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

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

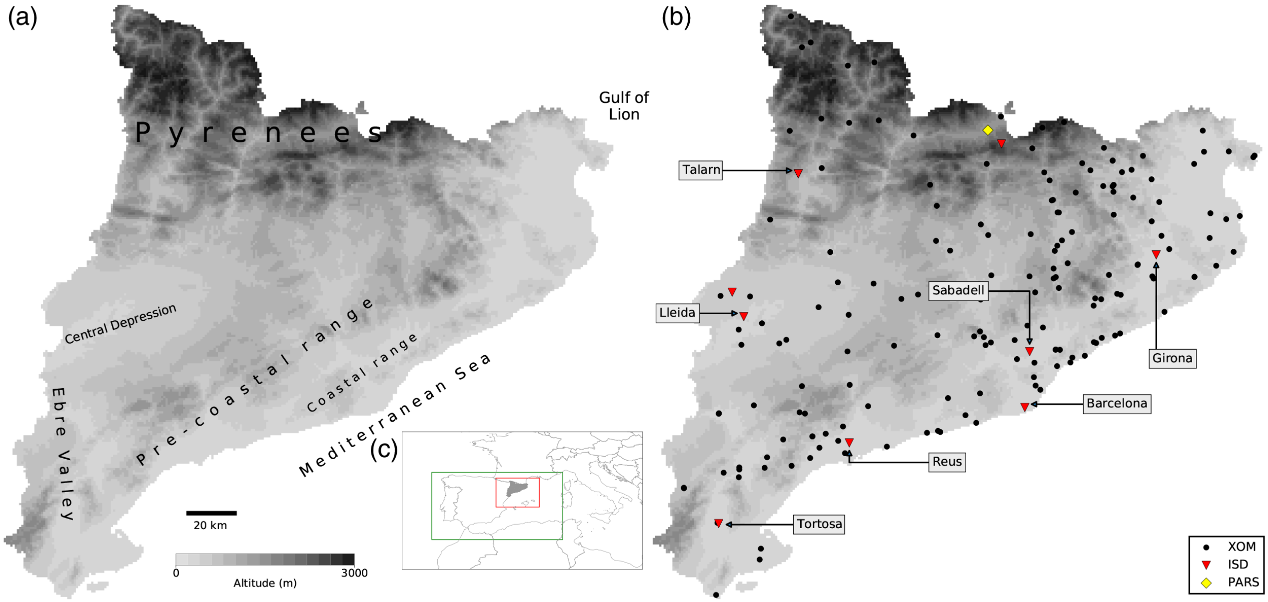


**FIGURE 1** Study area, including (a) the main orographic characteristics, (b) the location of the precipitation phase observation sources considered: Xarxa d'Observadors Meteorològics (XOM) and Integrated Surface Database (ISD), and (c) the WRF model 9 km and 3 km domains [Colour figure can be viewed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

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

where $\Delta I$ 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 $T_{\mathrm{a}}$, Delta $T_{\mathrm{d}}$, Delta $T_{\mathrm{w}}$, and Delta KS), hereafter MostDelta, while the other combines all eight algorithms (Delta $T_{\mathrm{a}}$, Delta $T_{\mathrm{d}}$, Delta $T_{\mathrm{w}}$, 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 $9 \times 9\,\mathrm{km}^2$ 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 $9 \times 9\,\mathrm{km}^2$ 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 $9 \times 9$ 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.

## 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).

**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.

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).

