually over time (e.g., Fig. 6). In this figure, notice that visibility (gray line) was generally low during the period of active icing and then high during the period of sublimation. Such changes in visibility are often associated with the presence and absence of clouds that are associated with icing (as in Portin et al. 2009). LOWICE's simple sublimation scheme was developed empirically, based on a combination of wind speed and relative humidity [see

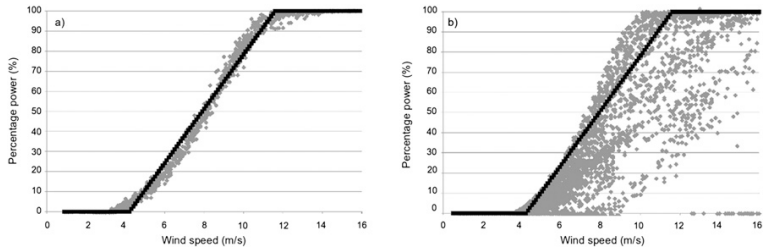


FIG. 7. Power curves as percent of full power (y axis) against wind speed (x axis). (a) Ice-free turbine power plotted against wind speed during an ice-free sample period in 2011 (gray dots), and the linear fit generated by wind turbine measurements of wind speed and power production (black line). (b) As in (a), but for a period with icing during 2011.

Eq. (4c)]. Naturally, sublimation occurs more quickly when winds speed (v) is large in subsaturated environments and, to a lesser extent, when RH is low:

$$\mathrm{SUB_{rate}} = 0.2[0.65\min\left(1.0, \frac{v}{10\,\mathrm{m\,s^{-1}}}\right) + 0.35(1.0 - \mathrm{RH_{map}})], \quad (4c)$$

where $\mathrm{SUB_{rate}}$ is in kilograms per hour, v is in meters per second, and $\mathrm{RH_{map}}$ is a fuzzy logic membership function based on RH (see Fig. 5). Equation (4c) is weighted more heavily toward v than RH. Because $\mathrm{RH_{map}} = 0.0$ for $\mathrm{RH} \leq 70\%$ and $\mathrm{RH_{map}} = 1.0$ for $\mathrm{RH} \geq 90\%$, thereby the RH part of the equation has no effect when $\mathrm{RH} \geq 90\%$ and a relatively strong effect when $\mathrm{RH} \leq 70\%$. The entire equation is multiplied by an ice-load change rate of $0.2\,\mathrm{kg\,h^{-1}}$, so the most ice that the system can sublimate in an hour is 0.2 kg (when $v \geq 10\,\mathrm{m\,s^{-1}}$ and $\mathrm{RH} \leq 70\%$). This rate may seem insubstantial, but the effects can be dramatic, especially when periods of icing are followed by extended periods of dry, windy, subfreezing weather. During such periods, ice cannot be removed by melting but sublimation can gradually deplete the ice (and its effects) over time. When icing is inactive ($\mathrm{ICE_{rate}} = 0\,\mathrm{kg\,h^{-1}}$), $\mathrm{ICE_{mass}}$ is decreased using the sublimation and melting equations described above. Without question, the sublimation and melting schemes presented here are an oversimplification of complex processes. The use of more complex melting and sublimation schemes from published literature should be tested in the future.

Another process that is important for removing ice from a wind turbine is the shedding process. Shedding tends to be stochastic in nature, with sudden events and effects that are very difficult to predict. Anecdotal evidence suggests that shedding occurs primarily during periods of melting but also during periods of high wind speeds, regardless of temperature. When blades are turning, increases in wind speed dramatically increase the centrifugal force on the blades, especially toward the blade tips, and can rapidly become strong enough to cause the ice to de-bond and shed.

3) WIND POWER ESTIMATIONS AT TURBINE—LOWICE V0, V1, AND V2

To estimate the potential impacts on power production, the power that the turbines are expected to generate when they are free of ice must first be estimated. To this end, measured wind speeds and power output were compared over ice-free portions of autumn at several wind farms. Using $0.1\,\mathrm{m\,s^{-1}}$ wind speed bins, a scatterplot was created, then a piecewise linear curve was fit to the measured power for each wind speed to approximate the relationship between wind speed and ice-free (clean) power ($\mathrm{PWR_{clean}}$; see Fig. 7a). During the winter season, the air-density effect is likely to increase the power production compared to autumn, as seen by the upward shift in maximum power production at moderate wind speeds in Fig. 7b. Correction for air-density will be considered in future LOWICE versions. Power curves like the one in Fig. 7a are applied to LAPS wind speeds to estimate $\mathrm{PWR_{clean}}$ at each point in time. Though results for only one site are presented in this paper, unique power curves were developed for several wind farms across Sweden, where testing was performed. As described earlier, it is well known that icing can have a dramatic effect on power production, but quantification of these effects has proven to be difficult (e.g., Durstewitz et al. 2008). One simple approach is to assume that as soon as any amount of ice is expected on the turbine, that turbine will not produce power. This is overly simplistic, since power can still be produced when ice accretions are present, though the production rate is often reduced. With this in mind, a simple first version of power loss code was developed for LOWICE (version V0), according to following model. If no ice load is present, then no loss of power is expected. As the ice load increases from 0.0 to $10.0\,\mathrm{kg\,m^{-1}}$, the power loss increases linearly from 0% to 100% [i.e., "iced power production" ($\mathrm{PWR_{iced}}$) decreases linearly from 100% to 0%]. Once the

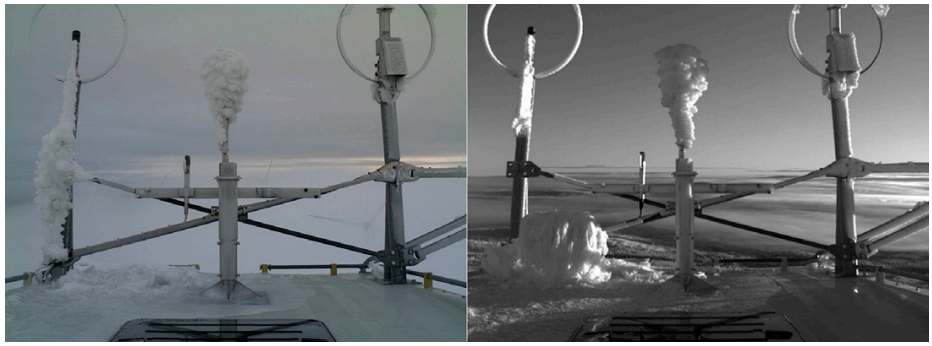


FIG. 8. Webcam images taken on (left) 7 and (right) 8 Jan 2013 from the top of the turbine nacelle, illustrating the visual assessment of icing and showing the ice load instrument with a cone of ice on it (middle of both images).

ice load reaches 10.0 kg m^{-1}, then a 100% power loss is assumed and PWR_{iced} goes to zero. This overly simple approach allows the turbine to produce power at a reduced rate until a substantial ice load is expected,

$$\text{Power loss}\,(\%) = 100\frac{PWR_{clean} - PWR_{iced}}{PWR_{clean}}. \quad (5)$$

However, intense scrutiny of real-time observations from a wide variety of icing events at wind farms in Sweden made it clear that icing-induced power losses were poorly correlated with both measured and predicted ice loads. This was true when examining all data across several icing seasons and for numerous subperiods, including periods where the measured ice loads seemed to be of particularly good quality (i.e., loads were present and ice was clearly evident in webcam images; e.g., Fig. 8). Scatterplots and statistical analysis comparing ice loads and power losses for long periods and case studies showed poor results. Only weak trends were found and *r*-squared values only reached 0.12 at best. It was also discovered that significant, if not full, power could be produced when significant measured ice loads were present. Further, it was quite obvious that tying power loss to ice load during periods of persistent load sometimes resulted in gross overestimates of power loss and underestimates of power production. In a climate like Scandinavia's, such loads can sometimes persist for weeks and even months at a time. Therefore, it was clearly evident that to better estimate icing-induced power losses, the losses have to be related to something other than ice load alone.

Further examination of observations indicated that other factors appeared to correlate more readily with power loss. In particular, power losses tended to be most evident when icing was active (i.e., ice was actively accreting at the site) and the intensity of the power losses was more strongly related to the icing rate than icing load. When icing is active, the ice is growing on the turbine blade, causing lift and drag penalties. As icing becomes inactive, the ice is often sublimated, melted, or shed from the blades fairly quickly as they rotate, resulting in a rapid return to power. The exception tends to be when turbine rotation has slowed dramatically or stopped altogether, sharply decreasing the chance for centrifugal shedding. Because of a relative lack of centrifugal force, ice on the slowly rotating vertical cylinders (used to measure loads) tends to remain in place, even when turbines shed ice and power returns. Thus, a much better correlation was found between power losses and icing rate than with icing load on a cylinder.

Time is also a critical element. First, as long as icing is active, its effect on power tends to increase or at least be maintained. Second, the longer the period of active icing is, the larger the effects tend to be. In LOWICE, this is called the ice "building effect." The building effect is essentially a function of icing rate, time, and temperature. Analysis of observations (presence/absence of active icing, ice load, temperature, relative humidity, wind speed, ceiling height, visibility, webcam images, and model output) during and after icing events clearly indicated that once icing becomes inactive, depletion effects, such as sublimation and melting, begin to take hold. The stronger those effects are and the longer they persist, the more rapidly the power tends to return. In LOWICE, this is called the ice "clearing effect." The clearing effect is a combination of the melting and sublimation effects, both of which are tied to their duration and intensity (sublimation rate and melting rate), plus the more stochastic and relatively instantaneous effect of shedding.

The equations provided below are empirical in nature and are derived from the examination of dozens of periods of active and inactive icing. They are meant to capture the essence of the effects of ice growth, decay, and their rates as associated with the presence and

absence of supercooled liquid water, temperature, relative humidity, and wind speed present when events and nonevents are occurring, though it will be demonstrated in this article that these approximations of the physical process that drive icing and its effects perform reasonably well when judged against independent observations. The equations below can most certainly be refined based on the scientific literature.

$$\text{TimeFactor} = 10\ \text{(h)}. \tag{6a}$$

$$\text{If}\quad T > 0°\text{C},\quad \text{MeltFactor} = \max\left(1, \frac{T}{5°\text{C}}\right),$$
$$\text{otherwise}\quad \text{MeltFactor} = 0. \tag{6b}$$

$$\text{SublimationEffect} = 2\frac{\left\{0.65\min\left(1, \frac{v}{10\,\text{m s}^{-1}}\right) + [0.35(1 - \text{RH}_{\text{map}})]\right\}}{\text{TimeFactor}}. \tag{6c}$$

$$\text{SheddingEffect} = \max\left[1, \left(\frac{v}{25\,\text{m s}^{-1}}\right)^2\right]. \tag{6d}$$

$$\text{If}\quad T > 0°\text{C},\quad \text{TempEffect} = 1. \tag{6e}$$

$$\text{If}\quad -5°\text{C} < T \le 0°\text{C},\quad \text{TempEffect} = 0.25 + 0.75\left(\frac{T + 5°\text{C}}{5°\text{C}}\right)^{1.5}. \tag{6f}$$

$$\text{If}\quad -25°\text{C} < T \le -5°\text{C},\quad \text{TempEffect} = 0.25\left(\frac{T + 25°\text{C}}{20°\text{C}}\right). \tag{6g}$$

$$\text{If}\quad T \le -25°\text{C},\quad \text{TempEffect} = 0. \tag{6h}$$

$$\text{V1BuildingEffectRate} = \min\left[1, \frac{\left(\frac{\text{ICE}_{\text{rate}_1}}{500}\right)}{\text{TimeFactor}}\right]. \tag{7a}$$

$$\text{If}\quad \text{MeltFactor} > 0,\quad \text{V1ClearingEffectRate} = \frac{5(\text{MeltFactor})}{\text{TimeFactor}}. \tag{7b}$$

$$\text{If}\quad \text{MeltFactor} = 0,\quad \text{V1ClearingEffectRate} = \frac{1}{\text{TimeFactor}}. \tag{7c}$$

$$\text{If}\quad \text{ICE}_{\text{rate}_1} > 0,\quad \text{V1Loss}_1 = \text{V1Loss}_0 + \text{V1BuildingEffectRate}. \tag{7d}$$

$$\text{If}\quad \text{ICE}_{\text{rate}_1} = 0,\quad \text{V1Loss}_1 = \text{V1Loss}_0 - \text{V1ClearingEffectRate}. \tag{7e}$$

$$\text{V2BuildingEffectRate} = \min\left[1, \frac{\left(\frac{\text{ICE}_{\text{rate}_1}}{250}\right)\text{TempEffect}}{\text{TimeFactor}}\right]. \tag{8a}$$

$$\text{V2ClearingEffectRate} = \min(1, \text{TempEffect} + \text{ShedEffect} + \text{SublimationEffect/TimeFactor}). \tag{8b}$$

$$\text{If}\quad \text{ICE}_{\text{rate}_1} > 0,\quad \text{V2Loss}_1 = \text{V2Loss}_0 + \text{V2BuildingEffectRate}. \tag{8c}$$

$$\text{If}\quad \text{ICE}_{\text{rate}_1} = 0,\quad \text{V2Loss}_1 = \text{V2Loss}_0 - \text{V2ClearingEffectRate}. \tag{8d}$$

With these concepts in mind, LOWICE's alternative power loss schemes (V1 and V2) combine the building and clearing effects described above [Eqs. (7), (8)]. At each time step, these factors are combined to determine the fractional expected loss factor ($\text{LOSS}_{\text{factor}}$; V1Loss or V2Loss, respectively), on a 0.0 (no power loss expected) to 1.0 (total power loss expected) scale. Iced power (PWR_{iced}) is then calculated accordingly:

$$\text{PWR}_{\text{iced}} = \text{PWR}_{\text{clean}}(1 - \text{LOSS}_{\text{factor}}). \tag{9}$$

Using this approach, icing-related losses can rapidly deplete power during periods of active icing, depending on the icing rate (i.e., intensity) and longevity of the event. Also, once active icing ceases, power can rapidly recover, depending on the rate of melting, sublimation, shedding, and the duration of icing inactivity.

3. Results and verification

As part of a wind pilot program, observations of power production and other fields (e.g., wind speed, temperature) have been examined for several wind farms across Sweden over five icing seasons: 2009–14. As a demonstration, the results are shown for one wind farm and for a single month (January 2013). This period has been chosen because it demonstrates numerous interesting features when observations were available continuously, including webcam images, turbine-mounted

TABLE 1. Statistics of the average values from the wind farm (a total of 18 turbines) during January 2013, where T is temperature (°C), U is wind speed ($m\,s^{-1}$), and Pwr loss is power production losses (%) due to icing on turbine. Estimated power (Pwr) loss is given for LOWICE versions V0, V1, and V2 in the table.

Turbine	T avg Obs	T avg LOWICE	T min Obs	T min LOWICE	T max Obs	T max LOWICE
Avg	−4.8	−7.2	−12.4	−15.3	0.6	−1.4
U avg Obs		U avg LOWICE		U max Obs		U max LOWICE
6.3		7.4		15.4		15.1
Pwr loss Obs		Pwr loss V0		Pwr loss V1		Pwr loss V2
19.5%		74.5%		18.3%		15.9%

instruments, and power measurements. Because of the confidentiality agreements with wind power companies, the actual production, wind farm names, and certain other data could not be included in this paper. The wind farm turbines were located at various elevations around 800 m, while the corresponding LOWICE model height (i.e., representative hub height) was about 10 m higher for this specific wind farm. There were 712 h of matched data during January 2013. Temperature and wind speed verification have been made against measurements from instruments mounted at the top of one of the wind turbines. Manual, hourly assessments of the presence of ice and the occurrence of active icing were determined by independent visual inspection of webcam images in combination with inspection of measurements of visibility, ceiling height, and the other fields described above. Though the presence of ice and active icing can be particularly difficult in some environments (e.g., when a thin glaze of icing may be present or lighting is poor), the manual assessment dataset is generally quite reliable. Results of the ice detection and estimated power are described in section 3b. In section 3c, a statistical evaluation of icing detection and power losses is presented.

a. Temperature and wind

Analyzed and observed temperature fields were in fairly good agreement during the test period, though there was a persistent cold bias, averaging 2.4°C (Table 1; Fig. 9a). This bias appears to have been driven by a cold bias in ECMWF surface temperature forecasts (i.e., first-guess field for LAPS), which were found to be about 2°C in the −13° to +5°C range for mountain stations (at height > 300 m above mean sea level; Fig. 10a). Even though the LAPS analysis generally performs very well, there are still outliers with stations being less corrected, giving a slight cold bias in the −13° to +5°C temperature range (Fig. 10b). Modeled and observed winds were also quite comparable, with only a slight positive wind bias of $1.1\,m\,s^{-1}$ (Table 1; Fig. 9b). At times, modeled winds were substantially higher than observed winds (e.g., 1–2, 17–19, and 24–26 January), but observed winds were sometimes falsely low due to icing effects on the anemometer (when the light green line is well below the gray lines in Fig. 9b). These false low "truth" values for wind contributed to the apparent $1.1\,m\,s^{-1}$ high wind speed bias.

b. Icing and power loss

The estimated observed power loss for January 2013 was 19.5% for all turbines at the site combined. Actual individual turbine power losses ranged from 3% to 34%. LOWICE expected power losses were quite reasonable for V1 (18.3%) and V2 (15.9%), but losses were grossly overestimated by the load-driven V0 (74.5%). Time series data shown in Fig. 11 reveal that V1 and V2 power losses tended to be quite similar and followed the trends of turbine observations quite well, while V0 frequently overestimated losses (e.g., 100% loss between 1 and 16 January), including when its iced power trends tracked somewhat similarly to V1 and V2 (25–31 January).

