UNIVERSITAT DE BARCELONA

Precipitation phase discrimination: diagnosing and nowcasting

Enric Casellas Masana

Enric Casellas Masana

Precipitation phase discrimination

Diagnosing and Nowcasting

Ph.D. Thesis submitted for the degree of
Doctor of Philosphy in Physics

UNIVERSITAT DE BARCELONA

University of Barcelona
Faculty of Physics
Department of Applied Physics-Meteorology

Universitat de Barcelona
Facultat de Física
Departament de Física Aplicada-Meteorologia

Precipitation phase discrimination

Diagnosing and nowcasting

Memòria presentada per optar al grau de doctor per la
Universitat de Barcelona

Programa de doctorat en Física

Autor: Enric Casellas Masana

Director: Joan Bech i Rustullet

Tutor: Joan Bech i Rustullet

Barcelona, gener 2022

Disseny de la portada inspirat en les edicions de l'editorial Periscopi.

Agraïments

Em toca escriure els agraïments d'aquesta tesi i no sé gaire per on començar ni què dir-hi. Mirant les tesis d'altres companys, totes comencen amb un agraïment al director de tesi. En el meu cas és en Joan Bech, i penso què dir-ne. Podria començar dient que em va dirigir el treball final de grau quan el vaig anar a buscar encaparrat que volia fer-lo sobre meteorologia. També podria afegir que em va dirigir el treball final de màster, però potser no cal tirar tan enrere. Hauria de centrar-me en aquesta etapa i dir que per a mi ha estat un dels pilars fonamentals d'aquesta tesi, agrair-li la confiança, el suport i l'empenta. Que qualsevol reunió es convertia en una injecció de moral i dir allò tan típic que no hauria pogut tenir un millor director i que en aquest cas és ben cert. Però clar, a part del director de tesi també he tingut dos supervisors al Servei Meteorològic de Catalunya, en Roger Veciana i en Josep Ramon Miró, i penso que també n'hauria de parlar al començament. D'en Roger hauria de començar dient que bona part del que he après de programació li dec a ell? O bé hauria d'encarar-ho des d'un vessant més personal i dir que m'ha contagiat la seva empenta, que m'ha ajudat i orientat sempre que ho he necessitat i que he tingut sempre la seva confiança? I d'en JR què n'hauria de dir? Començar parlant d'aquests darrers mesos en què s'ha convertit en una persona important en el meu dia a dia a la feina? O centrar-ho en el seu paper durant la tesi i dir que m'ha permès aprendre molt al seu costat i que a la vegada és capaç de fer-te les preguntes adients per continuar endavant i millorar? També podria afegir per tancar aquesta part que a casa s'ha convertit en el proveïdor oficial d'oli d'oliva, però potser això ja distaria una mica del que caldria posar en uns agraïments. Potser hauria de tancar el primer paràgraf parlant de l'Abdel Sairouni i agrair-li la confiança i el suport que m'ha mostrat sempre des del primer dia.

En aquest punt penso que hauria de fer referència al Francesc Gallart, al Jérôme Latron i a la Pilar Llorens per donar-me l'oportunitat de fer pràctiques al seu grup de l'IDAEA, on també vaig poder conèixer en Carles Cayuela, tot un exemple a seguir, i

endinsar-me en el món de la recerca. Estic convençut que sense aquella estada avui no estaria pensant què posar en uns agraïments d'una tesi. Però bé, potser això ho hauria de posar després de comentar els dos llocs on he fet el doctorat, el Servei Meteorològic de Catalunya i la Facultat de Física de la Universitat de Barcelona. Podria començar dient que ha estat un honor poder fer una part de la tesi a l'SMC i després agrair a tots els companys l'acollida, i potser fer menció d'aquells amb qui he compartit més temps. Seria millor començar pels de la cinquena planta o bé pels de la quarta? A la cinquena és on treballava i on he compartit taula amb el Sergio Castillo, el Francesc Roura i l'Anna Soler. On també he pogut conèixer al Nico, la Patri, el Tomeu, la Helen, l'Oriol, i després en Ferran, els quals m'han ajudat sempre que ho he necessitat i amb qui he compartit molts esmorzars. Però clar, a la cinquena planta també hi ha l'Equip de Predicció i també els hauria d'agrair l'acollida, en especial al Santi, a la Montse i al Sergio. A la quarta és on hi ha tota la gent de RAM, que també em van acollir i ajudar en el que calgués, i podria posar que en especial els Jordis, i fer una mica de broma dient que no són pas els activistes, sinó el Jordi Mercader i el Jordi Moré, però bé, potser una mica agafada amb pinces. A la quarta planta també hi ha la gent de SOM, sense els quals no hauria disposat de bona part de les dades d'aquesta tesi i fer una menció especial a la Xènia i al Ricard. Ara hauria de buscar una manera d'enllaçar amb la facultat i crec que ho podria fer a partir dels esmorzars. Dir que els divendres a l'SMC hi havia esmorzar, podríem dir especial, i llavors enllaçar-ho amb els super esmorzars de la Facultat i mencionar els increïbles pastissos del Jordi i la resta de postres i dolços que anaven portant tots els companys del DAM-Nyam. I així podria començar a agrair a tots els companys de departament la seva acollida, les bones estones a la saleta, que tinguessis el dia que tinguessis sempre t'animaven. I acabar fent menció de la Mireia, la Yolanda, el Jordi, el Joan, el Gabi, el Dani, la Chloé, el Richard, l'Àngela, el Javi, la Froila i la Marina. I llavors crec que així tancaria el segon paràgraf.

Al final veig que és un costum acabar amb els agraïments de caire més personal. I aquí em venen uns quants noms a qui els he d'agrair moltes coses. Potser podria començar pels *MeteoFreaks*, l'Anna, la Maria i l'Oriol, i dir que han estat uns referents per a mi i que amb qui compartir etapa de doctorand ha fet que la tesi sigui molt més portable. Per totes les estones i risses™ compartides. I em ve al cap que amb l'Anna hem compartit un congrés, el meu primer de fet, on vaig poder conèixer en Wael, que després va venir a Barcelona i amb qui vaig compartir despatx un estiu i una sortida de camp a la Cerdanya. Dit això podria seguir amb els amics de la carrera i dir que malgrat els anys, no ens hem deixat de veure. Tot i estar mig dispersats per Europa, aquesta Europa que ens mira, dir que els del *Tots entre tots* s'han interessat sempre per com m'anava tot.

Ara que hi penso, n'hi ha prou amb posar el nom del grup de WhatsApp o bé hauria d'anomenar la Maria, el Dani, la Jordina, la Judit, el Carles, el Lluís i la Saskia? De qui sí que n'hauria de fer menció especial és de la Irene, l'Elena i l'Andreu. Caldria posar-me intensito i dir que són un refugi per infinitat de coses que m'han passat i que espero que ho continuïn sent, que hi hagi més *Churchs*™? O bé hauria de simplement agrair-los el suport i el seu interès en la meva tesi? No ho sé, ja veuré. I crec que per tancar aquesta part i abans d'anar a la família, hauria de parlar d'aquella família que triem, dels amics de Calafell, els de tota la vida, els que hi han sigut sempre, en tots els casos, allò de les verdes i a les madures. I llavors dir que són el Max, el Roberto, l'Isma, el Fèlix, la Gwen, la Laia, la Clàudia i en especial el Victor, l'Arnau, el Jose i l'Omar.

I per tancar, hauria de fer menció a la família, als tiets i cosins, en especial a l'Anna amb qui anem compartint lectures i a l'Adrià que sempre ha estat una mica referent, que sempre s'han interessat per com m'anava, malgrat que a vegades no tingués gaires ganes de parlar-ne o bé no sabia explicar gaire bé què és el que feia. També hauria de fer menció a mun àvia, que malgrat haver d'explicar sovint en què consisteix la tesi amb no gaire sort, tot s'ha de dir, em dona suport en tot. I si bé al començament he dit que no sabia gaire què dir, aquí sí que ho tinc clar. Agrair a les persones més importants per a mi, els meus pares, tota la seva confiança i estima. Sense vosaltres no seria com soc ni hauria arribat on he arribat. I buscant alguna cosa per tancar definitivament aquests agraïments potser ho podria fer amb una frase d'Oques Grasses i dir que mentre fas la tesi *hi ha mil dies bons i dies de merda*.

I arribats a la tercera pàgina d'agraïments, potser tocaria deixar-ho aquí. Que en el procés de pensar què és allò que hauria d'incloure i com podria dir-ho, ja ho hauria dit tot. I fins aquí la part interessant de la tesi, tot el que ve a partir d'ara ja no ho és tant. I per últim, agrair a tothom qui hi ha estat, vingut, marxat o tornat i s'ha interessat per com m'anava tot, malgrat sempre respongués que nem fent.

Enric Casellas Masana
Universitat de Barcelona, 18 de gener de 2022

Abstract

Precipitation phase discrimination at ground level constitutes a fundamental variable in many meteorological and hydrological applications, including avalanche hazards, winter road safety, and flooding from rain on snow events. Discrimination of the precipitation phase at surface level has been widely studied following different approaches ranging from decision tree algorithms based on vertical temperature profiles parameters, to machine learning algorithms through schemes relying on microphysical parameterisations. However, precipitation phase discrimination is still challenging, specially at temperature close to freezing point. Several studies pointed out research gaps regarding this topic and the present thesis aims to make its small contribution to some of them. In addition, this thesis comes from the need to provide the Meteorological Service of Catalonia with an adjusted and verified precipitation phase discrimination product for diagnosing and nowcasting purposes. In order to achieve both kind of requirements six specific objectives were set and upon which this thesis was structured. These are the following:

- SO1. Obtention of a dynamic interpolation scheme suitable for complex terrain, and high spatial and temporal resolution.
- SO2. Evaluation and adjustment of different schemes and meteorological variables to diagnose discrimination of the surface precipitation phase.
- SO3. Assessment of citizen science and crowd sourced observations for monitoring snow events.
- SO4. Development and evaluation of different schemes to nowcast discrimination of the precipitation phase.
- SO5. Evaluation of ensemble techniques to nowcast discrimination of the precipitation phase.

- SO6. Implementation of a precipitation phase product in an operational chain.

The present thesis is based in a compendium of three scientific publications and three major blocks were defined following each publication. The storyline of the thesis is first based on obtaining spatial surface information from point meteorological observations. Then, the spatial information is used to estimate precipitation phase for diagnosing purposes. And finally, include extrapolation techniques and numerical weather prediction models to nowcast the precipitation phase with a forecast lead time of 180 minutes.

The first block of the thesis presents a methodology to interpolate high spatially and temporally resolved meteorological observations. Interpolation techniques have been widely studied and verified for daily and monthly observations, but limited for hourly or sub-hourly time scales. At these scales, observations tend to be more irregular and present higher variability as they are influenced by weather conditions, such as the presence of fog banks or thermal inversions. For this reason, an adaptive interpolation system was proposed. It is based on the combination of three elements: clustering, multiple linear regression, and residual correction. Meteorological observations are first divided in several clusters of variable size to separate areas prone to be affected by different weather conditions. A multiple linear regression is calculated for each cluster and then compared against an MLR that considers all data. It is in this step where the proposed system plays its role. The system, based on interpolation errors, decides which MLR uses in each cluster: that calculated using the stations of the cluster only or that using all stations available. The adaptive character of the system lays on using different number of clusters and test all them every time an interpolation is conducted. The system was successfully applied in three European regions, and results indicate a reduction of RMSE when the proposed interpolation system is used compared to using a single MLR considering all stations.

Once the step to interpolate point meteorological observations is achieved, the thesis focuses on discrimination of the precipitation phase in the following two blocks. The second block evaluates different precipitation phase discrimination schemes based on surface observations for diagnosing purposes. These schemes set thresholds on meteorological variables upon which precipitation is classified as rain, mixed or snow. In order to perform the evaluation of the schemes around 7700 quality-controlled observations of precipitation phase were gathered from different sources concerning Catalonia. According to the verification results, the schemes including air saturation conditions perform best, that is wet bulb temperature or combining air temperature with relative humidity. When analysing the schemes for specific snowfall events, a certain variability among the optimum thresholds was identified. This lead to suggest a range of thresh-

olds when monitoring snowfall events. In addition, apart from the quality-controlled observations, citizen science and crowd sourced observations were also collected and evaluated showing both advantages and limitations.

The third block of the thesis is also focused on precipitation phase determination, but in this case for nowcasting purposes. Apart from considering surface precipitation phase discrimination schemes, algorithms based on vertical temperature profiles, which play a key role on determining precipitation phase at ground level, were also considered. According to the threshold and performance variability observed when diagnosing precipitation phase and based on previous studies, combinations of algorithms were also taken into account in this block. The performance of the different algorithms and their combinations was assessed in eight low-altitude snowfall events reported in Catalonia between 2010 and 2021. Verification results showed that a combination of algorithms is preferable as it may provide a wide perspective to forecasters during precipitation phase transitions. In addition, this block included the implementation of a probabilistic methodology to nowcast the precipitation field.

The results obtained in the present thesis allowed to adjust and improve the real-time precipitation phase discrimination at the Meteorological Service of Catalonia. In addition, a nowcasting of precipitation phase product was also developed and operationally implemented. The results may also contribute to add a new verification dataset for precipitation phase discrimination purposes, together with the evaluation of precipitation phase schemes with interpolated meteorological variables and the development of spatially resolved products.

Resum

Discriminar el tipus de precipitació a la superfície és una informació fonamental en diverses aplicacions meteorològiques i hidrològiques, com ara el perill d'allaus, la seguretat a la carretera durant condicions hivernals, o les inundacions provocades per pluges sobre superfícies nevades. Una gran varietat d'estudis s'han centrat en la discriminació del tipus de precipitació a la superfície mitjançant diferents tècniques, des d'arbres de decisió basats en paràmetres de perfils verticals de temperatura fins a algoritmes basats en intel·ligència artificial, passant per esquemes amb parametritzacions de microfísica. Tot i així, la discriminació del tipus de precipitació a la superfície continua sent un repte, especialment quan la temperatura és propera al 0°C. Diversos estudis han enumerat un seguit de mancances relacionades amb la discriminació del tipus de precipitació i aquesta tesi pretén aportar el seu granet de sorra en algun d'ells. A més, la tesi també neix de la necessitat del Servei Meteorològic de Catalunya de disposar d'un producte ajustat i verificat per classificar la precipitació en pluja, aiguaneu i neu, tant a nivell de diagnosi com a predicció a molt curt termini. Per tal d'assolir els dos tipus de requeriment s'han definit un total de sis objectius específics (OE) a partir dels quals s'ha estructurat la tesi. Són els següents:

- OE1. Obtenció d'un sistema d'interpolació adaptatiu per a terreny complex, i per a altes resolucions temporals i espacials.
- OE2. Avaluació i ajust de diferents esquemes i variables meteorològiques per diagnosticar la discriminació del tipus de precipitació en la superfície.
- OE3. Avaluació de les observacions obtingudes a partir de la ciència ciutadana per al monitoratge de nevades a cotes baixes.
- OE4. Desenvolupament i avaluació de diferents esquemes per a la predicció a molt curt termini del tipus de precipitació.

- OE5. Avaluació de tècniques de conjunts per a la predicció a molt curt termini del tipus de precipitació.
- OE6. Implementació operativa d'un producte per discriminar el tipus de precipitació en la superfície.

Aquesta tesi està basada en un compendi de tres articles científics i, per tant, s'han definit tres blocs seguint cadascuna de les publicacions. El guió d'aquesta tesi comença amb l'obtenció d'informació espacial a partir d'observacions meteorològiques puntuals. Després, amb aquesta informació s'estima el tipus de precipitació per a un producte de diagnosi. Finalment, s'inclouen tècniques d'extrapolació i models numèrics de predicció per tal d'obtenir una predicció a molt curt termini (180 minuts) del tipus de precipitació.

El primer bloc de la tesi presenta una metodologia per a interpolar observacions meteorològiques de superfície a una alta resolució temporal i espacial. Hi ha multitud de tècniques d'interpolació i totes han estat a bastament estudiades per a escales diàries o mensuals. Ara bé, els estudis centrats en la verificació per a escales temporals més petites, com ara horàries o 30-minutals són més escassos. A aquestes escales les observacions tendeixen a presentar molta variabilitat, ja que es veuen notablement influenciades per les condicions meteorològiques, ja sigui la presència de bancs de boira o d'inversions tèrmiques. Per aquesta raó, s'ha proposat un sistema d'interpolació dinàmic. Està basat en la combinació de tres elements: agrupació (*clustering*), regressió lineal múltiple, i correcció dels residus. Primer, les observacions meteorològiques es divideixen en diferents grups de mides diferents amb la intenció de separar zones que tendeixen a presentar condicions meteorològiques diferents. Per a cada agrupació es calcula una regressió lineal múltiple i després es compara amb una regressió lineal múltiple obtinguda amb totes les estacions. És en aquest punt on entra en joc el sistema d'interpolació proposat. El sistema, tenint en compte els errors de la interpolació, decideix quina regressió lineal múltiple fa servir per a cada agrupació. O bé la calculada només amb les estacions que es troben dins l'agrupació, o bé la que considera totes les estacions. El caràcter adaptatiu del sistema recau en la utilització de diferents nombres d'agrupacions, testejant-les totes cada vegada que es vol fer una interpolació. El sistema s'ha verificat de forma satisfactòria a tres regions europees i els resultats indiquen una reducció de l'error quadràtic mitjà en comparació a fer servir una sola regressió lineal múltiple fent servir totes les estacions.

Una vegada s'ha obtingut una metodologia per a la interpolació d'observacions meteorològiques, la tesi se centra en la discriminació del tipus de precipitació en els següents dos blocs. El segon bloc avalua diferents esquemes basats en observacions de superfí-

cie per a discriminar el tipus de precipitació. Aquests esquemes estableixen llindars a variables meteorològiques a partir dels quals la precipitació es classifica en pluja, aiguaneu i neu. Per tal de fer aquesta avaluació s'han emprat al voltant de 7700 observacions del tipus de precipitació obtingudes de diverses fonts oficials. Els resultats de la verificació indiquen que els esquemes que inclouen informació respecte a les condicions de saturació de l'aire presenten un millor comportament, com ara la temperatura del termòmetre humit o la combinació de temperatura de l'aire i humitat relativa. Ara bé, quan s'analitzen els esquemes de discriminació per a esdeveniments concrets s'observa una certa variabilitat entre els llindars utilitzats. Això suggereix considerar un rang de llindars, i no només uns d'establerts, per al monitoratge de nevades. A més a més, a part de les observacions de fonts oficials, s'han tingut en compte observacions obtingudes a partir de ciència ciutadana mostrant-ne els avantatges i les limitacions.

El tercer bloc de la tesi també està enfocat a la discriminació del tipus de precipitació, però en aquest cas no per a la diagnosi sinó per a la predicció a molt curt termini. A part de considerar els esquemes de discriminació del tipus de precipitació basats en observacions de superfície, també es tenen en compte algoritmes basats en perfils verticals de temperatura. Aquests tenen un paper important a l'hora de determinar el tipus de precipitació en superfície. Tenint en compte la variabilitat observada al segon bloc pel que fa als llindars de discriminació i en estudis anteriors, també es consideren combinacions de diferents algoritmes. La verificació dels esquemes de forma individual i les seves combinacions es realitza a partir de vuit episodis de nevades a cotes baixes ocorreguts a Catalunya entre el 2010 i el 2021. Els resultats mostren que una combinació d'algoritmes és preferible, ja que pot proporcionar un punt de vista més ampli als predictors durant transicions de tipus de precipitació. A més a més, aquest bloc inclou la implementació d'una metodologia probabilística per a la predicció a molt curt termini del camp de precipitació.

Els resultats obtinguts en aquesta tesi han permès ajustar i millorar el producte de discriminació del tipus de precipitació en temps real del Servei Meteorològic de Catalunya. A més, també s'ha desenvolupat i implementat de forma operativa la predicció a molt curt termini del tipus de precipitació. Els resultats de la tesi també poden contribuir a afegir una nova verificació de metodologies per discriminar el tipus de precipitació, juntament amb l'ús de dades interpolades per a la discriminació del tipus de precipitació en comptes de dades d'observades.

Contents

Acronyms

AMS American Meteorological Society

AWS Automatic Weather Station

CoCoRaHS Community Collaborative Rain, Hail, and Snow network

DSD Drop Size Distribution

DT Dual Threshold

GPM Global Precipitation Measurement Mission

GSS Gerrity Skill Score

INCA Integrated Nowcasting through Comprehensive Analysis

ISD Integrated Surface Database

KS Koistinen-Saltikoff

MAE Mean Absolute Error

METAR Meteorological Terminal Air Report

MICA Meteorological field Interpolation based on Clustered data Analysis

MLR Multiple Linear Regression

mPING Meteorological Phenomena Identification Near the Ground

NIVOBS Nivo-meteorological Observations

NWP Numerical Weather Prediction

POD Probability Of Detection

PSS Peirce Skill Score

RMSE Root Mean Square Error

SMC Servei Meteorològic de Catalunya

SPP Surface Precipitation Phase

STEPS Short-Term Ensemble Prediction System

SYNOP Surface Synoptic Observations

WMO World Meteorological Organization

WRF Weather Research and Forecasting

XEMA Xarxa d'Estacions Meteorològiques Automàtiques

XOM Xarxa d'Observadors Meteorològics

Chapter 1

Introduction

1.1 Overview

Precipitation, as defined in the AMS Glossary (AMS, 2022), is *all hydrometeors formed in the atmosphere, including liquid, solid, or a combination of the two, that are large enough to fall as a result of gravity*. These hydrometeors play a critical role in the global water and energy cycle of Earth (Huntington, 2006; Schneider et al., 2014). It is precipitation the primary source of freshwater and drives the responses of hydrological, ecological, and atmospheric systems. Due to its importance, precipitation has long suscitated the interest of the scientific community and has been measured and observed in many different ways. Examples of instruments providing point observations are rain gauges, disdrometers and vertically pointing Doppler radars (Peters et al., 2002). On the other hand, single and dual polarisation weather radars provide spatial observations of precipitation (Bech & Chau, 2012), together with space-borne microwave sensors (Tropical Rainfall Measurement Mission, Kummerow et al. (1998); Global Precipitation Measurement Mission, Hou et al. (2014)). In most cases, precipitation is not studied as a whole but their associated characteristics, such as amount or distribution (Cortesi et al., 2012; Gonzalez & Bech, 2017), and their consequences, such as floods (Cortès et al., 2018; del Moral et al., 2020) or droughts (Livada & Assimakopoulos, 2007; Caloiero et al., 2018).

Phase is one of the associated characteristics of precipitation, and it is the one the present thesis is devoted to. Precipitation phase and its correct determination at ground level is critical in many meteorological, climatological and hydrological applications across manifold scientific fields. It is crucial for the correction of rain gauge mea-

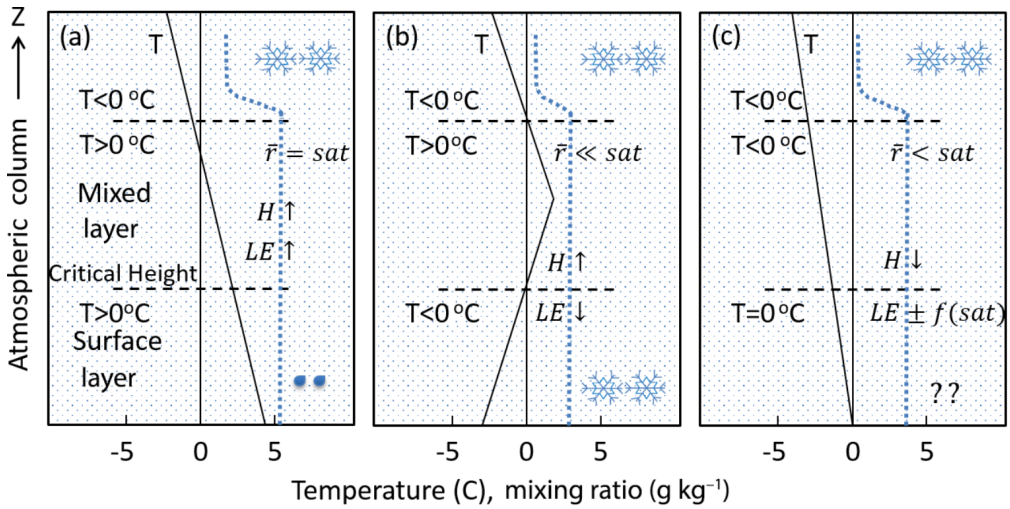

Fig. 1.1: Different temperature profiles and corresponding precipitation phase at ground level. (a) corresponds to rain, (b) to snow and (c) highlights the uncertainty involving the precipitation phase determination and the difficulty to predict it with only surface observations. The black line represents the temperature profile and the blue dotted one for the mixing ratio. H stands for sensible heat, LE for latent heat of evaporation, $f(sat)$ for function of saturation, and r for mixing ratio. [*Figure 2 from Harpold et al. (2017b)*]

surements in winter conditions (Rasmussen et al., 2012; Thériault et al., 2012; Buisán et al., 2017); for winter road safety (Andrey et al., 2003; Stewart et al., 2004; Papagiannaki et al., 2013); for the prediction of the timing and magnitude of surface runoff and streamflows (**?**); for the study of the surface energy budget (Box et al., 2012); and for climate change studies (Marty & Meister, 2012; Irannezhad et al., 2017; Hynčica & Huth, 2019).

Precipitation phase at ground level is influenced by multiple factors including vertical temperature profiles and precipitation characteristics (Bourgouin, 2000; Kain et al., 2000; Thériault et al., 2010), the elevation of a melting layer (Minder et al., 2011), the interaction between hydrometeors (Stewart, 1992; Kienzle, 2008), the wind speed (Matsuo & Sasyo, 1981; Harder & Pomeroy, 2013), or atmospheric stability (Thériault & Stewart, 2007), among others (Figure 1.1). Accounting for all the processes involved in the knowledge of the precise nature of precipitation phase, its discrimination among different types is still challenging. For this reason, different methodologies have been proposed to estimate precipitation phase at ground level including surface based observations, remote sensing observations and atmospheric information (Harpold et al.,

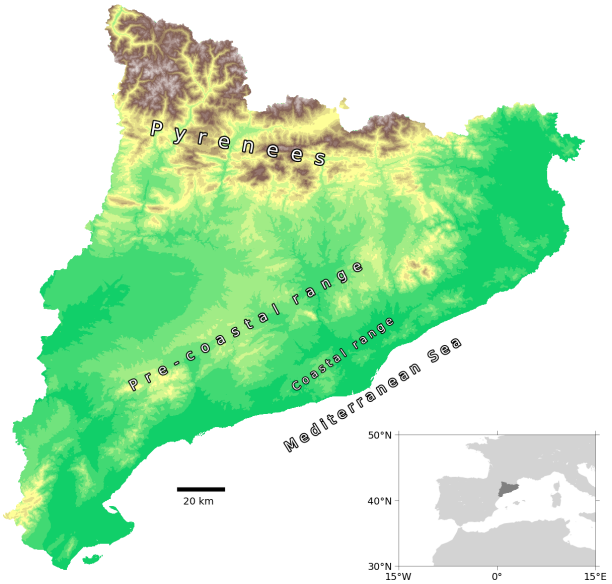

Fig. 1.2: Map of Catalonia showing the main orographic features and the position of the region in Europe.

2017b), and further details can be found in Section 1.2.3. One of these methodologies consists on establishing a threshold to a meteorological variable upon which precipitation is classified as rain, sleet or snow. It is the case of the Meteorological Service of Catalonia, which implemented the methodology proposed in Gjertsen & Ødegaard (2005) based on Koistinen & Saltikoff (1998).

Catalonia is located in the NE Iberian Peninsula and is bordered by the Mediterranean Sea to the E and S, the Ebre Valley to the W and the Pyrenees mountain range to the N (Figure 1.2). It is the Pyrenees and Pre-Pyrenees, with peaks exceeding 3,000 m above sea level, where snowfall commonly occurs from late autumn to early spring. However, it occasionally takes place in lower altitude areas where population is concentrated and major disruption to transport and socioeconomic activities can occur, as the 2010 heavy snowfall event (Bech et al., 2003; Llasat et al., 2014). As an example of the sporadic nature of these events, only 16 snowfall events were reported in Barcelona City between 1949 and 2009 (Aran et al., 2010). Fernandez (2016) reported 9 high societal impact snowfall events in Catalonia between 1981 and 2015. Although these events rarely occur, they can cause substantial turmoil due to vulnerable infrastructures and high population density areas being affected. For this reason, the Meteorological

Service of Catalonia decided to implement a methodology to discriminate precipitation phase at ground level to better monitor snowfall events. Still, its verification and adjustment have been very limited (Bech et al., 2014; Vidal et al., 2014).

The motivation of the present thesis under the *Industrial Doctorate Programme* framework comes from the need to evaluate and adjust the surface precipitation phase product available at the SMC as thresholds to discriminate precipitation phase changes across the Northern Hemisphere (Jennings et al., 2018). Furthermore, the *Industrial Doctorate Programme* project includes one step forward, the nowcasting of the precipitation phase to improve the monitoring of low altitude snowfall events. Apart from the two project deliverables, the present thesis also intends to do its part in three research gaps: to test precipitation discrimination schemes using regionalised meteorological variables rather than observed (Froidurot et al., 2014), the verification of current precipitation phase discrimination schemes with the inclusion of relative humidity information (Froidurot et al., 2014; Harpold et al., 2017b), and the development of spatially resolved products (Harpold et al., 2017b).

1.2 State of the art

1.2.1 Spatial interpolation of surface observations

Meteorological surface observations are usually measured with manned and automatic weather stations, commonly sparsely distributed over a region, conforming a network from which weather conditions can be obtained. These irregularly spaced stations provide information from specific locations, but continuous fields are generally necessary for a large number of subjects: climate analysis (Nastos et al., 2013; Mathbout et al., 2018), severe weather surveillance (Gjertsen & Ødegaard, 2005; Rogelis & Werner, 2013) or agrometeorological applications (Webb et al., 2016; Viggiano et al., 2019). For this reason, spatial interpolation techniques are implemented to obtain high resolution gridded fields.

The interpolation techniques can be classified into methods based only on the variable of interest and methods including additional information, which account for cause and effect relationships. Two examples of the first kind of interpolation methods are the so-called inverse of the distance (Cressman, 1959; Barnes, 1964; Lu & Wong, 2008) and ordinary kriging (Cressie, 1990). On the other hand, linear regression and kriging with external drift are two interpolation techniques that relate the variable to estimate with other parameters, such as precipitation or temperature with altitude (Kutner et al., 2005;

Table 1.1: List of selected hydrometeors and their description from WMO (2017b) and AMS (2022).

Hydrometeor	Description
Drizzle	Fairly uniform precipitation of very fine drops of water very close to one another that falls from a cloud not exceeding 0.5 mm of diameter.
Rain	Precipitation of drops of water that falls from a cloud equal or exceeding 0.5 mm of diameter.
Snow	Precipitation of ice crystals, singly or stuck together, that falls from a cloud.
Ice pellets	Precipitation of transparent ice particles that falls from a cloud. These particles are usually spheroidal or irregular, and rarely conical. Their diameter is less than 5 mm.
Hail	Precipitation of particles of ice (hailstones). These can be either transparent, or partly or completely opaque. They are usually spheroidal, conical or irregular in form, and generally 5–50 mm in diameter. The particles may fall from a cloud either separately or agglomerated in irregular lumps.

Lanfredi et al., 2015).

A further step in spatial interpolation is using residual or anomaly correction which forces the interpolated field values to be that of observations at their locations. This correction is not restricted to observation points, but a residual correction field can also be obtained through spatial interpolation which usually improves the performance of the interpolated field (Stahl et al., 2006; Joly et al., 2011; Szymanowski & Kryza, 2012; Piazza et al., 2015).

1.2.2 Present weather observations

Present weather is a description of the weather phenomena present at the time of observation (WMO, 2018). This kind of observations is qualitative and, according to WMO (2018) can be categorised into precipitation, atmospheric obscurity and suspensoids, and other weather events. Precipitation is labelled depending on its character, type and intensity. However, the present study focused on its type which includes rain, drizzle, snow, ice pellets, and hail, among others (WMO, 2017b). An example of this classification can be seen in Table 1.1.

Precipitation phase observations can be obtained either through trained human ob-

Fig. 1.3: OTT Parsivel 2 disdrometer from University of Barcelona (center) and an automated weather station from the SMC (right) located in Das, La Cerdanya (Pyrenees).

servers or automatic sensors specifically designed for this purpose, such as optical disdrometers (Bloemink & Lanzinger, 2005; Löffler-Mang & Blahak, 2001) (Figure 1.3). Regarding observations from trained observers, they are usually reported from airports and manned observatories. Both automatic and manned observations are generally compiled from more than 20,000 stations (METAR and SYNOP) in the Integrated Surface Database (ISD, Smith et al. (2011)).

Apart from the official sources of present weather observations, citizen science and crowdsourcing experienced an increasing role in recent years in Earth observation topics (Fritz et al., 2017). Examples of user friendly applications for citizens to report weather observations are CoCoRaHS (Reges et al., 2016) and mPING (Elmore et al., 2014). However, the use of social networks is starting to feature as a weather observation source. "Snowtweets" (King et al., 2009) or "Picking up Hailstones" (Farnell & Rigo, 2020) (Figure 1.4) are two examples of successful campaigns to encourage citizen participation in reporting snow and hail size, respectively. Nevertheless, non-conventional sources must be taken with care due to their inherent issues, such as location and time uncertainties (Schuster et al., 2005; Saltikoff et al., 2010; Reeves et al., 2014).

Fig. 1.4: "Picking up hailstones" Meteorological Service of Catalonia banner to encourage citizen participation in reporting hail size. [*Source: SMC*]

1.2.3 Precipitation phase discrimination

Precipitation phase discrimination at the surface level has been widely studied, that is, classifying precipitation among the liquid (e.g., rain and drizzle), freezing (e.g., freezing rain, and rain and snow mixed) or frozen (e.g., snow, hail and ice pellets) categories. Numerous approaches have been considered to address this problem. A rather simple but effective methodology to discriminate rain and snow is threshold-based schemes based on near-surface variables. However, more complex ones, which are able to classify a wider range of precipitation types, deal with hydrometeor mixing ratios at ground level aided by microphysics parameterisations (Thompson et al., 2008). Implicitly, these methodologies are also constrained on the available information. For example, instrumentation such as dual-polarisation weather radars (Chandrasekar et al., 2013; Thompson et al., 2014) or vertically pointing Doppler radars (Garcia-Benadí et al., 2020, 2021) allow to directly measure parameters from which precipitation phase can be estimated. On the other hand, if vertical temperature profiles, either modelled or observed, are available, different algorithms can be applied to infer the precipitation phase at surface level.

The present thesis, according to the available observations and instrumentation, focused on two specific types of methodologies: those based on surface observations and those on vertical temperature profiles.

Surface observation based methods

Methodologies relying only on surface observations present the limitation of being able to classify precipitation among three types: rain, snow and a mixture of both (hereafter, mixed). However, the simplicity and effectiveness of these methods make them a

suitable option to discriminate precipitation between rain and snow. But how do they work? These methodologies establish different thresholds to meteorological variables upon which precipitation phase is classified. For example, if one threshold is set and precipitation falls above it, precipitation is classified as rain, otherwise as snow. When two thresholds are set, then precipitation can be classified among rain, snow and mixed.

Several meteorological variables have been evaluated in different studies to discriminate precipitation phase: air temperature (Liu, 2008; Feiccabrino & Lundberg, 2009), dew point temperature (Feiccabrino & Lundberg, 2009; Marks et al., 2013), wet bulb temperature (Froidurot et al., 2014; Behrangi et al., 2018) or a combination of air temperature and relative humidity (Koistinen & Saltikoff, 1998). These studies are usually based on point observations, but their results can also be implemented to observations interpolated or NWP model fields and derive a precipitation phase classification over extensive areas. This is the case of Gjertsen & Ødegaard (2005) in Norway, Sims & Liu (2015) in the precipitation phase product of the GPM product and Tang et al. (2018) on the contiguous United States (CONUS).

These methodologies have the advantage of being easy to implement, but at the same time present the inconvenience that they do not account for what is happening to precipitation before reaching the ground. Bourgouin (2000), Kain et al. (2000) and Heymsfield et al. (2021) stated that vertical temperature profiles and processes affecting precipitation above the surface play a key role on determining the precipitation phase. However, if discrimination of precipitation phase is narrowed to rain and snow, these methodologies provide a reasonable trade-off between its simplicity and effectiveness of the results (Froidurot et al., 2014; Jennings et al., 2018; Ding et al., 2014).

Vertical temperature profile based methods

Methodologies including atmospheric information, such as vertical temperature profiles, allow to take into consideration the conditions affecting hydrometeors before they reach the ground, which present a key role on determining the precipitation phase at surface level (Bourgouin, 2000). These kinds of methodologies usually first determine which type of hydrometeors can be found or form in clouds, and then evaluate their phase change along their path to the surface. These phase changes depend on multiple factors, such as the position of the 0°C isotherm (Minder et al., 2011), the collision processes between hydrometeors (Stewart, 1992), and the latent heat exchange with the environment during the melting and freezing of precipitation (Thériault & Stewart, 2010).

Different approaches have been proposed to consider the processes affecting the

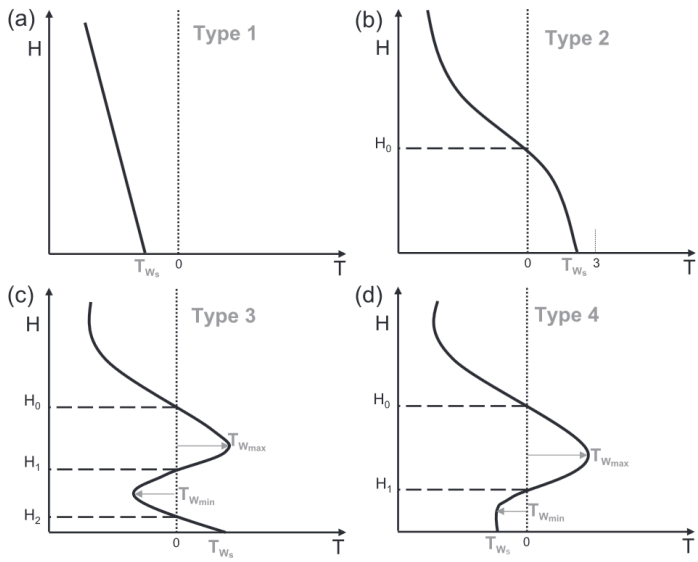

Fig. 1.5: Different types of wet bulb temperature profiles used in Schuur et al. (2012) to classify precipitation types. T corresponds to wet bulb temperature; H to height; H_0 depict the heights of the 0°C crossing points in the profiles; T_{w_s} to surface wet bulb temperature; $T_{w_{min}}$ and $T_{w_{max}}$ to the thresholds for the maximum and minimum acceptable temperature in the cold and warm layers, respectively. Type 1 (a) corresponds to snow, Type 2 (b) to rain or wet snow, Type 3 (c) to ice pellets or rain, and Type 4 (d) to freezing rain or ice pellets. [*Figure 4 from Schuur et al. (2012)]*

phase of precipitation. One of these approaches evaluates the thickness and average virtual temperature of different layers limited by four specific pressure surfaces to classify precipitation phase, together with surface information (Keeter & Cline, 1991; Zerr, 1997). However, it presents the drawback that information from all four pressure levels is mandatory, otherwise no precipitation phase estimation is provided. Other approaches focus on vertical thermodynamic profiles, either observed or modelled, which estimate first the type of hydrometeors at the precipitation generation layer and then calculates hydrometeor phase changes accounting for several effects and nature of the profiles. For example, an algorithm proposed in Ramer (1993) determines the initial ice fraction of hydrometeors being 0 for supercooled droplets and 1 for ice. Then, using wet bulb temperature profile, air pressure and relative humidity values establishes the ice fraction change of different layers from the precipitation generation level to the surface. A positive change indicates refreezing of hydrometeors, otherwise melting. Depending on the ice fraction reaching the surface, precipitation is classified. Other examples of

vertical thermodynamic profiles algorithms are those based on melting and refreezing energies, that is, accounting for the presence of warm and cold layers. This is the case of the algorithms proposed by Baldwin et al. (1994) based on wet bulb temperature profiles and by Bourgouin (2000) based on air temperature profiles. Schuur et al. (2012) proposed an algorithm using also wet bulb temperature profiles, but in this case based on the number of crossing points of 0°C and 2-m temperature (Figure 1.5).

Another approach is to use advanced microphysics parameterisations that explicitly resolve cloud, water vapor and precipitation processes. For example, in the WRF model there are different available schemes (Skamarock et al., 2019). These include the one proposed by Hong et al. (2004) which is a single-moment 3-class microphysics scheme that considers ice-phase and ice sedimentation parameterisations, and the one proposed by Thompson et al. (2008) which uses a microphysics bulk scheme with double-moment ice and rain and was developed for mid-latitude snowfall conditions in a convection-permitting grid scale.

1.2.4 Nowcasting precipitation and surface observations

Nowcasting is defined in WMO (2017a) as *forecasting with local detail, by any method, over a period from the present to 6 hours ahead, including a detailed description of the present weather*. Most of the nowcasting systems are focused on the prediction of precipitation and convective activity. Even meteorological variables such as temperature, relative humidity or wind have not been considered as themselves but as a mean for predicting convective developments (Haiden et al., 2011). However, nowcasting of wind speed and gust can be critical for aviation applications, such as determination of crosswinds, together with air temperature and relative humidity for visibility and fog related phenomena (Isaac et al., 2014b; Huang et al., 2012). Therefore, there is a need for nowcasting other meteorological variables apart from precipitation.

Most of the nowcasting systems are based on persistence or trends of observations which tend to result superior to NWP models for the first forecast hours (Bowler et al., 2006; Simonin et al., 2017). In case of precipitation, Eulerian or Lagrangian persistence and different advection schemes are the most common techniques, as done for the Nimrod system (Golding, 1998) or in STEPS (Seed, 2003), recently implemented in Python by Pulkkinen et al. (2019). In case of other meteorological variables, such as temperature or relative humidity, nowcasting is usually obtained considering an analysis field and the forecast variations of the analogous NWP fields as done in INCA (Haiden et al., 2011) and in Huang et al. (2012). In addition, methodologies including artificial intelligence and machine learning techniques are also emerging in the nowcasting field,

such as McGovern et al. (2017) for high-impact weather, Lagerquist et al. (2019) for prediction of synoptic-scale fronts, or Ravuri et al. (2021) for weather radar precipitation advection.

Another approach for nowcasting systems is the use of NWP models with data assimilation techniques. These may include, for example, radar information to relate reflectivity with hydrometeor properties of precipitation (Dowell et al., 2011). Performance of techniques using data assimilation is expected to improve that of extrapolation, except for the first steps of the forecast (Pierce et al., 2012; Sun et al., 2014).

1.3 Objectives

1.3.1 General objectives

Monitoring and surveillance of snowfall events in Catalonia (NE Iberian Peninsula), specially those affecting low altitude areas, require different tools to improve emergency management, together with the diagnostic and nowcasting of the precipitation phase during these events. Catalonia experienced snowfalls that caused high socio-economic impacts, such as in March 2010 (Bech et al., 2013; Llasat et al., 2014) or the so-called Storm Filomena in January 2021, where an accurate diagnostic and nowcasting surface precipitation phase product could have aided the forecasters during the monitoring of these events. The discrimination of precipitation phase has been addressed following different strategies including surface observations (Froidurot et al., 2014; Dai, 2008) and vertical temperature profiles derived from NWP models (Manikin, 2005; Wandishin et al., 2005). Therefore, these schemes should be evaluated in order to select the most appropriate one for two different purposes: diagnostic and nowcasting of the precipitation phase.

However, a robust and large database of precipitation phase observations is needed to obtain an accurate evaluation of the different schemes. These observations can be gathered from NIVOBS (Gavaldà et al., 2014; Apodaka et al., 2018), ISD (Smith et al., 2011), Parsivel observations (Bloemink & Lanzinger, 2005; Gonzalez et al., 2019) and observers from XOM (Ripoll et al., 2016). Apart from these sources, crowd-sourced data from social networks like Twitter, successfully used in other meteorological situations (Farnell & Rigo, 2020), and spotters from XOM (Ripoll et al., 2016) should also be explored.