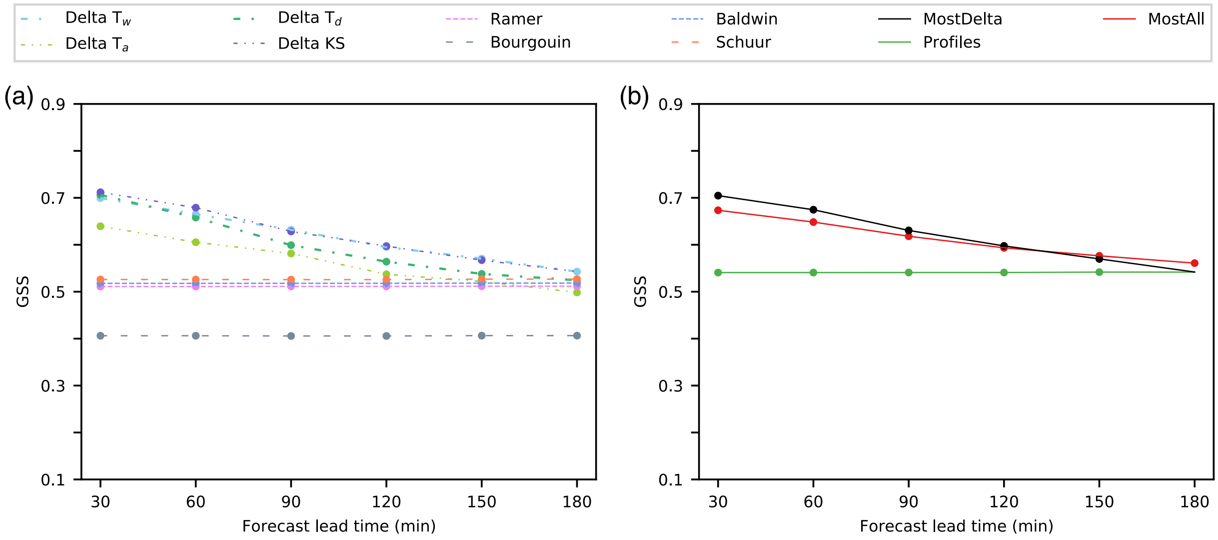


FIGURE 2 Mean Gerrity skill score (GSS) obtained from bootstrap sampling over the forecast lead time for each precipitation phase discrimination algorithm considering all available observations. (a) Scores for single precipitation phase algorithms. (b) Scores for combined precipitation phase algorithms: 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 $T_{\mathrm{w}}$ was the best in E3, E7, and E8, Delta $T_{\mathrm{d}}$ in E1 and E4, Delta $T_{\mathrm{a}}$ in E6, and Delta KS in E5 ($p$-value $< .001$). Regarding E2, Delta $T_{\mathrm{a}}$ and Delta KS presented a similar result and the null hypothesis could not be rejected; in contrast, Delta $T_{\mathrm{d}}$ 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 $T_{\mathrm{a}}$ 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 $T_{\mathrm{w}}$): 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 $T_{\mathrm{d}}$), E3 (between 30 and 90 min forecast lead time between MostDelta and Delta KS), or E7 (at 30 min forecast lead time between Delta $T_{\mathrm{a}}$ 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