LOWICE estimates of active icing ($ICE_{rate} \geq 10\,g\,h^{-1}$) and inactive icing (ice present but $ICE_{rate} < 10\,g\,h^{-1}$) are indicated on the color bars near the bottom of Fig. 11. The periods of active and inactive icing were generally captured very well, though active icing was indicated more often by LOWICE than by manual assessments. Most of these overwarnings were for short periods; however, there were a few prolonged periods when LOWICE indicated icing but none was observed (e.g., 1–2, 18–19, and 24–26 January in Fig. 11; compare red portions of the two color bars). The second of these expected events resulted in significant expected power losses. The power time series chart implies that LOWICE's iced power forecasts were still reasonable, but this was caused by a combination of overestimated wind speeds (see Fig. 9b) and overestimated icing effects compensating for one another. This is an example of how iced wind power models can get the right answer for the wrong reason and why statistical comparisons alone are not enough to determine the quality of system output. Despite this example, overall power losses for icing-rate-driven V1 and V2 were quite comparable to observed

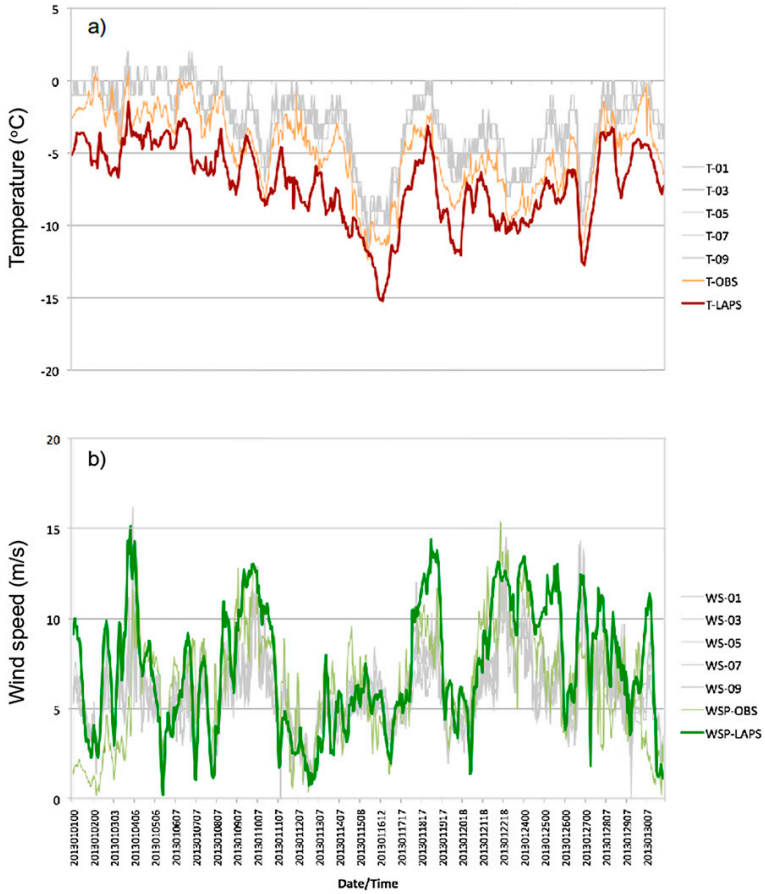


FIG. 9. Time series for January 2013 of (a) temperature and (b) winds speed from LOWICE [T-LAPS (red line) and WSP-LAPS (green line)]; observations of temperature and wind speed from standard instrumentation [T-OBS (tan line) and WSP-OBS (light green line)]; and individual turbine measurements, temperature and wind speed (T-01–T-09 and WS-01–WS-09) are shown with gray lines. All data are from the approximately same height (i.e., height of wind farm).

losses and far superior to those from the ice-load-driven V0 power loss code.

c. *Statistical results*

Though statistical results have their shortcomings, it is important to attempt to measure the quality of the icing and power fields produced by LOWICE and, indeed, any such system. To this end, standard statistical measures were used to determine the probability of detecting both positive (''Yes'') and negative (''No'') icing events and power loss events (PODy, PODn, respectively, following Mahoney et al. 2002; see Table 2). Both measures and their derivatives [e.g., critical success index (CSI) and true skill score (TSS)] are important because an icing and power assessment system must capture both the presence and absence of icing and its downstream effect on power to be useful to those making decisions regarding the operation of turbines in cold climates and the trading of power expected to be available to the grid.

1) PRESENCE OF ICE AND ACTIVE ICING

As described above, the presence and absence of ice, as well as whether the ice was actively growing, were assessed by subjective, manual inspection of a combination of webcam images and time history of automated measurements of temperature, dewpoint, visibility, ceiling height, and ice load. In addition, objective assessments were made using measured icing-related parameters and

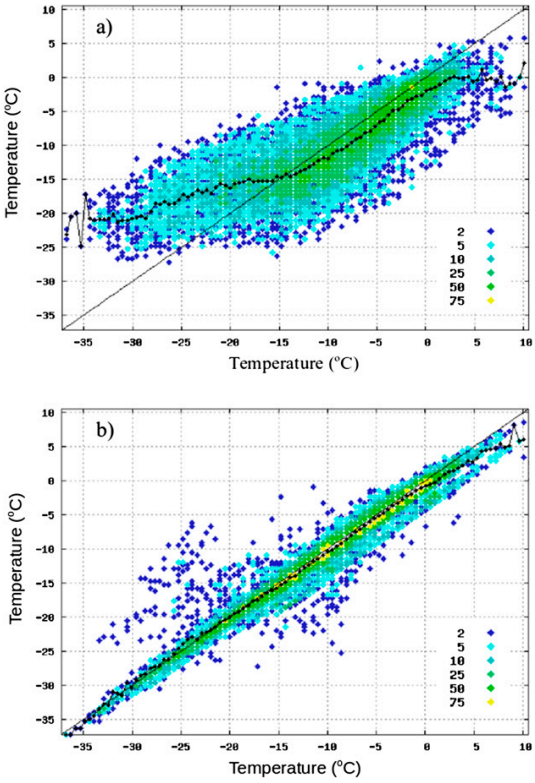


FIG. 10. Scatterplot of mountain stations' (i.e., >300 m MSL) 2-m temperature for January 2013. (a) ECMWF on *y* axis and observations on *x* axis and (b) LAPS analysis on *y* axis and observed values on *x* axis. Colored dots represent the number of observations. Black dotted line is the mean value of dataset. Included is the 1:1 best-fit curve (thin black curve).

derivatives thereof. Both approaches were used to produce Yes/No "ground truth" data on the presence of ice and active icing at the wind farm. These include increases in observed ice load over time (1-, 3-h periods) and the simultaneous presence of $T < 0$°C and either visibility < 1000 m or ceiling height < 250 m (see Table 3). An absolute Yes/No threshold that indicates that icing has occurred does not exist, so a series of thresholds is applied to reach the final answer. The LOWICE system uses similar parameters to determine the presence of ice (see Table 4; Fig. 12). Normally, a threshold of 0.1 N would be used for the observed presence of ice on the reference cylinder weighing gauge, but the presence of a bias in the observed loads at this site necessitated the use of a higher threshold (4.5 N).

For active icing, PODy values were all between 0.55 and 0.81, PODn values were all near 0.55, CSIs were 0.18–0.35, and TSSs were 0.11–0.36 (Figs. 12a–c). Overall, these results show some (but not particularly impressive) skill for this particular test site and period. However, it is important to note that the LOWICE system's results were best for comparisons with the most reliable observed field: manually assessed active icing (PODy: 0.81, PODn: 0.56, CSI: 0.25, TSS: 0.36). However, FAR was high at 0.73. For the presence of ice, LOWICE indicated positive loads throughout the month of January 2013, so PODy was 1.0, while PODn was 0.0, resulting in the point in the bottom-right portion of Fig. 12a. It is interesting to note that although the manual assessment indicated some hours with no ice presence, the icing load was measured to be >4.5 N throughout the month, so PODn and TSS were undefined for this field (which gives one less verification point in Figs. 12a,c). The CSI for manually assessed ice presence was 0.89.

2) STATISTICAL RESULTS—POWER LOSS

LOWICE's expected power loss estimates were verified against estimates of observed power loss. Again, there is no absolute Yes/No threshold that indicates that power loss has occurred or that it can be attributed to icing. However, a reasonable assessment of power loss can be gained by applying some simple thresholds. First, turbine-measured wind speeds must considered to be strong enough to be able to produce clean power that is at least 25% of the rated output from the turbine. This is estimated by sending the turbine-measured wind speed through LOWICE's piecewise linear approximation of the wind farm's power curve. The 25% criterion is used to be certain that an adequate amount of power is expected to allow for a meaningful estimate of power loss, as well as to avoid the inaccuracies of the piecewise linear power curve near the cut-in wind speed. Next, if at least 75% of the turbines met the clean power criterion mentioned above, then the measured power from each turbine was divided by the expected clean power to determine the percentage of expected clean power that is being produced. This value was subtracted from 1.0 to determine the power loss and then expressed as a percentage (see Table 5). Power loss thresholds (10%, 30%, 50%, 70%, and 90%) were then applied to the estimated losses described above. Each of these thresholds is used to determine when observed power loss was considered to be present (a Yes observation).

Among the turbines with adequate wind speeds to meet the 25% rated power output threshold described above, the number of turbines that met the power loss percentage threshold was counted. At least half of the turbines with adequate expected clean power had to meet the power loss threshold for a Yes observation of power loss to be indicated for that time. As stated above, for a Yes observation of power loss, multiple power loss thresholds (10%, 30%, . . . , 90%) were used and at least

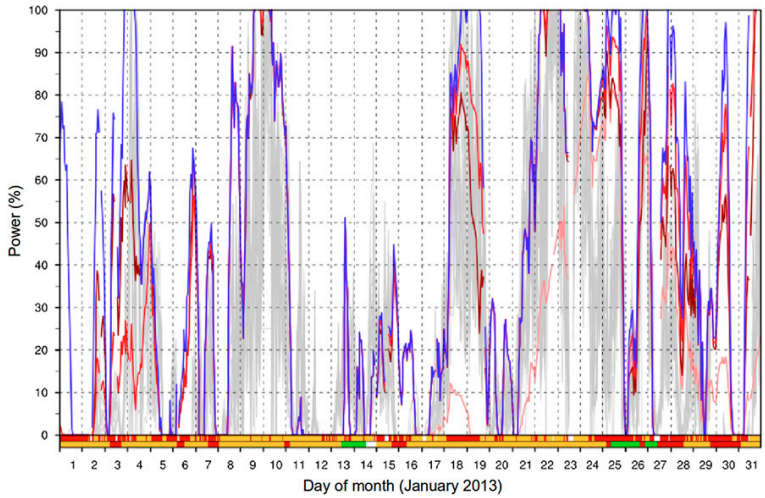


FIG. 11. Time series for January 2013 of the percentage of maximum power from wind turbines (gray), LOWICE clean power (blue), and iced power (red, brown, and pink lines for loss code versions V2, V1, and V0, respectively). These traces are not visible when power values equal zero. Color bars at the bottom provide an hour-by-hour assessment of the presence of icing as follows. Bar 1: LOWICE-expected icing [red represents active icing (rate $\geq 10\,g\,h^{-1}$), yellow denotes inactive icing (ice present but rate $<10\,g\,h^{-1}$), green denotes no ice, and white denotes no information]. Bar 2: manually assessed icing through webcam (red represents active icing, yellow denotes active icing possible but not confirmed, green denotes no active icing, and white denotes no information).

75% of the turbines had to meet the 25% expected power threshold.

For the LOWICE system output to produce a Yes, the system's expected iced power was compared to its expected clean power. Expected power loss was calculated by dividing the expected iced power by the expected clean power, and then the result was subtracted from 1.0 and expressed as a percentage (Table 5). Like observed power, power loss thresholds of 10%, 30%, 50%, 70%, and 90% were applied to convert expected losses from the system to a Yes/No answer. Yes and No forecasts and observations were compared for every combination of forecast and observed power loss threshold. For example, LOWICE expected losses of ≥10% were

TABLE 2. Statistical parameters and explanations thereof (Mahoney et al. 2002). PODy and PODn measure the percentage of times that the system correctly forecast the occurrence of positive (Yes) and negative (no) events. It is best to maximize the PODy and PODn values. FAR measures the frequency of Yes forecasts that did not verify. FAR should be minimized, because false alarms indicate overforecasting of positive (Yes) events. CSI measures the number of correct Yes forecasts relative to the total number of Yes forecasts and observations. CSI value should be maximized. TSS summarizes the ability of the forecasts, objectively measuring their skill, because it combines the PODy and PODn measures. Models are not rewarded that have high PODn by virtue of underforecasting events (and thus have a low PODy) or high PODy by virtue of overforecasting (and thus have a low PODn). Positive TSS value indicates skill, zero indicates no skill, and negative value indicates negative skill.

Field	Name	Equation
YY	Correct positive output	LOWICE = YES, observation = YES
YN	False positive output	LOWICE = YES, observation = NO
NY	False negative output	LOWICE = NO, observation = YES
NN	Correct positive output	LOWICE = NO, observation = NO
PODy	Probability of detection of positive events	YY/(YY + NY); perfect score = 1, no skill = 0, range = 0…1
PODn	Probability of detection of negative events	NN/(NN + YN); perfect score = 1, no skill = 0, range = 0…1
FAR	False alarm rate	YN/(YY + YN); perfect score = 0, no skill = 1, range = 0…1
CSI	Critical success index	YY/(YY + YN + NY); perfect score = 1, no skill = 0, range = 0…1
TSS	True skill score	(YY/(YY + NY)) + (NN/(YN + NN))−1 or PODy + PODn-1; perfect score = 1, no skill = 0, range = −1…1

TABLE 3. Observed parameters and thresholds used for objective assessment of the presence of ice and active icing. These include measured icing load (N), hourly and 3-hourly changes in icing load, the simultaneous presence of temperatures that were subfreezing or nearly so (T-factor) with either visibility less than 1000 m or ceiling height less than 250 m (vis-factor).

Observations; parameter	Values that gives a Yes answer
Icing load; ice presence	>0.45 N (normally >0.1 N would be used)
Icing load; dLoad/*dt* (hourly)	$>0\ \mathrm{N\,h^{-1}}$
Icing load; dLoad/*dt* (3-hourly)	$>0\ \mathrm{N\,h^{-1}}$
Temperature and visibility	T-factor × vis-factor ≥ 0.25 where T-factor = 1 (if $-40°\mathrm{C} < T < -1°\mathrm{C}$) T-factor = 0.5-T/1.5 (if $-1°\mathrm{C} < T < +0.5°\mathrm{C}$) T-factor = 0.0 (if $T < -40°\mathrm{C}$ or $T > +0.5°\mathrm{C}$) vis-factor = 0.0 (if vis > 1000 m) vis-factor = $([1000\ \mathrm{m} - \mathrm{vis}]/1000)^2$ (if 0 m < vis < 1000 m)
If visibility is missing, then ceiling height is used as substitute for vis-factor	vis-factor = 0.0 (if ceil > 250 m) vis-factor = $((250\ \mathrm{m} - \mathrm{ceil})/250)^2$ (if 0 m < ceil < 250 m)

compared to observed losses of ≥0%, 30%, 50%, 70%, and 90%. The same was true for LOWICE expected losses of ≥30%, 50%, 70%, and 90%. This resulted in a matrix of Yes and No forecasts and observations of power losses that were used to generate the plots in Figs. 12d–f.

Power loss results were generally quite good, especially for V1 and V2. Results are threshold dependent, but for V1 and V2 the PODn values were >0.65 in all cases, PODy values for most thresholds exceeded 0.3, and some exceeded 0.8. CSI and TSS values suggest that the system had good skill for many power loss thresholds. The load-based V0 system performed more poorly, with generally higher PODy (due to the persistent presence of positive loads) at the expense of lower PODn, higher FAR, and generally lower CSI and TSS (Figs. 12d–f).

4. Additional applications

Beyond the real-time power and icing assessments described in this paper, the LAPS–LOWICE system may also prove useful as part of the assessment of planned wind power sites and as an ice detection method.

Both wind and icing evaluations should be done as part of the evaluation of planned wind parks in order to assess their potential for power production and both the need for and potential benefits of an anti-icing or deicing system. Different ice detection measurement approaches give different information on frequency and duration of icing events, with none proving to be superior to the other (Tammelin et al. 2005). Hence, applying multiple methods would theoretically improve the reliability of icing information. Long-term estimations of the frequency of icing from a system such as LAPS–LOWICE could serve as one method for assessing the expected wind power and icing at a new site. In fact, LAPS–LOWICE has been used to generate 5-yr climatologies for multiple sites across Sweden using a combination of archived GFS model output and observations from satellite and surface stations. Because of their proprietary nature, the results of these studies cannot be included here.

In cold climate regions (i.e., where icing occurs) many turbines have heating systems, which need a control strategy. Proper and fast identification of icing events in real time is crucial to operating these systems properly and maximizing power production, since power production losses of 5%–15% can occur early in icing events before deicing systems are activated (Peltola et al. 1996). Usually, a basic control strategy includes an ice detection method that is used to activate the system. Detection strategies may rely upon 1) the difference between the wind speeds measured by heated and unheated anemometers that exceed a certain limit and 2) measured power that is lower than expected ice-free power for a given wind speed. Different ice detection methods give different results, and no method is accurate and reliable for all situations (Marjaniemi et al. 2000). Therefore, usage of several different ice detection methods may be beneficial for the control of wind turbines. LOWICE may prove to be useful for this purpose.

5. Discussions, conclusions, and future work

In this paper new methods on how to estimate wind power production and the effects of icing thereon have been described. Using LAPS 3D analyses of

TABLE 4. LOWICE output used for determining the presence of ice and active icing.

LOWICE field	Values that gives a Yes answer
Expected ice load; ice presence	>0.1 N
Expected icing rate	$>10\ \mathrm{g\,h^{-1}}$

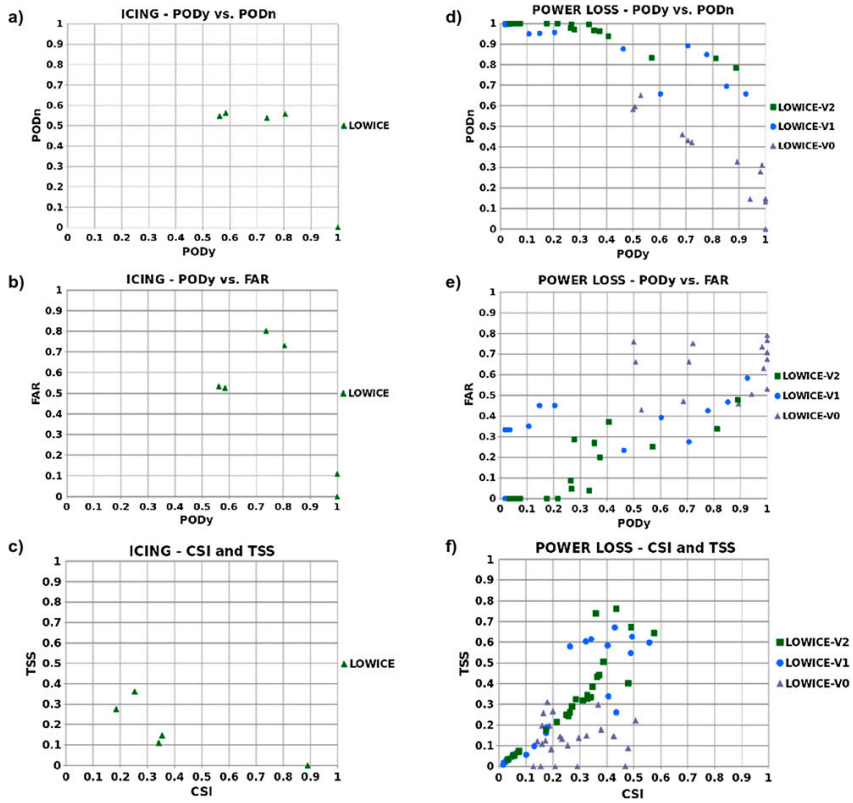


FIG. 12. Statistics for (a)–(c) icing and (d)–(f) power loss (all turbines) at the site for January 2013. For icing, PODy, PODn, FAR, CSI, and TSS were calculated by comparing LOWICE indications of the presence or absence of any ice (Yes when expected load was nonzero) and active icing (Yes when ICErate $> 10\,\mathrm{g\,h^{-1}}$) to observations, including (a) manual assessments (visual inspection) of the presence of ice, ice growth (frame-by-frame examination of webcam images), and measurements of (b) dLoad/*dt* (hourly; Yes when $> 0\,\mathrm{N\,h^{-1}}$), dLoad/*dt* (3 hourly; Yes when $>0\,\mathrm{N\,3\,h^{-1}}$), and "T-vis" (normally observed T and visibility, but T and ceiling height were used at this site because visibility was not measured there; Yes when $-40°\mathrm{C} < T < +0.125°\mathrm{C}$ and ceiling height $< 125\,\mathrm{m}$). For power loss, the same statistical fields were calculated by comparing the LAPS–LOWICE "expected" losses and the "observed" losses based on turbine measurements of wind speed and power production, and then applying 10%, 30%, 50%, 70%, and 90% power loss thresholds. A marker is plotted for each combination (e.g., 10% expected and 10% observed loss, 10% expected and 30% observed, etc.). Observed losses were estimated by comparing the observed power to the "expected clean power," which was calculated by passing the observed wind speed through the power curve.

temperature, winds, and clouds, the LOWICE system has proven to be able to estimate expected power from wind turbines, power losses that are associated with icing, and the recovery of power associated with the depletion of icing effects. Examination of data from the active icing month of January 2013 showed that LAPS–LOWICE provided realistic results for the fields described above. However, this paper does not focus only on good results to demonstrate the robustness of the system. Instead, it describes both good and poor results and some of the root causes of them, in an effort to move the science of wind power prediction forward. Examples include the anomalously cold LAPS temperature bias and both the over- and underdiagnosis of icing and their effects on expected production.

