Master Thesis
TVVR 26/5020

Control of Rainwater Harvesting Systems

Optical-Flow Radar based Nowcasting

Melker Jansson

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University

Control of Rainwater Harvesting Systems

Optical-Flow based Radar Nowcasting

By:
Melker Jansson

Master Thesis

Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University
Box 118
221 00 Lund, Sweden

Water Resources Engineering
TVVR-26/5020
ISSN 1101-9824

Lund 2026
www.tvrl.lth.se

Master Thesis
Division of Water Resources Engineering
Department of Building & Environmental Technology
Lund University

English title:	Control of Rainwater Harvesting Systems - Optical Flow-based Radar Nowcasting
Author:	Melker Jansson
Supervisors:	Jonas Olsson, SMHI and Lund University Johan Kjellin, Tyréns
Examiner:	Magnus Persson, Lund University
Language:	English
Year:	2025
Keywords:	rainwater harvesting; radar nowcasting; optical flow; forecast-based control; real-time control; PySTEPS; urban stormwater

Preface

This thesis marks the final step of my engineering degree at LTH.

I would like to thank my supervisors Jonas Olsson at SMHI and LTH, and Johan Kjellin at Tyréns, for their guidance and support throughout this work. I am also grateful to Remco van de Beek at SMHI for his substantial help with PySTEPS, and to Rasmus Møller Jess at Tyréns, whose preceding thesis on the feasibility of rainwater harvesting systems provided both context and a useful reference point. The ongoing discussions with Hedvig Sundstedt, who addressed the control side of the same problem in her thesis, were a valuable part of the process. Thanks also to Frida Heskebeck at the Department of Automatic Control at Lund University for her input on model predictive control, Johan Södling at SMHI for his advice regarding all things verification, Saeed Falahat at SMHI for his input on KNEP and Ulf Nordh at SMHI for crunching all of my requested composites.

I would also like to express my gratitude to the water group at Tyréns Malmö for hosting me during the thesis period, and to VA Syd for providing context, a study visit to the FBC system Pölen at the LTH campus, and the rain gauge data used in this study.

I thank Sveriges Ingenjörers Miljöfond, whose grant gave me access to workshops with LU Innovation that helped ground the project in its real-world context.

Above all, I thank my family and friends - without whom, none of this.

Melker Jansson
Lund, May 18, 2026

Abstract

Urban water systems in southern Sweden face increasing pressure from both short-duration intense rainfall, which can cause pluvial flooding and pollutant discharges, and more frequent dry periods, which motivate water reuse (IPCC 2023; Schimanke et al. 2022; Canedo Rosso et al. 2025). Rainwater harvesting (RWH) tanks can contribute to both goals (Campisano et al. 2017), but their detention benefit is reduced when storage is unavailable ahead of storms. Forecast-based control (FBC) can create capacity through pre-storm releases (Kerkez et al. 2016; Xu et al. 2022), but its viability depends on precipitation forecast skill at short lead times.

This thesis investigates whether high-resolution radar-based nowcasting can provide forecast information of sufficient quality to support FBC of RWH systems in the Malmö area. Using the `PySTEPS` framework (Pulkkinen et al. 2019), four extrapolation-based nowcasting configurations were generated from Swedish radar composites and evaluated for 19 intense precipitation events at 0.5 km / 5 min resolution, with additional verification at the 2 km / 15 min configuration used by SMHI's operational KNEP product (SMHI 2025a). Forecasts were compared against Eulerian and Lagrangian persistence baselines using continuous, categorical, and spatial verification metrics over lead times of 5–90 min.

Nowcast skill depends strongly on season and event duration. Winter events retain useful correlation beyond 60 min for 3 h durations, whereas convective summer events decorrelate within about 15–30 min, consistent with the conventional decorrelation framing based on Pearson correlation (Germann et al. 2002). At the lead times relevant to control, S-PROG at native resolution and ANVIL at KNEP-equivalent resolution maintain probability of detection above 0.7 and frequency bias close to unity; the latter supports the operational choice of ANVIL for KNEP (Falahat 2026). The direct effect of increasing radar resolution within a fixed method is modest, but the best-performing method changes with resolution: S-PROG performs best at native resolution, while ANVIL performs best at the coarser operational configuration.

Interpreted against required tank drawdown times from parallel emptying-time simulations (Sundstedt 2026), the results indicate that 30–90 min nowcast horizons can support operational pre-storm releases under permissive outflow strategies and urban-scale outflow constraints, while stricter constraints and strategies requiring a fully empty tank at event onset require longer horizons and motivate blending with numerical weather prediction.

Contents

Acronyms

Control and optimisation

FBC	Forecast Based Control
RTC	Real-Time Control

Hydrology and urban drainage

CSO	Combined Sewer Overflow
RWH	Rainwater Harvesting
SuDS	Sustainable Drainage Systems
UDS	Urban Drainage System

Meteorology and forecasting

dBZ	decibel-Z (radar reflectivity in logarithmic units)
KNEP	Korta NEderbördsPrognoser
NWP	Numerical Weather Prediction
UTC	Coordinated Universal Time
VSRF	Very Short-Range Forecasting

Nowcasting methods and model components

ANVIL	Autoregressive Nowcasting using Vertically Integrated Liquid
AR	Autoregressive
ARI	Autoregressive Integrated
DARTS	Deformation- and Rotation-based Tracking Scheme
EP	Eulerian Persistence
LK	Lucas–Kanade (optical flow method)
LINDA	Lagrangian INtegro-Difference equation model with Autoregression
LP	Lagrangian Persistence
S-PROG	Spectral PROGnosis
STEPS	Short-Term Ensemble Prediction System
VET	Variational Echo Tracking
VIL	Vertically Integrated Liquid

Data and software tools

API	Application Programming Interface
GUI	Graphical User Interface
IoT	Internet of Things
PySTEPS	Python framework for ensemble precipitation nowcasting

Organisations and institutions

IPCC	Intergovernmental Panel on Climate Change
IWA	International Water Association
SMHI	Swedish Meteorological and Hydrological Institute
UN	United Nations
WMO	World Meteorological Organization

Operational strategies

NOD	No outflow during rain event
OFD	Outflow continued during rain event

Computing

AI	Artificial Intelligence
NaN	Not a Number

Statistics and experimental design

IQR	Interquartile Range
OAT	One-At-a-Time (experimental design)

Image processing

LoG	Laplacian-of-Gaussian

Built environment

HVAC	Heating, Ventilation, and Air Conditioning
WC	Water Closet (toilet)

Units

MEUR	Million euros

Verification metrics

AUC	Area Under the Curve
BIAS	Frequency Bias
CDF	Cumulative Distribution Function
CRPS	Continuous Ranked Probability Score
CSI	Critical Success Index
FAR	False Alarm Ratio
FSS	Fractions Skill Score
MAE	Mean Absolute Error
POD	Probability of Detection
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic

List of Figures

List of Tables

1 Introduction

1.1 Context and motivation

Urban water systems in Sweden face a dual challenge. On the one hand, short-duration, high-intensity, often convective rainfall can trigger pluvial flooding, combined sewer overflows, and pollutant discharges when drainage capacity is exceeded (Sveriges miljömål 2022; Torelló-Sentelles et al. 2025). On the other hand, drought frequency and associated water stress are expected to increase, motivating more efficient use of available water resources (IPCC 2023; Canedo Rosso et al. 2025). In southern Sweden, observed and projected changes indicate increasing precipitation amounts and intensifying extremes (Schimanke et al. 2022; An et al. 2026). At the same time, many stormwater systems remain dimensioned for historical climate conditions, further increasing vulnerability to extreme events (Svenskt Vatten 2023).

Rainwater harvesting (RWH) systems offer a dual-purpose opportunity by storing roof runoff for non-potable uses (e.g. toilet flushing and irrigation) while also providing decentralised detention capacity (Campisano et al. 2017). However, RWH systems are commonly designed and evaluated using static water-balance approaches based on historical rainfall and simplified demand assumptions (Semaan et al. 2020). In practice, this can lead to suboptimal storage utilisation: tanks may be near full prior to intense rainfall, limiting their ability to mitigate peak inflows, or be emptied unnecessarily, reducing water-supply benefits.

Real-time control (RTC) has emerged as a promising approach to increase the performance of existing storage and conveyance without expanding infrastructure (Sweetapple et al. 2026). By combining sensors, communication, and actuators, RTC can actively regulate storage levels based on observed system states (Mullapudi et al. 2018). When precipitation forecasts are used to enable anticipatory operation, the approach is commonly referred to as forecast-based control (FBC) (Kerkez et al. 2016). For RWH and related decentralised storages, FBC typically aims to create available storage capacity through pre-storm releases prior to forecast rainfall (Xu et al. 2022). Existing applications indicate substantial potential, including reductions in overflow volumes while maintaining water-reuse performance in Swedish conditions (Knutsson 2025). However, realised benefit depends fundamentally on the quality, lead time, and resolution of the rainfall information used to trigger control actions (Knutsson 2025; Degadwala et al. 2024).

For short decision horizons in urban applications, precipitation nowcasting (0–6 h) is particularly relevant. Nowcasting targets lead times of minutes to a few hours and is often dominated by observation-driven methods, especially radar-based extrapolation (Prudden et al. 2020). In this regime, radar nowcasts can provide higher spatial and temporal detail than numerical weather prediction, but forecast skill can decay rapidly with lead time, especially for convective precipitation (Prudden et al. 2020; Imhoff et al. 2020). Operational products such as KNEP (Korta NEderbördsPrognoser) (SMHI 2025a) operate at 15 min / 2 km resolution. With recent advances in Swedish radar technology, higher-resolution radar input at 5 min / 0.5 km may provide meaningful additional skill at the spatial scales relevant to building-level RWH control. This remains an open question, explored in this thesis.

In this thesis, high-resolution radar-based nowcasts are generated using `PySTEPS`, an open-source Python library for short-term precipitation nowcasting that provides optical-flow-based motion estimation, extrapolation, and stochastic ensemble generation (Pulkkinen et al. 2019). Nowcast performance is evaluated against simple persistence baselines and the nowcasting part of the operational KNEP configuration across a set of intense rainfall events in the Malmö area, and the results are interpreted in terms of the forecast skill required for reliable forecast-based control of RWH systems.

Throughout this thesis, forecast skill is used in its verification sense: the quality of a nowcast measured against observations and benchmarked against reference (baseline) forecasts, rather than judged by absolute accuracy alone. The references adopted here are Eulerian and Lagrangian persistence, so a nowcast is said to possess skill insofar as it outperforms these baselines. Because no single measure captures all relevant aspects of a precipitation forecast, skill is assessed through a set of complementary metrics addressing intensity error, the occurrence and location of rainfall, spatial structure, and - for ensemble forecasts - probabilistic reliability and discrimination (Chapter 3.4). Skill generally degrades with lead time; the skilful lead time denotes the horizon up to which a nowcast retains useful skill relative to these baselines. In the forecast-based control setting that motivates this thesis, skill is regarded as sufficient when it remains adequate at the lead time required to initiate a pre-storm release.

1.2 Research purpose and questions

The purpose of this work is to investigate whether high-resolution radar-based nowcasting can provide forecast information of sufficient quality to support forecast-based control of rainwater harvesting systems in an urban context. Precipitation forecasts are generated using `PySTEPS` and evaluated against persistence (Eulerian and Lagrangian) baselines and the operational KNEP configuration across a sample of intense rainfall events in the Malmö area, stratified by event duration and season.

The work is guided by the following research questions:

(i) How does nowcast skill vary with lead time, event duration, and season for intense rainfall events in the Malmö area, and how does this compare to simple persistence baselines?

(ii) Does higher radar input resolution (0.5 km / 5 min) provide additional skill, relative to the operational KNEP configuration (2 km / 15 min), at the spatial scales relevant to urban RWH control?

(iii) Under what conditions does nowcast skill reach a level sufficient to justify operational pre-storm releases in forecast-based control of RWH systems?

2 Theoretical background

"Wouldn't you like to know, weather boy?"

- Feral boy, Fox News

This chapter provides the theoretical background for forecast-based control of rainwater harvesting systems. It begins with the combined pressures of pluvial flooding, climate change, and urbanisation, motivating the limitations of conventional pipe-based drainage and the role of sustainable drainage systems. Rainwater harvesting is then examined as both a water resource and a controllable storage element. The principles of real-time and forecast-based control are reviewed, followed by precipitation forecasting at lead times relevant to building-scale storage, including numerical weather prediction, observation-based nowcasting, and current approaches in probabilistic ensemble methods, deep learning, and blending with numerical weather model output. The open-source Python library `PySTEPS`, which forms the forecasting toolbox of this thesis, is introduced throughout the chapter.

2.1 Urban Stormwater Challenges

Climate change is intensifying precipitation extremes in southern Sweden (An et al. 2026; Myhre et al. 2019). Observations indicate that average annual precipitation has increased from approximately 600 to 700 mm year^{-1} since 1930 (Schimanke et al. 2022) (Fig. 2.1), with projected increases of 8–12% by the end of the century (An et al. 2026). At the same time, drought frequency and associated water stress are expected to rise (IPCC 2023; Canedo Rosso et al. 2025).

For urban flood risk, the most disruptive rainfall is often tied to short events, when intense convective storms deliver large volumes over small areas in the span of minutes to a few hours. The 2014 cloudburst in Malmö, which caused damages of about 60 MEUR, and the even more severe cloudburst that hit Copenhagen three years earlier are clear examples of how quickly such storms can overwhelm the urban drainage system (UDS) and lead to extensive flooding (Olsson et al. 2021). Together, these events illustrate how exposed cities are to short-duration rainfall extremes, whose strong spatial and temporal variability challenge both standard design practice and day-to-day operational management, especially in small urban catchments with fast runoff responses

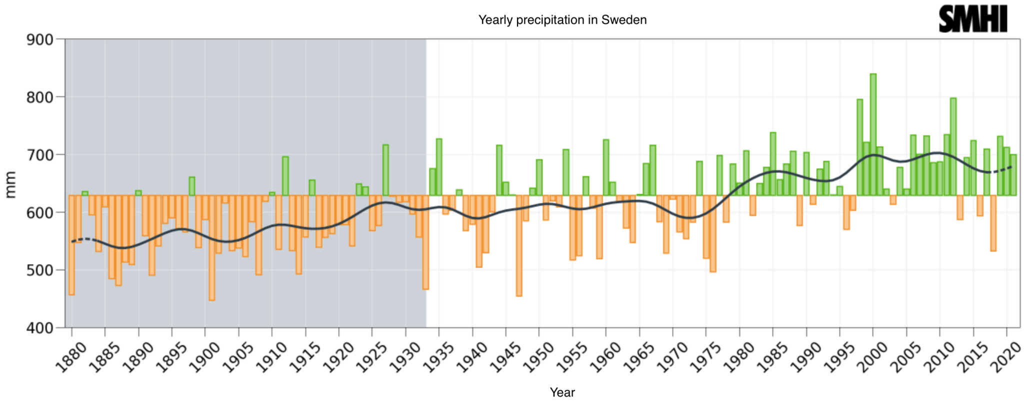

Figure 2.1: Annual accumulated precipitation. Green shows years with above-average and orange shows years with below-average precipitation relative to the 1961-1990 normal. The grey line represents a running mean calculated over ten years. Observations from before 1933 are deemed to have a lower accuracy than later observations - see the grey area in the chart. Adapted from (Schimanke et al. 2022).

(Torelló-Sentelles et al. 2025).

In parallel with climate change, global urbanisation has accelerated over the last century. While 20% of the world's population lived in cities in 1950, urban areas now host 45% of the world's 8.2 billion people, and most future population growth is projected to occur in cities (United Nations 2025). In Sweden, this growth is concentrated in metropolitan areas close to Stockholm, Gothenburg, and Malmö, and their commuter belts (Cedergren et al. 2025). Urbanisation alters catchment hydrology by reducing pervious surfaces, increasing peak flows, and shortening response times (Delleur 1982; Sillanpää et al. 2015).

The spatial expansion of impervious areas and densification further increase flood exposure (Feng et al. 2021; Tang et al. 2025). As a result, urban water systems increasingly need to manage both flood peaks during intense rainfall and water scarcity during dry periods. However, many Swedish stormwater systems remain dimensioned for historical climate conditions (Svenskt Vatten 2023), contributing to pluvial flooding, combined sewer overflows, and pollutant discharges (Sveriges miljömål 2022).

2.2 Rainwater Harvesting

Among decentralised source-control measures for urban stormwater management, sustainable drainage systems (SuDS) have gained increasing attention (Raimondi et al. 2023). SuDS aim to manage runoff close to its origin through infiltration, attenuation, and passive treatment, and include measures such as green roofs, swales, ponds, and rainwater harvesting (RWH) systems (Gimenez-Maranges et al. 2020).

Rainwater harvesting involves collecting and storing roof runoff for on-site use, typically to supply non-potable demands such as toilet flushing, laundry, and irrigation (e.g. for urban hydroponics (Sucozhañay et al. 2024)). By substituting potable water, RWH can reduce drinking-water demand and associated pressures on centralised supply systems (Campisano et al. 2017). Although sometimes framed as a modern technology, RWH has been applied for centuries to millennia in different forms, primarily for irrigation in water-scarce regions (Beckers et al. 2013; Bruins et al. 1986).

In Scandinavia, implemented RWH systems remain relatively uncommon. High freshwater availability and well-developed centralised water services have historically reduced the perceived need for decentralised reuse (Ivarsson et al. 2025), a situation reinforced in Sweden by unclear and fragmented legislative frameworks that lack clear definitions of permissible non-potable water reuse (Møller Jess 2025). However, recent summer droughts and groundwater stress - notably in Denmark (International Water Association 2022) and in southern Sweden (Canedo Rosso et al. 2025; An et al. 2026) - together with climate-adaptation ambitions, have increased interest in RWH as a complement to municipal services.

A recent assessment by Ivarsson et al. (2025) provides a systematic overview of implemented Scandinavian rainwater and stormwater reuse systems. They evaluated nine installations across Denmark, Sweden, and Norway using stakeholder interviews and a ranked-criteria sustainability analysis. The systems range from pilot to district scale and supply end-uses such as toilet flushing, laundry and irrigation. One key conclusion is that system performance is strongly design-dependent as simpler configurations with limited treatment generally achieved higher reliability and lower energy use than more complex systems with multiple treatment barriers.

RWH systems are commonly sized using water-balance simulations, where storage volume strongly influences performance but exhibits diminishing returns beyond moderate tank sizes (Mun et al. 2012). Sizing criteria involve trade-offs: designing for maximum rainwater use can lead to oversized systems, while targeting an intermediate efficiency can provide more balanced cost-performance outcomes. For example, an efficiency level of approximately 80% - defined as the fraction of non-potable demand supplied by harvested rainwater - has been suggested as a reasonable compromise between performance and cost in domestic applications (Santos et al. 2013). However, much of the sizing literature remains static, relying on historical rainfall and simplified demand assumptions. Such approaches do not explicitly account for climatic non-stationarity or for adaptive operation using forecast-informed control, where storage is actively managed in anticipation of rainfall events (Semaan et al. 2020).

In practice, this means that a passively operated RWH system will tend to either sacrifice water-supply performance by releasing stored water unnecessarily, or sacrifice detention capacity by retaining water ahead of intense rainfall. Neither outcome is avoidable through sizing adjustments alone. Real-time control, and in particular forecast-based control, offers a path to resolving this trade-off by enabling anticipatory storage management - a concept explored in the following section.

2.3 Real-Time Control of RWH Systems

Urban flooding and water-quality impacts have traditionally been addressed through the construction of additional storage capacity. However, such infrastructure expansions are often costly and offer limited flexibility under changing conditions (Sweetapple et al. 2026). An alternative approach is to increase the utilisation of existing storage by retrofitting it with sensors, communication systems, and actuators, enabling operational optimisation through smart control (Kerkez et al. 2016).

Real-time control (RTC) refers to the active regulation of hydraulic states, such as water levels and flows, through measurements and automated decision-making used to operate pumps, valves, and other control elements. Control actions can rely on observed system states and downstream constraints (Mullapudi et al. 2018), but can also incorporate rainfall forecasts to enable anticipatory operation. Such strategies are commonly referred to as forecast-based control (FBC) (Kerkez et al. 2016).

For rainwater harvesting (RWH) systems, the relevant spatial scale is typically individual buildings or small building clusters, where storage volumes respond rapidly to rainfall inputs and can reach capacity within a short period during high-intensity events. This creates a clear operational trade-off: maintaining stored water maximises water-supply benefits, whereas releasing water prior to rainfall maximises detention capacity and flood mitigation. Forecast-based control addresses this trade-off by triggering pre-event releases when rainfall is predicted, thereby creating available storage capacity in the critical lead-up to a storm (Kerkez et al. 2016).

A growing body of studies demonstrates the potential of integrating rainfall forecasts into RTC of decentralised stormwater storage systems (Table 2.1). A Swedish case study from Gothenburg showed that incorporating rainfall forecasts into the control of a retention basin could reduce overflow volumes by up to 50% (Fig. 2.2) compared with static rule-based operation, while still maintaining a substantial share of non-potable water reuse (Knutsson 2025). Similar concepts have been explored for decentralised storage systems connected to roofs. For example, distributed smart rain barrels equipped with IoT-based monitoring and control have been shown to reduce flooding and combined sewer overflows when operated using forecast information (Oberascher et al. 2021). At the household scale, studies on smart RWH tanks demonstrate that anticipatory releases triggered by rainfall forecasts can improve the hydraulic performance of UDSs and reduce peak flows (Xu et al. 2020a; Xu et al. 2022). Together, these studies bring forth the central promise of RTC-enabled RWH systems: by actively managing storage prior to rainfall events, it is possible to enhance flood mitigation while still preserving water-supply functionality.

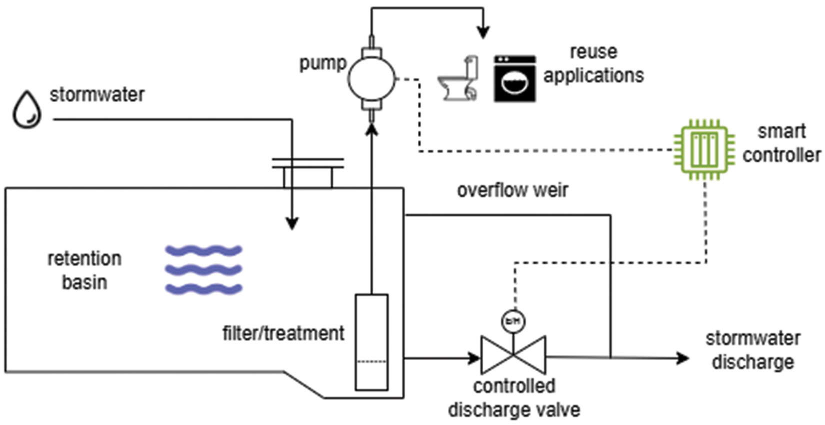

Figure 2.2: Graphical abstract showing the smart tank setup proposed by Knutsson (2025)

Several studies further suggest that extending the forecast horizon can improve control performance. For instance, Xu et al. (2020a) demonstrate that forecast windows of up to seven days allow earlier and more gradual pre-event releases from RWH tanks, thereby reducing the risk of overflow during subsequent rainfall events. At larger spatial scales, predictive control strategies applied to distributed stormwater storage or detention infrastructure have also shown improved system performance when rainfall forecasts are incorporated into optimisation frameworks (Shishegar et al. 2021; Sun et al. 2023). Recent research has also explored predictive control of roof-based storage systems, including smart internal drainage roofs operated using short-term rainfall predictions (Zhou et al. 2024).

In practice, however, most existing studies rely either on coarse-resolution numerical weather prediction products with daily (24 h+) temporal resolution or on simplified assumptions of perfect rainfall forecasts (Xu et al. 2020a; Sun et al. 2023; Zhou et al. 2024). Thus, the operational value of high-resolution rainfall information in the immediate hours before intense rainfall events remains largely unexplored. This limitation is particularly relevant for convective summer storms, which can develop rapidly and produce highly localised rainfall extremes.

Study	Asset / scale	Forecast input/horizon	Notes
Smart retention basin(Knutsson 2025)	Sweden, residential property scale	Prediction: hourly rainfall forecast; horizon not clearly specified	Forecast-based control of a retention basin. Demonstrated substantial overflow reduction while maintaining water reuse performance.
Smart rain barrels(Oberascher et al. 2021)	Austria, 15.2 ha	Prediction: 1, 4 or 12 h forecast horizon; hourly update	IoT-enabled distributed rain barrels (200–500 L) with forecast-based rule control. Micro-storage network reduces flooding and CSO volumes.
Smart tanks for RWH(Xu et al. 2020b)	Australia, household scale	Prediction: 1 d and 7 d NWP forecast horizons; 24 h update; control: 6 min	Household-scale RTC strategies for RWH tanks using NWP rainfall forecasts. Demonstrates that a 7-day forecast window substantially improves flood mitigation and delivers an outflow regime closer to natural streamflow compared to 1-day forecasts.
Controlled detention basins(Shishegar et al. 2021)	Canada, 311 ha	Prediction: up to 48 h; control: 5 min	System-level predictive optimisation of multiple detention basins. Coordinates spatially distributed basins to reduce peak flows, improve water quality through extended detention, and reduce receiving water erosion. Demonstrates model-based control strategies at UDS scale.
Smart storage tank(Sun et al. 2023)	China, 1.14 ha	Prediction: $\sim$2 h (24 $\times$ 5 min); control: 5 min	Predictive optimisation of stormwater storage tanks. Assumes perfect rainfall forecasts when evaluating performance improvements.
Smart internal drainage roofs (SIDR)(Zhou et al. 2024)	China, 1.3 km^2	Prediction: radar nowcasting (90 min); control: 5 min	Controlled roof drainage system for pluvial flood mitigation. Combines rule-based and predictive algorithms but assumes perfect rainfall input.

Table 2.1: Overview of studies applying RTC and FBC to decentralised storage systems and related infrastructure.

2.4 Rainfall Forecasting for FBC

Pre-storm release strategies depend on precipitation information that is both timely and actionable, as control actions are to be implemented before rainfall onset. At building or neighbourhood scale, effectiveness is primarily governed by forecast skill in the immediate hours ahead (Zhou et al. 2024). Forecast quality matters through its consequences for decisions: a missed event can cause avoidable overflows, while a false alarm triggers unnecessary releases that reduce water-supply benefits or saturate the UDS.

This section reviews precipitation forecasting methods relevant to forecast-based control of rainwater harvesting systems. It begins by situating nowcasting within the broader landscape of forecasting scales - spatial and temporal - before examining the two dominant paradigms: numerical weather prediction (NWP) and observation-based nowcasting. The relative strengths and limitations of each are discussed, with particular attention to the sub-2-hour decision horizon central to this thesis.

2.4.1 Temporal and Spatial Resolution

Forecasting models are commonly divided according to the spatial scale they represent. Global models simulate atmospheric processes at the planetary scale, mesoscale models resolve regional weather systems, and microscale approaches focus on local processes at the scale of cities or neighbourhoods (Guerova et al. 2022).

Precipitation forecasting is then further temporally divided into nowcasting, short-range, medium-range, and long-range forecasting. These categories are not strictly fixed in time, and their boundaries vary between operational definitions and research contexts.

The WMO Working Group on Nowcasting Research defines nowcasting 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" (World Meteorological Organization 2017). The term was originally coined by Browning (1981) in a narrower sense, referring to the detailed description of current weather and the prediction of changes over a few hours. Within the 0–6 h window, a further distinction is commonly made between nowcasting proper (0–2 h), where observation-driven methods dominate, and very short-range forecasting (VSRF, 2–12 h), where numerical weather prediction becomes increasingly relevant (Prudden et al. 2020). This distinction reflects practical differences in predictability and methodology. Forecasts within 0–2 h are typically dominated by observation-driven approaches, whereas forecasts beyond this range increasingly rely on numerical weather prediction (NWP) (Golding 1998).

In this thesis, the term *nowcasting* primarily refers to short-term, observation-driven precipitation forecasting within the 0–2 h range, while acknowledging the broader operational context extending beyond that.

2.4.2 Numerical Weather Prediction

Numerical weather prediction (NWP) forecasts precipitation by numerically solving the governing equations of atmospheric motion and thermodynamics on a discrete computational grid (Guerova et al. 2022). Forecast skill depends on the quality of the initial atmospheric state (typically obtained through data assimilation), model resolution and physical parameterisations, and the intrinsic predictability of the underlying processes (Pu et al. 2018).

NWP has advanced substantially since Richardson's 1922 estimate that a forecasting "factory" would require approximately 64,000 mathematicians to numerically solve the governing equations of atmospheric motion (Guerova et al. 2022; Lynch 2015). What began as a theoretical and computationally impractical idea became operational with the advent of electronic computers, marked by the first successful computer-based forecast in 1950. Since then, advances in computing, atmospheric observations, data assimilation, and model physics have transformed NWP into the foundation of modern weather forecasting (Pu et al. 2018), with forecasting improving across all timescales for the northern and southern hemisphere alike (Bauer et al. 2015)

In Sweden, operational short-range forecasting is based on the convection-permitting HARMONIE-AROME configuration within the ALADIN-HIRLAM NWP system, with a horizontal resolution of approximately 2.5 × 2.5 km and an update frequency of six hours (Bengtsson et al. 2017; SMHI 2025b). For the decision horizons relevant to RWH control - on the order of minutes to a few hours - this update frequency means that the most recent model initialisation may already be several hours old at the moment a control decision must be made. Beyond the update cycle, NWP forecasts at short lead times are further constrained by intrinsic predictability limits at convective, often sub-kilometre scales: small uncertainties in initial conditions can grow rapidly, and processes relevant for intense rainfall such as convective initiation are difficult to represent at operational grid spacings (Yano et al. 2018). Forecasts can also be degraded immediately after initialisation due to model spin-up (Cao et al. 2016), which is particularly problematic when the decision horizon is on the order of tens of minutes (Prudden et al. 2020).

2.4.3 Extrapolation-Based Nowcasting

Extrapolation-based nowcasting emerged during the late 1970s and early 1980s in an era when numerical weather prediction primarily focused on synoptic and mesoscale atmospheric processes (Prudden et al. 2020; Pierce et al. 2012). Rainfall forecasting during this period posed difficulties for NWP models because computational limitations constrained their spatial resolution. While NWP models were capable of representing mesoscale atmospheric features such as frontal systems, they were unable to resolve smaller-scale convective processes embedded within these systems (Prudden et al. 2020).