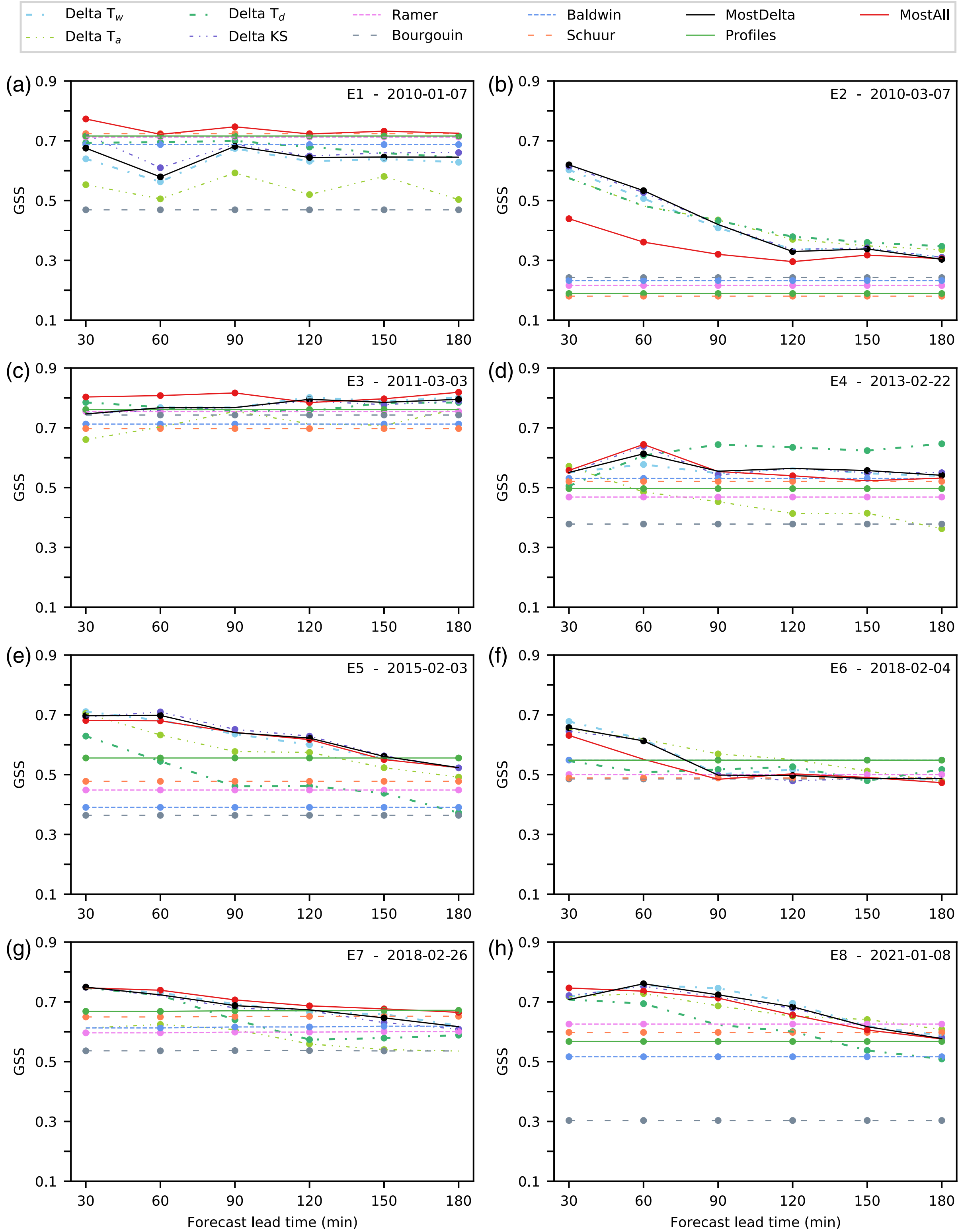


**FIGURE 3** Mean Gerrity skill score (GSS) obtained from bootstrap sampling over the forecast lead time for each precipitation phase discrimination algorithm and each event. Bullets indicate the statistical significance of score differences relative to all schemes for each lead time [Colour figure can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