ECMWF surface temperature forecasts were a potential cause of the overall negative temperature bias in LAPS output (see section 3a; Fig. 10). This may have been particularly important for power loss errors early in the month, when observed temperatures hovered around 0°C, yet icing loads persisted in the system. Another potential source of the cold bias may be due to

TABLE 5. Power loss assessments for LOWICE system output to produce a Yes answer.

System parameters	Field name	Values	Yes thresholds
Observed power loss (calculated at each turbine if expected clean power ≥ 25% of turbine rated power)	Obs-loss	1—(observed power/expected clean power)	10%, 30%, 50%, 70%, 90% loss Must be met for at least half of all turbines that had expected clean power ≥ 25% of turbine rated power
System expected power loss (calculated at each turbine if expected clean power ≥ 25% of turbine rated power	Sys-loss	1—(iced power/clean power)	10%, 30%, 50%, 70%, 90% loss

the assimilation of surface temperatures from nearby mountain valley stations, with locally very cold temperatures (i.e., valley inversion effects), which might not be representative of the local climate at an elevated wind farm. Also, it is possible that wind turbine temperature sensors may have been affected by heating from deicing equipment, which might have affected the validation results. At several sites, turbine-measured temperatures proved to be consistently warmer than independently measured temperatures at essentially the same height (Fig. 9a). Wind speed verification results were quite good, and both the timeliness and amplitude of important wind features were captured quite well by LAPS in time series analysis. The largest exceptions appeared to be associated with falsely low wind observations from iced anemometers. Uncertainties in measured temperatures and winds will be further investigated and reported in upcoming articles.

Verification of power estimates and icing effects on power indicated that V1 and V2 of the power loss code performed particularly well, providing consistent, reasonable estimates of power production. In contrast, V0 performed relatively poorly because its power losses are tied solely to the expected ice load, which was overestimated. The ice load has also proven to be poorly correlated with power loss.

Future developments of the LAPS–LOWICE system will include 1) the use of radar data for better cloud and precipitation analyses and SLWC adjustment and 2) an estimation of insolation to better assess the potential for ice shedding. Recent measurements at wind farms could be used to correct the LAPS–LOWICE temperatures, wind speeds, and clean power estimates. Also, observations of visibility, ceiling height, and icing (e.g., from Holooptics, ice-load trends, or ice detectors such as the ''Rosemount'' probe; e.g., Mughal and Virk 2013) could be used to improve system estimates of the presence and intensity of icing conditions at the site.

While real-time diagnoses of the wind speed, icing, and power from LOWICE clearly have utility, there is also great value in accurate and timely predictions of these parameters. Thus, a forecast version of LOWICE, known as ''FLOWICE,'' has been developed using numerical model output to generate forecasts of wind speed, icing, and wind power production out to 48 h. In an effort to improve initialization of icing and power loss fields at the start of each run, FLOWICE runs are initiated using the most recent LOWICE analysis fields, rather than ''cold start'' models that are initialized with no icing or model runs that are initiated from the final state of the previous model run, which can lead to compounding effects of previous poor forecasts. Corrections to system initialization and biases could also be gained from comparison to recent and historical observations. This concept was tested during the 2014–15 icing season and these adjustments appear to be helpful.

The analysis and forecasting of icing is very difficult, especially close to ground. There is a great value in real-time observations if they are applied correctly and blended effectively with model fields. In particular, all data (observations and model fields) related to clouds and precipitation (satellite, METAR, SYNOP, T and RH profiles, etc.) have proven to be of great interest and can have great value for wind turbine icing and power production diagnoses and forecasts. The developing LAPS–LOWICE and FLOWICE systems have great potential to provide highly valuable information to the wind power community.

Acknowledgments. We want to thank OX2, the Swedish Energy Agency (funding), Göran Ronsten, the Wind Pilot Program Reference Group, and the anonymous reviewers of this paper for their valuable input and/or support of this work.

REFERENCES

Albers, S. C., J. A. McGinley, D. L. Birkenheuer, and J. R. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Wea. Forecasting,* **11,** 273–287, doi:10.1175/1520-0434(1996)011<0273:TLAAPS>2.0.CO;2.

Bernstein, B. C., F. McDonough, M. K. Politovich, B. G. Brown, T. P. Ratvasky, D. R. Miller, C. A. Wolff, and G. Cunning, 2005: Current icing potential (CIP): Algorithm description and comparison with aircraft observations. *J. Appl. Meteor.,* **44,** 969–986, doi:10.1175/JAM2246.1.

——, ——, C. A. Wolff, M. K. Politovich, G. Cunning, S. Mueller, and S. Zednik, 2006: The new CIP icing severity product. *12th Conf. on Aviation, Range and Aerospace Meteorology,* Atlanta, GA, Amer. Meteor. Soc., P9.5 [Available online at http://ams.confex.com/ams/pdfpapers/102273.pdf.]

——, E. Gregow, I. Wittmeyer, and J. Hirvonen, 2011a: LOWICE: An example of the adaption of in-flight icing diagnosis concepts to the structural icing environment. *Proc. SAE 2011 Int. Conf. on Aircraft and Engine Icing and Ground Deicing,* Chicago, IL, SAI International and AIAA, 2011-38-0070. [Available online at http://papers.sae.org/2011-38-0070/.]

——, C. A. Wolff, and F. M. McDonough, 2011b: A regional comparison of icing conditions in boundary layer clouds. *Proc. 2011 SAE Int. Conf. on Aircraft and Engine Icing and Ground Deicing,* Chicago, IL, SAI International and AIAA, 2011-38-0021. [Available online at http://papers.sae.org/2011-38-0021/.]

Cattin, R., S. Kunz, A. Heimo, G. Russi, M. Russi, and M. Tiefgraber, 2007: Wind turbine ice throw studies in the Swiss Alps. *Proc. European Wind Energy Conf. and Exhibition 2007,* Milan, Italy, EWEA, BL3.269. [Available online at http://www.meteotest.ch/fileadmin/user_upload/Windenergie/pdfs/paper_ewec2007_cattin_final.pdf.]

Derrien, M., and H. Le Gléau, 2005: MSG/SEVIRI cloud mask and type from SAFNWC. *Int. J. Remote Sens.,* **26,** 4707–4732, doi:10.1080/01431160500166128.

Durstewitz, M., J. Dobschinski, and Z. Khadiri-Yazami, 2008: Wind power forecast accuracy under icing conditions—General approach, practical applications and options for considering effects of wind turbine icing. *Proc. Winterwind 2008,* Norrköping, Sweden, Swedish Wind Power Association [Available online at http://www.winterwind.se/2008/presentationer/17_Durstewitz_Winterwind_2008.pdf.]

Dybbroe, A., K.-G. Karlsson, and A. Thoss, 2005: NWCSAF AVHRR cloud detection and analysis using dynamic thresholds and radiative transfer modeling. Part II: Tuning and validation. *J. Appl. Meteor.,* **44,** 55–71, doi:10.1175/JAM-2189.1.

ECMWF, 2014: IFS documentation CY40r1. Accessed 8 May 2014. [Available online at http://old.ecmwf.int/research/ifsdocs/CY40r1/>.]

EWEA, 2013: Wind power growth expected to slow in 2013, but recovery predicted. Accessed 6 April 2014. [Available online at http://www.ewea.org/blog/2013/04/wind-power-growth-expected-to-slow-in-2013-but-recovery-predicted/.]

Finstad, K. F., E. P. Lozowski, and E. M. Gates, 1988: A computational investigation of water droplet trajectories. *J. Atmos. Oceanic Technol.,* **5,** 160–170, doi:10.1175/1520-0426(1988)005<0160:ACIOWD>2.0.CO;2.

GWEC, 2012: Global wind report 2012: Annual market update. 72 pp. [Available online at http://www.gwec.net/wp-content/uploads/2012/06/Annual_report_2012_LowRes.pdf.]

Haggerty, J. A., F. McDonough, J. Black, S. Landolt, C. Wolff, S. Mueller, P. Minnis, and W. L. Smith, 2008: Integration of satellite-derived cloud phase, cloud top height, and liquid water path into an operational aircraft icing nowcasting system. *13th Conf. on Aviation, Range and Aerospace Meteorology,* New Orleans, LA, Amer. Meteor. Soc., 3.13. [Available online at https://ams.confex.com/ams/pdfpapers/131893.pdf.]

Huffman, G. J., and G. A. Norman Jr., 1988: The supercooled warm rain process and the specification of freezing precipitation. *Mon. Wea. Rev.,* **116,** 2172–2182, doi:10.1175/1520-0493(1988)116<2172:TSWRPA>2.0.CO;2.

ISO, 2001: Atmospheric icing of structures. International Organization for Standarization Copyright Office ISO 12494: 2001, 56 pp. [Available online at http://www.iso.org/iso/catalogue_detail.htm?csnumber=32823.]

Koskinen, J. T., and Coauthors, 2011: The Helsinki Testbed: A mesoscale measurement, research, and service platform. *Bull. Amer. Meteor. Soc.,* **92,** 325–342, doi:10.1175/2010BAMS2878.1.

Le Bot, C., 2004: SIGMA: System of Icing Geographic Identification in Meteorology for Aviation. *11th Conf. on Aviation, Range and Aerospace Meteorology,* Hyannis, MA, Amer. Meteor. Soc., P6.5. [Available online at https://ams.confex.com/ams/11aram22sls/techprogram/paper_81704.htm.]

Mahoney, J., J. K. Henderson, B. G. Brown, J. E. Hart, A. Loughe, C. Fischer, and B. Sigren, 2002: The Real-Time Verification System (RTVS) and its application to aviation weather forecasting. Preprints, *10th Conf. on Aviation, Range and Aerospace Meteorology,* Portland OR, Amer. Meteor. Soc., 9.8. [Available online at https://ams.confex.com/ams/13ac10av/techprogram/paper_40728.htm.]

Makkonen, L., 2000: Models for the growth of rime, glaze, icicles and wet snow on structures. *Philos. Trans. Roy. Soc. London,* **358A,** 2913–2939, doi:10.1098/rsta.2000.0690.

Marjaniemi, M., H. Holttinen, J. Keinanen, E. Holttinen, L. Makkonen, E. Peltola, T. Maki, and K. O. Petersen, 2000: Wind turbines in light icing conditions—Experiences of the Pori 8 MW wind farm. *Proceedings of the BOREAS V Conference,* B. Tammelin et al., Eds., Finnish Meteorological Institute, 13 pp.

Mughal, U. N., and M. S. Virk, 2013: Atmospheric icing sensors—An insight. *SENSORCOMM 2013: The Seventh International Conference on Sensor Technologies and Applications,* S. Yurish and M. S. Virk, Eds., IARIA, 191–199. [Available online at http://www.thinkmind.org/index.php?view=article&articleid=sensorcomm_2013_8_20_10053.]

Peltola, E., M. Marjaniemi, J. Kaas, and E. Aarnio, 1996: Pyhätunturi operational experiences. *BOREAS III: Proceedings of an International Meeting,* B. Tammelin et al., Eds., Finnish Meteorological Institute, 131–146.

Politovich, M. K., and B. C. Bernstein, 1995: Production and depletion of supercooled liquid water in a Colorado winter storm. *J. Appl. Meteor.,* **34,** 2631–2648, doi:10.1175/1520-0450(1995)034<2631:PADOSL>2.0.CO;2.

Portin, H. J., M. Komppula, A. P. Leskinen, S. Romakkaniemi, A. Laaksonen, and K. E. J. Lehtinen, 2009: Observations of aerosol–cloud interactions at the Puijo semi-urban measurement station. *Boreal Env. Res.,* **14,** 641–653.

Rogers, R. R., and M. K. Yau, 1989: *A Short Course in Cloud Physics.* 3rd ed. International Series in Natural Philosophy, Vol. 113, Butterworth Heinemann, 304 pp.

SAFNWC, 2013: Algorithm theoretical basis document for "cloud products" (CMa-PGE01 v3.2, CT-PGE02 v2.2 & CTTH-PGE03 v2.2). SAFNWC SAF/NWC/CDOP2/MFL/SCI/ATBD/01, Issue 3, Rev. 2.1, 87 pp. [Available online at http://www.nwcsaf.org/indexScientificDocumentation.html.]

Tafferner, A., T. Hauf, C. Leifeld, T. Hafner, H. Leykauf, and U. Voigt, 2003: ADWICE: Advanced Diagnosis and Warning System for Aircraft Icing Environments. *Wea. Forecasting,* **18,** 184–203, doi:10.1175/1520-0434(2003)018<0184:AADAWS>2.0.CO;2.

Tammelin, B., M. Cavaliere, H. Holttinen, C. Morgan, H. Seifert, and K. Säntti, 2000: Wind energy production in cold climate (WECO). Finnish Meteorological Institute Rep. JOR3-CT95-0014, 38 pp. [Available online at http://cordis.europa.eu/documents/documentlibrary/47698271EN6.pdf.]

——, and Coauthors, 2005: Wind turbines in icing environment: Improvement of tools for siting, certification and operation. Finnish meteorological Institute NEW ICETOOLS Rep. 2005:6, 127 pp.

Reprinted, with permission, from
Hydrol. Earth Syst. Sci., 21, 267–279,
doi:10.5194/hess-21-267-2017

III

Hydrol. Earth Syst. Sci., 21, 267–279, 2017
www.hydrol-earth-syst-sci.net/21/267/2017/
doi:10.5194/hess-21-267-2017

Hydrology and Earth System Sciences
Open Access
EGU

Improving the precipitation accumulation analysis using lightning measurements and different integration periods

Erik Gregow[1], **Antti Pessi**[2], **Antti Mäkelä**[1], **and Elena Saltikoff**[1]

[1]Meteorological Research, Finnish Meteorological Institute, P.O. Box 503, 00101 Helsinki, Finland
[2]Applied Meteorology, Vaisala, 3 Lan Dr., Westford, MA 01886, USA

Correspondence to: Erik Gregow (erik.gregow@fmi.fi)

Received: 7 March 2016 – Published in Hydrol. Earth Syst. Sci. Discuss.: 17 March 2016
Revised: 24 November 2016 – Accepted: 17 December 2016 – Published: 11 January 2017

Abstract. The focus of this article is to improve the precipitation accumulation analysis, with special focus on the intense precipitation events. Two main objectives are addressed: (i) the assimilation of lightning observations together with radar and gauge measurements, and (ii) the analysis of the impact of different integration periods in the radar–gauge correction method. The article is a continuation of previous work by Gregow et al. (2013) in the same research field.

A new lightning data assimilation method has been implemented and validated within the Finnish Meteorological Institute – Local Analysis and Prediction System. Lightning data do improve the analysis when no radars are available, and even with radar data, lightning data have a positive impact on the results.

The radar–gauge assimilation method is highly dependent on statistical relationships between radar and gauges, when performing the correction to the precipitation accumulation field. Here, we investigate the usage of different time integration intervals: 1, 6, 12, 24 h and 7 days. This will change the amount of data used and affect the statistical calculation of the radar–gauge relations. Verification shows that the real-time analysis using the 1 h integration time length gives the best results.

1 Introduction

Accurate estimates of accumulated precipitation are needed for several applications such as flood protection, hydropower, road- and fire-weather models. In Finland, one of the most economically relevant users of precipitation is the hydropower industry. Between 10 and 20 % of Finnish annual electric power production comes from hydropower, depending on the amount of precipitation and water levels in dams and water reservoirs. In order to maintain correct calculation of the energy supplied to customers and to avoid (or at least minimize) the environmental risks and economical losses during extreme precipitation and flooding events, a profound analysis of the expected water amounts in dams and reservoirs from catchment areas is needed. The current hydropower strategy of Finland is to increase capacity by improving the efficiency of existing plants through technical adjustments. The maintenance and planning of proper dam structures need the most up-to-date information about the rain rates to be able to adjust the regulation functions of the dams, both for the current and the changing climatic conditions (IPCC-AR5, 2013).

Often, the accumulated precipitation values are based on pure radar analysis, unless there exists a surface gauge observation in the immediate surroundings. Radar echoes are related to rainfall rate and thereafter transformed into accumulation values. However, such conversions are based on general empirical relations which are not suitable for all meteorological cases (e.g., depending on precipitation type; Koistinen and Michelson, 2002). Radar reflectivity can, in some cases, suffer from poor quality, resulting from electronic miscalibration, beam blocking, clutter, attenuation and overhanging precipitation (Saltikoff et al., 2010), which results in poor estimations of the precipitation accumulation. In some cases, the radar can even be missing, e.g., during maintenance, upgrading or due to technical problems. Especially during thunderstorms, there is a potential of radar disturbances, either in the form of missing data due to interrup-

tions in electricity and telecommunication systems, or in the form of quality issues such as attenuation, due to intervening heavy precipitation.

The research of combining radar and surface observations, to perform corrections to precipitation accumulation, is well explored. Many have made developments in this field and much literature is available, for example, Sideris et al. (2014), Schiemann et al. (2011) and Goudenhoofdt and Delobbe (2009). In general, combining radar and rain gauge data is very difficult in the vicinity of heavy local rain cells (Einfalt et al., 2005). Recently, Jewell and Gaussiat (2015) compared performances of different merging schemas and noted a large difference between convective and stratiform situations. In their study, the nonparametric kriging with external drift outperformed other methods in an accumulation period of 60 min. Wang et al. (2015) developed a sophisticated method for urban hydrology, which preserves the non-normal characteristics of the precipitation field. They also noticed that common methods have a tendency to smooth out the important but spatially limited extremes of precipitation.

Comparing radars and gauges, an additional challenge arises from the different sampling sizes of the instruments. Radar measurement volume can be several kilometers wide and thick (a 1° beam is approximately 5 km wide at 250 km), while the measurement area of a gauge is 400 cm^2 (weighing gauges) or 100 cm^3 (optical instruments). Part of the disparateness of radar and gauge measurements is due to variability of the raindrop size distribution within the area of a single radar pixel. Jaffrain and Berne (2012) have observed variability up to 15 % of the rain rate in a 1×1 km pixel, with time steps of 1 min.

Lightning is associated with convective precipitation, but in areas where a large portion of precipitation is stratiform, lightning data alone are not adequate for precipitation estimation. Although convective events contribute only a fraction of the annual precipitation amount, they might be important during flooding events. However, lightning has been used to complement and improve other datasets. Morales and Agnastou (2003) combined lightning with satellite-based measurements to distinguish between convective and stratiform precipitation area and achieved a remarkable 31 % bias reduction, compared to satellite-only techniques. Lightning has also been assimilated to numerical weather prediction (NWP) models, using nudging techniques, or improving the initialization process of the model. This can be done by blending them with other remote sensing data to create heating profiles (e.g., estimating the latent heat release when precipitation is condensed). Papadopulos et al. (2005) used lightning data to identify convective areas and then modified the model humidity profiles, allowing the model to produce convection and release latent heat using its own convective parameterization scheme. They combined lightning with 6-hourly gauge data, within a mesoscale model in the Mediterranean area, and showed improvement in forecasts up to 12 h lead time. Pessi and Businger (2009) derived a lightning–convective rainfall relationship over the North Pacific Ocean and used it for latent heat nudging method in an NWP model. They were able to improve the pressure forecast of a North Pacific winter storm significantly.