Thereby, two general objectives (GO) were set from which the present thesis has been constructed.

- **GO1. Adjustment and implementation of a diagnostic surface precipitation phase product.** The SMC had implemented a product to discriminate the precipitation phase but it has never been validated nor adjusted. Therefore, the validation of the current product together with the evaluation of different precipitation phase schemes should result in an updated version of the product focusing on three aspects: the interpolation methodology, the discrimination scheme, and the thresholds used in the scheme. All this is done to obtain an accurate and robust precipitation phase product.

- **GO2. Development and implementation of a nowcasting surface precipitation phase product.** Based on the framework set during the implementation of the diagnostic product, focus on a nowcasting product can also be addressed. The work to be developed in this general objective comprises exploring different nowcasting tools not only regarding the precipitation phase but also the precipitation field. Here the implementation of an operational nowcasting precipitation phase product in the SMC is also considered.

1.3.2 Specific objectives

Based on the general objectives set, six specific objectives (SO) were defined to fulfill the work of the present thesis. These are presented below.

- **SO1. Obtention of a dynamic interpolation scheme suitable for complex terrain, and high spatial and temporal resolution.** Precipitation phase information can be obtained for specific locations by applying discrimination schemes to observations from an AWS network. However, if spatial information of precipitation phase is needed, an interpolation methodology must be set. This methodology should account for complex terrain characteristics, high temporal resolution observations, which can present large variability, and, at the same time, deal with high spatial resolution to properly capture precipitation phase transition areas.

- **SO2. Evaluation and adjustment of different schemes and meteorological variables to diagnose discrimination of the surface precipitation phase.** Different methodologies to diagnose precipitation phase based on surface observations have been proposed in past studies, usually based on establishing thresholds to meteorological variables. Although one of them (Koistinen & Saltikoff, 1998) has already been implemented in the SMC, verification studies have been very limited (Bech et al., 2014; Vidal et al., 2014). Therefore, a need for validation and

adjustment of threshold methodologies is required to provide the most accurate discrimination available regarding the precipitation phase among rain, sleet and snow.

- **SO3. Assessment of citizen science and crowd sourced observations for monitoring snow events.** Non-conventional observations, such as those obtained from citizen science have proved their value in reporting different meteorological phenomena, such as hail (Farnell & Rigo, 2020) or tornadoes (Rodríguez et al., 2021). Therefore, observations gathered from spotter networks or crowd-sourced from social networks should be evaluated addressing the precipitation phase discrimination problem. The SMC supports a weather surveillance spotter network (XOM) which can be evaluated together with observations gathered from Twitter social network.

- **SO4. Development and evaluation of different schemes to nowcast discrimination of the precipitation phase.** Vertical temperature profiles play a key role in determining the precipitation phase at surface level. The nowcasting of precipitation phase allows the inclusion of NWP models from which vertical temperature profiles can be derived. Past studies have proposed different algorithms based on temperature profiles to determine precipitation phase. Therefore, evaluation of these algorithms together with threshold schemes based on the extrapolation of surface meteorological observations is proposed.

- **SO5. Evaluation of ensemble techniques to nowcast discrimination of the precipitation phase.** Past studies noted that performance of precipitation phase discrimination algorithms based on vertical temperature profiles presents certain variability along the snowfall event or between different events. A combination of algorithms is suggested to account for this variability. This approach is also proposed in the present thesis combining not only vertical temperature related algorithms but also those based on extrapolated surface observations.

- **SO6. Implementation of a precipitation phase product in an operational chain.** Considering the applied nature of a PhD Industrial project the ultimate reason for the general and specific objectives is to operatively implement a diagnosing and nowcasting product in the SMC. This operational chain should account for the results obtained in the aforementioned objectives producing an accurate and robust product.

1.4 Structure of the thesis

The present thesis is built as a compendium of three publications and divided in five chapters.

Chapter 1 includes the present section which includes the introduction divided in an overview, the state of the art of the general topics addressed in the thesis and its general and specific objectives. The following three chapters comprise the major part of the thesis, including the three published papers. Chapter 2 addresses the interpolation of surface meteorological observations at high temporal and spatial resolution scales included in:

- Casellas, E., Bech, J., Veciana, R., Miró, J. R., Sairouni, A., & Pineda, N. (2020). A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, 246, 105103.

Chapter 3 focuses on the evaluation and adjustment of different precipitation phase scheme and their application for diagnosing purposes. This topic is addressed in:

- Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Miró, J. R., & Sairouni, A. (2021). Surface precipitation phase discrimination in complex terrain. *Journal of Hydrology*, 592, 125780.

Chapter 4 is analogous to Chapter 3 but in this case focusing on nowcasting purposes for both the precipitation field and precipitation phase. The paper comprising this study is:

- Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J. R., Moré, J., Rigo, T. & Sairouni, A. (2021). Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(739), 3135-3153.

The last chapter of the thesis, Chapter 5, includes the conclusions and gives answers to the general and specific objectives set in Chapter 1. In addition, possible future lines of research for the topics addressed in this thesis are sketched. Finally, two Appendix conclude the present thesis. Firstly, in Appendix A the contributions made in scientific journals, and poster presentations in congress and oral presentations in seminars are listed. Secondly, Appendix B comprises the evaluation and verification of different methodologies to refine air temperature and dew point temperature NWP model fields from a coarse to high spatial resolution. This assessment constituted one of the first steps to develop the study included in Chapter 4.

Chapter 2

Spatial interpolation of surface observations

2.1 A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain

2.1.1 Summary

In this paper the development of an interpolation scheme for sub-daily meteorological observations is presented. Different interpolation approaches have been proposed for daily and monthly observational data, but interpolation studies at hourly temporal scales are much more limited. High temporal scale typically implies more irregular data due to higher variability of weather conditions. For example, thermal inversions or fog banks yield added difficulties for interpolation, since conditions over a region differ whether they are affected by these phenomena or not. Therefore, the interpolation scheme presented in this paper aims to provide a meteorological analysis scheme suited for complex terrain and high temporal resolution observations to face local isolated meteorological phenomena.

How is this achieved? The presented scheme, "Meteorological field Interpolation based on Clustered data Analysis" (MICA), combines three elements: clustering, multiple linear regression (MLR) and residual correction. The first element divides data in different clusters of variable size aiming to separate regions prone to be concurrently affected by dissimilar weather conditions. Then, a MLR is calculated for each cluster and compared against a global MLR, which considers all data. Here is where MICA, based

on interpolation errors, decides which MLR uses for a cluster: the one calculated with data from the cluster only or that one calculated with all data. Therefore, the resultant interpolated field may be composed of a global MLR model alone, a combination of global MLR and cluster MLR models, or cluster MLR models only. Finally, a residual correction field is subtracted from the estimated field to account for local effects. This correction field is obtained by interpolating the difference between the estimated and observed values at weather stations.

This interpolation scheme was applied to three European regions: Catalonia (Spain), Emilia-Romagna (Italy) and Baden-Württemberg (Germany). The results, comparing the application of MICA against a global MLR model, indicate a reduction of cross-validation errors, being higher in Catalonia and Emilia-Romagna. This can be explained due to their complex orography and the density of their AWS network compared to those of Baden-Württemberg. In addition, study cases in the three regions showed that the interpolation scheme is effective during fog and inversion conditions which yielded a clear error reduction and a better representation of complex orography features. The increase of middle and high altitude AWS station weight in MLR model calculation and the increase of linearity when AWS data is clustered may explain the improvement provided by the presented interpolation scheme.

2.1.2 Article

Casellas, E., Bech, J., Veciana, R., Miró, J. R., Sairouni, A., & Pineda, N. (2020). A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, 246, 105103.

Atmospheric Research 246 (2020) 105103

Contents lists available at ScienceDirect

Atmospheric Research

journal homepage: www.elsevier.com/locate/atmosres

A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain

Enric Casellas[a,b,*], Joan Bech[a], Roger Veciana[b], Josep Ramon Miró[b], Abdel Sairouni[b], Nicolau Pineda[b]

[a] Department of Applied Physics - Meteorology, University of Barcelona, Barcelona, Spain
[b] Meteorological Service of Catalonia, Barcelona, Spain

ARTICLE INFO

Keywords:
Spatial interpolation
Clustering
Hourly analysis
Complex terrain
Regression model

ABSTRACT

An adaptive high-temporal resolution interpolation scheme for meteorological observations is presented. It stems from a combination of linear regression, anomaly correction and clustering. A number of approaches to tackle this problem for monthly and daily data have been proposed in the past, but interpolation studies at sub-daily temporal scales are much more limited. Hourly and sub-hourly observational datasets use to present high variability that may be related to different weather conditions. In the proposed methodology, rather than considering the whole data set to perform the interpolation, data are divided in different clusters of variable size, separating regions with potential dissimilar behaviour. A linear regression model is calculated for each cluster and compared against a global model obtained considering all the observations. Only those clusters whose regression model yields a reduction of errors compared to the global model are selected. The adaptive condition lays on that several numbers of clusters are tested and the one that performs the best, in terms of Root Mean Square Error, is selected every time an interpolation is conducted. The methodology presented provides gridded analysis fields of hourly and sub-hourly intervals at 250 m of horizontal resolution. It was originally developed for a complex terrain region (Catalonia, NE Spain), and it was also demonstrated in the German Land of Baden-Württemberg and in the Italian region of Emilia-Romagna. Results show a reduction of cross-validation errors using the leave-one-out method for air temperature and dew point temperature fields and a proper representation of complex orography features. The scheme presented is implemented in Python as pyMICA and it is available as open-source software.

1. Introduction

High resolution gridded air temperature and relative humidity fields are necessary for a large number of subjects, such as agrometeorological applications (e.g. Webb et al., 2016; Le Roux et al., 2017; Viggiano et al., 2019), climate analysis (e.g. Nastos et al., 2013; Liuzzo et al., 2017; Mathbout et al., 2018) or severe weather surveillance (e.g. Gjertsen and Ødegaard, 2005; Lowe et al., 2011; Bech et al., 2013; Rogelis and Werner, 2013; Fehlmann et al., 2018). Air temperature and relative humidity are usually measured over a region through a network of irregularly spaced stations, from which continuous fields can be obtained through spatial interpolation. Several interpolation techniques are usually implemented for meteorological analysis. They can be classified on methods based only on the variable of interest and methods considering additional data, accounting for cause and effect relationships. The former includes, for example, inverse of the distance schemes (Cressman, 1959; Barnes, 1964; Lu and Wong, 2008) and ordinary kriging (Cressie, 1990). The latter includes, among others, linear regression and kriging with external drift which often relates the variable of interest with topographical parameters, such as precipitation or air temperature with altitude (Kutner et al., 2005; Lanfredi et al., 2015).

Both methodologies can be implemented separately or in a complementary way, which usually implies first a regression calculation and then an interpolation of the regression residuals (Ninyerola et al., 2000; Szymanowski et al., 2013; Kormos et al., 2018). This technique, depending on the variable of interest to interpolate, tends to report better results as obtained for air temperature by Stahl et al. (2006), Joly et al. (2011) and Szymanowski and Kryza (2012). Interpolation performance is dependant on the temporal resolution and spatial density of the observations (Berndt and Haberlandt, 2018). Consequently, sub-daily or hourly time scales imply typically more irregular data than lower temporal resolution observations (i.e. monthly or daily data) due

* Corresponding author at: Department of Applied Physics – Meteorology, University of Barcelona, Barcelona.
E-mail address: enric.casellas@meteo.ub.edu (E. Casellas).

https://doi.org/10.1016/j.atmosres.2020.105103
Received 20 March 2020; Received in revised form 8 June 2020; Accepted 9 June 2020
Available online 19 June 2020

to higher variability of meteorological phenomena, such as thermal inversions or fog, which yield added difficulties for interpolation. Therefore, accounting for local effects is critical when interpolating high time resolution observations, particularly in complex topography areas. For this reason, residuals, which are the difference between the predicted and observed values, can be used to force the interpolated field values at station locations to be the same as that of observations. Therefore, an anomaly correction of the residuals is often applied to the resultant interpolated field, accounting for local and isolated effects that may be disguised in regression and kriging models (Steinacker et al., 2006). Note that the analysis of precipitation fields at hourly or sub-hourly scale, given their high spatial variability in complex terrain, requires not only point (rain gauge) measurements but also weather radar observations (see for example Velasco-Forero et al. (2009) or Sideris et al. (2014)). For this reason, precipitation analysis is not considered here.

The main goal of this study is to provide a meteorological analysis scheme suited for complex terrain and high temporal resolution observations able to deal with isolated local meteorological phenomena. For this purpose, the interpolation scheme designed, based on the combination of Multiple Linear Regression (MLR) and anomaly correction (Ninyerola et al., 2000; Stahl et al., 2006; Kormos et al., 2018), is aided by grouping weather stations in clusters similarly to Ninyerola et al. (2007) in Spain, Joly et al. (2011) in France and Fick and Hijmans (2017) at global scale. Therefore, the proposed strategy is threefold: (1) classify weather stations in different clusters; (2) apply and verify the interpolation methodology individually on the clusters; (3) dynamically select the clusters that performs the best each time an interpolation is conducted. The designed methodology is illustrated at three different regions presenting different orographic and station density characteristics: Catalonia in Spain, Baden-Württemberg in Germany and Emilia-Romagna in Italy.

This paper is organized as follows. Section 2 describes the three study regions considered and the data used to illustrate the methodology developed. Section 3 provides an overview of the interpolation methodology. In Section 4 the analysis scheme proposed in this study is presented in detail. The numbers of clusters applied, the air temperature and dew point temperature cross validation results, together with case studies, for Catalonia, Baden-Württemberg and Emilia-Romagna are presented in Section 5. Summary and conclusions are reported in Section 6.

2. Data

Meteorological data considered in this study included air temperature, dew point temperature and relative humidity observations. Since the linear regressions were computed using air and dew point temperature, data from stations reporting relative humidity observations were transformed to dew point values through Buck (1981). The different station networks considered are depicted in the following subsections for each region of interest.

2.1. Catalonia

Located in the North-East of the Iberian Peninsula and with an extension about 32,000 km^2 it is mostly characterized by a Mediterranean climate. It is encompassed between the Pyrenees range to the North, the Mediterranean Sea to the East and South and the Ebre Valley to the West. Despite its relatively small area, Catalonia presents a notable orographic variability. From the Pyrenees range summits above 3000 m ASL (above sea level) to the coastal ranges from 500 to 1000 m ASL, passing through extensive plains such as the Central Depression (Fig. 1a). These irregularities contribute to strong temperature and precipitation contrasts between interior regions (Martn-Vide et al., 2008), where mean annual temperatures range from 0 °C to 17 °C and mean annual precipitation from 400 mm to 1200 m depending on the zone.

Meteorological station data were obtained from the automatic weather stations (AWS) network, Xarxa d'Estacions Meteorolgiques Automtiques (XEMA), managed and quality controlled by the Meteorological Service of Catalonia (SMC) (Serra et al., 2016; Llabrés-Brustenga et al., 2019). XEMA consists of 184 AWS that provide, among other measurements, temperature and relative humidity data with 30-min time resolution. XEMA stations are not uniformly distributed over the region, with higher density (160 stations) in flat areas where population is concentrated and lower density as altitude increases (Fig. 1). A few number of AWS (12) are located in middle altitudes, between 1000 and 2000 m ASL and 12 stations are installed at high mountain areas (Fig. 1b and c). In addition, due to complex orography, there are several areas prone to the formation of cold pools in which the temperature and humidity behaviour are difficult to characterize, such as in La Cerdanya Valley (Pagès et al., 2017).

2.2. Baden-Württemberg

The Land of Baden-Württemberg is located in the South-West of Germany with an extension of ~36,000 km^2 and limited by the Rhine Valley to the East and the Alps' foothills to the South. The region does not present great orographic complexity except for the Black Forest region where there are mountains reaching up to 1493 m ASL (Fig. 2a). This Land is characterized by a humid continental climate with a mean annual air temperature of 8.3 °C and a mean annual precipitation of 1000 mm, according to the German Climate Atlas (Kaspar et al., 2013).

Meteorological data, including air temperature and dew point temperature observations were derived from the AWS network supported and maintained by the German Meteorological Service (DWD) (Kaspar et al., 2013) and distributed by their Open Data Service (DWD Climate Data Center, 2018). Baden-Württemberg holds a network of 62 AWS which are mostly located in flat areas below 500 m ASL and close to populated areas (40) (Fig. 2b and c), 21 stations in middle altitudes and 1 station above 1000 m ASL, installed in the Feldberg mountain in the Black Forest region (Fig. 2b and c).

2.3. Emilia-Romagna

Encompassed between the Po river and the Apennines, and limited to the East by the Adriatic Sea and with an extension of ~22,500 km^2, the Italian Emilia-Romagna region is characterized by a sharp orographic transition between the extensive Po Plain and the Apennines (Fig. 3a). The highest summit, Monte Cimone (2165 m ASL), is located in the southern part of the domain. The region presents a subcontinental and cool temperate climate (Nistor, 2016) with a mean annual air temperature range of 5 °C to 14 °C and a mean annual precipitation ranging of 600 mm to 2300 mm (Antolini et al., 2016).

Meteorological station data was provided by the Agenzia Regionale per la prevenzione, l'ambiente e l'energia dell'Emilia-Romagna database (ARPAE). ARPAE supports a weather station network including 188 air temperature sensors, considering agrometeorological, urban and hydrometeorological stations (Fig. 3b). They are mostly located between 0 and 250 m in accordance with the orography of the region (Fig. 3c). However, they are also densely deployed in the Apennine area, comprising altitudes from 250 m ASL to 1250 m ASL (Fig. 3b and c).

3. Overview of the interpolation methodology

3.1. Multiple linear regression (MLR)

A MLR model allows to linearly predict a response variable (dependant) using different explanatory variables (independent). The MLR model can be expressed as:

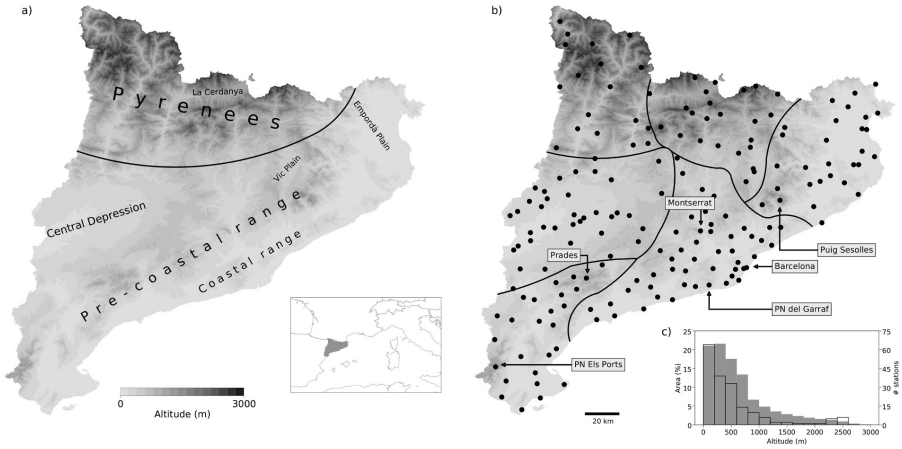

Fig. 1. a) Hypsometry and characteristic orographic features of Catalonia. b) Automatic weather stations used in this study. Black lines represent the cluster divisions considered in panels a) and b). c) Altitude distribution of the study region's relief (% of Catalonia's area, dark shaded histogram) and the automatic weather stations (non-shaded histogram).

$$y_i = \sum_k \beta_k x_{ik} + \varepsilon_i \quad (1)$$

where y_i is the dependent variable, in this case the meteorological variable of interest, β_k are the coefficients of linear regression, x_{ik} are the independent or explanatory variables and ε_i are the residuals of the regression, which are the difference between the predicted and observed values. The explanatory variables were introduced in the MLR in a forward stepwise procedure, likewise in Kurtzman and Kadmon (1999), Ninyerola et al. (2000) and Joly et al. (2011). Thereby, collinearity of explanatory variables is restricted (Joly et al., 2011).

The explanatory variables considered in this study for the MLR calculations were altitude, longitude, latitude and distance to the coast as proposed in previous studies (Hiebl et al., 2009; Brunetti et al., 2014). Elevation data was obtained from the 30-m spatial resolution Digital Elevation Model (DEM) provided by the NASA's Shuttle Radar Topography Mission and downloaded from Jarvis et al. (2008) for the three regions of interest. DEMs were resampled to a resolution of 250-m. Longitude, latitude and distance to coast were obtained considering the same extent and resolution as the DEMs. However, distance to the coast was only considered for Catalonia and Emilia-Romagna regions and, rather than considering the Euclidean distance, it was calculated through an exponential function (Eq. 2) following Ninyerola et al. (2007) and Joly et al. (2011):

$$d_{coast} = 1 - e^{-\frac{3 \times dist}{D}} \quad (2)$$

where d_{coast} is the resultant distance to coast of the function, *dist* is the Euclidean distance from a point to the coast line and D is the distance from which the influence of the sea is considered negligible. This

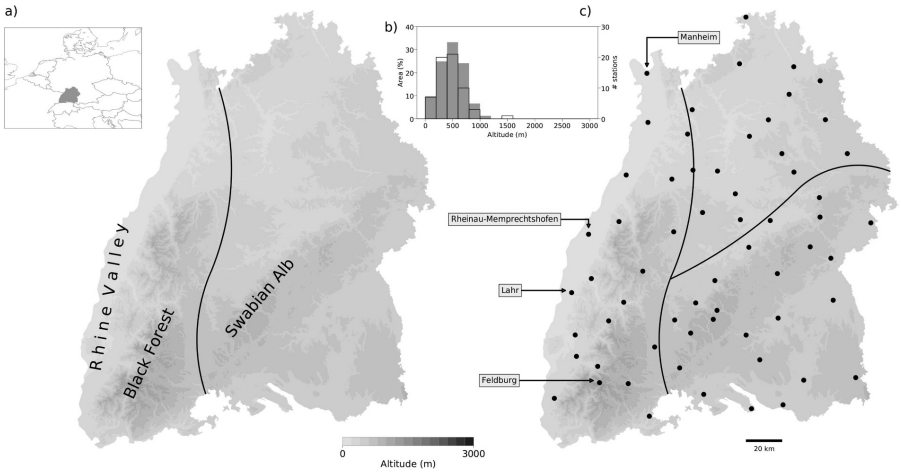

Fig. 2. As Fig. 1 for Baden-Württemberg Land.

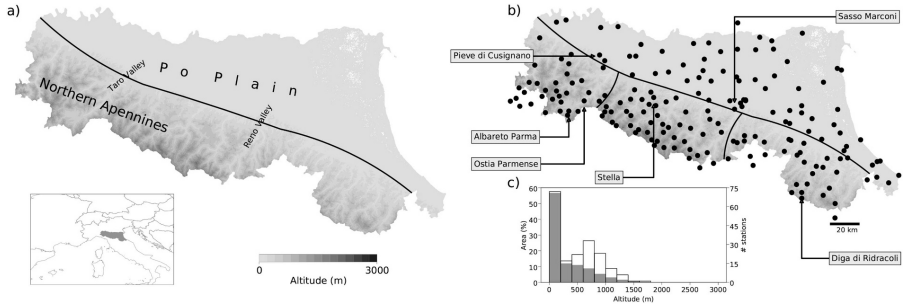

Fig. 3. As Fig. 1 for Emilia-Romagna region.

function is 0 at the coast line and ~1 when the Euclidean distance (*dist*) is equal or higher than a defined distance (*D*). Thereby, a higher weight is given to those stations that are closer to the coast. On the contrary, those that are far from the coast line have the same d_{coast} value as there is nearly no difference, in terms of sea influence, being 100 km or 150 km to the coast. For the present study, D was set to 100 km.

3.2. Residuals correction

Residuals correction consists on subtracting the difference between the estimated and observed values at weather stations points. That is, forcing predicted values at weather stations points to be the same as the observations. The correction is not restricted to station points, but a residuals correction field can be obtained through spatial interpolation, which generally improves the MLR performance (Stahl et al., 2006; Joly et al., 2011; Szymanowski and Kryza, 2012; Piazza et al., 2015). For this study, inverse of the 3D-distance interpolation was selected for its suitability for complex terrain regions (Frei, 2014; Lussana et al., 2018). The residuals are calculated at a certain point considering not only the horizontal distance between the point and all the stations, but also their altitude difference. Thereby, the higher the altitude difference between points the longer the distance is associated. For example, if two stations are at the same horizontal distance to a point, but one is at the same altitude and the other one is 1000 m higher, a greater distance will be assigned to the latter.

4. The MICA system

Meteorological field Interpolation based on Clustered data Analysis (MICA) scheme lays on dividing a weather station network in different clusters to separate areas of the same region that may be concurrently influenced by heterogeneous weather conditions.

4.1. Clustering stations

Defining clusters over a region allows classification of weather stations in different groups that may share more similar characteristics than considering one single group including all stations. In fact, splitting data into smaller groups may lead to an increase of linearity among the stations and therefore, be better modelled. Grouping stations can be performed using manifold criteria, for instance considering watersheds, different land use areas or climatic zones, but any way all these a priori classifications imply some degree of subjectivity. Thereby, similar to Frei (2014), Hiebl and Frei (2016) and Krähenmann et al. (2018) clusters were defined based on meteorological knowledge and aiming to enhance regional differentiation. In addition, two premises were followed: (1) the number of stations per cluster should be at least 20 for local model calibration (Joly et al., 2011; Szymanowski and Kryza, 2012); (2) the stations included in a cluster should be representative of the area covered by the cluster preventing unrealistic values (Szymanowski and Kryza, 2012; Kormos et al., 2018). Another option would be to consider clustering with unsupervised classification techniques, such as K-means (Lloyd, 1982), to obtain a first-guess station classification. This option may require further modification to fulfil the two aforementioned premises.

The strategy followed lead to use the clusters technique in two ways. Firstly, identifying two groups of stations that clearly present orographic contrast, such as separating a mountainous area from a flat area. Secondly, defining clusters capturing more local and finer characteristics of each region delimiting potential behaviour differences. For example, grouping stations located over extensive plain areas, deep and isolated valleys, or the same watershed or, alternatively, separating regions with different slope orientation, or different land use.

4.2. Cluster configuration selection process

The cluster configuration includes two aspects: the number of clusters (grouping of the stations) and the analysis method applied to each cluster. This configuration is selected using an iterative optimization process which requires two steps. Firstly, it calculates a MLR considering all the stations of the network (MLR_{Global}) as a preliminary analysis method and with the residuals of all of them extracts the Root Mean Square Error (RMSE) value ($RMSE_{Global}$) using the following equation:

$$RMSE = \sqrt{\frac{\sum_{k}^{N} \varepsilon_k^2}{N}} \tag{3}$$

where ε_k are the individual station errors and N the total number of stations. Secondly, MICA calculates a MLR for each cluster ($MLR_{Cluster}$) (Fig. 4) considering the stations included in it and also computes a RMSE value ($RMSE_{Cluster}$). MICA compares the RMSE values obtained by applying the $MLR_{Cluster}$ and the MLR_{Global} to the stations of the cluster. If the $RMSE_{Cluster}$ is lower than the $RMSE_{Global}$, then the $MLR_{Cluster}$ for that cluster is retained. Otherwise, the MLR_{Global} is chosen. Therefore, only those $MLR_{Cluster}$ that report an improvement against that of MLR_{Global} are selected. Finally, the resultant interpolated field may be composed solely of $MLR_{Cluster}$, MLR_{Global} or a combination of both as in Fig. 4d. The RMSE for that field is calculated accordingly.

Since MICA can be configured to use more than one number of clusters, the second step process is repeated for each number of clusters considered. Then, the number that minimizes the RMSE is selected. Using different number of clusters and configurations guarantees an adaptive character to the interpolation scheme appropriate for regions behaving diversely.

An example is presented in Table 1 and Fig. 4, where three

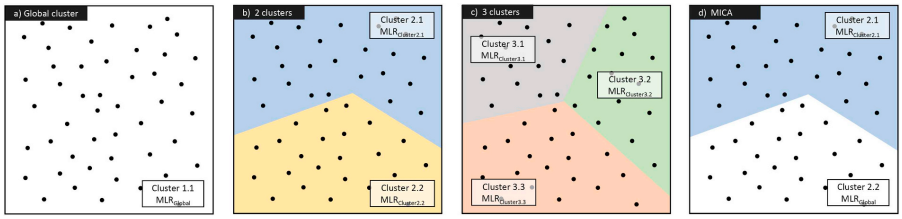

Fig. 4. Illustration of cluster configuration selection process. For each panel clusters and analysis methods are labelled: a) Global cluster (Cluster 1.1), b) 2 clusters (Cluster 2.1 and Cluster 2.2), c) 3 clusters (Cluster 3.1, Cluster 3.2 and Cluster 3.3) and d) final selection configuration by the MICA system (Cluster 2.1 with MLR Cluster 2.1, and Cluster 2.2 with MLR Global for the remaining area). Black dots represent weather stations and shaded areas the clusters domain.

configurations with different number of clusters are considered: 1 (Global cluster), 2 and 3. When two clusters are considered, the comparison between RMSE from individual clusters and Global resulted in the selection of the $MLR_{Cluster}$ for one cluster (Table 1, Cluster 2.1) and the default MLR_{Global} for the other (Table 1, Cluster 2.2). If three clusters are considered, the comparison shows that individual $MLR_{Cluster}$ models performed the best for two clusters (Table 1, Cluster 3.1, 3.2) and retained the MLR_{Global} for the third one (Table 1, Cluster 3.3). Then, if final RMSE values are compared, considering two clusters is the option that minimizes the error, which result from the combination of the default global interpolation (MLR_{Global}) and an individual $MLR_{Cluster}$ (Table 1 and Fig. 4d).

4.3. Merging of the clusters

Once the configuration of clusters that minimizes RMSE is selected, the field interpolation and cluster merging must be performed. For this purpose, the coefficients of the MLRs selected for each cluster are applied over their corresponding cluster area. Since the coefficients of these MLRs may differ depending on the cluster, a merging is required in order to disguise cluster boundaries avoiding abrupt transitions. This is achieved through the combination of a blurring and a weighting process. The blurring allows an overlapping region in contiguous clusters, similarly to Frei (2014), Hiebl and Frei (2016) and Fick and Hijmans (2017). The weighting is done using a Gaussian function that results in weights that range from 1 in the center of the cluster and 0 at the boundaries of it. Thereby, the values in the overlapping regions are calculated using the aforementioned weights (Hijmans et al., 2005).

5. Results and discussion

5.1. Cluster divisions

The strategy to define clusters in the three regions of study was similar. Firstly, identify two groups of stations that clearly present orographic contrast. Secondly, define clusters capturing more local and finer characteristics of each region delimiting potential behaviour differences. Thereby, two numbers of clusters was proposed in Catalonia. One separating the Pyrenees area from the rest of the region (Fig. 1a) and one focusing on specific characteristics, like separating the oriental and occidental part of the Pyrenees or isolating an area prone to the formation of fog banks (Metzger et al., 2005; Martínez et al., 2008) (Fig. 1b). In Baden-Württemberg, one division isolated the Rhine valley and the Black Forest from the rest of the region (Fig. 2a) and another implied further division like grouping stations in the Black Forest or the Swabian Alb with different orientation slopes (Fig. 2b) (Metzger et al., 2005; Schröder et al., 2006). Finally, weather stations in Emilia-Romagna were first grouped based on the orographic division between the Po plain and the Northern Appenines (Fig. 3a). The second group of clusters divided the mountainous regions in different areas including several valleys, but maintaining a single cluster the Po Plain (Fig. 3b) (Metzger et al., 2005; Nistor, 2016).

5.2. Cross-validation results

The quality of the interpolated meteorological fields was assessed by 'leave-one-out' cross validation technique (Hastie et al., 2001), hereafter CV. This technique consists of removing an observation of the dataset and predict the value at the point of that observation with a model, in this study a MLR model and an anomaly correction, built without it. Then, a prediction error for that point is obtained. Following this technique, the prediction errors were used to calculate the RMSE.

CV errors were calculated for air and dew point temperatures for 2017 for Catalonia and Baden-Württemberg at 30-min and hourly scale, respectively. For Emilia-Romagna only air temperature data at hourly resolution was considered. A comparison was made between interpolated fields obtained considering only one cluster (Global cluster) and the same approach but obtained through MICA, where several numbers of clusters were considered.

5.2.1. Catalonia

Considering the MICA scheme with two number of clusters, two and six, reported an improvement in air temperature interpolated fields. For 2017, median CV RMSE experienced a decrease of 0.2C from 1.4 °C to

Table 1
Example of cluster selection by RMSE cross-validation considering a maximum of three clusters. Number of clusters indicate the number of groups in which stations are classified. The Cluster ID is the label assigned to each cluster in a given number of clusters. The number of stations indicate how many stations are included in each cluster. RMSE columns show the error of the cluster, the same but considering the residuals obtained by a MLR calculated using all stations and the final error for the specific number of clusters.

Number of clusters	Cluster ID	Number of stations	$RMSE_{Cluster}$ (°C)	$RMSE_{Global}$ (°C)	$RMSE_{Final}$ (°C)
(Global cluster)	1.1	120	1.5	1.5	1.5
2	2.1	60	1.2 ✓	2.0	1.1
	2.2	60	1.2	1.0 ✓	
3	3.1	40	1.2 ✓	1.8	1.2
	3.2	40	1.2 ✓	1.5	
	3.3	40	2.1	1.3 ✓	

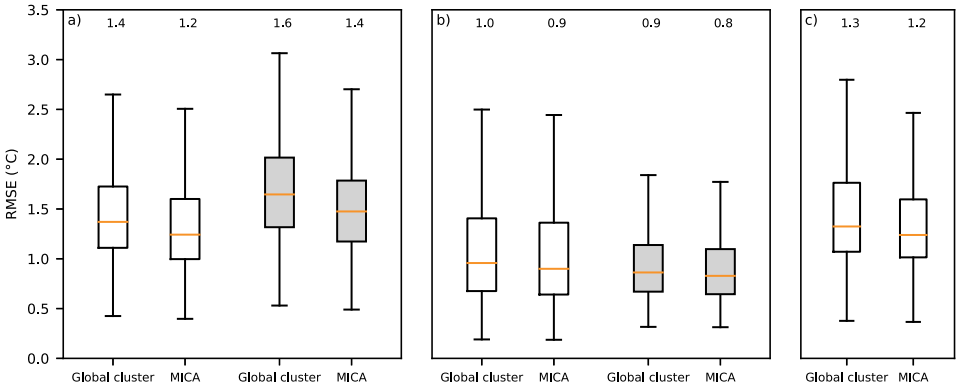

Fig. 5. Boxplots of cross-validation (CV) RMSE for air temperature (white) and dew point temperature (grey) considering the Global cluster and applying the MICA scheme. The median of each boxplot is indicated by the orange line and its value is shown at the top of the plot. The panels represents the three regions of study: a) Catalonia, b) Baden-Württemberg and c) Emilia-Romagna.

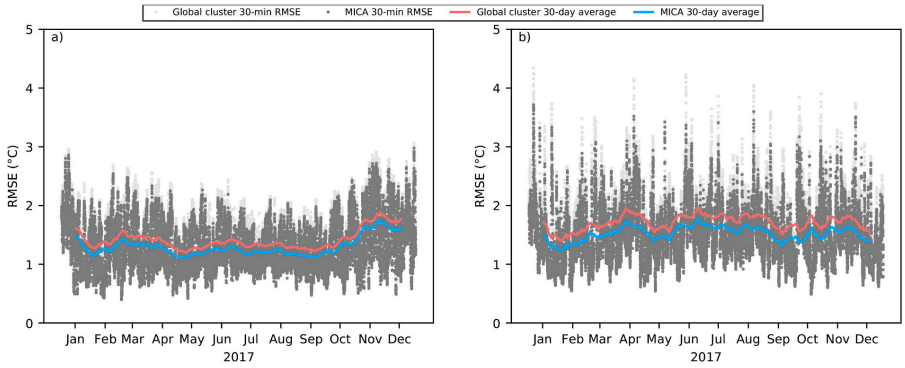

Fig. 6. Time series of the cross-validation (CV) RMSE errors in Catalonia at 30-min resolution for a) air temperature and b) dew point temperature. Light grey dots indicate Global cluster CV RMSE values and dark grey MICA CV RMSE values. Red line illustrates a running mean of 30-days using Global cluster and the same for the blue line considering MICA. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

1.2 °C. Although the inter-quartile remained unaltered at ~0.6 °C, a slight reduction of high error cases is reflected in the 0.2 °C difference between the superior whiskers (Fig. 5a). In Fig. 6a the time series of 30-min CV RMSE shows that highest uncertainties were concentrated in autumn and winter months with CV RMSE values up to 3.2 °C for the Global cluster and 3.1 °C for MICA. Spatial interpolation in autumn and winter months use to yield added difficulties, such as the formation of cold pools or thermal inversions, which increase the interpolation error (Lussana et al., 2018; Krähenmann et al., 2018). Mapping the CV RMSE over the area of study resulted in clearly separated regions: (1) plains and the metropolitan area of Barcelona, and (2) the Pyrenees range. The first of them concentrated the lowest uncertainty due to high station density and gentle orography. On the contrary, the second one presented greater CV RMSE average errors, expected from the combination of low station density and orographic complexity (Fick and Hijmans, 2017). However, most of the stations presenting high CV RMSE values are located at middle altitudes, between (600–1500 m ASL), such as Prades, Montserrat and PN Els Ports (Figs. 1, 7a and b). But at the same time, these stations present a clear CV RMSE reduction when MICA is applied: 17% for Prades, 20% for Montserrat and 31% for PN Els Ports (Fig. 7c). The reason for this decrease is that those stations are placed in isolated ranges and when clusters were considered, the weight of stations with similar altitudes increased in the MLR, being then better represented. Overall, individual station errors were reduced by at least 5% for 109 out of 183 stations and by at least 10% for 59 of them. However, the CV RMSE increased in 11 stations by at least 5% (Fig. 7c).

Regarding dew point temperature results, the median CV RMSE behaviour was similar to air temperature with a decrease of 0.2 °C, from 1.6 °C to 1.4 °C. However, in this case the interquartile decreased 0.1 °C and the highest errors (superior whiskers) dropped from 3.1 °C to 2.7 °C, reporting an improvement of 0.4 °C (Fig. 5a). Time series of CV RMSE for dew point temperature (Fig. 6b) did not fluctuate as air temperature (Fig. 6a) and there was not a clear season where errors were higher than others. Map of CV RMSE values showed analogous features as those observed for air temperature (Fig. 7d and e). Focusing on individual stations, 126 out of 183 presented a median CV RMSE reduction by at least 5%, in which 83 was at least 10% Fig. 7f). The highest reductions were found at stations located at middle altitude ranges, such as PN Els Ports (24%), Puig Sesolles (34%) and Prades (25%).

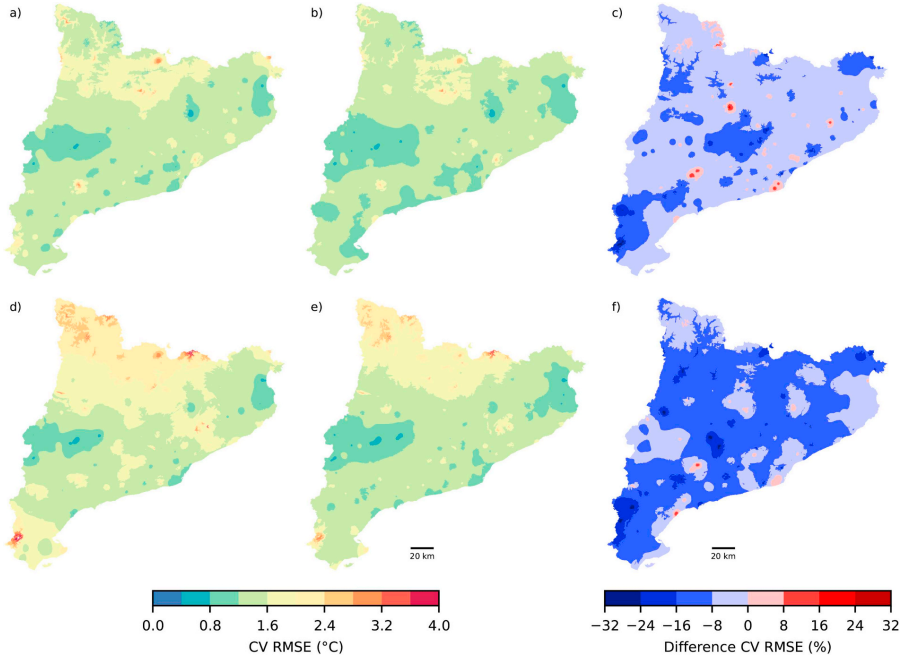

Fig. 7. Map of interpolated cross-validation (CV) RMSE values in Catalonia for air temperature (a, b and c) and dew point temperature (d, e, f). a) Illustrates the map of CV RMSE values obtained using the Global cluster and b) the same but using MICA scheme. c) Represents the percentage difference between MICA and Global cluster errors. The same for d), e) and f), respectively.

5.2.2. Baden-Württemberg

Similar results for air temperature and dew point temperature in terms of median CV RMSE were obtained either considering the Global cluster alone or applying the MICA scheme (Fig. 5b). The maximum number of clusters considered in this area was three due to the number and density of weather stations. The latter, together with a smooth orography may lead to a slight error reduction in the MICA scheme. The CV RMSE values were in agreement to those reported by Krähenmann et al. (2018) for approximately the same region. From an annual point of view the results were practically the same, however some differences arise when focusing on individual stations. CV median errors experienced a decrease of at least 5% for 18 out of 62 stations for air temperature, and for 19 stations for dew point temperature. One of the highest error reduction was reported in the Feldburg station (1493 m ASL) with 16% for air temperature and 23% for dew point temperature. This may be explained through the density of stations at that altitude, since the present case was the only station above 1000 m ASL. Therefore, when clusters were considered, the weight of stations with similar altitude in the regression model calculation was higher than if only one regression model would have been considered for all the German region.

5.2.3. Emilia-Romagna

Applying the MICA scheme for air temperature hourly measurements in Emilia-Romagna yielded a slight reduction of the median CV RMSE values, from 1.3 °C to 1.2 °C (Fig. 5c). However, a decrease of high error cases and the interquartile range is obtained with MICA, reflected in the whiskers and boxes of Fig. 5c. In a previous study, a Mean Absolute Error (MAE) of 1.3 °C was obtained at daily scale for minimum and maximum air temperature for the 1961–2010 period (Antolini et al., 2016), similarly to that obtained in this study for hourly observations, 1.0 °C. Focusing on individual station errors, 105 out of 188 showed an improvement of CV RMSE of at least 5%, being higher than 10% in 77 stations. However, 28 stations experienced an increase of the mean CV RMSE values by at least 5%. Still, the number of stations where a reduction of CV RMSE is present largely overweights those presenting an increment.

The general improvement is mostly reported in the mountainous areas, specially in the Northern Apennines. For example, the stations in the Taro Valley: Pieve Cusignano, Albareto Parma and Ostia Parmense experienced an error reduction of 43%, 35% and 28%, respectively (Figs. 3a, b and 8). In the Reno Valley the stations of Sasso Marconi and Porretta Terme the improvement was of 40% and 28%, respectively (Figs. 3a, b and 8).

Separating Emilia-Romagna in different clusters, being a region prone to thermal inversions and formation of fog banks (Mariani, 2009), yielded a better representation of stations in valley bottoms and allowed to separate two areas presenting dissimilar behaviour: the Po Plain and the Appenines.

5.3. Clusters and explanatory variables used

The MICA scheme selects the cluster configuration that minimizes the RMSE each time an interpolation is performed. Considering air temperature and dew point temperature observations recorded in 2017, 35,040 30-min analysis were done for Catalonia, 17,520 for Baden-Württemberg and 8760 Emilia-Romagna. In the Catalan region, the predominant number of clusters used was 6 either for air temperature

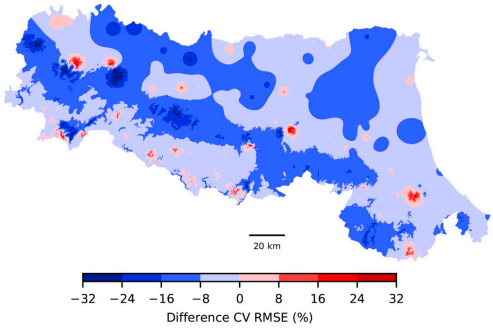

Fig. 8. Map of the cross-validation (CV) RMSE percentage difference between considering a single cluster and applying the MICA scheme in Emilia-Romagna.

or dew point temperature calculations. Due to complex orography, a tendency to select the highest number of clusters was found which may be explained by the linearity increase when station data is split into smaller groups. Furthermore, in Emilia-Romagna, a 4 cluster configuration was selected for 56% of the interpolation calculations. In this case, isolating the Po Plain from the Apennines mountainous area resulted effective to separate regions with clear behaviour differences. Regarding Baden-Württemberg, a 3 cluster configuration, the maximum considered for this region, was selected for analysis.