Extrapolation-based nowcasting addresses short lead times by extrapolating observed precipitation (e.g. radar or satellite) fields forward in time. Successive radar composites are used to estimate precipitation motion, after which precipitation structures are advected to generate short-term forecasts. The process rests on the assumption that precipitation structures move with a stationary motion field over short time intervals, an assumption commonly referred to as *Lagrangian persistence* (Zawadzki et al. 1994). Under this assumption, radar echoes are advected forward in time without explicitly modelling atmospheric dynamics. The approach enables forecasts with high spatial and temporal resolution (Golding 1998; Pulkkinen et al. 2019).

Denoting the precipitation field by Ψ and the displacement vector by $\alpha(\tau)$, the conservation of precipitation along the motion trajectory can be written as

$$\Psi(x_0, t + \tau) = \Psi(x_0 - \alpha(\tau), t). \tag{2.1}$$

In differential form this relationship can be expressed as the advection equation

$$\frac{\partial \Psi}{\partial t} + u\frac{\partial \Psi}{\partial x} + v\frac{\partial \Psi}{\partial y} = 0, \tag{2.2}$$

where (u, v) denote the horizontal components of the motion field.

The advection equation can be reframed as an optical flow constraint - a technique originating in computer vision (Horn et al. 1981) - by identifying radar reflectivity with image intensity. Under this interpretation, the condition that reflectivity remains approximately constant along precipitation trajectories is expressed as

$$\frac{\partial E}{\partial t} + u\frac{\partial E}{\partial x} + v\frac{\partial E}{\partial y} = 0 \tag{2.3}$$

where $E(x, y, t)$ denotes radar reflectivity at position (x, y) and time t, and (u, v) are the horizontal motion components. Estimating this motion field from sequential radar images constitutes one of the central challenges in extrapolation-based nowcasting (Prudden et al. 2020).

However, this equation alone is insufficient to uniquely determine the motion field at each location, a limitation known as the aperture problem. At any given point, the optical-flow constraint provides a single equation relating the two unknown motion components u and v, meaning the system is fundamentally underdetermined from local information alone. Intuitively, a local change in reflectivity is consistent with many different combinations of horizontal and vertical motion, and without additional information there is no way to distinguish between them (Horn et al. 1981). Practical optical-flow algorithms therefore introduce additional assumptions, most commonly spatial smoothness of the motion field, guaranteeing a unique solution for each (u, v) and allowing a motion field to be estimated from sequential radar images (Horn et al. 1981; Pulkkinen et al. 2019).

In the `PySTEPS` framework, several motion estimation methods are available, including the local Lucas–Kanade method, a global variational echo-tracking approach, and the spectral DARTS method (Pulkkinen et al. 2019).

Once the motion field has been estimated, the precipitation field can be advected forward in time using numerical advection schemes.

Despite their high spatial and temporal resolution, radar-based nowcasts are inherently limited by growth and decay of precipitation, displacement errors, and rapidly increasing uncertainty with lead time; forecast skill typically decreases beyond one to a few hours, particularly during convective events (Imhoff et al. 2020; Prudden et al. 2020). The relationship between observation-based nowcasting methods and NWP is well illustrated by Golding (1998) (Fig. 2.3):

“The downward trend in information content of the theoretical limit of predictability reflects the fact that in a chaotic system like the atmosphere, there is an inevitable loss of information as one looks further into the future. The potential information

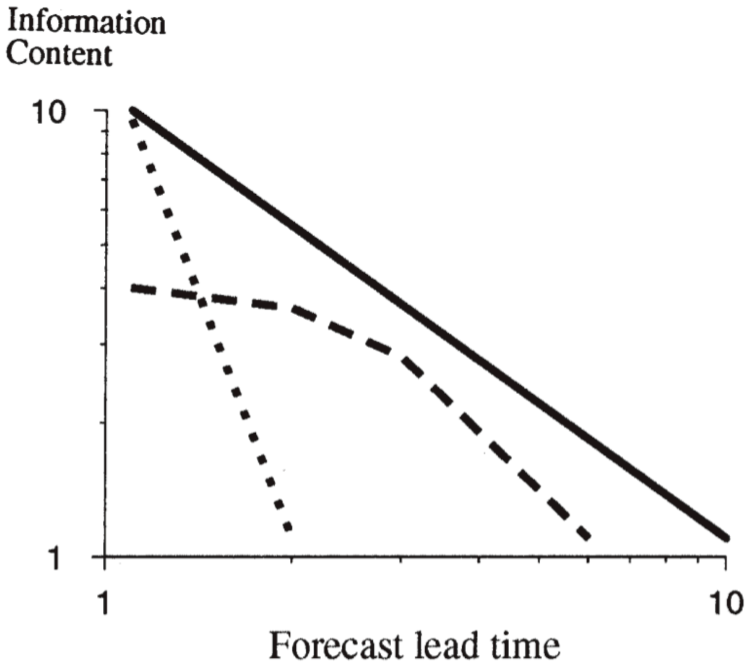

Figure 2.3: Forecast information loss as a function of lead time. The solid line represents the theoretical limit of predictability, the dotted line nowcasting methods, and the dashed line NWP models. Adapted from Golding (1998).

content will usually relate to larger scales as the forecast range becomes longer. For a nowcasting method, the initial information capture is 'perfect', or nearly so, but without representation of atmospheric physics, the information loss is very rapid with forecast time. By contrast, limited resolution and imperfect assimilation algorithms result in a relatively poor representation of the observed state in NWP models. This, however, deteriorates only slowly in the early part of the forecast owing to the good representation of larger scales, and then parallels the 'perfect' solution, the difference reflecting the imperfect representation of physics, and the residual effect of distortion of small scales in the initial representation of the atmosphere."

2.5 State of the Art in Radar-Based Nowcasting

This section reviews the state of the art in radar-based precipitation nowcasting. It begins with probabilistic ensemble methods, which generate multiple plausible forecast scenarios rather than a single prediction, before examining deep learning approaches and their current operational limitations. Blending strategies that combine nowcasting and NWP-VSRF are then discussed, as these extend the useful forecast horizon beyond what radar extrapolation alone can achieve.

2.5.1 Probabilistic Nowcasting

Since the mechanisms governing rainfall generation are inherently chaotic and non-linear, a probabilistic approach is preferable: deterministic model errors grow very quickly, particularly at small scales when convective events occur (De Luca et al. 2025; Hohenegger et al. 2007). Probabilistic nowcasting addresses this by generating ensembles of plausible future precipitation (see Fig 2.4), allowing uncertainty to be represented explicitly and enabling end users to make better-founded decisions (Krzysztofowicz 2001).

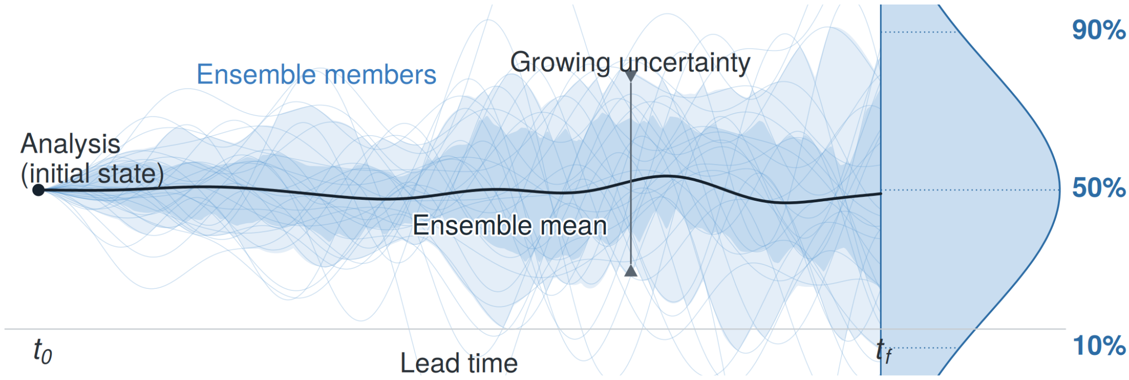

Figure 2.4: Schematic illustration of ensemble nowcasting. Starting from a single analysis state at t_0, an ensemble of N members evolves along plausible trajectories, with spread growing as small-scale features decorrelate with lead time. The ensemble mean (dark line) provides a deterministic summary, while the full ensemble collapses into a forecast distribution at the final lead time t_f, enabling probabilistic decision-making.

The STEPS framework (Short-Term Ensemble Prediction System) developed by Bowler et al. (2006) and further improved upon by Seed et al. (2013) formalises this through a cascade decomposition (Fig. 2.5): the precipitation field is decomposed into spatial scales using a bandpass filter, and random noise calibrated to observed forecast error growth at each scale is used to generate ensemble members. As small-scale features decorrelate faster than large-scale ones, STEPS produce ensembles whose spread widens realistically with lead time (Bowler et al. 2006; Seed et al. 2013). `PySTEPS` implements STEPS as its primary probabilistic module and has become a community standard for ensemble nowcasting research (Pulkkinen et al. 2019; Imhoff et al. 2020).

The more recent LINDA framework (Lagrangian INtegro-Difference equation model with Autoregression, Pulkkinen et al. (2021)) replaces the cascade decomposition with a physically motivated Lagrangian framework that better preserves small-scale precipitation structures. LINDA has shown skill advantages over STEPS for intense, localised convective cells, precisely the events most likely to overwhelm building-scale storage.

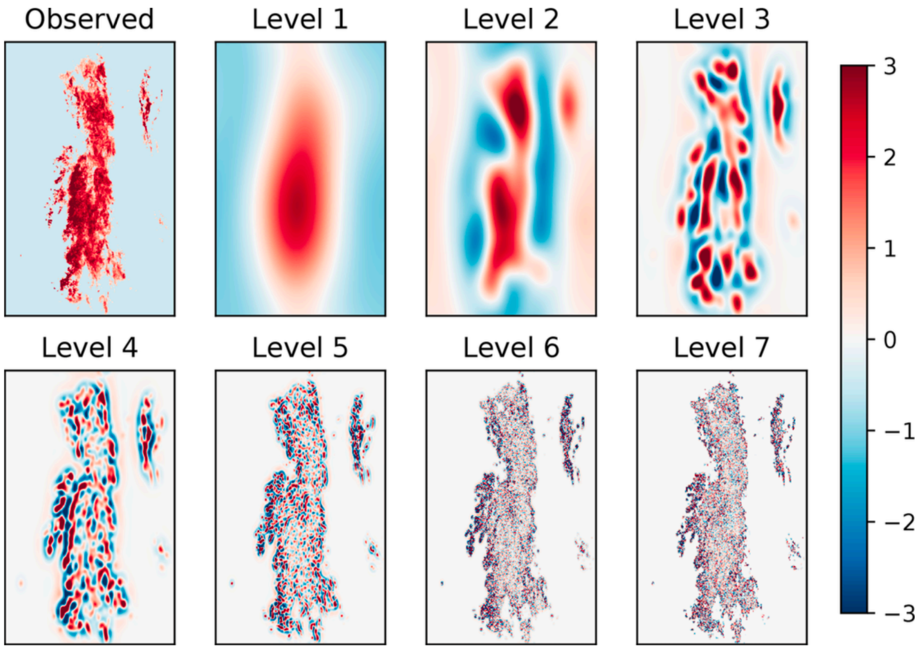

Figure 2.5: 7 level Spectral cascade following decomposition of input signal thorugh a bandpass filter. Reprinted from Pulkkinen et al. (2019).

This comes, however, at a greater computational cost compared to STEPS (Imhoff et al. 2020; Tounsi et al. 2023)

For FBC applications, the value of ensemble nowcasting lies not in improved point accuracy but in enabling probabilistic decision rules. A pre-storm release can be triggered when the ensemble probability of exceeding a rainfall threshold surpasses an defined threat level, trading the cost of unnecessary releases against the risk of overflow (Courdent et al. 2015). This probabilistic threshold-based approach has been shown to substantially reduce missed warnings at the cost of only a moderate increase in false alarms, representing a favourable trade-off compared to deterministic forecasts (Biondi et al. 2021).

2.5.2 Deep Learning Nowcasting

A growing body of research has applied deep learning, a specialized form of machine learning that uses several processing layers to learn complex abstractions (Misra et al. 2020), to precipitation nowcasting. These approaches learn to recognise patterns directly from large archives of historical radar observations, and have shown competitive performance on benchmarking datasets (Li et al. 2026). De Luca et al. (2025) states that deep learning is particularly well suited to this task for three reasons: (i) the sheer volume of data collected daily by radar and ground stations provides the large training datasets these methods require; (ii) modern graphics hardware allows predictions to be generated far faster than traditional physically-based models; and (iii) deep neural networks are capable of approximating highly complex, non-linear relationships, such as the chaotic ways in which rainfall fields evolve over time and space - a process known as parametrisation.

Precipitation nowcasting can be approached as one of three related pattern-recognition problems: predicting a future rainfall image from past ones, modelling how a sequence of observations evolves forward in time, or generating plausible future scenarios - each of which has inspired distinct families of model architecture (Li et al. 2026). Landmark examples include DGMR (Ravuri et al. 2021), a generative model trained on radar data from the United Kingdom that produces probabilistic nowcasts up to 90 minutes ahead, and NowcastNet (Zhang et al. 2023), which incorporates physical conservation principles to improve forecast skill for extreme precipitation events.

From an operational standpoint, however, deep learning nowcasting remains at an earlier stage of maturity than ensemble extrapolation methods. Models are typically trained on radar data from a specific region and may perform poorly when applied elsewhere, as precipitation patterns differ substantially across climates (Li et al. 2026; De Luca et al. 2025). The lack of shared, standardised datasets for evaluating models across different regions also makes it difficult to compare results between studies (Li et al. 2026). A deeper limitation is that deep learning models learn patterns from historical data without encoding the physical rules that govern how weather actually behaves. As a result, they can produce forecasts that are statistically plausible but physically implausible when confronted with conditions not well represented in their training data (De Luca et al. 2025; Li et al. 2026). A further practical problem concerns forecast sharpness: models trained to minimise average error tend to *"hedge"* their predictions, spreading rainfall over too large an area and underestimating peak intensities, while models designed to generate realistic-looking fields can occasionally produce rainfall in locations where none is warranted (Li et al. 2026; Ravuri et al. 2021). Both failure modes are problematic for control applications that depend on crossing a rainfall threshold to trigger action.

2.5.3 Blending with Numerical Weather Prediction

Neither pure radar extrapolation nor numerical weather prediction alone is optimal across all lead times. As described in Section 2.4.3, extrapolation-based methods are to be preferred at sub-hourly to two-hour lead times, while numerical weather prediction skill increasingly exceeds that of extrapolation beyond this range. Blending strategies exploit this complementarity by weighting contributions from each source as a function of lead time (Golding 1998; Seed et al. 2013; De Luca et al. 2025). One of the earliest examples of this approach is NIMROD, developed at the UK Met Office, which combined radar extrapolation with numerical weather prediction output to extend the useful forecast horizon (Golding 1998).

The INCA (Integrated Nowcasting through Comprehensive Analysis) system takes a similar but more structured approach: radar extrapolation is used exclusively for the first two hours, numerical weather prediction output is introduced progressively from two to six hours, and beyond that the numerical model takes over entirely (Haiden et al. 2011).

SMHI's KNEP used to follow the same hybrid logic as INCA: radar-based extrapolation is used for the shortest lead times, while numerical weather prediction is introduced progressively as forecast range increases. KNEP would advect the latest radar reflectivity field for up to six hours using an AROME wind field and then blends the result toward the AROME precipitation forecast with lead-time-dependent weights. The result would be an operational product that combines the strengths of extrapolation at very short lead times and NWP at longer lead times (SMHI 2025a). However, at date, KNEP is only using NWP for NaN-gap filling (Falahat 2026).

Operationally, blending ranges from simple weighted averaging to more sophisticated approaches. Simple linear blending can underperform both constituent forecasts during the transition period, as averaging two spatially displaced fields systematically dampens precipitation intensities (Hwang et al. 2015). More advanced alternatives such as saliency-based blending (Hwang et al. 2015) and methods based on ensemble Kalman filtering (Nerini et al. 2019) have been proposed to address this. `PySTEPS` implements several of these methods, with numerical weather prediction contributions progressively replacing extrapolation as lead time increases (Pulkkinen et al. 2019). Deep learning has also been applied within blended frameworks, both to post-process numerical weather prediction output and to embed physical constraints directly into data-driven architectures - an approach that shows particular promise for improving forecasts of heavy rainfall events (De Luca et al. 2025; Li et al. 2026).

3 Methods

This chapter is organised as follows. Section 3.1 describes the literature-review approach; Section 3.2 the gauge and radar data and the event-selection procedure that reduces the candidate set to the 19 analysed events; Section 3.3 the six nowcasting configurations evaluated; and Section 3.4 the deterministic and probabilistic verification metrics. The experiments (Section 3.5) proceed in three stages. A one-factor-at-a-time *sensitivity analysis* first fixes the model parameters used thereafter. The *full runs* then apply the five deterministic configurations to every event and initialisation across three input resolutions, and form the basis of the results for RQ(i) and RQ(ii). Finally, a smaller set of *exploratory probabilistic runs* gives an initial, deliberately limited assessment of ensemble nowcasting for FBC, informing RQ(iii); as these use a single fixed initialisation, they are not directly comparable to the full runs and are treated as indicative throughout.

3.1 Literature review methodology

The literature review was conducted to establish the scientific context for short-term precipitation forecasting, with a particular focus on radar-based nowcasting methods and their applicability to urban rain and stormwater management. The review was designed as a targeted, problem-driven literature study rather than a fully systematic review.

Relevant literature was identified through structured searches in major scientific databases, including Elsevier's *Scopus*, *Google Scholar*, and Lund University's *Finn*. Search queries combined keywords related to precipitation forecasting and nowcasting (e.g., *precipitation nowcasting*, *radar-based forecasting*, *optical flow*, *ensemble nowcasting*, *PySTEPS*) with terms associated with urban hydrology and decision support (e.g., *urban stormwater*, *real-time control*, *rainwater harvesting*). Backward and forward citation tracking was used to identify additional influential studies.

The literature selection focused on peer-reviewed journal articles, authoritative review papers, and widely cited methodological references. Particular emphasis was placed on studies that (i) describe the theoretical foundations of radar-based nowcasting, (ii) evaluate extrapolation-based methods, (iii) discuss ensemble approaches and uncertainty representation, and (iv) assess the strengths and limitations of nowcasting at short lead times. Foundational hydrological and meteorological references were included to support conceptual definitions and event-based analysis practices.

Studies were excluded if they primarily addressed medium- or long-range numerical weather prediction without relevance to short lead times, relied on data sources not comparable to weather radar, or focused on application domains outside the scope of urban precipitation-driven decision-making.

3.2 Data

The data used in this thesis fall into two main categories: gauge data for event selection, and radar data for nowcasting.

3.2.1 Gauge Data

Rainfall events were identified using tipping-bucket rain gauge data provided by VA Syd, the regional water and wastewater utility in the Malmö area. Data were obtained from three stations in the Malmö urban area: Turbinen PST, Limhamn PST, and Augustenborg (Fig. 3.1). Each record consists of individual tip events with exact timestamps and a fixed tip volume of 0.2 mm, covering the period 2016–2025. Prior to event selection, the raw tip data were resampled to 15-minute accumulated totals.

Figure 3.1: Location of the three VA Syd tipping-bucket rain gauge stations used for event selection: Turbinen PST, Limhamn PST, and Augustenborg. All stations are situated within the urban area. A subset of events were also sourced from the SMHI MetObs API from the Malmö A station.

Station availability varies across the study period: Turbinen PST and Limhamn PST provide continuous data through 2025, while Augustenborg covers 2016–2023. The three stations are located within approximately 5–10 km of each other, and events were selected from the combined network with cross-station deduplication applied to avoid counting the same rainfall event multiple times.

3.2.1.1 Event Selection Based on Gauge Observations

Rainfall events were selected from tipping-bucket rain gauge data provided by VA Syd, covering three stations in the Malmö area: Turbinen PST (55.60°N, 12.98°E), Limhamn PST (55.58°N, 12.92°E), and Augustenborg (55.58°N, 13.03°E). The data consist of individual 0.2 mm tip records with local-time timestamps, resampled to 15-minute totals prior to analysis. The study period spans 2017-03-01 to 2025-12-31, with Augustenborg available through 2023 only.

Events were defined as rolling totals over three durations, $d \in \{1, 3, 6\}$ h, selected to span short convective bursts (1 h) through longer mixed and frontal events (6 h). Stratifying events by duration in this way is consistent with established nowcasting-evaluation practice (Olsson et al. 2013; Imhoff et al. 2020; Tounsi et al. 2023).

Each rolling window was assigned to a season based on its start time: Summer (April–September) or Winter (October–March), following the convention that sub-hourly extremes in Sweden occur almost exclusively during the summer convective season, while longer-duration extremes may occur in any season (An et al. 2026).

A candidate window was kept only if it met several conditions. A dry buffer of 2 hours was required before and after the window, with a threshold of 0.2 mm per 15-min interval and a maximum pre-event accumulation of 0.5 mm over the buffer period in order to have a clear event onset. The event had to reach a minimum 15-minute peak intensity of 1.0 mm per 15-min interval. A minimum time gap of 2 hours between selected events was enforced for each duration and season, ensuring temporal independence (Imhoff et al. 2020).

Event selection was performed independently for each station. Candidates were scored based on total accumulated precipitation and ranked within each duration-season slot. Within-station deduplication was applied per duration class; overlapping events of the same duration were resolved by retaining the more extreme one, analogous to the approach of Olsson et al. (2013). Station pools were then merged and cross-station deduplication was applied, again per duration class. From the deduplicated pool, the six most extreme events per duration–season combination were retained, yielding a final dataset of 36 events (3 durations × 2 seasons × 6 events).

Return periods were estimated for each event by inverting the empirical formula of Dahlström (2010) (eq. 6), which relates design rain intensity to return period and duration:

$$R_{ij} \approx 190\,\tau_i^{1/3}\frac{\ln \Delta t_j}{\Delta t_j^{0.98}} + 2 \tag{3.1}$$

where R_{ij} is the rain intensity ($\mathrm{l\,s^{-1}\,ha^{-1}}$), τ_i the return period in months, and Δt_j the event duration in minutes. The observed gauge accumulation was converted to intensity via $R = P\,/\,(\Delta t \cdot 0.36)$, where P is the accumulated depth (mm) and Δt the duration in hours. Solving eq. (3.1) for τ_i gives:

$$\tau_i = \left(\frac{R - 2}{190 \cdot \dfrac{\ln \Delta t_j}{\Delta t_j^{0.98}}} \right)^3 \tag{3.2}$$

Events with $R \leq 2\,\mathrm{l\,s^{-1}\,ha^{-1}}$ fall below the frontal-rain baseline of the formula and are assigned no return period.

All events were visually verified in the BALTRAD GUI to confirm correspondence between gauge and radar observations (BALTRAD 2026).

3.2.1.2 Overview of Selected Events

Following the event selection methodology described in Section 3.2.1.1, 36 precipitation events were initially identified and requested from SMHI's radar archive. After data retrieval and quality control, events were excluded if radar composite data was not available, contained substantial gaps in the event window, or relied on fallback to lower-resolution radar products during portions of the event. Three additional events from an earlier training set were retained to broaden the seasonal and duration coverage.[1] After these filters, 19 events remained suitable for analysis at native PySTEPS resolution (0.5 km / 5 min).

Table 3.1 lists the 19 events used in the analysis. The full list of 36 originally selected events, including those excluded due to data quality issues, is provided in Appendix A3.

Onset frames and hyetographs for all analyzed events are included in Appendix A3 (Fig. A3, A4, A5, and A2).

[1] These events were used at an earlier stage to construct the sensitivity analysis. They were identified using the same methodology as the VA Syd events, but were sourced from SMHI's MetObs API from station *Malmö A (52350)*.

Table 3.1: Precipitation events included in the analysis. Date and time refer to event onset (local time, Europe/Stockholm). Gauge totals and peak intensities are from the VA Syd gauge record. T is the return period estimated using Dahlström (2010). Type indicates the broad character of the event: C = convective, S = stratiform, M = mixed. Added events from the initial SMHI screening are shown in blue.

#	Date & onset (local)	Season	Type	Dur. (h)	Gauge (mm)	Peak ($\mathrm{mm\,h^{-1}}$)	T (yr)	Station
2	2022-07-03 21:05	Summer	C	1	12.8	24.8	1.1	Augustenborg
5	2021-09-28 00:45	Summer	M	1	10.0	32.8	0.5	Turbinen PST
38	2022-05-31 11:22	Summer	C	1	11.5	17.6	0.8	(Malmö A)
7	2022-10-01 16:35	Winter	M	1	14.4	21.6	1.6	Limhamn PST
9	2019-10-13 20:00	Winter	S	1	4.8	16.8	0.1	Turbinen PST
10	2018-11-12 01:05	Winter	M	1	4.6	7.2	0.1	Turbinen PST
12	2020-11-02 02:00	Winter	S	1	4.0	6.4	0.1	Turbinen PST
15	2018-08-09 23:50	Summer	S	3	21.4	35.2	2.1	Turbinen PST
39	2019-06-12 21:55	Summer	M	3	23.5	36.4	2.9	(Malmö A)
40	2023-07-12 06:37	Summer	S	3	32.2	26.0	8.0	(Malmö A)
19	2022-10-01 13:45	Winter	M	3	15.8	20.0	0.7	Turbinen PST
21	2017-10-02 09:05	Winter	S	3	11.4	6.4	0.2	Turbinen PST
22	2023-02-03 05:55	Winter	M	3	11.0	6.4	0.2	Augustenborg
24	2021-12-01 21:00	Winter	M	3	8.2	4.0	0.1	Turbinen PST
28	2019-07-28 07:05	Summer	C	6	22.4	32.0	1.1	Turbinen PST
29	2018-08-09 23:50	Summer	M	6	21.6	35.2	1.0	Turbinen PST
30	2021-09-29 15:35	Summer	M	6	20.8	8.8	0.9	Augustenborg
33	2021-10-21 07:25	Winter	S	6	19.2	10.4	0.6	Turbinen PST
36	2021-12-01 17:05	Winter	S	6	16.0	4.0	0.3	Turbinen PST

3.2.2 Radar Data

High-resolution radar composites were obtained directly from SMHI. These composites have a spatial resolution of 0.5×0.5 km and a temporal resolution of 5 minutes, covering the entirety of Sweden.

The high resolution is an upgrade from the 2×2 km, 15-minute resolution used in SMHI's operational nowcasting product KNEP (Korta NEderbörds Prognoser), originally requested by the Flight Meteorologists at SMHI in Uppsala for aviation use to better resolve convective precipitation (Nordh 2026). The underlying radar technology to produce data at higher resolution is not new, but the operational adoption has been constrained by both the substantial increase in data volume - the 0.5 km / 5-min product contains $16 \times 3 = 48$ times as many data points as the KNEP input - and the absence of explicit demand for finer resolution from internal users or external customers, which together have meant that the established 2 km / 15-min standard has remained in operational use (Van De Beek 2026).

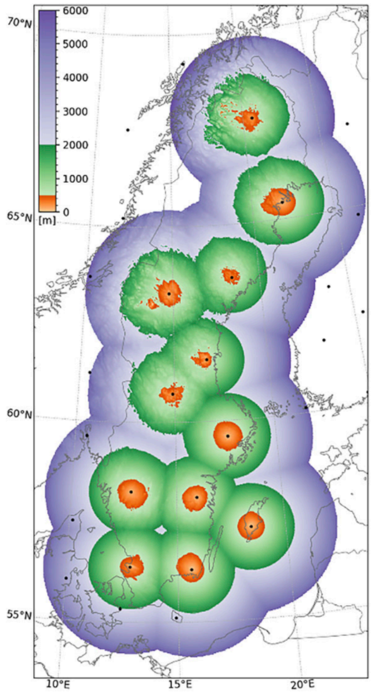

Figure 3.2: Weather radar coverage map of Sweden. The twelve weather radar stations are marked with black dots. The colour scales indicate the altitude ranges of radar measurements. Reproduced with permission from SMHI. Originally published in (Björck et al. 2025).

Radar of this resolution has to the author's extent of knowledge not been been used for nowcasting in the literature.

The Swedish national weather radar network consists of twelve C-band Doppler weather radar stations (Hosseini et al. 2022) operating at 5.6 GHz with a maximum range of 240-250 km per station (SMHI 2026). The network employs dual-polarization technology and is part of the BALTRAD collaboration, which merges scans from Sweden, Norway, Finland, Estonia, and Latvia to create composite coverage over nearly all of Northern Europe and the Baltic Sea (Michelson et al. 2018). (Fig. 3.2).

The composite product is generated by merging individual radar scans into a single nationwide field, with quality control applied to reduce artefacts such as ground clutter and beam blockage. For each selected event, radar data were extracted covering the event window with six hours of padding before and after.

Prior to data delivery, the national radar composites were cropped to a regional subdomain covering approximately 7.7–24.1°E and 53.8–62.9°N, encompassing southern Sweden, north Eastern Denmark and the surrounding sea. This reduced to the spatial extent to the area relevant for the Malmö study site while substantially decreasing data flow and subsequent computational load.

3.2.3 Radar reflectivity and rainfall rate

Weather radar measures the radar reflectivity factor Z, which represents the returned signal from hydrometeors (e.g. raindrops, snow or hail) within the radar sampling volume. Because Z spans several orders of magnitude, it is commonly expressed on a logarithmic scale as

$$dBZ = 10 \log_{10} \left(\frac{Z}{Z_0} \right), \tag{3.3}$$

where Z_0 is a reference reflectivity.

For hydrological applications, reflectivity must be converted to rainfall intensity R ($\mathrm{mm\,h^{-1}}$). This is typically done using the empirical Marshall-Palmer (Marshall et al. 1947) Z–R relationships of the form

$$Z = a \cdot R^b, \tag{3.4}$$

where a and b are empirically derived parameters. In this thesis the standard values $a = 200$ and $b = 1.6$ is used, in accordance with Andersson et al. (2022) and Marshall et al. (1948).

3.3 PySTEPS Framework

`PySTEPS` is an open-source Python library for short-term precipitation nowcasting in the 0–6 h range (Pulkkinen et al. 2019). It provides tools for motion field estimation using optical flow, radar-based extrapolation, stochastic ensemble generation, and optional blending with numerical weather prediction (Fig. 3.3). In this thesis, `PySTEPS` is used to generate deterministic and probabilistic radar-based nowcasts, evaluated against simple persistence baselines and the operational KNEP product.

3.3.1 Nowcasting configurations

`PySTEPS` is based on the assumption of Lagrangian persistence. Over short lead times, precipitation structures are assumed to move with a quasi-stationary velocity field, while their intensity remains approximately constant (Pulkkinen et al. 2019). The forecast is therefore generated by advecting the observed radar field forward in time using a semi-Lagrangian advection scheme.