Our situation is different from the above-mentioned experiments because lightning activity is usually low in Finland, compared to warmer climates (Mäkelä et al., 2011). Also, our analysis area already has a good radar coverage and a relatively evenly distributed network of 1 h gauge measurements. However, if we want to enlarge the analysis area, we will soon go to either sea areas or neighboring countries where availability of radar data and frequent gauge measurements is low. We also anticipate the usefulness of lightning data as a backup plan in the occasions when radar data are either missing or of deteriorated quality. Even though these occasions are rare, they often occur on days when detailed precipitation estimates are of great interest. Thunderstorms producing heavy localized rainfall are also often producing heavy winds, causing unavailability of radar data due to breaks on electricity and data communications. Our principal goal is to have as good an analysis as possible, which is different from having a best analysis to start a model.

Gregow et al. (2013) have demonstrated the benefit of assimilating different data sources (radars and gauges) in precipitation estimation. The largest uncertainties were observed during heavy convective rainfall. These are the situations when lightning occurs. The accumulation process is based on the radar reflectivity field, where gauges correct the initial field; e.g., if there is no reflectivity field, there is no accumulation (gauges are not used alone). To improve the spatially accurate real-time precipitation analysis, new methods are adopted by fusion of weather radar, lightning observations and rain gauge information in novel ways. This leads to better possibilities in estimating convective rainfall events (i.e., $> 5\ mm\ h^{-1}$) and the accumulated precipitation for the benefit of hydropower management and other related application areas. The work reported here has been performed using the Local Analysis and Prediction System (LAPS), which is used operationally in the weather service of the Finnish Meteorological Institute (FMI). Testing new approaches in an operational system has its challenges. For example, it is not possible to exclude a large amount of independent reference stations. Also, the possibilities to rerun cases with different settings have been limited. The major benefit of working in an operational environment is that we can be sure that we only use data and methods which are operationally available and feasible.

In this article, the observational datasets are described in Sect. 2. New methods on how to calculate the precipitation accumulation are handled in Sect. 3, and the results and discussion are shown in Sects. 4 and 5, respectively.

2 Observations and instrumentation

Here, we describe the three data sources employed in this study (rain gauge, radar and lightning observations) and the verification periods used in this study.

2.1 Rain gauge observations

Rain gauges provide point observations of the accumulation. They are usually considered more accurate than radar as point values and are frequently used to correct the radar field (Wilson and Brandes, 1979). The surface precipitation network (in total, 472 stations) consists of standard weighting gauges and optical sensors mounted on road-weather masts. Since 2015, FMI has managed 102 stations instrumented with the weighting gauge OTT Messtechnik Pluvio2. The Finnish Transport Agency (FTA) runs 370 road-weather stations with optical sensor measurements (Vaisala Present Weather Detectors models PWD22 and, to some extent, PWD11). The precipitation intensity is measured in different time intervals which are summed up to 1 h precipitation accumulation information. Uncertainties and more detailed information can be found in Gregow et al. (2013). If measurements consistently indicate poor data quality, either manually identified from station error logs or by inspecting the data, those stations are blacklisted within the LAPS process and do not contribute to the precipitation accumulation analysis. Hereafter, in this article, the weighting gauges and road-weather measurements are indistinctly called gauges and their placement in Finland is shown in Fig. 1a.

2.2 The radar data

As of summer 2016, FMI operates 10 C-band Doppler radars (with the newest one operational since late 2015). All but one station (VIM in western Finland; see Fig. 1b) are dual-polarization radars. At the moment, the quantitative precipitation estimation based on dual polarization is not used operationally in FMI, but the polarimetric properties contribute to the improved clutter cancellation (i.e., removal of non-meteorological echoes, especially sea clutter, birds and insects). In southern Finland, the distance between radars is 140–200 km, but in the north, the distance between stations LUO and UTA is 260 km. The location of the radars and the coverage is shown in Fig. 1b. As Finland has no high mountains, the horizon of all the radars is near zero elevation with no major beam blockage, and, in general, the radar coverage is very good except in the most northern part of the country. The Finnish radar network does have a very high system utilization rate (e.g., no interruption). During the years 2014 and 2015, the utilization rate was $>99\,\%$. Further details of the FMI radar network and processing routines are described in Saltikoff et al. (2010).

The basic radar volume scan consists of 13 plan position indicator (PPI) sweeps. The FMI-operated LAPS version (hereafter FMI-LAPS) is using the six lowest elevations: 0.3 (alternative 0.1 or 0.5, depending on site location), 0.7, 1.5, 3.0, 5.0 and 9.0, which are scanned out to 250 km, and repeated every 5 min. These radar volume scans are further used in LAPS routines for the rain rate calculations but also as proxy data to the lightning data assimilation (LDA) method (see Sect. 3.2).

2.3 The Lightning Location System (LLS)

The Lightning Location System (LLS) of FMI is part of the Nordic Lightning Information System (NORDLIS). The system detects cloud-to-ground (CG) and intracloud (IC) strokes in the low-frequency (LF) domain. Finland is situated between 60–70° N and 19–32° E, and thunderstorm season begins usually in May and lasts until September. During the period 1960–2007, on average, 140 000 ground flashes occurred during approximately 100 days per year (Tuomi and Mäkelä, 2008). The present modern LLS was installed in summer 1997 (Tuomi and Mäkelä, 2007; Mäkelä et al., 2010, 2016). The system consists of Vaisala Inc. sensors of various generations, and the sensor locations in 2015 and the efficient network coverage area can be seen in Fig. 2. Lightning location sensors detect the electromagnetic (EM) signals emitted by lightning return strokes, and measure the signal azimuth and exact time (GPS). Sensors send this information to the central processing computer in real time which combines them, optimizes the most probable strike point and outputs this information to the end user. More detailed information of LLS principles is described in Cummins et al. (1998).

2.4 Verification periods

The verification periods are limited to summer season (the active convective season in Finland) where two long periods were included in the verification: (a) 1 April to 1 September 2015 and (b) 1 May to 26 July 2016. These long verification periods include many cases of stratiform precipitation with no lightning, and therefore the effective impact by lightning is diluted (e.g., no influence by the LDA method). Hence, two subsets of two lightning intensive cases (e.g., situations with heavy rain and strong convection), datasets (c) and (d), were used to explicitly find the lightning-induced impacts. The dataset (c) includes full days (24 h periods) with more than 100 CG strokes per day. The dataset (d) includes only the stations and time intervals affected by lightning (defined as stations with maximum distance of 30 km to the lightning position and within the 1 h accumulation time interval, hereafter called the scaled dataset). An early dataset from 2014, dataset (e), consists of 4 days (3, 23, 24 and 30 July 2014) with more than 100 CG strokes per day. This dataset was used to perform several autonomous experiments with the FMI-LAPS LDA system in the early stage of the development of the LDA method.

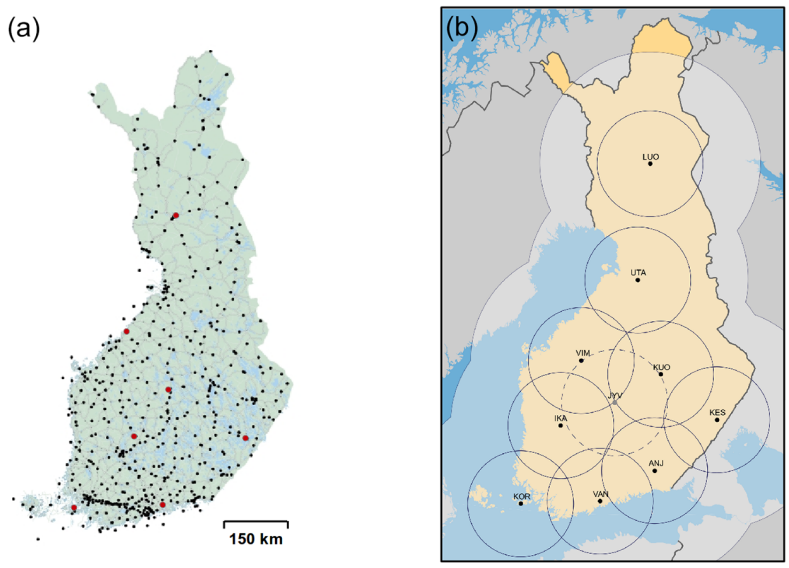


Figure 1. In panel **(a)**, the Finnish surface gauge stations are shown (as dots on the map); these are used to measure the hourly precipitation accumulation. The red dots indicate the position of the seven independent stations used for the verification. In panel **(b)**, the outer rectangular frame of the map depicts the LAPS analysis domain. The black dots represent the 10 Finnish radar stations and the outer black curved lines display their coverage. The thin circles surrounding each radar represent the areas where measurements are performed below 2 km height. The dashed circle indicates radar station JYV, which was not included in the radar network during summer 2015.

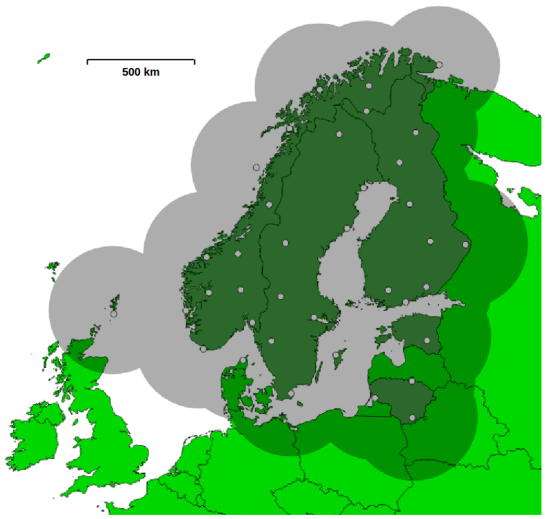


Figure 2. The LLS sensor locations (white dots) and coverage (grey circular areas) as of the year 2015.

3 Methods

The systems used to assimilate radar, gauge and lightning measurements are described in Sect. 3.1 and 3.2. The impact of different integration time periods on the regression and Barnes (RandB) method is shown in Sect. 3.3 and 3.4 and the verification methods in Sect. 3.5.

3.1 The Local Analysis and Prediction System (LAPS)

The LAPS produces 3-D analysis fields of several different weather parameters (Albers et al., 1996). LAPS performs a high-resolution spatial analysis where observational input from several sources is fitted to a coarser background model first-guess field (e.g., ECMWF forecast model). Additionally, high-resolution topographical data are used when creating the final analysis fields. The FMI-LAPS products are mainly used for nowcasting purposes (i.e., what is currently happening and what will happen in the next few hours), which is of critical interest for end users who demand near-real-time products.

The FMI-LAPS use a pressure coordinate system including 44 vertical levels distributed with a higher resolution (e.g., 10 hPa) at lower altitudes and decreasing with height. The horizontal resolution is 3 km and the temporal resolution is 1 h. The domain used in this article covers the whole country of Finland and some parts of the neighboring coun-

tries (Fig. 1b). LAPS highly relies on the existence of high-resolution observational network, in both space and time, and especially on remote sensing data. The FMI-LAPS is able to process several types of in situ and remotely sensed observations (Koskinen et al., 2011), among which radar reflectivity, weighting gauges and road-weather observations are used for calculating the precipitation accumulation. The Finnish radar volume scans are read into LAPS as NetCDF format files; thereafter, the data are remapped to the LAPS internal Cartesian grid and the mosaic process combines data of the different radar stations (Albers et al., 1996). The rain rates are calculated from the lowest levels of the LAPS 3-D radar mosaic data via the standard Z-R formula (Marshall and Palmer, 1948), which is then used for precipitation accumulation calculations (see Sect. 3.2). Other information on observational usage, first-guess fields, the coordinate system etc. is described in Gregow et al. (2013).

In this study, the lightning data are ingested into the FMI-LAPS. Modifications have been made to the software in order to use it together with FMI operational radar input data and the new lightning algorithms.

3.2 The LAPS lightning data assimilation (LDA) method

A lightning data assimilation (hereafter LDA) system has been developed by Vaisala and distributed as open and free software (Pessi and Albers, 2014). The LDA method is constructed to build up statistical relationships between radar and lightning measurements. The lightning information used for the LAPS LDA method is the location data (e.g., time, longitude and latitude) for each CG lightning stroke. LDA counts the amount of CG lightning strokes and converts lightning rates into vertical radar reflectivity profiles within each LAPS grid cell. The radar reflectivity–lightning (hereafter Rad-Lig) relationship profiles may differ depending on the local geographical regime and climate. A set of default profiles are included within the LDA package, which were derived over the eastern United States with the use of radar data from NEXRAD network and lightning data from the GLD360 network (Pessi, 2013; Said et al., 2010). These profiles can be used as a first guess if profiles for the local climate are not available.

For this study over Finland, climatological Rad-Lig reflectivity relationship profiles were estimated using NORDLIS-LLS lightning information and operational radar volume data from the Finland area during summer 2014. A total of approximately 220 000 lightning strokes were used for this calibration. The FMI-LAPS LDA used a 5 min interval of lightning and radar data, within a LAPS grid box of 3×3 km resolution. The collected strokes are divided into binned categories using an exponential division (i.e., 2^n ... 2^{n+1}), following the same method used in Pessi (2013). This results in six different lightning categories (e.g., with 1, 2–3, 4–7, 8–15, 16–31 and 32–63 strokes) for the NORDLIS-LLS dataset. For each of these six categories, the average reflectivity is calculated at each grid point for each level and gives the average Rad-Lig profiles (Fig. 3a), which is the baseline method. There is a good correlation ($R^2 = 0.95$) between the maximum reflectivity of profile and number of lightning strokes (Fig. 3b; results shown for the average Rad-Lig profiles). We extend this method to also calculate the third quartile (i.e., 75 % percentile) and the variable quartile Rad-Lig profiles, for each category. The variable quartile method uses a range between the 50 % percentile (for the lower dBZ values) and the 95 % percentile (for the highest dBZ values). The specific percentiles used for the six categories are the 50, 50, 60, 75, 90 and 95 % percentiles, respectively. The reasoning is to take into account the uncertainties in the low categories (due to larger spread and bias in the collected datasets) and, on the other hand, rely on the high percentiles for the high categories (since these have less spread). The profiles from the two categories with largest amount of strokes have the least data, because they are the rarest categories. All datasets suffer from missing data at some height levels, but these two categories are more sensitive due to the overall small data amounts. This can sometimes create artificial peaks of reflectivity values that are too low. This was especially seen at high altitudes, which can partly be explained by the radar measurement geometry. Therefore, these two reflectivity profiles have been manually smoothed to have the same shape as the other profiles.

The Rad-Lig reflectivity profiles can be used either independently or merged with the radar data in the LAPS accumulation analysis. When merging the two sources, radar and lightning reflectivity values are compared at each grid point both horizontally and vertically. The data source giving the highest reflectivity value will be used in that LAPS grid point. The logic behind this is that the radars are more likely to underestimate than overestimate the precipitation (due to attenuation, beam blocking or the nearest radar missing from the network; e.g., Battan, 1973; Germann, 1999), especially in thunderstorm situations. This is an approximation, aiming to compensate for the most serious radar error sources, which could be a subject for further improvement in future developments (especially if independent quality estimates of the radar data become available). LAPS then uses the generated 3-D volume reflectivity field in a similar manner, as it would use the regular volume radar data, for example, to adjust hydrometeor fields and rainfall.

The reflectivity (Z; $\mathrm{mm^6\,m^{-3}}$) parameter, measured by the radar or estimated by LDA method, is converted to precipitation intensity (R; $\mathrm{mm\,h^{-1}}$) within LAPS, using a pre-selected Z-R equation (Marshall and Palmer, 1948) as of the type

$$Z = A \cdot R^b, \tag{1}$$

where A and b are empirical factors describing the shape and size distribution of the hydrometeors. In FMI-LAPS's implementation, $A = 315$ and $b = 1.5$ for liquid precipitation,

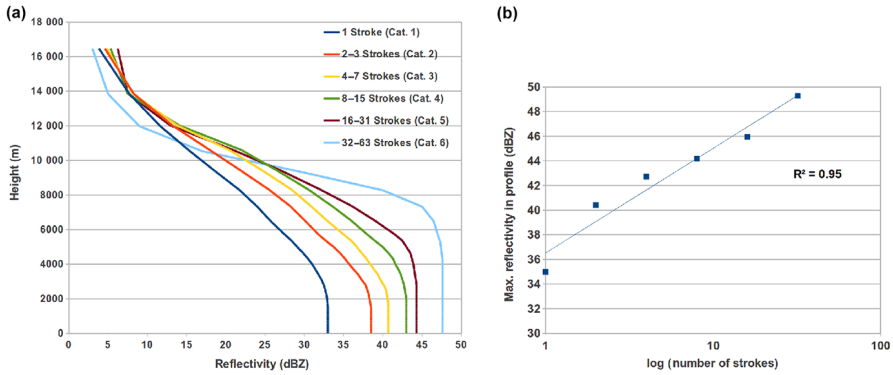


Figure 3. Panel **(a)** shows Rad-Lig relationship profiles (smoothed) from Finland NORDLIS-LLS, calculated using the dataset from summer 2014. Profiles are divided into binned categories of strokes, with a temporal resolution of 5 min and spatial resolution of 3 km. Panel **(b)** shows profiles' max reflectivity values vs. lightning rate (logarithmic scale of bins).

which is relevant in this study since it is carried out during the summer period. These static values introduce a gross simplification, since the drop size and particle shapes vary according to the weather situation (drizzle/convective, wet snow/snow grain). Challenging situations include both convective showers, with heavy rainfall, and the opposite event of drizzle, with little precipitation (Uijlenhoet, 2001). On the other hand, the same static factors have been used for many years in FMI's other operational radar products, and looking at long-term averages, the radar accumulation data do match the gauge accumulation values within reasonable accuracy (Aaltonen et al., 2008). The intensity field (R; Eq. 1) is calculated at every 5 min, and the 1 h accumulation is thereafter obtained by accumulating 5 min intervals. Gires et al. (2014) have shown that the scale difference has an effect on verification measures (such as normalized bias, e.g., RMSE) but it decreases with growing accumulation time (e.g., from 5 to 60 min). In our study, the 60 min accumulation period is smoothing some of the differences.

The following FMI-LAPS precipitation accumulation products are calculated based on radar (hereafter Rad_Accum), LDA (hereafter LDA_Accum) and the combined radar and LDA (hereafter Rad_LDA_Accum) precipitation accumulation.

3.3 The FMI-LAPS regression and Barnes (RandB) analysis method

The FMI-LAPS RandB method corrects the precipitation accumulation estimates using radar and gauge datasets. The first step in this method is to make the radar–gauge correction using the regression method. Data of hourly accumulation values are derived from the radar–gauge pairs within the LAPS grid (i.e., from the same location and time), and from this a linear regression function can be established. The corrections from the regression method are applied to the whole radar accumulation field and thereafter used as input for the second step, the Barnes analysis. Within LAPS routines, the Barnes interpolation converge the radar field towards gauge accumulation measurements at smaller areas (i.e., for gauge station surroundings). Several iterative correction steps are performed within the Barnes analysis, adjusting the final accumulation. The FMI-LAPS RandB method is described in more details in Gregow et al. (2013).

In this article, the RandB method is used to calculate the precipitation accumulation with the use of radar, gauges, lightning and the combination of radar–lightning. This gives the additional three FMI-LAPS accumulation products: Rad_RandB, LDA_RandB and Rad_LDA_RandB, respectively.