The most explanatory variable used in the three regions considered is altitude, which was selected in all the air temperature and dew point temperature interpolation calculations. In Catalonia, the second most selected variable was distance to coast, followed by latitude and longitude. On the contrary, in Emilia-Romagna, distance to coast was relegated as the fourth most frequently selected variable, overpassed by longitude and latitude. This may be related to a more gradual influence of distance to coast due to small orography changes (Fig. 3a) compared to Catalonia with the presence of the Coastal and Pre-Coastal mountainous ranges (Fig. 1a). Therefore, this gradual influence may better captured when longitude is considered, rather than an exponential function (Eq. 2). In Baden-Württemberg, latitude was selected over 60% either for air temperature and dew point temperature.

5.4. Case studies

In this section interpolated fields of air temperature and relative humidity are shown. Case studies were selected to exhibit MICA performance over concurrent different meteorological situations in a region with respect to considering a unique MLR model.

5.4.1. Persistent fog and strong inversion in Catalonia

Two examples with different meteorological conditions concurrently affecting Catalonia were selected. The first one illustrates a winter case of persistent fog on 21st February 2017 12.00 UTC. The NASA Worldview (Behnke et al., 2019) image (Fig. 9a) showed a fog area over the Central Depression that was consistent with relative humidity values (Fig. 9b and c). Relative humidity field was obtained through air temperature and dew point temperature interpolated fields. Both fields exhibited a decrease of CV RMSE, from 2.1 °C to 1.5 °C for air temperature and from 1.4 °C to 1.2 °C for dew point temperature when six clusters were considered (Fig. 9c) compared to global interpolation (Fig. 9b). Thereby, a more plausible relative humidity field was obtained, where the fog area boundaries that were hardly followed by the global model (Fig. 9b) were better defined by MICA (Fig. 9c). In this case, using physically informed clusters allowed to separate regions that clearly behaved differently due to their dissimilar meteorological conditions.

The second example is related to a strong inversion situation on 11th March 2017 06.00 UTC. Global interpolation was unable to capture the meteorological situation properly (Fig. 9d), but applying MICA with six clusters lead to a reduction of artificial local extremes produced by residuals interpolation across the region (Fig. 9e). This is reflected in the low correlation coefficient obtained for altitude when considering Global interpolation. Although the reduction of CV RMSE was from 2.5 °C to 2.2 °C, dividing the region in different clusters aided the MLR calculation with more linear groups. Focusing on individual stations, it is worth mentioning the error reduction in Montserrat from 5.4 °C to 2.9 °C, in PN Els Ports from 3.7 °C to 0.9 °C and in PN del Garraf from 3.5 °C to 2.4 °C.

5.4.2. Persistent fog in Baden-Württemberg

A fog case was selected in Baden-Württemberg to demonstrate MICA performance. The present case took place on 25th September 2017, when fog was present on the Rhine Valley in the western part of the Land (Fig. 10a, NASA Worldview (Behnke et al., 2019)). Fig. 10 illustrates relative humidity fields for 25th September 2017 11.00 UTC obtained considering a single MLR for the region (Fig. 10b) and applying the MICA scheme (Fig. 10c). From both fields it can be seen that the extension of the fog (Fig. 10a) coincided with maximum relative humidity values. Although the CV RMSE differences between MICA and a single MLR were marginal, they were better represented when the former was applied, since the relative humidity field below the fog area was nearly homogeneous (Fig. 10c). In addition, the expected strong gradient between fog and no fog areas was better represented using MICA. On the contrary, unrealistic "bull-eyes" were present in the field obtained using a single regression (Fig. 10b). If individual Rhine Valley station errors are analysed, a notable decrease can be found in air temperature and dew point temperature errors. Lahr station experienced a T_a improvement of 1.2 °C, from 3.1 °C to 1.9 °C, Rheinau-Memprechtshofen error decreased from 2.8 °C to 2.0 °C. Similarly, Manheim station error decreased 0.6 °C for T_a.

5.4.3. Strong inversion and fog bank in Emilia-Romagna

Two examples of the application of MICA scheme in Emilia-Romagna are presented in Fig. 11. The first one corresponds to a strong horizontal temperature gradient case due to the presence of a dense fog bank in the Po Plain on 23rd February 2017 (Fig. 11a and b). Isolating the Po Plain from the Apennines area using two clusters, allowed to decrease the CV RMSE from 1.9 °C to 1.4 °C. In addition, a better representation of the valleys in the Apennines was obtained, leading to a reduction of "bull-eyes" interpolation artifacts (Fig. 11c). Furthermore, a clearer separation between fog and no fog areas is obtained (Fig. 11b). Regarding the case on 16th December 2017, a strong inversion was present in the alpine area. In this case, similarly to the latter, the CV RMSE decreased from 3.2 °C to 2.5 °C when two clusters were used. Strong residuals in the Po Plain and the Apennines valleys were clearly reduced (Fig. 11c and d). It is worth mentioning the clearer representation of Taro and Reno valleys when MICA is applied (Fig. 11d) compared to global interpolation (Fig. 11c). Thereby, CV individual station errors in these valleys plummeted, for example: from 6.7 °C to 0.9 °C in Pieve di Cusignano, from 2.6 °C to 0.8 °C in La Stella and from 5.4 °C to 0.7 °C in Diga di Ridracoli.

6. Summary and conclusions

A spatial interpolation scheme based on clustering weather stations data, multiple linear regressions and anomaly corrections (MICA) was presented. The scheme was designed to work with hourly and sub-hourly time scales, which are generally highly variable and more affected by weather conditions than daily or monthly means.

Over the three regions considered, a reduction of CV RMSE was obtained. However, MICA exhibited a better performance on Catalonia

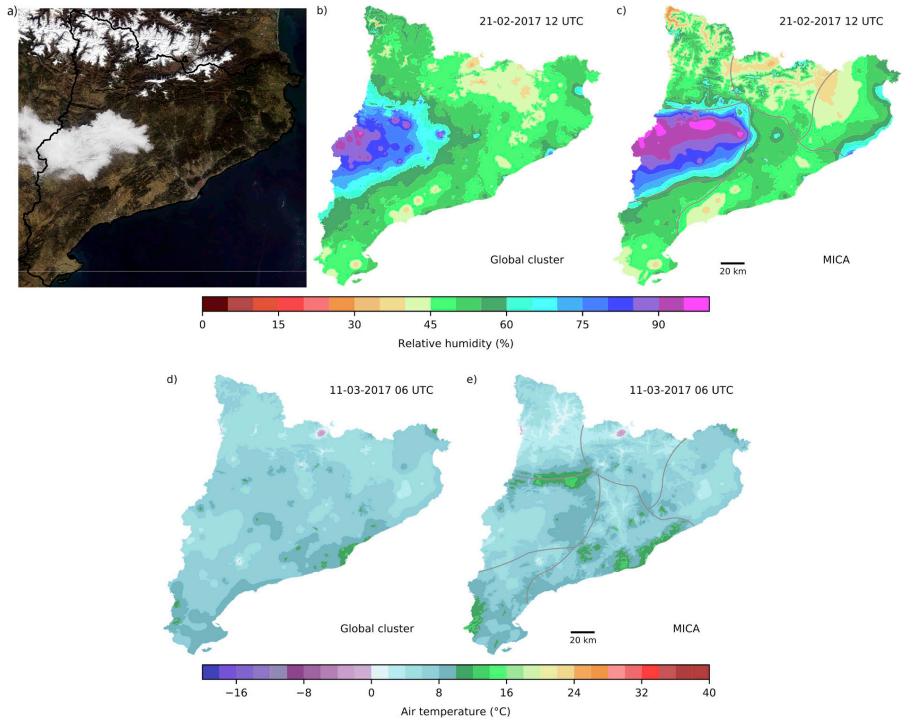

Fig. 9. Fog bank on 21 February 2017 12 UTC (a, b, c) and air temperature inversion on 11 March 2017 06 UTC (d, e) study cases in Catalonia. a) satellite image from NASA Worldview where fog over Central Depression can be easily seen. b) and c) are the relative humidity fields obtained using the Global cluster and MICA, respectively, for the fog case. d) and e) are the air temperature interpolated fields obtained using the Global cluster and MICA, respectively, for the inversion case.

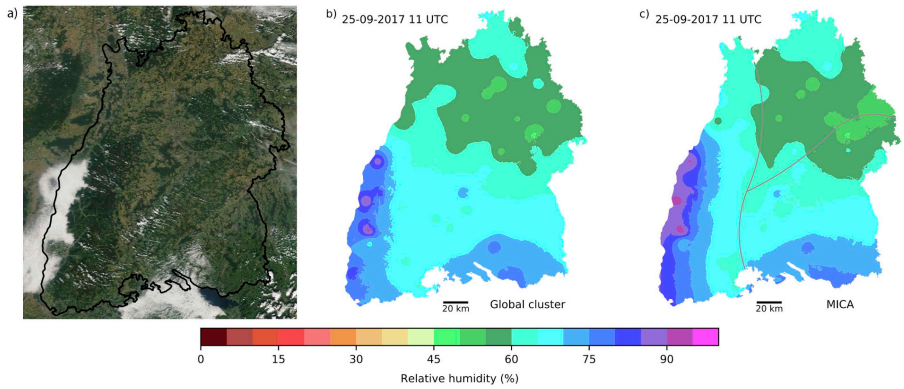

Fig. 10. Fog study case on 25th September 2017 11.00 UTC in Baden-Württemberg Land. a) Satellite image from NASA Worldview where fog over Rhine Valley can be easily seen. b) and c) are the relative humidity interpolated fields obtained using the Global cluster and MICA, respectively, through the air temperature and dew point temperature fields.

and Emilia-Romagna compared to Baden-Württemberg, which can be related to station density and orography complexity present in the former regions. Therefore, this suggests that MICA is suited mostly for complex terrain regions, but limited by weather station availability. However, it can also be applied to more gentle orographic regions where added value can be found during specific meteorological

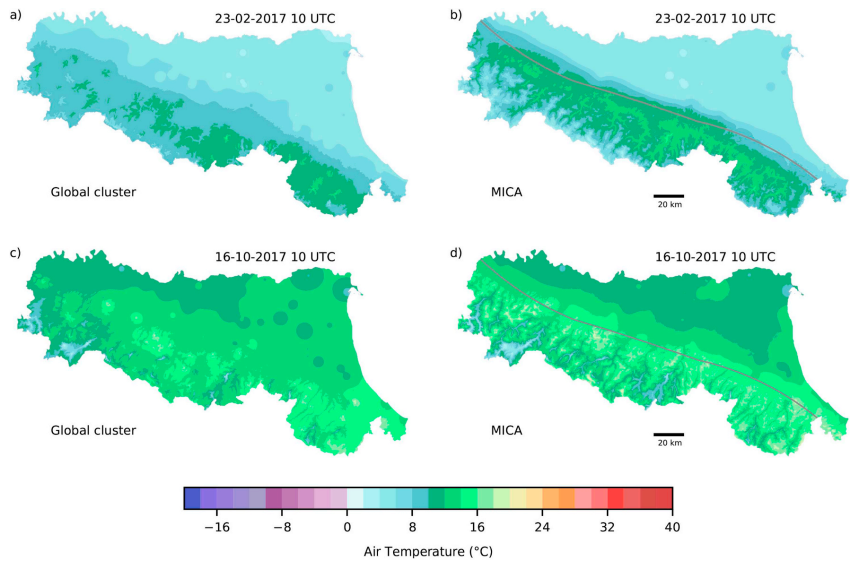

Fig. 11. Fog bank on 23 February 2017 10 UTC (a, b) and air temperature inversion on 16 October 2017 07 UTC (c, d) study cases in Emilia-Romagna. a) and b) are the air temperature fields obtained using the Global cluster and MICA, respectively, for the fog case. d) and e) analogously to a) and b) for the inversion case.

conditions, such as fog or air temperature inversions.

Clustering stations allowed to reduce middle and high altitude station errors, together with a better representation of complex orography features, such as valley bottoms, found in Catalonia and Emilia-Romagna. This was mainly due to the increase of middle and high altitude station weight on MLR models and also benefiting from a linearity increase when stations are grouped in smaller groups rather than considering them as a whole. In fact, in the three regions considered, a tendency to select configurations with the highest number of clusters was found.

Case studies showed more plausible air temperature and relative interpolated fields with a clear drop of RMSE, as the fog cases examples presented for Catalonia, Baden-Württemberg and Emilia-Romagna with more homogeneous fields under fog area. Regarding the strong inversion cases in Catalonia and Emilia-Romagna, apart from the RMSE reduction, a clear disguising of interpolation residuals was observed across the regions. These examples showed that MICA, with an informed cluster definition, was useful to isolate a part of a region prone to specific meteorological conditions, such as fog. At the same time, MICA exhibited its usefulness when the observations do not follow the expected trends, such as air temperature inversions. Still, MICA is based on linear regression calculations and poor correlations may be found at high-temporal scales or when interpolating other meteorological variables such as precipitation. The latter, due to its high spatial variability at high temporal resolution may require further information than time invariant explanatory variables (altitude, distance to coast) to obtain a successful interpolation, such as weather radar data. In these cases, geostatistical interpolation or methodologies relying on observations only should be considered.

The interpolation scheme presented in this study is available as an open-source library implemented in Python, pyMICA (Casellas et al., 2019).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by the Industrial Doctorate Projects framework, partly founded by the Government of Catalonia (Project DI-053/2017), performed under the framework of the HyMeX program and was carried out with partial support from the Spanish Projects CGL2015-65627-C3-2-R (MINECO/FEDER), CGL2016-81828-REDT/AEI and RTI2018-098693-B-C32 (MINECO). Thanks are given to ARPAE (Italy) and DWD-Climate Data Centre (Germany) for the data provided through respective open data services. We also acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov/), part of the NASA Earth Observing System Data and Information System (EOSDIS). Authors also acknowledge an anonymous reviewer for providing detailed and constructive comments that helped to improve the final form of this study.

References

Antolini, G., Auteri, L., Pavan, V., Tomei, F., Tomozeiu, R., Marletto, V., 2016. A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961–2010. Int. J. Climatol. 36, 1970–1986.

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

Bech, J., Pineda, N., Rigo, T., Aran, M., 2013. Remote sensing analysis of a Mediterranean thundersnow and low-altitude heavy snowfall event. Atmos. Res. 123, 305–322.

Behnke, J., Mitchell, A., Ramapriyan, H., 2019. NASAs earth observing data and information system–near-term challenges. Data Sci. J. 18.

Berndt, C., Haberlandt, U., 2018. Spatial interpolation of climate variables in Northern GermanyInfluence of temporal resolution and network density. J. Hydrol. 15, 184–202.

Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., Spinoni, J., 2014. High-resolution temperature climatology for Italy: Interpolation method intercomparison. Int. J.

Climatol. 34, 1278–1296. https://doi.org/10.1002/joc.3764.
Buck, A.L., 1981. New equations for computing vapor pressure and enhancement factor. J. Appl. Meteorol. 20, 1527–1532.
Casellas, E., Veciana, R., Bech, J., 2019. pyMICA: Meteorological Variable Interpolation Based on Clustered Data Analysis. URL. https://github.com/meteocat/pymica.
Cressie, N., 1990. The origins of kriging. Math. Geol. 22, 239–252.
Cressman, G.P., 1959. An operational objective analysis system. Mon. Weather Rev. 87, 367–374.
DWD Climate Data Center, 2018. Historical hourly station observations of 2m air temperature and humidity for Germany, version v006.
Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M., Stoffel, M., 2018. Estimating the snowfall limit in alpine and pre-alpine valleys: a local evaluation of operational approaches. Atmos. Res. 204, 136–148.
Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315.
Frei, C., 2014. Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. Int. J. Climatol. https://doi.org/10.1002/joc.3786.
Gjertsen, U., Ødegaard, V., 2005. The water phase of precipitationa comparison between observed, estimated and predicted values. Atmos. Res. 77, 218–231.
Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning. Springer Series in Statistics. New York, NY, USA : Springer New York Inc.
Hiebl, J., Frei, C., 2016. Daily temperature grids for Austria since 1961concept, creation and applicability. Theor. Appl. Climatol. 124, 161–178.
Hiebl, J., Auer, I., Böhm, R., Schöner, W., Maugeri, M., Lentini, G., Spinoni, J., Brunetti, M., Nanni, T., Tadić, M.P., et al., 2009. A high-resolution 1961–1990 monthly temperature climatology for the greater Alpine region. Meteorol. Z. 18, 507–530.
Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978.
Jarvis, A., Reuter, H., Nelson, A., Guevara, E., 2008. Hole-Filled Seamless SRTM Data V4: International Centre for Tropical Agriculture (CIAT). srtm.csi.cgiar.org.
Joly, D., Brossard, T., Cardot, H., Cavailhes, J., Hilal, M., Wavresky, P., 2011. Temperature interpolation based on local information: the example of France. Int. J. Climatol. https://doi.org/10.1002/joc.2220.
Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., Deutschländer, T., 2013. Monitoring of climate change in Germany–data, products and services of Germany's National Climate Data Centre. Adv. Sci. Res. 10, 99–106.
Kormos, P.R., Marks, D.G., Seyfried, M.S., Havens, S.C., Hedrick, A., Lohse, K.A., Sandusky, M., Kahl, A., Garen, D., 2018. 31 years of hourly spatially distributed air temperature, humidity, and precipitation amount and phase from Reynolds critical Zone Observatory. Earth Syst. Sci. Data 10, 1197–1205.
Krähenmann, S., Walter, A., Brienen, S., Imbery, F., Matzarakis, A., 2018. High-resolution grids of hourly meteorological variables for Germany. Theor. Appl. Climatol. 131, 899–926.
Kurtzman, D., Kadmon, R., 1999. Mapping of temperature variables in Israel: a comparison of different interpolation methods. Clim. Res. 13, 33–43.
Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., et al., 2005. Applied linear statistical models Volume 5. McGraw-Hill Irwin Boston.
Lanfredi, M., Coppola, R., D'Emilio, M., Imbrenda, V., Macchiato, M., Simoniello, T., 2015. A geostatistics-assisted approach to the deterministic approximation of climate data. Environ. Model. Softw. 66, 69–77.
Le Roux, R., De Resseguier, L., Corpetti, T., Jégou, N., Madelin, M., Van Leeuwen, C., Quénol, H., 2017. Comparison of two fine scale spatial models for mapping temperatures inside winegrowing areas. Agric. For. Meteorol. 247, 159–169.
Liuzzo, L., Bono, E., Sammartano, V., Freni, G., 2017. Long-term temperature changes in Sicily, Southern Italy. Atmos. Res. 198, 44–55.
Llabrés-Brustenga, A., Rius, A., Rodrguez-Solà, R., Casas-Castillo, M.C., Redaño, A., 2019. Quality control process of the daily rainfall series available in Catalonia from 1855 to the present. Theor. Appl. Climatol. 1–15.
Lloyd, S., 1982. Least squares quantization in pcm. IEEE Trans. Inf. Theory 28, 129–137.
Lowe, D., Ebi, K.L., Forsberg, B., 2011. Heatwave Early warning systems and adaptation advice to reduce human health consequences of heatwaves. Int. J. Environ. Res. Public Health 8, 4623–4648. http://www.mdpi.com/1660-4601/8/12/4623 https://doi.org/10.3390/ijerph8124623.
Lu, G.Y., Wong, D.W., 2008. An adaptive inverse-distance weighting spatial interpolation technique. Comput. Geosci. 34, 1044–1055.
Lussana, C., Tveito, O.E., Uboldi, F., 2018. Three-dimensional spatial interpolation of 2 m temperature over Norway. Q. J. R. Meteorol. Soc. https://doi.org/10.1002/qj.3208.
Mariani, L., 2009. Fog in the Po valley: some meteo-climatic aspects. Ital. J. Agrometeorol. 14, 35–44.
Martínez, D., Cuxart, J., Cunillera, J., 2008. Conditioned climatology for stably stratified nights in the Lleida area. Tethys 5, 13–24.
Martn-Vide, J., Raso Nadal, J., Morera Palacios, A., 2008. Atles climàtic de Catalunya, perode 1961–1990: termopluviometria. *Institut Cartogràfic de Catalunya i Servei Meteorològic de Catalunya*. Generalitat de Catalunya, Barcelona, Spain.
Mathbout, S., Lopez-Bustins, J.A., Martin-Vide, J., Bech, J., Rodrigo, F.S., 2018. Spatial and temporal analysis of drought variability at several time scales in Syria during 1961–2012. Atmos. Res. 200, 153–168.
Metzger, M.J., Bunce, R.G.H., Jongman, R.H., Mücher, C.A., Watkins, J.W., 2005. A climatic stratification of the environment of Europe. Glob. Ecol. Biogeogr. 14, 549–563.
Nastos, P.T., Politi, N., Kapsomenakis, J., 2013. Spatial and temporal variability of the Aridity Index in Greece. Atmos. Res. 119, 140–152.
Ninyerola, M., Pons, X., Roure, J.M., 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. Int. J. Climatol. 20, 1823–1841.
Ninyerola, M., Pons, X., Roure, J.M., 2007. Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. Int. J. Climatol. 27, 1231–1242.
Nistor, M.M., 2016. Spatial distribution of climate indices in the Emilia-Romagna region. Meteorol. Appl. 23, 304–313.
Pagès, M., Pepin, N., Miró, J., 2017. Measurement and modelling of temperature cold pools in the Cerdanya valley (Pyrenees), Spain. Meteorol. Appl. 24, 290–302.
Piazza, A.D., Conti, F.L., Viola, F., Eccel, E., Noto, L.V., 2015. Comparative analysis of spatial interpolation methods in the Mediterranean area: application to temperature in Sicily. Water 7, 1866–1888.
Rogelis, M. C., & Werner, M. G. F. (2013). Spatial interpolation for real-time rainfall field estimation in areas with complex topography. J. Hydrometeorol., 14, 85–104. URL: https://doi.org/10.1175/JHM-D-11-0150.1. doi: https://doi.org/10.1175/JHM-D-11-0150.1.
Schröder, W., Schmidt, G., Hasenclever, J., 2006. Geostatistical analysis of data on air temperature and plant phenology from Baden-Württemberg (Germany) as a basis for regional scaled models of climate change. Environ. Monit. Assess. 120, 27–43.
Serra, A., Merc, B., Vendrell, R., 2016. Automatic weather stations network (XEMA) of the Meteorological Service of Catalonia (SMC). In *WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (CIMO TECO 2016):* Ensuring Sustained High-Quality Meteorological Observations from Sea, Land and Upper Atmosphere in a Changing World. WMO. URL. https://library.wmo.int/index.php?lvl=notice_display&id=19676#~~=~~EWdHvZ7m70.
Sideris, I., Gabella, M., Erdin, R., Germann, U., 2014. Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. Q. J. R. Meteorol. Soc. 140, 1097–1111.
Stahl, K., Moore, R., Floyer, J., Asplin, M., McKendry, I., 2006. Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. Agric. For. Meteorol. 139, 224–236.
Steinacker, R., Ratheiser, M., Bica, B., Chimani, B., Dorninger, M., Gepp, W., Lotteraner, C., Schneider, S., Tschannett, S., 2006. A mesoscale data analysis and downscaling method over complex terrain. Mon. Weather Rev. 134, 2758–2771.
Szymanowski, M., Kryza, M., 2012. Local regression models for spatial interpolation of urban heat islandan example from Wrocław, SW Poland. Theor. Appl. Climatol. 108, 53–71.
Szymanowski, M., Kryza, M., Spallek, W., 2013. Regression-based air temperature spatial prediction models: an example from Poland. Meteorol. Z. 22, 577–585.
Velasco-Forero, C.A., Sempere-Torres, D., Cassiraga, E.F., Gómez-Hernández, J.J., 2009. A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. Adv. Water Resour. 32, 986–1002.
Viggiano, M., Busetto, L., Cimini, D., Di Paola, F., Geraldi, E., Ranghetti, L., Ricciardelli, E., Romano, F., 2019. A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements. Agric. For. Meteorol. 276, 107590.
Webb, M.A., Hall, A., Kidd, D., Minansy, B., 2016. Local-scale spatial modelling for interpolating climatic temperature variables to predict agricultural plant suitability. Theor. Appl. Climatol. 124, 1145–1165.

Chapter 3

Diagnosing the precipitation phase

3.1 Surface precipitation phase discrimination in complex terrain

3.1.1 Summary

This chapter presents the validation and implementation of a surface precipitation phase (SPP) diagnose product in Catalonia to classify precipitation among rain, sleet, and snow. Estimating precipitation phase at ground level is fundamental in numerous meteorological and hydrological applications. Several studies have addressed this issue considering SPP discrimination schemes based on establishing thresholds to surface meteorological variable observations, such as air temperature or dew point temperature. For example, if precipitation is observed above a certain air temperature value, it is classified as rain; otherwise, as snow.

In this study, the evaluation of eight SPP schemes based on different meteorological variables was conducted considering 7,702 precipitation phase observations. These were collected between 2010 and 2018 from different sources, but unlike other studies most of the precipitation phase observations did not concurrently report meteorological variable observations such as air temperature and relative humidity. This could be seen as an inconvenience, but at the same time it was an opportunity to test the application of SPP schemes at catchment and regional scales using interpolated fields of meteorological variables rather than observations. Therefore, the performance evaluation of SPP schemes was conducted by comparing precipitation phase observations against interpolated fields of air temperature and dew point temperature, among other meteorological

variables. The evaluation of SPP schemes consisted in obtaining optimised thresholds for each scheme considering all observations. Results indicate that schemes based on meteorological variables including air saturation conditions, such as wet bulb temperature, exhibited a better performance compared to those considering air temperature alone, which is a widely used SPP scheme. Once the best thresholds for each SPP scheme were found, their application to two low altitude snowfall events in Catalonia was tested. In addition, optimised thresholds for each event were also calculated. This revealed certain variability among the thresholds set and SPP scheme performance, not only between the two considered events, but inside the same event. Therefore, taking into account different SPP schemes or a range of discrimination thresholds was suggested when monitoring snowfall events. Finally, citizen science and crowd sourced observations were also collected and evaluated proving their potential and limitations. Both provided important information since they are mostly concentrated in populated areas and close to critical infrastructures complementing other data from conventional observational networks.

Despite the notable performance of the SPP schemes, two limitations have been identified: (i) exclusion of vertical temperature profiles, which play a key role on determining precipitation phase at surface level, and (ii) no single SPP scheme is able to capture the precise nature of precipitation phase discrimination, which could be overcome by combining different SPP schemes or adjusting the thresholds of a single SPP scheme depending on the event.

3.1.2 Article

Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Miró, J. R., & Sairouni, A. (2021). Surface precipitation phase discrimination in complex terrain. *Journal of Hydrology*, 592, 125780.

Journal of Hydrology 592 (2021) 125780

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Surface precipitation phase discrimination in complex terrain

Enric Casellas [a,b,*], Joan Bech [b], Roger Veciana [a], Nicolau Pineda [a], Tomeu Rigo [a], Josep Ramon Miró [a], Abdel Sairouni [a]

[a] Meteorological Service of Catalonia, Barcelona, Spain
[b] Department of Applied Physics – Meteorology, University of Barcelona, Barcelona, Spain

ARTICLE INFO

Keywords:
Precipitation phase
Threshold temperature
Wet-bulb temperature
Weather radar
Citizen science

ABSTRACT

Surface precipitation phase (SPP) discrimination at the ground level (rain or snow) is a key step in numerous meteorological and hydrological applications. Previous studies have undertaken this by comparing surface observations, such as air temperature, relative humidity and wet-bulb temperature, with concurrent present weather observations of the precipitation phase to derive thresholds for discrimination purposes. The first objective of this study was to examine schemes for precipitation phase discrimination at the ground level, using data interpolated from a network of automatic weather stations covering an area of complex terrain. The second objective was to combine the SPP interpolated fields with precipitation estimates from single-polarisation weather radar, which provide precipitation occurrence information but not precipitation phase type, to obtain a real-time SPP product. Finally, the third objective was to evaluate the role of citizen science and crowd sourced observations in the monitoring of snowfall events with SPP schemes. Results from nine cold seasons (Oct-May) indicated that out of the seven SPP schemes tested against 7,702 quality-controlled present weather observations in Catalonia (NE Spain), those including information on air saturation conditions provided the best results. A wet-bulb temperature threshold of 0.7 °C produced the best discrimination for snow vs no snow, with a Pierce skill score of 0.77. Finally, the SPP product was used with two case studies, demonstrating its added value and pending challenges for real-time applications.

1. Introduction

Surface precipitation phase (SPP) discrimination at the ground level (rain, snow or a mixture of both, hereafter referred to as sleet) is fundamental in many meteorological and hydrological applications such as the surveillance of heavy rainfall, precipitation-runoff modelling, and snow avalanche forecasting. Several studies have undertaken SPP discrimination using different methodologies and predictors (Harpold et al., 2017). These strategies vary in the type of information used to classify the precipitation phase. Some use surface observations (e.g. Gjertsen and Ødegaard, 2005; Dai, 2008), others are based on additional atmospheric characteristics such as microphysics schemes from numerical weather prediction models (e.g. Fernández-González et al., 2015; Tapiador et al., 2019) or vertical temperature profiles (e.g. Bourgouin, 2000; Wandishin et al., 2005), while some include data from dual-polarisation weather radars (e.g. Ryzhkov and Zrnic, 1998; Schuur et al., 2012).

The methodology involving surface observations, used in this study, has produced notable results across different regions (Gjertsen and Ødegaard, 2005; Chen et al., 2014; Froidurot et al., 2014) and involve the use of atmospheric variables such as air temperature or dew point temperature to classify the precipitation phase by establishing a predetermined threshold. For example, if air temperature exceeds a selected threshold, the precipitation is labelled as rain; otherwise it is classified as snow. A preliminary air temperature threshold could be 0.0 °C, since it corresponds to the freezing point, but snow is usually reported at warmer air temperatures. Two physical effects, melting and evaporation, explain the presence of snow at warmer temperatures (Harpold et al., 2017). When snowflakes fall through a layer with saturated air conditions where the temperature is above freezing, the melting of the snowflakes extracts heat from their surroundings, thereby cooling the environment (Kain et al., 2000). However, under unsaturated air conditions, the dominant effect is evaporation. In this case, there is a flux of latent heat to the snowflake from the surrounding air, cooling the temperature of the air adjacent to the snowflake (Kain et al., 2000). When evaporation ceases, the temperature of the snowflake is that of the

* Corresponding author at: Department of Applied Physics – Meteorology, University of Barcelona, Barcelona, Spain.
E-mail address: enric.casellas@meteo.ub.edu (E. Casellas).

https://doi.org/10.1016/j.jhydrol.2020.125780
Received 6 March 2020; Received in revised form 12 November 2020; Accepted 15 November 2020
Available online 4 December 2020

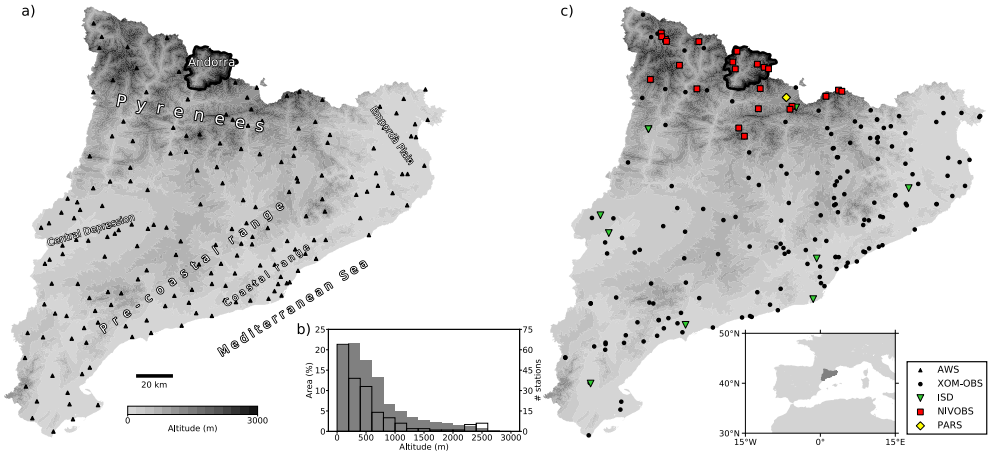

Fig. 1. Map of Catalonia and Andorra (black line) showing (a) the main orographic features and the automatic weather stations used in this study, (b) the distribution of the altitudes of the automatic weather stations (non-shaded histogram) and the region's relief (% of the area of Catalonia; dark shaded histogram), and (c) locations of the quality-controlled present weather observations and the position of the region in Europe.

wet-bulb temperature (Lumb, 1961). In both cases, the probability of solid precipitation particles maintaining their phase throughout their path to the surface increases. For this reason, the following variables that take into account the moisture content of air have also been considered when establishing thresholds to discriminate the precipitation phase: relative humidity together with air temperature (Koistinen and Saltikoff, 1998), the dew point temperature (Feiccabrino and Lundberg, 2009), and the wet-bulb temperature (Froidurot et al., 2014).

These schemes, especially those based on air temperature, are widely used in hydrological models to calculate the fraction of precipitation that falls as snow and rain, such as the Penn State Integrated Hydrologic Modeling System (PHIM) (Jepsen et al., 2016), the Distributed Hydrological Model (DHM) (Shrestha et al., 2014) and the Hydrologiska Byråns Vattenbalansavdelnin (HBV) (Sikorska and Seibert, 2018). In addition to its application in hydrological models, SPP information can be merged with data from weather radars when data from dual-polarisation radars are not available. These combined data indicate the extent, intensity and spatial distribution of the precipitation together with its phase, becoming a key tool in the monitoring of snowfall events (Koistinen and Saltikoff, 1998; Schmid and Mathis, 2004; Gjertsen and Ødegaard, 2005; Saltikoff, 2012). This diagnostic tool can provide critical information to meteorological services and decision makers, especially in regions where there is occasional snowfall at low altitudes, such as Catalonia (NE Spain) (Fig. 1). This NW Mediterranean region is characterised by complex terrain and areas with high population densities, where low-altitude snowfall can have a high socio-economic impact. For instance, the events of March 2010 (Bech et al., 2013; Llasat et al., 2014), February 2015 and February 2018 led to 58,000, 1,090 and 3,608 phone calls, respectively, from the general public to the emergency services due to power outages, trapped vehicles, collapsed roofs, and public transport problems, among other reasons. Therefore, this study aimed to determine the best schemes and thresholds for precipitation phase discrimination in order to provide an SPP product that can improve emergency management, as well as the diagnostic and surveillance tasks of winter weather events.

Different sources of surface-based precipitation phase observations were used to achieve the following objectives: (1) performance analysis of the different schemes and thresholds for precipitation phase discrimination at the surface level with quality-controlled surface-based observations; (2) evaluation of the role of citizen science and social network observations in the monitoring of snow events; and (3) implementation of a real-time SPP product providing the necessary tools for diagnosing and surveillance tasks. The results are shown for the study region, Catalonia (NE Spain), using data from the Meteorological Service of Catalonia (SMC). The methodology implemented can be applied to other regions that have surface observations and radar data available, using the open-source software developed for this purpose, pyPROS (Casellas et al., 2019).

The rest of the paper is organised as follows. The study region and the data employed are described in Section 2, while the precipitation phase discrimination methodologies and the validation strategy are presented in Section 3. Section 4 provides the validation results of the different sources of precipitation phase observations, including those from citizen science, with threshold-based schemes using information on air moisture content outperforming the rest. Two low-altitude snowfall events were examined in depth regarding the performance of the SPP schemes, from which the following question arose: should the threshold parameters be adjusted at the event scale? This is discussed in Section 5, which also describes the implementation of a real-time SPP product as well as considerations for future studies.

2. Data

2.1. Study site

The study region, Catalonia, is located in the NE Iberian Peninsula (SW Europe) and is bordered by the Pyrenees mountain range to the N, the Ebre Valley to the W and the Mediterranean Sea to the E and S (Fig. 1). The region has an area of ~32,000 km^2 and presents marked orographic variability, which yields strong contrasts of precipitation and temperature. The highest mountains exceed 3,000 m above sea level (a.s.l.) in the north, contrasting with the lower coastal mountain ranges (500~1,000 m a.s.l.) and some inland plains. Snowfalls occur regularly from late autumn to early spring, concentrating mainly in the mountainous regions. However, they are occasionally reported to occur at lower altitudes, below 700 m a.s.l., where the socio-economic impact may be high due to vulnerable infrastructures and the high population densities, e.g., cities containing up to 250,000 inhabitants and the

Table 1
Total number of precipitation phase observations from different sources and, in the parentheses, the observations used for the verification that are within the air temperature range [−7.5 °C to 5 °C]. NIVOBS, nivo-meteorological observations; XOM-OBS and XOM-SPO, Xarxa d'Observadors Meteorològics observers and spotters, respectively; ISD, Integrated Surface Database; PARS, Parsivel disdrometer.

Source	Period	Rain		Sleet		Snow	
NIVOBS	2010–2014	244	(200)	31	(30)	900	(769)
XOM-OBS	2010–2018	249955	(2640)	331	(314)	1769	(1675)
ISD	2010–2018	28719	(1292)	51	(34)	315	(304)
PARS	2016–2018	519	(203)	42	(37)	267	(204)
XOM-SPO	2010–2018	2864	(107)	602	(556)	3680	(3658)
Twitter	2018	24	(0)	33	(0)	1211	(1188)

Barcelona metropolitan area containing ~5 million inhabitants (Bech et al., 2013; Llasat et al., 2014).

2.2. Surface observations

2.2.1. Temperature and humidity fields

The methodologies used in this study to discriminate between rain, sleet and snow (Section 3.1) require surface meteorological data. The surface observations of air temperature (T_a) and dew point temperature (T_d) were obtained from 185 automatic weather stations (AWS, Fig. 1a) of the Xarxa d'Estacions Meteorològiques Automàtiques (XEMA) network, which have an irregular spatial distribution and a 30-min temporal resolution (Serra et al., 2016) that reports the mean value for each 30-min period (0 to 30 min past the hour, and 30 min past the hour to 0). Most of the XEMA stations (160) are located in flat areas. As the altitude increases, the density of the AWS decreases, with 12 stations located at medium altitudes and 12 in high mountainous areas (>2,000 m a.s.l., Fig. 1a, b). However, the precipitation phase was not reported at the locations of the AWS. Therefore, it was necessary to interpolate temperature and humidity fields to estimate their values. This was achieved using a spatial interpolation methodology that is based on linear regression with anomaly correction and clustering, using both discrete meteorological observations and a digital terrain model (Casellas et al., 2020). Interpolated fields were obtained with a spatial and temporal resolution of 250 m and 30 min, respectively, and included the following meteorological variables: T_a, T_d, relative humidity (RH) and wet-bulb temperature (T_w). RH was derived from T_a and T_d (Lawrence, 2005) and T_w from T_a, T_d, and air pressure (Sadeghi et al., 2013), since altitude differences, although small, yield variations in T_w values.

2.2.2. Weather radar data

Weather radar observations were obtained from the Xarxa de Radars Meteorològics (XRAD), which consists of four single-polarisation C-band Doppler weather radars located at specific locations to minimise radar beam blockage by orographic features (Bech et al., 2003; Trapero et al., 2009). The product used in this study is the lowest height of the composite reflectivity of the four weather radars obtained after performing a number of automatic quality control procedures to optimise the accuracy of the surface precipitation mass estimates (Sánchez-Diezma et al., 2002; Bech et al., 2005; Franco et al., 2006; Altube et al., 2015). The product has a temporal resolution of 6 min and a spatial resolution of 1 km.

2.2.3. Precipitation phase

Present weather observations of the precipitation phase (rain, sleet and snow) were used in this study. Available sources included automatic instruments, such as disdrometers, and manual reports from observers (trained meteorological staff, meteorological spotters, and social networks). The quality of the source was determined by assessing the following: location uncertainty, temporal uncertainty, and the quality control procedures applied. The number of observations and the available period of each source are presented in Table 1, including NIVOBS, XOM-OBS, XOM-SPO, ISD, PARS and Twitter, which are defined below.

NIVOBS (Xarxa d'Observadors Nivometeorològics de Catalunya) provides at least one present weather observation per day as well as other measurements such as the cloud base height, snow drift and fresh snow accumulation. Stations are located across the Pyrenees range (Fig. 1c, red squares) and their data are quality controlled by the Cartographic and Geological Institute of Catalonia (ICGC). Data are presented according to the international NIMET code, which is the standard codification used for nivo-meteorological observations in the Pyrenees mountains (Andorra, France and Spain) resembling that of SYNOP (Gavaldà et al., 2014; Apodaka et al., 2018).

XOM (Xarxa d'Observadors Meteorològics) consists of 233 trained meteorological volunteers, who are grouped as observers (XOM-OBS, 137 members) or spotters (XOM-SPO, 94 members) (Ripoll et al., 2016). The observers report different meteorological observations, including present weather, from manual weather stations every day at 07:00 UTC. The spotters provide information about in situ relevant meteorological phenomena (not necessarily always at a fixed location), such as snow, hail, or fog. The observers and spotters, although part of XOM, are considered separately (as explained in Section 3.2) since the locations of the observers are fixed (Fig. 1c, black dots), whereas those of the spotters may vary, with only the town centre provided in their reports. In addition, information from the XOM-OBS is quality controlled with daily, monthly and annual protocols by SMC staff, ensuring consistency and comparing against the meteorological situation of that moment with remote sensing data (satellite and weather radar products) and information from the XEMA network.

ISD stands for Integrated Surface Database (Smith et al., 2011) and compiles hourly airport reports (METARs) and synoptic surface observations (SYNOPs) from more than 20,000 stations around the globe. The present study included data produced by the Spanish Meteorological Agency (AEMET) from eight stations between 2010 and 2018, five from the Catalan airports (Barcelona-El Prat, Sabadell, Reus, Girona-Costa Brava and Lleida-Alguaire) and three from manned observatories (Talarn, Tortosa and Lleida) (see Fig. 1c, green triangles). ISD data were filtered according to the quality of the observation parameter provided in the datasets (Lott, 2004). Finally, PARS stands for Parsivel observations. Parsivel is a laser disdrometer that provides not only automated present weather precipitation phase records, but also the ranges of the precipitation particle size and fall speed (Löffler-Mang and Blahak, 2001; Bloemink and Lanzinger, 2005). The University of Barcelona has one unit installed (Gonzalez et al., 2019) at an altitude of 1,100 m a.s.l. in the Pyrenees (Fig. 1c, yellow diamond). Parsivel data were filtered with a sensor status parameter indicating the quality of the observation.

Twitter, an online social network service, was also used. As can be seen in Table 1, the amount of reports or "tweets" regarding snow largely exceed those for rain and sleet, mainly because low-altitude snowfall is a relatively unusual event in the region. Two types of tweets were considered: reports with a geolocation and those containing the name of the town or location (a ski resort, a neighbourhood or the name of a hotel). Both types were processed to assign the coordinates of the indicated location to each observation.

NIVOBS, Twitter and XOM observations correspond mostly to daytime hours while ISD and PARS are reported every hour. Considering the latter two sources no diurnal cycle of precipitation phase was found in our dataset. Based on the location and temporal uncertainty as well as

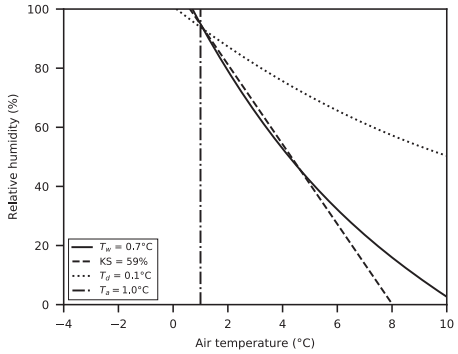

Fig. 2. Methodologies for precipitation phase discrimination presented in the relative humidity vs air temperature plane, including the air temperature (T_a), dew point temperature (T_d), wet-bulb temperature (T_w) and Koistinen–Saltikoff (KS) schemes.