Six nowcasting configurations are evaluated in this study, representing increasing levels of methodological complexity:

Eulerian Persistence (EP) uses the most recent radar composite as the forecast for all lead times, with no advection or dynamical processing. It serves as the minimum baseline against which all other methods are compared.

Lagrangian Persistence (LP) advects the most recent radar composite forward in time using the estimated motion field, but applies no cascade decomposition or stochastic perturbations. It isolates the contribution of motion estimation alone.

S-PROG (Spectral PROGnosis), developed by Seed (2003), extends LP by applying a multi-scale cascade decomposition and autoregressive (AR) modelling at each spatial scale, allowing small-scale features to dissipate more rapidly than large-scale ones. The result is a single deterministic nowcast field with a more physically plausible intensity evolution than LP.

ANVIL (Autoregressive Nowcasting using Vertically Integrated Liquid) method, developed by Pulkkinen et al. (2020), is a deterministic extension of the Lagrangian persistence paradigm designed to better capture the growth and decay of precipitation structures (Tounsi et al. 2023). While the method was originally designed for use with volumetric radar observations of Vertically Integrated Liquid (VIL), it is here applied to surface rainfall intensity (SRI) fields as implemented within the PySTEPS framework (Annella et al. 2025; Pulkkinen et al. 2021).

The primary innovation of ANVIL is the replacement of the standard second-order autoregressive AR(2) process - used in models like S-PROG - with an autoregressive integrated ARI process (Pulkkinen et al. 2021). This approach explicitly models the temporal evolution of the differences between successive precipitation fields, which allows the model to preserve small-scale variability and sharp gradients that are otherwise smoothed out by the scale-filtering processes in S-PROG (Annella et al. 2025; Pulkkinen et al. 2021). Furthermore, ANVIL utilizes a spatially localized parameter estimation scheme, enabling it to effectively predict regional growth and dissipation patterns rather than relying on domain-wide averages (Pulkkinen et al. 2021).

STEPS (Short-Term Ensemble Prediction System) builds on S-PROG by adding

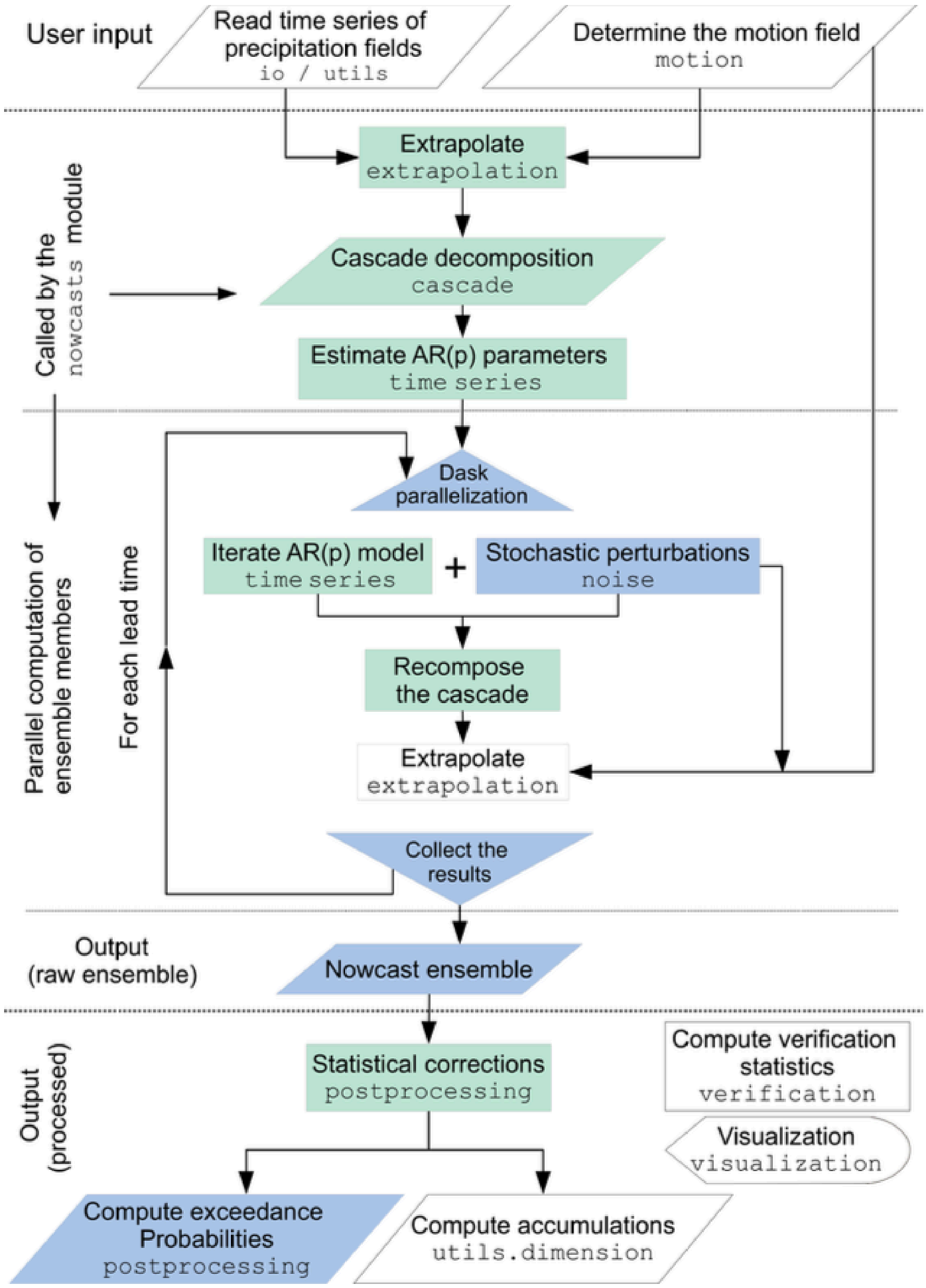

Figure 3.3: Workflow for computing probabilistic precipitation nowcasts using `PySTEPS`. Reprinted from (Pulkkinen et al. 2019).

stochastic perturbations calibrated to observed forecast error growth at each cascade level, generating an ensemble of plausible future precipitation fields (Bowler et al. 2006; Seed et al. 2013).

LINDA As an exploratory comparison, nowcasts were also generated using the LINDA framework (Lagrangian INtegro-Difference equation model with Autoregression) (Pulkkinen et al. 2021). LINDA replaces the cascade decomposition of STEPS with a physically motivated Lagrangian framework that explicitly models the growth and decay of individual precipitation objects. Rather than decomposing the rainfall field into spatial scales and applying AR models at each level, LINDA tracks the evolution of precipitation structures directly in Lagrangian coordinates.

LINDA has shown skill advantages over STEPS for intense, localised convective cells (Pulkkinen et al. 2021; Imhoff et al. 2020; Tounsi et al. 2023; Annella et al. 2025) - precisely the event type most likely to challenge extrapolation-based nowcasting and most relevant to the FBC context of this study.

In all configurations, motion fields were estimated using the Lucas-Kanade (LK) optical-flow method, applied to the four most recent radar frames prior to each nowcast initialisation. LK was selected as the simplest and computationally lightest optical-flow scheme available in `PySTEPS`. This choice is supported by Pulkkinen et al. (2019), who report that the alternative schemes (VET and DARTS) differ by less than 2% in skill.

Nowcasts were produced at the native radar resolution of 0.5×0.5 km with a temporal resolution of 5 minutes and a lead time of 90 minutes. An overview of the four configurations is given in Table 3.2.

Table 3.2: Nowcasting configurations evaluated in this study, ordered by increasing model complexity. Columns indicate whether each method applies Lagrangian advection of the precipitation field, a multi-scale spectral cascade with autoregressive temporal evolution (AR), and stochastic perturbations to generate an ensemble. Methods: EP - Eulerian Persistence; LP - Lagrangian Persistence; S-PROG - Spectral Prognosis; ANVIL - Autoregressive Nowcasting of Vertically Integrated Liquid; STEPS - Short-Term Ensemble Prediction System; LINDA - Lagrangian Integro-Difference equation model with Autoregression.

Method	**Advection**	**Cascade + AR**	**Ensemble**
EP	No	No	No
LP	Yes	No	No
S-PROG	Yes	Yes	No
ANVIL	Yes	Yes	No
STEPS	Yes	Yes	Yes
LINDA	Yes	No	Yes

3.4 Verification

3.4.1 Verification metrics for deterministic nowcasts

Deterministic and mean probabilistic nowcasts were evaluated using a complementary set of metrics capturing intensity errors, spatial pattern skill, event detection capability, and scale dependence. All metrics were computed as a function of lead time and aggregated over the full spatial domain unless stated otherwise.

Pixel-wise intensity errors were quantified using the Mean Absolute Error (MAE)(Wilks 2005).

Pearson's correlation coefficient $\rho(t)$ was computed between the forecast field $F(\mathbf{x}, t)$ and the corresponding observed field $O(\mathbf{x}, t)$ across all grid cells $\mathbf{x}$ at each lead time t:

$$\rho(t) = \frac{\sum_{\mathbf{x}}\big(F(\mathbf{x},t) - \bar{F}(t)\big)\big(O(\mathbf{x},t) - \bar{O}(t)\big)}{\sqrt{\sum_{\mathbf{x}}\big(F(\mathbf{x},t) - \bar{F}(t)\big)^2 \sum_{\mathbf{x}}\big(O(\mathbf{x},t) - \bar{O}(t)\big)^2}}, \tag{3.5}$$

where $\bar{F}(t)$ and $\bar{O}(t)$ denote the spatial means of the forecast and observed fields. The decorrelation time τ is defined as the lead time at which $\rho(t)$ falls below $1/e \approx 0.368$, interpreted as the limit of useful predictability (Germann et al. 2002; Imhoff et al. 2020), and estimated by linear interpolation between the two bracketing lead times.

Event detection skill was evaluated using categorical verification metrics at a rainfall intensity threshold of $\theta = 0.1\ \mathrm{mm\,h^{-1}}$. Forecasts and observations were converted to binary exceedance fields, from which the metrics listed in Table 3.3 were derived (Wilks

2005).

To account for spatial displacement errors, the Fractions Skill Score (FSS) (Roberts et al. 2008) was employed. For a given neighbourhood size n and lead time t, FSS is defined as

$$\mathrm{FSS}(n,t) = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\big(f_i(n,t) - o_i(n,t)\big)^2}{\frac{1}{N}\sum_{i=1}^{N}\big(f_i^2(n,t) + o_i^2(n,t)\big)}, \tag{3.6}$$

where $f_i(n,t)$ and $o_i(n,t)$ denote the fractions of forecast and observed grid cells exceeding θ within a neighbourhood of size n centred at grid cell i. FSS ranges from 0 to 1, with higher values indicating better spatial agreement.

Table 3.3: Categorical verification metrics derived from the contingency table of forecast and observed binary exceedance fields at threshold $\theta = 0.1$ mm h^{-1}. Here, a denotes hits, b false alarms, and c misses.

Metric	Abbreviation	Formula
Critical Success Index	CSI	$\dfrac{a}{a+b+c}$
Probability of Detection	POD	$\dfrac{a}{a+c}$
False Alarm Ratio	FAR	$\dfrac{b}{a+b}$
Frequency Bias	BIAS	$\dfrac{a+b}{a+c}$

3.4.2 Verification metrics for probabilistic nowcast ensembles

The ensemble nowcasts were evaluated using a set of probabilistic verification metrics (Pulkkinen et al. 2019), applied consistently at the intensity threshold $\theta = 0.1$ mm h^{-1} with ensemble size N varied to assess sensitivity.

Ensemble calibration was assessed using the rank histogram (Talagrand diagram), in which the observation is ranked among the N ensemble members for each grid cell and lead time. A well-calibrated ensemble produces a uniform rank distribution; characteristic departures from uniformity diagnose underdispersion, overdispersion, or systematic bias (Hamill 2001; Wilks 2005) (see Fig. 3.4).

Overall probabilistic skill was quantified using the Continuous Ranked Probability Score (CRPS), which generalises the MAE to probabilistic forecasts by comparing the cumulative distribution functions of the forecast ensemble and the observation (Imhoff

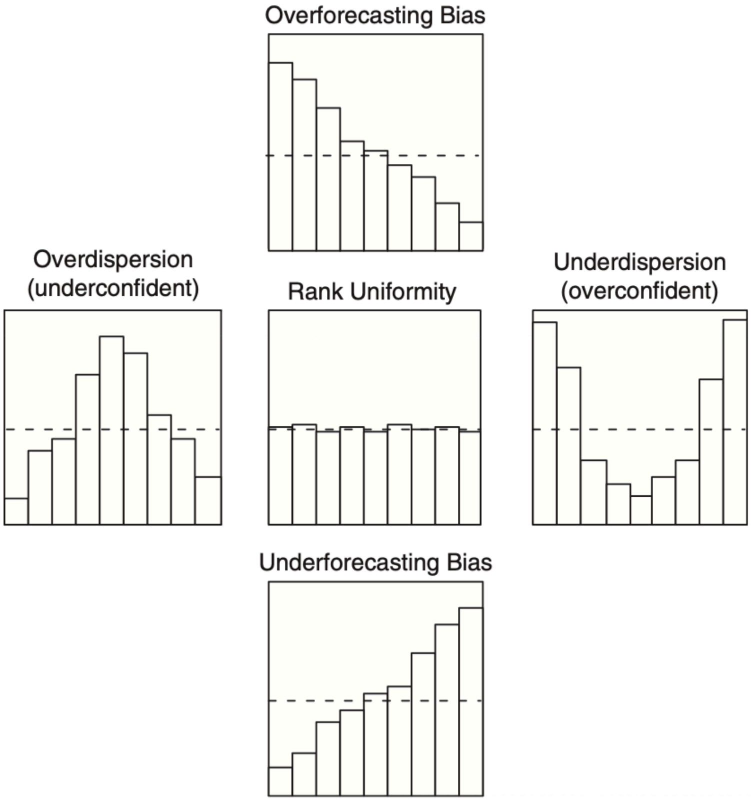

Figure 3.4: Verification rank histograms for hypothetical ensembles of 8 members, illustrating characteristic ensemble dispersion and bias errors. Reprinted from Wilks (2005).

et al. 2020):

$$\mathrm{CRPS} = \int_{-\infty}^{\infty} \big(F(y) - \mathbf{1}(y \geq x)\big)^2 \, \mathrm{d}y, \tag{3.7}$$

where $F(y)$ is the forecast CDF and $\mathbf{1}(y \geq x)$ is the unit step (Heaviside) function at the observed value x. Lower values indicate better skill. Crucially, the CRPS reduces to the MAE for deterministic forecasts, enabling direct comparison between deterministic and probabilistic methods. CRPS was computed as a function of lead time to characterise the temporal decay of ensemble skill.

Forecast calibration was assessed using the reliability diagram (Fig. 3.5), which plots observed relative frequency of threshold exceedance against the forecast probability for each probability bin. A perfectly calibrated ensemble follows the 1:1 diagonal; systematic departures indicate over- or underforecasting, or miscalibrated confidence (Wilks 2005). The accompanying refinement distribution reflects the sharpness of the ensemble.

Discrimination ability was assessed using the ROC curve, which evaluates whether

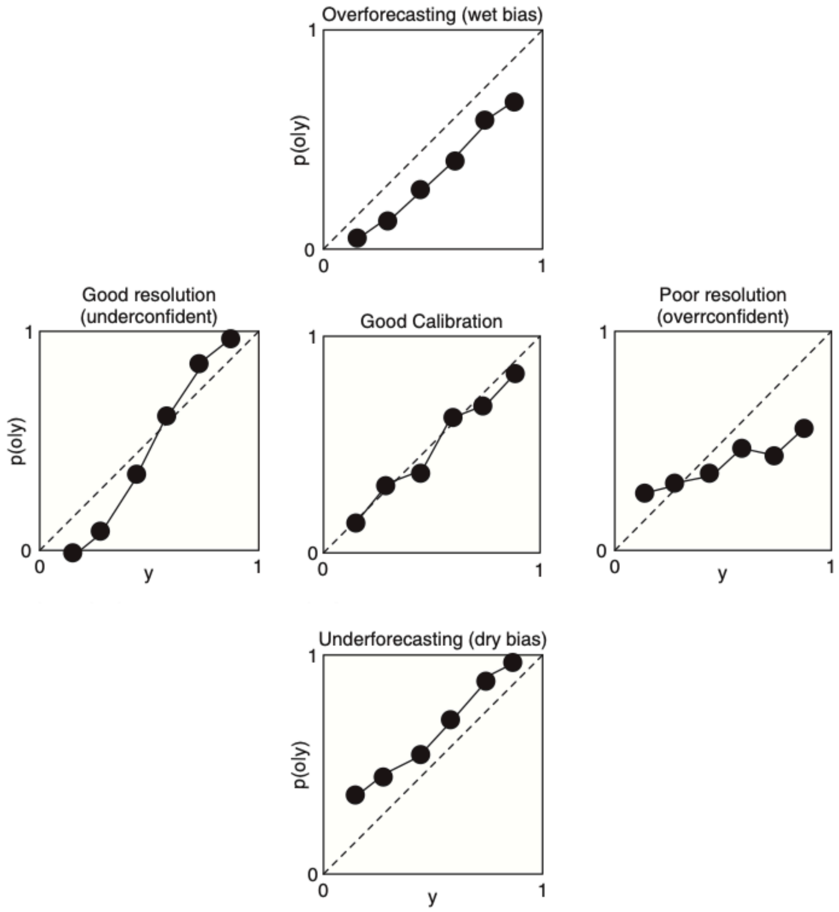

Figure 3.5: Reliability diagrams illustrating systematic forecast biases and differences in forecast confidence, based on hypothetical ensemble forecasts. Adapted from Wilks (2005).

ensemble exceedance probabilities consistently rank rain-producing grid cells above dry ones, independently of calibration. The area under the curve (AUC) summarises discrimination in a single value, where 1.0 indicates perfect discrimination and 0.5 corresponds to no skill (Wilks 2005).

3.5 Experimental Setup

3.5.1 Sensitivity Analysis

A sensitivity analysis was conducted to assess how key model configuration parameters influence nowcast performance. All experiments were performed locally on a consumer-grade laptop (Apple M1 Air, 8-core CPU with 4 performance cores and 4 efficiency cores, 16 GB unified memory), which imposed practical constraints on computational cost. The analysis therefore balanced parameter coverage against feasible runtimes rather than exhaustively exploring the full parameter space.

3.5.1.1 Deterministic

S-PROG and ANVIL. The deterministic sensitivity experiments for S-PROG and ANVIL were carried out across all 19 events. The parameters investigated includethe number of cascade levels used in the multiscale decomposition, and the autoregressive (AR) model order.

A one-factor-at-a-time approach was adopted, where each parameter was varied while keeping all others fixed at baseline values (AR order 2, 8 cascade levels, native 0.5 km / 5 min resolution).

The AR order was varied between 1, 2, and 3, and the number of cascade levels was tested at 3, 8, and 16. This approach allows for clear attribution of performance changes, although it does not capture potential interactions between parameters.

LINDA-D. The sensitivity analysis for LINDA-D was conducted on all 19 events (covering 1 h, 3 h, and 6 h durations across both seasons) under the same OAT design, with all other settings held at their PySTEPS defaults. Three LINDA-specific parameters were examined:

- **Localization radius** (g_w): the standard deviation of the Gaussian window used to estimate spatially localised model parameters. Values of 50, ∼100 (PySTEPS default, $0.2 \times \min(m, n)$), and 200 grid pixels were tested, corresponding to approximately 25 km, 50 km, and 100 km at native 0.5 km resolution.
- **Blob-detection threshold** ($b_{\min}$): the minimum scale parameter (`min_sigma`) in the Laplacian-of-Gaussian (LoG) feature detector. A low value (1.5) retains more, finer-scale precipitation features, while a high value (6.0) retains only larger structures.
- **AR input order** (p): the autoregressive order of the LINDA model (`ari_order` = 1, 2, or 3), which determines how many prior radar frames are used to fit the local AR coefficients. The baseline uses `ari_order` = 2.

For each event, nowcasts were generated using a rolling-initialisation strategy. The rolling-initialisation principle follows Imhoff et al. (2020); the specific schedule used here - 25 nowcasts per event, issued every 5 min from 90 min before event onset to 30 min after - was chosen so that verification statistics reflect performance across a range of atmospheric states within each event, rather than depending on a single initialisation (Fig. 3.6). Each nowcast extends 90 minutes into the future at the native radar resolution and 5-minute temporal steps, yielding 18 forecast lead times per initialisation. Only lead times falling within the event window contribute to verification; lead times outside the event window are discarded. Verification metrics were aggregated across all valid lead times and initialisations to produce a single score per event, parameter configuration, and metric. This rolling approach ensures that verification statistics reflect forecast performance across a range of atmospheric states within each event, rather than depending on a single initialisation.

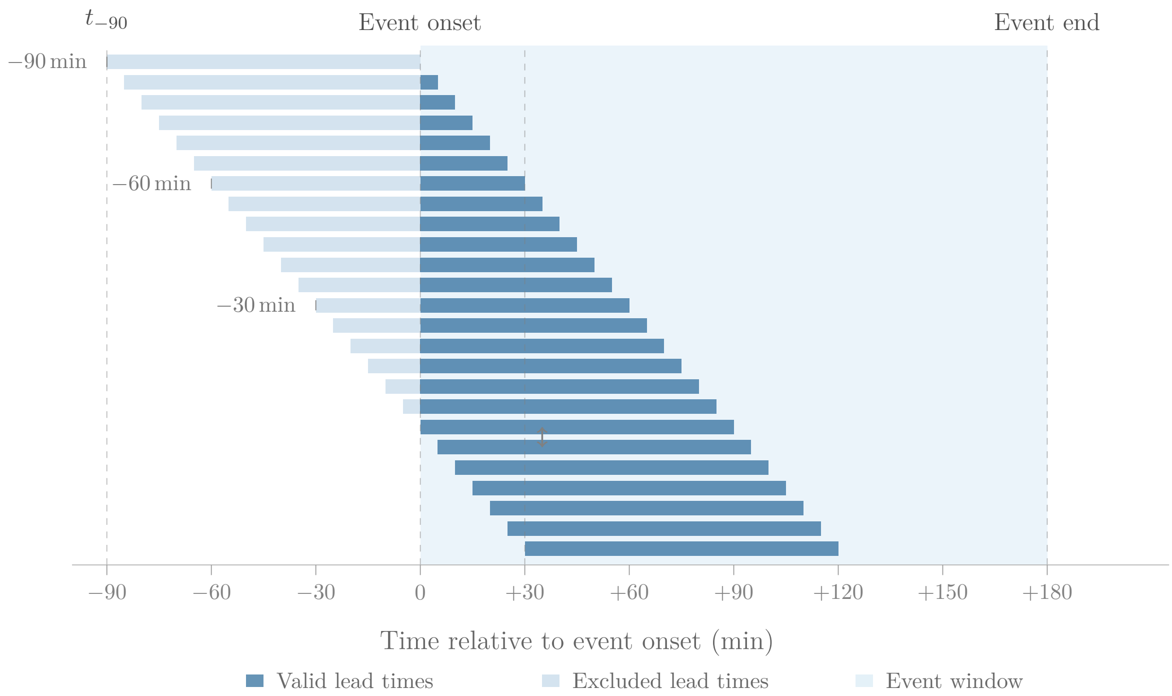

Figure 3.6: Schematic illustration of the rolling initialisation strategy applied in this study, inspired by Imhoff et al. (2020) and Tounsi et al. (2023). A total of 25 nowcasts are issued per event, initialised every 5 minutes starting 90 minutes prior to event onset and ending 30 minutes into the event. Each nowcast has a lead time of 90 minutes. Only lead times falling within the event window contribute to verification (dark blue); lead times outside the event window are excluded (light blue). The figure illustrates a representative 3-hour precipitation event.

While the use of local hardware limits the scale of the experiments, the relative differences observed between configurations are considered robust, as all simulations were performed under consistent computational conditions.

3.5.2 Full Runs

All nowcasts for the full runs were produced using `PySTEPS` v1.20.0 running locally on an Apple M1 Air (8 cores, 16 GB RAM).

Rolling initialisations were utilised for the full runs as well, as described in Section 3.5.1.1.

Nowcasts were computed on a cropped version of the regional subdomain delivered by SMHI (approximately 11.0–16.0°E and 54.5–57.0°N), which provides sufficient spatial context for motion field estimation while limiting computation time. For verification, a smaller domain centred on the Malmö urban area was used, covering approximately 12.75–13.5°E and 55.5–56.0°N (see Fig. 3.7). This domain encompasses the rain gauge stations used for event selection and comprises 113 × 98 grid cells at native resolution (approximately 57 × 49 km), providing sufficient spatial support for robust verification statistics and FSS computation at neighbourhood scales of 5–10 km while remaining representative of the urban area relevant to the intended FBC application. The exact domain boundaries were not optimised, but since all configurations are verified against the same domain, relative comparisons remain robust.

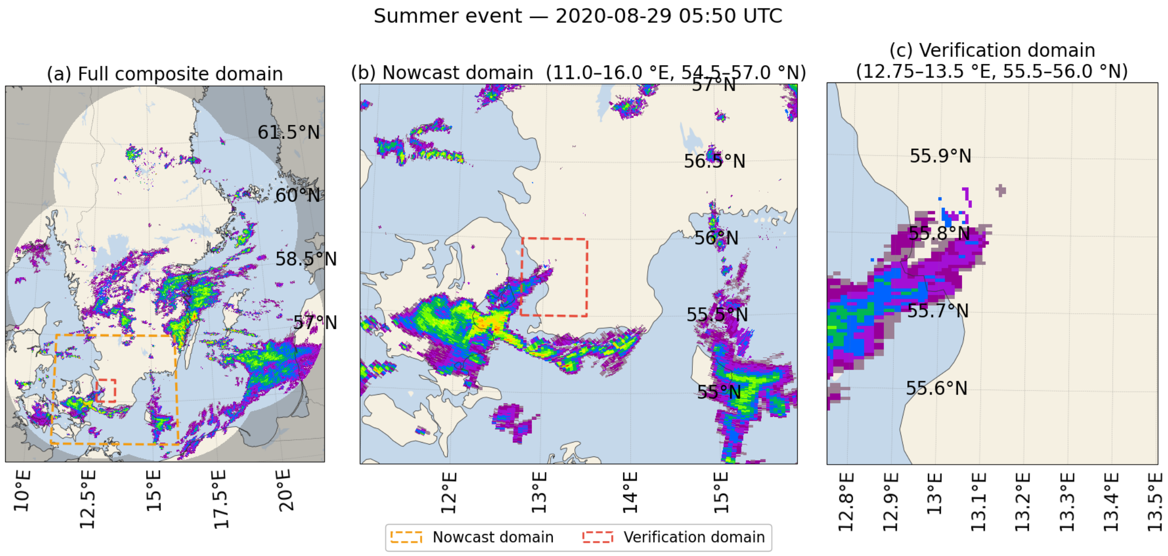

Figure 3.7: Spatial domains used for nowcasting and verification, illustrated on a Summer event (2020-08-29 05:50 UTC). (a) Full SMHI south-Sweden radar composite. The orange dashed box marks the nowcast domain and the red dashed box marks the verification domain. (b) Nowcast domain (11.0–16.0°E, 54.5–57.0°N), on which motion estimation and ensemble nowcasts were computed. The red dashed box indicates the verification domain. (c) Verification domain (12.75–13.5°E, 55.5–56.0°N), centred on the Malmö urban area.

Five deterministic nowcasting configurations - Eulerian persistence (EP), Lagrangian persistence (LP), S-PROG, ANVIL, and LINDA-D - were applied to every event and initialisation time. Based on the sensitivity analysis, S-PROG and ANVIL were run using the PySTEPS default parameters (AR/ARI order 2, 8 cascade levels), and LINDA with AR order 1.

Motion fields were estimated using the Lucas-Kanade optical flow method applied to the four most recent radar composites prior to each initialisation time (Section 3.3.1). Radar reflectivity fields were converted to rainfall rates using the Marshall–Palmer Z–R relationship ($a = 200$, $b = 1.6$) prior to nowcast generation.

To assess sensitivity to input resolution, two coarser configurations were produced alongside the native 0.5 km / 5 min resolution: 1 km / 10 min and 2 km / 15 min. Each was generated by block-averaging the native radar fields over the corresponding spatial window (2×2 and 4×4 grid cells) and subsampling the resulting timeseries at the target temporal stride ($2\times$ and $3\times$ the native 5 min step).

3.5.3 Exploratory Probabilistic Runs

To provide an initial assessment of probabilistic nowcasting for FBC applications, exploratory runs were performed using STEPS and LINDA with 16 ensemble members. These runs used a single fixed initialisation at 30 min before event onset, in contrast to the rolling initialisation strategy used for the deterministic analysis. Results are therefore not directly comparable to the deterministic verification and should be interpreted as indicative of the potential added value of ensemble forecasts for FBC decision-making.

4 Results

4.1 Sensitivity Analysis

4.1.1 S-PROG & ANVIL

Figure 4.1 summarises the deterministic sensitivity analysis of autoregression order and cascade levels, as relative change from baseline for each parameter–metric combination at 60 min lead time. The baseline configuration uses AR order 2, 8 cascade levels, and native resolution (0.5 km / 5 min).

S-PROG is robust to both parameters. Deviations from baseline remain within ±6 % for both metrics, with most falling below ±4 %.

ANVIL shows a qualitatively different pattern. AR order 3 produces substantial degradation across all metrics, with MAE worsening by 31.2 % and CSI by 6.7 %. This performance drop suggests that the optical-flow-based advection used by ANVIL may be sensitive to higher-order autoregressive terms. Sensitivity to the number of cascade levels remains modest for ANVIL.

Overall, inter-event variability far exceeds parameter sensitivity for both models, indicating that nowcast skill at 60 min lead time is governed primarily by the meteorological situation rather than by configuration choices.

Based on these results, all subsequent deterministic runs use the baseline configuration (AR order 2, 8 cascade levels). Notably, SMHI's operational ANVIL configuration (AR order 2, 3 cascade levels, 2 km/15 min resolution) (Falahat 2026) falls within the robust parameter space identified here, albeit at a coarser resolution.