3.4 RandB method and the integration time period

The original FMI-LAPS RandB method uses radar and gauge data from the recent hour. Using only the latest hour, the gauge observational dataset can suffer from too few observations and thereby affect the quality and robustness of the regression and Barnes calculations. As a further investigation in this article, we use a selection of longer time periods (e.g., the previous 6, 12, 24 h and 7 days of data) in order to build up a larger radar–gauge dataset. These datasets are thereafter used to make the correction within the RandB method.

We have limited our studies to compare how the occurring synoptic weather situation, i.e., frontal or convective situation (1 to 12 h), and the medium-time-range information (24 h to 7 days) impact the accumulation analysis. The longer the integration time, the less information on the situational weather occurring at analysis time; i.e., the dataset is getting more smoothed and extremes might disappear.

Verification was done for the summer 2015 period using the input from radar and lightning, and gives the following resulting accumulation products: Rad_LDA_RandB (i.e.,

dataset collected within the last 1 h), Rad_LDA_RandB_6hr, Rad_LDA_RandB_12hr, Rad_LDA_RandB_24hr and Rad_LDA_RandB_7d, respectively.

3.5 Verification methods

The hourly accumulation results have been verified against surface gauge observations, both dependent and independent stations. The dependent station data are included in the FMI-LAPS analysis calculating the 1 h precipitation accumulation; i.e., the analysis is depending on the station information used as input. There are seven independent stations which are excluded from the LAPS analysis. Note that, in the Rad_Accum and Rad_LDA_Accum products, the gauge data have not been used; therefore, all gauge stations are independent references for their verification. In this study, we apply a filter to the verification datasets where hourly accumulation data less than 0.3 mm are discarded (due to the lowest threshold value of surface gauge measurements from the FMI database). In a separate verification exercise for the 2016 data, only stations located more than 100 km and more than 150 km from the nearest radar station were used to demonstrate the potentially deteriorating quality of radar data with distance to the radar due to, e.g., attenuation and beam broadening (a 1° beam is 5 km wide at a distance of 250 km).

The validation of the different analysis methods is based on the logarithmic standard deviation (SD; Eq. 2), root mean square deviation (RMSE; Eq. 3) and Pearson's correlation coefficient (CORR; Eq. 4):

$$\mathrm{SD} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\log\left(\frac{\mathrm{Analysis}}{\mathrm{Gauge}}\right)_i - \overline{\log\left(\frac{\mathrm{Analysis}}{\mathrm{Gauge}}\right)}\right)^2 \quad (2)$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}\left((\mathrm{Analysis}-\mathrm{Gauge})_i\right)^2}{N-1}} \quad (3)$$

$$\mathrm{CORR} = \frac{\sum_i\left(\left(\mathrm{Gauge}_i - \overline{\mathrm{Gauge}}\right)\left(\mathrm{Analysis}_i - \overline{\mathrm{Analysis}}\right)\right)}{\sqrt{\sum_i\left(\mathrm{Gauge}_i - \overline{\mathrm{Gauge}}\right)^2\sum_i\left(\mathrm{Analysis}_i - \overline{\mathrm{Analysis}}\right)^2}}. \quad (4)$$

SD quantifies the amount of variation (i.e., spread) of a dataset. A low SD indicates that the data points tend to be close to the mean value of the dataset. Here, we use the logarithm of the quotients, in order to get the datasets closer to be normally distributed. RMSE is a quadratic scoring rule which measures the average magnitude of the error. Since the errors are squared before they are averaged, RMSE gives a relatively high weight to large errors. CORR gives a measure of the linear relationship (both strength and direction) between two quantities.

4 Results

Verification results using lightning data are presented in Sect. 4.1 and the impact from different integration time intervals in Sect. 4.2.

4.1 FMI-LAPS LDA results

The verification for the entire summer of 2015, i.e., using verification dataset (a) including days with no thunderstorms, assures that introducing lightning data has no significant impact on the overall performance of the system. The impact of using the LDA method for estimating the precipitation accumulation is neutral for this long verification period (shown in Fig. 4, where the data are from dependent stations). The same result is seen in the scores of RMSE, SD and CORR values (not included here). Since the data have been much influenced by weather situations not related to lightning, the focus will be on the subsets, i.e., datasets (c) and (d), the 25-day periods of intense lightning days of both 2015 and 2016, respectively.

The 25-day period with frequent thunderstorms during summer 2015, verification dataset (c), for which we used the average method to calculate the Rad_Lig profiles, shows an inconsistent result using lightning data (see Table 1, left column). For the independent dataset, the Rad_LDA_Accum has a slightly improved result (lower RMSE value) when compared with Rad_Accum. On the other hand, Rad_LDA_RandB gets worse results, as can be seen from the RMSE and CORR. The dependent data show almost neutral impact (RMSE is slightly better for Rad_LDA_RandB) with the use of the LDA method and average calculated Rad-Lig profiles.

Figure 5 shows the results using verification dataset (e), where different Rad-Lig profiles are compared (e.g., average, third quartile and variable quartile profiles) and validated against Rad_Accum. The precipitation accumulation estimates are improved at high accumulation values (> 5 mm) using either third or variable quartile profiles. Simultaneously, they both add to the overestimate in low accumulation values (< 5 mm). The third quartile profiles give the largest overestimate over the whole accumulation scale. The variable quartile gives the overall best result, with improved estimates for high accumulation values and only slight overestimation at low values.

The results, from the scaled dataset (d) and the dependency of distance to radar location, reveal the positive impact of using the lightning data as input for the LAPS-LDA model. Hence, using the variable quartile profiles in the accumulation analysis for the 25-day dataset of summer 2016 has a positive impact on the accumulation estimates (see Table 1, right column). Even if the improved scores are relatively small (the largest reduction in RMSE being 6.3 %), the LDA method shows a consistent correction of the results. The independent verification gives decreased RMSE

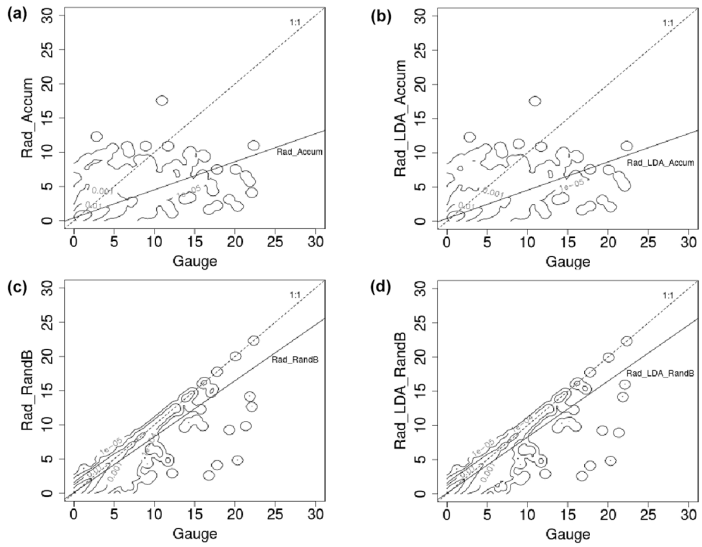


Figure 4. The FMI-LAPS precipitation accumulation (described in plots with density isolines of hourly accumulation values in millimeters) calculated using four different methods. Fit in solid line (see regression equations), the perfect solution would align on the 1 : 1 dashed line: **(a)** Rad_Accum ($y = 0.410x + 0.398$), **(b)** Rad_LDA_Accum ($y = 0.413x + 0.396$), **(c)** Rad_RandB ($y = 0.817x + 0.093$) and **(d)** Rad_LDA_RandB ($y = 0.819x + 0.091$). Results are from the dependent gauge dataset during summer 2015, i.e., verification dataset (a).

Table 1. Precipitation accumulation results from summer of 2015 (i.e., dataset c, left column) and 2016 (i.e., dataset d, right column), for periods of the 25 intensive lightning days (e.g., > 100 CG strokes per day) during both years. Precipitation results are shown for radar (Rad_Accum) and radar merged with lightning data (Rad_LDA_Accum), together with and without gauge measurements included with the RandB method (Rad_RandB and Rad_LDA_RandB, respectively). In the lowest panels, only data from more than 100 or 150 km from the nearest radar are used. Verification is performed against both independent and dependent stations, i.e., those used or left out from the gauge analysis.

Summer 2015 (average scheme)					Summer 2016 (variable quartile scheme)				
Independent	Rad_ Accum	Rad_LDA_ Accum	Rad_ RandB	Rad_LDA_ RandB	Independent	Rad_ Accum	Rad_LDA_ Accum	Rad_ RandB	Rad_LDA_ RandB
No. obs	3206	3332	256	256	No. obs	1320	1333	74	74
SD	0.27	0.27	0.11	0.11	SD	0.32	0.32	0.12	0.11
RMSE	1.66	1.64	0.58	0.70	RMSE	2.62	2.60	0.92	0.89
CORR	0.67	0.67	0.97	0.96	CORR	0.64	0.65	0.96	0.96
Dependent					Dependent				
No. obs			3566	3567	No. obs			1364	1376
SD			0.12	0.12	SD			0.14	0.13
RMSE			0.77	0.76	RMSE			1.27	1.19
CORR			0.93	0.93	CORR			0.93	0.94
> 100 km					> 100 km				
No. obs					No. obs	656	656	694	698
SD					SD	0.34	0.34	0.15	0.15
RMSE					RMSE	2.44	2.39	1.03	1.01
CORR					CORR	0.66	0.67	0.95	0.95
> 150 km					> 150 km				
No. obs					No. obs	153	153	168	171
SD					SD	0.39	0.39	0.20	0.20
RMSE					RMSE	2.46	2.42	1.47	1.43
CORR					CORR	0.33	0.35	0.80	0.81

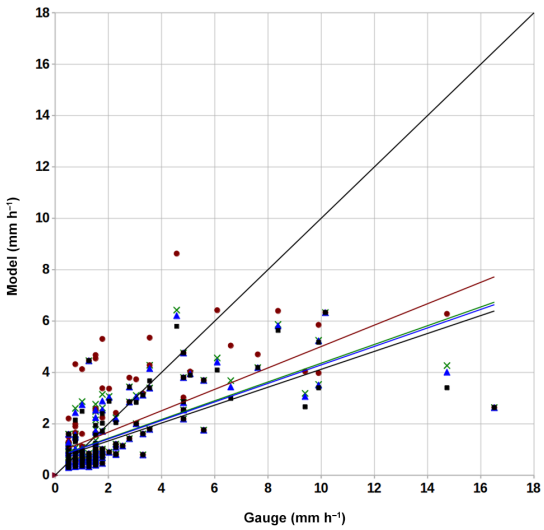


Figure 5. Verification of hourly accumulation values for Rad_Accum (black squares, with regression line equation $y = 0.349x + 0.638$) and LDA_Accum (triangle, cross and circular markers), using three different methods to calculate the relationship profiles: average (blue triangles, $y = 0.360x + 0.691$), third quartile (red circles, $y = 0.417x + 0.844$) and the variable quartile (green crosses, $y = 0.365x + 0.710$) accumulation estimates. The corresponding regression lines (see equations) are represented with same color as the markers for each method. Data are for the 4-day period in summer 2014, i.e., verification dataset (e). The best-fit curve (i.e., the 1 : 1 fit) is shown as a black solid line.

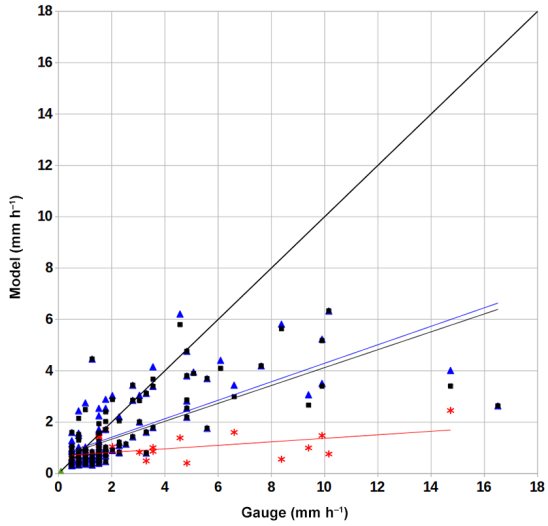


Figure 6. Verification of hourly accumulation values for LDA_Accum (red stars, with regression line equation $y = 0.068x + 0.685$) and the merged Rad_LDA_Accum (blue triangles, $y = 0.360x + 0.691$), compared to Rad_Accum (black boxes, $y = 0.349x + 0.638$). The corresponding regression lines (see equations) are represented with same color as the markers for each method. Data are for the 4-day period in summer 2014, i.e., verification dataset (e). The black solid line is the best-fit line (1 : 1 fit).

and increased CORR values for Rad_LDA_Accum compared to Rad_Accum. Also, Rad_LDA_RandB gets smaller errors than Rad_RandB (see SD and RMSE in Table 1, most upper-right panel). For the dependent stations, all scores are improved using the LDA method, especially the RMSE (as seen in Table 1, right column, second panel). The verification of distance dependencies, i.e., for observations further away than 100 and 150 km from the nearest radar stations, shows improved accumulation estimates when using the LDA method (see Table 1, right column, two last panels). The RMSE and CORR scores for Rad_LDA_Accum and Rad_LDA_RandB are better than Rad_Accum and Rad_RandB, respectively. Here, only dependent gauges are available for verification.

Comparing accumulation results from the 4-day period, i.e., verification dataset (e), for radar alone (Rad_Accum; black markers in Fig. 6) and lightning alone (LDA_Accum; red markers in Fig. 6), it is clear that the use of LDA_Accum is less accurate than Radar_Accum results. Figure 6 also shows that the Rad_LDA_Accum estimates (using the baseline method, with average Rad-Lig profiles) are amplified over the whole range of precipitation values, compared to Rad_Accum (Fig. 6; compare the blue with the black markers). For the high accumulation values ($> 5\,\mathrm{mm\,h^{-1}}$), this is a positive effect, while in the lower range ($< 5\,\mathrm{mm\,h^{-1}}$) there is an overestimation of the results.

4.2 RandB method and impact from different integration periods

The plotted results of different time sampling periods are seen in Fig. 7, where the density of points are drawn as isolines in the scatter plot, with verification against the independent stations from verification dataset (a). The Rad_LDA_RandB (i.e., using observations from the latest 1 h) does give the best result, when compared to Rad_LDA_Accum, Rad_LDA_RandB, Rad_LDA_RandB_6hr, Rad_LDA_RandB_12hr, Rad_LDA_RandB_24hr and the Rad_LDA_RandB_7d output. The statistical scores shown in Table 2 also imply the same result. The Rad_LDA_Accum (e.g., a method not using RandB) is included as a reference when comparing the results of different integration periods.

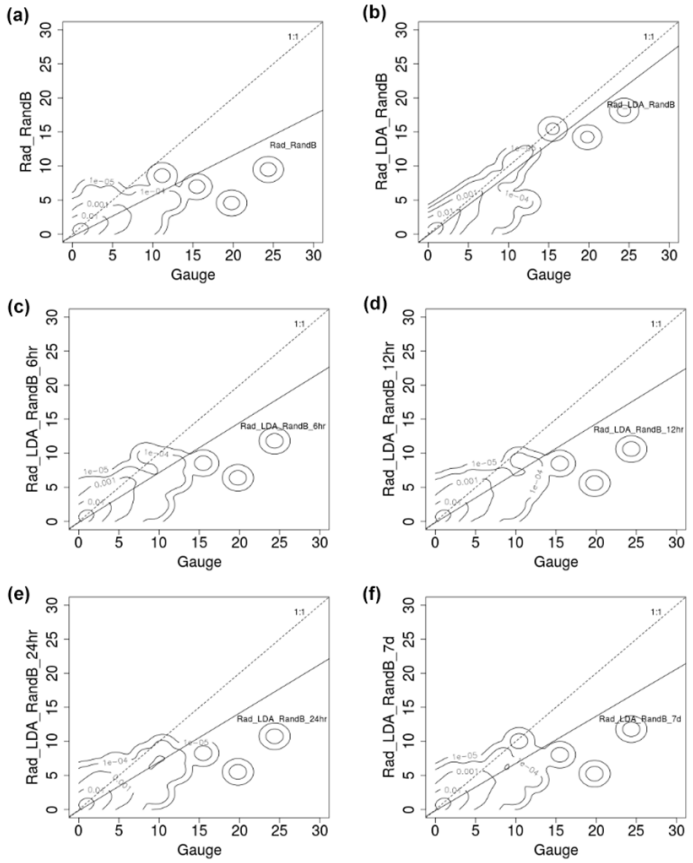


Figure 7. Impact of changing the integration time length, with verification for the independent gauges, using verification dataset (a) from summer 2015. Accumulation plots with density isolines of hourly values in millimeters: **(a)** Rad_LDA_Accum (with regression line equation $y = 0.594x - 0.312$), **(b)** Rad_LDA_RandB ($y = 0.891x - 0.147$), **(c)** Rad_LDA_RandB_6hr ($y = 0.732x - 0.160$), **(d)** Rad_LDA_RandB_12hr ($y = 0.725x - 0.169$), **(e)** Rad_LDA_RandB_24hr ($y = 0.715x - 0.167$) and **(f)** Rad_LDA_RandB_7d ($y = 0.692x - 0.166$). The fit is shown in solid lines (see regression equations); the perfect solution would align on the 1 : 1 dashed line.

Table 2. Impact of the integration time length on the RandB method for the dependent and independent stations datasets during summer 2015, i.e., dataset (a). The Rad_LDA_Accum (e.g., a method not using RandB) is included as a reference.

Dependent	Rad_LDA_ Accum	Rad_LDA_ RandB_1hr	Rad_LDA_ RandB_6hr	Rad_LDA_ RandB_12hr	Rad_LDA_ RandB_24hr	Rad_LDA_ RandB_7d
No. of observations	13 200	16 311	10 956	10 917	10 915	11 033
SD ($\log(R/G)$)	0.25	0.13	0.13	0.13	0.14	0.14
RMSE	1.20	0.52	0.67	0.71	0.72	0.72
CORR	0.64	0.93	0.91	0.90	0.89	0.89
Independent						
No. of observations	1177	1492	1028	1013	1005	1014
SD ($\log(R/G)$)	0.25	0.15	0.22	0.22	0.22	0.22
RMSE	1.38	0.68	1.16	1.23	1.24	1.24
CORR	0.39	0.92	0.79	0.77	0.77	0.77

5 Discussions and conclusions

The aim of this article is to describe new methods on how to improve the hourly precipitation accumulation estimates, especially for heavy rainfall events (> 5 mm) and as much as possible for the low-valued ranges (< 5 mm).

The strength of the LDA method is that the radar and lightning information can be merged and complement each other. This is especially important in areas of poor or even non-existent radar coverage, where the lightning information will improve the reflectivity field and thereby the hourly precipitation accumulation analysis. It is important to recall that, in the LAPS accumulation process, the reflectivity field is the first step, which is then corrected with gauges (e.g., if there is no reflectivity field, gauges will not be used and there will be no accumulation field). The results in this article are limited to Finland but should this area be extended to include Scandinavia, the LDA method will become even more useful. There are also other LAPS users in other parts of the world, whom we want to encourage to continue this work.

The whole summer periods of 2015 and 2016 show neutral impact on the results using the LDA method; scores are not included here but Fig. 4 shows the graphs for verification dataset (a). It is important to make long-term verification in order to see that the system is robust and does not generate any bad data during any weather situation, i.e., perform a sanity check of the system. However, in order to narrow down our analysis to areas and times where lightning did occur (i.e., exclude stratiform precipitation), we focused our results on the subset of 25 lightning intensive days for both 2015 and 2016, datasets (c) and (d), respectively. The subset of 2015, using the average method, gave inconsistent results and no unambiguous conclusions could be drawn (Table 1, left column).