Table 2
Three strategies for surface precipitation phase discrimination. X can be replaced by any of the following meteorological variables: air temperature (T_a), dew point temperature (T_d) or wet-bulb temperature (T_w). $p(snow)$ refers to the logistic equation (Eq. 1).

Method	Rain	Sleet	Snow
Single threshold	$X > X_{th}$	–	$X <= X_{th}$
Double threshold	$X >= X_{th_{rain}}$	$X_{th_{snow}} < X < X_{th_{rain}}$	$X <= X_{th_{snow}}$
Koistinen–Saltikoff	$p(snow) <= \%_{rain}$	$\%_{rain} < p(snow) < \%_{snow}$	$p(snow) >= \%_{snow}$

the quality control procedures, data from the NIVOBS, XOM-OBS, ISD and PARS were considered to be of high quality in this study.

3. Methodology

3.1. Precipitation phase classification

Three types of classification strategies based on surface meteorological information were used, depending on how the precipitation phase discrimination was performed: (1) one using a single threshold for physical variables such as air temperature (Liu, 2008); (2) another approach using two thresholds for physical variables (Liu et al., 2018); and (3) one involving a logistic formula to derive probabilities (Koistinen and Saltikoff, 1998), on which thresholds were also applied.

The single threshold (ST) approach is the simplest one and classifies precipitation into two types, rain and snow. For example, when considering the air temperature, if this temperature is cooler (equal or warmer) than the threshold, the precipitation is labelled as snow (rain) (Fig. 2). A similar strategy is followed by the dual threshold (DT) approach, but with three precipitation phases classified (rain, snow and sleet). This procedure yields more realistic results than the single threshold approach, since a diffuse boundary is expected between rain and snow. Finally, the third methodology considers results from an empirical formulation that takes into account T_a and RH, indicating the probability that the precipitation is snow ($p(snow)$), defined in Eq. (1) as follows:

$$p(snow) = 1 - \frac{1}{1 + e^{(22 - 2.7 \times T_a - 0.2 \times RH)}} \quad (1)$$

This equation was originally developed in Finland (Koistinen and Saltikoff (1998); DT-KS hereafter) to determine the probability of rain ($p(rain) = 1 - p(snow)$), and was later applied and tested in other regions such as Norway (Gjertsen and Ødegaard, 2005), the Swiss Alps (Froidurot et al., 2014), Spain (Bech et al., 2014) and China (Chen et al., 2014). For practical reasons, it is formulated here as the probability of the precipitation being snow at a given air temperature and relative humidity. To account for three different precipitation phases, two probability thresholds were set, initially 33% and 66% (Koistinen and Saltikoff, 1998). The three approaches for precipitation phase discrimination are summarised in Table 2.

3.2. Adjustment and validation strategy

The validation was separated into two parts depending on the source of the precipitation phase observations, and an additional analysis was performed using two case studies.

3.2.1. Threshold determination

The first part of the validation consisted of using only quality-controlled present weather observations, which included data from NIVOBS, ISD (METARS and SYNOPS) and XOM-OBS. Each observation was compared with the nearest pixel of the estimated precipitation phase field obtained through interpolation (Section 2.2.1) with a spatial resolution of 250 m. Since the interpolated fields were built using 30-min averages of the AWS observations and as the present weather observations were not averaged but mostly reported either at 00 or 30 min past the hour, the mean of the preceding and subsequent interpolated fields including the time of the observation was used.

The different methodologies presented in Section 3.1 were compared in terms of verifying deterministic event forecasts, that is, defining a contingency table comparing all possible cases of the observation and the forecast (Table A1). For example, the discrimination of snow individually was evaluated by comparing observations of snow vs no snow, that is, snow against sleet and rain events. Two skill scores, suitable for multi-categorical forecasts (Jolliffe and Stephenson, 2012), were selected to determine the thresholds that separated the positive and negative events: the Pierce skill score (PSS) (Table A2) for the single threshold schemes and the Gerrity skill score (GSS) (Gerrity, 1992) for the dual threshold schemes (Appendix A).

Each precipitation phase considered was evaluated individually and compared against the others. Threshold determination consisted of evaluating each methodology (Table 2) by varying the corresponding meteorological variable within a range from −5 °C to 5 °C in 0.1 °C steps for single and dual threshold schemes and from 0% to 100% in 1% steps for the Koistinen–Saltikoff method. The thresholds that maximised the PSS and GSS scores were selected. In addition, three skill scores were selected to characterise the schemes for precipitation phase discrimination: probability of detection (POD), false alarm rate (FAR) and frequency bias index (FBI) (Jolliffe and Stephenson, 2012) (Table A2).

3.2.2. Citizen science and crowd sourced data value assessment

Once the best performing thresholds and schemes were selected, the second part of the validation evaluated the reliability and role of the XOM-SPO data and observations on social networks. Both sources have some issues that must be taken into account. First, the time of the observation may not exactly coincide with the time the reported event occurred. Second, the location of the observation can be misleading, since the spotter may be reporting an event that is occurring at their vicinity and not at the indicated location, for instance, being at the bottom of a valley, but reporting snow at a higher altitude. Third, the precipitation phase reported could be a misclassification. Therefore, the assessment procedure resembled that of the other sources (Section 3.2.1), but including spatial fuzzy verification (Ebert, 2008). Thus, different areas around the locations of the observations were considered. If a single pixel inside the defined area agreed with the precipitation phase observed, this was considered a hit.

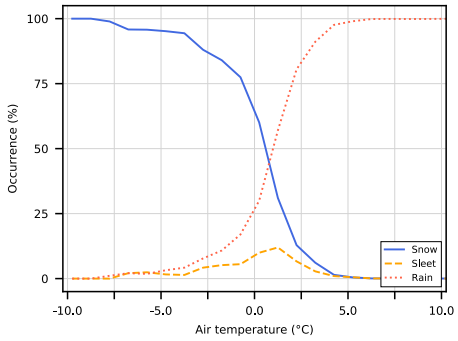

Fig. 3. Occurrence of the precipitation phase based on quality-controlled present weather observations (NIVOBS, ISD and XOM-OBS) as a function of the estimated air temperature.

3.2.3. Case studies

Two low-altitude snowfall events were selected to analyse the performance of the schemes at the event scale and to examine how snow/rain and rain/snow transitions were described by the precipitation phase estimates. In addition to comparing the precipitation phase observed with the estimated one, the coverage of the present weather observations of the precipitation fields derived from the weather radar (XRAD) was also evaluated. As the radar observations have a temporal resolution of 6 min, a temporal fuzzy scheme was adopted. For a single observation and the time when it was reported, the coinciding as well as the two preceding and the subsequent radar images were selected. This approach tried to account for different effects such as the time that the precipitation takes to reach the ground from the moment it is observed by the radar, especially if it is in the form of snow, and possible advection below the radar observation (Mittermaier et al., 2004; Lauri et al., 2012; Saltikoff, 2012).

4. Results

4.1. Precipitation phase and air temperature

Before evaluating the different schemes for precipitation phase discrimination described in Section 3.1, the present weather observations were narrowed down based on their T_a, like in Ding et al. (2014) and Froidurot et al. (2014). To establish the T_a range of the study, each precipitation phase observation was linked to its corresponding T_a estimation, obtained from the nearest pixel of the interpolated air temperature field. Fig. 3 shows the precipitation phase when the T_a is between −10 °C and 10 °C. The range where the precipitation phases overlap falls approximately between −7.5 °C and 5.0 °C, which was adopted as the range of this study. This range included 7,702 observations, of which 2,952 were snow, 415 sleet and 4,335 rain. The range selected was consistent with those of previous studies: −3 °C to 5 °C in Froidurot et al. (2014), −4 °C to 8 °C in Liu et al. (2018), −8 °C to 8 °C in Jennings et al. (2018) and −10 °C to 10 °C in Ding et al. (2014).

4.2. Threshold determination

4.2.1. Single threshold

A PSS score of 0.77 was obtained for the ST-T_w scheme, using T_w = 0.7 °C. T_w was the best performing meteorological variable among the single threshold methodologies (ST in Table 3). The POD and FAR showed high and low values, respectively (Table 3), for both snow and rain. The T_w threshold obtained was similar to that of Sims and Liu (2015) over land (1.0 °C) and of Froidurot et al. (2014) (~1 °C). It also fell within the range of thresholds reported by Behrangi et al. (2018) (0.6 °C – 1.0 °C). Regarding the T_a and T_d, the latter performed better in terms of the PSS and exhibited the highest POD for snow (Table 3). The best T_a threshold obtained in this study (1.3 °C) was in agreement with the threshold of 1.0 °C found for Sweden (Feiccabrino and Lundberg, 2009) and Northern USA (Feiccabrino et al., 2012) and that of 1.1 °C reported for the Czech Republic (Hynčica and Huth, 2019). It was also similar to the values reported by Sims and Liu (2015) (1.6 °C) and Liu (2008) (2.0 °C). Furthermore, it was comparable to those used in different hydrological models, such as the threshold of 1.0 °C reported in Quéno et al. (2018) and Azam et al. (2019). The T_d threshold in this study was similar to those of other studies, such as the threshold of 0.0 °C used by (Marks et al., 2013) in a mountain basin in the USA and 0.1 °C applied by (Feiccabrino and Lundberg, 2009) in Sweden.

4.2.2. Dual threshold

The results from the dual threshold methodologies (DT, Table 3) showed that the best schemes discriminating rain, sleet and snow were DT-KS and DT-T_w, with a GSS of 0.73, compared to that of 0.68 and 0.70 for DT-T_a and DT-T_d, respectively. The DT-T_w and DT-KS methodologies produced practically the same results for the PSS. The largest differences between the two methodologies arose in the POD and FAR scores for the rain and snow categories. The DT-KS scheme showed a higher POD for rain (0.87) compared to that of DT-T_w (0.84) (Table 3). The DT-T_w scheme exhibited a POD of 0.89 for snow, which was four points higher than that of DT-KS (Table 3). The small differences between the methodologies in terms of the GSS and PSS could be due to the similar boundaries in the T_a vs *RH* plane that can be seen in Fig. 2. By contrast, the DT-T_d scheme was completely different, presenting a lower slope (Fig. 2) and consequently yielding the highest and lowest POD for snow and rain, respectively.

Based on the results presented in Table 3, the best PSS and POD scores were obtained with the single threshold methodologies. However, they do not account for the mixed phase transition between rain and

Table 3
Skill scores for the different schemes for precipitation phase discrimination (single threshold (ST), dual threshold (DT) and Koistinen–Saltikoff (DT-KS)) and their corresponding optimal thresholds. Scores stratified by the precipitation phase (rain, sleet and snow). The highest scores for each precipitation phase are shown in bold. POD, probability of detection; FAR, false alarm rate; PSS, Pierce skill score; FBI, frequency bias index; GSS, Gerrity skill score.

Scheme	Threshold/s	GSS	Rain/No rain				Sleet/No sleet				Snow/No snow			
			POD	FAR	PSS	FBI	POD	FAR	PSS	FBI	POD	FAR	PSS	FBI
ST-T_a	1.3 °C	–	**0.88**	0.10	0.72	**0.98**	–	–	–	–	0.84	0.18	0.72	**1.03**
ST-T_d	0.1°C	–	0.83	**0.06**	0.74	0.89	–	–	–	–	**0.90**	0.22	0.74	1.16
ST-T_w	0.7 °C	–	**0.88**	0.08	**0.77**	0.95	–	–	–	–	0.88	**0.17**	**0.77**	1.06
DT-T_a	1.0 °C 1.7 °C	0.68	0.82	0.12	0.68	0.93	0.11	**0.81**	0.06	0.59	0.81	**0.19**	0.69	**1.00**
DT-T_d	0.1 °C 0.4 °C	0.70	0.79	**0.09**	0.69	0.87	0.03	0.90	0.01	0.37	**0.90**	0.26	0.71	1.23
DT-T_w	0.7 °C 1.0 °C	**0.73**	0.84	**0.09**	**0.72**	0.93	0.14	0.84	0.09	1.29	0.89	0.22	**0.73**	1.13
DT-KS	40% 58%	**0.73**	**0.87**	0.11	**0.72**	**0.98**	**0.17**	0.84	**0.13**	**1.14**	0.85	**0.19**	**0.73**	1.05

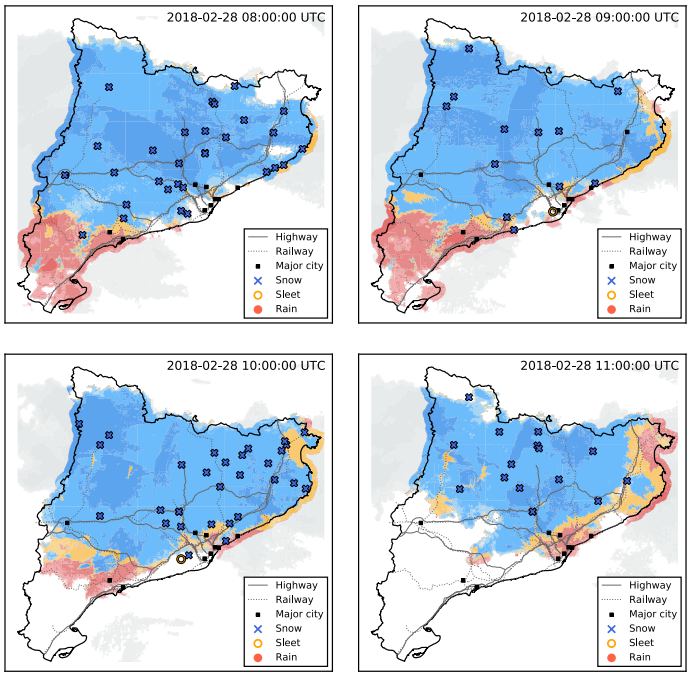

Fig. 4. Surface precipitation phase obtained with the Koistinen–Saltikoff (DT-KS) scheme merged with weather radar precipitation fields on 28th February 2018. Shaded colours represent rain (reddish), sleet (yellowish), snow (blueish) and undefined for areas outside the boundaries of Catalonia (greyish). Symbols represent citizen science and crowd sourced observations: snow (blue crosses), sleet (yellow empty circles) and rain (red solid circles). Grey lines represent major highways (continuous line) and railways (dotted line), and black squares major cities (>100,000 inhabitants). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

snow. For this purpose, sleet observations were valuable to define the boundaries between rain and snow. Among the seven schemes considered, ST-T_w, DT-T_w and DT-KS stood above the rest.

4.3. Value assessment of citizen science and crowd sourced data

The interest from the general public for severe weather is increased when weather warnings are issued, as found for the United States and the word "tornado" on Twitter (Ripberger et al., 2014). Studying human responses to physical stimuli, Weber and Ross (1978) found that the greater the level of change in the stimulus compared to a given background value, the greater the human response. This principle was applied to a climate change study which found that people's sensitivity to extreme weather events was proportional to the situation they normally experienced: i.e., the larger the change, the greater the impact (Sisco et al., 2017). Thus, this study took advantage of the exceptionality of low-altitude snow events in Catalonia by gathering precipitation phase observations from Twitter, together with those from the SMC spotter network. However, there were far fewer rain and sleet observations on Twitter and on the spotter network compared to snow observations. Therefore, only the snow observations were considered. For this reason, only the POD score was calculated. When the nearest pixel (250-m) was considered, the POD obtained was 0.77 for the XOM-SPO observations and 0.70 for Twitter. When the area was expanded to 3 km^2, the POD was 0.90 for XOM-SPO and 0.84 for Twitter. Both of these POD scores are notable, highlighting the importance and the value of non-conventional real-time observations in weather surveillance. Moreover, if the area considered for the fuzzy verification was extended to ~30 km^2, similar to the mean area of the Catalan towns, the POD values obtained were 0.94 for XOM-SPO and 0.89 for Twitter. Fig. 4 shows an example of citizen science and crowd sourced observations for an event on 28th February 2018. Although a few observations did not agree with the estimated form of precipitation, the vast majority coincided with the estimates (Fig. 4), including the progressive displacement of the snowfall towards the north of the region.

Therefore, both citizen science and crowd sourced observations are reliable and provide important information since they cover a wider area than the fixed ISD and XOM-OBS observations. In addition, they are mostly concentrated in populated areas and close to critical infrastructures such as highways, train lines and schools. Nevertheless, care must be taken when using crowd sourced data due to their inherent issues (Schuster et al., 2005; Saltikoff et al., 2010; Reeves et al., 2014).

4.4. Case studies

4.4.1. 7th–8th March 2010

In March 2010, one of the most important low-altitude snowfall events in recent years in Catalonia took place, not only from a meteorological point of view (Aran et al., 2010), but also for the damages it caused (Llasat et al., 2014). This event has been widely studied from different points of view: the social impact (Amaro et al., 2010), the weather warnings issued (Vilaclara et al., 2010), the meteorological conditions (Aran et al., 2010), and the occurrence of thundersnow (Bech et al., 2013).

The event was characterised by the passage of a cold and dry air mass from Central Europe and a low air pressure centre that formed SE of the Catalan coast, which advected warmer and humid air from the Mediterranean Sea, common in low-altitude snowfall events in the region (Bech et al., 2013). The situation is illustrated in Fig. 5a, b, based on ERA5-Reanalysis (Hersbach et al., 2020). At the coast (sea level), snow depths ranged from 4 to 15 cm, reaching 60 cm at 500 m a.s.l. on the coastal hills. These depths had not been reported since 1992 (Llasat

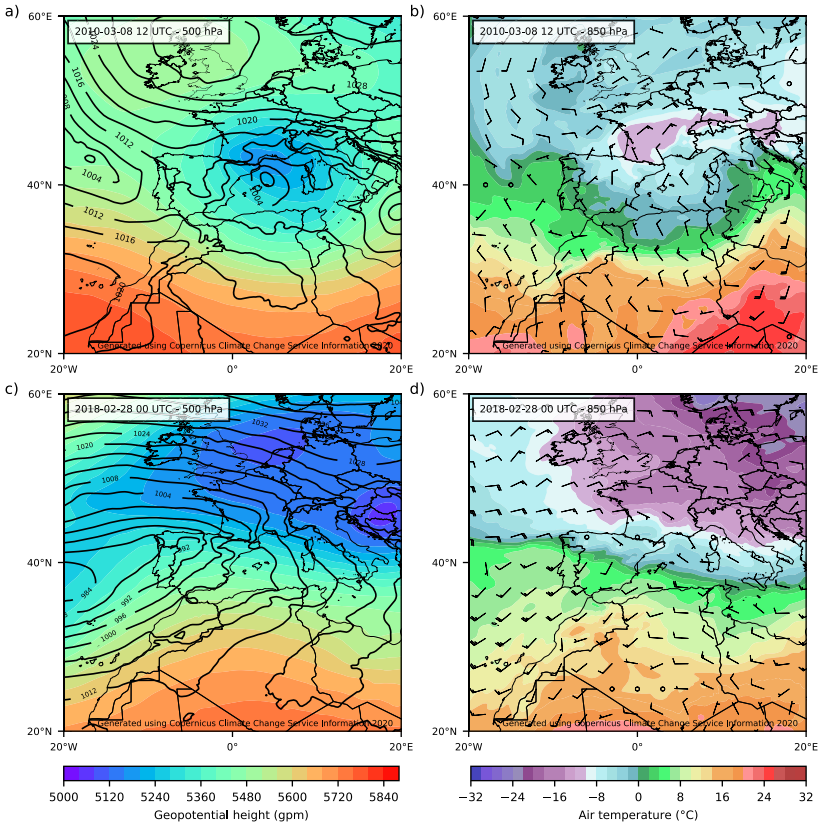

Fig. 5. ERA5 reanalysis showing the geopotential height at 500 hPa (shaded colours), as well as the mean sea level pressure (black contours) in the left column and air temperature (shaded colours) and mean wind (barbs) at 850 hPa in the right column. First row corresponds to 12:00 UTC on 8th March 2010, while the second row corresponds to 00:00 UTC on 28th February 2018.

et al., 2014). The event started in the afternoon of 7th March in the SW of Catalonia, progressively spreading to the whole region during 8th March before finally being restricted to the NE sector.

The ISD and XOM-OBS gathered 277 observations from this event, of which 132 corresponded to snow, 29 to sleet and 116 to rain. These, together with their air temperature and relative humidity estimates, are shown in Fig. 6a. Fig. 6b displays the precipitation phase field on 8th March 2010 at 07:00 UTC based on the DT-T_w scheme, with a threshold of 0.7 °C for snow and 1.0 °C for rain. There was general agreement between the precipitation phase observations and the estimated field, even for the spatial precipitation phase transition (snow, sleet and rain) observed in the surroundings of Reus and Barcelona. However, there were some observations that clearly did not fit with the estimates, especially the locations with rain observations that were surrounded extensively by snow areas (Fig. 6b).

The DT-KS and DT-T_w schemes produced a PSS of 0.66 and 0.65, respectively, and a relatively high POD for snow vs no snow observations. When using the event-optimised thresholds, the snow threshold shifted from 40% to 32% for the DT-KS scheme and from 0.7 °C to 0.9 °C for the DT-T_w scheme (Table 4). These changes increased the POD and PSS scores for both schemes (Table 4), suggesting that despite applying predetermined thresholds, different meteorological conditions at the event scale may lead to variability in the thresholds used to discriminate between rain and snow. One reason for this could be the limitation of the schemes to capture variations in vertical temperature profiles (Feiccabrino et al., 2015).

The weather reports from the four airports of Barcelona-El Prat, Girona-Costa Brava, Sabadell and Reus (Fig. 6c) indicated that at least one rain-to-snow transition occurred during the event, allowing us to test the ability of the diagnostic scheme to deal with changing conditions. In Sabadell, there was a slow transition from rain to snow with several hours of sleet that was only interrupted by some rain. The DT-KS scheme (represented by the black line and the shaded colours in Fig. 6c, Sabadell) captured the sleet observations, together with the transient change to rain. Even though it managed to identify sleet, the DT-KS scheme estimated the occurrence of snow too early. At the Girona-Costa Brava and Reus airports (Fig. 6c, Girona-Costa Brava and Reus), the precipitation phase estimates were mostly in agreement with the observations, including the snow-sleet-rain-snow transition in Girona-Costa Brava at around 06:00 UTC on 8th March 2010. Finally, at the Barcelona-El Prat airport, the scheme failed to identify the rain-to-snow transition properly because the KS values were between the thresholds

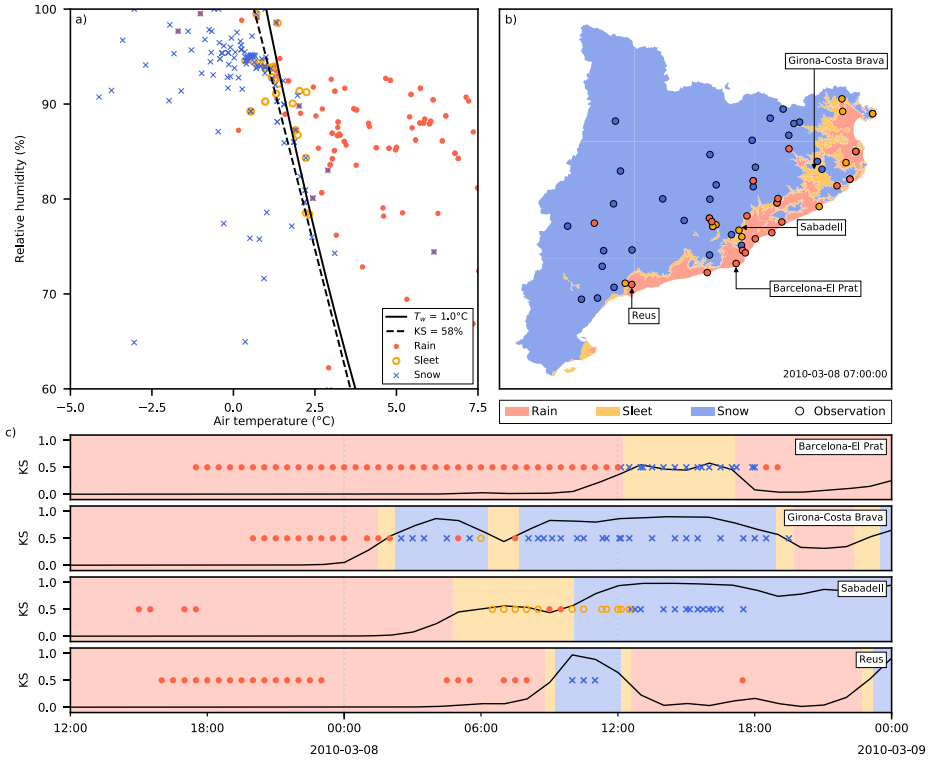

Fig. 6. Summary of the low-altitude snowfall event of 7th and 8th March 2010. (a) Precipitation phase observations with corresponding air temperature and relative humidity estimates, together with the Koistinen–Saltikoff (DT-KS) (dashed line) and dual threshold wet-bulb temperature (DT-T_w) (solid line) threshold for snow [based on (Froidurot et al., 2014)]. (b) Precipitation phase field obtained with the DT-*Tw* scheme using precipitation phase observations for 07:00 UTC on 8th March 2010. (c) Precipitation phase observations (coloured symbols as in panel (a)) and KS probability of snow values (black line) during the event, together with the precipitation phase estimates represented by the shaded colours (red for rain, yellow for sleet and. blue for snow). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Skill scores for snow vs no snow observations for the 7th-8th March 2010 episode using general thresholds. The scores for the event-optimised thresholds are indicated with (Adjusted).

Scheme	Thresholds	PSS	POD	FAR
DT-KS	40% 58%	0.60	0.76	0.14
DT-KS (Adjusted)	32% 50%	0.65	0.83	0.14
DT-T_w	0.7 °C 1.0 °C	0.59	0.75	0.14
DT-T_w (Adjusted)	0.9 °C 1.3 °C	0.66	0.84	0.14

for rain and snow (indicated by the yellow shade in Fig. 6c) during the observations of snow. Nevertheless, the misclassification was, in general, marginal and the KS value was close to the established thresholds compared to the reports.

The results also indicated that 98% of the present weather observations were covered by the precipitation fields detected by the weather radar when compared to the nearest pixel of the composite reflectivity. Thus, 271 of the 277 quality-controlled observations were included in the SPP product combining the precipitation phase and mask fields.

4.4.2. 26th–28th February 2018

Between 26th and 28th February 2018, two low-altitude snowfall events took place in Catalonia. On the 26th and 27th February, during the first snowfall event, the snow was restricted to the NE sector of the region, with snow accumulations at 500 m a.s.l. of up to 26 cm. This was due to the arrival of a polar continental air mass that did not cover the entire Iberian Peninsula, but led to air temperature values of around −7 °C at the 850 hPa level (Fig. 5c, d). On 28th February, the second snowfall event occurred. This was a more general and widespread snowfall event associated with the passage of a warm front (Bullón and Fernández, 2019). Advection of humid air from the south (Fig. 5c, d), favoured by an Atlantic low pressure located west of the Iberian Peninsula and warmer than the Siberian one, facilitated substantial precipitation across Catalonia, especially in the Pyrenees. The snowfall followed a SW to NE path, with snow occurring at around 100 m a.s.l. or even 0 m a.s.l. in the early morning and at 1,000 m a.s.l. during the day. Snow accumulations were 5 to 10 cm in the coastal and pre-coastal ranges, more than 40 cm on the peaks of the Pyrenees, and 20 to 30 cm in the Pre-Pyrenees range and in the central part of Catalonia.

This event was evaluated by considering 431 observations from between 26th and 28th February, of which 191 corresponded to snow, 17 to sleet and 223 to rain (shown in the T_a vs *RH* plane in Fig. 7a). Both the

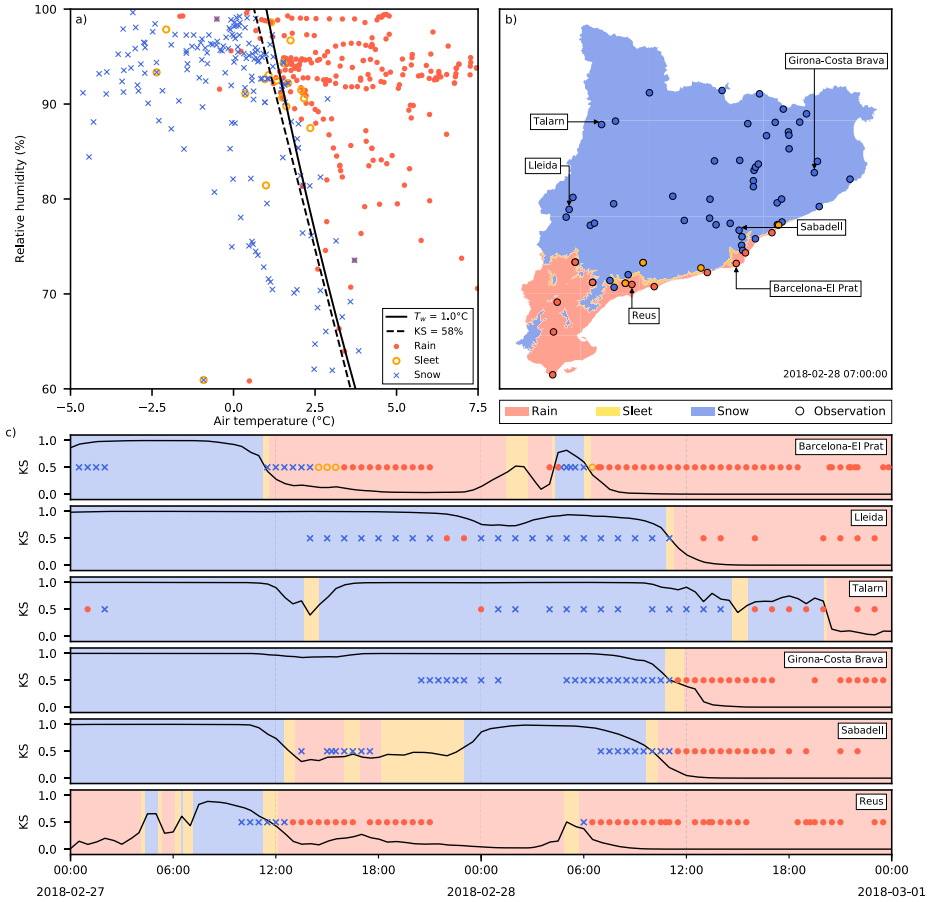

Fig. 7. As in Fig. 6, for the event of 26th to 28t.h February 2018.

Table 5
Skill scores for snow vs no snow observations for the 27th-28th February 2018 episode using general thresholds. The scores for the event-optimised thresholds are indicated with (Adjusted).

Scheme	Thresholds	PSS	POD	FAR
DT-KS	40% 58%	0.74	0.85	0.10
DT-KS (Adjusted)	39% 46%	0.76	0.90	0.13
DT-T_w	0.7 °C 1.0 °C	0.74	0.84	0.10
DT-T_w (Adjusted)	1.0 °C 1.1 °C	0.75	0.92	0.15

DT-KS and DT-T_w methodologies produced a similar PSS of 0.78 for rain and 0.74 for snow (Table 5). As in Section 4.4.1, the event-optimised thresholds were adjusted. This optimisation increased the FAR, but this was compensated by the increase in the POD and PSS scores for snow vs no snow observations (Table 5). Fig. 7b presents an example of the DT-KS field for this event for 28th February 2018 at 07:00 UTC. It can be seen that more than half of the region was experiencing snowfall, except for the central coastline and the southern sector. Despite some exceptions, there was generally good agreement between the DT-KS field and the observations, especially in the boundaries between rain and snow.

Transitions from rain to snow and vice versa are shown in Fig. 7c. For example, in Barcelona, there were three of these transitions. First, on 27th February at 12:00 UTC, the DT-KS and DT-T_w (not shown) schemes failed to identify the transition. Indeed, the schemes estimated it three hours too early. In this case, it snowed at relatively warm temperatures (2.0 °C to 3.0 °C) and relatively high *RH* (80% to 90%). The black line in the Barcelona-El Prat graph is far from the thresholds for rain and snow, suggesting that the scheme clearly attributed the precipitation phase to rain. However, this did not occur in Sabadell or Reus. In Sabadell, it estimated rain and sleet instead of the snow that was observed. However, the DT-KS value was closer to the thresholds for Sabadell than for Barcelona. By contrast, the snow-to-rain transition was mostly captured by the DT-KS scheme in Reus.

Fig. 7c shows that all the present weather observations reported during 28th February included precipitation phase transitions. In Lleida, Girona-Costa Brava and Sabadell, the DT-KS scheme was able to

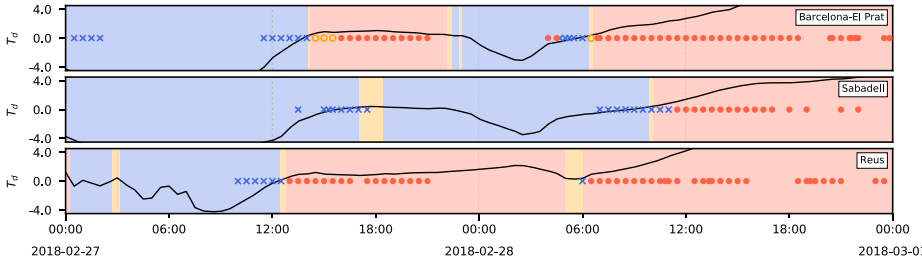

Fig. 8. The low-altitude snowfall event of 26th to 28th February 2018. Evolution of the precipitation phase observations and T_d values (black line) during the event, together with the precipitation estimates from the dual threshold dew point temperature (DT-Td) scheme represented by shaded colours (red for rain, yellow for sleet and blue for snow) for Barcelona-El Prat, Sabadell and Reus. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

reproduce approximately the observed transitions. In Talarn, despite the DT-KS scheme not estimating the time of the transition correctly, the snow and rain thresholds worked reasonably well. The rain-snow-rain transition at the Barcelona-El Prat airport on 28th February at 06:00 UTC was captured almost perfectly by the DT-KS scheme. In this case, however, the DT-T_d scheme failed to identify the sharp transition in Barcelona-El Prat, but performed better for Reus and similarly to DT-KS in Sabadell (Fig. 8).

The results indicated that 75% of the present weather observations were covered by the precipitation fields detected by the weather radar when compared to the nearest pixel of the composite reflectivity. If an area of 5 x 5 px (1 px is ~1 km^2) surrounding the location of the observations was taken into account, this increased to 82%.

During the first hours of the event, the DT-KS and DT-T_w (not shown) schemes failed to capture properly the precipitation phase due to a combination of high air temperatures and high *RH*. However, the DT-T_d scheme captured this well, since it diagnoses snow at warmer air temperatures (Figs. 2 and 8). However, for 28th February, the DT-KS and DT-T_w schemes performed better than the DT-T_d scheme. This suggests that there is no unique methodology that fully covers the precise nature of rain-to-snow transitions, indicating that these schemes are unable to capture variations in vertical temperature profiles (Feiccabrino et al., 2015). In this case, the DT-T_d scheme performed well in the first part of the event, but failed in the middle and final parts of it.

4.5. Implementation of a real-time SPP product

The implementation of a real-time SPP product for monitoring and surveillance must provide sufficient and clear information to decision makers. Analysis of individual events, as described in Section 4.4, can lead to different optimal thresholds, with differences not only at the event scale, but also in the intra-event periods. Therefore, the real-time SPP product must reflect the variability observed, allowing decision makers to modify the predetermined thresholds within a restricted range or schemes to discriminate between snow, sleet and rain, in accordance with the precipitation phase observations.

Furthermore, the implementation must take into account the end-users. Thresholds that maximise the validation scores may not meet all the needs of the end-users, since decision makers may be less concerned about the FAR if this is compensated by an increase in the POD (Manzato, 2007).

The methodologies described have been implemented for real-time operations at the Meteorological Service of Catalonia, where schemes using the wet-bulb temperature and variable thresholds have been used. This can be easily adopted elsewhere with the open-source software package pyPROS (standing for Precipitation type: Rain Or Snow). Scripts, tutorials and documentation can be found in Casellas et al. (2019).

Table 6
Pierce skill score (PSS) for the wet-bulb temperature dual threshold (DT-T_w) and Koistinen-Saltikoff (DT-KS) schemes for the Integrated Surface Database (ISD) observations only. Observed refers to the scores obtained when using the observed air temperature and dew point temperature values, whereas the interpolated results refer to those obtained from the interpolation of the T_a and T_d observations.

Scheme	Thresholds	PSS (Rain/No rain)	PSS (Snow/No snow)
DT-KS (Observed)	10% 15%	0.74	0.75
DT-KS (Interpolated)	10% 20%	0.77	0.78
DT-T_w (Observed)	1.0 °C 1.3 °C	0.74	0.74
DT-T_w (Interpolated)	0.9 °C 1.3 °C	0.77	0.77

5. Discussion

Froidurot et al. (2014) suggested the need to analyse schemes for precipitation phase discrimination using estimated (instead of measured) data, allowing applications at a catchment scale. In this study, only estimated T_a, T_d, *RH* and T_w values were used. The T_a and T_d fields were obtained from the spatial interpolation of AWS observations, while *RH* and T_w were calculated from the aforementioned fields. The results showed that including information on humidity clearly increased performance (Table 3), indicating that schemes that take into account air saturation conditions play an important role in precipitation phase discrimination. The obtained results (Table 3) are comparable with those reported in previous studies. For example, Gjertsen and Ødegaard (2005) obtained a POD of 0.85 (0.97) and a FAR of 0.05 (0.15) for rain (snow), Schmid and Mathis (2004) reported a POD of 0.98 (0.89) and a FAR of 0.19 (0.01) for rain (snow), while Froidurot et al. (2014) obtained a CSI of 0.84 for rain, compared to the 0.78 identified in this study.

One of the limitations of the study was the relatively low number of simultaneous observations of SPP, T_a and T_d, which were only available in the SYNOP and METAR locations. These data were gathered from the ISD and used to assess the agreement between the thresholds and the performance of the schemes using observed and interpolated data. The results demonstrated that the optimised thresholds and the corresponding PSS values for the ISD locations (Fig. 1c, green triangles) were similar for the interpolated and observed data for the DT-T_w and DT-KS schemes (Table 6). The other schemes provided comparable results (not shown). One of the reasons behind these results could be the interpolation method applied, including the residual correction, which led to the interpolated field values being equal to those of the observations at their locations. Therefore, this enabled the schemes to be tested with

Table A1
A 3x3 contingency table.

		Observed		
		Rain	Snow	Sleet
Forecast	Rain	a	b	e
	Snow	c	d	f
	Sleet	g	h	i

estimated values and provided added value through the use of information on humidity when applying the schemes at a catchment or regional scale.

Despite the reasonably good performance of the SPP schemes, there were some limitations. First, the vertical temperature profiles play a key role in determining the precipitation phase at the surface level (Bourgouin, 2000) through the effects of melting, evaporation and refreezing. These processes are not considered when only surface information is used. Second, some threshold variability was observed, not only at the inter-event scale, but also at the intra-event scale, suggesting that the parameters of the schemes should be adjusted according to the event. This is in agreement with the findings of Cortinas et al. (2002) and Wandishin et al. (2005), who noted that no vertical temperature profile-based algorithm discriminating the SPP outperformed the other algorithms all the time.

This variability was addressed in the implementation of the real-time SPP product by using different thresholds. Despite the simplicity of the schemes used, they produced good results that could be improved in further studies by including vertical temperature profiles, which would provide key information about possible warm or cold layers above the surface, as well as a more extensive database of case studies to check for possible relationships between the scheme thresholds and event characteristics. An option to enlarge this precipitation phase observation database could be achieved through citizen science and crowdsourcing, whose role experienced in recent years an increasing trend in Earth observation topics (Fritz et al., 2017), such as heavy precipitation events (Spruce et al., 2020), natural disasters (Guan and Chen, 2014) and catchment modelling (Starkey et al., 2017). mPING (Elmore et al., 2014), CoCoRaHS (Reges et al., 2016) and WOW (wow.metoffice.gov.uk) are examples of well established user friendly applications that citizens can use to report weather observations. Another option to gather observations is the use of social networks, such as Twitter or Facebook. In this study, precipitation phase observations were retrieved from Twitter exhibiting good results. Still, more observations could have been gathered with an active campaign, such as "Snowtweets" (King et al., 2009) or "Picking up Hailstones" (Farnell and Rigo, 2020) campaigns, which encouraged citizens to use specific hashtags to report snow and hail size, respectively. These are examples of how people can provide valuable information using applications not specifically developed for citizen science purposes. Therefore, active campaigns using a specific communication strategy, such as comparing a forecast to a well-known past event, can engage and motivate citizens to report weather-related observations during particular events (Lambrecht et al., 2019). In addition, it is also important to assure that the observations reported by citizens are needed, appreciated and acknowledged (Reges et al., 2016). The combination of citizen-science and the installation of present weather sensors at key locations could improve the management of risks, transport infrastructure and road safety.

Table A2
Skill scores for the Rain/No Rain case used in the validation strategy, their formula and perfect value: POD (Probability of Detection), FAR (False Alarm Rate), FBI (Frequency Bias Index) and PSS (Pierce Skill Score). *a, b, c, d, e, f, g, h* and *i* refer to letters defined in the contingency table in Table A1. When sleet observations are not considered *e, f, g, h* and *i* equal to 0.

Skill score	Formula	Perfect value
POD	$\frac{a}{a+(c+g)}$	1
FAR	$\frac{(b+e)}{a+(b+e)}$	0
FBI	$\frac{a+(b+e)}{a+(c+g)}$	1
PSS	$\frac{a\cdot(d+i)-(b+e)\cdot(c+g)}{(a+c+g)\cdot(b+e+d+i)}$	1

6. Summary and conclusions

This paper studied the application of different schemes for precipitation phase discrimination to evaluate their performances and suitability for a real-time SPP product. Compared to previous studies that developed and analysed schemes based on concurrent meteorological measurements and precipitation phase observations, the present study used interpolated air temperature and dew point estimates. In addition, the role of citizen science and crowd sourced data regarding precipitation classification among rain, sleet and snow was assessed. Finally, a proposal on the implementation of a real-time SPP product was presented. The following conclusions arose from the study:

- Out of all the schemes considered in this study, those incorporating information on air humidity, even that estimated through interpolation, performed better than those considering air temperature alone.
- The single threshold air temperature scheme (ST-T_a), widely used in hydrological models, resulted in a threshold of 1.3 °C.
- The two best schemes from a statistical point of view when three types of precipitation (rain, sleet and snow) are considered were the dual threshold wet-bulb temperature (DT-T_w) and Koistinen–Saltikoff (DT-KS) schemes, with both including information on humidity and temperature.
- Citizen science and crowd sourced observations presented a high probability of detection (POD), indicating their potential in the monitoring of low-altitude snow events.
- The analyses of individual events showed that small modifications in the defined thresholds may improve the schemes, suggesting that they should be adapted to specific events.
- The implementation of a real-time surface precipitation phase product should consider variability in the optimum thresholds observed at the inter- and intra-event scales.

CRediT authorship contribution statement

Enric Casellas: Methodology, Software, Validation, Writing - original draft. **Joan Bech:** Conceptualization, Supervision, Writing - review & editing, Funding acquisition, Project administration. **Roger Veciana:** Methodology, Software, Conceptualization. **Nicolau Pineda:** Project administration, Funding acquisition, Supervision. **Tomeu Rigo:** Data curation, Writing - review & editing. **Josep Ramon Miró:** Conceptualization, Formal analysis, Writing - review & editing. **Abdel Sairouni:** Project administration, Funding acquisition, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was performed under the framework of the HyMeX program, and supported by the Industrial Doctorate Projects framework partly founded by the Government of Catalonia (Project DI-053/2017), and with partial support from the Spanish Projects CGL2015-65627-C3-2-R (MINECO/FEDER), CGL2016-81828-REDT/AEI and RTI2018-098693-B-C32 (MINECO). The authors acknowledge the Cartographic

and Geological Institute of Catalonia (ICGC), the Meteorological Service of Catalonia (SMC) staff, and the XOM observers and spotters for providing present weather observations. We also acknowledge two anonymous reviewers for providing constructive suggestions that improved the final form of this paper.

Appendix A. Skill scores calculation

Multi-categorical binary events (rain, sleet and snow), can be represented by a contingency table (Table A1), accounting for the correctly classified, correctly rejected, missed and false alarm events. For example, for the Rain/No Rain case *a* corresponds to correctly classified, *d* and *i* to correctly rejected, *c* and *g* to missed, and *b* and *e* to false alarm occurrences.