4.1.2 LINDA-D

Figure 4.2 summarises the deterministic sensitivity of LINDA-D to three parameters, evaluated on the 19 events at 60 min lead time: the *localisation radius*, which controls the spatial extent over which the convolution kernel and ARI parameters are estimated locally rather than globally; the *blob threshold*, which sets the detection sensitivity for identifying rain cells around which local parameters are fitted; and the *motion frame AR order*, which specifies the order of the autoregressive integrated process. LINDA-D is most sensitive to the localisation radius. The default radius is automatically deter-

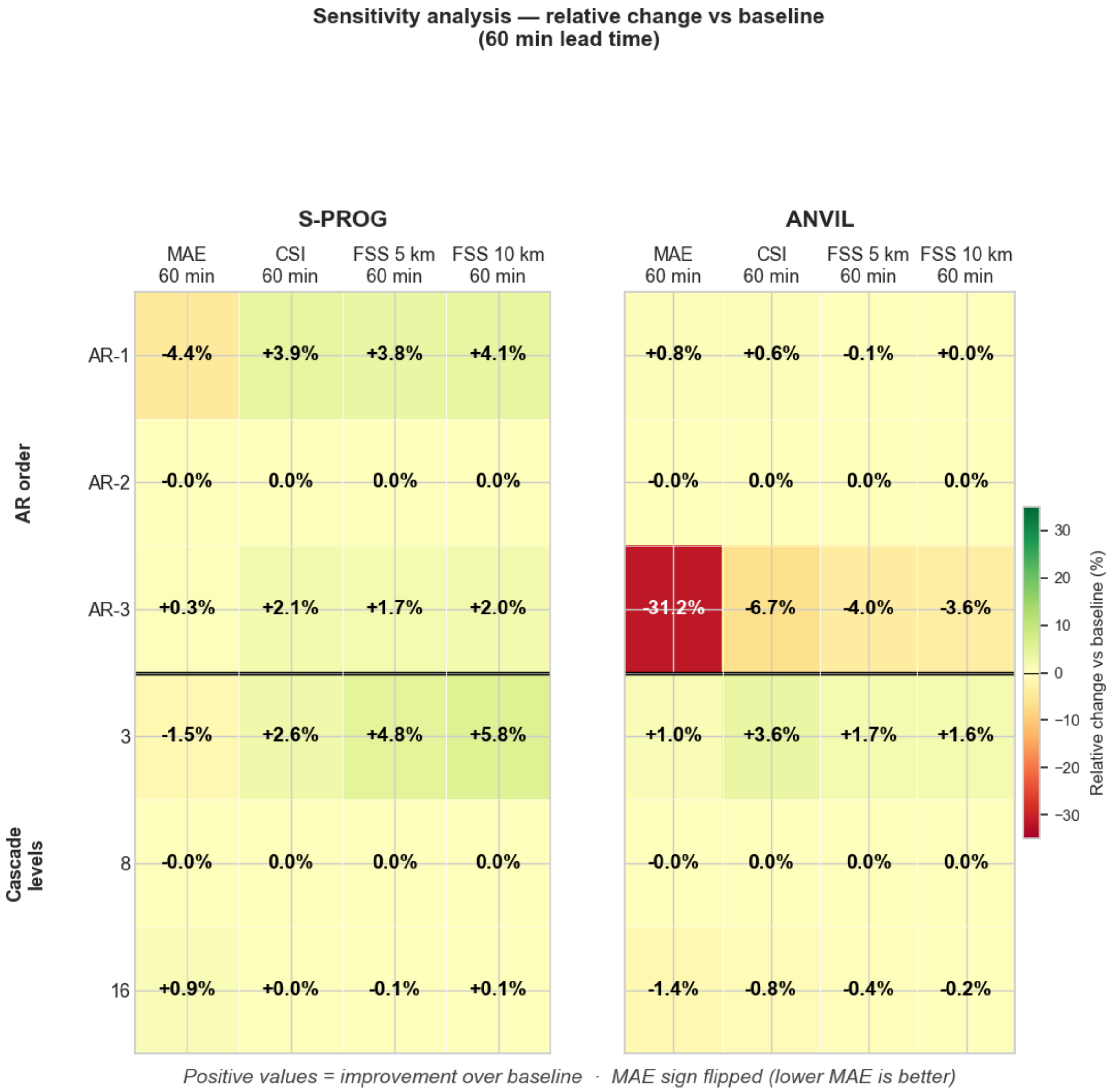

Figure 4.1: Sensitivity of S-PROG and ANVIL to AR order and cascade levels. Values show relative change from baseline (AR order 2, 8 cascade levels). The MAE sign is flipped so that green consistently indicates improvement across all metrics.

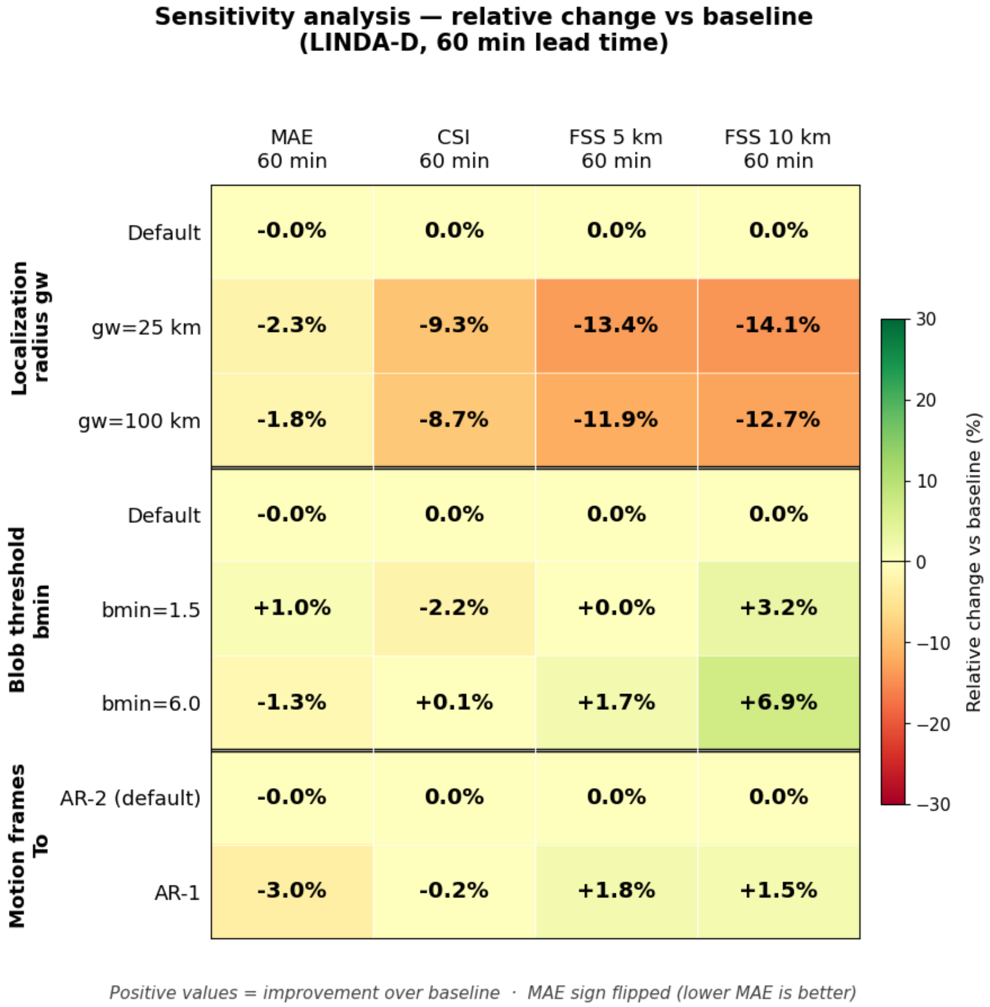

Figure 4.2: Sensitivity of LINDA-D to localisation radius, blob threshold, and motion frame AR order at 60 min lead time, evaluated on 19 events. Values show relative change from the default configuration. The MAE sign is flipped so that green consistently indicates improvement.

mined by PySTEPS as $0.2 \times \min(Height, Width)$ of the nowcasting window, yielding approximately 56 km for the domain used in this study. Reducing the radius to 25 km degrades all metrics substantially, with FSS at 5 and 10 km dropping by 13–14 % and CSI by 9.3 %. Increasing the radius to 100 km produces a similar but slightly weaker degradation. The default value thus sits near the optimum, balancing local adaptivity against spatial smoothness.

Blob threshold and motion frame AR order show more modest effects. A higher blob threshold (bmin = 6.0) improves FSS at both scales (+1.7 % and +6.9 %) at the cost of slightly worse MAE (−1.3 %), indicating that fewer, larger blobs produce spatially smoother forecasts that verify better at the neighbourhood scale. Reducing the motion frame AR order from 2 to 1 shows a similar trade-off: improved FSS (+1.5–1.8 %) but degraded MAE (−3.0 %). AR order 3 was also tested, but did not produce a nowcast due to an unresolved implementation issue; these runs are therefore excluded.

Based on these results, all subsequent LINDA-D runs use the default configuration, as no alternative consistently improves all metrics.

4.2 Deterministic Nowcast Performance

This section presents the verification of the five deterministic nowcasting methods (S-PROG, LP, EP, ANVIL, LINDA-D) at native PySTEPS resolution (0.5 km / 5 min) over all 19 events. S-PROG and ANVIL are configured with the baseline parameters identified in Section 4.1.1 (AR order 2, 8 cascade levels); LINDA-D uses the default configuration identified in Section 4.1.2. LP and EP serve as Lagrangian and Eulerian persistence baselines. All forecasts are evaluated at lead times from 5 to 90 min in 5-min steps. Verification at the coarser configuration used by SMHI's KNEP product is presented separately in Section 4.3.

4.2.1 Overall verification

Before turning to aggregated metrics, Figure 4.3 illustrates the qualitative behaviour of the five deterministic configurations for a representative event (initialised at 2021-09-27 22:45 UTC; 2021-09-28 00:45 local time, Europe/Stockholm). The persistence baselines (EP and LP) preserve the structure and intensity of the initial radar field but, by construction, cannot represent growth or decay; LP additionally advects the field with the estimated motion. S-PROG produces visibly smoothed forecasts in which small-scale features dissipate progressively with lead time, while ANVIL and LINDA-D retain sharper precipitation structures throughout the 90 min horizon. These visual differences foreshadow the bias and intensity-error patterns quantified below.

Figure 4.4 aggregates verification across all 19 events for the eight metrics computed in this study. S-PROG performs best or tied-best in the categorical (CSI, POD), spatial (FSS at 5 and 10 km), and bias metrics, with LP consistently second. ANVIL and LINDA-D occupy the middle range, scoring below S-PROG and LP across CSI, POD, and FSS at all lead times. EP serves as the lower bound across all metrics, with Pearson r approaching zero by 30 min and CSI degrading steadily throughout the lead-time range.

The MAE panel reveals that S-PROG retains the lowest absolute errors throughout the lead-time range (1.0–2.3 mm/h at 5–90 min), followed by LP. ANVIL and LINDA-D produce higher MAE than the persistence baseline LP, with LINDA-D approaching EP's error level beyond 60 min lead time. The smoother S-PROG fields visible in Figure 4.3 are consistent with this lower pixel-wise error.

The bias panel is informative for distinguishing between intensity preservation and detection skill. S-PROG maintains bias close to unity throughout the lead-time range

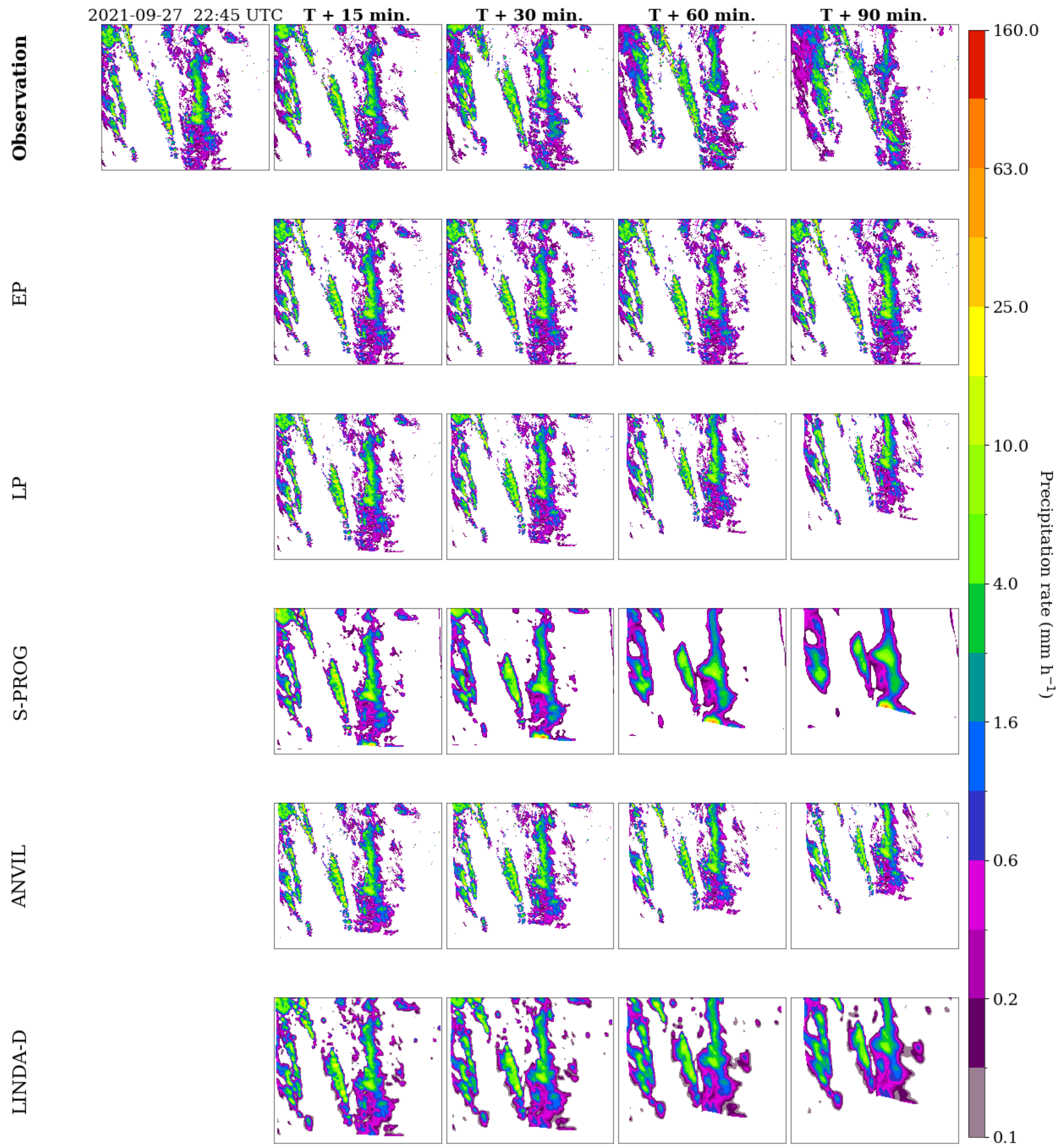

Figure 4.3: Forecast evolution at lead times of 15, 30, 60, and 90 min for the five deterministic configurations, initialised at 2021-09-27 22:45 UTC. Top row shows the corresponding observed radar fields. EP holds the initial field constant; LP advects it along the estimated motion field; S-PROG progressively smooths small-scale features through cascade-AR filtering; ANVIL and LINDA-D retain sharper structures through their alternative scale treatments.

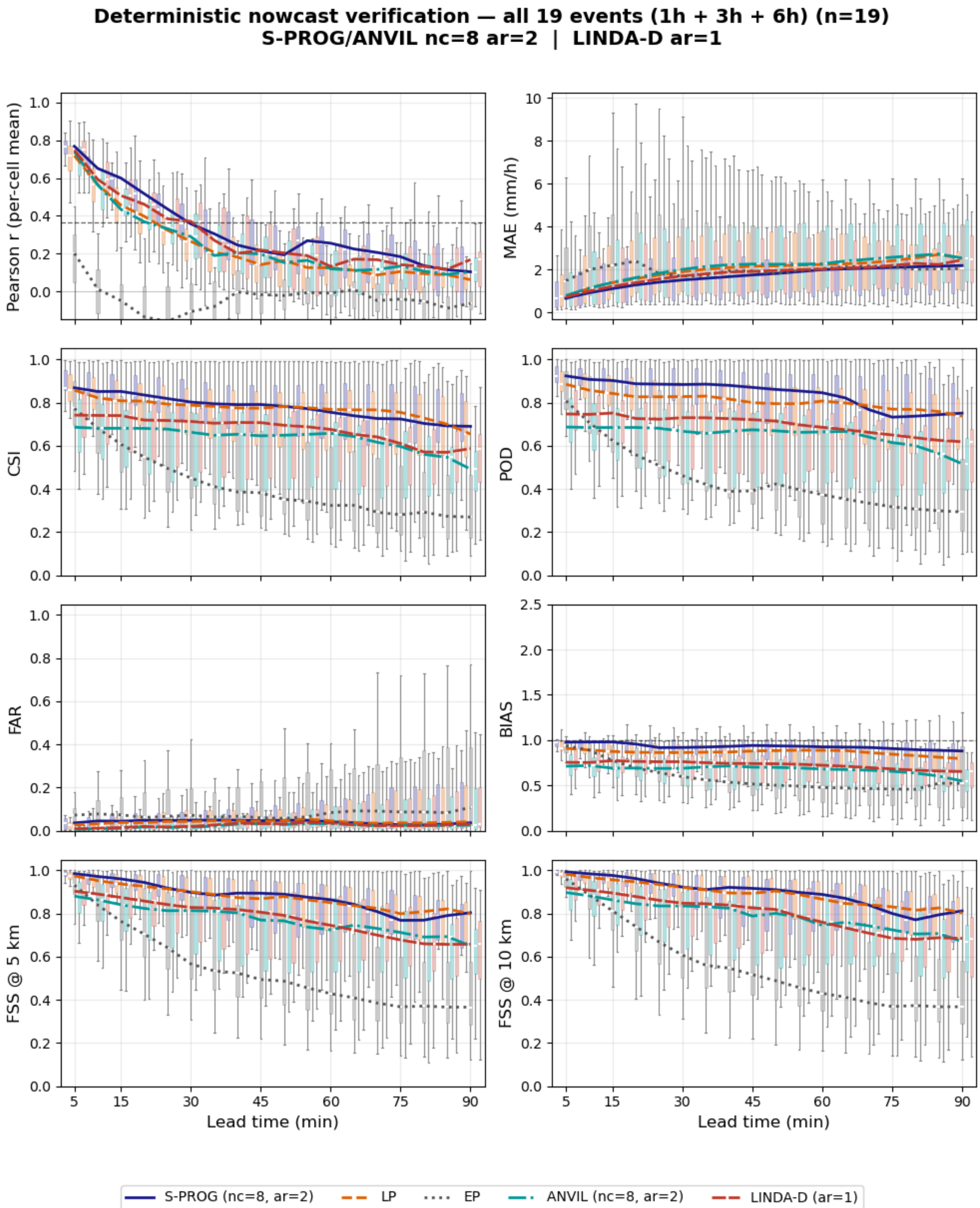

Figure 4.4: Deterministic verification across all 19 events at native resolution (0.5 km / 5 min). Lines show median values per model; boxes show inter-event variability (25th–75th percentile); whiskers extend to the 5th and 95th percentiles. The dashed line in the Pearson panel marks $r = 1/e \approx 0.368$.

(0.95 at 5 min, 0.87 at 90 min), while LP, ANVIL, and LINDA-D systematically underpredict (bias 0.55–0.75 by 90 min). EP shows the largest bias drift, falling to approximately 0.55 at 90 min.

False-alarm ratio remains low and broadly comparable across all skilful methods (FAR < 0.10 for S-PROG, LP, ANVIL, and LINDA-D throughout 5–90 min).

4.2.2 Decorrelation time and skilful lead time

Pearson correlation stratified by season and event duration is shown in Figure 4.5. The dashed horizontal line marks $r = 1/e \approx 0.368$, the threshold conventionally used to define maximum skilful lead time (Germann et al. 2002).

Two patterns are evident. First, winter events are consistently more predictable than summer events of the same duration. For 1-hour events, S-PROG crosses 1/e at approximately 20 min in summer but maintains $r > 1/e$ until 40 min in winter. Simillar separation holds for 3- and 6-hour events, with 30/50 respectively 30/70 min.

Second, S-PROG retains the longest skilful lead time across most (season $\times$ duration) categories. The crossing of the 1/e threshold for S-PROG occurs at approximately 17, 30, and 30 min in summer for 1, 3, and 6-hour events, and at 41, 47 and 46 min in winter. The exception from the dominance of S-PROG is at 1h winter and 3h summer events where LINDA-D overtakes by a slight measure.

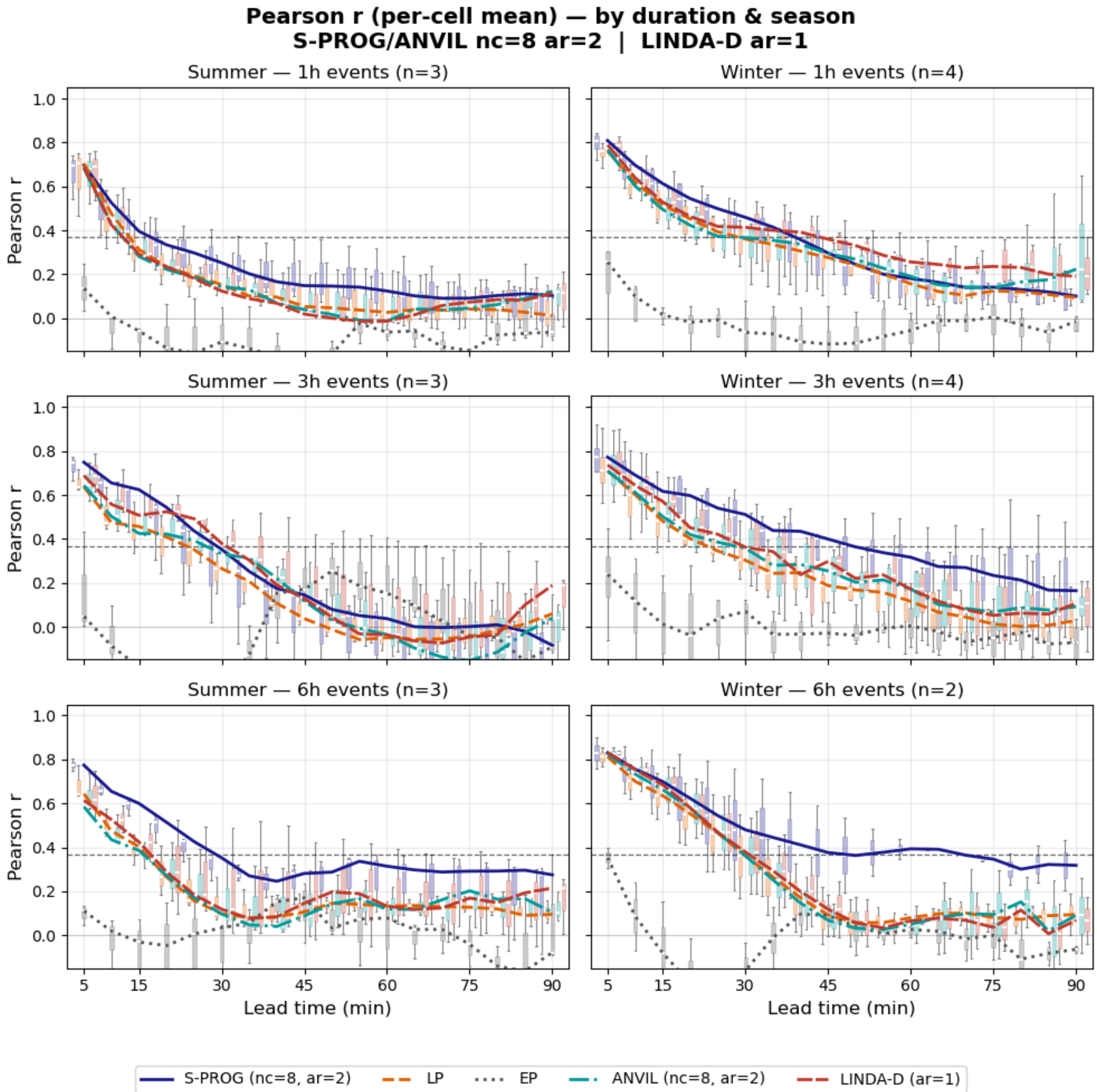

Figure 4.5: Pearson correlation between forecast and observation versus lead time, stratified by season (columns) and event duration (rows). The dashed line at $r = 1/e \approx 0.368$ marks the decorrelation threshold conventionally used to define maximum skilful lead time. Box plots show inter-event variability; lines show median per model. Sample sizes (n) are indicated in each panel title.

4.2.3 Seasonal and duration dependence in intensity errors

Figure 4.6 shows MAE stratified by season, event duration, and lead-time window (5–30, 35–60, 65–90 min). The figure isolates how absolute forecast errors vary with both meteorological context and forecast horizon.

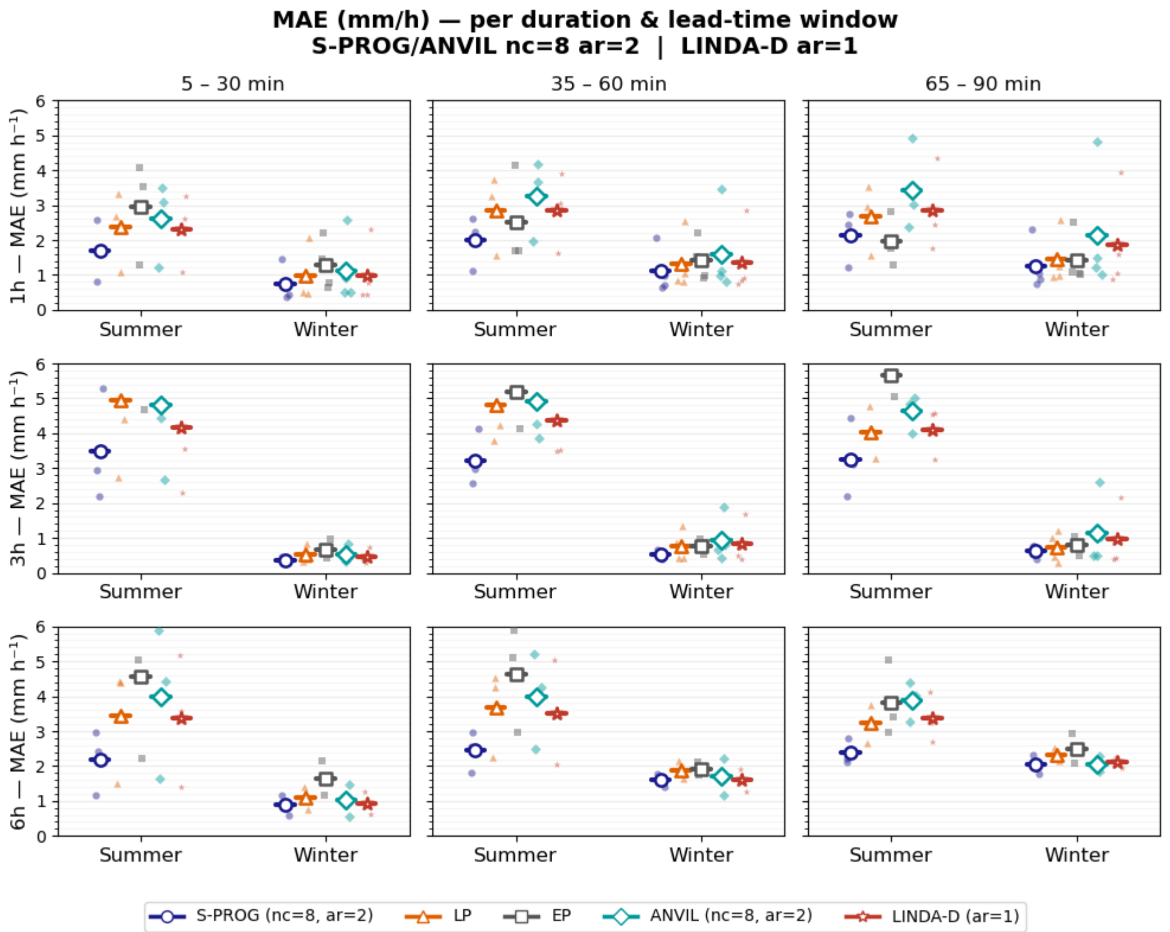

Figure 4.6: Mean absolute error stratified by event duration (rows), lead-time window (columns), and season (within-panel). Open markers show median values; small markers show individual events. Note the factor-of-3 difference in MAE between summer and winter at all duration–lead-time combinations.

Winter MAE is approximately three times lower than summer MAE across all duration and lead-time combinations. For 3-hour events at 35–60 min lead time, S-PROG MAE is 2 mm/h in summer versus 1 mm/h in winter; for 6-hour events at the same lead-time window, the corresponding values are 2.7 and 1.3 mm/h. The seasonal contrast is largest for 3-hour events, where summer MAE exceeds winter MAE by a factor of 5–7 across all lead-time windows.

The relative ranking of methods is preserved across seasons: S-PROG produces the lowest MAE in all 18 (duration $\times$ season $\times$ lead-time-window) combinations, followed by LP. ANVIL and LINDA-D track each other closely and exceed LP's MAE in most categories. EP exceeds the skilful methods in summer but is comparable to LP, ANVIL, and LINDA-D in winter, where the absolute error magnitudes are too small for the methods to differentiate strongly.

MAE growth with lead time is most pronounced for 1-hour summer events (1.8 mm/h at 5–30 min growing to 2.5 mm/h at 65–90 min for S-PROG) and weakest for 6-hour winter events (0.7 to 1.7 mm/h over the same windows). This pattern is consistent with the Pearson-based decorrelation analysis: events with the steepest correlation decay also seem to show the steepest absolute error growth.

4.3 Comparative runs at 1 km/10 min and 2 km/15 min

To assess how nowcast performance varies with input resolution, S-PROG, ANVIL, and LINDA-D were rerun at two coarser configurations (1 km / 10 min and 2 km / 15 min) in addition to the native 0.5 km / 5 min results presented in Section 4.2. Figure 4.7 summarises performance at 30 min lead time across seven verification metrics, normalised across all methods and resolutions. Figure 4.9 shows the full lead-time evolution of Pearson r stratified by season and event duration. See Appendix A2 for summary of all runs at all resolutions, as well as spider plots at 60 min lead time.