New methods to calculate the Rad-Lig profiles were tested and reveal that the variable quartile method improves the estimates for the large accumulation (i.e., > 5 mm), though with some overestimation in low accumulation (Fig. 5). The third quartile approach has the highest impact on the whole accumulation field, which results in large overestimates for the low accumulation values (i.e., 0–5 mm). The average method smoothes out the small-scale variances, which are observed in heavy convection. Hence, the collected radar reflectivity profiles are less representative, and therefore the calculated Rad-Lig profiles will have values that are too low in these cases. As a result, the average method will have a low impact on the final precipitation accumulation estimates, compared to the use of the third quartile and variable quartile methods (Fig. 5). One should also mention that there is an overall uncertainty due to instrumental errors and the collocation between observations within the LDA method. This could potentially result in dislocation and bad quality of the received radar and lightning measurements, which would affect the calculated Rad-Lig profiles (for example, in the event of radar attenuation, where strong rainfall weakens some part

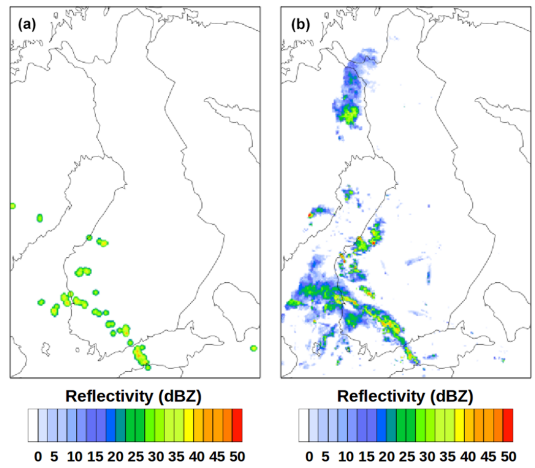


Figure 8. Reflectivity field simulated from lightning data alone (left) and, for verification, from radar data alone (right) 30 July 2014 at 16:00 UTC. The reflectivity color scale is shown below plots.

of the reflectivity field). Here, the collected radar profiles will have reflectivity values that are too low and give underestimated Rad-Lig profiles, especially when using the average method.

The newest results from 2016 and the 25-day subset show that there is a benefit to using the LDA (variable quartile) method. Mainly, all scores are becoming better and few are unchanged when lightning information is used to estimate the precipitation accumulation (see Table 1, right column). Verifying the dataset with distance to radar stations (i.e., gauges situated further away than 100 and 150 km) also shows the same results; the accumulation product is improved with the LDA method. The impact on scores is mainly in the second decimal, but they are consistent, and clearly show the tendency of improvement by using the LDA method with the variable quartile profiles. One reason we do not see a larger impact by the LDA method could be that the Finnish radar network does have a very high quality and system utilization rate and therefore is less impacted by the LDA method. In an upcoming version of FMI-LAPS, the verification will be focusing on including areas with poor (or non-existent) radar coverage where gauges are available.

The accumulation products generated from the RandB method are corrected using gauge information. This process influences the final accumulation results much more than the contribution from the LDA method (seen in Fig. 4 results from the dependent dataset, where a, c and b, d panels, respectively, are almost identical). The same result was seen for the independent dataset (not shown here). Nonetheless, we have proven that if there were no radar data (for example, if the radar is malfunctioning), precipitation accumulation information would be available from lightning data and

add value to the final product. This is shown in Fig. 6, where accumulation would be generated from the LDA method (as seen in Fig. 6; red markers) and also visualized through the example in Fig. 8, where the radar and Rad-Lig lowest reflectivity fields are plotted for one analysis time: 16:00 UTC, 30 July 2014. This case study also demonstrates how the LDA method can reconstruct the highest reflectivities, but areas with weak precipitation are missing.

In the RandB method, the regression is used to correct for large-scale multiplicative biases between radar and gauge data. In this article, we introduce lightning into the RandB method as an additional data source. However, lightning errors are likely to be different from those of radar and gauges, and this could have an effect on the methodology used here. In future developments, after collecting longer time series to quantify the nature of uncertainty of lightning-based precipitation estimates, we intend to improve the analysis in this direction.

In the present analysis area, we mainly anticipate the usefulness of lightning data as a backup plan of rare but significant cases. Due to the rare nature of such events, it is not possible to collect a statistically representative dataset in a few years; even though attenuation of radar signals or completely missing data are observed several times a summer, it is not so often that such events happen just over a rain gauge station. However, our overall analysis shows that when we include the lightning data every day at every point, they make, on average, a small improvement, and they are there as a safety network waiting for the cases where radars fail.

For the near-real-time accumulation product, data used from the recent hour of analysis time do give the best precipitation accumulation result (Table 2 and Fig. 7). We see correlation peaking at the 1 h integration period and decreasing already for the 6 h period. Therefore, according to the results in this study, the use of long time integration periods for the RandB method (until 7 days in this case) does not improve the hourly precipitation accumulation analysis. Berndt et al. (2014) compared data resolutions from 10 min to 6 h and reported a large improvement in the correlation (from 10 min to 1 h, the correlation increased 0.37 to 0.57). From 1 to 6 h, the corresponding increase was 0.57 to 0.62, respectively. In Norway, Abdella and Alfredsen (2010) have shown that the use of average monthly adjustment factors leads to less than optimal results. One could speculate that there is an intermediate choice of temporal resolution that would improve the results in this article. For example, there could be better results using periods of 2 to 5 h. This has not been investigated in this article but will be considered in future studies.

6 Data availability

LAPS source code, including the LDA method, is available from NOAA (2017). The materials and data used in this article are available upon request from Finnish Meteorological Institute (or main author).

Acknowledgements. We thank NOAA ESRL/GSD and Vaisala for their support of LAPS-LDA developments, Marco Gabella for his encouraging words and Asko Huuskonen for helping in the final and critical stage of evaluating the results.

Edited by: P. Molnar
Reviewed by: H. Leijnse, F. Marra, S. Ochoa Rodriguez, and one anonymous referee

References

Aaltonen, J., Hohti, H., Jylhä, K., Karvonen, T., Kilpeläinen, T., Koistinen, J., Kotro, J., Kuitunen, T., Ollila, M., Parvio, A., Pulkkinen, S., Silander, J., Tiihonen, T., Tuomenvirta, H., and Vajda, A.: Strong precipitation and urban floods (Rankkasateet ja taajamatulvat RATU), Finnish Environment Institute (Suomen ympäristökeskus), Helsinki, Finland, 80 pp., 2008.

Abdella, Y. and Alfredsen, K.: Long-term evaluation of gauge-adjusted precipitation estimates from a radar in Norway, Hydrol. Res., 41, 171–192, 2010.

Albers, S. C., McGinley, J. A., Birkenheuer, D. L., and Smart, J. R.: The local analysis and prediction system (LAPS): Analyses of clouds, precipitation, and temperature, Weather Forecast., 11, 273–287, 1996.

Battan, L. J.: Radar Observation of the Atmosphere, University of Chicago Press, Chicago, USA, 1973.

Berndt, C., Rabiei, E., and Haberlandt, U.: Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios, J. Hydrology, 508, 88–101, 2014.

Cummins, K. L., Murphy, M. J., Bardo, E. A., Hiscox, W. L., Pyle, R. B., and Pifer, A. E.: A combined TOA/MDF technology upgrade of the U.S. National Lightning Detection Network, J. Geophys. Res., 103, 9035–9044, doi:10.1029/98JD00153, 1998.

Einfalt, T., Jessen, M., and Mehlig, B.: Comparison of radar and raingauge measurements during heavy rainfall, Water Sci. Technol., 51, 195–201, 2005.

Germann, U.: Radome attenuation – a serious limiting factor for quantitative radar measurements?, Meteorol. Z., 8, 85–90, 1999.

Gires, A., Tchiguirinskaia, I., Schertzer, D., Schellart, A., Berne, A., and Lovejoy, S.: Influence of small scale rainfall variability on standard comparison tools between radar and rain gauge data, Atmos. Res., 138, 125–138, doi:10.1016/j.atmosres.2013.11.008, 2014.

Goudenhoofdt, E. and Delobbe, L.: Evaluation of radar-gauge merging methods for quantitative precipitation estimates, Hydrol. Earth Syst. Sci., 13, 195–203, doi:10.5194/hess-13-195-2009, 2009.

Gregow, E., Saltikoff, E., Albers, S., and Hohti, H.: Precipitation accumulation analysis – assimilation of radar-gauge measurements and validation of different methods, Hydrol. Earth Syst. Sci., 17, 4109–4120, doi:10.5194/hess-17-4109-2013, 2013.

IPCC-AR5: Climate change 2013, available at: http://www.ipcc.ch/report/ar5/wg1/ (last access: 9 January 2017), 2013.

Jaffrain, J. and Berne, A.: Influence of the subgrid variability of the raindrop size distribution on radar rainfall estimators, J. Appl. Meteorol. Clim., 51, 780–785, doi:10.1175/JAMC-D-11-0185.1, 2012.

Jewell, S. A. and Gaussiat, N.: An assessment of kriging-based rain-gauge-radar merging techniques, Q. J. Roy. Meteor. Soc., 141, 2300–2313, doi:10.1002/qj.2522, 2015.

Koistinen, J. and Michelson, D. B.: BALTEX weather radar-based precipitation products and their accuracies, Boreal Environ. Res., 7, 253–263, 2002.

Koskinen, J. T., Poutiainen, J., Schultz, D. M., Joffre, S., Koistinen, J., Saltikoff, E., Gregow, E., Turtiainen, H., Dabberdt, W. F., Damski, J., Eresmaa, N., Göke, S., Hyvärinen, O., Järvi, L., Karppinen, A., Kotro, J., Kuitunen, T., Kukkonen, J., Kulmala, M., Moisseev, D., Nurmi, P., Pohjola, H., Pylkkö, P., Vesala, T., and Viisanen, Y.: The Helsinki Testbed: A mesoscale measurement, research, and service platform, B. Am. Meteorol. Soc., 92, 325–342, doi:10.1175/2010BAMS2878.1, 2011.

Mäkelä, A., Tuomi, T. J., and Haapalainen, J.: A decade of high-latitude lightning location: Effects of the evolving location network in Finland, J. Geophys. Res., 115, D21124, doi:10.1029/2009JD012183, 2010.

Mäkelä, A., Rossi, P., and Schultz, D. M.: The daily cloud-to-ground lightning flash density in the contiguous United States and Finland, Mon. Weather Rev., 139, 1323–1337, doi:10.1175/2010MWR3517.1, 2011.

Mäkelä, A., Mäkelä, J., Haapalainen, J., and Porjo, N.: The verification of lightning location accuracy in Finland deduced from lightning strikes to trees, Atmos. Res., 172, 1–7, 2016.

Marshall, J. S. and Palmer, W. M.: The Distribution of raindrops with size, J. Meteorol., 5, 165–166, 1948.

Morales, C. A. and Anagnostou, E. N.: Extending the capabilities of high-frequency rainfall estimation from geostationary-based satellite infrared via a network of long-range lightning observations, J. Hydrometeorol., 4, 141–159, 2003.

NOAA: Latest release of LAPS software, available at: http://laps.noaa.gov/cgi/LAPS_SOFTWARE.cgi, last access: 11 January 2017.

Papadopoulos, A., Chronis, T. G., and Anagnostou, E. N.: Improving convective precipitation forecasting through assimilation of regional lightning measurements in a mesoscale model, Mon. Weather Rev., 133, 1961–1977, 2005.

Pessi, A.: Characteristics of Lightning and Radar Reflectivity in Continental and Oceanic Thunderstorms, 93th Annual American Meteorological Society Meeting, Austin, Texas, USA, 6–10 January 2013, availabe at: https://ams.confex.com/ams/93Annual/webprogram/Paper215562.html (last access: 9 January 2017), 2013.

Pessi, A. and Albers, S.: A Lightning Data Assimilation Method for the Local Analysis and Prediction System (LAPS): Impact on Modeling Extreme Events, 94th Annual American Meteorological Society Meeting, Atlanta, Georgia, USA, 1–6 February, 2014, available at: https://ams.confex.com/ams/94Annual/webprogram/Paper238715.html (last access: 9 January 2017), 2014.

Pessi, A. and Businger, S.: The Impact of Lightning Data Assimilation on a Winter Storm Simulation over the North Pacific Ocean, Mon. Weather Rev., 137, 3177–3195, 2009.

Said, R. K., Inan, U. S., and Cummins, K. L.: Long-range lightning geolocation using a VLF radio atmospheric waveform bank, J. Geophys. Res., 115, D23108, doi:10.1029/2010JD013863, 2010.

Saltikoff, E., Huuskonen, A., Hohti, H., Koistinen, J., and Järvinen, H.: Quality assurance in the FMI Doppler Weather radar network, Boreal Environ. Res., 15, 579–594, 2010.

Schiemann, R., Erdin, R., Willi, M., Frei, C., Berenguer, M., and Sempere-Torres, D.: Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland, Hydrol. Earth Syst. Sci., 15, 1515–1536, doi:10.5194/hess-15-1515-2011, 2011.

Sideris, I. V., Gabella, M., Erdin, R., and Germann, U.: Real-time radar-raingauge merging using spatiotemporal co-kriging with external drift in the alpine terrain of Switzerland, Q. J. Roy. Meteor. Soc., 140, 1097–1111, 2014.

Tuomi, T. J. and Mäkelä, A.: Lightning observations in Finland, Reports, Finnish Meteorological Institute, Helsinki, Finland, 2007:5, 49 pp, 2007.

Tuomi, T. J. and Mäkelä, A.: Thunderstorm climate of Finland 1998–2007, Geophysica, 44, 29–42, 2008.

Uijlenhoet, R.: Raindrop size distributions and radar reflectivity–rain rate relationships for radar hydrology, Hydrol. Earth Syst. Sci., 5, 615–628, doi:10.5194/hess-5-615-2001, 2001.

Wang, L.-P., Ochoa-Rodríguez, S., Onof, C., and Willems, P.: Singularity-sensitive gauge-based radar rainfall adjustment methods for urban hydrological applications, Hydrol. Earth Syst. Sci., 19, 4001–4021, doi:10.5194/hess-19-4001-2015, 2015.

Wilson, J. W. and Brandes, E. A.: Radar Measurement of Rainfall – A Summary, B. Am. Meteorol. Soc., 60, 1048–1058, 1979.

Reprinted, with permission, from
Bulletin of the American Meteorological Society, 94, 847–858,
DOI:10.1175/BAMS-D-12-00039.1

COLD-SEASON THUNDERSTORMS IN FINLAND AND THEIR EFFECT ON AVIATION SAFETY

BY A. MÄKELÄ, E. SALTIKOFF, J. JULKUNEN, I. JUGA, E. GREGOW, AND S. NIEMELÄ

One day in October 2011, 10 commercial planes took off into a convective storm in Finland, triggering lightning that temporarily blinded some of the pilots—demonstrating the need to improve warnings for this unseasonable weather.

It is estimated that every commercial airplane is struck on average once per year by lightning (Uman and Rakov 2003), and several studies have been focused on the "triggering effect" of an airplane to lightning (Clifford and Kasemir 1982; Mazur 1989; Moreau et al. 1992). According to Rakov and Uman (2005), an airplane hit is typically a single event, and only rarely are several planes hit within the same storm, for example, in Los Angeles on 24 February 1987 when at least six airplanes were hit within only a couple of hours. We will show a similar case with 10 hits during a single evening.

Despite the high peak current of a lightning flash, it is a very rare case that an airplane is severely damaged by the flash (Cherington and Mathys 1995); the lightning protection system of the airplane prevents the lightning current from entering the critical parts—say, fuel tanks—of the plane. However, minor damages, such as small holes, are reported (Plumer and Robb 1982; Uman and Rakov 2003).

Maybe the most famous and important, regarding the development of safety regulations, airplane accident by lightning occurred on 8 December 1963 in Maryland (Civil Aeronautics Board 1965); ▶

Detail of Fig. 1. See p. 848 for more information.

PanAm flight 214 exploded and crashed when gasoline fumes were ignited by a lightning flash, killing all 81 people onboard. After the incidence, the Federal Aviation Administration (FAA) regulated all airplanes flying in the U.S. airspace to have special lightning rods installed.

Other reported lightning-caused incidents include the fuel tank explosion of the Iranian commercial aircraft over Spain in 1976 (National Transportation Safety Board 1976) with the death of 17 persons; the Lineas Aéreas Nacionales Sociedad Anonima (LANSA) flight 508 in Peru in 1971, which caused the death of about 90 persons; an accident in Germany in 1988, when an aircraft lost its wing after being struck by lightning (21 casualties); and the glider accident of 1999 in Bedfordshire, United Kingdom. In the latter case, the glider was apparently literally blown apart by a high-peak current positive ground flash (AAIB 1999).

On 19 October 2011, several commercial aircrafts were struck by lightning during their approach or departure in the surroundings of the terminal control area (TMA) of Helsinki–Vantaa airport in southern Finland. Ten aircrafts, of three different aircraft types, reported a lightning strike and several departing aircrafts had to return to the airport because of a minor technical problem or because of a momentary blindness or deafness. Some of the pilots reported the whole windshield to be illuminated by electricity (St. Elmo's fire). A photograph showing one of the hits is depicted in Fig. 1. In December 2011, there were three more incidences.

The continuation of the thunderstorm season up to December was caused by a very mild and humid weather pattern. Because of cool southwesterly airstream from the Baltic Sea, the air mass gained energy from the warm sea surface. Air masses were not particularly unstable, but the upper troposphere was cold. Therefore cumulonimbus (Cb) towers were comparatively thin (their tops at 6–7 km AGL). Because of the relatively warm sea surface temperatures and coastal convergence in southwesterly flow, conditions were favorable for forced lifting.

Wintertime thunderstorms, defined as lightning-producing storms occurring between October and April with a ground temperature of 0°C or below, in Finland have been described by Rinne (2009). The results show that thunderstorms are possible almost every month in the wintertime with minor maximums in November and February. Most of the winter the thunderstorms occur within the occlusion front in warm advection, with convection available potential energy (CAPE) values near zero. In almost all of the cases, the sea areas were open without ice cover.

A 30-yr climatology of thundersnow events in the contiguous United States was studied by Market et al. (2002). Their results show the typical characteristics,

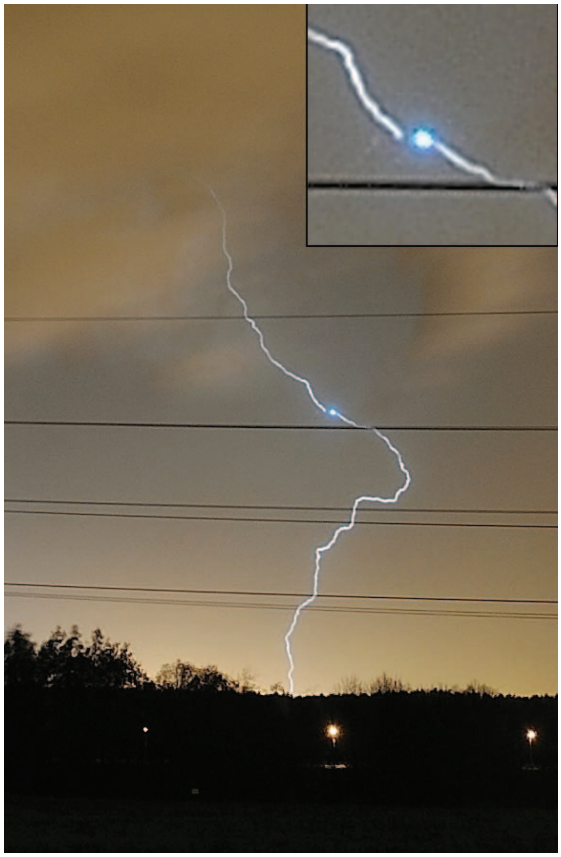


Fig. 1. Lightning strike to an airplane near Helsinki–Vantaa airport at 1736:32 UTC 19 Oct 2011. Pilot's comment: "It looked like a bucketload of sparkles had been thrown to my cockpit window." (Courtesy: Pavel Shatylovich.)