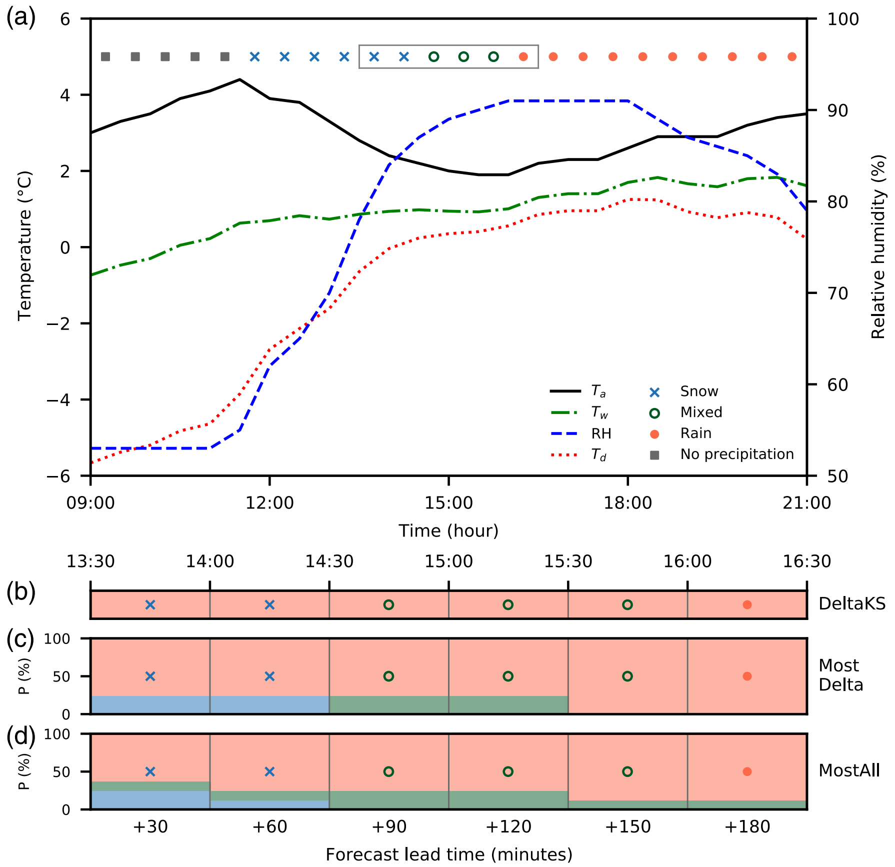


**FIGURE 4** (a) Air temperature ($T_{\mathrm{a}}$), dew-point temperature ($T_{\mathrm{d}}$), wet-bulb temperature ($T_{\mathrm{w}}$), and relative humidity (*RH*) observations on February 27, 2018, between 0900 and 2100 UTC for Barcelona. Markers represent the actual precipitation phase observed (snow, 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

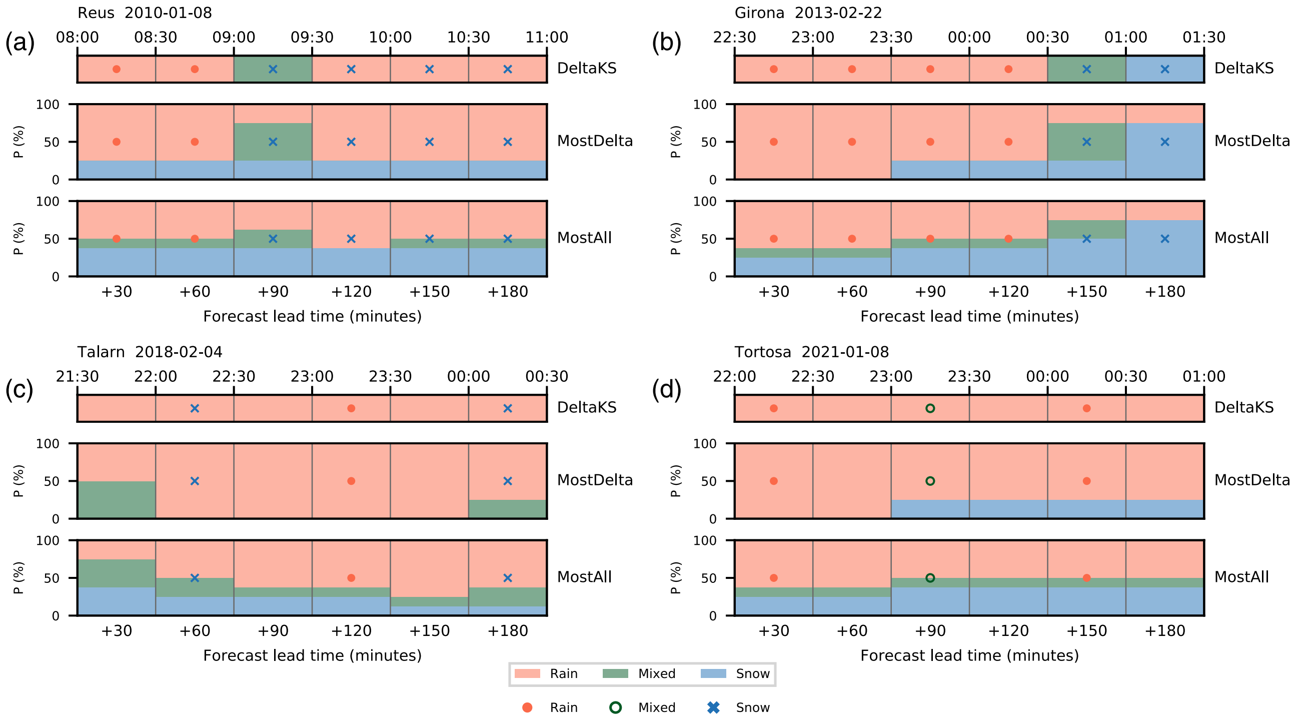


FIGURE 5 Comparison between Delta KS, MostDelta, and MostAll precipitation-phase nowcasting schemes over the forecast lead time for different locations and events. The location and nowcasting initialisation time are shown above each panel. Shaded colours indicate the probability of each precipitation 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 $T_w$,