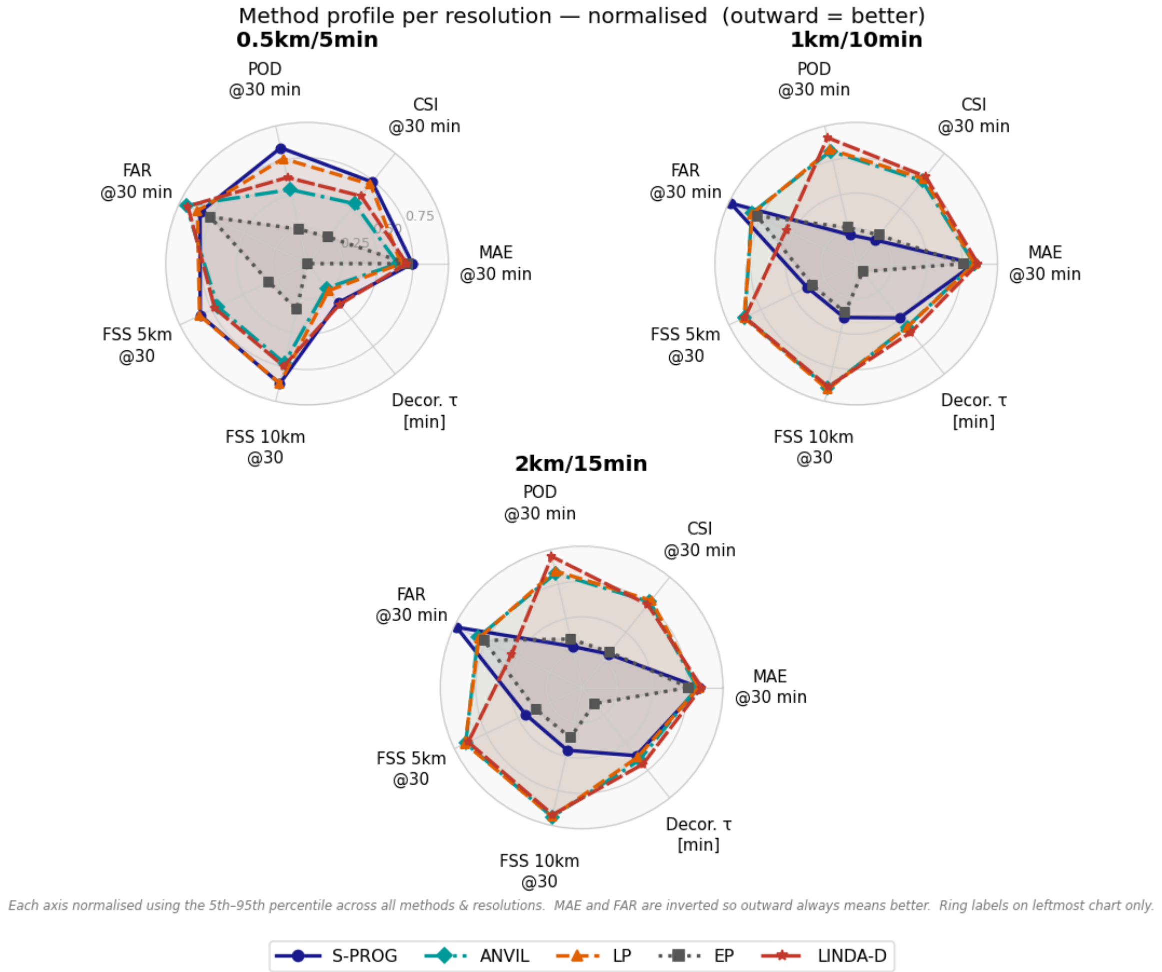

Figure 4.7: Method profile per resolution at 30 min lead time, aggregated across all 19 events. Each axis shows a verification metric normalised so that 0 corresponds to the bottom 5th percentile and 1 to the top 5th percentile across all methods and resolutions; outward always indicates better performance. MAE and FAR are inverted accordingly.

The spider plots show distinct method rankings at each resolution. At native 0.5 km / 5 min (left panel), S-PROG and LP occupy the outermost positions across most axes, with ANVIL and LINDA-D forming an inner band and EP serving as the lower bound. At 1 km / 10 min (centre panel) the configuration shifts: LINDA-D extends outward on POD, CSI, and both FSS axes, while S-PROG retains the MAE and FAR axes. At 2 km / 15 min (right panel) the inversion is fully developed: LINDA-D dominates POD, CSI, FSS at both 5 and 10 km, and decorrelation time, while ANVIL traces a similar but slightly inner profile. S-PROG retains the outermost MAE and FAR positions but has collapsed inward on the categorical and spatial metrics.

Figure 4.8 shows how frequency bias at 30 and 60 min lead time evolves across the three input resolutions. The methods diverge sharply. S-PROG and LP both maintain bias close to unity at native resolution (approximately 0.9) but drop to 0.4–0.5 at 1 km / 10 min and 2 km / 15 min, indicating systematic under-prediction of rain area at coarser configurations. LINDA-D shows the opposite trajectory, rising from 0.7 at native resolution to approximately 1.1 at the two coarser configurations and remaining slightly above unity. ANVIL occupies an intermediate range (0.7–0.9) with comparatively flat behaviour across resolutions. EP remains stable at approximately 0.5 throughout, a level that S-PROG and LP both fall to at the coarser configurations.

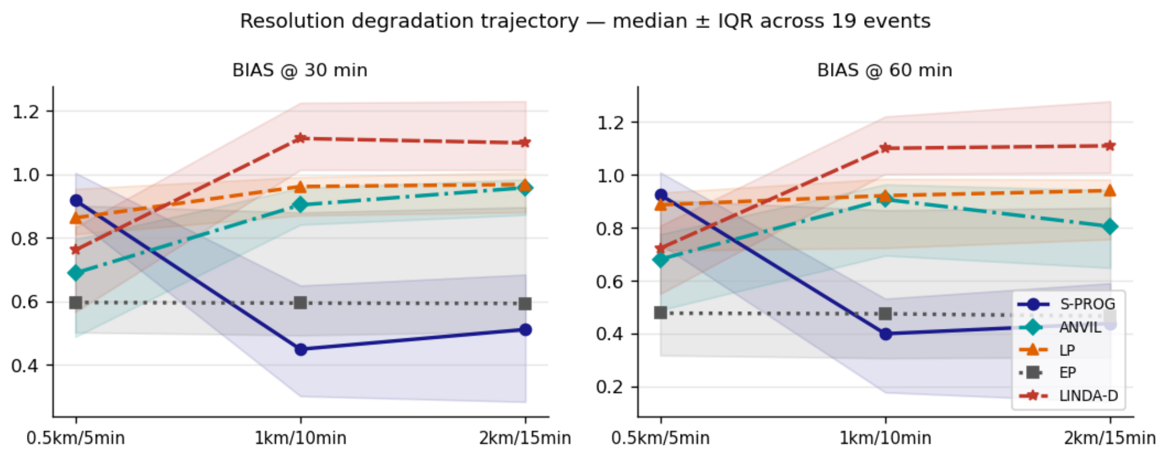

Figure 4.8: Frequency bias a t30 and 60 min lead time as a function of input resolution, aggregated across all 19 events. Lines show median values per method; shaded bands show inter-event IQR. Values close to unity indicate correct prediction of rain area; values below unity indicate systematic under-prediction.

The Pearson stratification in Fig. 4.9 reveals two consistent patterns. First, the native 0.5 km / 5 min configuration produces the lowest correlation values across all three methods in nearly every season–duration panel, with ANVIL and LINDA-D at native resolution falling below 0.2 within 30 min for most summer cases. Second, the 1 km / 10 min and 2 km / 15 min configurations track each other closely and lie consistently above the native curves, often by 0.1–0.3 in correlation at matched lead times. The separation between native and coarser resolutions is largest for short summer events (1 h and 3 h durations, left column), where native ANVIL and LINDA-D approach zero correlation by 60 min while their coarser counterparts retain $r \approx 0.2$–0.3.

The decorrelation behaviour against the $r = 1/e$ threshold differs by configuration. For 6h winter events, all models at 1 km / 10 min and 2 km / 15 min remain above 1/e throughout the full 90 min lead-time range. For summer events, crossings occur earliest at native resolution (within 20 min for 1 h events) and latest at the coarser configurations (typically 50–70 min). The 6 h summer panel (n=3) shows the most resolution- dependent spread, with S-PROG at coarse resolution remaining above 1/e throughout while native S-PROG, ANVIL, and LINDA-D fall well below the threshold by 20-30 min.

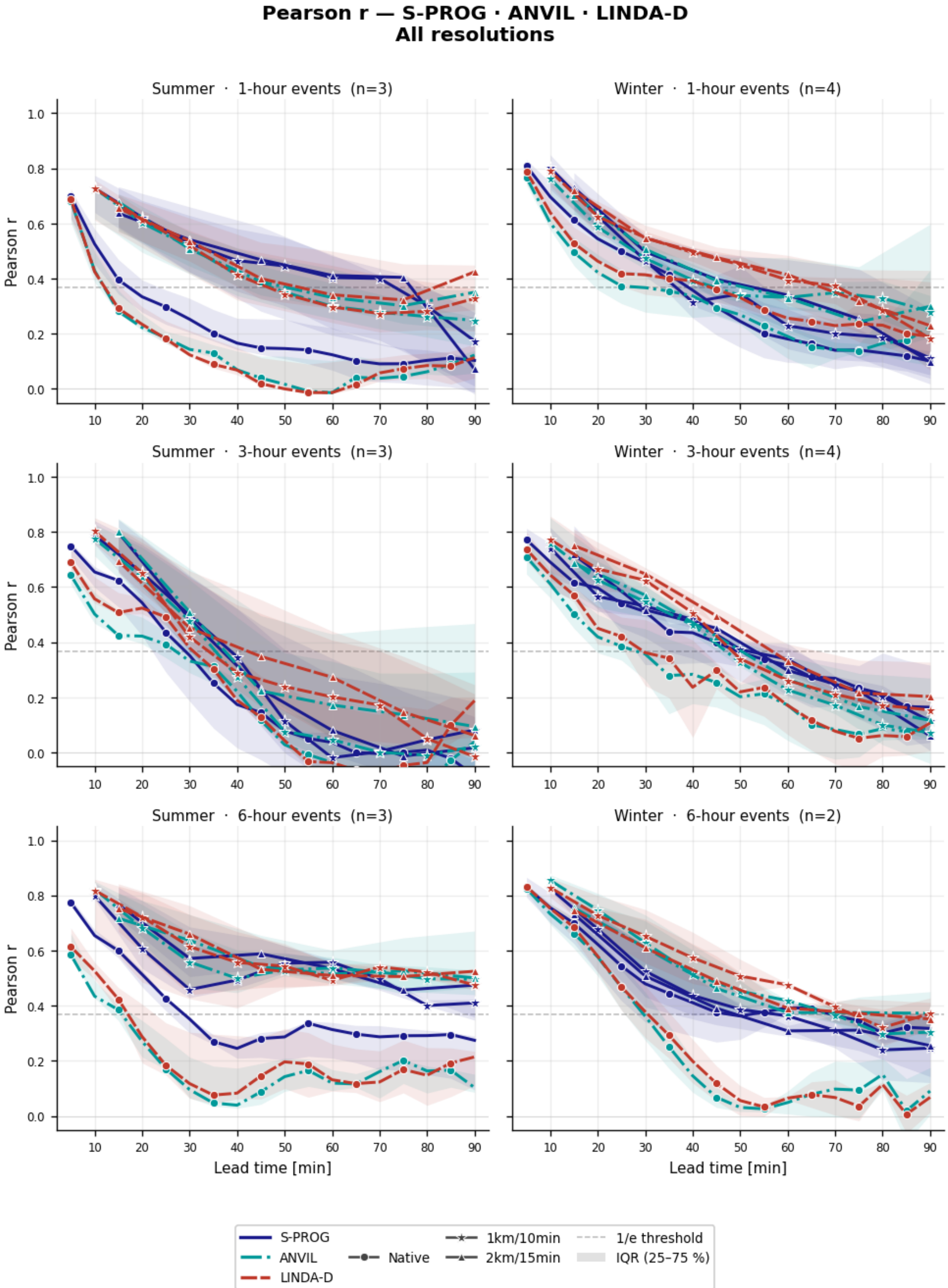

Figure 4.9: Pearson correlation between forecast and observation versus lead time, stratified by season (columns) and event duration (rows), shown for S-PROG, ANVIL, and LINDA-D at all three resolutions. Marker style indicates resolution: circles for 0.5 km / 5 min, stars for 1 km / 10 min, and triangles for 2 km / 15 min. Shaded bands show the inter-event IQR (25th–75th percentile). The dashed horizontal line marks $r = 1/e \approx 0.368$. Sample sizes (n) are indicated in each panel title.

4.4 STEPS versus LINDA-P

This section compares STEPS and LINDA-P on a stratified subset of five events: convective and stratiform summer (#38, #15), stratiform winter (#9), mixed summer frontal (#30), and a long-duration weak-skill convective summer case (#28). Both methods were run with 16 ensemble members at native resolution (0.5 km / 5 min) over 18 lead times (5–90 min) using the same fixed-onset-minus-60-minutes initialisation. The verification threshold was set to 0.1 mm/h.

4.4.1 Ensemble calibration

Aggregated rank histograms (Fig. 4.10, top) indicate calibration difference. STEPS produced a near-uniform distribution with mild excess in the highest bin (8.9 % versus 6.3 % expected). LINDA-P showed a pronounced spike in the lowest bin (27 % of observations) and a monotonic decay across the remaining bins, with a small secondary excess in the highest bin (4.8 %).

The reliability diagrams (Fig. 4.10, bottom) confirm that the methods are biased in opposite directions at this threshold: the STEPS curve lies consistently above the diagonal (under-prediction) and the LINDA-P curve consistently below (over-prediction).

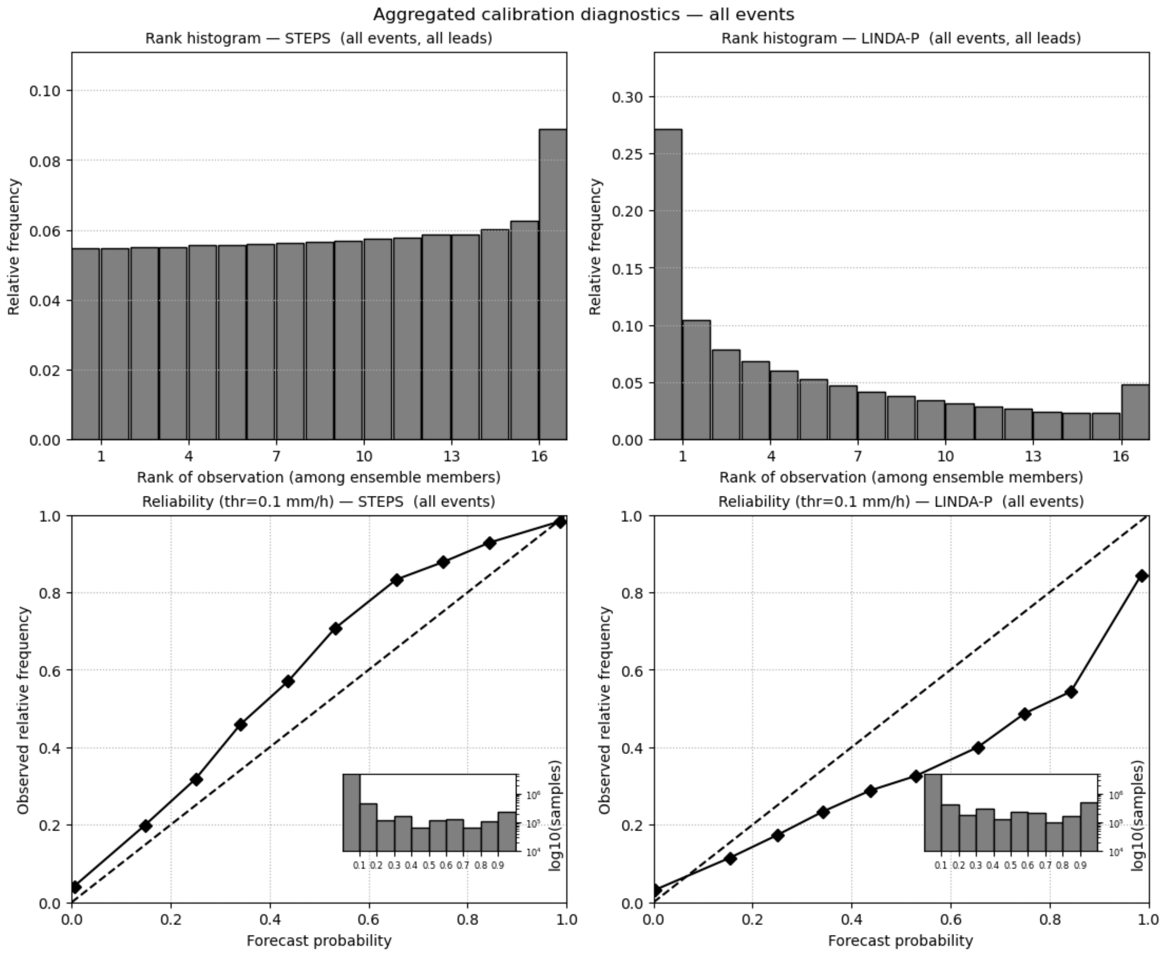

Figure 4.10: Aggregated rank histograms (top) and reliability diagrams at the 0.1 mm/h threshold (bottom) for STEPS (left column) and LINDA-P (right column), combined across all five events and 18 lead times. STEPS shows near-uniform rank distribution with mild excess in the highest bin (under-dispersion at the upper tail). LINDA-P shows a pronounced spike in the lowest rank bin (27 % of observations) indicating systematic wet bias. The reliability diagrams confirm opposite-direction biases at the rain/no-rain boundary: STEPS under-predicts (curve above diagonal), LINDA-P over-predicts (curve below diagonal). Insets show forecast-probability bin populations.

4.4.2 Probabilistic skill and discrimination

CRPS values lie within approximately 0.2 mm/h between the two methods for four of the five events across all lead times (Fig. 4.11).

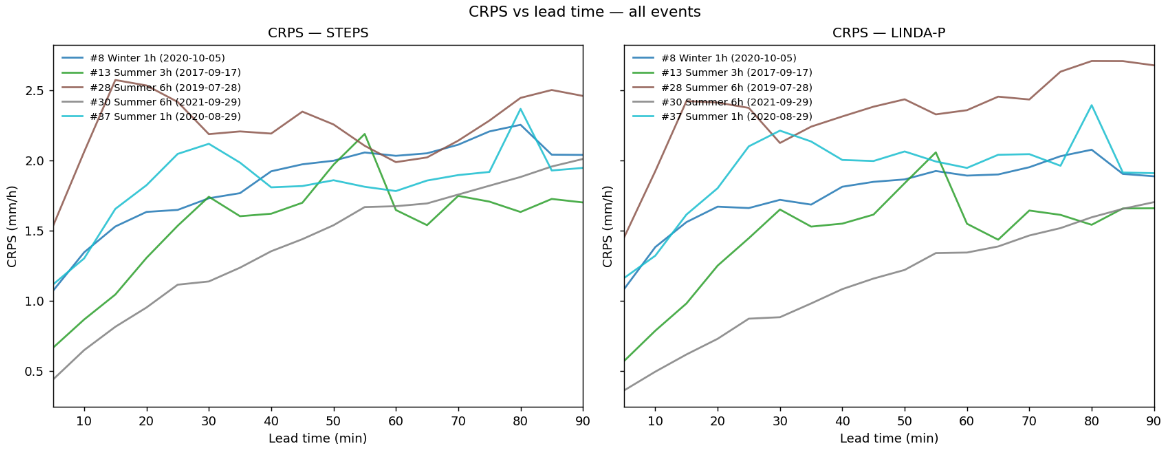

Figure 4.11: CRPS versus lead time per event for STEPS (left) and LINDA-P (right). Values are within approximately 0.2 mm/h between the methods for four of the five events. Event #28 (long-duration weak-skill summer) shows consistently higher CRPS for LINDA-P at lead times beyond 30 min.

The exception is event #28, where LINDA-P CRPS exceeds STEPS by 0.1–0.3 mm/h at lead times beyond 30 min. Aggregated across all five events and all lead times, ROC AUC at the 0.1 mm/h threshold is 0.851 for STEPS and 0.864 for LINDA-P (Fig. 4.12), indicating broadly equivalent discrimination ability between the two methods.

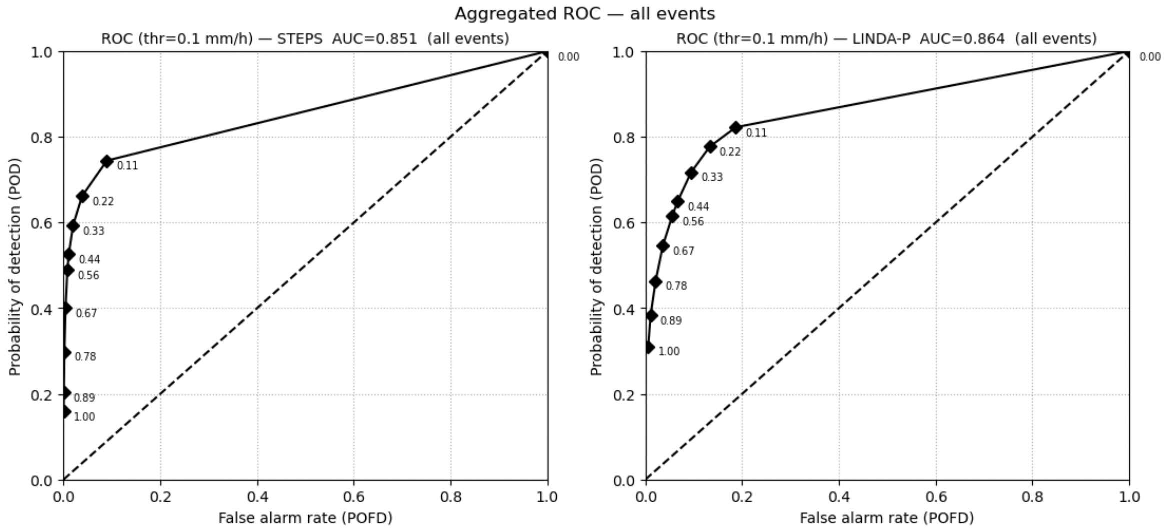

Figure 4.12: Aggregated ROC curves at the 0.1 mm/h threshold for STEPS (left) and LINDA-P (right), combined across all five events and 18 lead times. AUC values are 0.851 (STEPS) and 0.864 (LINDA-P), indicating broadly equivalent discrimination ability.

4.5 Computational cost

Figure 4.13 reports wall-clock computation time for the nowcast generation step, excluding data loading and motion estimation.

The deterministic methods span a factor of approximately 35 in cost: S-PROG completes in 5 s per forecast and ANVIL in 13 s, while LINDA-D requires 183 s - the highest single-realisation cost among the methods evaluated. The probabilistic methods, generating 16 ensemble members each, scale approximately linearly with member count from their underlying single-realization cost: STEPS at 106 s ($\sim$7 s member^{-1}) and LINDA-P at 842 s ($\sim$53 s member^{-1}). All five methods, except LINDA-P complete within a single 5 min forecast cycle on the hardware used.

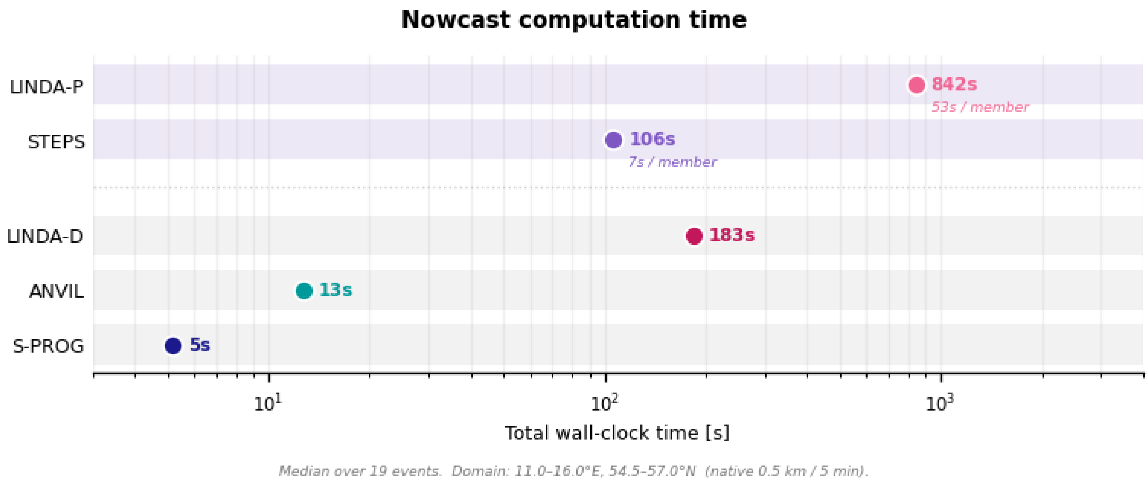

Figure 4.13: Wall-clock computation time for the nowcast generation step, excluding data loading and motion estimation. Deterministic methods (S-PROG, ANVIL, LINDA-D) produce a single forecast field; probabilistic methods (STEPS, LINDA-P) generate 16 ensemble members each, with per-member times shown in italics. Values are medians over 19 events at 0.5 km / 5 min resolution ODIM radar composite, cropped to 11.0–16.0°E and 54.5–57.0°N, with 18 lead times (90 min).

5 Discussion

The verification results characterise the skill of four optical flow-based nowcasting configurations (plus EP) across 19 precipitation events in the Malmö area, evaluated at lead times relevant for forecast-based control (FBC) of rainwater harvesting (RWH) systems. This chapter interprets those results in light of the three research questions: how skill varies with lead time, event duration, and season (Section 5.1); whether higher radar resolution provides additional skill relative to the extrapolation part of the operational KNEP configuration (Section 5.2); and under what conditions nowcast skill is sufficient to justify operational pre-storm releases (Section 5.5).

5.1 Event characteristics and forecast skill

This section addresses the first research question, RQ(i): how nowcast skill varies with lead time, event duration, and season. The comparison against the Eulerian and Lagrangian persistence baselines that RQ(i) also calls for is developed in Section **??**, where the contribution of motion estimation is isolated.

Nowcast skill depends strongly on season and event duration. Winter MAE is consistently lower than summer MAE across all duration–lead-time combinations (Fig. 4.6), but the magnitude of the seasonal contrast varies sharply with duration: at the 0.5 km baseline and 30 min lead time, winter values are a factor of ∼3–4 lower than summer for 1 h events, ∼2.5–3 lower for 6 h events, and as much as a factor of seven lower for 3 h events (e.g. LP: 4.53 vs. 0.60 mm h^{-1}).

Maximum skilful lead time, τ, exhibits a corresponding seasonal pattern but with a strong resolution dependence that we return to below. At 0.5 km/5 min, median τ for the advection-based methods falls to roughly 12–17 min for 1 h summer events, the steepest decay observed in the sample, while 3 h winter events reach 22–47 min and approach the 50 min mark only at coarser resolutions (Fig. 4.5). The 1 h summer collapse occurs because the lifetime of individual convective cells becomes comparable to the forecast horizon itself, leaving little coherent structure for Lagrangian extrapolation to propagate. The winter contrast reflects the well-established difference between frontal, stratiform precipitation - whose spatial coherence and slow evolution are captured reasonably well by advection - and convective summer events, whose growth and decay lie outside what advection-based methods can represent (Germann et al. 2002; Prudden et al. 2020).

Comparison with published values requires care because τ is strongly resolution-dependent in our results. At 0.5 km/5 min, our maximum skilful lead times are shorter than those reported by Tounsi et al. (2023) for the New York City area (25, 45, and 56 min for 2-, 4-, and 6 h events at 1 km/10 min) and shorter than the values in Imhoff et al. (2020) for Dutch catchments (25, 40, and 55 min for 1-, 3-, and 6 h events at 1 km/5 min). Resampling to 1 km/10 min brings the medians into close agreement with, or slightly above, both references: 1 h τ rises from ∼13 min to 36–50 min, and 3 h τ from ∼27 min to 28–48 min depending on season. Two factors plausibly account for the resolution sensitivity. First, a 0.5 km field contains substantially structure that has no advectable analogue at the next time step, so pixel-level correlation decays faster even when the underlying motion is identical. Second, the verification domain (∼57 ×49 km), while embedded in a larger cast domain that eliminates inflow edge effects, samples a smaller fraction of each convective cell's footprint than the regional domains used in those studies; pixel-level displacement errors therefore weigh more heavily relative to the verified wet area. The implication for KNEP and PySTEPS configurations is that lead-time benchmarks reported in the literature cannot be transferred between resolutions without rescaling.

For FBC, the season-duration error pattern carries an uncomfortable implication. The events for which nowcast skill is highest are those for which the operational consequences of a missed forecast are smallest: long-duration winter events rarely overwhelm building-scale storage. The short, intense convective summer events that pose the greatest overflow risk are precisely those where forecast skill decays most rapidly. This is a property of the underlying predictability rather than a limitation of the methods, and it implies that threshold-based release rules should not assume uniform forecast quality across the seasonal cycle (Section 5.5).

In direct answer to the first part of RQ(i), skilful lead time is governed jointly by season and duration: it ranges from roughly 12–17 min for convective 1 h summer events - the steepest decay in the sample - to 22–47 min for stratiform 3 h winter events.

5.2 Resolution dependence: S-PROG versus ANVIL

S-PROG and ANVIL seem to change ranks depending on input resolution. At 0.5 km / 5 min (the native PySTEPS resolution), S-PROG has the lowest MAE and the best categorical and spatial scores, and its bias stays close to one (Fig. 4.4). At 2 km / 15 min (the resolution used in KNEP), the order flips: ANVIL reaches clearly higher POD, CSI, and FSS scores and keeps its bias near one at short lead times - where S-PROG is already missing about half of the rain area (Fig. 4.8).

The bias panel shows where the problem comes from. S-PROG's frequency bias drops to around 0.4–0.6 already at the first lead time, depending on season and event duration. This early collapse points to the cascade–AR machinery itself, rather than to gradual information loss with increasing lead time. The likely reason is the following. At 0.5 km / 5 min), the smallest cascade scales capture sub-cell variability, which is what the AR damping is designed to suppress. At 2 km / 15 min, with the same number of cascade levels (eight in all configurations), larger coherent precipitation features now fall into the lowest cascade scales. The AR process then pulls them toward the near-zero domain mean, so rain area itself is damped away rather than only small-scale noise (Fig. 5.1). ANVIL estimates its AR parameters locally in space (Pulkkinen et al. 2020; Pulkkinen et al. 2021), so it keeps regional growth and decay information instead of relying on cascade-mean reversion, and is therefore less sensitive to coarsening. The slightly lower MAE and FAR that S-PROG keeps at coarse resolution follow from the same effect: a model that forecasts less rain has smaller absolute errors over dry pixels, and fewer positive forecasts that can be counted as false alarms.