AFFILIATIONS: Mäkelä, Saltikoff, Julkunen, Juga, Gregow, and Niemelä—Finnish Meteorological Institute, Helsinki, Finland
CORRESPONDING AUTHOR: Antti Mäkelä, Finnish Meteorological Institute, P.O. Box 503, FIN-00101 Helsinki, Finland
E-mail: antti.makela@fmi.fi

The abstract for this article can be found in this issue, following the table of contents.

DOI:10.1175/BAMS-D-12-00039.1

A supplement to this article is available online (10.1175/BAMS-D-12-00039.2)

In final form 9 October 2012

including the synoptic environment, wind directions, and surface and dewpoint temperatures, of the events. Comparing their results to those of Rinne (2009) mentioned above, it seems that a thundersnow event is often a very localized and short-lived phenomenon. However, because of the relatively high temperatures (e.g., in December +3°–8°C) in the cases analyzed in this paper, our cases here cannot be qualified as winter thunderstorms (or thundersnow) but rather as *cold-season* thunderstorms.

The variability of wintertime weather conditions is large in Finland. Basically, whenever there is a wide low pressure area west of Scandinavia and a mild southwesterly flow prevailing in Finland, the conditions may be favorable for thunderstorms. This is especially true if there is an upper trough or frontal system coming from the southwest. During mild winters such situations are quite common. However, during a cold anticyclonic weather type (the "real winter conditions"), there may be long periods without any potential for cold-season thunderstorms.

The average number of thunderstorm days during the cold season (October–April) is shown in Fig. 2. Most of these storm days occur in early October, when the sea is warm. Interestingly, a maximum is found over the mainland near Helsinki; one explanation for this maximum is that in that area stands one of the highest radio masts of Finland, the Kivenlahti mast (325 m). This suggests that the mast increases the number of thunderstorm days during the cold season. Apparently, the physical explanation is that when convective clouds, not yet producing lightning, move over the Kivenlahti mast, the mast triggers a flash; this effect has resemblance to the lightning ignition by an airplane in favorable conditions.

Schultz (1999) investigated lake-effect snowstorms with and without lightning in two locations, northern Utah and western New York. The results show that the most useful parameters for forecasting lightning during lake-effect snowstorms are low-tropospheric temperatures and lifted index. The cases with lightning have substantially higher temperatures and dewpoints in the lower troposphere and lower lifted indices than the cases without lightning. The CAPE, which is often used as an indicator of warm-season thunderstorm potential, was not a useful predictor of lightning during wintertime lake-induced snowstorms.

In Finland, which is situated in northern Europe between latitudes 60° and 70°N, the thunderstorm season is highly concentrated in the summertime (May–September). Thunderstorms occur also outside this period, but their effect to the annual sum

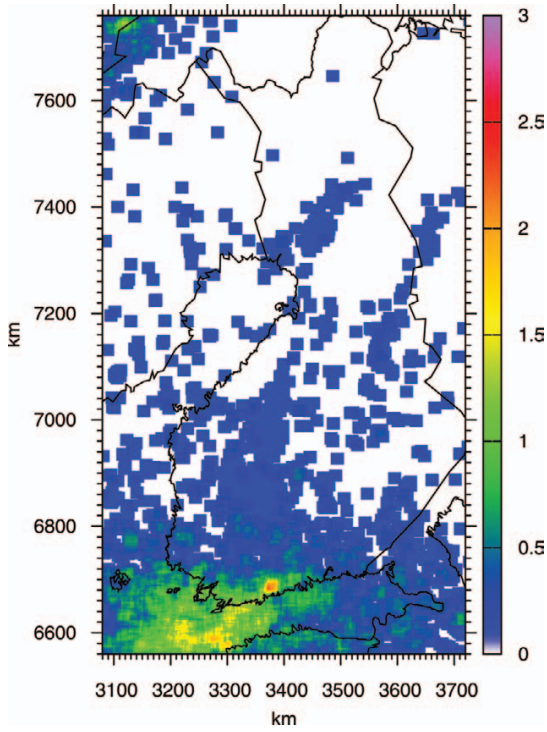


Fig. 2. Average number of cold-season thunderstorm days in 2002–11. The unit is thunderstorm days per cold season. The *x*- and *y*-axis values are kilometers to the east and north, respectively, based on the Finnish Uniform Coordinate System.

of flashes is practically negligible; however, they do affect the number of thunderstorm days, because a single flash is enough for a thunderstorm day. Cold-season thunderstorms are especially interesting because of the following two reasons:

1) Their forecasting is difficult because the warning algorithms are designed for summertime convection.
2) People are less prepared for winter lightning because of the rarity of the phenomenon.

Regarding reason 2, cold-season thunderstorms may be even more dangerous than summer storms because of their infrequent and sudden nature. The same conclusion can be found in Gough et al. (2009; summarized in Hemink et al. 2010), who showed statistics and prediction methods of cold-season thunderstorm around the Schiphol international airport in Amsterdam, the Netherlands. Their study indicates that actually most of the lightning encounters by airplanes occur during the cold season (October–

April), although the warm season (May–September) is clearly the most abundant period of thunderstorms and lightning. According to Gough et al. (2009), the situation is the same in the United States.

The motivation of this paper is to show the statistics of cold-season thunderstorms in southern Finland in 2011, which caused 13 hits to commercial airplanes near the Helsinki–Vantaa airport; 10 of the hits occurred during a single thunderstorm. The hit percentage is large, considering the overall low number of occurred strokes, about 130 per day within 50 km from the airport. Although no serious damages occurred, the incidents raise two fundamental questions:

1) What was the primary cause for the large number of the lightning-strike incidences in 2011?
2) How could have the incidences been prevented?

We examine these questions with the available meteorological observations, and with the information obtained by interviewing the pilots of the lightning-struck airplanes.

MATERIALS AND METHODS. *Lightning location system.* The Lightning Location System (LLS) of the Finnish Meteorological Institute (FMI) is part of the Nordic Lightning Information System (NORDLIS; Mäkelä et al. 2010); each participating country shares the raw sensor data from all of the NORDLIS sensors and processes the lightning data independently. The system detects primarily ground-to-ground strokes in the low-frequency (LF) domain. However, some of the located events are classified as cloud flashes, according to the peak-to-zero time of the lightning waveform (Schulz et al. 2005; Mäkelä et al. 2010).

The estimated flash detection efficiency of NORDLIS in southern Finland is above 90%, and the median location accuracy is about 500 m (Mäkelä et al. 2010). The relatively good performance and large coverage of NORDLIS is essentially due to the Nordic cooperation; without it, for example, the FMI LLS would have much smaller coverage and poorer efficiency.

Weather radar. The main weather radar serving Helsinki Vantaa airport is the Vantaa radar, located 8 km from the end of runway. It is a C-band dual-polarization radar; technical details and the measurement program are described in Saltikoff and Nevvonen (2011). Location of the radar is indicated in Fig. 3. The aviation forecaster has several radar products available, and a subset of them has been used in this study: constant level reflectivity image [constant altitude plan position indicator (CAPPI)] at 500-m altitude, maximum height of +20-dBZ threshold (TOPS; labeled as "risk of thunder" for the aviation users), vertical profile of average wind and reflectivity in 30-km cylinder around the radar [volume velocity processing (VVP)], and a range–height indicator (RHI) north and south of the radar. The TOPS product indicates the intensity of the updraft, therefore also showing the higher possibilities for the production of lightning. The TOPS product has been used for years with success in the nowcasting of thunderstorms in Finland. Hydrometeor classification based on dual-polarization parameters is shown on vertical RHI and conical plain position indicator (PPI) surfaces.

Mesoscale analysis system. FMI operates the Local Analysis and Prediction System (LAPS; http://laps.fsl.noaa.gov/; Albers et al. 1996; Toth et al. 2011) for production of 3D analysis fields of different weather parameters. Within LAPS observations are fitted to the coarser first-guess background field from the global numerical weather prediction model of the European Centre for Medium-Range Weather Forecasts (ECMWF) by using mainly a multiscale successive correction method, and high-resolution topographical datasets are taken into account while creating the final high-resolution analysis fields.

Pilot interview background. A few days after the incidences on 19 October, FMI was in contact with the airlines and pilots of the lightning-struck airplanes. Discussions revealed that the case was indeed extraordinary and that it should be investigated further. Also, many of the pilots informed the authors of their interest to receive more information regarding the synoptic situation in order to be better prepared if something similar happens in the future. Therefore, it was decided that the FMI would collect information and feedback via a questionnaire from the pilots. A total of 6 replies out of 10 were received. The questions and answers are shown in the "Results" section.

Numerical weather prediction (NWP). FMI uses two operational limited-area NWP models. High-Resolution Limited-Area Model (HIRLAM; Undén et al. 2002) is a hydrostatic primitive equation model covering all of Europe with a 16.5-km grid size. In addition to traditional larger-scale models, FMI operates the nonhydrostatic mesoscale NWP model HIRLAM–Aire Limitée Adaptation Dynamique Développement International (ALADIN) Research on

Mesoscale Operational NWP in Euro–Mediterranean Partnership (HARMONIE). This model can provide high-resolution precipitation data, both in space (2.5 km) and time (hourly or less), and it has a more detailed description of precipitation physics (Seity et al. 2011) than previous NWPs, empowering better simulations of heavy rainfall episodes (Niemelä 2009; Bengtsson and Niemelä 2008; Niemelä et al. 2007). Furthermore, HIRLAM is designed to be used in scales where all convective flow structures need to be parameterized (Kain and Fritch 1990; Kain 2004). In HARMONIE the deep convective flow structures are assumed to be resolved explicitly (kilometer scale), leaving the parameterization problem only for nonprecipitating shallow convection (Siebesma et al. 2007).

Other. Upper-air soundings from Jokioinen, Finland, were used to analyze low-tropospheric winds and shear, and to derive the vertical temperature differences and different stability indices in the studied cases. The sounding closest to the event in time was used. The Jokioinen observation and sounding station is located about 100 km northwest of the Helsinki–Vantaa airport (Fig. 3).

RESULTS. In this section we first show the general synoptic scenario of all four cases of 2011: 19 October, and 4, 14, and 26 December. Then we concentrate on the case of 19 October, which is the most interesting regarding the influence it had on the Helsinki–Vantaa airport.

Synoptic situation of cold-season thunderstorms in 2011. Table 1 shows information about four thunderstorm events that occurred late 2011. The first one (19 October 2011) was the most active, with lightning activity lasting for several hours during the late afternoon and evening along a southwest–northeast-oriented area passing the Helsinki–Vantaa airport (Fig. 3). The synoptic setting of this case is described more thoroughly in the section "Case 19 October 2011."

The first three cases (19 October, and 4 and 14 December) were quite similar with a wide low pressure area west of Finland and a southwesterly airstream prevailing in southern Finland. The upper-air lapse rates show surprisingly similar values, for example, the temperature difference between surface and 700 hPa being 17°–18°C. These values are equivalent to those found by Schultz (1999) during wintertime thunderstorm events in western New York. The lifted index (LI) and total totals index (TOTL) for the three cases are quite similar, indicating high probability of showers and thunder (especially the TOTL values); the CAPE values were very low, which is also in good agreement with the results by Schultz (1999). The temporal evolution of the TOTL index and the vertical wind profile for two of the cases according to LAPS are shown in Fig. 4. The left panel contains all the cases; however, it is complemented with a nonthundery case from 27 December 2011, when organized convection was formed but without lightning. The cases seem to be organized into two groups, the one showing potential risk for thunderstorms (TOTL values above 50) and the other showing slightly lower values.

The fourth case, 26 December, was somewhat different: the lapse rates and stability indices show less instability than in the three previous cases. In this

Fig. 3. (left) Finland and the surrounding areas and a (right) zoom-in image showing the position of the Helsinki–Vantaa airport, Vantaa weather radar, and a 25-km-radius range circle around it, the Jokioinen atmospheric sounding station, and the located lightning (crosses) on 19 Oct 2011.

case, the thunderstorm was probably associated with the passage of a trough, preceded by warm advection and very strong vertical wind shear in the lower troposphere, which can also be seen in the analyzed wind profile from LAPS (Fig. 4). Based on the 0000 UTC 26 December sounding from Jokioinen, the surface and850-hPa wind directions and speeds were 210° at 9 m s^{-1} and 240°at 34 m s^{-1}, respectively. This situation resulted in stormy wind gusts at the surface later in the day, causing a lot of forest damage and electricity cuts. The common feature in all four cases is that a moderate or brisk southwesterly wind prevailed (Table 1, winds at Helsinki–Vantaa airport), advecting heat and moisture from the relatively warm Gulf of Finland toward the inland, thus promoting convective development.

Table 1. Four convective events with lightning in southern Finland during late autumn and early winter in 2011. (a) Data from Jokioinen upper-air sounding station (source: the University of Wyoming), (b) surface observations at Helsinki–Vantaa airport (complemented with the most severe METAR report), (c) detected lightning strokes, and (d) Vantaa radar data within a 100-km radius: maximum height of +20-dBZ isoline ("thunderstorm risk indicator") and the largest reflectivity (dBZ).

Case	1200 UTC 19 Oct 2011	1200 UTC 4 Dec 2011	1200 UTC 14 Dec 2011	0000 UTC 26 Dec 2011
(a) Sounding data from Jokioinen				
P_s (hPa)	995	975	996	986
$Wind_s$ [direction (°)/speed (m s^{-1})]	210/8	210/5	200/5	210/9
$Wind_{850}$ [direction (°)/speed (m s^{-1})]	225/19	230/16	215/20	240/34
$Wind_{700}$ [direction (°)/speed (m s^{-1})]	225/16	230/16	215/18	260/34
T_s (°C)	7.3	4.0	3.6	5.0
T_{700} (°C)	−10.7	−13.3	−14.3	−5.7
T_{500} (°C)	−30.9	−34.3	−34.9	−22.5
$\Delta T_{s\text{-}700}$ (°C)	18.0	17.3	17.9	10.7
$\Delta T_{700\text{-}500}$ (°C)	20.2	21.0	20.6	16.8
$\Delta T_{s\text{-}500}$ (°C)	38.2	38.3	38.5	27.5
KI (°C)	27.30	17.30	9.90	18.30
LI (°C)	0.40	−0.30	0.55	9.62
TOTL (°C)	59.20	58.60	56.80	43.40
CAPE (J kg^{-1})	10.83	26.87	0.47	0
(b) Surface observations at Helsinki–Vantaa airport	1750Z 23012G24KT 200V260 6000 TSGS SCT014 BKN022 BKN030CB	041850Z 19011KT 9999-SHRA FEW008 SCT013 SCT025CB BKN030 04/03	141550Z 18013KT 9999-SHRA FEW010 SCT013 FEW030CB	260250Z 23026G38KT 9999-SHRA FEW013 BKN023 FEW030CB 08/06
T_s (°C)	10.1	3.3	5.3	4.6
T_d (°C)	4.9	2.3	3.5	3.5
Wind direction (°)	200	210	200	210
Wind speed (m s^{-1})	11	5	8	13
(c) Observed lightning strikes within 50 km of Helsinki–Vantaa airport	1500–1900 UTC	1700–1900 UTC	1400–1600 UTC	0200–0400 UTC
No. of lightning strokes within 50 km/ whole of Finland	132/691	3/3	6/30	12/20
(d) Radar data (UTC)	1500–1900	1700–1900	1400–1600	0200–0400
Maximum height of +20 dBZ (km)	8.0	7.5	6.0	5.0
Largest reflectivity (dBZ)	55	51	51	59

Case 19 October 2011. The 1800 UTC synoptic weather map and weather radar CAPPI image from southern Finland on 19 October are shown in Fig. 5. A wide low pressure area was located over Scandinavia, moving slowly northeast. A quite-strong southwesterly airstream prevailed in Finland, surface winds being around 10 m s^{-1} at the southwestern coastal areas and the 850/700-hPa winds being almost 20 m s^{-1} based on the Jokioinen sounding (Table 1) and LAPS analysis (Fig. 4). The incoming air mass had a long fetch over the relatively warm water (10°–13°C) of the Baltic Sea, resulting in a couple of degrees higher surface air temperatures in Finland's southwestern coast than farther inland. At upper levels, a tongue of cold air had pushed in over southern Scandinavia and western Finland. Based on the Jokioinen sounding, the temperature at the 500-hPa level was about –31°C, resulting in a vertical temperature difference of 38°C between the surface and 500-hPa level (Table 1). The relative humidity was high (80%–100%) up to the 650-hPa level and above that level the air was drier. Circumstances were quite favorable for convective development, which was indicated, for example, by the high TOTL index value, 59.2, based on the Jokioinen sounding and LAPS analysis. During the day, well-organized

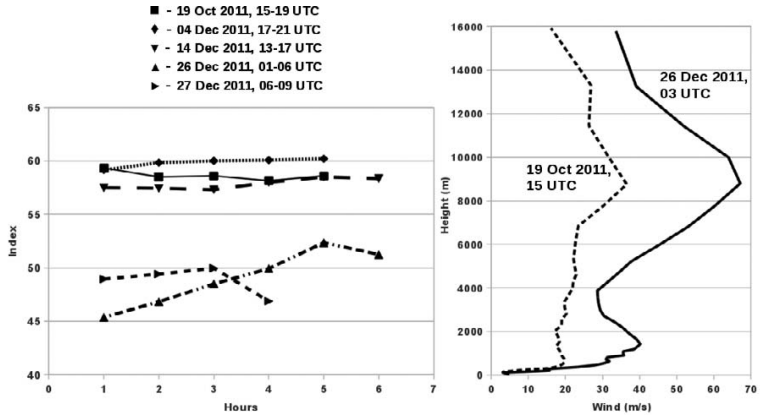


Fig. 4. (left) Stability index (total totals) at Vantaa airport. Values above 50 indicate risk of moderate–severe thunderstorm. (right) LAPS wind profiles at Vantaa airport. Dashed and solid lines show profiles for 1500 UTC 19 Oct 2011and 0300 UTC 26 Dec 2011, respectively.