From these values, skill scores calculated are listed in Table A2 for the Rain/No Rain case. For the Snow/No Snow and Sleet/No Sleet cases, formulas in Table A2 must be changed accordingly.

When only two categories are considered, such as rain and snow, sleet column and row values are set to null, that is, collapsing the contingency table from 3 x 3 to 2 x 2. Therefore, *e*, *f*, *g*, *h* and *i* in skill score formulas (Table A2) are set to 0 for the Rain/No Rain case.

Gerrity Skill Score (GSS), indicated for multi-categorical forecasts (Jolliffe and Stephenson, 2012), is used solely for dual threshold schemes to find the best thresholds that properly separate the three precipitation phase considered. GSS is a modification of a score defined by Gandin and Murphy (Gandin and Murphy, 1992) introduced by Gerrity (1992) which uses a scoring matrix based on the sample probabilities making the score truly equitable. A characteristic of using GSS is that penalizes misclassification differently depending on the distance from the correct category. A complete description of GSS calculation can be found in Jolliffe and Stephenson (2012).

References

Altube, P., Bech, J., Argemí, O., Rigo, T., 2015. Quality control of antenna alignment and receiver calibration using the sun: Adaptation to midrange weather radar observations at low elevation angles. Journal of Atmospheric and Oceanic Technology 32, 927–942.

Amaro, J., Llasat, M., Aran, M., 2010. The social impact of the snowfall of 8 March 2010 in Catalonia. In: 12th Plinius Conference on Mediterranean Storms, Corfu Island, Greece.

Apodaka, J., Pons, M., Trapero, L., Margalef, A., Albalat, A., Gallego, N., López-Moreno, J.I., Furdada, G., 2018. Evaluation of 30 years of nivo-meteorological and avalanche data in Andorra. In: International Snow Science Workshop, Innsbruck, Austria.

Aran, M., Rigo, T., Bech, J., Brucet, C., Vilaclara, E., 2010. Analysis of the hazardous low-altitude snowfall, 8th March 2010, in Catalonia. In: 12th Plinius Conference on Mediterranean Storms, held September 1-4, 2010 in Corfu Island, Greece. http://meetings. copernicus. org/plinius12, id. 77.

Azam, M.F., Wagnon, P., Vincent, C., Ramanathan, A., Kumar, N., Srivastava, S., Pottakkal, J., Chevallier, P., 2019. Snow and ice melt contributions in a highly glacierized catchment of Chhota Shigri Glacier (India) over the last five decades. Journal of Hydrology 574, 760–773.

Bech, J., Codina, B., Lorente, J., Bebbington, D., 2003. The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. Journal of Atmospheric and Oceanic Technology 20, 845–855.

Bech, J., Rigo, T., Pineda, N., Segalà, S., Vilaclara, E., Sánchez-Diezma, R., Sempere, D., Velasco, E., 2005. Implementation of the EHIMI software package in the weather radar operational chain of the Catalan Meteorological Service. In: Proceedings, 32nd International Conference on Radar Meteorology, Alburquerque, NM, USA.

Bech, J., Pineda, N., Rigo, T., Aran, M., 2013. Remote sensing analysis of a Mediterranean thundersnow and low-altitude heavy snowfall event. Atmospheric Research 123, 305–322.

Bech, J., Vidal, V., Ortiz, J., Pineda, N., Veciana, R., 2014. Real-time estimation of surface precipitation type merging weather radar and automated station observations. 17th International Road Weather Conference, SIRWEC, 12, 8.

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., Ye, H., 2018. On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance. Quarterly Journal of the Royal Meteorological Society 144, 89–102.

Bloemink, H. I., Lanzinger, E., 2005. Precipitation type from the Thies disdrometer. In: WMO Technical Conference on Instruments and Methods of Observation (TECO-2005), Bucharest, Romania (pp. 4–7).

Bourgouin, P., 2000. A method to determine precipitation types. Weather and Forecasting 15, 583–592.

Bullón, J.C., Fernández, J., 2019. Inusual episodio de nevadas de frente cálido en la costa cantábrica y catalana. Agencia Estatal de Meteorología.

Casellas, E., Veciana, R., Bech, J., 2019. pyPROS: Precipitation type: Rain or Snow. URL https://github.com/meteocat/pypros.

Casellas, E., Bech, J., Veciana, R., Miró, J.R., Sairouni, A., Pineda, N., 2020. A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. Atmospheric Research 246, 105103.

Chen, R.-S., Liu, J.-F., Song, Y.-X., 2014. Precipitation type estimation and validation in China. Journal of Mountain Science 11, 917–925.

Cortinas, J., Brill, K., Baldwin, M., 2002. Probabilistic forecasts of precipitation type. In: Preprints, 16th Conf. on Probability and Statistics in the Atmospheric Sciences, Orlando, FL, Amer. Meteor. Soc. vol. 3.

Dai, A., 2008. Temperature and pressure dependence of the rain-snow phase transition over land and ocean. Geophysical Research Letters 35.

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., He, X., 2014. The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. Journal of Hydrology 513, 154–163.

Ebert, E.E., 2008. Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. Meteorological Applications 15, 51–64.

Elmore, K.L., Flamig, Z., Lakshmanan, V., Kaney, B., Farmer, V., Reeves, H.D., Rothfusz, L.P., 2014. mPING: Crowd-sourcing weather reports for research. Bulletin of the American Meteorological Society 95, 1335–1342.

Farnell, C., Rigo, T., 2020. The lightning jump, the 2018 "picking up hailstones" campaign and a climatological analysis for catalonia for the 2006–2018 period. Tethys 17, 10–20.

Feiccabrino, J., Lundberg, A., 2009. Precipitation phase discrimination in Sweden. In Eastern Snow Conference: 28/05/2008-30/05/2008, pp. 239–254.

Feiccabrino, J., Lundberg, A., Gustafsson, D., 2012. Improving surface-based precipitation phase determination through air mass boundary identification. Hydrology Research 43, 179–191.

Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., Gustafsson, D., 2015. Meteorological knowledge useful for the improvement of snow rain separation in surface based models. Hydrology 2, 266–288.

Fernández-González, S., Valero, F., Sánchez, J., Gascón, E., López, L., García-Ortega, E., Merino, A., 2015. Numerical simulations of snowfall events: Sensitivity analysis of physical parameterizations. Journal of Geophysical Research: Atmospheres 120.

Franco, M., Sánchez-Diezma, R., Sempere-Torres, D., 2006. Improvements in weather radar rain rate estimates using a method for identifying the vertical profile of reflectivity from volume radar scans. Meteorologische Zeitschrift 15, 521–536.

Fritz, S., Fonte, C.C., See, L., 2017. The role of Citizen Science in Earth Observation. Remote Sensing 9. https://doi.org/10.3390/rs9040357.

Froidurot, S., Zin, I., Hingray, B., Gautheron, A., 2014. Sensitivity of precipitation phase over the Swiss Alps to different meteorological variables. Journal of Hydrometeorology 15, 685–696.

Gandin, L.S., Murphy, A.H., 1992. Equitable skill scores for categorical forecasts. Monthly Weather Review 120, 361–370.

Gavaldà, J., Benet, J., de Carreteres, D.G., 2014. Ten years of avalanche forecastin on Bonaigua and Beret roads, Aran valley, Spain. 17th International Road Weather Conference, SIRWEC, p. 7.

Gerrity Jr, J.P., 1992. A note on Gandin and Murphy's equitable skill score. Monthly Weather Review 120, 2709–2712.

Gjertsen, U., Ødegaard, V., 2005. The water phase of precipitation—a comparison between observed, estimated and predicted values. Atmospheric Research 77, 218–231.

Gonzalez, S., Bech, J., Udina, M., Codina, B., Paci, A., Trapero, L., 2019. Decoupling between precipitation processes and mountain wave induced circulations observed with a vertically pointing K-band doppler radar. Remote Sensing 11, 1034.

Guan, X., Chen, C., 2014. Using social media data to understand and assess disasters. Natural hazards, 74, 837–850.

Harpold, A., Rajagopal, S., Crews, J., Winchell, T., Schumer, R., 2017. Relative humidity has uneven effects on shifts from snow to rain over the western US. Geophysical Research Letters 44, 9742–9750.

Harpold, A.A., Kaplan, M.L., Klos, P.Z., Link, T., McNamara, J.P., Rajagopal, S., Schumer, R., Steele, C.M., 2017. Rain or snow: hydrologic processes, observations, prediction, and research needs. Hydrology and Earth System Sciences 21, 1–22. https://doi.org/10.5194/hess-21-1-2017.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al., 2020. The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society 146, 1999–2049.

Hynčica, M., Huth, R., 2019. Long-term changes in precipitation phase in Europe in cold half year. Atmospheric Research 227, 79–88.

Jennings, K.S., Winchell, T.S., Livneh, B., Molotch, N.P., 2018. Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. Nature Communications 9, 1148.

Jepsen, S., Harmon, T., Meadows, M., Hunsaker, C., 2016. Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, USA. Journal of Hydrology 533, 332–342.

Jolliffe, I.T., Stephenson, D.B., 2012. Forecast Verification: A Practitioner's Guide in Atmospheric Science. John Wiley & Sons.

Kain, J.S., Goss, S.M., Baldwin, M.E., 2000. The melting effect as a factor in precipitation-type forecasting. Weather and Forecasting 15, 700–714.

King, J., Cabrera, A., Kelly, R., 2009. The Snowtweets Project: Communicating snow depth measurements from specialists and non-specialists via mobile communication technologies and social networks. AGUFM, 2009, ED11A–0562.

Koistinen, J., Saltikoff, E., 1998. Experience of customer products of accumulated snow, sleet and rain. COST75 Advanced Weather Radar Systems, pp. 397–406.

Lambrecht, K., Hatchett, B., Walsh, L., Collins, M., Tolby, Z., 2019. Improving visual communication of weather forecasts with rhetoric. Bulletin of the American Meteorological Society 100, 557–563.
Lauri, T., Koistinen, J., Moisseev, D., 2012. Advection-based adjustment of radar measurements. Monthly Weather Review 140, 1014–1022.
Lawrence, M.G., 2005. The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. Bulletin of the American Meteorological Society 86, 225–234.
Liu, G., 2008. Deriving snow cloud characteristics from CloudSat observations. Journal of Geophysical Research: Atmospheres 131.
Liu, S., Yan, D., Qin, T., Weng, B., Lu, Y., Dong, G., Gong, B., 2018. Precipitation phase separation schemes in the Naqu River basin, eastern Tibetan plateau. Theoretical and Applied Climatology 131, 399–411.
Llasat, M., Turco, M., Quintana-Seguí, P., Llasat-Botija, M., 2014. The snow storm of 8 March 2010 in Catalonia (Spain): a paradigmatic wet-snow event with a high societal impact. Natural Hazards and Earth System Sciences 14, 427.
Löffler-Mang, M., Blahak, U., 2001. Estimation of the equivalent radar reflectivity factor from measured snow size spectra. Journal of Applied Meteorology 40, 843–849.
Lott, J., 2004. The quality control of the Integrated Surface Hourly Database, paper presented at the 14th Conference on Applied Climatology. Am. Meteorol. Soc., Seattle, Wash, pp. 10–16.
Lumb, F.E., 1961. The problem of forecasting: the downward penetration of snow. Meteor. Mag., pp. 310–319.
Manzato, A., 2007. A note on the maximum Peirce skill score. Weather and Forecasting 22, 1148–1154. https://doi.org/10.1175/WAF1041.1.
Marks, D., Winstral, A., Reba, M., Pomeroy, J., Kumar, M., 2013. An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. Advances in Water Resources 55, 98–110.
Mittermaier, M.P., Hogan, R.J., Illingworth, A.J., 2004. Using mesoscale model winds for correcting wind-drift errors in radar estimates of surface rainfall. Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography 130, 2105–2123.
Quéno, L., Vionnet, V., Cabot, F., Vrécourt, D., Dombrowski-Etchevers, I., 2018. Forecasting and modelling ice layer formation on the snowpack due to freezing precipitation in the Pyrenees. Cold Regions Science and Technology 146, 19–31.
Reeves, H.D., Elmore, K.L., Ryzhkov, A., Schuur, T., Krause, J., 2014. Sources of uncertainty in precipitation-type forecasting. Weather and Forecasting 29, 936–953.
Reges, H.W., Doesken, N., Turner, J., Newman, N., Bergantino, A., Schwalbe, Z., 2016. CoCoRaHS: The evolution and accomplishments of a volunteer rain gauge network. Bulletin of the American Meteorological Society 97, 1831–1846.
Ripberger, J.T., Jenkins-Smith, H.C., Silva, C.L., Carlson, D.E., Henderson, M., 2014. Social media and severe weather: do tweets provide a valid indicator of public attention to severe weather risk communication? Weather, Climate, and Society 6, 520–530. https://doi.org/10.1175/WCAS-D-13-00028.1 arXiv:https://journals.ametsoc.org/wcas/article-pdf/6/4/520/4622490/wcas-d-13-00028_1.pdf.
Ripoll, R., del Amo, X., Vendrell, R., 2016. The weather observers network of the meteorological service of Catalonia. In: WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation TECO-2016. WMO.
Ryzhkov, A., Zrnic, D., 1998. Discrimination between rain and snow with a polarimetric radar. Journal of Applied Meteorology 37, 1228–1240.
Sadeghi, S.-H., Peters, T.R., Cobos, D.R., Loescher, H.W., Campbell, C.S., 2013. Direct calculation of thermodynamic wet-bulb temperature as a function of pressure and elevation. Journal of Atmospheric and Oceanic Technology 30, 1757–1765. https://doi.org/10.1175/JTECH-D-12-00191.1 arXiv:https://doi.org/10.1175/JTECH-D-12-00191.1.
Saltikoff, E., 2012. Measuring snow with weather radar. In: Doppler Radar Observations-Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications. IntechOpen.
Saltikoff, E., Tuovinen, J.-P., Kotro, J., Kuitunen, T., Hohti, H., 2010. A climatological comparison of radar and ground observations of hail in finland. Journal of Applied Meteorology and Climatology 49, 101–114.
Sánchez-Diezma, R., Sempere-Torres, D., Bech, J., Velasco, E., 2002. Development of a hydrometeorological flood warning system (EHIMI) based on radar data. In: Second European Conference on Radar Meteorology, Delft, Netherlands.
Schmid, W., Mathis, A., 2004. Validation of methods to detect winter precipitation and retrieve precipitation type. Annalen der Meteorologie 40.
Schuster, S.S., Blong, R.J., Speer, M.S., 2005. A hail climatology of the greater sydney area and new south wales, australia. International Journal of Climatology: A Journal of the Royal Meteorological Society 25, 1633–1650.
Schuur, T.J., Park, H.-S., Ryzhkov, A.V., Reeves, H.D., 2012. Classification of precipitation types during transitional winter weather using the ruc model and polarimetric radar retrievals. Journal of Applied Meteorology and Climatology 51, 763–779.
Serra, A., Barnolas, M., Vendrell, R., 2016. Automatic weather stations network (XEMA) of the Meteorological Service of Catalonia (SMC). In: WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (CIMO TECO 2016): Ensuring sustained high-quality meteorological observations from sea, land and upper atmosphere in a changing world. WMO. https://library.wmo.int/index.php?lvl=notice_display&id=19676#.XEWdHvZ7m70.
Shrestha, M., Wang, L., Koike, T., Tsutsui, H., Xue, Y., Hirabayashi, Y., 2014. Correcting basin-scale snowfall in a mountainous basin using a distributed snowmelt model and remote-sensing data. Hydrology and Earth System Sciences 18, 747–761.
Sikorska, A.E., Seibert, J., 2018. Value of different precipitation data for flood prediction in an alpine catchment: A Bayesian approach. Journal of Hydrology 556, 961–971.
Sims, E.M., Liu, G., 2015. A parameterization of the probability of snow–rain transition. Journal of Hydrometeorology 16, 1466–1477.
Sisco, M.R., Bosetti, V., Weber, E.U., 2017. When do extreme weather events generate attention to climate change? Climatic Change 143, 227–241.
Smith, A., Lott, N., Vose, R., 2011. The integrated surface database: Recent developments and partnerships. Bulletin of the American Meteorological Society 92, 704–708.
Spruce, M., Arthur, R., Williams, H., 2020. Using social media to measure impacts of named storm events in the united kingdom and Ireland. Meteorological Applications 27 e1887.
Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., Gibson, C., 2017. Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. Journal of Hydrology 548, 801–817.
Tapiador, F.J., Sánchez, J.-L., García-Ortega, E., 2019. Empirical values and assumptions in the microphysics of numerical models. Atmospheric Research 215, 214–238.
Trapero, L., Bech, J., Rigo, T., Pineda, N., Forcadell, D., 2009. Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. Atmospheric Research 93, 408–418.
Vilaclara, E., Segalá, S., Andrés, A., Aran, M., 2010. Operational warnings issued by the SMC in the 8th March snow event in Catalonia. In: 12th Plinius Conference on Mediterranean Storms, held September 1-4, 2010 in Corfu Island, Greece. http://meetings. copernicus. org/plinius12, id. 79.
Wandishin, M.S., Baldwin, M.E., Mullen, S.L., Cortinas Jr, J.V., 2005. Short-range ensemble forecasts of precipitation type. Weather and Forecasting 20, 609–626.
Weber, E.H., Ross, H.E., 1978. The Sense of Touch. Academic Press for the Experimental Psychology Society.

Chapter 4

Nowcasting the precipitation phase

4.1 Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts

4.1.1 Summary

This chapter presents the development and implementation of a nowcasting system to estimate the precipitation phase. As mentioned in the previous paper, diagnosing the precipitation phase at ground level is critical in manifold meteorological and hydrological applications, as well as its nowcasting which helps to prevent risks and harm to people and critical infrastructures.

The determination of the precipitation phase for nowcasting purposes is usually achieved through a combination of extrapolated observational data and NWP model output. It is a topic that has been widely studied considering different algorithms, including surface observation threshold-based schemes, decision tree algorithms based on vertical temperature profiles, and machine learning methods. However, nowcasting the precipitation phase is still a challenging problem, especially when air temperature is close to freezing point. In fact, some studies reported that a single algorithm may not capture all the variability involved in precipitation phase discrimination at surface level, suggesting a combination of algorithms.

Therefore, in this study, the nowcasting of precipitation phase is not only addressed considering eight individual algorithms (four of them based on surface extrapolated observations and the other four on modelled vertical temperature profiles), but also three

ensembles of algorithms. Firstly, an ensemble that has already been tested in other studies which consists of the combination of the four vertical temperature profile algorithms (hereafter, Profiles) was considered. Secondly, an ensemble of the surface extrapolated observations algorithms (hereafter, MostDelta). Thirdly, the combination of the previously considered eight algorithms (hereafter, MostAll). The two last ensembles are tested for the first time in this study.

As a test bank for the verification of the different algorithms and ensembles, eight low-altitude snowfall events reported in Catalonia between 2010 and 2021 are considered. The verification includes not only the evaluation of the algorithms and ensembles to discriminate precipitation phase, but also their ability to capture precipitation phase transitions. The verification results show that threshold algorithms based on surface extrapolated observations exhibit the best performance in most of the events. However, in some events or parts of them, algorithms based on modelled vertical temperature profiles show a better performance. Therefore, and as reported in previous studies in other regions, certain performance variability is observed in the algorithms' performance depending on the event which makes an ensemble of algorithms advisable.

Among the three ensembles considered, MostAll performs best in three out of eight events and presents a similar performance to the best-performing algorithm in the rest of the events. MostDelta also shows a similar behaviour to MostAll, but the latter includes information on vertical temperature profiles, which play a key role in determining the precipitation phase at ground level. Another advantage of using an ensemble rather than a single algorithm is that agreement or disagreement between its members can be interpreted as forecast uncertainty. The latter provides a wider perspective to forecasters and may be helpful in precipitation phase transitions and in situations where the air temperature is close to freezing point and the difficulty in determining the precipitation phase is greater.

4.1.2 Article

Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J. R., Moré, J., Rigo, T. & Sairouni, A. (2021). Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(739), 3135-3153.

Received: 18 September 2020 | Revised: 16 May 2021 | Accepted: 23 June 2021 | Published on: 28 July 2021

DOI: 10.1002/qj.4121

Quarterly Journal of the Royal Meteorological Society RMetS

RESEARCH ARTICLE

Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts

Enric Casellas[1,2] | **Joan Bech**[1] | **Roger Veciana**[2] | **Nicolau Pineda**[2] | **Josep Ramon Miró**[2] | **Jordi Moré**[2] | **Tomeu Rigo**[2] | **Abdel Sairouni**[2]

[1]Department of Applied Physics–Meteorology, University of Barcelona, Barcelona, Spain

[2]Meteorological Service of Catalonia, Barcelona, Spain

Correspondence
E. Casellas, Department of Applied Physics–Meteorology, University of Barcelona, Barcelona, Spain.
Email: enric.casellas@meteo.ub.edu

Funding information
Industrial Doctorate Projects (Government of Catalonia) MINECO/FEDER MINECO, Grant/Award Numbers: DI-053/2017, CGL2015-65627-C3-2-R, CGL2016-81828-REDT/AEI, RTI2018-098693-B-C32

Abstract

Heavy snowfall events can cause substantial transport disruption and exert a negative socioeconomic impact, particularly in low-altitude and midlatitude regions, where it seldom snows. Such problems may be exacerbated if there are rapid transitions between different precipitation phases within the same event. Previous studies have addressed this issue using precipitation-phase nowcasting techniques, often focusing on critical infrastructures such as airports. Very short-range forecasts are usually based on trends of observations and numerical weather prediction models. Nowcasting schemes considering the precipitation phase generally merge extrapolated surface observations, modelled vertical temperature profiles, and extrapolated weather radar precipitation fields. In this study, a precipitation-phase nowcasting scheme was developed and evaluated, initially using eight different algorithms to classify precipitation into rain, sleet or snow, together with a probabilistic weather radar data extrapolation technique. In addition, three combinations of the previous algorithms were also evaluated. The nowcasting scheme was applied to a midlatitude region in the Northwestern Mediterranean to assess its performance during eight snowfall events. Single and combined algorithms were compared to determine their suitability in conditions close to freezing point, when there is increased uncertainty about the precipitation phase. The results indicate that, although single and combined algorithms perform similarly, the latter can provide valuable information during event monitoring. Precipitation phase transitions were also analysed, finding that on average they can be forecast correctly with a lead time of 120 min. The proposed methodology can be readily applied to other regions where ground-based observations, weather radar data, and model forecasts are available.

KEYWORDS

nowcasting, precipitation phase, snow, weather radar

1 | INTRODUCTION

Severe winter weather conditions can affect transport infrastructures, road safety, road maintenance, and everyday activities (Schmidlin, 1993; Andrey *et al.*, 2003; Papagiannaki *et al.*, 2013; Malin *et al.*, 2019). Nowcasting the precipitation phase can provide key information for meteorological services or emergency managers, who can then activate specific protocols to prevent or manage risks and harm to people and critical infrastructures (Vilaclara *et al.*, 2010; Bonelli *et al.*, 2011; Kann *et al.*, 2015). For example, airports are critical infrastructures sensitive to snowfall, and several ad hoc nowcasting systems have been developed to address this issue specifically, such as the Canadian Airport Nowcasting system (CAN-Now: Isaac *et al.*, 2014), Probabilistic Nowcasting of Winter Weather for Airports (PNOWWA: Saltikoff *et al.*, 2018), and Winter Hazards in Terminal Environment for Munich Airport (WHITE: Keis, 2015). Recent progress has also been reported regarding the use of medium-range forecasts (Gascón *et al.*, 2018; Fehlmann *et al.*, 2019).

Discrimination of the precipitation phase at the surface level has been widely studied using several approaches, including near-surface air-temperature threshold-based schemes (e.g., Koistinen and Saltikoff, 1998; Froidurot *et al.*, 2014; Behrangi *et al.*, 2018), decision-tree algorithms based on implicit assumptions about microphysical processes and vertical temperature profiles (e.g., Bourgouin, 2000; Schuur *et al.*, 2012), explicit algorithms based on hydrometeor mixing ratios at the ground level provided by complex microphysics parameterisations (e.g., Ikeda *et al.*, 2013; Benjamin *et al.*, 2016), and machine-learning algorithms (e.g., McGovern *et al.*, 2017; 2019). However, surface precipitation phase discrimination remains challenging, especially at temperatures close to freezing point. Therefore, a single methodology may not capture all variability involved in precipitation phase discrimination at surface level. In fact, Cortinas *et al.* (2002) and Wandishin *et al.* (2005) reported that no single precipitation phase algorithm based on vertical temperature profiles outperformed the rest all the time, suggesting that a combination of algorithms was superior to the single algorithm that performed the best. A similar conclusion was drawn when considering surface information alone, where threshold-based strategies may present very high spatial and temporal variability (Zhong *et al.*, 2018; Casellas *et al.*, 2021). In addition, in some models, thresholds for single events can be changed or adjusted dynamically (Pomeroy *et al.*, 2007).

Precipitation-phase nowcasting schemes generally form part of a nowcasting system, which tends to combine extrapolated observational data and numerical weather prediction (NWP) model output. Nowcasting is usually achieved by means of persistence or trends of observations, since their performance surpasses that of NWP models (Golding, 1998; Haiden *et al.*, 2011). Examples of surface precipitation type nowcasting include combined ground-temperature trend observations and NWP vertical temperature profiles in Integrated Nowcasting through Comprehensive Analysis (INCA: Haiden *et al.*, 2011), a decision-tree algorithm based on modelled vertical temperature profiles in CAN-Now (Isaac *et al.*, 2014), surface temperature conditions to classify different types of snow in PNOWWA (Saltikoff *et al.*, 2018), and features of modelled vertical temperature profiles and trends of observations in WHITE (Keis, 2015). All the aforementioned nowcasting schemes implement a weather radar extrapolation technique for precipitation, since this tends to outperform NWP models for the first 1.5–2 hr (Simonin *et al.*, 2017). In INCA, CAN-Now, and WHITE, a deterministic extrapolation is used, whereas PNOWWA uses a probabilistic approach.

For the sake of simplicity, and similarly to Ikeda *et al.* (2017), only three different precipitation types are considered here, namely rain, snow, and a mixture of both, hereafter mixed. Other precipitation types, such as freezing rain and ice pellets, were not considered, due to the lack of observations in the datasets analysed. Therefore, strictly speaking, this study focused on nowcasting precipitation phase rather than precipitation type. The proposed methodology was applied in the Northwest Mediterranean region, where snowfall at low altitudes can be considered exceptional or rare. Thus, the correct forecast of these events is critical, compared with higher altitude or latitude regions where snowfall is much more common. Consequently, we considered several precipitation phase algorithms based on modelled vertical temperature profiles and trends of surface observations. We evaluated the individual performance of eight selected precipitation phase schemes in eight low-altitude snowfall events that took place in the study area between January 2010 and January 2021. In addition, we evaluated two new proposed combinations of algorithms. Our analysis focused on precipitation phase classification and precipitation phase transitions throughout the region and at specific locations, including airports. This evaluation raised the following question: is there a single precipitation phase discrimination scheme capable of capturing the variability of low-altitude snowfall events, or is a combination of algorithms preferable?

The rest of the article is organised as follows. Section 2 describes the study area and the data sources considered. Section 3 presents the nowcasting scheme designed. Then, Section 4 provides validation results for single and combined precipitation phase discrimination algorithms in eight different low-altitude snowfall events in the study

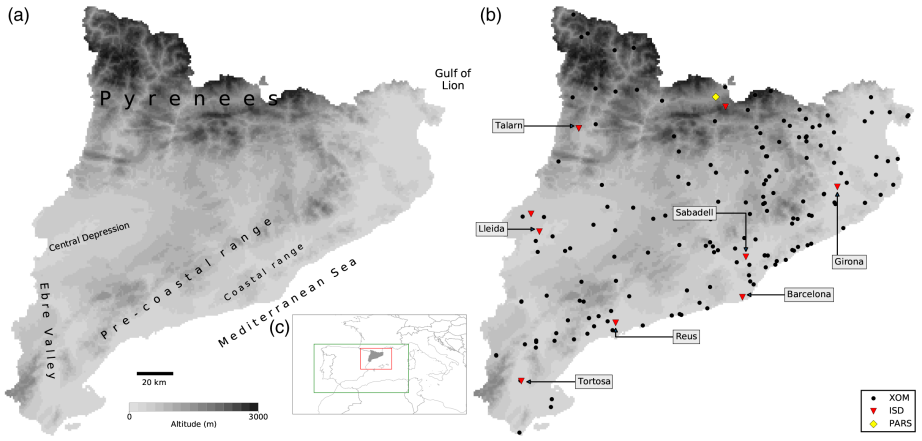

FIGURE 1 Study area, including (a) the main orographic characteristics, (b) the location of the precipitation phase observation sources considered: Xarxa d'Observadors Meteorològics (XOM) and Integrated Surface Database (ISD), and (c) the WRF model 9 km and 3 km domains [Colour figure can be viewed at wileyonlinelibrary.com]

area. Section 5 addresses the question of whether a combination of algorithms is preferable to a single algorithm, and, lastly, the study is summarised and the conclusions presented in Section 6.

2 | DATA

2.1 | Study area

The study area, Catalonia, is located in the northeast (NE) of the Iberian Peninsula and is bordered by the Pyrenees mountain range to the north, the Mediterranean Sea to the east and south, and the Ebre Valley to the west. It presents marked orographic complexity, with mountains exceeding 3,000 m above sea level (a.s.l.), mountainous ranges (~500–1,000 m a.s.l) close to the coastline, such as the coastal and precoastal ranges, and extensive plains (Figure 1a). Snowfall episodes generally occur between late autumn and early spring and are mainly restricted to mountainous areas. However, occasional snowfall at low altitudes has been reported below 700 m a.s.l, and can generate a significant socioeconomic impact because of infrastructure vulnerability and high population density (Bech *et al.*, 2013; Llasat *et al.*, 2014).

2.2 | Precipitation phase observations

Snow, mixed, and rain observations were collected from several sources. Eight events were selected between January 2010 and January 2021, applying the criteria of social impact and over 100 precipitation phase ground observations. These events, their identification code, and the number of observations are listed in Table 1. They included the March 8, 2010 event, which notably affected the city of Barcelona and the NE of Catalonia (Bech *et al.*, 2013; Llasat *et al.*, 2014), a three-day event in February 2018 and the so-called Storm *Filomena* in January 2021, which affected extensive parts of the Iberian Peninsula.

Two precipitation phase observation sources were used: the Integrated Surface Database (ISD) and Xarxa d'Observadors Meteorològics (XOM). The ISD (Smith *et al.*, 2011) includes airport (METAR) reports and synoptic surface observations (SYNOP) corresponding to nine locations in the study area, while the XOM consists of a trained volunteer meteorological observers' network, which reports the precipitation phase in addition to other meteorological variables. Observations are routinely measured at 0700, 1200, and 1800 UTC (Ripoll *et al.*, 2016). Figure 1b shows the location of observational sites and Table 2 gives the number of observations available from each source. Precipitation phase observations were filtered according to their air temperature values, excluding those that fell outside a (−7.5 to 5.0 °C) interval (Casellas *et al.*, 2021).

2.3 | Temperature and humidity observations

Some of the precipitation phase discrimination schemes considered here are based on interpolated surface

Event	Start date	Finish date	Duration (hr)	# Observations
E1	2010-01-07 06:00	2010-01-08 18:00	36	218
E2	2010-03-07 17:30	2010-03-08 19:30	26	275
E3	2011-03-03 06:00	2011-03-04 17:00	35	118
E4	2013-02-22 06:00	2013-02-23 08:30	27	124
E5	2015-02-03 06:00	2015-02-05 08:00	50	204
E6	2018-02-04 07:00	2018-02-06 17:00	58	189
E7	2018-02-26 14:30	2018-03-01 02:00	60	376
E8	2021-01-08 20:00	2021-01-11 18:00	70	379
		Total	362	1,883

TABLE 1 Summary of selected events including their identification, start, and finish dates (in YYYY-mm-dd HH:MM UTC format), duration, and the number of precipitation phase observations

TABLE 2 Precipitation phase observations from the Integrated Surface Database (ISD) and the Xarxa d'Observadors Meteorològics (XOM)

Source	Rain	Mixed	Snow	Total
ISD	657	40	263	960
XOM	407	87	429	923
Total	1,064	127	692	1,883

temperature and humidity fields (see Section 3.1.1). These fields were obtained using dry-bulb air temperature (T_a) and dew-point temperature (T_d) observations acquired from the automatic weather station (AWS) network of the Meteorological Service of Catalonia (Xarxa d'Estacions Meteorològiques Automàtiques (XEMA): Serra *et al.*, 2016) with a 30-min temporal resolution. Interpolated fields of T_a and T_d with a horizontal resolution of 1 km were calculated using multiple linear regression, anomaly correction, and clustering (Casellas *et al.*, 2020). Then, relative humidity (RH) and wet-bulb temperature (T_w) fields were derived from the T_a and T_d fields following Lawrence (2005) and Sadeghi *et al.* (2013), respectively.

2.4 | Weather radar data

The Xarxa de Radars Meteorològics (XRAD), a meteorological radar network consisting of four single-polarisation C-band Doppler weather radars located at specific locations to reduce radar beam blockage (Bech *et al.*, 2003; Trapero *et al.*, 2009), provided weather radar observations covering the study area. The lowest height of the radar reflectivity composite was used after the application of automatic quality-control procedures (Sánchez-Diezma *et al.*, 2002; Bech *et al.*, 2005; Franco *et al.*, 2006; Altube *et al.*, 2015; 2017). The temporal resolution of the data was 6 min and the spatial resolution was 1 km.

2.5 | Temperature and humidity profiles

The modelled vertical temperature and humidity profiles required for some of the algorithms used were provided by the Weather Research Forecasting (WRF) model with the ARW (Advanced Research WRF) solver, version 3.5.1 (Skamarock *et al.*, 2008). The use of soundings was ruled out because of their lack of suitability for nowcasting purposes, since only twice-daily data from one location (Barcelona) were available for the study area. The WRF model was configured in three nested domains: horizontal grid lengths of 27, 9, and 3 km with 31 sigma levels and applying one-way nesting (Figure 1c). The set up for the main physical parameterisations was as follows: WRF Single-Moment (WSM) five-class microphysics scheme (Hong *et al.*, 2004; 2006), Kain–Fritsch (new Eta) cumulus scheme (Kain, 2004), Dudhia scheme (Dudhia, 1989) for short-wave radiation, RRTM (Mlawer *et al.*, 1997) for long-wave radiation, YSU scheme (Hong *et al.*, 2006) for PBL, and NOAH land-surface model (Chen and Dudhia, 2001). ECMWF-IFS data were employed as initial and boundary conditions for the coarsest domain. For the innermost domain at 3 km, WRF was updated every 3 hr (0000, 0300, 0600, 0900, 1200, 1500, 1800, 1900, and 2100 UTC), taking advantage of the improved initial conditions obtained with a data assimilation cycle based on three-dimensional variational analysis (3DVAR) conducted by the WRF Data Assimilation System (WRF-DA). In particular, reflectivity and radial wind velocity from the weather radar network (XRAD) and surface observation data (XEMA and METAR) were assimilated in a 3-hr cycle. The resulting WRF simulations at 3 km (namely WRF-3DVAR, hereafter WRF) provided 12-hr forecasts at an hourly resolution that

were linearly interpolated to a 30-min temporal resolution, following Isaac *et al.* (2014) and Keis (2015), to render them consistent with AWS network observations.

3 | NOWCASTING SCHEME

The proposed precipitation-phase nowcasting scheme consists of two parts: classification of the hydrometeor phase at surface level and temporal extrapolation of the precipitation field. The former is performed by considering extrapolated surface observations and NWP model vertical temperature profiles, while the latter is obtained using weather radar data and a probabilistic advection scheme. Hence, precipitation phase nowcasting involves forecasting both the hydrometeor classification and the spatial evolution of the precipitation field. The scheme was designed to produce 3-hr lead time forecasts with a 30-min temporal resolution, a 1 km spatial resolution, and 30-min frequency updates. From now on, the forecast lead time refers to the lead time of the nowcasting scheme.

The process to obtain a precipitation phase nowcast involved the following steps:

1. calculation of the precipitation phase at surface level for each lead time;
2. calculation of the precipitation field extrapolation using a probabilistic approach; and
3. combination of the precipitation phase field calculated in step 1 and the extrapolated precipitation field calculated in step 2.

Each step of the process is detailed in the following subsections.

3.1 | Step 1: Precipitation phase classification

Eleven precipitation phase discrimination algorithms were considered in this study and are presented in this section. Thus, we analysed eight single algorithms (four based on using surface information only, and four based on modelled vertical temperature profiles) and three schemes that combined different algorithms. A description of each algorithm is provided below.

3.1.1 | Surface information

Precipitation phase classification using surface information alone was performed using threshold-based schemes relying on physical variables. For example, if the air temperature is cooler (warmer) than a threshold, precipitation is classified as snow (rain). In this study, we considered four dual threshold (DT) schemes, which establish two thresholds to classify precipitation into three different types: rain, mixed, and snow. The selected DT schemes are based on different meteorological variables: air temperature (DT-T_a: Liu, 2008), dew-point temperature (DT-T_d: Marks *et al.*, 2013), wet-bulb temperature (DT-T_w: Behrangi *et al.*, 2018), and the Koistinen and Saltikoff (1998) formula (DT-KS). The final of these gives the probability that precipitation will be in the form of snow ($p(snow)$) and is based on an empirical formula using air temperature (T_a) and relative humidity (RH) as follows:

$$p(snow) = 1 - \frac{1}{1 + e^{22 - 2.7 \cdot T_a - 0.2 \cdot RH}}. \quad (1)$$

Based upon this probability, the thresholds used were $< 40\%$ for rain, $> 58\%$ for snow, and between the two for mixed. Similarly, the rest of the schemes were based on establishing two thresholds, one to discriminate between snow and mixed and another to discriminate between mixed and rain. In the case of wet-bulb temperature (DT-T_w), 0.7 and 1.0 °C were used as snow–mixed and mixed–rain thresholds, respectively. The thresholds used for the other two schemes were 0.1 and 0.4 °C for DT-T_d and 1.0 and 1.4 °C for DT-T_a. The selection of these schemes and their corresponding thresholds was based on an analysis carried out in the study area (Casellas *et al.*, 2021).

Diagnostic DT schemes were built using interpolated T_a, T_d, T_w, and RH fields from T_a and T_d observations (see Section 2.3). Extrapolation of these fields to a 3-hr lead time was performed considering the forecast variations of the analogous WRF model fields, adapted from Haiden *et al.* (2011) (hereafter, DeltaForecast). In addition, we blended the extrapolated fields with those of the WRF model. The process to obtain the extrapolated T_a fields at a future time step *i+1* can be expressed by adding the model tendency between time step *i* and time step *i+1* as

$$T_{a_{i+1}}^{\text{Delta}} = T_{a_i}^{\text{Delta}} + (T_{a_{i+1}}^{\text{WRF}} - T_{a_i}^{\text{WRF}}), \quad (2)$$

when $i = 0$, T_{a_o} corresponds to the observed T_a. Then, the blending with the WRF output was achieved through

$$T_{a_i}^{\text{DeltaForecast}} = g(i) \cdot T_{a_i}^{\text{Delta}} + (1 - g(i)) \cdot T_{a_i}^{\text{WRF}}, \quad (3)$$

where $g(i)$ is a weight function ($e^{-i/8}$) and i is the lead time (in hours). The process was repeated for T_d, while T_w and RH were obtained from the T_a and T_d fields following Sadeghi *et al.* (2013) and Lawrence (2005), respectively. From now on, forecast DT schemes are termed Delta T_a for air temperature, Delta T_d for dew-point temperature, Delta

T_w for wet-bulb temperature, and Delta KS for Koistinen and Saltikoff (1998).

Since the spatial resolution of the WRF model output is 3 km, it was downscaled to the resolution of the interpolated fields of surface observations, that is, from 3 to 1 km. Downscaling of T_a and T_d was achieved using the so-called elevation correction (Sheridan *et al.*, 2010; Rouf *et al.*, 2020). This accounts for differences between the orography used by the WRF model and that used by a finer source. Generally, this method is applied to regions with a complex orography where differences between the two orography sources are marked, such as deep valleys or high mountain ridges (Haiden *et al.*, 2011).

3.1.2 | Vertical thermodynamic profiles

We studied four precipitation type algorithms based on vertical thermodynamic information from the WRF model: Ramer (Ramer, 1993), Baldwin (Baldwin *et al.*, 1994), Bourgouin (Bourgouin, 2000) and Schuur (Schuur *et al.*, 2012). The algorithms consider wet-bulb temperature profiles, melting and freezing energies, and warm and cold layer depths, which are used to determine the precipitation type at ground level. A general overview of the algorithms is provided here and more detailed descriptions are given in the above references and in Cortinas *et al.* (2002) and Wandishin *et al.* (2005).

The Ramer algorithm (Ramer, 1993) is based on the ice fraction of hydrometeors (I) and its changes between a defined precipitation generation layer and the surface. The precipitation generation layer corresponds to the highest saturated layer, and the T_w at that level determines the hydrometeor characteristics: supercooled droplets or ice. The ice fraction (I) is 1 for ice and 0 for supercooled droplets. Then, changes in I from the precipitation generation layer to the surface are calculated layer by layer as follows:

$$\Delta I = \frac{(0 - T_w)(\ln(p_{z-1}) - \ln(p_z))}{0.045 \cdot RH}, \quad (4)$$

where ΔI corresponds to the ice fraction change, T_w to the layer mean wet-bulb temperature, RH to the layer mean relative humidity, and p to air pressure. z identifies each layer, starting from the precipitation generation layer and using a top-down approach until the surface layer is reached. If $I < 0.04$ at the lowest model level, precipitation is classified as rain when $T_w > 0\,°C$ and as freezing rain when $T_w \leq 0\,°C$. If $I > 0.85$, ice pellets are assumed, but snow if $I = 1$. If $0.04 \leq I \leq 0.85$, precipitation is classified as freezing rain when $T_w < 0\,°C$ at the lowest model level or as a mixture of types when $T_w \geq 0\,°C$.

Similarly to Ramer (1993), the Baldwin algorithm (Baldwin *et al.*, 1994) classifies the initial phase of the hydrometeors at the precipitation generation level using T_w. Solid precipitation at the highest saturated level is considered when $T_w < -4°C$, otherwise the liquid phase is assumed. Precipitation type at ground level is obtained using a decision-tree algorithm based on the depth and strength of cold and warm layers. These are calculated from the area between the vertical T_w profile and $T_w = 0\,°C$ for all layers from the highest saturated level to the surface. Then, phase changes of the established initial phase are evaluated considering calculated melting (warm layers) and freezing (cold layers) potential together with surface air temperature. If the melting area is large enough, rain is diagnosed. Otherwise, a decision tree is followed to classify precipitation into snow, ice pellets, or freezing rain, which is ultimately determined using surface T_a. For example, snow is diagnosed if $T_w < -4\,°C$ is reported in any layer between the surface and the highest saturated level and the area between $-4\,°C$ and the T_w profile does not exceed an empirically derived value (Baldwin *et al.*, 1994).

The Schuur algorithm (Schuur *et al.*, 2012) consists of two steps. In the first one, the precipitation type is retrieved from the T_w model profile, while the second one is a refinement of the previous classification based on dual-polarisation weather radar data. Here, only the first step was implemented. The algorithm is based on the form of vertical T_w profiles considering the number of crossing points of 0 °C. Four types of profile are defined depending on the number of crossings. Precipitation type is determined for each profile taking into account its type and the depth of melting and freezing layers. In addition, surface T_w is also used to classify precipitation type. For example, if surface $T_w \geq 3°C$, then precipitation is classified as rain. In contrast, if T_w is below 0 °C at the surface and also for the entire vertical profile, precipitation is classified as snow.

The Bourgouin algorithm (Bourgouin, 2000) considers vertical temperature (T_a) profiles and the melting and freezing energies available to change the phase of hydrometeors. The energy corresponds to areas calculated according to the departure of the observed or modelled vertical T_a profile from $T_a = 0\,°C$. If the T_a profile is below 0 °C, a cold layer is diagnosed and contributes to the freezing energy. Otherwise, a warm layer is identified and contributes to the melting energy. Precipitation type at the surface level is determined by comparing the melting and freezing energies. In addition, surface T_a is used to discriminate between freezing rain and rain.