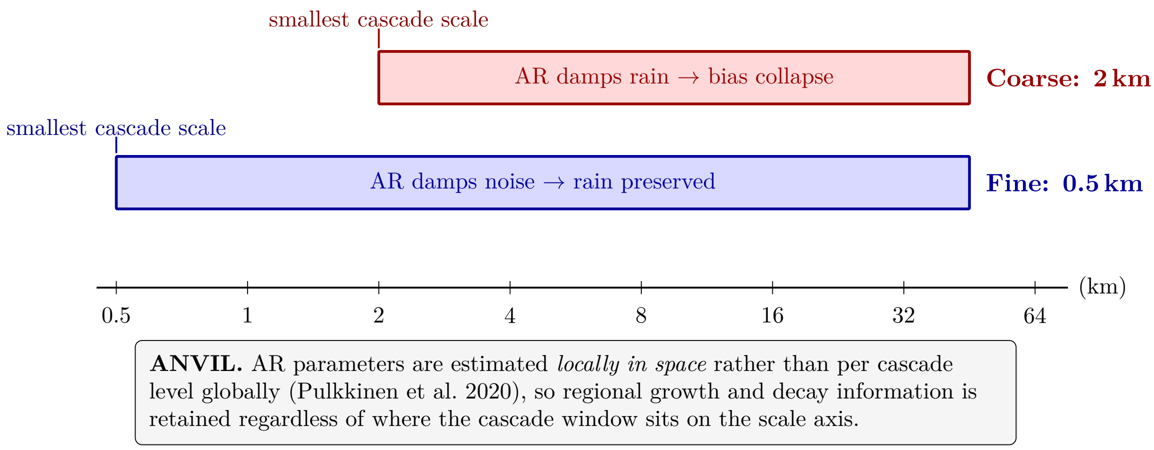

Figure 5.1: Schematic of why S-PROG and ANVIL change ranks with input resolution. With a fixed number of cascade levels, the smallest cascade scale shifts with the input grid: at 0.5 km it falls in the noise regime, which AR damping is designed to suppress, while at 2 km it falls within the scales of coherent precipitation features, which S-PROG's global AR then damps toward the domain mean - producing the early bias collapse. ANVIL's spatially local AR estimation avoids this effect.

For RQ(ii), the main resolution effect is not the finer pixels themselves, but the interaction between input resolution and the choice of nowcasting method best suited to each configuration. SMHI's choice of ANVIL for KNEP at 2 km / 15 min is well supported by these results, and a future operational service based on the higher-resolution radar product cannot simply use the same method without re-examining this trade-off.

5.3 The role of motion estimation

This section completes the answer to RQ(i) by isolating how the configurations compare against the persistence baselines. The Eulerian and Lagrangian persistence runs are not merely lower bounds: the gap between them quantifies how much of the available skill comes from motion estimation alone.

At native resolution (0.5 km / 5 min) LP tracks S-PROG closely in categorical and spatial scores, though with higher MAE and shorter decorrelation times (Fig. 4.4). EP decorrelates within roughly 5 min and serves as a clear lower bound for the advection-based methods at this resolution. ANVIL and LINDA-D drop close to EP for winter events at 0.5 km - consistent with the resolution sensitivity discussed in Section 5.2.

Two implications follow. First, most of the forecast value from the more elaborate configurations comes from the motion field itself; the cascade–AR machinery contributes mainly through bias correction and intensity calibration, with a smaller additional gain in temporal coherence. Second, the gap between EP and LP confirms that even a simple advection scheme is needed to extract the available predictability - motion estimation is a prerequisite, not an optional refinement, for any nowcast intended to support pre-storm release decisions.

Operationally, the small gap between LP and S-PROG suggests that simpler implementations may carry much of the practical value of more sophisticated configurations. This is relevant for FBC adoption at the building or neighbourhood scale, where computational and maintenance resources may be limited.

5.4 Probabilistic forecasting for FBC

Although exploratory, the probabilistic comparison speaks directly to RQ(iii). For forecast-based control the operative question is not point accuracy but whether ensemble probabilities can be trusted to trigger a release rule; calibration, not ranking, is therefore the property that determines operational usability.

The probabilistic comparison rests on five events with a single fixed initialisation per event and is therefore exploratory; it nonetheless shows patterns relevant for FBC.

First, STEPS produces a near-uniform rank histogram with mild excess in the highest bin and a reliability curve slightly above the diagonal (Fig. 4.10), indicating modest under-prediction at the upper tail. LINDA-P shows a pronounced spike of 27 % in the lowest rank bin (present in every event, ranging 20–34 %) and a reliability curve consistently below the diagonal, indicating systematic wet bias. Despite this calibration difference, the two methods discriminate rain-producing from dry pixels comparably (ROC AUC 0.851 versus 0.864; Fig. 4.12). The biases reside in the mapping from forecast probability to observed frequency, not in the underlying ranking.

For FBC this distinction matters more than headline skill scores. A release rule of the form "release stored water when the ensemble probability of exceeding θ over the next T minutes exceeds $p^\star$" depends directly on probability calibration: a wet-biased ensemble triggers releases at lower true exceedance probabilities than $p^\star$ implies, leading to systematically more unnecessary discharges. STEPS probabilities can plausibly be used directly with mild adjustment; the LINDA-P probabilities evaluated here would require calibration. This puts emphasis on the broader point made by Courdent et al. (2015) and Biondi et al. (2021) that the operational value of ensemble nowcasting lies in enabling probabilistic decision rules whose missed-event/false-alarm trade-off can be tuned to system characteristics, rather than in delivering improved point accuracy.

5.5 Implications for forecast-based control

Translating meteorological skill into operational implications for FBC of RWH systems requires a shift in perspective: what matters is not absolute accuracy but whether the forecast crosses the right threshold, at the right time, with the right confidence to justify a pre-storm release. Three skill properties matter most. The first is detection capability at intensities relevant for tank inflow, captured by POD at appropriate thresholds. The second is frequency bias close to unity - a model that systematically under-predicts rain area will systematically miss the release window. The third is lead time exceeding the tank's drawdown time, which depends on system geometry and the chosen operational strategy; this is informed here by emptying-time simulations from a parallel MSc thesis conducted within the same overarching project (Sundstedt 2026). We draw on those results in the comparison that follows.

Two operational strategies are considered there: outflow continued during the rain event (OFD), in which the controlled outlet remains open through the storm, and no outflow during the rain event (NOD), in which the outlet is closed at event onset so that the full tank volume must be available beforehand. Required emptying time, T_{req}, under NOD are therefore necessarily larger than under OFD for any given outflow constraint.

On the first two criteria, the results suggest a realistic operational window of approximately 30–60 min for convective summer events. For winter events, the slower observed decorrelation supports extending this window toward 90 min, although direct verification in this study was limited to 60 min lead time. S-PROG at native resolution maintains POD above 0.7 and bias above 0.85 throughout this window; ANVIL at the coarser KNEP-equivalent configuration delivers comparable skill at 30 min, though with faster degradation in both POD and bias as lead time approaches 60 min. Either represents a reasonable starting point for FBC implementation in the Malmö area.

Comparing this skilful window against the required emptying times in Table 5.1 sharpens the conditional answer to RQ(iii). Sundstedt (2026) simulated emptying times for three outflow constraints - pre-development runoff, urban sewer capacity (Godset), and rural sewer capacity (Ale) - each under two strategies that either continue (OFD) or halt (NOD) outflow during the event. *Pre-dev.* constrains the outflow to not exceed pre-development runoff; *Godset* and *Ale* use local sewer capacity in a densely and less densely populated area, respectively, as the outflow constraint. Constraints are calibrated to a reference RWH installation with 600 m^2 roof catchment and 70 m^3 tank storage (initially full), yielding allowable outflows of 1.37, 7.81, and 2.66 l/s for the pre-development, Godset, and Ale cases respectively.

The figures presented here represent an upper bound on T_{req} in the sense that the reference tank is assumed initially full; under a continuously operated FBC system, partial pre-emptying from earlier forecasts would reduce the actual lead time required at any given event onset. The urban constraint produces T_{req} within the 90 min horizon under both strategies across the full range of return periods considered, with values remaining below 18 min under OFD and below 60 min under NOD. The rural constraint is compatible with the nowcast horizon under OFD up to 10-year return periods, but under NOD exceeds 90 min from 5-year return periods at 60 min duration upward, with T_{req} reaching 163–176 min at 10-year return periods for 60 and 180 min durations. Constraining outflow to pre-development runoff is feasible under nowcast skill alone for return periods of 1 and 2 years; at 5 years T_{req} reaches 75–80 min for 30–60 min design durations, and up to 87 min for the 180 min duration; at 10 years it exceeds the horizon (107–119 min across the design durations considered).

Table 5.1: Required tank emptying time before a rain event, T_{req} (min), for three outflow capacity cases and (where applicable) two operational strategies (*OFD* = outflow continued during rain event; *NOD* = no outflow during rain event). Values transcribed from (Sundstedt 2026) for the subset of return periods and durations matching the event climatology used in this study. Ranges within cells reflect variation across design rain types, i.e. peak position. *Pre-dev.* constrains the outflow to not exceed pre-development runoff; *Godset* and *Ale* use local sewer capacity in a densely and less densely populated area, respectively, as the outflow constraint. Constraints are calibrated to a reference RWH installation with 600 m^2 roof catchment and 70 m^3 tank storage (initially full), yielding allowable outflows of 1.37, 7.81, and 2.66 l/s for the pre-development, Godset, and Ale cases respectively.

		Pre-dev.	Godset (urban)		Ale (rural)	
Return (yr)	Dur. (min)	OFD (min)	OFD (min)	NOD (min)	OFD (min)	NOD (min)
1	30	35–37	0–4	17–18	21–25	50–51
1	60	22–40	0–2	19–22	4–28	60–63
1	180	0–39	0	24–28	0–23	82–86
2	30	51–52	0–6	21–22	33–35	62–63
2	60	41–54	0–6	25–27	19–39	76–78
2	180	0–57	0	32–35	0–37	103–106
5	30	80	0–11	28–29	55–56	84–85
5	60	75–80	0–13	35–36	45–57	103–105
5	180	5–87	0–3	45–49	0–59	138–142
10	30	107–108	7–15	36	76–78	106
10	60	109–110	0–18	44–45	71–76	163–164
10	180	49–119	0–9	57–60	0–82	173–176

The season-duration stratification carries a direct implication for control rule design that has not, to the author's knowledge, been addressed in the existing FBC-RWH literature. Convective summer events combine the highest peak intensities with the shortest skilful lead times: a tank near capacity at the onset of a 1 h summer cell can overflow within minutes. For these events, a permissive release threshold may be operationally justified despite the higher false-alarm rate. Winter events, with longer skilful lead times and lower peak intensities, can tolerate a more conservative threshold without material loss of detention performance. A control strategy applying a single fixed threshold across the seasonal cycle, as is typical in existing FBC-RWH studies that assume perfect or season-independent forecasts (Xu et al. 2020b; Sun et al. 2023; Zhou et al. 2024), would be sub-optimal in both directions. The probabilistic framing discussed in Section 5.4 provides the natural mechanism for implementing such season-conditional rules.

Two limits on this interpretation should be stated. The operational window derived above describes meteorological skill only and does not account for valve actuation time or physical drawdown dynamics, both of which subtract from the lead time available to the controller. And the emptying times in Table 5.1 describe a single reference installation (600 m^2 roof catchment, 70 m^3 tank, initially full) under three outflow constraints scaled from representative downstream pipe capacities; results for installations of different roof area, tank volume, or initial fill level may differ, and a coupled hydraulic–control simulation across a representative ensemble of RWH installations in Malmö is suggested for future work.

With respect to RQ(iii), the conditions under which nowcast skill is sufficient to justify operational pre-storm releases are event-type and strategy dependent rather than uniform. Deterministic skill at 30–60 min lead time supports FBC for short convective summer events under operational strategies that maintain outflow during rain (OFD) and use urban or moderate-capacity sewer constraints; the full 90 min horizon extends this coverage to longer events and to the rural constraint under OFD across the return-period range considered. For NOD strategies under strict outflow constraints, and for the pre-development constraint at high return periods, the meteorological lead time becomes the binding constraint, and operational viability requires extending the forecast horizon through NWP blending.

5.6 Limitations

The event sample is modest: 19 events distributed across three duration classes and two seasons yields only a small number of cases per season–duration group, and individual events therefore have substantial influence on the stratified statistics. Aggregated metrics across all 19 events are more robust. The seasonal and duration patterns interpreted in Section 5.1 should be treated as indicative rather than statistically established at the per-cell level: several stratification cells contain only three or four events (e.g. $n = 2$ for 6 h winter events), and confidence intervals on the median Pearson r for these cells span much of the y-axis range in Figure 4.5.

Verification was performed against the same radar composites used as nowcast input, introducing a degree of circularity: systematic radar artefacts at initialisation will also appear in the verification field. Gauge-based verification was not pursued because the limited number of gauge locations provides insufficient spatial coverage for the FSS and categorical metrics, and because gauge–radar discrepancies introduce additional scale-mismatch artefacts. The relative ranking of methods is unlikely to be affected, but absolute skill values may be optimistic.

Reflectivity was converted to rainfall rate using fixed Marshall–Palmer parameters ($a = 200$, $b = 1.6$). For relative comparison this is unproblematic since the same conversion is applied across all configurations, but absolute MAE and bias values could shift meaningfully under a different choice, particularly one calibrated to the local meteoreology

Gauge-radar correspondence was verified visually in the BALTRAD GUI for each event but not quantified. A systematic gauge-radar comparison would require addressing scale mismatch between point gauges and the 0.5 km radar grid cell, and was outside the scope of this work. The absolute calibration of radar-derived rain rates against the gauge climatology used for event selection therefore remains unverified, although the relative ranking of nowcasting methods is unaffected by any uniform Z–R offset.

The probabilistic comparison was deliberately exploratory: five events with a single fixed initialisation per event, and therefore not directly comparable to the deterministic verification. The calibration patterns identified are sufficiently pronounced to be unlikely artefacts of the small sample, but their generalisability remains to be established.

Finally, NWP blending was not evaluated, despite occupying a central place in the theoretical background (Section 2.5.3). The deterministic results are best understood as characterising the radar-only contribution to a forecast that, operationally, would be supplemented by NWP at longer lead times; implementing and testing such blending is identified as a priority for future work.

5.7 Conclusions

This thesis evaluated four `PySTEPS` nowcasting configurations (plus EP) against radar composites for 19 precipitation events in the Malmö area, with verification framed for the requirements of forecast-based control (FBC) of rainwater harvesting (RWH) systems. The three research questions can now be answered in turn.

RQ(i): Variation of skill with lead time, duration, and season, and comparison with persistence. PySTEPS skill varies systematically with season and event duration. Winter events sustain useful Pearson correlation ($r > 1/e$) of 30–50 min depending on duration and resolution; convective summer events of 3 h duration decorrelate within 25–30 min, and 1 h summer events within approximately 15 min. The seasonal MAE contrast spans a factor of three to seven depending on duration. These skilful lead times align with those reported in larger-domain studies (Imhoff et al. 2020; Tounsi et al. 2023). Against the persistence baselines, most of the forecast value comes from the motion field itself: Lagrangian persistence already captures the majority of the gain over the Eulerian baseline, which decorrelates within roughly 5 min, while the cascade–AR machinery contributes mainly bias correction and intensity calibration.

RQ(ii): Additional skill from higher resolution. The headline finding on resolution is the inversion of method ranking between configurations. S-PROG produces the lowest MAE and highest spatial scores at native 0.5 km / 5 min, while ANVIL retains bias close to unity and higher categorical and spatial scores at the coarse 2 km / 15 min configuration used by KNEP, where S-PROG already under-predicts rain area by approximately a factor of two. SMHI's choice of ANVIL for the operational KNEP product is well supported by these results, and the corollary is that resolution and method must be co-selected: the coarse-resolution preference for ANVIL does not automatically transfer to a future native-resolution service.

RQ(iii): Conditions for sufficient skill. Drawing on emptying-time simulations from a parallel MSc thesis (Sundstedt 2026), the viability of nowcast-driven FBC is conditional rather than uniform. Within these bounds, S-PROG at native resolution and ANVIL at the coarser KNEP-equivalent configuration each maintain POD above

0.7 and bias close to unity over operationally relevant lead times. Requested nowcast skill sits at lead times of 30–90 min, which is sufficient under operational strategies that maintain outflow during the rain event combined with urban-scale sewer capacity, across the full return-period range, and also sufficient for the rural capacity case under continued outflow up to 10-year return periods. Skill is insufficient, motivating extension via NWP blending, for strategies requiring a fully empty tank at event onset under strict outflow constraints, and for the pre-development constraint at return periods of 10 years and above. Operational viability is therefore as much a function of control strategy and outflow constraint as of meteorological skill, provided that the season–duration stratification of skill is reflected in any control rule design.

As a concrete operational recommendation, an FBC implementation in southern Sweden should pair the method to the available radar resolution: S-PROG where the native 0.5 km / 5 min product is available, and ANVIL where the operational 2 km / 15 min KNEP configuration is used. Both maintain detection and frequency-bias skill adequate for pre-storm release decisions over the 30–90 min horizon, provided the release threshold is allowed to vary with season and event duration.

5.7.1 Recommended future work

The most consequential extension for operational FBC is integration of nowcasting with NWP. The skilful lead times established here (15–90 min depending on event type) cover only the lower end of operational relevance; several control strategies require horizons of 100–180 min, within the canonical nowcasting–NWP blending window. Implementation within the PySTEPS framework using a high-resolution NWP product is the obvious next step, and would permit re-examination of the operational KNEP configuration in light of genuine blended skill rather than the NaN-filling currently used to extend ANVIL fields beyond the radar horizon.

A coupled simulation chain - hydrological loading, hydraulic tank model, controller, forecast input - would enable end-to-end propagation of forecast uncertainty into operationally meaningful regret metrics (overflow volume, retained yield, missed releases) and direct comparison against rule-based controllers using gauge observations alone. Such a chain would accommodate valve actuation and physical drawdown dynamics excluded from the present analysis and could be tailored to specific RWH installations rather than generic configurations.

Beyond simulation, real-world deployment remains the ultimate test. A pilot at an instrumented RWH site in the Malmö–Skåne region, ingesting either the operational KNEP product or a higher resolution nowcast in real time and triggering pre-storm releases, would surface the practical constraints - latency, missing data, operator overrides, regulatory compliance - that neither verification studies nor coupled simulations capture, closing the loop between the meteorological work presented here and the operational question that motivated it.

Although this study addresses RWH tanks specifically, the framework generalises to other sustainable drainage systems (SuDS) along two axes: whether the system has a controllable element that a forecast can drive, and its response time, which sets the required lead time. Controllable detention ponds occupy the same axis as RWH tanks, differing mainly in a longer required emptying time and therefore a greater reliance on NWP blending; passive features such as swales and green roofs have no actuator to drive, so a nowcast informs their design rather than their real-time operation. The skill-versus-lead-time results presented here are thus transferable to any actively controlled storage, while the operational thresholds remain specific to the system geometry.

"The end of labor is to gain leisure."

- Aristotle

References

An, Dong, Jonas Olsson, Yiheng Du, Johanna Sörensen, Cintia B. Uvo, and Peter Berg (Jan. 2026). "Future projections of wet and dry spells in southern Sweden: The impact of climate model resolution". In: *Atmospheric Research* 330, p. 108547. DOI: 10.1016/j.atmosres.2025.108547. URL: http://dx.doi.org/10.1016/j.atmosres.2025.108547.

Andersson, Jafet C. M., Jonas Olsson, Remco (C. Z.) van de Beek, and Jonas Hansryd (Dec. 2022). "OpenMRG: Open data from Microwave links, Radar, and Gauges for rainfall quantification in Gothenburg, Sweden". In: *Earth System Science Data* 14.12, pp. 5411–5426. DOI: 10.5194/essd-14-5411-2022. URL: http://dx.doi.org/10.5194/essd-14-5411-2022.

Annella, Clizia, Edoardo Raparelli, Gianfranco Vulpiani, Vincenzo Capozzi, Elisa Adirosi, Luca Baldini, Giorgio Budillon, Raffaele Lidori, and Mario Montopoli (2025). "A Strategy to set up Test Dataset and Evaluation Benchmark for Radar Nowcasting of Precipitation in Italy". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 18, pp. 20240–20261. DOI: 10.1109/jstars.2025.3568185. URL: http://dx.doi.org/10.1109/JSTARS.2025.3568185.

BALTRAD (2026). *BALTRAD Maps Viewer*. URL: https://se.baltrad.eu/ (visited on 04/03/2026).

Bauer, Peter, Alan Thorpe, and Gilbert Brunet (Sept. 2015). "The quiet revolution of numerical weather prediction". In: *Nature* 525.7567, pp. 47–55. DOI: 10.1038/nature14956. URL: http://dx.doi.org/10.1038/nature14956.

Beckers, Brian, Jonas Berking, and Brigitta Schütt (Sept. 2013). "Ancient Water Harvesting Methods in the Drylands of the Mediterranean and Western Asia". In: *eTopoi. Journal for Ancient Studies* 2.

Bengtsson, Lisa, Ulf Andrae, Trygve Aspelien, Yurii Batrak, Javier Calvo, Wim de Rooy, Emily Gleeson, Bent Hansen-Sass, Mariken Homleid, Mariano Hortal, Karl-Ivar Ivarsson, Geert Lenderink, Sami Niemelä, Kristian Pagh Nielsen, Jeanette Onvlee, Laura Rontu, Patrick Samuelsson, Daniel Santos Muñoz, Alvaro Subias, Sander Tijm, Velle Toll, Xiaohua Yang, and Morten Ødegaard Køltzow (Apr. 2017). "The HARMONIE–AROME Model Configuration in the ALADIN–HIRLAM NWP System". In: *Monthly Weather Review* 145.5, pp. 1919–1935. DOI: 10.1175/mwr-d-16-0417.1. URL: http://dx.doi.org/10.1175/MWR-D-16-0417.1.

Biondi, Daniela, Ezio Todini, and Angela Corina (Oct. 2021). "A parsimonious post-processor for uncertainty evaluation of ensemble precipitation forecasts: an application to quantitative precipitation forecasts for civil protection purposes". In: *Hydrology Research* 52.6, pp. 1405–1422. DOI: 10.2166/nh.2021.045. URL: http://dx.doi.org/10.2166/nh.2021.045.

Björck, Johan and Margaret McNamee (Jan. 2025). "Evaluation of Wildland Fire Detection Methods Using Expert Input". In: *Fire Technology* 61.4, pp. 2547–2569. DOI: 10.1007/s10694-024-01696-5. URL: http://dx.doi.org/10.1007/s10694-024-01696-5.

Bowler, Neill E., Clive E. Pierce, and Alan W. Seed (Oct. 2006). "STEPS: A probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP". In: *Quarterly Journal of the Royal Meteorological Society* 132.620, pp. 2127–2155. DOI: 10.1256/qj.04.100. URL: http://dx.doi.org/10.1256/qj.04.100.

Browning, K. A. (1981). "Forward". In: *Nowcasting: Mesoscale Observations and Short-range Prediction*. Ed. by B. Battrick and J. Mort. Vol. 16. ESA SP. Proceedings of an International Symposium, Hamburg, Germany, 25–28 August. European Space Agency.

Bruins, H.J., M. Evenari, and U. Nessler (Jan. 1986). "Rainwater-harvesting agriculture for food production in arid zones: the challenge of the African famine". In: *Applied Geography* 6.1, pp. 13–32. DOI: 10.1016/0143-6228(86)90026-3. URL: http://dx.doi.org/10.1016/0143-6228(86)90026-3.

Campisano, Alberto, David Butler, Sarah Ward, Matthew J. Burns, Eran Friedler, Kathy DeBusk, Lloyd N. Fisher-Jeffes, Enedir Ghisi, Ataur Rahman, Hiroaki Furumai, and Mooyoung Han (May 2017). "Urban rainwater harvesting systems: Research, implementation and future perspectives". In: *Water Research* 115, pp. 195–209. DOI: 10.1016/j.watres.2017.02.056. URL: http://dx.doi.org/10.1016/j.watres.2017.02.056.

Canedo Rosso, Claudia, Lars Nyberg, and Ilias Pechlivanidis (Nov. 2025). "Drought hazard assessment across Sweden's diverse hydro-climatic regimes". In: *Natural Hazards and Earth System Sciences* 25.11, pp. 4577–4592. DOI: 10.5194/nhess-25-4577-2025. URL: http://dx.doi.org/10.5194/nhess-25-4577-2025.

Cao, Zuohao and Da-Lin Zhang (Feb. 2016). "Analysis of Missed Summer Severe Rainfall Forecasts". In: *Weather and Forecasting* 31.2, pp. 433–450. DOI: 10.1175/waf-d-15-0119.1. URL: http://dx.doi.org/10.1175/WAF-D-15-0119.1.

Cedergren, Elin, Hilma Salonen, and Lumi Tomren (2025). *Rethinking Population Shrinkage: Smart Adaptation for Nordic Municipalities and Regions*. Tech. rep. Nordregio Policy Brief 2025:19. Stockholm: Nordregio. URL: https://pub.nordregio.org/pb-2025-19-rethinking-population-shrinkage.

Courdent, Vianney, Luca Vezzaro, Peter Steen Mikkelsen, Ane Loft Mollerup, and Morten Grum (Apr. 2015). "Using ensemble weather forecast in a risk based real time optimization of urban drainage systems". In: *La Houille Blanche* 101.2, pp. 101–107. DOI: 10.1051/lhb/20150025. URL: http://dx.doi.org/10.1051/lhb/20150025.

Dahlström, Bengt (2010). *Regnintensitet – En molnfysikalisk betraktelse*. Tech. rep. 2010-05. Svenskt Vatten AB. URL: http://vav.griffel.net/filer/Rapport_2010-05.pdf.

De Luca, Davide Luciano, Francesco Napolitano, Dongkyun Kim, Christian Onof, Daniela Biondi, Li-Pen Wang, Fabio Russo, Elena Ridolfi, Benedetta Moccia, and Flavia Marconi (May 2025). "Rainfall nowcasting models: state of the art and possible future perspectives". In: *Hydrological Sciences Journal* 70.9, pp. 1419–1438. DOI: 10.1080/02626667.2025.2490780. URL: http://dx.doi.org/10.1080/02626667.2025.2490780.

Degadwala, Sheshang and Vedant Sukhadia (2024). “Advances and Challenges in Weather Nowcasting – A Comprehensive Review of Modern Techniques and Models”. In: *International Journal of Scientific Research in Computer Science Engineering and Information Technology* 10.5, pp. 113–126. doi: 10.32628/CSEIT24105810. url: https://doi.org/10.32628/CSEIT24105810.

Delleur, J. W. (Jan. 1982). *Introduction to Urban Hydrology and Stormwater Management.* doi: 10.1029/wm007p0001. url: http://dx.doi.org/10.1029/WM007p0001.

Falahat, Saeed (2026). *Personal communication.* Personal communication, SMHI, April 2026.

Feng, Boyu, Ying Zhang, and Robin Bourke (Jan. 2021). “Urbanization impacts on flood risks based on urban growth data and coupled flood models”. In: *Natural Hazards* 106.1, pp. 613–627. doi: 10.1007/s11069-020-04480-0. url: http://dx.doi.org/10.1007/s11069-020-04480-0.

Germann, Urs and Isztar Zawadzki (Dec. 2002). “Scale-Dependence of the Predictability of Precipitation from Continental Radar Images. Part I: Description of the Methodology”. In: *Monthly Weather Review* 130.12, pp. 2859–2873. doi: 10.1175/1520-0493(2002)130<2859:sdotpo>2.0.co;2. url: http://dx.doi.org/10.1175/1520-0493(2002)130%3C2859:SDOTPO%3E2.0.CO;2.

Gimenez-Maranges, Marc, Jürgen Breuste, and Angela Hof (May 2020). “Sustainable Drainage Systems for transitioning to sustainable urban flood management in the European Union: A review”. In: *Journal of Cleaner Production* 255, p. 120191. doi: 10.1016/j.jclepro.2020.120191. url: http://dx.doi.org/10.1016/j.jclepro.2020.120191.

Golding, B W (Mar. 1998). “Nimrod: a system for generating automated very short range forecasts”. In: *Meteorological Applications* 5.1, pp. 1–16. doi: 10.1017/s1350482798000577. url: http://dx.doi.org/10.1017/S1350482798000577.

Guerova, Guergana and Tzvetan Simeonov (2022). “GNSS and numerical weather prediction models”. In: *Global Navigation Satellite System Monitoring of the Atmosphere.* Elsevier, pp. 65–84. doi: 10.1016/b978-0-12-819152-1.00006-4. url: http://dx.doi.org/10.1016/B978-0-12-819152-1.00006-4.

Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber (Apr. 2011). “The Integrated Nowcasting through Comprehensive Analysis (INCA) System and Its Validation over the Eastern Alpine Region”. In: *Weather and Forecasting* 26.2, pp. 166–183. doi: 10.1175/2010waf2222451.1. url: http://dx.doi.org/10.1175/2010WAF2222451.1.

Hamill, Thomas M. (Mar. 2001). “Interpretation of Rank Histograms for Verifying Ensemble Forecasts”. In: *Monthly Weather Review* 129.3, pp. 550–560. doi: 10.1175/1520-0493(2001)129<0550:iorhfv>2.0.co;2. url: http://dx.doi.org/10.1175/1520-0493(2001)129%3C0550:IORHFV%3E2.0.CO;2.

Hohenegger, Cathy and Christoph Schar (Nov. 2007). “Atmospheric Predictability at Synoptic Versus Cloud-Resolving Scales”. In: *Bulletin of the American Meteorological Society* 88.11, pp. 1783–1794. doi: 10.1175/bams-88-11-1783. url: http://dx.doi.org/10.1175/BAMS-88-11-1783.

Horn, Berthold K.P. and Brian G. Schunck (Aug. 1981). "Determining optical flow". In: *Artificial Intelligence* 17.1–3, pp. 185–203. DOI: 10.1016/0004-3702(81)90024-2. URL: http://dx.doi.org/10.1016/0004-3702(81)90024-2.

Hosseini, S. H., H. Hashemi, R. Larsson, R. Berndtsson, A. Bengtsson, S. Bengtsson, S. Dahl, J. Olsson, R. van de Beek, N. South, M. Palo, H. Hammarlund, H. Aspegren, S. Granath, S. Steen Kronborg, and E. Falk (Oct. 2022). "Kombinerade X-bandsanläggningar som väderradar. Jämförelse med etablerade metoder för regnmätning". In.