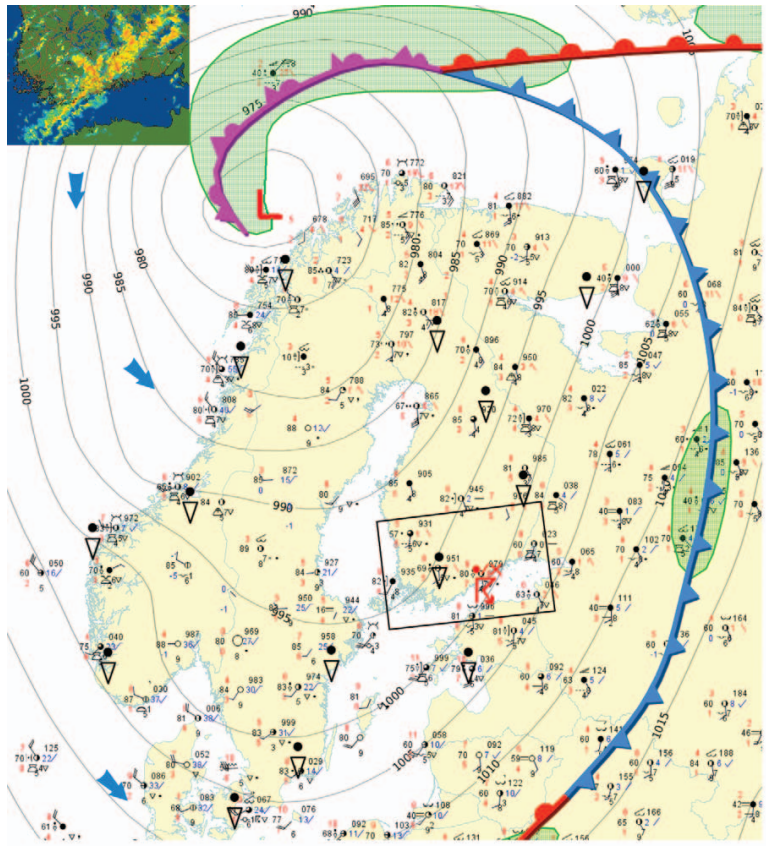


Fig. 5. Synoptic situation in northern Europe at 1800 UTC 19 Oct 2011 (analysis by FMI). (top left) Weather radar inset from southern Finland (rectangle) at 1800 UTC.

southwest–northeast-oriented convective lines formed over the western Gulf of Finland and at the coast west of the city of Helsinki. The individual convective cells moved to the northeast by the midtropospheric flow, while the area of heaviest convection moved very slowly east during the evening. (An animation is available as supplemental material online at http://dx.doi.org/10.1175/BAMS-D-12-00039.2.)

In the top panel of Fig. 6, we see the isolines of radar reflectivity averaged in a 30-km cylinder around the Vantaa radar; the bottom panel indicates the 15-min lightning rate within 25 km from the airport. The 20-dBZ isoline corresponds to a product used for thunderstorm warning, and it rises before a thunderstorm is reported in the aviation routine weather reports (METARs). Similar behavior was observed in the other cases, too. The lightning data show three peaks, at about 1415, 1700, and 1900 UTC. It is highly possible that some of the reported lightning strikes to the airplanes (black stars in the bottom panel) were actually triggered by the plane itself, like the flash (two strokes) at 1530–1545 UTC.

The hydrostatic model HIRLAM and the nonhydrostatic mesoscale model HARMONIE show a large difference in their analysis (Fig. 7). HIRLAM indicates only some cloud water in the lower levels and is not able to predict well the vertical structure of the storm, while HARMONIE seems to capture it largely in the same way as the weather radar sees it (Fig. 7).

Interestingly, a Finnish storm chaser, Pavel Shatylovich, took a photograph of one of the airplanes being hit (Fig. 1). The time of the photograph matches to a flash located at 1736:32 UTC. This single-stroke negative polarity flash has an estimated peak current of 7.4 kA (i.e., relatively low peak current). The airplane appears in the photograph as a bright spot in the middle of the lightning channel.

Table 2 summarizes the replies received from the pilots of the lightning-struck airplanes on 19 October. We have filtered the replies to some extent to save space.

The main findings of Table 2 are the following:

- Most of the pilots had experienced lightning hit before.
- During the hit, practically all of the hit planes were approximately at the cloud-base height inside the cloud. One pilot reported the plane to be about 2 km away from the cloud.
- Typically, a bright light was seen and a loud bang was heard. One pilot reported the whole windshield to be illuminated by St. Elmo's fire.
- Prior to the hit, interference was observed in the radio (apparently due to the electricity of the cloud).
- The forecast was not specific enough about the risk of thunderstorms of this magnitude.
- More rapid response and warnings from the traffic control and airlines are encouraged.
- Real-time lightning location data available at the cockpit would be useful.

DISCUSSION. We return to the questions raised earlier. First, what was the primary cause for the large number of the lightning-strike incidences in 2011? The answer is a combination of four ingredients:

1) *The exceptionally warm early winter.* This makes the occurrence of thunderstorms possible, but their forecasting (especially the magnitude) highly difficult.
2) *The direction of the wind at the Helsinki–Vantaa airport during a convective weather type.* The direction of wind dictates the runway to be used. Especially in the 19 October case, many of the planes took off right into the core of the storm. In the worst case, three consecutive departing planes got hit by lightning.
3) *The lack of information.* The convective situation of this magnitude in the cold season at high latitudes surprised both air traffic control and the pilots.

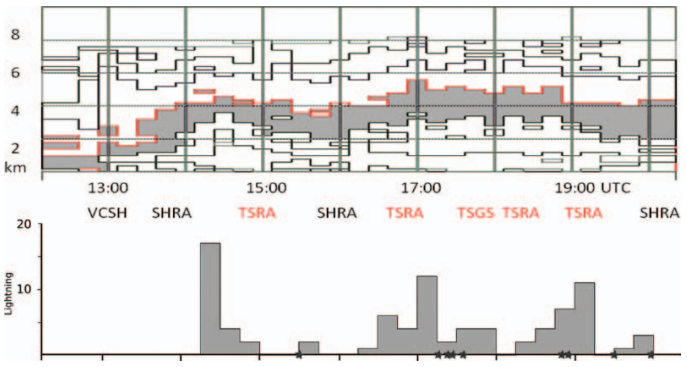


Fig. 6. (top) Average reflectivity in 30-km cylinder around Vantaa radar 19 Oct and isolines at intervals of 5 dBZ. Height of +20 dBZ (red) is used as thunder indicator, when it reaches 6 km in summer. Interval +15–20 dBZ is shaded. (middle) Prevailing weather at the airport: SHRA—rain showers, TSRA—thunderstorm, TSGS—with graupel. (bottom) Lightning flash rate (strokes per 15 min). Stars indicate times when planes were hit.

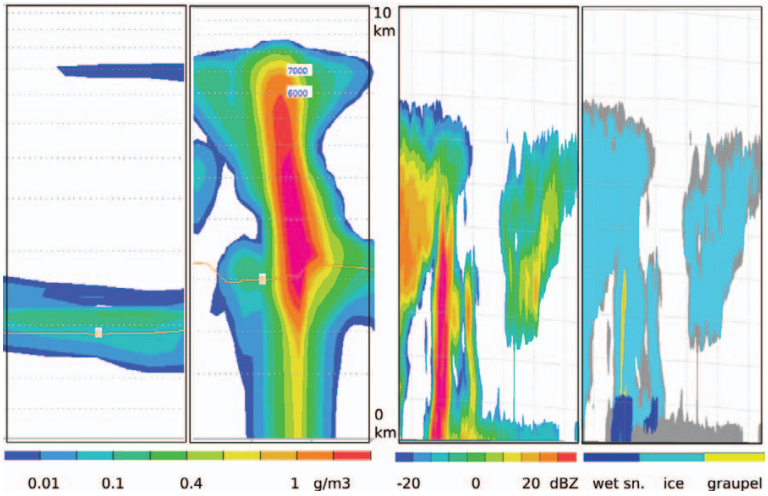


Fig. 7. The most intensive cell around 1830 UTC. (from the left) Cloud condensate in HIRLAM, clouds and hydrometeors in HARMONIE, radar reflectivity, and hydrometeor classification in radar (gray values indicate echoes classified as nonmeteorological). For each panel, width is 60 km and height is 10 km. The vertical axis is linear in radar images, but it is different for the two models because of the number of model levels in the boundary layer (13 and 20 model levels in the lowest 1,000 m, respectively).

4) *The lack of air traffic control procedures*. It seems there are no exact procedures on how to react in similar situations. The options are to delay the flights until the convective situation is over or to stay on the schedule and take the risk.

Second, how could have the incidences been prevented? One of the pilots stated ironically that this would have been easy by staying at home. Actually, the statement is quite accurate: it may well be that similar situations will occur anyway, because a rare cold-season thunderstorm is difficult to predict, and the delaying of flights is not economically profitable to the airlines. However, regarding small and isolated cold-season thunderstorms, which are the majority, for example, in the Amsterdam Schiphol Airport (Gough et al. 2009; Hemink et al. 2010), certain avoidance procedures can be successfully used. The problem with smaller cells is that usually the first flash is actually the one triggered by an airplane; this means that the thunderstorm area cannot be monitored prior to the hit with, for example, a lightning location system.

According to the analyzed data, we believe that the extent of similar situations can be highly reduced by training, better nowcasting tools, and with the aid of common procedures at the airports. Regarding the nowcasting tools based on weather radar, the height of +20-dBZ isoline (see the section "Weather radar") was the best at indicating the active thunderstorms and separating thundery and nonthundery cases. Hydrometeor classification showed graupel associated with the thunderstorms, typically forming narrow and tall vertical pillars. As these pillars did not always extend to the lowest weather radar measurement, and as their diameter is typically only a few radar pixels, finding them in the standard images is a challenge. The lightning–ignition process is highly increased when an airplane flies into the convective core. When the first flash of the storm has occurred, a real-time lightning location system may give important information for pilots, forecasters, and air traffic control.

The operational models showed remarkable differences. The high-resolution HARMONIE model was superior at predicting the convective nature of the event compared to the coarser-resolution HIRLAM. HIRLAM was able to represent only the low-level cloud cover because of to its lower resolution and parameterized deep convection. On the contrary, high-resolution HARMONIE can simulate more realistic deep convective structures with high and narrow cloud/precipitation band. In this case, the mesoscale model with explicit treatment of deep convection and five-species prognostic microphysics parameterization clearly outperforms HIRLAM.

Stability indices such as the *K* index (KI), LI, and TOTL can provide useful information, from either upper-air sounding or analysis models, such as the LAPS system.

Finally, we note that it is possible that these kinds of cold-season weather phenomena will be more probable in the future if similar weather types occur more often because of climate change.

CONCLUSIONS. Cold-season thunderstorms are rare events, which increase their threat to aviation safety; both pilots and forecasters can be surprised when they occur. Furthermore, especially at high

Table 2. Pilot questionnaire. (TCAS stands for traffic collision avoidance system.)

	Pilot 1	Pilot 2	Pilot 3	Pilot 4	Pilot 5	Pilot 6
Aircraft type	A330 or A340	A340	E170	A320	E170	E190
How many times have been hit by lightning before?	~10	2	~10	~5	None	None
Depart/arrive	Departure	Departure	Arrival	Arrival	Arrival	Departure
Inside/outside the cloud?	Inside but not in the middle of the Cb cores	Outside, ~2 km south of the cloud, at cloud-base height	Inside, at the height of the cloud base	Inside but aside from the most intense cores; cloud-base height	Inside	Inside at the height of ~1.5–2 km
Distance to the Cb?	Hard to say. The flash was not from the most intense cell.	~2 km	—	~1 n mi	—	—
Other observations?	The flash was clearly visible before the hit. It probably hit our right wing. A weak sound was heard, but not by passengers	A heavy bang and a bright flash of light	Intense downdraft	St. Elmo's fire in the windows. About 10 cm long, 1 cm thick. Purple discharge at the icing sensor	Heavy blast immediately after a bright flash	Like a branch of threads of light to the left engine and to the cockpit. It resembled like someone had thrown a bucket full of big sparks to the wind shield.
Technical problems?	Nothing	TCAS unserviceable after the hit.	No problems. The radar in the plane is not the best.	Nothing	Nothing	Some failure signals were displayed. Nothing serious.
Did you receive information about the previous hits?	No	No. We saw the previous departed plane being hit.	No	No	No. Prior to the hit, we heard some radio interference typical for Cb's.	No. There were three planes departing, we were the first. All three got hit and had to land for inspection.
Ideas for improvements?	More rapid response and communication between the traffic control and pilots.	To emphasize the meaning of similar weather scenarios at trainings.	Training and better information regarding cold-season storms.	Because of previous hits to other airplanes, traffic control or airline could give some warning.	If flashes have been observed, it would be wise to inform about it.	A lot of radio interference by electricity and plenty of lightning. I doubt if anything could have been done for us. Later on, the arriving planes were warned about lightning.
Forecast quality?	Normal	Forecast did not mention the risk of thunder-storms.	Misleading	Thunderstorms were not mentioned in the forecast.	Nothing that would have helped to be cautious about Cbs.	The possibility of thunderstorms was not emphasized enough.
Would it help to have lightning location data at the plane?	No	Yes	Yes	Possibly yes	Not necessarily	Yes, it would give more details for making decisions.

latitudes, these thunderstorms occur often in the dark, which adds to their physiological effect on the pilots.

Some of the forecasting tools used in the warm season can also be used in the cold season. New and improved tools such as nonhydrostatic models and dual polarization radars have additional value in these cases. Three-dimensional radar data are useful in thunderstorm analysis even in the cold season. Besides forecasters, pilots also need training. The Lightning Location System is an excellent tool for pinpointing in real time where, when, and to what rate lightning is occurring.

Based on the analyzed cases, the most useful indicators of cold-season lightning in southern Finland are the following: the vertical temperature difference between the surface and midtroposphere (700/500 hPa); low-tropospheric wind shear; southwesterly flow (impact of warm sea water); and some convective indices, such as the "total totals" index and the lifted index. CAPE was not a useful predictor here.

Forecasters can have more realistic information on the characteristics of cold-season convective episodes from high-resolution mesoscale NWP than from coarser-resolution models (e.g., global models).

REFERENCES

AAIB, 1999: Schleicher ASK 21 two seat glider. AAIB Bull. 12/99, Reference EW/C99/04/02, 15 pp. [Available online at www.aaib.gov.uk/sites/aaib/cms_resources/dft_avsafety_pdf_500699.pdf.]

Albers, S., J. McGinley, D. Birkenheuer, and J. Smart, 1996: The Local Analysis and Prediction System (LAPS): Analyses of clouds, precipitation, and temperature. *Wea. Forecasting,* **11,** 273–287.

Bengtsson, L., and S. Niemelä, 2008: Nested AROME simulations of convective precipitation. *HIRLAM Newsletter,* No. 54, HIRLAM, De Bilt, Netherlands, 37–47. [Available online http://hirlam.org/index.php?option=com_docman&Itemid=70.]

Cherington, M., and K. Mathys, 1995: Deaths and injuries as a result of lightning strikes to aircraft. *Aviat. Space Environ. Med.,* **66,** 687–689.

Civil Aeronautics Board, 1965: Pan American World Airways, Inc., Boeing 707-121, N709PA, near Elkton, Maryland, December 8, 1963. Aircraft Accident Report File 1-0015, 22 pp.

Clifford, D. W., and H. W. Kasemir, 1982: Triggered lightning. *IEEE Trans. Electromagn. Compat.,* **24,** 112–122.

Gough, W. R., J. Hemink, S. Niemeijer, and T. H. Fahey, 2009: The prediction and occurrence of aircraft lightning encounters at Amsterdam–Schiphol Airport. *Proc. 47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition,* Orlando, Florida, AIAA, ASE-5. [Available online at http://arc.aiaa.org/doi/abs/10.2514/6.2009-553.]

Hemink, H. J., S. Niemeijer, W. R. Gough, and T. H. Fahey, 2010: Prediction and occurrence of aircraft lightning encounters around Amsterdam Airport Schiphol. KNMI Research Triennial Scientific Rep. 2007–2009, 96–100. [Available online at www.knmi.nl/research/biennial/07-09_H16.pdf.]

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.,* **43,** 170–181.

——, and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.,* **47,** 2784–2802.

Mäkelä, A., T. J. Tuomi, and J. Haapalainen, 2010: A decade of high-latitude lightning location: Effects of the evolving location network in Finland. *J. Geophys. Res.,* **115**, D21124, doi:10.1029/2009JD012183.

Market, P. S., C. E. Halcomb, and R. L. Ebert, 2002: A climatology of thundersnow events over the contiguous United States. *Wea. Forecasting,* **17,** 1290–1295.

Mazur, V., 1989: A physical model of lightning initiation on aircraft in thunderstorms. *J. Geophys. Res.,* **94** (D3), 3326–3340.

Moreau, J.-P., J.-C. Alliot, and V. Mazur, 1992: Aircraft lightning initiation and interception from in situ electric measurements and fast video observations. *J. Geophys. Res.,* **97** (D14), 15 903–15 912.

National Transportation Safety Board, 1976: Special investigation report: Wing failure of Boeing 747-131, near Madrid, Spain, May 9, 1976. Bureau of Technology, 38 pp.

Niemelä, S., 2009: Verification of high-resolution precipitation forecasts by using the SAL method. *HIRLAM Newsletter,* No. 55B, HIRLAM, De Bilt, Netherlands, 41–47. [Available online at http://hirlam.org/index.php?option=com_docman&Itemid=70.]

——, U. Andrae, B. Sass, and B. Anderssen, 2007: Convective precipitation in AROME. *HIRLAM Newsletter,* No. 52, HIRLAM, De Bilt, Netherlands, 33–38. [Available online at http://hirlam.org/index.php?option=com_docman&Itemid=70.]

Plumer, J. A., and J. D. Robb, 1982: The direct effects of lightning on aircraft. *IEEE Trans. Electromagn. Compat.,* **24,** 158–172.

Rakov, V. A., and M. A. Uman, 2005: *Lightning: Physics and Effects.* Cambridge University Press, 687 pp.

Rinne, J., 2009: *Talviukkoset Suomessa 2003–2007* (Winter thunderstorm in Finland 2003–2007). M.S. thesis, Dept. of Physics, University of Helsinki, 72 pp.

Saltikoff, E., and L. Nevvonen, 2011: First experiences of the operational use of a dual-polarisation weather radar in Finland. *Meteor. Z.,* **20,** 323–333.

Schultz, D. M., 1999: Lake-effect snowstorms in northern Utah and western New York with and without lightning. *Wea. Forecasting,* **14,** 1023–1031.

Schulz, W., K. Cummins, G. Diendorfer, and M. Dorninger, 2005: Cloud-to-ground lightning in Austria: A 10-year study using data from a lightning

location system. *J. Geophys. Res.,* **110,** D09101, doi:10.1029/2004JD005332.

Seity, Y., P. Brousseau, S. Malardel, G. Hello, P. Bénard, F. Bouttier, C. Lac, and V. Masson, 2011: The AROME-France convective-scale operational model. Mon. *Wea. Rev.,* **139, 3,** 976–991.

Siebesma, A. P., P. M. M. Soares, and J. Teixeira, 2007: A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *J. Atmos. Sci.,* **64,** 1230–1248.

Toth, Z., S. C. Albers, and Y. Xie, 2011: Analysis of fine-scale weather phenomena. *Bull. Amer. Meteor. Soc.,* **93,** ES35–ES38.

Uman, M. A., and V. A. Rakov, 2003: The interaction of lightning with airborne vehicles. *Prog. Aerosp. Sci.,* **39,** 61–81.

Undén, P., and Coauthors, 2002: HIRLAM-5 scientific documentation. HIRLAM-5 project, SMHI, 144 pp. [Available online at http://hirlam.org/index.php?option=com_content&view=article&id=64&Itemid=101.]

FINNISH METEOROLOGICAL INSTITUTE
Erik Palménin aukio 1
P.O. Box 503
FI-00101 HELSINKI
tel. +358 29 539 1000
WWW.FMI.FI

FINNISH METEOROLOGICAL INSTITUTE
CONTRIBUTIONS No. 142
ISBN 978-952-336-042-6 (paperback)
ISSN 0782-6117
Erweko
Helsinki 2018

ISBN 978-952-336-043-3 (pdf)
Helsinki 2018