As explained in the Introduction, owing to the limited number of different precipitation type observations (Table 2), only rain, snow and mixed were considered.

Therefore, the outcome of the algorithms described above was narrowed to rain, mixed, and snow. Consequently, ice pellets (Ramer, Bourgouin, Schuur, Baldwin) were reclassified as snow, and freezing rain (Ramer, Bourgouin, Schuur, Baldwin) and wet snow (Schuur) as mixed (Ikeda *et al.*, 2017).

3.1.3 | Combination of algorithms

Besides the eight algorithms presented above, three additional combinations of algorithms were also considered, resulting in 11 different strategies to classify precipitation into rain, mixed or snow. The first combination has already been tested in previous studies (Cortinas *et al.*, 2002; Manikin, 2005; Wandishin *et al.*, 2005) and consists of considering the ensemble of vertical temperature profile algorithms (Baldwin, Bourgouin, Ramer, and Schuur), hereafter Profiles. The other two combinations are proposed in this study. One includes schemes based on extrapolated surface observations only (Delta T_a, Delta T_d, Delta T_w, and Delta KS), hereafter MostDelta, while the other combines all eight algorithms (Delta T_a, Delta T_d, Delta T_w, Delta KS, Baldwin, Bourgouin, Ramer, and Schuur), hereafter MostAll. For each of these strategies, the most probable precipitation phase from among the algorithms considered was selected for every pixel in the field.

3.2 | Step 2: Nowcasting of precipitation fields

The nowcasting of precipitation fields was conducted using the extrapolation of weather radar data, because, depending on the weather situation, it generally outperforms NWP forecasts for the first forecasting hours (Berenguer *et al.*, 2012; Simonin *et al.*, 2017). We selected a probabilistic approach, the Short-Term Ensemble Prediction System (STEPS), which was developed by the UK Met Office and the Australian Bureau of Meteorology (Bowler *et al.*, 2006). STEPS is based on perturbing deterministic Lagrangian extrapolation of weather radar data fields through stochastic noise to account for unpredictable precipitation growth and decay processes (Seed, 2003). In this study, it was applied using pySTEPS (Pulkkinen *et al.*, 2019), a Python library that implements different precipitation nowcasting approaches, including STEPS.

STEPS was configured to calculate 20-member ensemble, 3-hr lead time forecasts with 6-min temporal resolution output fields. The nowcast was produced using a nonparametric noise generator and the motion field was estimated using the Lucas–Kanade method. The spatial resolution of the output fields was that of the native weather radar data, 1 km. The outcome of the precipitation nowcasting scheme was processed to obtain the probability fields of exceeding 0.1 mm of precipitation in 30 min.

3.3 | Step 3: Outcome of the precipitation-phase nowcasting scheme

Once the precipitation phase and the precipitation probability forecast fields had been calculated, they were merged onto the same geographic coordinate with a 1 km spatial resolution. Then, they were combined to obtain 30-min temporal resolution fields including the probability of exceeding 0.1 mm of precipitation, the precipitation phase alone, and the precipitation phase masked with precipitation probability.

4 | RESULTS

Validation of the precipitation-phase nowcasting scheme focused on the problem of distinguishing the precipitation phase rather than on the occurrence of precipitation, following the approach of Wandishin *et al.* (2005). Therefore, we determined the capacity of different strategies to discriminate the precipitation phase independently of the occurrence of forecast precipitation for eight selected low-altitude snowfall events.

Validation was conducted considering a spatial fuzzy verification scheme with a $9 \times 9\,km^2$ neighbourhood area around the nearest pixel of the observation (Ebert, 2008). Therefore, point-to-grid comparison limits potential misclassification due to a short grid distance between an observation and the nearest pixel (Ebert, 2008; Jolliffe and Stephenson, 2012). The comparison between a point observation and grid forecast was performed as follows: (a) selection of the 30-min forecast window that matched the observation time, (b) calculation of the nearest pixel in the grid to the observation location and selection of $9 \times 9\,km^2$ neighbourhood area, and (c) comparison of the present weather observation and the precipitation phase estimation from the nowcasting scheme grid forecast to evaluate the capacity of the scheme to discriminate between snow, mixed, and rain. For example, the capacity to discriminate snow individually was evaluated by comparing observations of snow versus no snow, that is, snow against rain and mixed observations. Since a fuzzy verification scheme was employed, we selected the most frequent precipitation phase in the 9×9 km neighbourhood area. Each step was repeated for each 30-min interval of the 180-min forecast lead time. The WRF model run used in each nowcasting calculation was the last one available at the time the nowcasting was conducted.

Performance of the nowcasting schemes was evaluated using a contingency table and calculating different verification scores. The Gerrity skill score (GSS: Gerrity, 1992) was estimated to evaluate three precipitation phases simultaneously. The value of GSS is dependent on the order of the classes, which in this study were ordered according to a "physical" criterion (Elmore *et al.*, 2015): rain, mixed, and snow. Therefore, a misclassification of rain as mixed is penalised less than a misclassification as snow. In addition, we calculated the Peirce skill score (PSS), which measures the capacity to discriminate between events and non-events, the probability of detection (POD), and the false-alarm rate (FAR: Jolliffe and Stephenson, 2012). The best score for GSS, PSS, and POD is 1 and the worst is 0, whereas for FAR it is the opposite.

A bootstrapping technique (Efron and Tibshirani, 1994; Jolliffe, 2007) was used with 1,000 iterations to obtain a sampling distribution of the scores, defining their mean and 95% confidence level. Then, as in McCabe *et al.* (2016) and Moon and Kim (2020), Wilcoxon tests for paired samples (Wilks, 2011) were performed to assess the statistical significance of score differences between schemes. The tests were performed for all lead times individually and for the average scores of each scheme. In this study, differences were considered significant if the *p*-value was below .001.

TABLE 3 Mean verification skill scores and 95% confidence interval obtained from bootstrap sampling for the precipitation phase discrimination strategies

Scheme	GSS	PSS snow	PSS rain
MostDelta	**0.62 (0.61–0.63)**	**0.63 (0.61–0.64)**	0.61 (0.60–0.63)
MostAll	0.61 (0.60–0.62)	0.61 (0.60–0.62)	0.61 (0.60–0.62)
Profiles	0.54 (0.53–0.55)	0.55 (0.54–0.57)	0.53 (0.52–0.54)
Schuur	0.53 (0.51–0.54)	0.52 (0.50–0.53)	0.53 (0.52–0.55)
Baldwin	0.51 (0.50–0.53)	0.53 (0.52–0.54)	0.51 (0.49–0.52)
Bourgouin	0.41 (0.39–0.42)	0.41 (0.39–0.42)	0.40 (0.39–0.41)
Ramer	0.51 (0.50–0.52)	0.52 (0.51–0.53)	0.50 (0.49–0.52)
Delta T_w	**0.62 (0.61–0.63)**	0.62 (0.60–0.63)	0.62 (0.61–0.63)
Delta T_d	0.60 (0.59–0.61)	0.61 (0.60–0.62)	0.59 (0.57–0.60)
Delta KS	**0.62 (0.61–0.63)**	**0.63 (0.61–0.64)**	**0.62 (0.61–0.63)**
Delta T_a	0.56 (0.55–0.58)	0.54 (0.53–0.55)	0.59 (0.58–0.60)

Note: Scores correspond to the mean value over a 180-min lead time for all events. The Gerrity skill score (GSS) evaluates each strategy from a multicategorical perspective (3 × 3 contingency table), while the Peirce skill score (PSS) evaluates dichotomous events (2 × 2 contingency table), applied to snow versus no snow and rain versus no rain observations. Wilcoxon tests indicate that score differences between all schemes shown in this table are significant. Bold values indicate the best performing discrimination strategies for each score.

4.1 | Overview of performance of precipitation phase algorithms

According to mean GSS values considering all events, the algorithms based on extrapolated surface observations (Delta T_a, Delta T_d, Delta T_w, and Delta KS) outperformed those based on NWP model vertical temperature profiles (Baldwin, Bourgouin, Ramer, and Schuur: see Table 3). Among the schemes including surface information only, Delta KS and Delta T_w exhibited similar performance, but presented some statistically significant differences. The performance of Delta KS and Delta T_w was superior to that of Delta T_a and Delta T_d. Regarding NWP model vertical temperature profile based algorithms, Schuur showed the best performance (0.53), closely followed by Ramer and Baldwin (0.51), and then Bourgouin with a GSS of 0.41 (Figure 2a and Table 3). An analysis of the evolution of mean GSS over forecast lead time showed that extrapolated surface observation schemes presented a decreasing trend, starting at ∼0.7 for the first 30 min of lead time and reaching ∼0.5 at 180 min of lead time. Delta T_a and Delta T_d were surpassed by the NWP model-based algorithms beyond 120 min of lead time for Delta T_a and 150 min of lead time for Delta T_d (Figure 2a). Results for the NWP model-based algorithms are shown as straight lines, since the same model forecast was used for different lead times in the nowcasting scheme, providing constant verification scores. Therefore, they simply provide a reference benchmark for the performance of NWP model-based algorithms (Figure 2a).

The mean GSS verification results of combining different algorithms are presented in Table 3. MostDelta exhibited the best performance (0.62), closely followed by MostAll (0.61), but with a statistically significant difference. As for Profiles, this exhibited better performance than any single NWP model-based scheme, but did not outperform the combination of surface extrapolated observation schemes (Table 3). Regarding the evolution of these combinations of schemes over lead time, MostDelta was superior during the first 120 min, and from then on MostAll was the best (Figure 2b).

The Peirce skill score (PSS) was used to evaluate algorithm performance for individual precipitation phases, obtaining similar results to those for GSS. PSS values showed that single and combined dual threshold schemes and their combination with vertical temperature profiles exhibited the best performance for either snow or rain observations (Table 3). Poor performance was obtained for the mixed precipitation phase, with PSS values below 0.1 (not shown). All methods obtained a very high POD for rain and snow observations, and FAR scores were ∼0.20 for all the schemes (not shown).

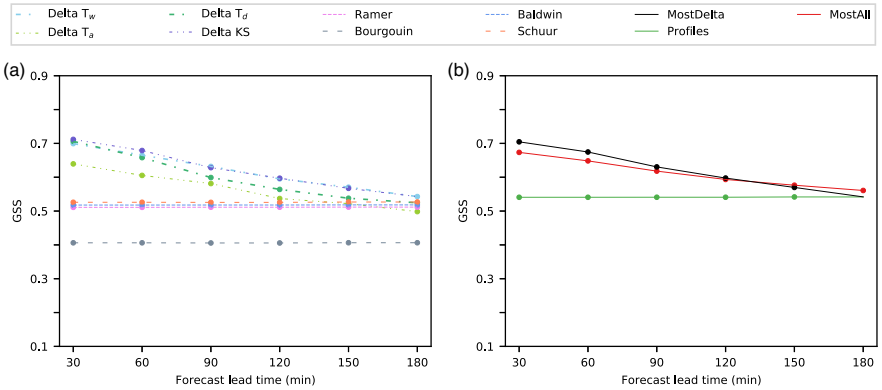

FIGURE 2 Mean Gerrity skill score (GSS) obtained from bootstrap sampling over the forecast lead time for each precipitation phase discrimination algorithm considering all available observations. (a) Scores for single precipitation phase algorithms. (b) Scores for combined precipitation phase algorithms: MostDelta for extrapolated surface information schemes, Profiles for vertical temperature profile algorithms, and MostAll for a combination of both. Bullets indicate statistical significance of score differences relative to all schemes for each lead time [Colour figure can be viewed at wileyonlinelibrary.com]

4.2 | Event-to-event variability of algorithm performance

The performance of the single and combined algorithms for each event is presented in Tables 4 and 5, including the mean GSS and 95% confidence interval obtained by means of bootstrap sampling. Additionally, Figure 3 illustrates the evolution of mean GSS over the forecast lead time.

A high event-to-event variability was observed among algorithms, including those using surface information only, since none of them exhibited the best performance in each of the eight events (Table 4 and 5) at all times (Figure 3). For example, Delta T_w was the best in E3, E7, and E8, Delta T_d in E1 and E4, Delta T_a in E6, and Delta KS in E5 (p-value $< .001$). Regarding E2, Delta T_a and Delta KS presented a similar result and the null hypothesis could not be rejected; in contrast, Delta T_d showed statistically significantly better performance.

NWP model vertical profile schemes showed the same variability. Schuur presented the best performance in three out of the eight events (E1, E5, and E7), Ramer (E3 and E8) and Baldwin (E4 and E6) in two of them, and Bourgouin outperformed all other NWP model schemes in E2 (Tables 4 and 5). As regards the combinations of algorithms, Profiles was the best combination in E6, but MostAll and MostDelta dominated the rest of the events (Table 4 and 5). Although MostDelta showed the best performance of all combination schemes in E2, it was similar to Delta T_a or Delta KS and the null hypothesis could not be rejected. In E8, however, MostDelta was statistically significantly better than the other combinations. Meanwhile, MostAll presented the best performance in three events (E1, E3, E7), not only among combinations of algorithms but also compared with single algorithms.

Figure 3 shows the mean GSS of the algorithms for each event over the forecast lead time. As can be seen, in five out of the eight events considered, some of the algorithms based on vertical temperature profiles and their combination (Profiles) outperformed the two best extrapolated surface observation schemes (Delta KS and Delta T_w): over the entire forecast lead time in E1 (Figure 3a) and partially in E5, E6, E7, and E8 (Figure 3e,f,g,h, respectively). Figure 3 also indicates the statistical significance of score differences between a scheme and the rest with a marker symbol (otherwise, differences were not statistically significant). Although some of the differences between schemes were not statistically significant for some of the forecast lead times, as in the case of E1 (at 180 min lead time between MostDelta and Delta T_d), E3 (between 30 and 90 min forecast lead time between MostDelta and Delta KS), or E7 (at 30 min forecast lead time between Delta T_a and Baldwin), their differences can be statistically significant when all forecast lead times of an event are considered (Tables 4 and 5). Profiles outperformed all single NWP model-based algorithms in three events, but was surpassed in the others. However, the

TABLE 4 Mean Gerrity skill score (GSS) and 95% confidence interval obtained from bootstrap sampling for the precipitation phase discrimination strategies in events E1–E4

Scheme	E1 2010-01-07	E2 2010-03-07	E3 2011-03-03	E4 2013-02-22
MostDelta	0.65 (0.60–0.70)	0.43 (0.38–0.47)*	0.78 (0.72–0.85)	0.56 (0.50–0.64)
MostAll	**0.74 (0.69–0.78)**	0.34 (0.29–0.39)	**0.81 (0.75–0.86)**	0.56 (0.49–0.64)
Profiles	0.72 (0.67–0.77)	0.19 (0.13–0.25)	0.76 (0.70–0.83)	0.50 (0.40–0.58)
Schuur	0.73 (0.68–0.78)	0.18 (0.13–0.23)	0.70 (0.63–0.77)	0.52 (0.44–0.60)
Baldwin	0.69 (0.64–0.74)	0.23 (0.17–0.30)	0.71 (0.65–0.78)*	0.53 (0.46–0.61)
Bourgouin	0.47 (0.40–0.53)	0.24 (0.19–0.30)	0.75 (0.67–0.82)	0.38 (0.28–0.48)
Ramer	0.71 (0.67–0.76)	0.22 (0.15–0.28)	0.76 (0.70–0.83)	0.47 (0.39–0.55)
Delta T_w	0.63 (0.58–0.68)	0.42 (0.37–0.47)	0.78 (0.72–0.85)	0.56 (0.48–0.63)
Delta T_d	0.68 (0.63–0.73)	**0.43 (0.38–0.48)**	0.77 (0.72–0.83)*	**0.61 (0.54–0.68)**
Delta KS	0.67 (0.62–0.72)	0.43 (0.38–0.48)*	0.77 (0.71–0.84)*	0.57 (0.50–0.64)
Delta T_a	0.54 (0.49–0.60)	0.43 (0.38–0.47)*	0.72 (0.66–0.78)*	0.45 (0.38–0.53)

Note: Scores correspond to the mean value over a 180-min forecast lead time. The asterisk superscript (*) indicates that differences between this scheme and at least another one are not significant ($p > .001$). Bold values indicate the best performing discrimination strategies for each score.

TABLE 5 As Table 4, for E5–E8

Scheme	E5 2015-02-03	E6 2018-02-04	E7 2018-02-26	E8 2021-01-08
MostDelta	0.62 (0.58–0.67)	0.54 (0.45–0.63)	0.68 (0.66–0.71)	0.68 (0.64–0.71)
MostAll	0.62 (0.57–0.67)	0.52 (0.43–0.63)	**0.70 (0.68–0.73)**	0.67 (0.63–0.71)*
Profiles	0.56 (0.50–0.62)	0.55 (0.46–0.64)*	0.67 (0.64–0.70)	0.57 (0.52–0.61)
Schuur	0.48 (0.42–0.54)	0.49 (0.40–0.58)	0.65 (0.63–0.68)	0.60 (0.56–0.64)
Baldwin	0.39 (0.33–0.45)	0.55 (0.47–0.64)	0.62 (0.59–0.64)	0.52 (0.47–0.57)
Bourgouin	0.36 (0.30–0.42)	0.49 (0.39–0.59)	0.54 (0.51–0.57)	0.30 (0.26–0.35)
Ramer	0.45 (0.39–0.51)	0.50 (0.43–0.59)	0.60 (0.57–0.63)	0.63 (0.59–0.67)
Delta T_w	0.62 (0.57–0.67)	0.55 (0.46–0.64)*	0.69 (0.66–0.71)	**0.69 (0.65–0.72)**
Delta T_d	0.49 (0.43–0.54)	0.51 (0.42–0.61)	0.64 (0.62–0.67)	0.61 (0.57–0.65)
Delta KS	**0.63 (0.58–0.68)**	0.54 (0.45–0.62)	0.68 (0.65–0.70)	0.68 (0.64–0.71)
Delta T_a	0.59 (0.54–0.64)	**0.56 (0.48–0.65)**	0.58 (0.55–0.61)	0.67 (0.64–0.70)*

Note: Bold values indicate the best performing discrimination strategies for each score.

Profiles combination was superior to any single NWP model-based algorithm when all events were considered (Table 3).

4.3 | Precipitation phase transitions

During the eight events considered, different precipitation phase transitions took place and some of them are analysed in the following subsections regarding correct determination of the precipitation phase and its timing.

4.3.1 | Precipitation phase determination: Single and combination of algorithms

The behaviour of one of the best-performing schemes according to Figure 2 and Table 3, Delta KS, was compared against two combinations of algorithms, MostDelta and MostAll. For example, Figure 4 shows a snow to rain transition in Barcelona on February 27, 2018, between 1330 and 1630 UTC. It also includes several meteorological variables between 0900 and 2100 UTC, together with precipitation phase observations. Interestingly, it can be seen that snow was observed at a warmer air temperature than rain (Figure 4a). Focusing on the precipitation phase transition (grey box in Figure 4a), rain was forecast by Delta KS during the entire forecast lead time. At the same time, the most probable precipitation phase according to MostDelta and MostAll forecasts was also rain, but both reported small percentages of snow and mixed.

Other precipitation phase transitions are shown in Figure 5. During E1 on January 8, 2010, a transition from rain to snow occurred in Reus, and, according to Figure 5a, Delta KS, MostDelta, and MostAll all failed to

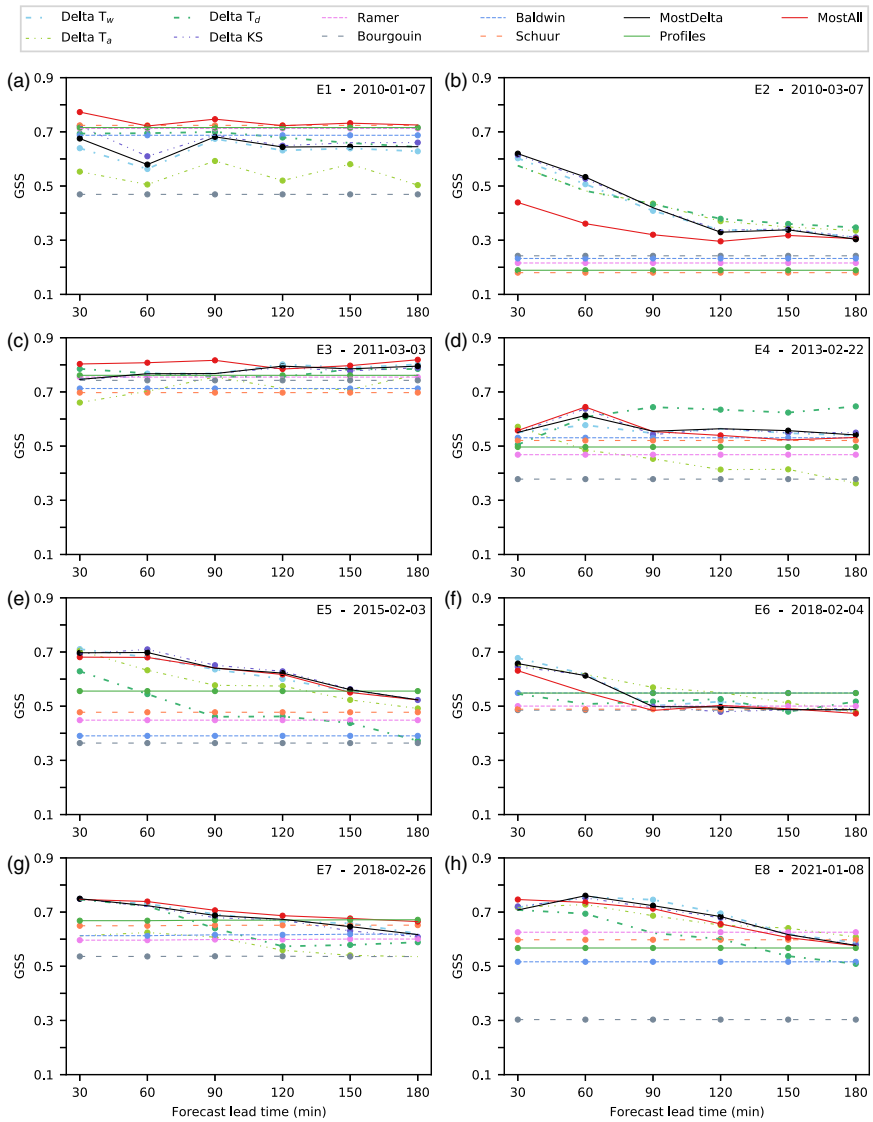

FIGURE 3 Mean Gerrity skill score (GSS) obtained from bootstrap sampling over the forecast lead time for each precipitation phase discrimination algorithm and each event. Bullets indicate the statistical significance of score differences relative to all schemes for each lead time [Colour figure can be viewed at wileyonlinelibrary.com]

forecast it correctly. On the one hand, Delta KS forecast rain throughout the entire forecast lead time, except for 30 min when mixed precipitation was forecast. On the other hand, MostAll presented a wide disparity between algorithms, with ~50% probability of rain, ~40% of snow, and ~10% of mixed. MostDelta showed high probabilities

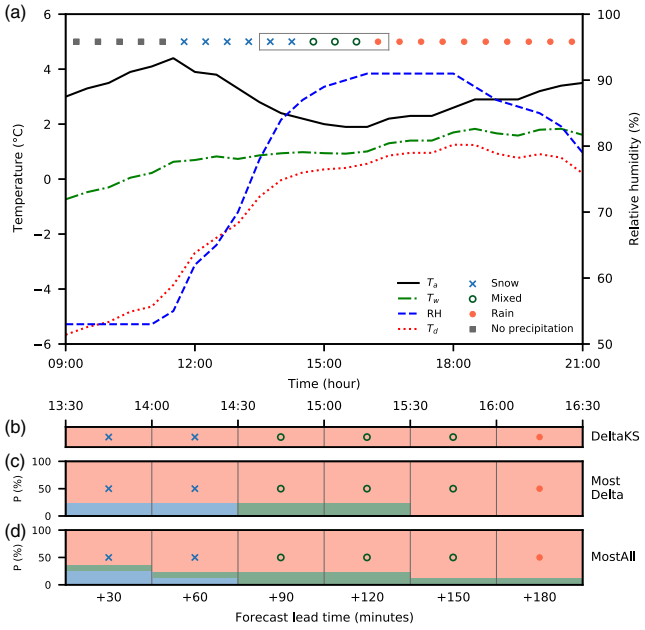

FIGURE 4 (a) Air temperature (T_a), dew-point temperature (T_d), wet-bulb temperature (T_w), and relative humidity (RH) observations on February 27, 2018, between 0900 and 2100 UTC for Barcelona. Markers represent the actual precipitation phase observed (snow, mixed, rain and no precipitation). (b,c,d) Nowcasting for Barcelona initialised at 1300 UTC (valid for the grey box in (a)), for (b) Delta KS, (c) MostDelta, and (d) MostAll schemes. Shaded colours indicate the probability of each precipitation phase and markers represent the actual precipitation phase observed [Colour figure can be viewed at wileyonlinelibrary.com]

for rain, but some mixed and snow probabilities were also reported. Another example took place in Girona during E4 on February 22, 2013. The three schemes correctly forecast the rain to snow transition with a lead time of 150 min (Figure 5b). In this case, Delta KS showed 30 min of mixed precipitation before estimating snow, while MostDelta and MostAll forecast an increasing probability of snow over the forecast lead time. The last two examples in Figure 5 show two forecasts where rapid precipitation phase transitions took place: in Figure 5c (event E6), rain was observed between two observations of snow, and in Figure 5d (event E8), mixed precipitation was reported between two rain observations. In both cases, Delta KS forecast rain for the entire lead time, whereas MostDelta forecast some mixed and snow probabilities and MostAll reported similar probabilities for rain and snow.

4.3.2 | Application of a combination of algorithms and a precipitation field

The examples above concerned precipitation phase estimates at specific locations without considering precipitation occurrence information. In this section, we examine the combination resulting from merging the precipitation phase field and precipitation field forecast for three events. In this case, we shall focus on application of the MostAll algorithm. For example, on March 8, 2010, at 1100 UTC, a widespread snowfall affected most of the study area, but at that time it was raining in Barcelona (Figure 6a). Two hours later, rain transitioned to snow, as shown by the crosses in Figure 6b,c. The nowcasting scheme based on the MostAll combination forecast a precipitation phase transition from rain to snow at 1230 UTC, which was actually observed. Therefore, the transition from rain to snow was forecast correctly with a lead time of 120 min. At the same time, in Girona, which is located in the NE of the region, snow was observed but rain was forecast (Figure 6a,b).

Another example of precipitation phase transition occurred during Storm *Filomena* on January 9, 2021 (E8). At 0900 and 1000 UTC snow was observed in Lleida, but then it transitioned to rain at 1100 UTC (Figure 6d,e). This phase change was forecast correctly by the nowcasting scheme with a 150-min lead time, with a probability exceeding 50% for rain among the eight algorithms included in MostAll (Figure 6f). In this nowcast, the precipitation area forecast missed some observations in the southern part of the region at 150-min lead time.

The last example corresponds to the second part of E7, which took place on February 28, 2018. This was associated with the passage of a warm front, which is an

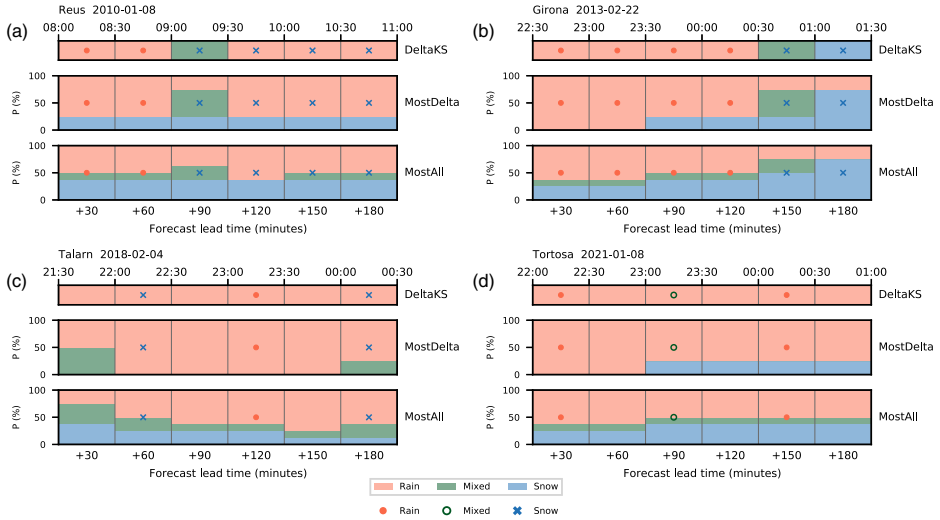

FIGURE 5 Comparison between Delta KS, MostDelta, and MostAll precipitation-phase nowcasting schemes over the forecast lead time for different locations and events. The location and nowcasting initialisation time are shown above each panel. Shaded colours indicate the probability of each precipitation phase: snow, mixed, and rain. Markers represent the actual precipitation phase observed [Colour figure can be viewed at wileyonlinelibrary.com]

unusual situation for a snowfall in the region (Bullón and Fernández, 2019). At 1000 UTC, Catalonia was mostly under snowfall conditions, with the precipitation phase being highly influenced by the warm front position. Thus, the snow level in the northern part of the region was 0 m a.s.l at 1000 UTC, whereas it was ~300 m a.s.l in the southern part (Figure 7a). The warm front moved towards the north and, as indicated by precipitation phase observations, the snow level rose accordingly (Figure 7a–f). The nowcasting scheme failed to forecast the transition from snow to rain correctly in Sabadell (Figure 7g), since this was forecast 60 min earlier (1030 UTC) than when it actually occurred (1130 UTC). Nevertheless, the forecast probability of rain was not 100% and some algorithms in the MostAll combination forecast snow (Figure 7g). However, in Girona, the nowcasting scheme forecast the transition correctly with a lead time of 120 min (Figure 7h).

5 | DISCUSSION

Regardless of the algorithm used, the performance of precipitation phase classification algorithms generally presented a slight decrease from 30 to 180 min forecast lead time (Figure 2). This result is consistent with Wandishin *et al.* (2005) and Elmore *et al.* (2015), who found that precipitation phase algorithms are relatively insensitive to model errors over the forecast lead time. The mean GSS values presented in Table 3 for 180-min lead time and all events suggest that single extrapolated surface observation threshold-based schemes (Delta T_a, Delta T_d, Delta T_w, and Delta KS) perform better than vertical temperature profile-based algorithms (Baldwin, Bourgouin, Ramer and Schuur).

Above all, considering GSS and PSS values for snow and rain observations (Figure 2 and Table 3), the best-performing algorithms were Delta T_w and Delta KS. Both obtained a mean GSS value of 0.62, but Delta KS presented statistically significantly better performance, albeit the differences were marginal. Nevertheless, when the algorithms were analysed for individual events, some variability was observed (Figure 3, Tables 4 and 5). This result is in accordance with Cortinas *et al.* (2002), Manikin (2005), and Wandishin *et al.* (2005), who indicated that no algorithm performed better than all others at all times and therefore suggested using a combination of algorithms.

In this study, single NWP model-based algorithms were outperformed by their combination, Profiles, in terms of average performance (Figure 2 and Table 3). On the other hand, MostAll and MostDelta did not surpass the best-performing single algorithms, Delta KS or Delta T_w,

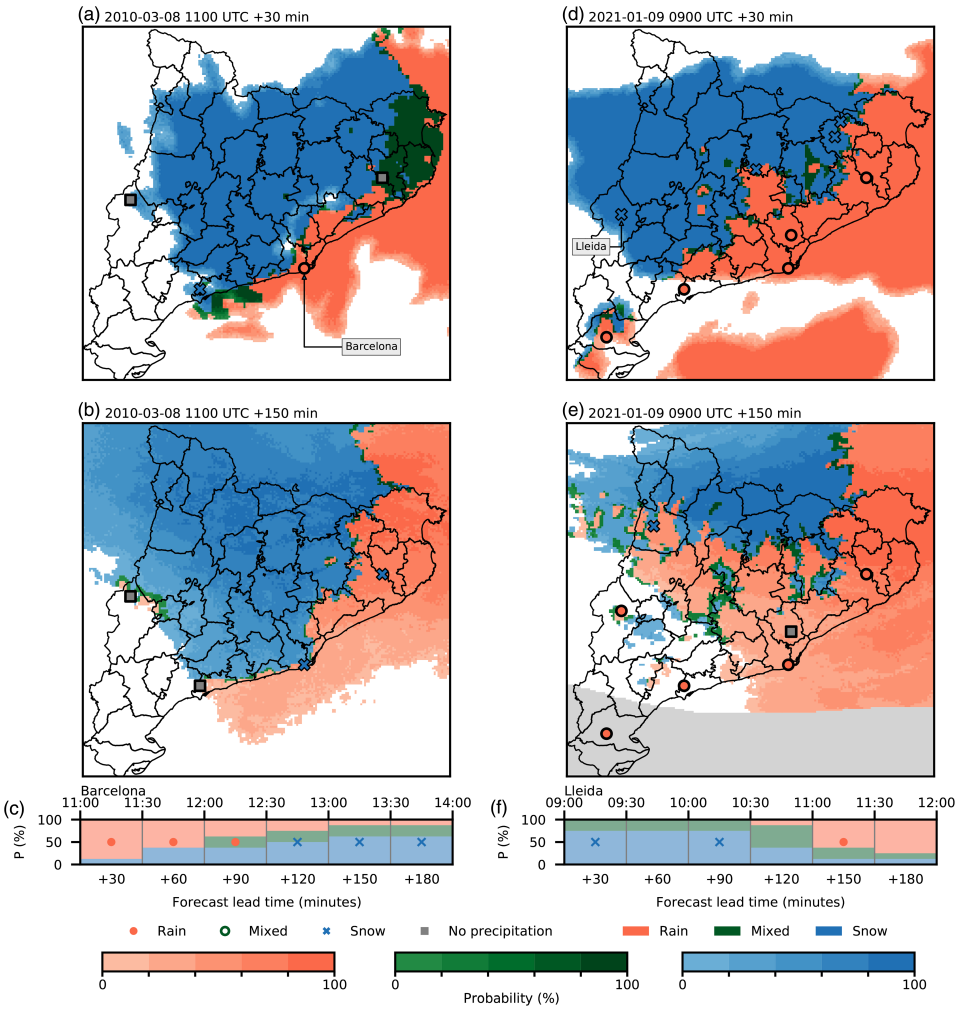

FIGURE 6 Most probable precipitation phase (MostAll) nowcasting masked with probability of precipitation exceeding 0.1 mm·hr^{-1} at (a, d) 30 and (b, e) 150 min forecast lead time. The lowest area in (e) indicates where the extrapolated precipitation field is not available. (c, f) Nowcasting for a specific location. The left column corresponds to a nowcast initialised on March 8, 2010, at 1100 UTC, and the right column to January 9, 2021, at 0900 UTC. Precipitation phases shown are rain, mixed, and snow. Shaded colours in (c, f) indicate the probability of each precipitation phase and markers represent the actual precipitation phase observed in Barcelona (c) and in Lleida (f) [Colour figure can be viewed at wileyonlinelibrary.com]

presenting statistically significant differences but a similar mean GSS (Table 3). Therefore, a reasonable conclusion would be to select either Delta KS or Delta T_w. However, the performance variability observed in different events cannot be captured using a single algorithm (Tables 4 and 5). MostAll was the best scheme in three out of eight events, and, except for E2, it presented similar performance to the best-performing algorithm in each

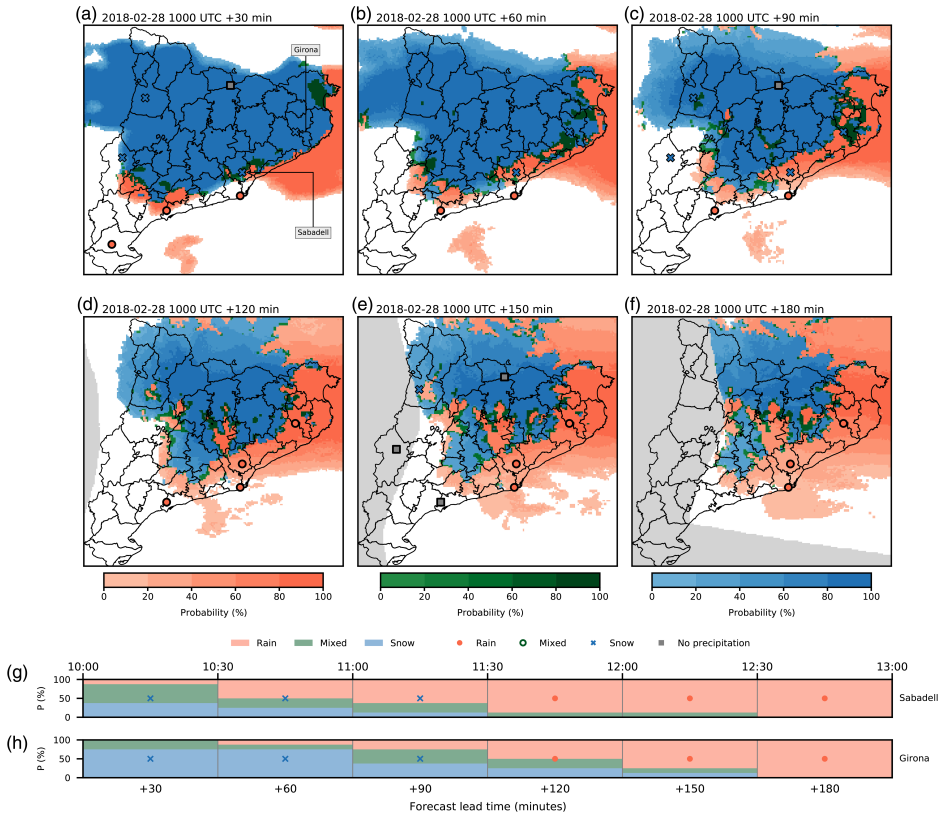

FIGURE 7 Most probable precipitation phase (MostAll) nowcasting masked with the probability of precipitation exceeding 0.1 mm·hr^{-1} on February 28, 2018, at 1000 UTC, for (a–f) 30–180 min lead time. Precipitation phases shown are rain, mixed, and snow. Nowcasting for (g) Sabadell and (h) Girona: shaded colours indicate the probability of each precipitation phase and markers represent the actual precipitation phase observed [Colour figure can be viewed at wileyonlinelibrary.com]

event. Furthermore, in some situations, a combination of algorithms can provide more information than just one. Figures 4 and 5 show that MostDelta and MostAll gave the same most probable precipitation phase estimates as Delta KS, but, at the same time, probability information can assist operational forecasters in situations of high uncertainty or difficulty in determining the precipitation phase. As an example, it is worth highlighting the case of February 27, 2018, in Barcelona (Figure 4), where air saturation conditions played a key role in determining the precipitation phase at surface level. During the snowfall, considerably low relative humidity values, together with calm winds, produced solid precipitation (snowflakes) near the ground through evaporative cooling, an effect described in previous studies (Kain *et al.*, 2000). The snow to rain transition started with humid air advection from the Mediterranean Sea, increasing relative humidity values and consequently enhancing snowflake melting despite the decrease in air temperature (Matsuo and Sasyo, 1981). The precise nature of this transition was not well captured by Delta KS, which estimated rain, or by MostDelta or MostAll. However, the latter two schemes showed some non-negligible snow and mixed probabilities, illustrating qualitatively the discrepancy between algorithms, which may provide guidance about the overall uncertainty of the event.

Considering the examples presented above and from a statistical point of view, the differences in performance between the two best single extrapolated surface observation schemes and the MostDelta and MostAll combinations were narrow, albeit statistically significant. However, using a combination of algorithms provides three advantages compared with using only one. First, the amount of information provided to an operational forecaster surpasses that of a single scheme. This can be critical in cases close to freezing point, where a bias or an error in the temperature forecast can affect precipitation phase algorithms (Bailey *et al.*, 2014) substantially. Therefore, considering more than one algorithm can mitigate this potential error. Second, if MostAll is chosen, the inclusion of NWP model-based discrimination algorithms makes it possible to consider vertical temperature profile information, which plays a key role in determining the precipitation phase at the surface level (Bourgouin, 2000). Third, agreement or disagreement between schemes can be interpreted as forecast uncertainty (Cortinas *et al.*, 2002; Wandishin *et al.*, 2005) and can provide key information if the output of the precipitation-phase nowcasting scheme is applied to subsequent forecasting systems such as hydrological models (Rossa *et al.*, 2010).

6 | SUMMARY AND CONCLUSIONS

This study evaluated the application of a precipitation-phase nowcasting scheme in the Western Mediterranean region. The scheme was constructed by merging precipitation phase discrimination algorithms and extrapolated weather radar precipitation fields. The main findings were as follows.

- Two algorithms based on extrapolated surface information only, Delta T_w and Delta KS, were the best precipitation phase discrimination schemes in terms of average performance.
- Precipitation phase algorithm performance exhibited some event-to-event variability, suggesting that an equally weighted ensemble of algorithms is advisable.
- The two proposed combinations of precipitation phase discrimination schemes, MostDelta and MostAll, outperformed the combination of NWP model vertical temperature profiles (Profiles).
- Application of the nowcasting scheme to the study cases reported here revealed notable performance in terms of determining the precipitation phase and the timing of precipitation phase transitions.
- The MostAll combination included extrapolated surface observations and vertical temperature profiles. Agreement and disagreement among the eight schemes included in MostAll can assist operational forecasters by providing a wider perspective, and may be helpful in situations where the temperature is close to freezing point and the difficulty in determining the precipitation phase is greater.

As the observations included only three different precipitation types (in this case, precipitation phases), rain, mixed, and snow, the vertical temperature profile algorithms were narrowed to these three categories. Future studies should consider more events with other precipitation types, such as freezing rain or ice pellets. In addition, with a larger precipitation type observation database, relationships between the performance of different algorithms regarding event conditions could be analysed in greater detail. Nevertheless, the methodology described here can be applied readily to other regions where ground-based observations, weather radar data, and model forecasts are available.

AUTHOR CONTRIBUTIONS

Enric Casellas: conceptualization; dataCuration; investigation; methodology; Software; validation; visualization; writingOriginalDraft; writingReviewEditing. **Joan Bech:** conceptualization; fundingAcquisition; Projectadministration; resources; supervision; writingOriginalDraft; writingReviewEditing. **Roger Veciana:** conceptualization; fundingAcquisition; Projectadministration; Software. **Nicolau Pineda:** fundingAcquisition; Projectadministration; resources; supervision. **Josep Ramon Miró:** investigation; methodology; supervision; validation; writingReviewEditing. **Jordi Moré:** dataCuration; resources; writingReviewEditing. **Tomeu Rigo:** dataCuration; resources; writingReviewEditing. **Abdel Sairouni:** fundingAcquisition; Projectadministration; resources; supervision.

ACKNOWLEDGEMENTS

This study was performed under the framework of the HyMeX program, and supported by the Industrial Doctorate Program partly funded by the Government of Catalonia (Project DI-053/2017), and the Spanish Projects CGL2015-65627-C3-2-R (MINECO/FEDER), CGL2016-81828-REDT/AEI, and RTI2018-098693-B-C32 (MINECO). The authors declare no conflict of interest. The authors acknowledge Jordi Mercader (Meteorolgical Service of Catalonia) for providing WRF model data. We also acknowledge three anonymous reviewers for their constructive suggestions that improved the final form of this article.

ORCID

Enric Casellas https://orcid.org/0000-0001-6894-0283
Joan Bech https://orcid.org/0000-0003-3597-7439
Nicolau Pineda https://orcid.org/0000-0002-2507-8424
Josep Ramon Miró https://orcid.org/0000-0003-2838-6083
Tomeu Rigo https://orcid.org/0000-0002-4520-4176

REFERENCES

Altube, P., Bech, J., Argemí, O. and Rigo, T. (2015) Quality control of antenna alignment and receiver calibration using the sun: adaptation to midrange weather radar observations at low elevation angles. *Journal of Atmospheric and Oceanic Technology*, 32, 927–942.

Altube, P., Bech, J., Argemí, O., Rigo, T., Pineda, N., Collis, S. and Helmus, J. (2017) Correction of dual-PRF Doppler velocity outliers in the presence of aliasing. *Journal of Atmospheric and Oceanic Technology*, 34, 1529–1543.