Hwang, Yunsung, Adam J. Clark, Valliappa Lakshmanan, and Steven E. Koch (Oct. 2015). "Improved Nowcasts by Blending Extrapolation and Model Forecasts". In: *Weather and Forecasting* 30.5, pp. 1201–1217. DOI: 10.1175/waf-d-15-0057.1. URL: http://dx.doi.org/10.1175/WAF-D-15-0057.1.

Imhoff, R. O., C. C. Brauer, A. Overeem, A. H. Weerts, and R. Uijlenhoet (July 2020). "Spatial and Temporal Evaluation of Radar Rainfall Nowcasting Techniques on 1, 533 Events". In: *Water Resources Research* 56.8. DOI: 10.1029/2019wr026723. URL: http://dx.doi.org/10.1029/2019WR026723.

International Water Association (2022). *Dealing with 'Sins of the Past' to Protect our Future Drinking Water*. International Water Association. URL: https://iwa-network.org/dealing-with-sins-of-the-past-to-protect-our-future-drinking-water/ (visited on 03/03/2026).

IPCC (2023). *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. DOI: 10.59327/IPCC/AR6-9789291691647. URL: https://www.ipcc.ch/report/ar6/syr/.

Ivarsson, Adam, Johan Åström, Johan Kjellin, Olof Jonasson, and Godecke-Tobias Blecken (Dec. 2025). "Experiences and sustainability aspects of rainwater harvesting systems in Scandinavia". In: *Water Practice & Technology* 20.12, pp. 2520–2536. DOI: 10.2166/wpt.2025.172. URL: http://dx.doi.org/10.2166/wpt.2025.172.

Kerkez, Branko, Cyndee Gruden, Matthew Lewis, Luis Montestruque, Marcus Quigley, Brandon Wong, Alex Bedig, Ruben Kertesz, Tim Braun, Owen Cadwalader, Aaron Poresky, and Carrie Pak (July 2016). "Smarter Stormwater Systems". In: *Environmental Science & Technology* 50.14, pp. 7267–7273. DOI: 10.1021/acs.est.5b05870. URL: http://dx.doi.org/10.1021/acs.est.5b05870.

Knutsson, Jesper (Apr. 2025). "Real-time control of retention basins for stormwater management and Rainwater Reuse". In: *Urban Water Journal* 22.6, pp. 706–717. DOI: 10.1080/1573062x.2025.2496958. URL: http://dx.doi.org/10.1080/1573062X.2025.2496958.

Krzysztofowicz, Roman (Aug. 2001). "The case for probabilistic forecasting in hydrology". In: *Journal of Hydrology* 249.1–4, pp. 2–9. DOI: 10.1016/s0022-1694(01)00420-6. URL: http://dx.doi.org/10.1016/S0022-1694(01)00420-6.

Li, Dawei, Jingnan Wang, Kefeng Deng, Di Zhang, Chengwu Zhao, Hongze Leng, Yingfang Wen, Yudi Liu, Kaijun Ren, and Junqiang Song (Apr. 2026). "Review on deep learning quantitative precipitation nowcasting: Advances and challenges". In:

Expert Systems with Applications 305, p. 130775. DOI: 10.1016/j.eswa.2025.130775. URL: http://dx.doi.org/10.1016/j.eswa.2025.130775.

Lynch, Peter (Oct. 2015). “Richardson’s Fantastic Forecast Factory”. In: *European Meteorological Society.* Accessed: 2026-03-09. URL: https://www.emetsoc.org/richardsons-fantastic-forecast-factory/.

Marshall, J. S., R. C. Langille, and W. Mc K. Palmer (1947). “Measurement of Rainfall by Radar”. In: *Journal of Atmospheric Sciences* 4.6, pp. 186–192.

Marshall, J. S. and W. Mc K. Palmer (1948). “The Distribution of Raindrops with Size”. In: *Journal of Atmospheric Sciences* 5.4, pp. 165–166.

Michelson, Daniel, Anders Henja, Sander Ernes, Günther Haase, Jarmo Koistinen, Katarzyna Ośródka, Tuomas Peltonen, Maciej Szewczykowski, and Jan Szturc (Mar. 2018). “BALTRAD Advanced Weather Radar Networking”. In: *Journal of Open Research Software* 6.1, p. 12. DOI: 10.5334/jors.193. URL: http://dx.doi.org/10.5334/jors.193.

Misra, Siddharth and Hao Li (2020). “Deep neural network architectures to approximate the fluid-filled pore size distributions of subsurface geological formations”. In: *Machine Learning for Subsurface Characterization.* Elsevier, pp. 183–217. DOI: 10.1016/b978-0-12-817736-5.00007-7. URL: http://dx.doi.org/10.1016/B978-0-12-817736-5.00007-7.

Møller Jess, Rasmus (2025). “Assessing the Potential of Forecast-Controlled Rain and Stormwater Harvesting for Dual Benefits: Flood Mitigation and Water Reuse”. ISSN 1650-6553, Nr 690. Degree Project E1 in Earth Science. Uppsala University, Department of Earth Sciences.

Mullapudi, Abhiram, Matthew Bartos, Brandon Wong, and Branko Kerkez (July 2018). “Shaping Streamflow Using a Real-Time Stormwater Control Network”. In: *Sensors* 18.7, p. 2259. DOI: 10.3390/s18072259. URL: http://dx.doi.org/10.3390/s18072259.

Mun, J.S. and M.Y. Han (Jan. 2012). “Design and operational parameters of a rooftop rainwater harvesting system: Definition, sensitivity and verification”. In: *Journal of Environmental Management* 93.1, pp. 147–153. DOI: 10.1016/j.jenvman.2011.08.024. URL: http://dx.doi.org/10.1016/j.jenvman.2011.08.024.

Myhre, G., K. Alterskjær, C. W. Stjern, Ø. Hodnebrog, L. Marelle, B. H. Samset, J. Sillmann, N. Schaller, E. Fischer, M. Schulz, and A. Stohl (Nov. 2019). “Frequency of extreme precipitation increases extensively with event rareness under global warming”. In: *Scientific Reports* 9.1. DOI: 10.1038/s41598-019-52277-4. URL: http://dx.doi.org/10.1038/s41598-019-52277-4.

Nerini, Daniele, Loris Foresti, Daniel Leuenberger, Sylvain Robert, and Urs Germann (Mar. 2019). “A Reduced-Space Ensemble Kalman Filter Approach for Flow-Dependent Integration of Radar Extrapolation Nowcasts and NWP Precipitation Ensembles”. In: *Monthly Weather Review* 147.3, pp. 987–1006. DOI: 10.1175/mwr-d-18-0258.1. URL: http://dx.doi.org/10.1175/MWR-D-18-0258.1.

Nordh, Ulf (2026). *Personal communication.* Personal communication, SMHI, April 2026.

Oberascher, Martin, Carolina Kinzel, Ulrich Kastlunger, Manfred Kleidorfer, Christoph Zingerle, Wolfgang Rauch, and Robert Sitzenfrei (May 2021). “Integrated urban water management with micro storages developed as an IoT-based solution – The smart rain barrel". In: *Environmental Modelling & Software* 139, p. 105028. DOI: 10.1016/j.envsoft.2021.105028. URL: http://dx.doi.org/10.1016/j.envsoft.2021.105028.

Olsson, Jonas, Yiheng Du, Dong An, Cintia B. Uvo, Johanna Sörensen, Erika Toivonen, Danijel Belušić, and Andreas Dobler (July 2021). "An Analysis of (Sub-)Hourly Rainfall in Convection-Permitting Climate Simulations Over Southern Sweden From a User's Perspective". In: *Frontiers in Earth Science* 9. DOI: 10.3389/feart.2021.681312. URL: http://dx.doi.org/10.3389/feart.2021.681312.

Olsson, Jonas, Lennart Simonsson, and Martin Ridal (Nov. 2013). "Rainfall nowcasting: predictability of short-term extremes in Sweden". In: *Urban Water Journal* 11.7, pp. 605–615. DOI: 10.1080/1573062x.2013.847465. URL: http://dx.doi.org/10.1080/1573062X.2013.847465.

Pierce, Clive, Alan Seed, Sue Ballard, David Simonin, and Zhihong Li (2012). "Nowcasting". In: *Doppler Radar Observations - Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications*. Ed. by Joan Bech. InTech. Chap. 4. URL: http://www.intechopen.com/books/doppler-radar-observations-weather-radar-wind-profiler-ionospheric-radar-and-other-advanced-applications/nowcasting.

Prudden, Rachel, Samantha Adams, Dmitry Kangin, Niall Robinson, Suman Ravuri, Shakir Mohamed, and Alberto Arribas (2020). "A review of radar-based nowcasting of precipitation and applicable machine learning techniques". In: *arXiv preprint*. arXiv:2005.04988. URL: https://arxiv.org/abs/2005.04988.

Pu, Zhaoxia and Eugenia Kalnay (2018). "Numerical Weather Prediction Basics: Models, Numerical Methods, and Data Assimilation". In: *Handbook of Hydrometeorological Ensemble Forecasting*. Springer Berlin Heidelberg, pp. 1–31. DOI: 10.1007/978-3-642-40457-3_11-1. URL: http://dx.doi.org/10.1007/978-3-642-40457-3_11-1.

Pulkkinen, Seppo, V. Chandrasekar, Annakaisa von Lerber, and Ari-Matti Harri (Nov. 2020). "Nowcasting of Convective Rainfall Using Volumetric Radar Observations". In: *IEEE Transactions on Geoscience and Remote Sensing* 58.11, pp. 7845–7859. DOI: 10.1109/tgrs.2020.2984594. URL: http://dx.doi.org/10.1109/TGRS.2020.2984594.

Pulkkinen, Seppo, V. Chandrasekar, and Tero Niemi (Dec. 2021). "Lagrangian Integro-Difference Equation Model for Precipitation Nowcasting". In: *Journal of Atmospheric and Oceanic Technology* 38.12, pp. 2125–2145. DOI: 10.1175/jtech-d-21-0013.1. URL: http://dx.doi.org/10.1175/JTECH-D-21-0013.1.

Pulkkinen, Seppo, Daniele Nerini, Andrés A. Pérez Hortal, Carlos Velasco-Forero, Alan Seed, Urs Germann, and Loris Foresti (Oct. 2019). "Pysteps: an open-source Python library for probabilistic precipitation nowcasting (v1.0)". In: *Geoscientific Model Development* 12.10, pp. 4185–4219. DOI: 10.5194/gmd-12-4185-2019. URL: http://dx.doi.org/10.5194/gmd-12-4185-2019.

Raimondi, Anita, Ruth Quinn, Gopinathan R. Abhijith, Gianfranco Becciu, and Avi Ostfeld (Apr. 2023). "Rainwater Harvesting and Treatment: State of the Art and

Perspectives". In: *Water* 15.8, p. 1518. doi: 10.3390/w15081518. url: http://dx.doi.org/10.3390/w15081518.

Ravuri, Suman, Karel Lenc, Matthew Willson, Dmitry Kangin, Remi Lam, Piotr Mirowski, Megan Fitzsimons, Maria Athanassiadou, Sheleem Kashem, Sam Madge, Rachel Prudden, Amol Mandhane, Aidan Clark, Andrew Brock, Karen Simonyan, Raia Hadsell, Niall Robinson, Ellen Clancy, Alberto Arribas, and Shakir Mohamed (Sept. 2021). "Skilful precipitation nowcasting using deep generative models of radar". In: *Nature* 597.7878, pp. 672–677. doi: 10.1038/s41586-021-03854-z. url: http://dx.doi.org/10.1038/s41586-021-03854-z.

Roberts, Nigel M. and Humphrey W. Lean (Jan. 2008). "Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts of Convective Events". In: *Monthly Weather Review* 136.1, pp. 78–97. doi: 10.1175/2007mwr2123.1. url: http://dx.doi.org/10.1175/2007MWR2123.1.

Santos, C. and F. Taveira-Pinto (Feb. 2013). "Analysis of different criteria to size rainwater storage tanks using detailed methods". In: *Resources, Conservation and Recycling* 71, pp. 1–6. doi: 10.1016/j.resconrec.2012.11.004. url: http://dx.doi.org/10.1016/j.resconrec.2012.11.004.

Schimanke, Semjon, Magnus Joelsson, Sandra Andersson, Thomas Carlund, Lennart Wern, Sverker Hellström, and Erik Kjellström (2022). *Observerad klimatförändring i Sverige 1860–2021.* Tech. rep. 69. Norrköping: SMHI.

Seed, A. W. (Mar. 2003). "A Dynamic and Spatial Scaling Approach to Advection Forecasting". In: *Journal of Applied Meteorology* 42.3, pp. 381–388. doi: 10.1175/1520-0450(2003)042<0381:adassa>2.0.co;2. url: http://dx.doi.org/10.1175/1520-0450(2003)042%3C0381:ADASSA%3E2.0.CO;2.

Seed, Alan W., Clive E. Pierce, and Katie Norman (Oct. 2013). "Formulation and evaluation of a scale decomposition-based stochastic precipitation nowcast scheme: Formulation of a Scale Decomposition Nowcast Scheme". In: *Water Resources Research* 49.10, pp. 6624–6641. doi: 10.1002/wrcr.20536. url: http://dx.doi.org/10.1002/wrcr.20536.

Semaan, Mary, Susan D. Day, Michael Garvin, Naren Ramakrishnan, and Annie Pearce (May 2020). "Optimal sizing of rainwater harvesting systems for domestic water usages: A systematic literature review". In: *Resources, Conservation & Recycling: X* 6, p. 100033. doi: 10.1016/j.rcrx.2020.100033. url: http://dx.doi.org/10.1016/j.rcrx.2020.100033.

Shishegar, Shadab, Sophie Duchesne, Geneviève Pelletier, and Reza Ghorbani (Jan. 2021). "A smart predictive framework for system-level stormwater management optimization". In: *Journal of Environmental Management* 278, p. 111505. doi: 10.1016/j.jenvman.2020.111505. url: http://dx.doi.org/10.1016/j.jenvman.2020.111505.

Sillanpää, Nora and Harri Koivusalo (Feb. 2015). "Impacts of urban development on runoff event characteristics and unit hydrographs across warm and cold seasons in high latitudes". In: *Journal of Hydrology* 521, pp. 328–340. doi: 10.1016/j.jhydrol.2014.12.008. url: http://dx.doi.org/10.1016/j.jhydrol.2014.12.008.

SMHI (2025a). *Korta nederbördsprognoser – KNEP*. Sveriges meteorologiska och hydrologiska institut. URL: https://www.smhi.se/kunskapsbanken/meteorologi/meteorologiska-modeller/korta-nederbordsprognoser---knep (visited on 03/31/2025).

— (2025b). *Meteorological Forecast Model (AROME) – Historical Analysis Data.* Dataset. Historical model data from the Harmonie-Arome forecast model. The dataset covers Scandinavia, the northern part of the continent, and the Baltic states. Data is provided in GRIB format and updated continuously. Coverage area coordinates: Southwest 51.85N 0.24E, Northwest 71.91N 15.32W, Southeast 51.47N 32.18E, Northeast 71.16N 49.61E. Grid resolution: 2.5 x 2.5 km. SMHI. URL: https://opendata-download-grid-archive.smhi.se/feed/5 (visited on 03/09/2026).

— (2026). *Radar.* Sveriges meteorologiska och hydrologiska institut. URL: https://www.smhi.se/kunskapsbanken/meteorologi/radar (visited on 04/03/2026).

Sucozhañay, Adrián, Juan Pesántez, Rigoberto Guerrero-Coronel, Rodrigo Guerrero-Coronel, David Peña, and Rolando Célleri (Mar. 2024). "Rainwater harvesting as a sustainable solution for the production of urban hydroponic crops". In: *Water Reuse* 14.2, pp. 177–189. DOI: 10.2166/wrd.2024.151. URL: http://dx.doi.org/10.2166/wrd.2024.151.

Sun, Lanxin, Jun Xia, Dunxian She, Qizhong Guo, Yuming Su, and Wenyucheng Wang (May 2023). "Integrated intra-storm predictive analysis and real-time control for urban stormwater storage to reduce flooding risk in cities". In: *Sustainable Cities and Society* 92, p. 104506. DOI: 10.1016/j.scs.2023.104506. URL: http://dx.doi.org/10.1016/j.scs.2023.104506.

Sundstedt, Hedvig (2026). "Control of Rainwater Harvesting: Subtitle". MSc thesis in progress, personal communication. MA thesis. Lunds universitet, Lunds Tekniska Högskola.

Svenskt Vatten (2023). *Klimatanpassningsutredningen föreslår viktiga förtydliganden för framtidens dagvattenhantering.* URL: https://www.svensktvatten.se/om-oss/nyheter-och-press/nyheter/svenskt-vatten-klimatanpassningsutredningen-foreslar-viktiga-fortydliganden-for-framtidens-dagvattenhantering/.

Sveriges miljömål (2022). *Dagvattenhantering i ny eller ändrad bebyggelse.* URL: https://www.sverigesmiljomal.se/etappmalen/dagvattenhantering-i-ny-eller-andrad-bebyggelse/.

Sweetapple, Chris, Anna Hastings, and Peter Melville-Shreeve (Apr. 2026). "Developing strategies for real-time control of stormwater systems: A review". In: *Water Research* 293, p. 125399. DOI: 10.1016/j.watres.2026.125399. URL: http://dx.doi.org/10.1016/j.watres.2026.125399.

Tang, Ziyi, Pin Wang, Yao Li, Chaohui Chen, Yihan Lou, and Tangao Hu (July 2025). "Contributions of urbanization and population growth to changes in urban flood exposure in recent decades". In: *Sustainable Cities and Society* 130, p. 106592. DOI: 10.1016/j.scs.2025.106592. URL: http://dx.doi.org/10.1016/j.scs.2025.106592.

Torelló-Sentelles, Herminia, Gabriele Villarini, Marika Koukoula, and Nadav Peleg (Aug. 2025). "Impacts of urban dynamics and thermodynamics on convective rainfall across different urban forms". In: *Urban Climate* 62, p. 102499. DOI: 10.1016/j.uclim.2025.102499. URL: http://dx.doi.org/10.1016/j.uclim.2025.102499.

Tounsi, Achraf, Marouane Temimi, Mohamed Abdelkader, and Jonathan J. Gourley (Oct. 2023). "Assessment of deterministic and probabilistic precipitation nowcasting techniques over New York metropolitan area". In: *Environmental Modelling & Software* 168, p. 105803. DOI: 10.1016/j.envsoft.2023.105803. URL: http://dx.doi.org/10.1016/j.envsoft.2023.105803.

United Nations (2025). *World Urbanization Prospects 2025: Summary of Results*. Tech. rep. UN DESA/POP/2025/TR/NO. 12. New York: United Nations, Department of Economic and Social Affairs, Population Division. URL: https://population.un.org.

Van De Beek, Remco (2026). *Personal communication*. Personal communication, SMHI, April 2026.

Wilks, Daniel S (Jan. 2005). *Statistical methods in the atmospheric sciences: Volume 100*. en. 2nd ed. International Geophysics. San Diego, CA: Academic Press.

World Meteorological Organization (2017). *Guidelines for Nowcasting Techniques*. Tech. rep. WMO-No. 1198. Geneva: World Meteorological Organization.

Xu, Wei D., Matthew J. Burns, Frédéric Cherqui, Sophie Duchesne, Geneviève Pelletier, and Tim D. Fletcher (Nov. 2022). "Real-time controlled rainwater harvesting systems can improve the performance of stormwater networks". In: *Journal of Hydrology* 614, p. 128503. DOI: 10.1016/j.jhydrol.2022.128503. URL: http://dx.doi.org/10.1016/j.jhydrol.2022.128503.

Xu, Wei D., Matthew J. Burns, Frédéric Cherqui, and Tim D. Fletcher (Dec. 2020a). "Enhancing stormwater control measures using real-time control technology: a review". In: *Urban Water Journal* 18.2, pp. 101–114. DOI: 10.1080/1573062x.2020.1857797. URL: http://dx.doi.org/10.1080/1573062X.2020.1857797.

Xu, Wei D., Tim D. Fletcher, Matthew J. Burns, and Frédéric Cherqui (Aug. 2020b). "Real Time Control of Rainwater Harvesting Systems: The Benefits of Increasing Rainfall Forecast Window". In: *Water Resources Research* 56.9. DOI: 10.1029/2020wr027856. URL: http://dx.doi.org/10.1029/2020WR027856.

Yano, Jun-Ichi, Michał Z. Ziemiański, Mike Cullen, Piet Termonia, Jeanette Onvlee, Lisa Bengtsson, Alberto Carrassi, Richard Davy, Anna Deluca, Suzanne L. Gray, Víctor Homar, Martin Köhler, Simon Krichak, Silas Michaelides, Vaughan T. J. Phillips, Pedro M. M. Soares, and Andrzej A. Wyszogrodzki (Apr. 2018). "Scientific Challenges of Convective-Scale Numerical Weather Prediction". In: *Bulletin of the American Meteorological Society* 99.4, pp. 699–710. DOI: 10.1175/bams-d-17-0125.1. URL: http://dx.doi.org/10.1175/BAMS-D-17-0125.1.

Zawadzki, I., J. Morneau, and R. Laprise (Dec. 1994). "Predictability of Precipitation Patterns: An Operational Approach". In: *Journal of Applied Meteorology* 33.12, pp. 1562–1571. DOI: 10.1175/1520-0450(1994)033<1562:poppao>2.0.co;2. URL: http://dx.doi.org/10.1175/1520-0450(1994)033%3C1562:POPPAO%3E2.0.CO;2.

Zhang, Yuchen, Mingsheng Long, Kaiyuan Chen, Lanxiang Xing, Ronghua Jin, Michael I. Jordan, and Jianmin Wang (July 2023). "Skilful nowcasting of extreme precipitation with NowcastNet". In: *Nature* 619.7970, pp. 526–532. DOI: 10.1038/s41586-023-06184-4. URL: http://dx.doi.org/10.1038/s41586-023-06184-4.

Zhou, Huaiyu, Yao Qu, Hailong Liu, and Guangheng Ni (Aug. 2024). “Smart roofs featuring predictive control: An upgrade for mitigating precipitation extreme-induced pluvial floods”. In: *Journal of Environmental Management* 365, p. 121504. DOI: 10.1016/j.jenvman.2024.121504. URL: http://dx.doi.org/10.1016/j.jenvman.2024.121504.

A Appendix

A1 Declaration regarding usage of AI tools

AI assistance played a key role in the programming work required for this thesis. It was used primarily to support the development, debugging, and restructuring of Python code for the nowcast evaluation workflow. In particular, it supported the transition from exploratory single-case runs to a structured evaluation pipeline suitable for repeated, iterative execution without sacrificing computation time to non-optimised loops etc. The tools used was Microsoft Copilot and Google Gemini under the Lund University licences.

A1.1 Limitations and Human Oversight

All AI-assisted outputs were reviewed, tested, and revised by the author before use. The author remained fully responsible for all code, analyses, and written content included in the thesis. Scientific decisions, including the formulation of the research problem, methodological choices, interpretation of results, and conclusions, were made independently by the author. AI-assisted tools were used to support technical implementation, but not to replace engineering judgement.

A2 Summary of all deterministic runs

Table A1: Median verification scores per method and resolution, stratified by event duration and season. Subscripts denote lead time in minutes; FSS superscripts denote neighbourhood radius in km.

Method	**Resolution**	$\mathbf{MAE_{30}}$ $(\mathbf{mm\,h^{-1}})$	$\mathbf{MAE_{60}}$ $(\mathbf{mm\,h^{-1}})$	$\mathbf{CSI_{30}}$	$\mathbf{CSI_{60}}$	$\mathbf{POD_{30}}$	$\mathbf{POD_{60}}$	$\mathbf{FAR_{30}}$	$\mathbf{FAR_{60}}$	$\mathbf{BIAS_{30}}$	$\mathbf{BIAS_{60}}$	r_{30}	r_{60}	$\mathbf{FSS^{5}_{30}}$	$\mathbf{FSS^{5}_{60}}$	$\mathbf{FSS^{10}_{30}}$	$\mathbf{FSS^{10}_{60}}$	τ **(min)**
*Summer – **1h** events*																		
S-PROG	0.5km/5min	2.07	2.42	0.78	0.67	0.90	0.85	0.13	0.20	0.95	0.97	0.25	0.12	0.94	0.86	0.96	0.89	17
S-PROG	1km/10min	1.07	1.30	0.54	0.35	0.54	0.35	0.01	0.03	0.54	0.35	0.53	0.40	0.74	0.54	0.77	0.59	50
S-PROG	2km/15min	1.07	1.31	0.54	0.36	0.54	0.36	0.01	0.04	0.54	0.38	0.54	0.41	0.75	0.57	0.78	0.61	51
ANVIL	0.5km/5min	3.90	3.37	0.61	0.42	0.67	0.47	0.06	0.15	0.69	0.59	0.14	-0.01	0.83	0.69	0.87	0.74	12
ANVIL	1km/10min	1.47	1.60	0.72	0.56	0.85	0.72	0.11	0.16	0.90	0.83	0.52	0.30	0.93	0.81	0.95	0.85	47
ANVIL	2km/15min	1.34	1.52	0.72	0.56	0.87	0.73	0.11	0.14	0.93	0.84	0.51	0.33	0.92	0.80	0.95	0.84	51
LP	0.5km/5min	3.31	3.26	0.70	0.57	0.82	0.71	0.10	0.18	0.86	0.87	0.15	0.03	0.92	0.82	0.94	0.86	13
LP	1km/10min	1.44	1.59	0.71	0.56	0.87	0.74	0.12	0.16	0.93	0.85	0.51	0.30	0.92	0.80	0.94	0.84	48
LP	2km/15min	1.37	1.53	0.71	0.55	0.89	0.74	0.11	0.14	0.95	0.85	0.51	0.33	0.91	0.79	0.94	0.84	51
EP	0.5km/5min	2.98	1.67	0.40	0.20	0.44	0.23	0.22	0.40	0.74	0.61	-0.11	-0.06	0.57	0.32	0.61	0.34	5
EP	1km/10min	1.95	2.23	0.41	0.18	0.45	0.23	0.21	0.42	0.72	0.58	-0.02	-0.21	0.57	0.31	0.61	0.35	10
EP	2km/15min	1.93	2.23	0.43	0.19	0.46	0.25	0.20	0.41	0.74	0.60	-0.02	-0.26	0.57	0.32	0.62	0.36	15
LINDA-D	0.5km/5min	3.21	2.80	0.63	0.46	0.68	0.54	0.05	0.15	0.71	0.70	0.12	-0.01	0.85	0.72	0.88	0.77	12
LINDA-D	1km/10min	1.33	1.50	0.68	0.62	0.92	0.92	0.28	0.36	1.22	1.27	0.52	0.30	0.88	0.82	0.92	0.86	36
LINDA-D	2km/15min	1.24	1.40	0.68	0.63	0.96	0.96	0.26	0.35	1.17	1.28	0.54	0.34	0.92	0.83	0.95	0.88	43
*Winter – **1h** events*																		
S-PROG	0.5km/5min	0.72	0.92	0.91	0.84	0.96	0.93	0.05	0.08	0.99	0.97	0.46	0.18	0.98	0.94	0.99	0.95	41
S-PROG	1km/10min	0.95	1.08	0.34	0.18	0.34	0.18	0.00	0.01	0.34	0.18	0.47	0.23	0.53	0.26	0.55	0.26	36
S-PROG	2km/15min	0.94	1.08	0.36	0.14	0.36	0.14	0.00	0.02	0.36	0.14	0.50	0.34	0.55	0.23	0.58	0.24	41
ANVIL	0.5km/5min	0.88	1.18	0.75	0.69	0.77	0.73	0.02	0.06	0.78	0.78	0.37	0.19	0.91	0.87	0.92	0.90	30
ANVIL	1km/10min	0.82	1.00	0.87	0.79	0.90	0.85	0.04	0.07	0.95	0.91	0.47	0.33	0.97	0.93	0.98	0.95	43
ANVIL	2km/15min	0.79	1.03	0.89	0.64	0.92	0.69	0.04	0.06	0.96	0.74	0.51	0.33	0.97	0.82	0.98	0.84	41
LP	0.5km/5min	0.85	1.03	0.87	0.80	0.91	0.86	0.04	0.09	0.94	0.93	0.36	0.16	0.97	0.93	0.98	0.95	29
LP	1km/10min	0.83	1.00	0.89	0.81	0.92	0.87	0.05	0.09	0.96	0.96	0.46	0.31	0.98	0.94	0.99	0.96	42
LP	2km/15min	0.79	0.97	0.89	0.81	0.93	0.87	0.04	0.09	0.97	0.96	0.49	0.38	0.98	0.93	0.99	0.95	61
EP	0.5km/5min	1.27	1.19	0.78	0.65	0.87	0.78	0.12	0.21	0.94	0.97	-0.06	-0.06	0.92	0.81	0.94	0.83	5
EP	1km/10min	1.19	1.20	0.76	0.63	0.85	0.76	0.12	0.22	0.92	0.94	-0.12	-0.10	0.91	0.79	0.93	0.81	10
EP	2km/15min	1.18	1.18	0.77	0.62	0.86	0.76	0.12	0.23	0.94	0.95	-0.14	-0.12	0.91	0.78	0.93	0.80	15
LINDA-D	0.5km/5min	0.76	0.99	0.78	0.71	0.80	0.75	0.03	0.07	0.82	0.81	0.41	0.26	0.92	0.88	0.94	0.91	44
LINDA-D	1km/10min	0.72	0.88	0.88	0.84	0.93	0.96	0.07	0.13	1.03	1.13	0.55	0.39	0.96	0.94	0.97	0.95	69
LINDA-D	2km/15min	0.72	0.88	0.90	0.84	0.96	0.96	0.07	0.14	1.05	1.14	0.55	0.41	0.97	0.92	0.98	0.93	63
*Summer – **3h** events*																		
S-PROG	0.5km/5min	2.98	3.57	0.85	0.89	0.88	0.91	0.04	0.03	0.92	0.94	0.35	0.04	0.90	0.90	0.92	0.92	29
S-PROG	1km/10min	2.80	3.18	0.71	0.64	0.72	0.65	0.01	0.02	0.72	0.66	0.50	-0.02	0.77	0.78	0.79	0.80	39
S-PROG	2km/15min	2.78	3.12	0.72	0.68	0.72	0.69	0.01	0.02	0.73	0.71	0.48	0.08	0.82	0.83	0.84	0.85	37
ANVIL	0.5km/5min	4.42	4.41	0.66	0.66	0.67	0.66	0.01	0.01	0.67	0.67	0.33	-0.04	0.70	0.72	0.72	0.74	27
ANVIL	1km/10min	3.02	4.13	0.86	0.89	0.91	0.92	0.05	0.04	0.96	0.97	0.48	0.05	0.92	0.93	0.94	0.95	30
ANVIL	2km/15min	3.02	4.14	0.88	0.73	0.93	0.77	0.06	0.07	0.99	0.83	0.51	0.17	0.95	0.86	0.97	0.88	32
LP	0.5km/5min	4.53	4.27	0.81	0.85	0.83	0.87	0.03	0.02	0.85	0.89	0.26	-0.05	0.84	0.85	0.87	0.87	24
LP	1km/10min	3.04	4.09	0.86	0.90	0.91	0.94	0.06	0.04	0.97	0.99	0.42	0.03	0.92	0.94	0.95	0.95	28
LP	2km/15min	2.96	4.07	0.88	0.91	0.93	0.95	0.06	0.05	1.01	0.98	0.43	0.11	0.94	0.94	0.97	0.96	29
EP	0.5km/5min	5.59	4.96	0.44	0.30	0.46	0.30	0.01	0.01	0.51	0.32	-0.28	0.15	0.57	0.35	0.59	0.36	5
EP	1km/10min	4.97	4.33	0.44	0.30	0.47	0.30	0.01	0.01	0.52	0.31	-0.18	-0.05	0.60	0.36	0.62	0.36	10
EP	2km/15min	4.98	4.13	0.51	0.36	0.53	0.36	0.01	0.01	0.59	0.38	-0.14	-0.06	0.60	0.38	0.63	0.39	15
LINDA-D	0.5km/5min	3.46	4.06	0.74	0.70	0.75	0.71	0.02	0.02	0.76	0.72	0.38	-0.04	0.75	0.75	0.76	0.76	30
LINDA-D	1km/10min	2.76	3.92	0.80	0.82	0.98	0.98	0.19	0.16	1.20	1.16	0.42	0.20	0.86	0.92	0.88	0.93	31
LINDA-D	2km/15min	2.92	4.02	0.80	0.83	1.00	1.00	0.19	0.16	1.23	1.19	0.45	0.27	0.86	0.92	0.87	0.93	37