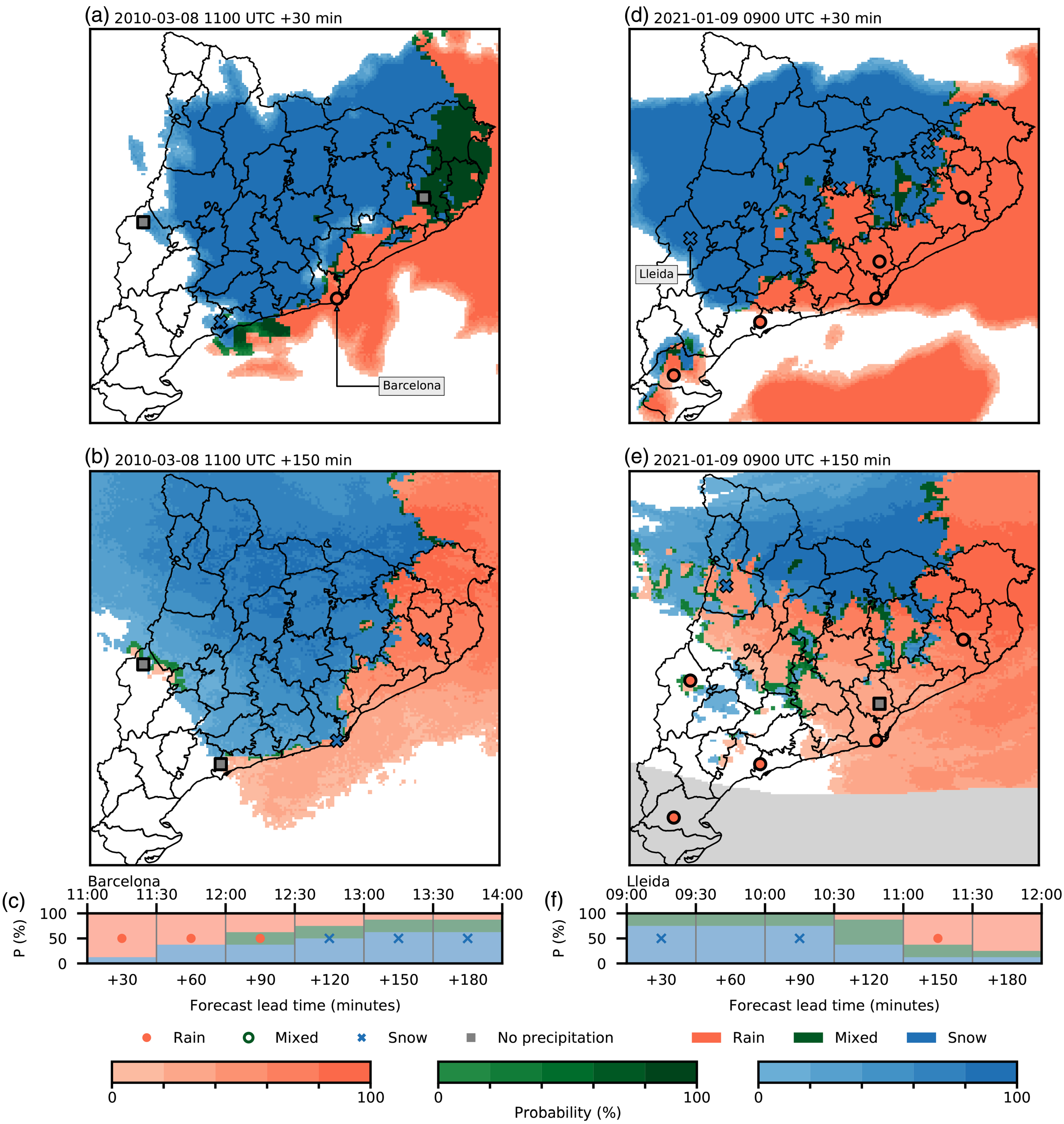


**FIGURE 6** Most probable precipitation phase (MostAll) nowcasting masked with probability of precipitation exceeding 0.1 mm·hr$^{-1}$ at (a, d) 30 and (b, e) 150 min forecast lead time. The lowest area in (e) indicates where the extrapolated precipitation field is not available. (c, f) Nowcasting for 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