Andrey, J., Mills, B., Leahy, M. and Suggett, J. (2003) Weather as a chronic hazard for road transportation in Canadian cities. *Natural Hazards*, 28, 319–343.

Bailey, M.E., Isaac, G.A., Gultepe, I., Heckman, I. and Reid, J. (2014) Adaptive blending of model and observations for automated short-range forecasting: examples from the Vancouver 2010 olympic and paralympic winter games. *Pure and Applied Geophysics*, 171, 257–276.

Baldwin, M., Treadon, R. and Contorno, S. (1994). Precipitation type prediction using a decision tree approach with NMCs mesoscale ETA model, Portland, OR. In: Preprints, 10th Conference on Numerical Weather Prediction.

Bech, J., Codina, B., Lorente, J. and Bebbington, D. (2003) The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. *Journal of Atmospheric and Oceanic Technology*, 20, 845–855.

Bech, J., Pineda, N., Rigo, T. and Aran, M. (2013) Remote sensing analysis of a Mediterranean thundersnow and low-altitude heavy snowfall event. *Atmospheric Research*, 123, 305–322.

Bech, J., Rigo, T., Pineda, N., Segalà, S., Vilaclara, E., Sánchez-Diezma, R., Sempere, D. and Velasco, E. (2005). Implementation of the EHIMI software package in the weather radar operational chain of the Catalan Meteorological Service. In: Proceedings 32nd International Conference on Radar Meteorology. Alburquerque, NM. American Meteorological Society.

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D. and Ye, H. (2018) On distinguishing snowfall from rainfall using near-surface atmospheric information: comparative analysis, uncertainties and hydrologic importance. *Quarterly Journal of the Royal Meteorological Society*, 144, 89–102.

Benjamin, S.G., Brown, J.M. and Smirnova, T.G. (2016) Explicit precipitation-type diagnosis from a model using a mixed-phase bulk cloud–precipitation microphysics parameterization. *Weather and Forecasting*, 31, 609–619.

Berenguer, M., Surcel, M., Zawadzki, I., Xue, M. and Kong, F. (2012) The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: intercomparison among numerical models and with nowcasting. *Monthly Weather Review*, 140, 2689–2705.

Bonelli, P., Lacavalla, M., Marcacci, P., Mariani, G. and Stella, G. (2011) Wet snow hazard for power lines: a forecast and alert system applied in Italy. *Natural Hazards and Earth System Sciences*, 11, 2419–2431.

Bourgouin, P. (2000) A method to determine precipitation types. *Weather and Forecasting*, 15, 583–592.

Bowler, N.E., Pierce, C.E. and Seed, A.W. (2006) STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, 132, 2127–2155.

Bullón, J.C. and Fernández, J. (2019). Inusual episodio de nevadas de frente cálido en la costa cantábrica y catalana. *Sexto Simposio Nacional de Predicción, Memorial Antonio Mestre*, Madrid: Agencia Estatal de Meteorología, pp. 487–496. [In Spanish].

Casellas, E., Bech, J., Veciana R., Miró, J.R., Sairouni, A., Pineda, N. (2020) A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, 246, 105103. http://dx.doi.org/10.1016/j.atmosres.2020.105103.

Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Miró, J.R. and Sairouni, A. (2021) Surface precipitation phase discrimination in complex terrain. *Journal of Hydrology*, 592, 125780.

Chen, F. and Dudhia, J. (2001) Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: model implementation and sensitivity. *Monthly Weather Review*, 129, 569–585.

Cortinas, J., Brill, K. and Baldwin, M. (2002). Probabilistic forecasts of precipitation type, Orlando, FL. In: Preprints, 16th Conference on Probability and Statistics in the Atmospheric Sciences p. 3.

Dudhia, J. (1989) Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of Atmospheric Sciences*, 46, 3077–3107.

Ebert, E.E. (2008) Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications*, 15, 51–64.

Efron, B. and Tibshirani, R.J. (1994) *An Introduction to the Bootstrap*. London: CRC Press.

Elmore, K.L., Grams, H.M., Apps, D. and Reeves, H.D. (2015) Verifying forecast precipitation type with mPING. *Weather and Forecasting*, 30, 656–667.

Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M. and Stoffel, M. (2019) Improving medium-range forecasts of rain-on-snow events in prealpine areas. *Water Resources Research*, 55, 7638–7661.

Franco, M., Sánchez-Diezma, R. and Sempere-Torres, D. (2006) Improvements in weather radar rain rate estimates using a method for identifying the vertical profile of reflectivity from volume radar scans. *Meteorologische Zeitschrift*, 15, 521–536.

Froidurot, S., Zin, I., Hingray, B. and Gautheron, A. (2014) Sensitivity of precipitation phase over the Swiss Alps to different meteorological variables. *Journal of Hydrometeorology*, 15, 685–696.

Gascón, E., Hewson, T. and Haiden, T. (2018) Improving predictions of precipitation type at the surface: description and verification of two new products from the ECMWF ensemble. *Weather and Forecasting*, 33, 89–108.

Gerrity, J.P. (1992) A note on Gandin and Murphy's equitable skill score. *Monthly Weather Review*, 120, 2709–2712.

Golding, B. (1998) Nimrod: a system for generating automated very short range forecasts. *Meteorological Applications*, 5, 1–16.

Haiden, T., Kann, A., Wittmann, C., Pistotnik, G., Bica, B. and Gruber, C. (2011) The Integrated Nowcasting through

Comprehensive Analysis (INCA) system and its validation over the Eastern Alpine region. *Weather and Forecasting*, 26, 166–183.

Hong, S.-Y., Dudhia, J. and Chen, S.-H. (2004) A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly Weather Review*, 132, 103–120.

Hong, S.-Y., Noh, Y. and Dudhia, J. (2006) A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, 134, 2318–2341.

Ikeda, K., Steiner, M., Pinto, J. and Alexander, C. (2013) Evaluation of cold-season precipitation forecasts generated by the hourly updating High-Resolution Rapid Refresh model. *Weather and Forecasting*, 28, 921–939.

Ikeda, K., Steiner, M. and Thompson, G. (2017) Examination of mixed-phase precipitation forecasts from the High-Resolution Rapid Refresh model using surface observations and sounding data. *Weather and Forecasting*, 32, 949–967.

Isaac, G.A., Bailey, M., Boudala, F.S., Burrows, W.R., Cober, S.G., Crawford, R.W., Donaldson, N., Gultepe, I., Hansen, B., Heckman, I., Huang, L.X., Ling, A., Mailhot, J., Milbrandt J.A., Reid, J., Fournier, M. (2014) The Canadian Airport Nowcasting System (CAN-Now). *Meteorological Applications*, 21(1), 30–49. http://dx.doi.org/10.1002/met.1342.

Jolliffe, I.T. (2007) Uncertainty and inference for verification measures. *Weather and Forecasting*, 22, 637–650.

Jolliffe, I.T. and Stephenson, D.B. (2012) *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. Oxford, UK: John Wiley & Sons.

Kain, J.S. (2004) The Kain–Fritsch convective parameterization: an update. *Journal of Applied Meteorology*, 43, 170–181.

Kain, J.S., Goss, S.M. and Baldwin, M.E. (2000) The melting effect as a factor in precipitation-type forecasting. *Weather and Forecasting*, 15, 700–714.

Kann, A., Kršmanc, R., Habrovskú, R., Šajn, A.Š., Bujňák, R., Schmid, F., Tarjáni, V., Wang Y., Wastl, C., Bica, B., Meirold-Mautner, I. (2015) High-resolution nowcasting and its application in road maintenance: experiences from the INCA Central European area project. *IET Intelligent Transport Systems*, 9(5), 539–546. http://dx.doi.org/10.1049/iet-its.2014.0102.

Keis, F. (2015) WHITE–Winter hazards in terminal environment: an automated nowcasting system for Munich Airport. *Meteorologische Zeitschrift*, 24, 61–82.

Koistinen, J. and Saltikoff, E. (1998) Experience of customer products of accumulated snow, sleet and rain. *COST75 Advanced Weather Radar Systems*, 18567, 397–406.

Lawrence, M.G. (2005) The relationship between relative humidity and the dewpoint temperature in moist air: a simple conversion and applications. *Bulletin of the American Meteorological Society*, 86, 225–234.

Liu, G. (2008) Deriving snow cloud characteristics from CloudSat observations. *Journal of Geophysical Research: Atmospheres*, 113, D00A09.

Llasat, M., Turco, M., Quintana-Seguí, P. and Llasat-Botija, M. (2014) The snow storm of 8 March 2010 in Catalonia (Spain): a paradigmatic wet-snow event with a high societal impact. *Natural Hazards and Earth System Sciences*, 14, 427.

Malin, F., Norros, I. and Innamaa, S. (2019) Accident risk of road and weather conditions on different road types. *Accident Analysis & Prevention*, 122, 181–188.

Manikin, S.G. (2005). An overview of precipitation type forecastingusing NAM and SREF data. Preprints,21st Conf. on Weather-Analysis and Forecasting/17th Conf. on Numerical WeatherPrediction, Washington, DC, Amer. Meteor. Soc., 8A.6.

Marks, D., Winstral, A., Reba, M., Pomeroy, J. and Kumar, M. (2013) An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Advances in Water Resources*, 55, 98–110.

Matsuo, T. and Sasyo, Y. (1981) Melting of snowflakes below freezing level in the atmosphere. *Journal of the Meteorological Society of Japan. Series II*, 59, 10–25.

McCabe, A., Swinbank, R., Tennant, W. and Lock, A. (2016) Representing model uncertainty in the Met Office convection-permitting ensemble prediction system and its impact on fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, 142, 2897–2910.

McGovern, A., Elmore, K.L., Gagne, D.J., Haupt, S.E., Karstens, C.D., Lagerquist, R., Smith, T. and Williams, J.K. (2017) Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, 98, 2073–2090.

McGovern, A., Lagerquist, R., John Gagne, D., Jergensen, G.E., Elmore, K.L., Homeyer, C.R. and Smith, T. (2019) Making the black box more transparent: understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, 100, 2175–2199.

Mlawer, E., Taubman, S., Brown, P., Iacono, M. and Clough, S. (1997) Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long wave. *Journal of Geophysical Research*, 102, 16.

Moon, S.-H. and Kim, Y.-H. (2020) An improved forecast of precipitation type using correlation-based feature selection and multinomial logistic regression. *Atmospheric Research*, 240, 104928.

Papagiannaki, K., Lagouvardos, K. and Kotroni, V. (2013) A database of high-impact weather events in Greece: a descriptive impact analysis for the period 2001–2011. *Natural Hazards and Earth System Sciences*, 13, 727–736.

Pomeroy, J., Gray, D., Brown, T., Hedstrom, N., Quinton, W., Granger, R. and Carey, S. (2007) The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes: An International Journal*, 21, 2650–2667.

Pulkkinen, S., Nerini, D., Pérez Hortal, A.A., Velasco-Forero, C., Seed, A., Germann, U. and Foresti, L. (2019) Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1. 0). *Geoscientific Model Development*, 12, 4185–4219.

Ramer, J. (1993). An empirical technique for diagnosing precipitation type from model output, Vienna, VA. In: 5th International Conference on Aviation Weather Systems. American Meteorological Society, pp. 227–230.

Ripoll, R., del Amo, X. and Vendrell, R. (2016). The weather observers network of the Meteorological Service of Catalonia, WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (CIMO TECO 2016), 27–30 September 2016, Madrid, Spain, P2(57), 2016.

Rossa, A., Haase, G., Keil, C., Alberoni, P., Ballard, S., Bech, J., Germann, U., Pfeifer, M. and Salonen, K. (2010) Propagation of

uncertainty from observing systems into NWP: COST-731 working group 1. *Atmospheric Science Letters*, 11, 145–152.

Rouf, T., Mei, Y., Maggioni, V., Houser, P. and Noonan, M. (2020) A physically based atmospheric variables downscaling technique. *Journal of Hydrometeorology*, 21, 93–108.

Sadeghi, S.-H., Peters, T.R., Cobos, D.R., Loescher, H.W. and Campbell, C.S. (2013) Direct calculation of thermodynamic wet-bulb temperature as a function of pressure and elevation. *Journal of Atmospheric and Oceanic Technology*, 30, 1757–1765.

Saltikoff, E., Hagen, M., Juntti, H., Kaltenböck, R. and Pulkkinen, S. (2018) Nowcasting snow for airports at heterogeneous terrain. *Geophysica*, 53, 29–41.

Sánchez-Diezma, R., Sempere-Torres, D., Bech, J. and Velasco, E. (2002). Development of a hydrometeorological flood warning system (EHIMI) based on radar data. 2nd European Radar Conference. Copernicus Gesellschat, Delft, Holland.

Schmidlin, T.W. (1993) Impacts of severe winter weather during December 1989 in the Lake Erie snowbelt. *Journal of Climate*, 6, 759–767.

Schuur, T.J., Park, H.-S., Ryzhkov, A.V. and Reeves, H.D. (2012) Classification of precipitation types during transitional winter weather using the RUC model and polarimetric radar retrievals. *Journal of Applied Meteorology and Climatology*, 51, 763–779.

Seed, A. (2003) A dynamic and spatial scaling approach to advection forecasting. *Journal of Applied Meteorology*, 42, 381–388.

Serra, A., Mercè, B. and Vendrell, R. (2016). Automatic weather stations network (XEMA) of the Meteorological Service of Catalonia (SMC). https://library.wmo.int/index.php?lvl=notice_display&id=19676#.XEWdHvZ7m70.

Sheridan, P., Smith, S., Brown, A. and Vosper, S. (2010) A simple height-based correction for temperature downscaling in complex terrain. *Meteorological Applications*, 17, 329–339.

Simonin, D., Pierce, C., Roberts, N., Ballard, S.P. and Li, Z. (2017) Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, 143, 2862–2873.

Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Duda, M.G., Huang, X.-Y., Wang, W. and Powers, J.G. (2008). A description of the advanced research WRF version 3, Rep., 113 pp., National Center for Atmospheric Research, Boulder, Colorado, USA.

Smith, A., Lott, N. and Vose, R. (2011) The integrated surface database: recent developments and partnerships. *Bulletin of the American Meteorological Society*, 92, 704–708.

Trapero, L., Bech, J., Rigo, T., Pineda, N. and Forcadell, D. (2009) Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. *Atmospheric Research*, 93, 408–418.

Vilaclara, E., Segalà, S., Andrés, A. and Aran, M. (2010). Operational warnings issued by the SMC in the 8th March snow event in Catalonia. http://meetings.copernicus.org/plinius12.

Wandishin, M.S., Baldwin, M.E., Mullen, S.L. and Cortinas, J.V. (2005) Short-range ensemble forecasts of precipitation type. *Weather and Forecasting*, 20, 609–626.

Wilks, D.S. (2011) *Statistical Methods in the Atmospheric Sciences* Vol. 100. Amsterdam, Boston: Academic Press.

Zhong, K., Zheng, F., Xu, X., Qin, C. (2018) Discriminating the precipitation phase based on different temperature thresholds in the Songhua River Basin, China. *Atmospheric Research*, 205, 48–59. http://dx.doi.org/10.1016/j.atmosres.2018.02.002.

How to cite this article: Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J.R., Moré, J. *et al.* (2021) Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(739), 3135–3153. Available from: https://doi.org/10.1002/qj.4121

Chapter 5

Conclusions

In this chapter the conclusions resulting from the research done during the predoctoral period are presented. The general objectives, previously outlined in Chapter 1, Section 1.3, upon the present thesis was structured are the following:

- **GO1. Adjustment and implementation of a diagnostic surface precipitation phase product.**
- **GO2. Development and implementation of a nowcasting surface precipitation phase product.**

The present thesis also contains six specific objectives which allowed the accomplishment of the two general objectives. Therefore, this section summarizes the actions and results that give answers to each of the defined objectives. Furthermore, future perspectives and working lines on how to improve the diagnostic and nowcasting of surface precipitation phase are provided.

5.1 Final conclusions

SO1. Obtention of a dynamic interpolation scheme suitable for complex terrain, and high spatial and temporal resolution. The proposed interpolation scheme, named Meteorological field Interpolation based on Clustered data Analysis (MICA), includes clustering weather stations data, multiple linear regressions and anomaly corrections. The combination of multiple linear regression and anomaly correction is a common spatial interpolation strategy, but the clustering of weather stations data confers an adaptive

character to the interpolation scheme. How? A linear regression model is calculated for each cluster of weather stations and compared against a model calculated with all weather stations (global). Only the clusters whose regression model yields an error reduction compared to that of the global model are selected. It is the testing of several numbers of clusters every time an interpolation is conducted where lays the adaptive condition of the scheme.

The results of applying MICA scheme for air temperature and dew point temperature observations to three European regions (Catalonia, Baden-Württemberg and Emilia-Romagna) yielded a reduction of cross-validation errors compared to using an MLR model followed by an anomaly correction without clusters. This was mainly due to the linearity increase when stations are grouped in smaller groups rather than considering them all in a MLR model. In addition, a clear reduction in middle and high altitude station air temperature and dew point temperature errors was obtained benefiting from the increase of those stations weight on MLR models.

Apart from the general results, case studies exhibited the usefulness of MICA during specific meteorological conditions. For example, interpolated fields with more homogeneous representation of areas under fog conditions, where an informed cluster definition was useful to isolate a part of a region prone to fog. In addition, grouping stations in different clusters allowed to reduce interpolation errors when observations did not follow the expected trends, such as strong temperature inversions.

SO2. Evaluation and adjustment of different schemes and meteorological variables to diagnose discrimination of the surface precipitation phase. Seven threshold-based schemes were evaluated regarding the discrimination of the surface precipitation phase. Three of them were based on a single threshold and four on dual thresholds. Among the single threshold discrimination schemes, wet bulb temperature exhibited the best validation scores for snow and rain observations with Pierce skill score (PSS) values of 0.77, followed by dew point temperature and air temperature. For dual threshold schemes, which include three types of precipitation observations (rain, sleet and snow), a similar trend was found. Wet bulb temperature (DT-T_w) and Koistinen-Saltikoff (DT-KS) schemes performed best and air temperature (DT-Ta) worst. These results indicate the suitability of meteorological variables including information about air saturation conditions compared to using air temperature alone for precipitation phase discrimination.

One of the most common schemes in hydrological models regarding precipitation phase classification is the single threshold based on air temperature. For Catalonia, after the validation and adjustment with more than 7000 quality-controlled precipitation

observations, this threshold was set to 1.3°C. The single threshold that performed best, based on wet bulb temperature, was set to 0.7°C. In case of dual threshold schemes, a wet bulb temperature of 0.7°C was found to better distinguish between snow and sleet and 1.0°C between sleet and rain.

Although the single and double threshold schemes to discriminate the precipitation phase were already tested in other regions, the added value of this study is that it only used interpolated surface information rather than actual observations. The results indicate that the regionalisation of relative humidity observations has a positive impact on the precipitation phase discrimination schemes, which may be translated into an hydrological modelling improvement.

SO3. Assessment of citizen science and crowd sourced observations for monitoring snow events. Precipitation phase observations reported by the SMC spotter network (XOM) exhibited a notable probability of detection (POD) score (0.90) when compared to precipitation phase estimates. Moreover, precipitation phase observations gathered from Twitter social network showed a similar POD (0.84). Both scores prove that valuable information can be obtained from citizen science and crowd-sourcing to monitor low-altitude snowfall events.

The exceptionality of low-altitude snowfall events in Catalonia increases the number of observations of snow when these events occur. Thus, the present study took advantage of this exceptionality to test around 1200 snow observations collected from Twitter during February and March 2018. This amount of observations, compared to the rest of precipitation phase sources, opens up the possibility to operatively consider Twitter or social networks as a reliable source of information for monitoring snowfall events. However, a higher degree of filtering and quality control checks should be considered with these observations due to their inherent issues.

After all and despite the implicit location and temporal uncertainty of these kind of sources, the results highlight the importance and value of these non-conventional real-time observations in weather surveillance.

SO4. Development and evaluation of different schemes to nowcast discrimination of the precipitation phase. The precipitation phase discrimination schemes based on extrapolated surface observations exhibited a superior performance compared to the vertical temperature profile-based algorithms. Schemes based on the extrapolation of wet bulb temperature (Delta T_w) and Koistinen-Saltikoff (KS, Koistinen & Saltikoff (1998)) presented best results among the selected algorithms when all events are considered

with GSS values of 0.62. Regarding the vertical temperature profile-based algorithms, Schuur (Schuur et al., 2012), was the best (GSS=0.53).

The best performing algorithms, either those based on extrapolated surface observations or vertical temperature profiles, included humidity information. Delta T_a, including only air temperature observations, and Bourgouin, based only on the air temperature profile, obtained the least good performance, highlighting the importance of the inclusion of air saturation information in precipitation phase discrimination schemes.

However, some performance variability among precipitation phase discrimination schemes was observed when the algorithms were analysed for individual events. That is, no algorithm performed better than all others at all times. This conclusion is in accordance with previous studies in other regions, which suggested to combine the algorithms as an equally weighted ensemble.

SO5. Evaluation of ensemble techniques to nowcast discrimination of the precipitation phase. Combination of precipitation phase discrimination schemes was already tested in other studies and regions, but including only vertical temperature profile-based algorithms (Profiles). In the present study two new combinations were evaluated: one including only extrapolated surface observations schemes (MostDelta) and one including all the schemes (MostAll). Profiles outperformed single NWP model-based algorithms in terms of average performance among all events. However, MostDelta and MostAll exhibited a similar behaviour to Delta KS or Delta T_w (based on extrapolated observations only), but considering a combination of algorithms resulted in three advantages. First, the amount of information provided surpasses that of a single scheme, which can be critical in cases close to freezing point and a small bias in the temperature forecast can affect precipitation phase algorithms. Second, forecast uncertainty can be interpreted from agreement or disagreement between schemes. And third, if MostAll is considered, it allows to include extrapolated observations together with vertical temperature profile information, which plays a key role in precipitation phase at surface level.

Apart from correctly determining the precipitation phase, it is also important to estimate the timing of possible phase transitions. It is true that this can be achieved with a single scheme, but in the present study, results show that when uncertainty is high or rapid precipitation phase transitions occur, an ensemble approach is advisable. When the precise nature of a transition cannot be correctly determined and schemes estimate different precipitation phase, an ensemble technique qualitatively illustrates the discrepancy among algorithms and, at the same time, may provide guidance about

the overall uncertainty of the transition.

SO6. Implementation of a precipitation phase product in an operational chain. The final objective of this thesis is to implement a diagnosis and nowcasting operational chain to estimate the precipitation phase in the Meteorological Service of Catalonia.

The implemented precipitation phase diagnosing product is based on a dual threshold scheme based on wet bulb temperature. Threshold values correspond to those obtained in SO2 and are applied to a wet bulb temperature interpolated field obtained with the resultant interpolation scheme from SO1. In addition, a second set of thresholds was selected to account for the threshold variability observed during the validation process. Thereby, the operational forecasters can adapt the precipitation phase product to different weather situations. This can be useful especially for low altitude snowfall events where a small change in threshold temperature greatly drives the potential area affected by snow conditions. The estimation of the precipitation phase is combined with weather radar composite fields providing not only the precipitation phase but also the precipitation occurrence.

Regarding the nowcasting operational chain, an ensemble of precipitation phase discrimination schemes was selected according to the results obtained from SO4 and SO5. Eight different schemes were combined to obtain not only the most probable precipitation phase but also the probability of each phase. Precipitation occurrence information was derived from weather radar reflectivity composite fields extrapolated with a stochastic nowcasting technique. The combination of both sources results in a three hour lead time forecast including an estimation of the most probable precipitation phase, the probability of each phase and the occurrence of precipitation with an update every thirty minutes.

The diagnose product was made publicly available on the website of the Meteorological Service of Catalonia. On the other hand, an ad-hoc internal map viewer was designed for the nowcasting product with the aim to make it public in the future.

5.2 Future work

The present thesis addressed three issues which future work should focus on: interpolation of surface meteorological observations, and diagnosing and nowcasting of the precipitation phase.

In regard to interpolation of meteorological observations and for the time being, the MICA scheme developed in the present thesis is only available for air temperature and

dew point temperature variables and should be expanded to other key meteorological variables such as precipitation. In this sense, MICA should not only rely on linear regressions, but could also include more sophisticated methodologies such as kriging (Haberlandt, 2007; Moral, 2010). Another aspect that could be improved in MICA is the way clusters are selected. For now, the clusters are designed beforehand based on orographic features and separating regions prone to potential behaviour differences. A possible upgrade would be to dynamically cluster stations depending on the weather conditions or the circulation type using machine learning techniques.

One of the limitations of the present thesis is the low number of precipitation phase observations compared to other studies in other regions for a similar period of time. It is true that precipitation in form of snow in Catalonia is mainly limited to mountainous regions and barely affects low altitude areas. However, the occasional snowfall events below 700 m a.s.l can cause a high socio-economic impact (Bech et al., 2013; Llasat et al., 2014). For this reason an effort on collecting precipitation phase information should be considered in key locations. An option would be to deploy a network of disdrometers (Pickering et al., 2019) that not only automatically reports the precipitation phase but also other critical information, such as drop size distributions (DSD) to calibrate weather radars (Leinonen et al., 2012; Adirosi et al., 2018). In the present thesis, citizen-science and crowdsourced data were assessed as valid sources for precipitation phase information, highlighting their importance in snowfall events. Therefore, advantage should be taken on social networks, such as Twitter, with active campaigns and specific communications to motivate their users to report information regarding snow events following previous successful campaigns ("Snowtweets" by King et al. (2009) or "Picking up Hailstones" by Farnell & Rigo (2020)). However, engagement in these campaigns is sometimes difficult and other strategies may be followed. In the present thesis, a scraping of Twitter with the aid of its API, that may be labelled as a passive campaign, provided notable results as seen in SO3 and its automatisation and operational implementation could be considered as a future work. All these strategies share the aim of collecting a larger number of precipitation phase observation that would contribute to a better understanding of the precipitation phase behaviour in the region and to a better calibration of discrimination schemes.

Finally, regarding the diagnosing and nowcasting of the precipitation phase, new methodologies to discriminate the precipitation phase can be incorporated, including the emerging methodologies based on machine learning (McGovern et al., 2017, 2019). In fact, the design of the nowcasting of the precipitation phase product in this thesis was conceived as a modular product, that is, a product expected to be improved in the

future in three aspects: nowcasting of meteorological variables to discriminate precipitation phase, nowcasting of precipitation fields, and discrimination of precipitation phase methodologies. For example, the nowcasting of precipitation fields was implemented using the STEPS methodology implemented in pySTEPS. However, a recent study (Ravuri et al., 2021) presented a new methodology based on AI that could improve the current implemented methodology in the product. Therefore, future work should also be centered on testing new methodologies and judge whether they yield an improvement to the whole product.

Bibliography

Adirosi, E., Roberto, N., Montopoli, M., Gorgucci, E., & Baldini, L. (2018). Influence of disdrometer type on weather radar algorithms from measured DSD: Application to Italian climatology. *Atmosphere*, *9*(9), 360.

Altube, P., Bech, J., Argemí, O., & Rigo, T. (2015). Quality control of antenna alignment and receiver calibration using the sun: Adaptation to midrange weather radar observations at low elevation angles. *Journal of Atmospheric and Oceanic Technology*, *32*(5), 927–942.

Altube, P., Bech, J., Argemí, O., Rigo, T., Pineda, N., Collis, S., & Helmus, J. (2017). Correction of dual-PRF Doppler velocity outliers in the presence of aliasing. *Journal of Atmospheric and Oceanic Technology*, *34*(7), 1529–1543.

Amaro, J., Llasat, M., & Aran, M. (2010). The social impact of the snowfall of 8 March 2010 in Catalonia. In *12th Plinius Conference on Mediterranean Storms, Corfu Island, Greece*.

AMS (2022). Glossary of meteorology.
URL `http://glossary.ametsoc.org`

Andrey, J., Mills, B., Leahy, M., & Suggett, J. (2003). Weather as a chronic hazard for road transportation in Canadian cities. *Natural Hazards*, *28*(2-3), 319–343.

Antolini, G., Auteri, L., Pavan, V., Tomei, F., Tomozeiu, R., & Marletto, V. (2016). A daily high-resolution gridded climatic data set for Emilia-Romagna, Italy, during 1961–2010. *International Journal of Climatology*, *36*(4), 1970–1986.

Apodaka, J., Pons, M., Trapero, L., Margalef, A., Albalat, A., Gallego, N., López-Moreno, J. I., & Furdada, G. (2018). Evaluation of 30 years of nivo-meteorological and avalanche data in Andorra. In *International Snow Science Workshop, Innsbruck, Austria*.

Aran, M., Rigo, T., Bech, J., Brucet, C., & Vilaclara, E. (2010). Analysis of the hazardous low-altitude snowfall, 8th March 2010, in Catalonia. In *12th Plinius Conference on Mediterranean Storms, held September 1-4, 2010 in Corfu Island, Greece. http://meetings. copernicus. org/plinius12, id. 77*.

Azam, M. F., Wagnon, P., Vincent, C., Ramanathan, A., Kumar, N., Srivastava, S., Pottakkal, J., & Chevallier, P. (2019). Snow and ice melt contributions in a highly glacierized catchment of Chhota Shigri Glacier (India) over the last five decades. *Journal of Hydrology*, *574*, 760–773.

Bailey, M. E., Isaac, G. A., Gultepe, I., Heckman, I., & Reid, J. (2014). Adaptive blending of model and observations for automated short-range forecasting: Examples from the Vancouver 2010 Olympic and Paralympic Winter Games. *Pure and Applied Geophysics*, *171*(1-2), 257–276.

Baldwin, M., Treadon, R., & Contorno, S. (1994). Precipitation type prediction using a decision tree approach with NMCs mesoscale eta model. In *Preprints, 10th Conf. on Numerical Weather Prediction, Portland, OR, Amer. Meteor. Soc*, (pp. 30–31).

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

Bech, J., & Chau, J. L. (2012). *Doppler Radar Observations: Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*. BoD–Books on Demand.

Bech, J., Codina, B., Lorente, J., & Bebbington, D. (2003). The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. *Journal of Atmospheric and Oceanic Technology*, *20*(6), 845–855.

Bech, J., Pineda, N., Rigo, T., & Aran, M. (2013). Remote sensing analysis of a Mediterranean thundersnow and low-altitude heavy snowfall event. *Atmospheric Research*, *123*, 305–322.

Bech, J., Rigo, T., Pineda, N., Segalà, S., Vilaclara, E., Sánchez-Diezma, R., Sempere, D., & Velasco, E. (2005). Implementation of the EHIMI software package in the

weather radar operational chain of the catalan meteorological service. In *Proceedings 32 nd International Conference on Radar Meteorology, Alburquerque, NM, USA*.

Bech, J., Vidal, V., Ortiz, J., Pineda, N., & Veciana, R. (2014). Real-time estimation of surface precipitation type merging weather radar and automated station observations. *17th International Road Weather Conference, SIRWEC*, *12*, 8.

Behnke, J., Mitchell, A., & Ramapriyan, H. (2019). NASA's Earth Observing Data and Information System–Near-Term Challenges. *Data Science Journal*, *18*(1).

Behrangi, A., Yin, X., Rajagopal, S., Stampoulis, D., & Ye, H. (2018). On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance. *Quarterly Journal of the Royal Meteorological Society*, *144*, 89–102.

Benjamin, S. G., Brown, J. M., & Smirnova, T. G. (2016). Explicit precipitation-type diagnosis from a model using a mixed-phase bulk cloud–precipitation microphysics parameterization. *Weather and Forecasting*, *31*(2), 609–619.

Berenguer, M., Surcel, M., Zawadzki, I., Xue, M., & Kong, F. (2012). The diurnal cycle of precipitation from continental radar mosaics and numerical weather prediction models. Part II: Intercomparison among numerical models and with nowcasting. *Monthly Weather Review*, *140*(8), 2689–2705.

Berndt, C., & Haberlandt, U. (2018). Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. *Journal of Hydrology: Regional Studies*, *15*, 184–202.

Bloemink, H. I., & Lanzinger, E. (2005). Precipitation type from the Thies disdrometer. In *WMO Technical Conference on Instruments and Methods of Observation (TECO-2005), Bucharest, Romania*, (pp. 4–7).

Bonelli, P., Lacavalla, M., Marcacci, P., Mariani, G., & Stella, G. (2011). Wet snow hazard for power lines: a forecast and alert system applied in Italy. *Natural Hazards and Earth System Sciences*, *11*(9), 2419–2431.

Bourgouin, P. (2000). A method to determine precipitation types. *Weather and Forecasting*, *15*(5), 583–592.

Bowler, N. E., Pierce, C. E., & Seed, A. W. (2006). STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quarterly Journal of the Royal Meteorological Society*, *132*(620), 2127–2155.

Box, J., Fettweis, X., Stroeve, J., Tedesco, M., Hall, D., & Steffen, K. (2012). Greenland ice sheet albedo feedback: thermodynamics and atmospheric drivers. *The Cryosphere*, *6*(4), 821–839.

Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., & Spinoni, J. (2014). High-resolution temperature climatology for Italy: Interpolation method intercomparison. *International Journal of Climatology*, *34*(4), 1278–1296.

Buck, A. L. (1981). New equations for computing vapor pressure and enhancement factor. *Journal of Applied Meteorology and Climatology*, *20*(12), 1527–1532.

Buisán, S. T., Earle, M. E., Collado, J. L., Kochendorfer, J., Alastrué, J., Wolff, M., Smith, C. D., & López-Moreno, J. I. (2017). Assessment of snowfall accumulation underestimation by tipping bucket gauges in the spanish operational network. *Atmospheric Measurement Techniques*, *10*(3), 1079–1091.

Bullón, J. C., & Fernández, J. (2019). Inusual episodio de nevadas de frente cálido en la costa cantábrica y catalana. *Sexto Simposio Nacional de Predicción, Memorial Antonio Mestre*, (pp. 487–496 [In Spanish]).

Caloiero, T., Veltri, S., Caloiero, P., & Frustaci, F. (2018). Drought analysis in Europe and in the Mediterranean basin using the standardized precipitation index. *Water*, *10*(8), 1043.

Casellas, E., Bech, J., Veciana, R., Mir, J. R., Sairouni, A., & Pineda, N. (2020). A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, *246*.

Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J. R., Moré, J., Rigo, T., & Sairouni, A. (2021a). Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, *147*(739), 3135–3153.

Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Miró, J. R., & Sairouni, A. (2021b). Surface precipitation phase discrimination in complex terrain. *Journal of Hydrology*, *592*, 125780.

Casellas, E., Veciana, R., & Bech, J. (2019a). pyMICA: Meteorological variable Interpolation based on Clustered data Analysis.
URL https://github.com/meteocat/pymica

Casellas, E., Veciana, R., & Bech, J. (2019b). pyPROS: Precipitation type: Rain or Snow.
URL https://github.com/meteocat/pypros

Casellas, E., Veciana, R., Sairouni, A., Pineda, N., & Bech, J. (2019c). First results of the Meteorological field Interpolation based on Clustered Analysis (MICA). In *7th International Meeting on Meteorology and Climatology of the Mediterranean*, (pp. 24 – 25).

Chandrasekar, V., Keränen, R., Lim, S., & Moisseev, D. (2013). Recent advances in classification of observations from dual polarization weather radars. *Atmospheric Research*, *119*, 97–111.

Chen, F., & Dudhia, J. (2001). Coupling an advanced land surface–hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 569–585.

Chen, R.-s., Liu, J.-f., & Song, Y.-x. (2014). Precipitation type estimation and validation in China. *Journal of Mountain Science*, *11*(4), 917–925.

Cortès, M., Llasat, M. C., Gilabert, J., Llasat-Botija, M., Turco, M., Marcos, R., Vide, J. P. M., & Falcón, L. (2018). Towards a better understanding of the evolution of the flood risk in Mediterranean urban areas: the case of Barcelona. *Natural Hazards*, *93*(1), 39–60.

Cortesi, N., González-Hidalgo, J. C., Brunetti, M., & Martin-Vide, J. (2012). Daily precipitation concentration across Europe 1971–2010. *Natural Hazards and Earth System Sciences*, *12*(9), 2799–2810.

Cortinas, J., Brill, K., & Baldwin, M. (2002). Probabilistic forecasts of precipitation type. In *Preprints, 16th Conf. on Probability and Statistics in the Atmospheric Sciences, Orlando, FL, Amer. Meteor. Soc*, vol. 3.

Cressie, N. (1990). The origins of kriging. *Mathematical geology*, *22*(3), 239–252.

Cressman, G. P. (1959). An operational objective analysis system. *Monthly Weather Review*, *87*(10), 367–374.

Dai, A. (2008). Temperature and pressure dependence of the rain-snow phase transition over land and ocean. *Geophysical Research Letters*, *35*(12).

del Moral, A., del Carmen Llasat, M., & Rigo, T. (2020). Connecting flash flood events with radar-derived convective storm characteristics on the northwestern Mediterranean coast: knowing the present for better future scenarios adaptation. *Atmospheric Research*, *238*, 104863.

Ding, B., Yang, K., Qin, J., Wang, L., Chen, Y., & He, X. (2014). The dependence of precipitation types on surface elevation and meteorological conditions and its parameterization. *Journal of hydrology*, *513*, 154–163.

Dowell, D. C., Wicker, L. J., & Snyder, C. (2011). Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Monthly Weather Review*, *139*(1), 272–294.

Dudhia, J. (1989). Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model. *Journal of Atmospheric Sciences*, *46*(20), 3077–3107.

DWD Climate Data Center (2018). Historical hourly station observations of 2m air temperature and humidity for Germany, version v006.

Ebert, E. E. (2008). Fuzzy verification of high-resolution gridded forecasts: a review and proposed framework. *Meteorological Applications*, *15*(1), 51–64.

Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press.

Elmore, K. L., Flamig, Z., Lakshmanan, V., Kaney, B., Farmer, V., Reeves, H. D., & Rothfusz, L. P. (2014). mPING: Crowd-sourcing weather reports for research. *Bulletin of the American Meteorological Society*, *95*(9), 1335–1342.

Elmore, K. L., Grams, H. M., Apps, D., & Reeves, H. D. (2015). Verifying forecast precipitation type with mPING. *Weather and Forecasting*, *30*(3), 656–667.

Farnell, C., & Rigo, T. (2020). The Lightning Jump, the 2018" Picking up Hailstones" Campaign and a Climatological Analysis for Catalonia for the 2006-2018 Period. *Tethys*, *17*, 10–20.

Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M., & Stoffel, M. (2018). Estimating the snowfall limit in alpine and pre-alpine valleys: A local evaluation of operational approaches. *Atmospheric Research*, *204*, 136–148.

Fehlmann, M., Gascón, E., Rohrer, M., Schwarb, M., & Stoffel, M. (2019). Improving medium-range forecasts of rain-on-snow events in prealpine areas. *Water Resources Research*, *55*(9), 7638–7661.

Feiccabrino, J., Graff, W., Lundberg, A., Sandström, N., & Gustafsson, D. (2015). Meteorological knowledge useful for the improvement of snow rain separation in surface based models. *Hydrology*, *2*(4), 266–288.

Feiccabrino, J., & Lundberg, A. (2009). Precipitation phase discrimination in Sweden. In *Eastern Snow Conference: 28/05/2008-30/05/2008*, (pp. 239–254).

Feiccabrino, J., Lundberg, A., & Gustafsson, D. (2012). Improving surface-based precipitation phase determination through air mass boundary identification. *Hydrology Research*, *43*(3), 179–191.

Fernandez, A. (2016). Meteorological characterization of exceptional snowfall event. *Treball Final de Grau. Facultat de Física. Universitat de Barcelona.*.

Fick, S. E., & Hijmans, R. J. (2017). Worldclim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315.

Franco, M., Sánchez-Diezma, R., & Sempere-Torres, D. (2006). Improvements in weather radar rain rate estimates using a method for identifying the vertical profile of reflectivity from volume radar scans. *Meteorologische Zeitschrift*, *15*(5), 521–536.

Frei, C. (2014). Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *International Journal of Climatology*.

Fritz, S., Fonte, C. C., & See, L. (2017). The role of citizen science in earth observation. *Remote Sensing*, *9*(4), 357.

Froidurot, S., Zin, I., Hingray, B., & Gautheron, A. (2014). Sensitivity of precipitation phase over the Swiss Alps to different meteorological variables. *Journal of Hydrometeorology*, *15*(2), 685–696.

Gandin, L. S., & Murphy, A. H. (1992). Equitable skill scores for categorical forecasts. *Monthly Weather Review*, *120*(2), 361–370.

Gao, L., Bernhardt, M., Schulz, K., & Chen, X. (2017). Elevation correction of ERA-Interim temperature data in the Tibetan Plateau. *International Journal of Climatology*, *37*(9), 3540–3552.

Garcia-Benadí, A., Bech, J., Gonzalez, S., Udina, M., & Codina, B. (2021). A new methodology to characterise the radar bright band using doppler spectral moments from vertically pointing radar observations. *Remote Sensing*, *13*(21), 4323.

Garcia-Benadí, A., Bech, J., Gonzalez, S., Udina, M., Codina, B., & Georgis, J.-F. (2020). Precipitation type classification of micro rain radar data using an improved doppler spectral processing methodology. *Remote Sensing*, *12*(24), 4113.

Gascón, E., Hewson, T., & Haiden, T. (2018). Improving predictions of precipitation type at the surface: description and verification of two new products from the ECMWF ensemble. *Weather and Forecasting*, *33*(1), 89–108.

Gavaldà, J., Benet, J., & de Carreteres, D. G. (2014). Ten years of avalanche forecasting on Bonaigua and Beret roads, Aran Valley, Spain. *17th International Road Weather Conference, SIRWEC*.

Gerrity Jr, J. P. (1992). A note on Gandin and Murphy's equitable skill score. *Monthly Weather Review*, *120*(11), 2709–2712.

Gjertsen, U., & Ødegaard, V. (2005). The water phase of precipitation—a comparison between observed, estimated and predicted values. *Atmospheric Research*, *77*(1-4), 218–231.

Golding, B. (1998). Nimrod: A system for generating automated very short range forecasts. *Meteorological Applications*, *5*(1), 1–16.

Gonzalez, S., & Bech, J. (2017). Extreme point rainfall temporal scaling: a long term (1805–2014) regional and seasonal analysis in spain. *International Journal of Climatology*, *37*(15), 5068–5079.

Gonzalez, S., Bech, J., Udina, M., Codina, B., Paci, A., & Trapero, L. (2019). Decoupling between precipitation processes and mountain wave induced circulations observed with a vertically pointing K-Band Doppler radar. *Remote Sensing*, *11*(9), 1034.

Guan, X., & Chen, C. (2014). Using social media data to understand and assess disasters. *Natural hazards*, *74*(2), 837–850.

Haberlandt, U. (2007). Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *Journal of Hydrology*, *332*(1-2), 144–157.

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

Harder, P., & Pomeroy, J. (2013). Estimating precipitation phase using a psychrometric energy balance method. *Hydrological Processes*, *27*(13), 1901–1914.

Harpold, A., Rajagopal, S., Crews, J., Winchell, T., & Schumer, R. (2017a). Relative humidity has uneven effects on shifts from snow to rain over the western US. *Geophysical Research Letters*, *44*(19), 9742–9750.

Harpold, A. A., Kaplan, M. L., Klos, P. Z., Link, T., McNamara, J. P., Rajagopal, S., Schumer, R., & Steele, C. M. (2017b). Rain or snow: hydrologic processes, observations, prediction, and research needs. *Hydrology and Earth System Sciences*, *21*(1), 1–22.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). *The Elements of Statistical Learning*. Springer Series in Statistics. New York, NY, USA: Springer New York Inc.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999–2049.

Heymsfield, A. J., Bansemer, A., Theis, A., & Schmitt, C. (2021). Survival of snow in the melting layer: Relative humidity influence. *Journal of the Atmospheric Sciences*, *78*(6), 1823–1845.

Hiebl, J., Auer, I., Bohm, R., Schoner, W., Maugeri, M., Lentini, G., Spinoni, J., Brunetti, M., Nanni, T., Tadic, M. P., et al. (2009). A high-resolution 1961-1990 monthly temperature climatology for the greater Alpine region. *Meteorologische Zeitschrift*, *18*(5), 507.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *25*(15), 1965–1978.

Hong, S.-Y., Dudhia, J., & Chen, S.-H. (2004). A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Monthly weather review*, *132*(1), 103–120.