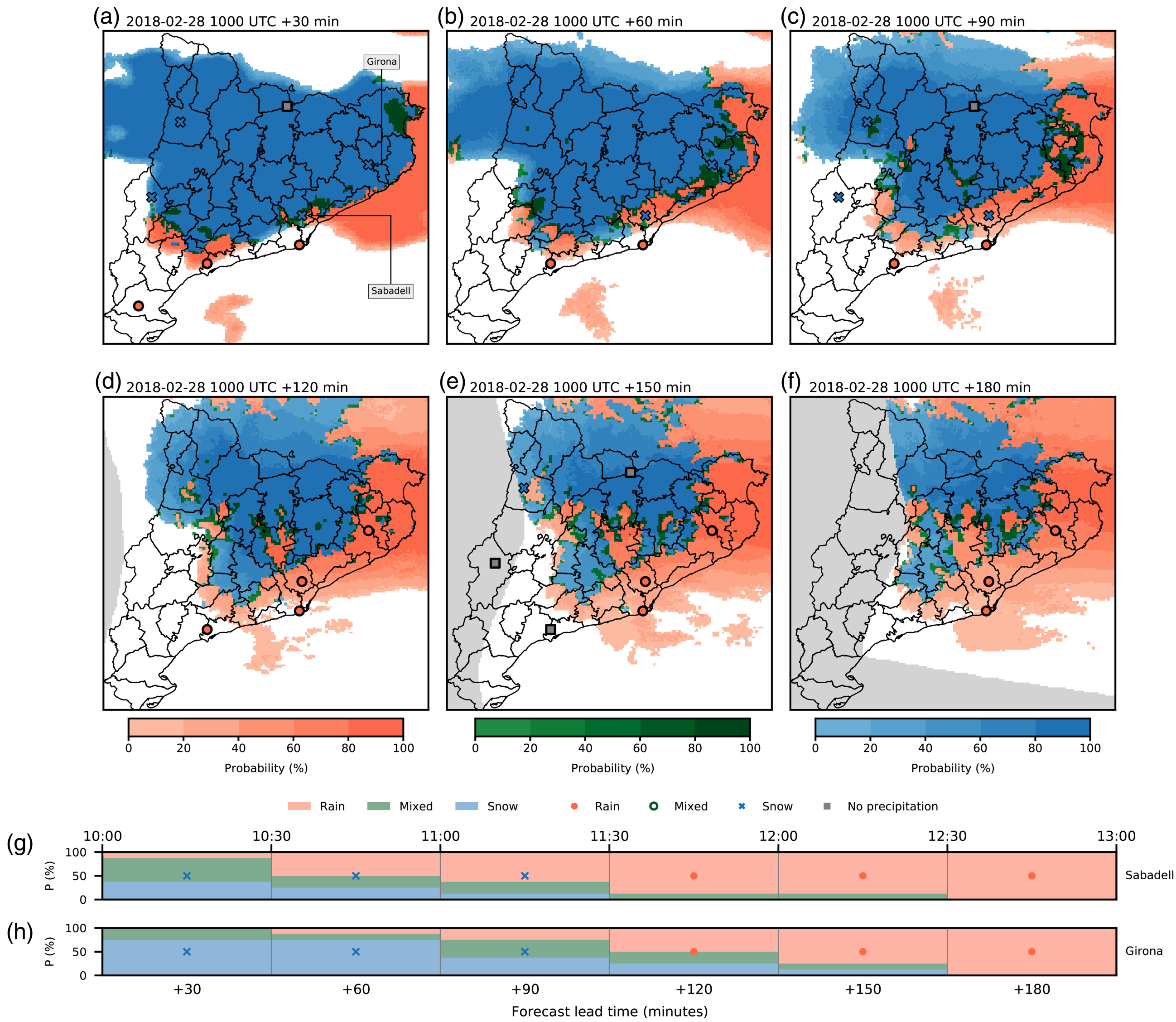


**FIGURE 7** Most probable precipitation phase (MostAll) nowcasting masked with the probability of precipitation exceeding 0.1 mm·hr$^{-1}$ on February 28, 2018, at 1000 UTC, for (a–f) 30–180 min lead time. Precipitation phases shown are rain, mixed, and snow. Nowcasting for (g) Sabadell and (h) Girona: 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 WeatherAnalysis 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