Hong, S.-Y., Noh, Y., & Dudhia, J. (2006). A new vertical diffusion package with an explicit treatment of entrainment processes. *Monthly Weather Review*, *134*(9), 2318–2341.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., & Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701–722.

Huang, L. X., Isaac, G. A., & Sheng, G. (2012). Integrating nwp forecasts and observation data to improve nowcasting accuracy. *Weather and forecasting*, *27*(4), 938–953.

Huntington, T. G. (2006). Evidence for intensification of the global water cycle: Review and synthesis. *Journal of Hydrology*, *319*(1-4), 83–95.

Hynčica, M., & Huth, R. (2019). Long-term changes in precipitation phase in Europe in cold half year. *Atmospheric Research*, *227*, 79–88.

Ikeda, K., Steiner, M., Pinto, J., & Alexander, C. (2013). Evaluation of cold-season precipitation forecasts generated by the hourly updating High-Resolution Rapid Refresh model. *Weather and forecasting*, *28*(4), 921–939.

Ikeda, K., Steiner, M., & Thompson, G. (2017). Examination of mixed-phase precipitation forecasts from the High-Resolution Rapid Refresh model using surface observations and sounding data. *Weather and forecasting*, *32*(3), 949–967.

Irannezhad, M., Ronkanen, A.-K., Kiani, S., Chen, D., & Kløve, B. (2017). Long-term variability and trends in annual snowfall/total precipitation ratio in Finland and the role of atmospheric circulation patterns. *Cold Regions Science and Technology*, *143*, 23–31.

Isaac, G. A., Bailey, M., Boudala, F. S., Burrows, W. R., Cober, S. G., Crawford, R. W., Donaldson, N., Gultepe, I., Hansen, B., Heckman, I., et al. (2014a). The Canadian Airport Nowcasting System (CAN-Now). *Meteorological Applications*, *21*(1), 30–49.

Isaac, G. A., Joe, P., Mailhot, J., Bailey, M., Bélair, S., Boudala, F., Brugman, M., Campos, E., Carpenter, R., Crawford, R., et al. (2014b). Science of Nowcasting

Olympic Weather for Vancouver 2010 (SNOW-V10): a world weather research programme project. *Pure and Applied Geophysics*, *171*(1), 1–24.

Jarvis, A., Reuter, H., Nelson, A., & Guevara, E. (2008). Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture CIAT. *srtm.csi.cgiar.org*.

Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. *Nature communications*, *9*(1), 1148.

Jepsen, S., Harmon, T., Meadows, M., & Hunsaker, C. (2016). Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, USA. *Journal of Hydrology*, *533*, 332–342.

Jolliffe, I. T. (2007). Uncertainty and inference for verification measures. *Weather and Forecasting*, *22*(3), 637–650.

Jolliffe, I. T., & Stephenson, D. B. (2012). *Forecast verification: a practitioner's guide in atmospheric science*. John Wiley & Sons.

Joly, D., Brossard, T., Cardot, H., Cavailhes, J., Hilal, M., & Wavresky, P. (2011). Temperature interpolation based on local information: the example of France. *International Journal of Climatology*, *31*(14), 2141–2153.

Kain, J. S. (2004). The Kain–Fritsch convective parameterization: an update. *Journal of Applied Meteorology*, *43*(1), 170–181.

Kain, J. S., Goss, S. M., & Baldwin, M. E. (2000). The melting effect as a factor in precipitation-type forecasting. *Weather and Forecasting*, *15*(6), 700–714.

Kann, A., Habrovsky R., Slak, A. Š., Bujňák, R., Schmid, F., Tarjáni, V., Wang, Y., Wastl, C., Bica, B., Meirold-Mautner, I., et al. (2015). High-resolution nowcasting and its application in road maintenance: experiences from the INCA Central European area project. *IET Intelligent Transport Systems*, *9*(5), 539–546.

Kaspar, F., Müller-Westermeier, G., Penda, E., Mächel, H., Zimmermann, K., Kaiser-Weiss, A., & Deutschländer, T. (2013). Monitoring of climate change in Germany–data, products and services of Germany's National Climate Data Centre. *Advances in Science and Research*, *10*(1), 99–106.

Keeter, K. K., & Cline, J. W. (1991). The objective use of observed and forecast thickness values to predict precipitation type in North Carolina. *Weather and forecasting*, *6*(4), 456–469.

Keis, F. (2015). WHITE–Winter hazards in terminal environment: An automated nowcasting system for Munich Airport. *Meteorologische zeitschrift*, *24*(1), 61–82.

Kienzle, S. W. (2008). A new temperature based method to separate rain and snow. *Hydrological Processes: An International Journal*, *22*(26), 5067–5085.

King, J., Cabrera, A., & Kelly, R. (2009). The snowtweets project: Communicating snow depth measurements from specialists and non-specialists via mobile communication technologies and social networks. In *AGU Fall Meeting Abstracts*, vol. 2009, (pp. ED11A–0562).

Koistinen, J., & Saltikoff, E. (1998). Experience of customer products of accumulated snow, sleet and rain. *COST75 Advanced Weather Radar Systems*, (pp. 397–406).

Kormos, P. R., Marks, D. G., Seyfried, M. S., Havens, S. C., Hedrick, A., Lohse, K. A., Sandusky, M., Kahl, A., & Garen, D. (2018). 31 years of hourly spatially distributed air temperature, humidity, and precipitation amount and phase from Reynolds Critical Zone Observatory. *Earth System Science Data*, *10*(2), 1197–1205.

Krähenmann, S., Walter, A., Brienen, S., Imbery, F., & Matzarakis, A. (2018). High-resolution grids of hourly meteorological variables for Germany. *Theoretical and Applied Climatology*, *131*(3-4), 899–926.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The tropical rainfall measuring mission (TRMM) sensor package. *Journal of atmospheric and oceanic technology*, *15*(3), 809–817.

Kurtzman, D., & Kadmon, R. (1999). Mapping of temperature variables in Israel: a comparison of different interpolation methods. *Climate Research*, *13*(1), 33–43.

Kutner, M. H., Nachtsheim, C. J., Neter, J., Li, W., et al. (2005). Applied linear statistical models. *McGraw-Hill New York*.

Lagerquist, R., McGovern, A., & Gagne II, D. J. (2019). Deep learning for spatially explicit prediction of synoptic-scale fronts. *Weather and Forecasting*, *34*(4), 1137–1160.

Lambrecht, K., Hatchett, B., Walsh, L., Collins, M., & Tolby, Z. (2019). Improving visual communication of weather forecasts with rhetoric. *Bulletin of the American Meteorological Society*, *100*(4), 557–563.

Lanfredi, M., Coppola, R., D'Emilio, M., Imbrenda, V., Macchiato, M., & Simoniello, T. (2015). A geostatistics-assisted approach to the deterministic approximation of climate data. *Environmental Modelling & Software*, *66*, 69–77.

Lauri, T., Koistinen, J., & Moisseev, D. (2012). Advection-based adjustment of radar measurements. *Monthly Weather Review*, *140*(3), 1014–1022.

Lawrence, M. G. (2005). The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. *Bulletin of the American Meteorological Society*, *86*(2), 225–234.

Le Roux, R., De Resseguier, L., Corpetti, T., Jégou, N., Madelin, M., Van Leeuwen, C., & Quénol, H. (2017). Comparison of two fine scale spatial models for mapping temperatures inside winegrowing areas. *Agricultural and Forest Meteorology*, *247*, 159–169.

Leinonen, J., Chandrasekar, V., & Moisseev, D. (2012). A Bayesian Algorithm for Tangential Deconvolution of Weather Radar Images. In *European Conference on Radar in Meteorology and Hydrology (ERAD)*.

Liu, G. (2008). Deriving snow cloud characteristics from CloudSat observations. *Journal of Geophysical Research: Atmospheres*, *113*(D8).

Liu, S., Yan, D., Qin, T., Weng, B., Lu, Y., Dong, G., & Gong, B. (2018). Precipitation phase separation schemes in the Naqu River basin, eastern Tibetan plateau. *Theoretical and applied climatology*, *131*(1-2), 399–411.

Liuzzo, L., Bono, E., Sammartano, V., & Freni, G. (2017). Long-term temperature changes in Sicily, Southern Italy. *Atmospheric Research*, *198*, 44–55.

Livada, I., & Assimakopoulos, V. (2007). Spatial and temporal analysis of drought in Greece using the Standardized Precipitation Index (SPI). *Theoretical and applied climatology*, *89*(3), 143–153.

Llabrés-Brustenga, A., Rius, A., Rodríguez-Solà, R., Casas-Castillo, M. C., & Redaño, A. (2019). Quality control process of the daily rainfall series available in Catalonia from 1855 to the present. *Theoretical and applied climatology*, *137*(3), 2715–2729.

Llasat, M., Turco, M., Quintana-Seguí, P., & Llasat-Botija, M. (2014). The snow storm of 8 March 2010 in Catalonia (Spain): a paradigmatic wet-snow event with a high societal impact. *Natural Hazards and Earth System Sciences*, *14*(2), 427.

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

Lowe, D., Ebi, K. L., & Forsberg, B. (2011). Heatwave early warning systems and adaptation advice to reduce human health consequences of heatwaves. *International Journal of Environmental Research and Public Health*, *8*(12), 4623–4648. URL http://www.mdpi.com/1660-4601/8/12/4623

Lu, G. Y., & Wong, D. W. (2008). An adaptive inverse-distance weighting spatial interpolation technique. *Computers & geosciences*, *34*(9), 1044–1055.

Lumb, F. E. (1961). The problem of forecasting: the downward penetration of snow. *Meteor. Mag.*, (90), 310–319.

Lussana, C., Tveito, O. E., & Uboldi, F. (2018). Three-dimensional spatial interpolation of 2 m temperature over Norway. *Quarterly Journal of the Royal Meteorological Society*.

Malin, F., Norros, I., & Innamaa, S. (2019). Accident risk of road and weather conditions on different road types. *Accident Analysis & Prevention*, *122*, 181–188.

Manikin, S. G. (2005). An overview of precipitation type forecasting using NAM and SREF data. In *21st Conference on Weather Analysis and Forecasting/17th Conference on Numerical Weather Prediction*.

Manzato, A. (2007). A note on the maximum Peirce skill score. *Weather and Forecasting*, *22*(5), 1148–1154.

Mariani, L. (2009). Fog in the Po valley: some meteo-climatic aspects. *Ital. J. Agrometeorol*, *3*, 35–44.

Marks, D., Winstral, A., Reba, M., Pomeroy, J., & Kumar, M. (2013). An evaluation of methods for determining during-storm precipitation phase and the rain/snow transition elevation at the surface in a mountain basin. *Advances in Water Resources*, *55*, 98–110.

Martín-Vide, J., Raso Nadal, J., & Morera Palacios, A. (2008). Atles climàtic de Catalunya, període 1961-1990: termopluviometria. *Institut Cartogràfic de Catalunya i Servei Meteorològic de Catalunya, Generalitat de Catalunya, Barcelona, Spain*.

Martınez, D., Cuxart, J., Cunillera, J., & de Mallorca, P. (2008). Conditioned climatology for stably stratified nights in the Lleida area. *Journal of Weather & Climate of the Western Mediterranean*, *5*, 13–24.

Marty, C., & Meister, R. (2012). Long-term snow and weather observations at Weissfluhjoch and its relation to other high-altitude observatories in the Alps. *Theoretical and Applied Climatology*, *110*(4), 573–583.

Mathbout, S., Lopez-Bustins, J. A., Martin-Vide, J., Bech, J., & Rodrigo, F. S. (2018). Spatial and temporal analysis of drought variability at several time scales in Syria during 1961–2012. *Atmospheric Research*, *200*, 153–168.

Matsuo, T., & Sasyo, Y. (1981). Melting of snowflakes below freezing level in the atmosphere. *Journal of the Meteorological Society of Japan. Ser. II*, *59*(1), 10–25.

McCabe, A., Swinbank, R., Tennant, W., & Lock, A. (2016). Representing model uncertainty in the Met Office convection-permitting ensemble prediction system and its impact on fog forecasting. *Quarterly Journal of the Royal Meteorological Society*, *142*(700), 2897–2910.

McGovern, A., Elmore, K. L., Gagne, D. J., Haupt, S. E., Karstens, C. D., Lagerquist, R., Smith, T., & Williams, J. K. (2017). Using artificial intelligence to improve real-time decision-making for high-impact weather. *Bulletin of the American Meteorological Society*, *98*(10), 2073–2090.

McGovern, A., Lagerquist, R., John Gagne, D., Jergensen, G. E., Elmore, K. L., Homeyer, C. R., & Smith, T. (2019). Making the black box more transparent: Understanding the physical implications of machine learning. *Bulletin of the American Meteorological Society*, *100*(11), 2175–2199.

Metzger, M. J., Bunce, R. G. H., Jongman, R. H., Mücher, C. A., & Watkins, J. W. (2005). A climatic stratification of the environment of Europe. *Global ecology and biogeography*, *14*(6), 549–563.

Minder, J. R., Durran, D. R., & Roe, G. H. (2011). Mesoscale controls on the mountainside snow line. *Journal of Atmospheric Sciences*, *68*(9), 2107–2127.

Mittermaier, M. P., Hogan, R. J., & Illingworth, A. J. (2004). Using mesoscale model winds for correcting wind-drift errors in radar estimates of surface rainfall. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(601), 2105–2123.

Mlawer, E., Taubman, S., Brown, P., Iacono, M., & Clough, S. (1997). Radiative transfer for inhomogeneous atmosphere: RRTM, a validated correlated-k model for the long wave. *J. Geophys. Res*, *102*, 16.

Moon, S.-H., & Kim, Y.-H. (2020). An improved forecast of precipitation type using correlation-based feature selection and multinomial logistic regression. *Atmospheric Research*, *240*, 104928.

Moral, F. J. (2010). Comparison of different geostatistical approaches to map climate variables: application to precipitation. *International Journal of Climatology*, *30*(4), 620–631.

Nastos, P. T., Politi, N., & Kapsomenakis, J. (2013). Spatial and temporal variability of the Aridity Index in Greece. *Atmospheric Research*, *119*, 140–152.

Ninyerola, M., Pons, X., & Roure, J. M. (2000). A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *20*(14), 1823–1841.

Ninyerola, M., Pons, X., & Roure, J. M. (2007). Objective air temperature mapping for the Iberian Peninsula using spatial interpolation and GIS. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *27*(9), 1231–1242.

Nistor, M. M. (2016). Spatial distribution of climate indices in the Emilia-Romagna region. *Meteorological applications*, *23*(2), 304–313.

Pagès, M., Pepin, N., & Miró, J. (2017). Measurement and modelling of temperature cold pools in the Cerdanya valley (Pyrenees), Spain. *Meteorological Applications*, *24*(2), 290–302.

Papagiannaki, K., Lagouvardos, K., & Kotroni, V. (2013). A database of high-impact weather events in Greece: a descriptive impact analysis for the period 2001–2011. *Natural Hazards and Earth System Sciences*, *13*(3), 727–736.

Peters, G., Fischer, B., & Andersson, T. (2002). Rain observations with a vertically looking micro rain radar (mrr). *Boreal environment research*, *7*(4), 353–362.

Piazza, A. D., Conti, F. L., Viola, F., Eccel, E., & Noto, L. V. (2015). Comparative analysis of spatial interpolation methods in the Mediterranean area: application to temperature in Sicily. *Water*, *7*(5), 1866–1888.

Pickering, B. S., Neely III, R. R., & Harrison, D. (2019). The Disdrometer Verification Network (DiVeN): A UK network of laser precipitation instruments. *Atmospheric Measurement Techniques*, *12*(11), 5845–5861.

Pierce, C., Seed, A., Ballard, S., Simonin, D., & Li, Z. (2012). Nowcasting. *Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications, Joan Bech and Jorge Luis Chau, IntechOpen*.

Pomeroy, J., Gray, D., Brown, T., Hedstrom, N., Quinton, W., Granger, R., & Carey, S. (2007). The cold regions hydrological model: a platform for basing process representation and model structure on physical evidence. *Hydrological Processes: An International Journal*, *21*(19), 2650–2667.

Pulkkinen, S., Nerini, D., Pérez Hortal, A. A., Velasco-Forero, C., Seed, A., Germann, U., & Foresti, L. (2019). Pysteps: an open-source python library for probabilistic precipitation nowcasting (v1. 0). *Geoscientific Model Development*, *12*(10), 4185–4219.

Quéno, L., Vionnet, V., Cabot, F., Vrécourt, D., & Dombrowski-Etchevers, I. (2018). Forecasting and modelling ice layer formation on the snowpack due to freezing precipitation in the Pyrenees. *Cold Regions Science and Technology*, *146*, 19–31.

Ramer, J. (1993). An empirical technique for diagnosing precipitation type from model output. In *International Conference on Aviation Weather Systems, 5 th, Vienna, VA*, (pp. 227–230).

Rasmussen, R., Baker, B., Kochendorfer, J., Meyers, T., Landolt, S., Fischer, A. P., Black, J., Thériault, J. M., Kucera, P., Gochis, D., et al. (2012). How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bulletin of the American Meteorological Society*, *93*(6), 811–829.

Ravuri, S., Lenc, K., Willson, M., Kangin, D., Lam, R., Mirowski, P., Fitzsimons, M., Athanassiadou, M., Kashem, S., Madge, S., Prudden, R., Mandhane, A., Clark, A., Brock, A., Simonyan, K., Hadsell, R., Robinson, N., Clancy, E., Arribas, A., &

Mohamed, S. (2021). Skilful precipitation nowcasting using deep generative models of radar. *Nature*, *597*, 672–677.

Reeves, H. D., Elmore, K. L., Ryzhkov, A., Schuur, T., & Krause, J. (2014). Sources of uncertainty in precipitation-type forecasting. *Weather and forecasting*, *29*(4), 936–953.

Reges, H. W., Doesken, N., Turner, J., Newman, N., Bergantino, A., & Schwalbe, Z. (2016). CoCoRaHS: The evolution and accomplishments of a volunteer rain gauge network. *Bulletin of the American Meteorological Society*, *97*(10), 1831–1846.

Ripberger, J. T., Jenkins-Smith, H. C., Silva, C. L., Carlson, D. E., & Henderson, M. (2014). Social media and severe weather: Do tweets provide a valid indicator of public attention to severe weather risk communication? *Weather, Climate, and Society*, *6*(4), 520–530.

Ripoll, R., del Amo, X., & Vendrell, R. (2016). The weather observers network of the Meteorological Service of Catalonia. In *WMO Technical conference on meteorological and environmental intruments and methods of observation TECO-2016*. WMO.

Rodríguez, O., Bech, J., Arus, J., Castan, S., Figuerola, F., & Rigo, T. (2021). An overview of tornado and waterspout events in Catalonia (2000–2019). *Atmospheric Research*, *250*, 105415.

Rogelis, M., & Werner, M. (2013). Spatial interpolation for real-time rainfall field estimation in areas with complex topography. *Journal of Hydrometeorology*, *14*(1), 85–104.

Rossa, A., Haase, G., Keil, C., Alberoni, P., Ballard, S., Bech, J., Germann, U., Pfeifer, M., & Salonen, K. (2010). Propagation of uncertainty from observing systems into NWP: COST-731 Working Group 1. *Atmospheric Science Letters*, *11*(2), 145–152.

Rouf, T., Mei, Y., Maggioni, V., Houser, P., & Noonan, M. (2020). A physically based atmospheric variables downscaling technique. *Journal of Hydrometeorology*, *21*(1), 93–108.

Ryzhkov, A., & Zrnic, D. (1998). Discrimination between rain and snow with a polarimetric radar. *Journal of Applied Meteorology*, *37*(10), 1228–1240.

Sadeghi, S.-H., Peters, T. R., Cobos, D. R., Loescher, H. W., & Campbell, C. S. (2013). Direct Calculation of Thermodynamic Wet-Bulb Temperature as a Function

of Pressure and Elevation. *Journal of Atmospheric and Oceanic Technology*, *30*(8), 1757–1765.

Saltikoff, E. (2012). Measuring snow with weather radar. In *Doppler Radar Observations-Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*. IntechOpen.

Saltikoff, E., Hagen, M., Juntti, H., Kaltenböck, R., & Pulkkinen, S. (2018). Nowcasting snow for airports at heterogeneous terrain. *Geophysica*, *53*(1), 29–41.

Saltikoff, E., Tuovinen, J.-P., Kotro, J., Kuitunen, T., & Hohti, H. (2010). A climatological comparison of radar and ground observations of hail in finland. *Journal of applied meteorology and climatology*, *49*(1), 101–114.

Sánchez-Diezma, R., Sempere-Torres, D., Bech, J., & Velasco, E. (2002). Development of a hydrometeorological flood warning system (EHIMI) based on radar data. In *Second European Conference on Radar Meteorology, Delft, Netherlands*.

Schmid, W., & Mathis, A. (2004). Validation of methods to detect winter precipitation and retrieve precipitation type. *Annalen der Meteorologie*, *40*.

Schmidlin, T. W. (1993). Impacts of severe winter weather during December 1989 in the Lake Erie snowbelt. *Journal of climate*, *6*(4), 759–767.

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, *115*(1), 15–40.

Schröder, W., Schmidt, G., & Hasenclever, J. (2006). Geostatistical analysis of data on air temperature and plant phenology from Baden-Württemberg (Germany) as a basis for regional scaled models of climate change. *Environmental monitoring and assessment*, *120*(1), 27–43.

Schuster, S. S., Blong, R. J., & Speer, M. S. (2005). A hail climatology of the greater Sydney area and New South Wales, Australia. *International Journal of Climatology*, *25*(12), 1633–1650.

Schuur, T. J., Park, H.-S., Ryzhkov, A. V., & Reeves, H. D. (2012). Classification of precipitation types during transitional winter weather using the RUC model and

polarimetric radar retrievals. *Journal of Applied Meteorology and Climatology*, *51*(4), 763–779.

Seed, A. (2003). A dynamic and spatial scaling approach to advection forecasting. *Journal of Applied Meteorology*, *42*(3), 381–388.

Serra, A., Mercè, B., & Vendrell, R. (2016). Automatic weather stations network (XEMA) of the Meteorological Service of Catalonia (SMC). In *WMO Technical Conference on Meteorological and Environmental Instruments and Methods of Observation (CIMO TECO 2016): Ensuring sustained high-quality meteorological observations from sea, land and upper atmosphere in a changing world*. WMO.
URL https://library.wmo.int/index.php?lvl=notice_display&id=19676#.XEWdHvZ7m70

Sheridan, P., Smith, S., Brown, A., & Vosper, S. (2010). A simple height-based correction for temperature downscaling in complex terrain. *Meteorological Applications*, *17*(3), 329–339.

Shrestha, M., Wang, L., Koike, T., Tsutsui, H., Xue, Y., & Hirabayashi, Y. (2014). Correcting basin-scale snowfall in a mountainous basin using a distributed snowmelt model and remote-sensing data. *Hydrology and Earth System Sciences*, *18*(2), 747–761.

Sideris, I., Gabella, M., Erdin, R., & Germann, U. (2014). Real-time radar–rain-gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Quarterly Journal of the Royal Meteorological Society*, *140*(680), 1097–1111.

Sikorska, A. E., & Seibert, J. (2018). Value of different precipitation data for flood prediction in an alpine catchment: A Bayesian approach. *Journal of Hydrology*, *556*, 961–971.

Simonin, D., Pierce, C., Roberts, N., Ballard, S. P., & Li, Z. (2017). Performance of Met Office hourly cycling NWP-based nowcasting for precipitation forecasts. *Quarterly Journal of the Royal Meteorological Society*, *143*(708), 2862–2873.

Sims, E. M., & Liu, G. (2015). A parameterization of the probability of snow–rain transition. *Journal of Hydrometeorology*, *16*(4), 1466–1477.

Sisco, M. R., Bosetti, V., & Weber, E. U. (2017). When do extreme weather events generate attention to climate change? *Climatic change*, *143*(1), 227–241.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., & Powers, J. G. (2008). G.: A description of the advanced research wrf version 3. In *NCAR Tech. Note NCAR/TN-475+ STR*.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., et al. (2019). A description of the advanced research WRF model version 4. *National Center for Atmospheric Research: Boulder, CO, USA*, (p. 145).

Smith, A., Lott, N., & Vose, R. (2011). The integrated surface database: Recent developments and partnerships. *Bulletin of the American Meteorological Society*, *92*(6), 704–708.

Spruce, M., Arthur, R., & Williams, H. (2020). Using social media to measure impacts of named storm events in the United Kingdom and Ireland. *Meteorological Applications*, *27*(1), e1887.

Stahl, K., Moore, R., Floyer, J., Asplin, M., & McKendry, I. (2006). Comparison of approaches for spatial interpolation of daily air temperature in a large region with complex topography and highly variable station density. *Agricultural and forest meteorology*, *139*(3-4), 224–236.

Starkey, E., Parkin, G., Birkinshaw, S., Large, A., Quinn, P., & Gibson, C. (2017). Demonstrating the value of community-based ('citizen science') observations for catchment modelling and characterisation. *Journal of Hydrology*, *548*, 801–817.

Steinacker, R., Ratheiser, M., Bica, B., Chimani, B., Dorninger, M., Gepp, W., Lotteraner, C., Schneider, S., & Tschannett, S. (2006). A mesoscale data analysis and downscaling method over complex terrain. *Monthly weather review*, *134*(10), 2758–2771.

Stewart, R. E. (1992). Precipitation types in the transition region of winter storms. *Bulletin of the American Meteorological Society*, *73*(3), 287–296.

Stewart, T. R., Pielke Jr, R., & Nath, R. (2004). Understanding user decision making and the value of improved precipitation forecasts: Lessons from a case study. *Bulletin of the American Meteorological Society*, *85*(2), 223–236.

Sun, J., Xue, M., Wilson, J. W., Zawadzki, I., Ballard, S. P., Onvlee-Hooimeyer, J., Joe, P., Barker, D. M., Li, P.-W., Golding, B., et al. (2014). Use of NWP for nowcasting

convective precipitation: Recent progress and challenges. *Bulletin of the American Meteorological Society*, *95*(3), 409–426.

Szymanowski, M., & Kryza, M. (2012). Local regression models for spatial interpolation of urban heat island—an example from Wrocław, SW Poland. *Theoretical and applied climatology*, *108*(1), 53–71.

Szymanowski, M., Kryza, M., & Spallek, W. (2013). Regression-based air temperature spatial prediction models: an example from Poland. *Meteorologische Zeitschrift*, *22*(5), 577–585.

Tang, G., Behrangi, A., Ma, Z., Long, D., & Hong, Y. (2018). Downscaling of ERA-Interim Temperature in the Contiguous United States and Its Implications for Rain-Snow Partitioning. *Journal of Hydrometeorology*, *19*(7), 1215–1233.

Tapiador, F. J., Sánchez, J.-L., & García-Ortega, E. (2019). Empirical values and assumptions in the microphysics of numerical models. *Atmospheric research*, *215*, 214–238.

Thériault, J. M., Rasmussen, R., Ikeda, K., & Landolt, S. (2012). Dependence of snow gauge collection efficiency on snowflake characteristics. *Journal of applied meteorology and climatology*, *51*(4), 745–762.

Thériault, J. M., & Stewart, R. (2007). On the effects of vertical air velocity on winter precipitation types. *Natural Hazards and Earth System Sciences*, *7*(2), 231–242.

Thériault, J. M., & Stewart, R. E. (2010). A parameterization of the microphysical processes forming many types of winter precipitation. *Journal of the Atmospheric Sciences*, *67*(5), 1492–1508.

Thériault, J. M., Stewart, R. E., & Henson, W. (2010). On the dependence of winter precipitation types on temperature, precipitation rate, and associated features. *Journal of applied meteorology and climatology*, *49*(7), 1429–1442.

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

Thompson, G., Field, P. R., Rasmussen, R. M., & Hall, W. D. (2008). Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II:

Implementation of a new snow parameterization. *Monthly Weather Review*, *136*(12), 5095–5115.

Trapero, L., Bech, J., Rigo, T., Pineda, N., & Forcadell, D. (2009). Uncertainty of precipitation estimates in convective events by the Meteorological Service of Catalonia radar network. *Atmospheric Research*, *93*(1-3), 408–418.

Velasco-Forero, C. A., Sempere-Torres, D., Cassiraga, E. F., & Gómez-Hernández, J. J. (2009). A non-parametric automatic blending methodology to estimate rainfall fields from rain gauge and radar data. *Advances in Water Resources*, *32*(7), 986–1002.

Vidal, V., Bech, J., Pineda, N., & Veciana, R. (2014). Verification of a weather radar derived surface precipitation type product. *8th European Conference on Radar in Meteorology and Hydrology. Garmisch, Germany.*.

Viggiano, M., Busetto, L., Cimini, D., Di Paola, F., Geraldi, E., Ranghetti, L., Ricciardelli, E., & Romano, F. (2019). A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements. *Agricultural and Forest Meteorology*, *276*, 107590.

Vilaclara, E., Segalà, S., Andrés, A., & Aran, M. (2010). Operational warnings issued by the SMC in the 8th March snow event in Catalonia. In *12th Plinius Conference on Mediterranean Storms, held September 1-4, 2010 in Corfu Island, Greece. http://meetings. copernicus. org/plinius12, id. 79*.

Wandishin, M. S., Baldwin, M. E., Mullen, S. L., & Cortinas Jr, J. V. (2005). Short-range ensemble forecasts of precipitation type. *Weather and Forecasting*, *20*(4), 609–626.

Wang, R., Kumar, M., & Link, T. E. (2016). Potential trends in snowmelt-generated peak streamflows in a warming climate. *Geophysical Research Letters*, *43*(10), 5052–5059.

Webb, M. A., Hall, A., Kidd, D., & Minansy, B. (2016). Local-scale spatial modelling for interpolating climatic temperature variables to predict agricultural plant suitability. *Theoretical and applied climatology*, *124*(3-4), 1145–1165.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences*, vol. 100. Academic press.

WMO, S., Geneva (2017a). Guidelines for Nowcasting Techniques.
URL https://library.wmo.int/doc_num.php?explnum_id=3795

WMO, S., Geneva (2017b). International Cloud Atlas: Manual on the observation of clouds and other meteors.
URL https://cloudatlas.wmo.int/en/home.html

WMO, S., Geneva (2018). Guide to Instruments and Methods of Observation.
URL https://library.wmo.int/doc_num.php?explnum_id=10616

Zerr, R. J. (1997). Freezing rain: An observational and theoretical study. *Journal of Applied Meteorology*, *36*(12), 1647–1661.

Zhong, K., Zheng, F., Xu, X., & Qin, C. (2018). Discriminating the precipitation phase based on different temperature thresholds in the Songhua River Basin, China. *Atmospheric Research*, *205*, 48–59.

Appendix A

Contributions

This chapter includes a list of contributions made during the pre-doctoral period to scientific journals, congress, and seminars.

A.1 Papers

- Llorens, P., Gallart, F., Cayuela, C., Planasdemunt, M. R., Casellas, E., Molina, A. J., Moreno de las Heras, M., Bertran, G., Sánchez-Costa, E. & Latron, J. (2018). What have we learnt about Mediterranean catchment hydrology? 30 years observing hydrological processes in the Vallcebre research catchments. *Cuadernos de Investigación Geográfica/Geographical Research Letters*, (44), 475-502.

- Casellas, E., Latron, J., Cayuela, C., Bech, J., Udina, M., Sola, Y., Keun-Ok, L. & Llorens, P. (2019). Moisture origin and characteristics of the isotopic signature of rainfall in a Mediterranean mountain catchment (Vallcebre, eastern Pyrenees). *Journal of Hydrology*, 575, 767-779.

- Casellas, E., Bech, J., Veciana, R., Miró, J. R., Sairouni, A., & Pineda, N. (2020). A meteorological analysis interpolation scheme for high spatial-temporal resolution in complex terrain. *Atmospheric Research*, 246, 105103.

- Soula, S., Pineda, N., Georgis, J. F., Leroy, A., Vanpoucke, I., Montanya, J., Casellas, E., González, S. & Bech, J. (2021). On the conditions for winter lightning at the Eagle Nest Tower (2537 m asl) during the Cerdanya-2017 field experiment. *Atmospheric Research*, 247, 105208.

- Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Miró, J. R., & Sairouni, A. (2021). Surface precipitation phase discrimination in complex terrain. *Journal of Hydrology*, 592, 125780.
- Casellas, E., Bech, J., Veciana, R., Pineda, N., Miró, J. R., Moré, J., Rigo, T. & Sairouni, A. (2021). Nowcasting the precipitation phase combining weather radar data, surface observations, and NWP model forecasts. *Quarterly Journal of the Royal Meteorological Society*, 147(739), 3135-3153.

A.2 Conference presentations/posters

- Casellas, E., Bech, J., Veciana, R., Pineda, N., Rigo, T., Moré, J., González, S., Paci, A. & Codina, B. (2018). Multisource data verification of a weather radar surface precipitation type product. *10th European Conference on Radar in Meteorology and Hydrology, Ede, The Netherlands.*
- Pineda, N., Montanya, J., Romero, D., Bech, J., Casellas, E., & González, S. (2018). Meteorological aspects of winter upward lightning from an instrumented tower in the Pyrenees. *34th International Conference on Lightning Protection (ICLP) (pp. 1-7). IEEE.*
- Casellas, E., Bech, J., Veciana, R., Pineda, N. & Sairouni, A. (2019). First results of the Meteorological field Interpolation based on Clustered Analysis (MICA). *7th International Conference on Meteorology and Climatology of the Mediterranean (MetMed), Palma, Spain.*
- Bech, J., Udina, M., Codina, B., Gonzalez, S., Garcia, A., Altube, P., Mercader, J., Callado, A., Arús, J., Rodríguez, O., Casellas, E., Roura-Adserias, F., Rosell, A., Polls, F., Kosovic, B., Montornès, A., Escribà, P., Trapero, L., Paci, A., Boudevillain, B. & Tokay, A. (2021). An overview of the Analysis of Precipitation Processes in the Eastern Ebro Subbasin (WISE-PreP) Project. *8th International Conference on METeorology and climatology of the MEDiterranean (MetMed). May 25-27, 2021, The University of the Balearic Islands, Palma, Spain.*
- Bech, J., Udina, M., Codina, B., Gonzalez, S., Garcia, A., Altube, P., Mercader, J., Callado, A., Arús, J., Rodríguez, O., Casellas, E., Roura-Adserias, F., Rosell, A., Polls, F., Kosovic, B., Montornès, A., Escribà, P., Trapero, L., Paci, A., Boudevillain, B. & Tokay, A. (2021). Preliminary results of the Analysis of

Precipitation Processes in the Eastern Ebro Subbasin (WISE-PreP) Field Campaign within HILIAISE. *EMS Annual Meeting Abstracts: 18th EMS European Meteorological Society Annual Meeting.*

- Bech, J., Udina, M., Codina, B., Gonzalez, S., Garcia, A., Altube, P., Mercader, J., Callado, A., Arús, J., Rodríguez, O., Casellas, E., Roura-Adserias, F., Rosell, A., Polls, F., Kosovic, B., Montornès, A., Escribà, P., Trapero, L., Paci, A., Boudevillain, B. & Tokay, A. (2021). Analysis of Precipitation Processes in the Eastern Ebro Subbasin. *LIAISE Workshop, March 8-9, 2021. National Center for Meteorological Research (CNRM)/Météo-France), Toulouse, France.*

A.3 Seminars

- Casellas, E., Bech, J., Veciana, R. & Pineda, N. (2019). Estimación remota del tipo de precipitación invernal. *Seminario Final de la Red Temática "Winter Precipitation and Strong Winds: Observational Studies (WiPSWis)". Universidad Complutense de Madrid, Madrid, Spain.*

Appendix B

Elevation correction of NWP model forecasts

B.1 Overview

The development of a nowcasting system to discriminate precipitation phase involves manifold processes. One of them, as shown in Chapter 4 and according to the methodologies selected to classify precipitation phase, is the forecast of air temperature and dew point temperature. Forecasts can be obtained through the Weather Research and Forecasting (WRF) model that it is operatively run at the Meteorological Service of Catalonia. However, as explained in Chapter 4, the spatial resolution of the nowcasting product is 1 km, whereas that of WRF is 3 km. Taking into account that some precipitation phase discrimination strategies are based on establishing thresholds to meteorological variables dependant on air temperature and dew point temperature values, these are usually highly influenced by altitude. Therefore, the spatial resolution difference between WRF and the nowcasting product matters. Thereby, one of the objectives was to refine the 3 km spatial resolution air temperature and dew point temperature fields to 1 km following a different approach than nearest neighbours or bilinear interpolation.

A common kind of methodology to increase the spatial resolution of an air temperature or dew point temperature field is to apply an elevation correction. This consists on taking into account the altitude differences between the orography at 3 km spatial resolution compared to 1 km, similarly to Figure B.1.

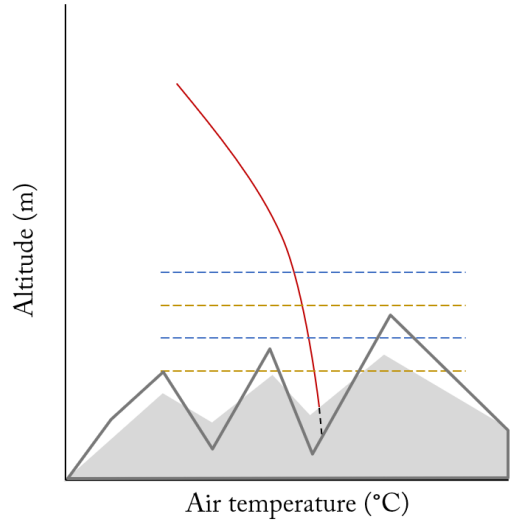

Fig. B.1: Extrapolation of surface temperature value from native NWP orography grid (light grey shape) to desired orography (dark grey line). Red line represents the vertical temperature profiles and dashed lines represent different pressure levels.

B.2 Methodologies

There are different methodologies to conduct an elevation correction to air temperature and dew point temperature fields. Most of them are based on applying a temperature lapse rate to altitude differences between the NWP native orography and the desired one following the next equation:

$$T_{1km} = T_{3km} - \gamma \cdot (z_{1km} - z_{3km}), \tag{B.1}$$

where T is temperature, γ is a temperature lapse rate, z_{1km} is the altitude given by a digital terrain model with 1 km gridlength, and z_{3km} the native NWP model altitude at 3 km resolution.

Altitude differences are usually marked in valley bottoms and high mountainous ridges. The methodologies evaluated in the present thesis are summarized in Table B.1.

A different approach was followed to obtain the elevation correction of dew point temperature fields. Rather than calculating the lapse rates, the dew point temperature depression was calculated at the native grid scale, then interpolated to 1 km spatial scale using bilinear interpolation, and finally this field was subtracted to elevation corrected

Table B.1: Descriptions of gradient calculation for each elevation correction methodology considered.

Methodology	Gradient calculation
Gao et al. (2017)	Methodology based on surface and upper-air information. The gradient is calculated considering the temperature differences between 925 hPa and 700 hPa pressure levels for each pixel of the grid.
INCA (Haiden et al., 2011)	Methodology based on surface and upper-air information. The gradient is calculated considering the maximum temperature lapse rate between the closest pressure level to the surface and a specified height above the surface among successive intervals for each pixel of the grid.
Fixed	Methodology based on surface information. The gradient considered to use in the refinement process is a constant one, 6.5°C/km for each pixel of the grid.
Regression	Methodology based on surface information. The gradient is calculated through a linear regression obtained for each pixel of the grid considering the air temperature and altitude pairs from a neighborhood around each pixel.

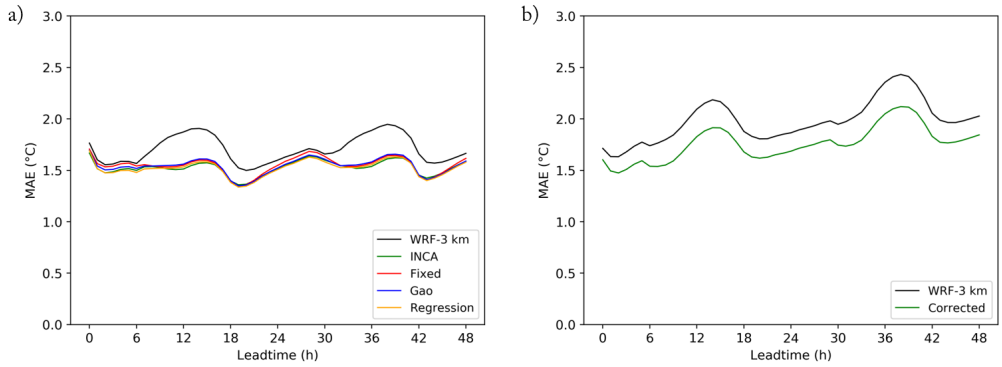

Fig. B.2: Lead time Mean Absolute Error (MAE) of (a) the different elevation correction methodologies considered for air temperature and (b) the correction applied to dew point temperature.

air temperature field.

B.3 Results

A verification study was done in order to asses the performance of each elevation correction methodology and to make an informed decision about which one would be selected for the nowcasting product. The dataset used to verify the refined air temperature and dew point temperature fields was hourly observations from XEMA network (Serra et al., 2016) for 2018. Observation and estimated values were compared considering the closest pixel of the refined field to the observation location, and Mean Absolute Errors (MAE) were computed.

Figure B.2a presents the MAE values of each methodology and those of WRF-3km without applying the refinement for air temperature. All methodologies considered exhibited a similar performance, but Regression (yellow line) presents the lowest MAE in most parts of the forecast lead time. Nevertheless, the performance differences were narrow and a decision was made accounting also for the amount of information needed to conduct the refinement of the 3 km spatial resolution field. Thereby, a gradient calculated following the Regression methodology was selected because it does not need vertical temperature information. This is an advantage because this methodology can be applied to all air temperature fields without the need of having information in all vertical levels. For example, WRF model is operatively run at the Meteorological Service of Catalonia, but other NWP models may be obtained from other sources and usually

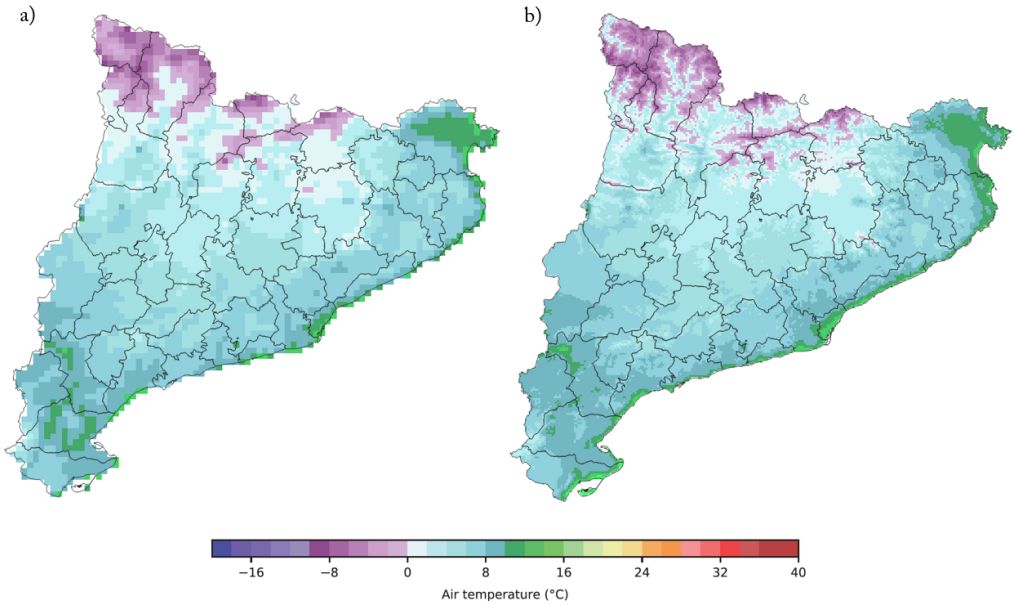

Fig. B.3: Air temperature fields corresponding to (a) WRF 3 km output and (b) elevation corrected field on 1st May 2018 00 UTC.

only some fields and specific pressure levels are available.

Regarding dew point temperature results, the refinement methodology used also yielded a reduction of MAE, similarly to air temperature as it can be seen in Figure B.2b. An example of the comparison between the 3 km air temperature field obtained from the WRF model and that of 1 km obtained through a refinement methodology is shown in Figure B.3.