(continues on next page)

(continued from previous page)

Method	**Resolution**	**MAE_{30} (mm h^{-1})**	**MAE_{60} (mm h^{-1})**	**CSI_{30}**	**CSI_{60}**	**POD_{30}**	**POD_{60}**	**FAR_{30}**	**FAR_{60}**	**$BIAS_{30}$**	**$BIAS_{60}$**	r_{30}	r_{60}	**FSS^{5}_{30}**	**FSS^{5}_{60}**	**FSS^{10}_{30}**	**FSS^{10}_{60}**	**τ (min)**
*Winter – **3h** events*																		
S-PROG	0.5km/5min	0.45	0.59	0.81	0.74	0.93	0.78	0.04	0.05	1.04	0.86	0.51	0.32	0.91	0.79	0.93	0.79	47
S-PROG	1km/10min	0.61	0.79	0.27	0.28	0.27	0.28	0.00	0.00	0.27	0.29	0.53	0.34	0.37	0.34	0.39	0.35	48
S-PROG	2km/15min	0.75	0.85	0.26	0.23	0.26	0.24	0.01	0.00	0.26	0.25	0.52	0.30	0.36	0.30	0.37	0.31	53
ANVIL	0.5km/5min	0.62	0.76	0.72	0.70	0.72	0.72	0.02	0.03	0.74	0.74	0.36	0.17	0.84	0.84	0.85	0.86	31
ANVIL	1km/10min	0.44	0.64	0.87	0.85	0.89	0.88	0.03	0.03	0.92	0.91	0.55	0.23	0.96	0.95	0.97	0.96	46
ANVIL	2km/15min	0.49	0.64	0.88	0.78	0.90	0.80	0.02	0.04	0.92	0.84	0.57	0.27	0.95	0.89	0.97	0.91	50
LP	0.5km/5min	0.60	0.69	0.87	0.83	0.89	0.86	0.03	0.04	0.93	0.89	0.30	0.12	0.95	0.93	0.96	0.95	22
LP	1km/10min	0.50	0.69	0.88	0.86	0.92	0.89	0.04	0.03	0.96	0.92	0.54	0.20	0.97	0.95	0.98	0.96	45
LP	2km/15min	0.53	0.68	0.90	0.84	0.92	0.87	0.03	0.04	0.95	0.90	0.54	0.27	0.97	0.94	0.97	0.95	45
EP	0.5km/5min	0.71	0.77	0.73	0.64	0.77	0.68	0.04	0.05	0.79	0.69	0.07	0.00	0.82	0.70	0.83	0.70	5
EP	1km/10min	0.58	0.67	0.72	0.63	0.76	0.67	0.05	0.05	0.78	0.68	0.02	0.02	0.82	0.69	0.83	0.70	10
EP	2km/15min	0.65	0.71	0.73	0.63	0.76	0.68	0.04	0.05	0.79	0.69	0.03	0.04	0.83	0.70	0.84	0.71	15
LINDA-D	0.5km/5min	0.52	0.65	0.74	0.74	0.75	0.75	0.02	0.03	0.77	0.78	0.36	0.17	0.86	0.87	0.88	0.89	34
LINDA-D	1km/10min	0.40	0.57	0.87	0.89	0.93	0.97	0.06	0.09	1.00	1.05	0.62	0.26	0.96	0.96	0.97	0.97	48
LINDA-D	2km/15min	0.42	0.58	0.91	0.90	0.95	0.98	0.05	0.08	1.01	1.02	0.65	0.33	0.97	0.96	0.98	0.97	55
*Summer – **6h** events*																		
S-PROG	0.5km/5min	2.98	2.27	0.79	0.66	0.82	0.69	0.04	0.04	0.86	0.83	0.35	0.31	0.87	0.69	0.88	0.72	29
S-PROG	1km/10min	1.27	1.51	0.41	0.45	0.41	0.47	0.01	0.02	0.45	0.51	0.46	0.56	0.48	0.46	0.49	0.48	83
S-PROG	2km/15min	1.34	1.64	0.46	0.44	0.47	0.44	0.01	0.02	0.51	0.48	0.57	0.54	0.55	0.48	0.56	0.48	81
ANVIL	0.5km/5min	4.42	4.07	0.45	0.46	0.45	0.50	0.01	0.01	0.46	0.60	0.10	0.12	0.55	0.57	0.56	0.60	16
ANVIL	1km/10min	1.12	1.61	0.75	0.63	0.78	0.79	0.05	0.04	0.87	0.97	0.56	0.54	0.86	0.78	0.88	0.82	–
ANVIL	2km/15min	1.16	1.64	0.81	0.63	0.85	0.65	0.06	0.07	0.98	0.69	0.64	0.52	0.94	0.75	0.96	0.79	51
LP	0.5km/5min	4.53	3.79	0.68	0.59	0.70	0.70	0.04	0.02	0.85	0.89	0.10	0.13	0.81	0.71	0.83	0.75	16
LP	1km/10min	1.17	1.62	0.74	0.64	0.77	0.81	0.06	0.04	0.97	0.99	0.51	0.54	0.86	0.77	0.88	0.81	–
LP	2km/15min	1.17	1.69	0.81	0.69	0.85	0.80	0.06	0.05	0.99	1.00	0.59	0.53	0.94	0.80	0.96	0.84	–
EP	0.5km/5min	5.59	3.61	0.38	0.24	0.42	0.25	0.06	0.03	0.47	0.28	0.03	0.08	0.49	0.36	0.50	0.37	5
EP	1km/10min	1.62	2.07	0.37	0.20	0.40	0.21	0.06	0.03	0.47	0.23	0.20	0.13	0.46	0.22	0.48	0.23	10
EP	2km/15min	1.74	2.28	0.39	0.22	0.44	0.22	0.06	0.02	0.47	0.25	0.20	0.13	0.50	0.24	0.52	0.25	15
LINDA-D	0.5km/5min	3.46	3.47	0.51	0.50	0.51	0.56	0.02	0.02	0.53	0.68	0.12	0.13	0.61	0.62	0.63	0.65	17
LINDA-D	1km/10min	1.00	1.43	0.80	0.76	0.94	0.95	0.19	0.21	1.20	1.21	0.61	0.49	0.86	0.82	0.88	0.83	16
LINDA-D	2km/15min	1.08	1.51	0.80	0.77	0.97	0.98	0.19	0.23	1.23	1.29	0.66	0.51	0.86	0.83	0.87	0.83	24
*Winter – **6h** events*																		
S-PROG	0.5km/5min	1.25	1.80	0.80	0.74	0.82	0.75	0.03	0.01	0.84	0.76	0.48	0.39	0.87	0.74	0.88	0.75	46
S-PROG	1km/10min	1.09	1.71	0.43	0.22	0.43	0.22	0.00	0.00	0.43	0.22	0.53	0.36	0.46	0.22	0.47	0.22	41
S-PROG	2km/15min	1.20	1.73	0.41	0.19	0.41	0.19	0.00	0.00	0.41	0.19	0.51	0.31	0.46	0.20	0.48	0.20	40
ANVIL	0.5km/5min	1.43	1.81	0.71	0.70	0.72	0.70	0.02	0.01	0.73	0.71	0.36	0.05	0.80	0.76	0.82	0.77	30
ANVIL	1km/10min	0.99	1.46	0.78	0.71	0.79	0.71	0.02	0.01	0.81	0.72	0.63	0.42	0.87	0.74	0.89	0.75	50
ANVIL	2km/15min	1.05	1.45	0.75	0.70	0.77	0.70	0.03	0.01	0.80	0.71	0.61	0.38	0.87	0.76	0.89	0.78	48
LP	0.5km/5min	1.56	2.05	0.78	0.79	0.79	0.79	0.02	0.01	0.81	0.80	0.36	0.08	0.88	0.86	0.90	0.87	30
LP	1km/10min	1.08	1.55	0.78	0.73	0.79	0.74	0.03	0.01	0.82	0.74	0.59	0.41	0.87	0.76	0.89	0.76	48
LP	2km/15min	1.14	1.55	0.75	0.72	0.77	0.73	0.03	0.01	0.80	0.73	0.58	0.38	0.87	0.78	0.89	0.79	47
EP	0.5km/5min	1.78	2.32	0.49	0.37	0.50	0.38	0.05	0.04	0.53	0.40	-0.16	0.03	0.57	0.47	0.58	0.47	5
EP	1km/10min	1.36	1.81	0.50	0.36	0.51	0.37	0.05	0.04	0.54	0.39	0.16	0.20	0.56	0.43	0.57	0.43	14
EP	2km/15min	1.42	1.84	0.47	0.33	0.48	0.33	0.04	0.04	0.50	0.35	0.11	0.12	0.54	0.40	0.55	0.40	15
LINDA-D	0.5km/5min	1.32	1.78	0.73	0.69	0.74	0.70	0.02	0.01	0.76	0.70	0.38	0.07	0.83	0.72	0.85	0.73	31
LINDA-D	1km/10min	0.90	1.35	0.86	0.79	0.94	0.81	0.09	0.03	1.03	0.83	0.65	0.48	0.93	0.80	0.95	0.80	69
LINDA-D	2km/15min	0.93	1.35	0.86	0.82	0.93	0.85	0.09	0.04	1.02	0.88	0.61	0.39	0.94	0.89	0.95	0.90	50

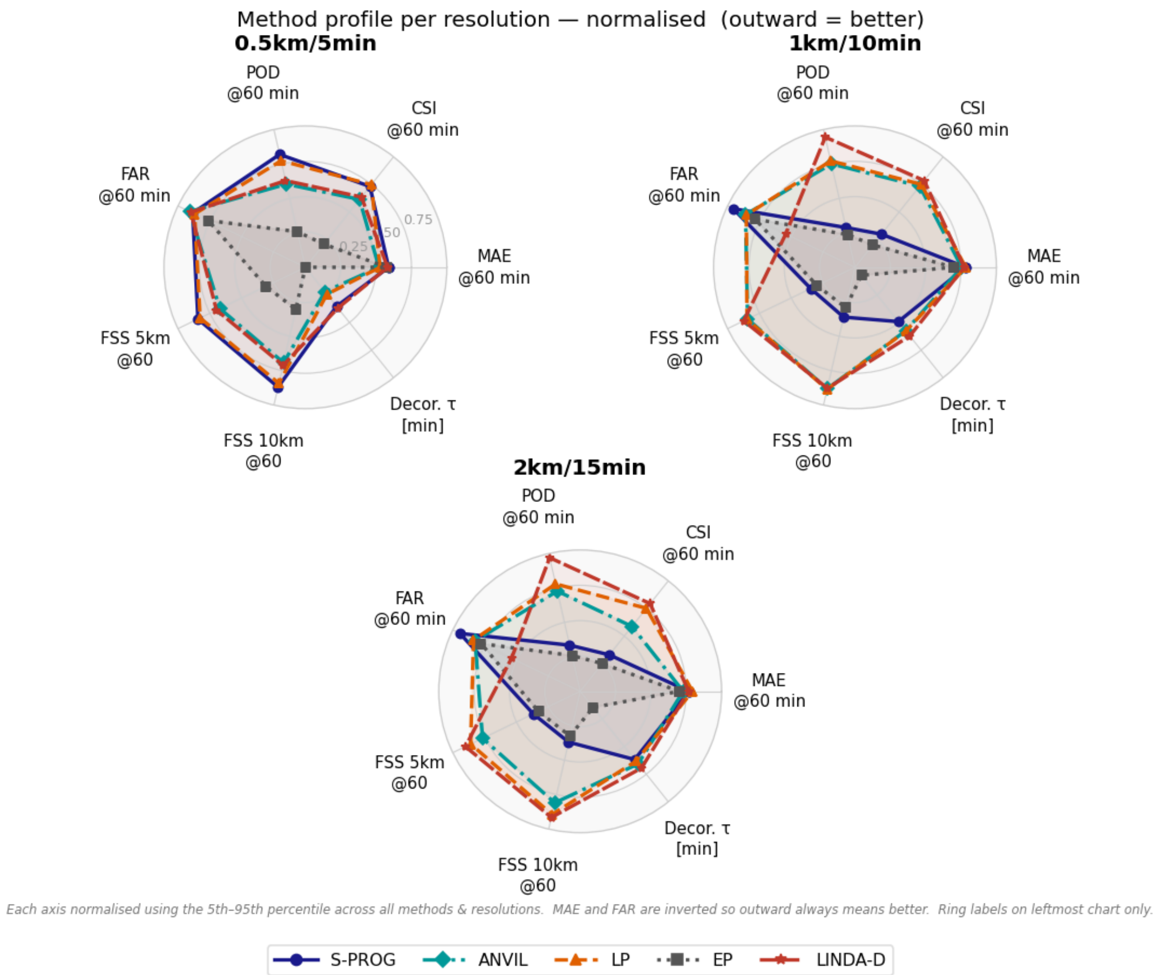

Figure A1: Method profile per resolution at 360 min lead time, aggregated across all 19 events. Each axis shows a verification metric normalised so that 0 corresponds to the bottom 5th percentile and 1 to the top 5th percentile across all methods and resolutions; outward always indicates better performance. MAE and FAR are inverted accordingly.

A3 Complete List of Selected Events

Table A2: Complete list of all 39 selected events. Red shaded rows indicate events that were excluded from the final analysis due to unavailable radar composite data, substantial gaps in the event window, or fallback to lower-resolution radar products during portions of the event. Events 37–39 (blue) were retained from an earlier sensitivity analysis using SMHI's MetObs station *Malmö A (52350)* to broaden seasonal and duration coverage.

#	Date (local)	Season	Dur. (h)	Gauge (mm)	Peak ($\mathrm{mm\,h^{-1}}$)	T (yr)	Station
1	2023-06-17	Summer	1	16.2	42.4	2.4	Augustenborg
2	2022-07-03	Summer	1	12.8	24.8	1.1	Augustenborg
3	2018-04-19	Summer	1	10.6	40.0	0.6	Turbinen PST
4	2025-09-10	Summer	1	10.2	20.8	0.5	Turbinen PST
5	2021-09-27	Summer	1	10.0	32.8	0.5	Turbinen PST
6	2024-07-10	Summer	1	9.2	32.0	0.4	Augustenborg
7	2022-10-01	Winter	1	14.4	21.6	1.6	Limhamn PST
8	2020-10-06	Winter	1	7.0	17.6	0.2	Turbinen PST
9	2019-10-13	Winter	1	4.8	16.8	0.1	Turbinen PST
10	2018-11-12	Winter	1	4.6	7.2	0.1	Turbinen PST
11	2025-10-04	Winter	1	4.0	7.2	0.1	Turbinen PST
12	2020-11-02	Winter	1	4.0	6.4	0.1	Turbinen PST
13	2017-09-17	Summer	3	38.4	31.2	14.0	Turbinen PST
14	2023-07-12	Summer	3	30.8	25.6	6.9	Augustenborg
15	2018-08-09	Summer	3	21.4	35.2	2.1	Turbinen PST
16	2023-08-19	Summer	3	18.8	24.8	1.4	Limhamn PST
17	2020-06-19	Summer	3	18.2	20.8	1.2	Limhamn PST
18	2023-06-17	Summer	3	16.6	42.4	0.9	Augustenborg
19	2022-10-01	Winter	3	15.8	20.0	0.7	Turbinen PST
20	2023-11-16	Winter	3	15.2	12.0	0.7	Limhamn PST
21	2017-10-02	Winter	3	11.4	6.4	0.2	Turbinen PST
22	2023-02-03	Winter	3	11.0	6.4	0.2	Augustenborg
23	2022-10-02	Winter	3	10.4	10.4	0.2	Augustenborg
24	2021-12-01	Winter	3	8.2	4.0	0.1	Turbinen PST
25	2017-09-17	Summer	6	38.6	31.2	7.8	Turbinen PST
26	2023-07-12	Summer	6	31.2	25.6	3.8	Augustenborg
27	2022-05-30	Summer	6	24.2	25.6	1.5	Limhamn PST
28	2019-07-28	Summer	6	22.4	32.0	1.1	Turbinen PST
29	2018-08-09	Summer	6	21.6	35.2	1.0	Turbinen PST
30	2021-09-29	Summer	6	20.8	8.8	0.9	Augustenborg
31	2023-11-16	Winter	6	31.2	18.4	3.8	Limhamn PST
32	2021-10-05	Winter	6	19.8	6.4	0.7	Augustenborg
33	2021-10-21	Winter	6	19.2	10.4	0.6	Turbinen PST

(continued on next page)

(continued from previous page)

#	Date (local)	Season	Dur. (h)	Gauge (mm)	Peak ($\mathrm{mm\,h^{-1}}$)	T (yr)	Station
34	2022-10-01	Winter	6	18.4	20.0	0.5	Turbinen PST
35	2025-10-04	Winter	6	17.8	7.2	0.5	Turbinen PST
36	2021-12-01	Winter	6	16.0	4.0	0.3	Turbinen PST
38	2022-05-31	Summer	1	11.5	17.6	0.8	(Malmö A)
39	2019-06-12	Summer	3	23.5	36.4	2.9	(Malmö A)
40	2023-07-12	Summer	3	32.2	26.0	8.0	(Malmö A)

A4 Onset Frames

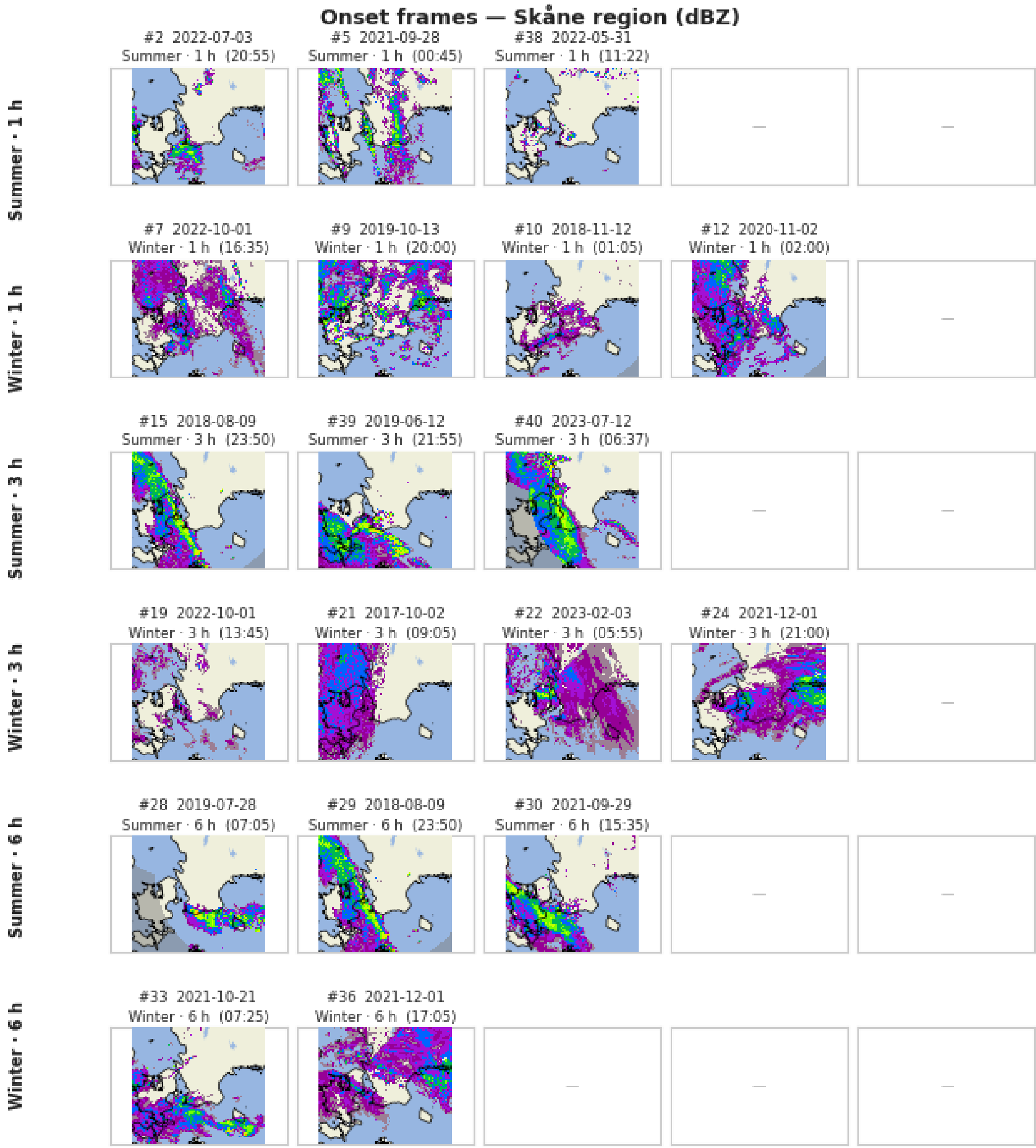

Figure A2: Onset frames for all 19 events, stratified after duration and season.

A5 Hyetographs

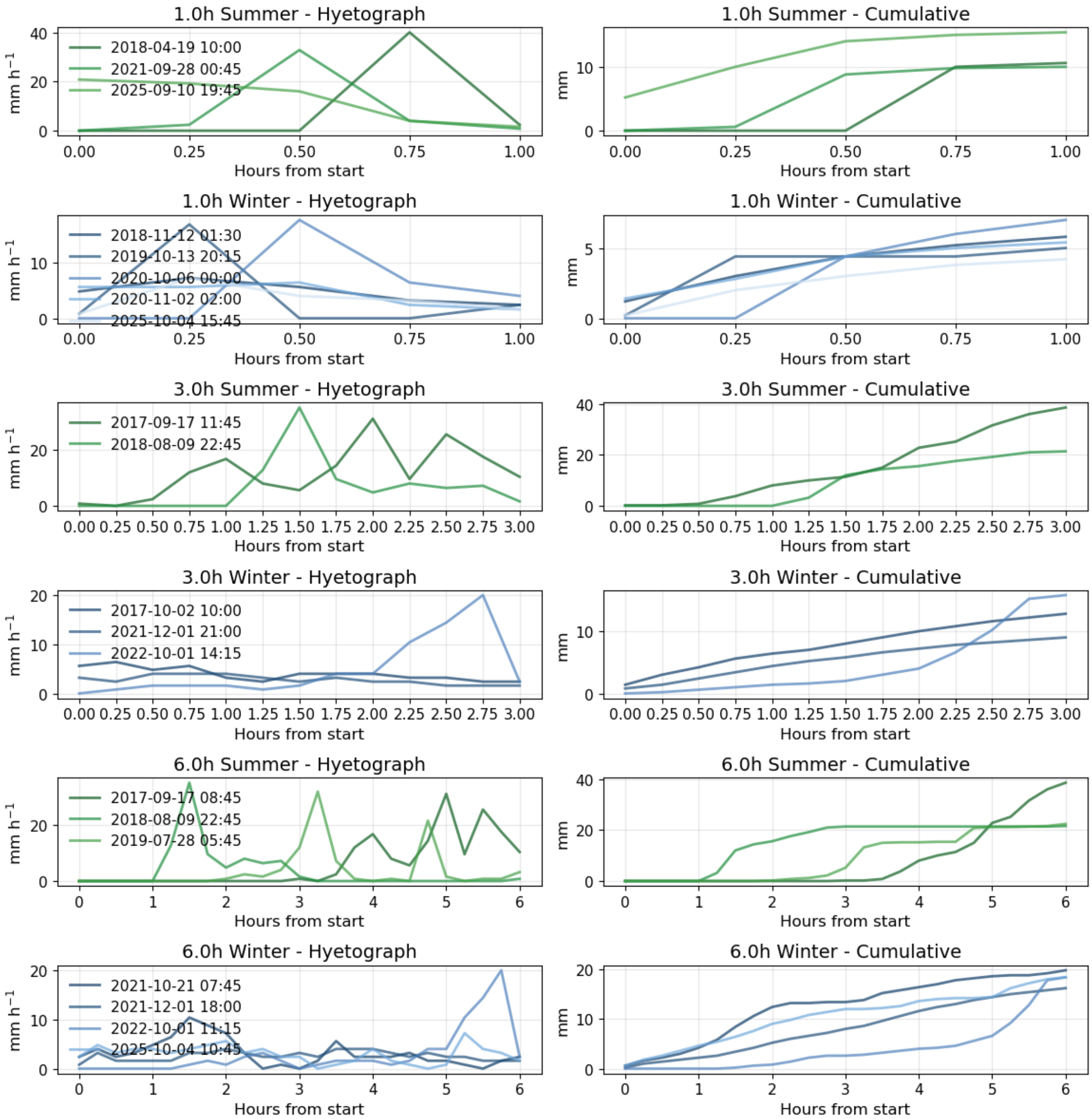

Figure A3: Hyetographs for events chosen from station Turbinen PST.

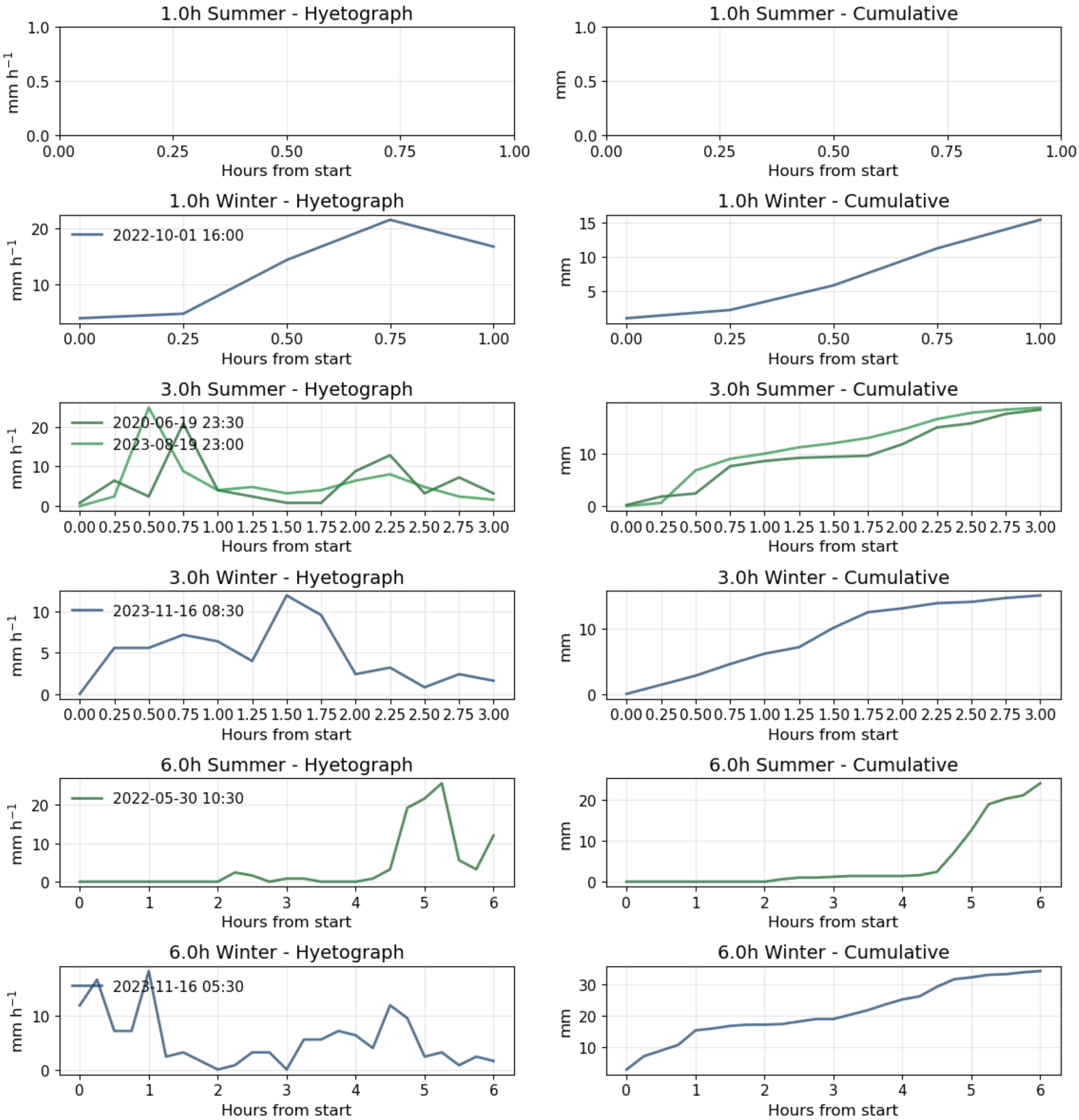

Figure A4: Hyetographs for events chosen from station Limhamn PST.

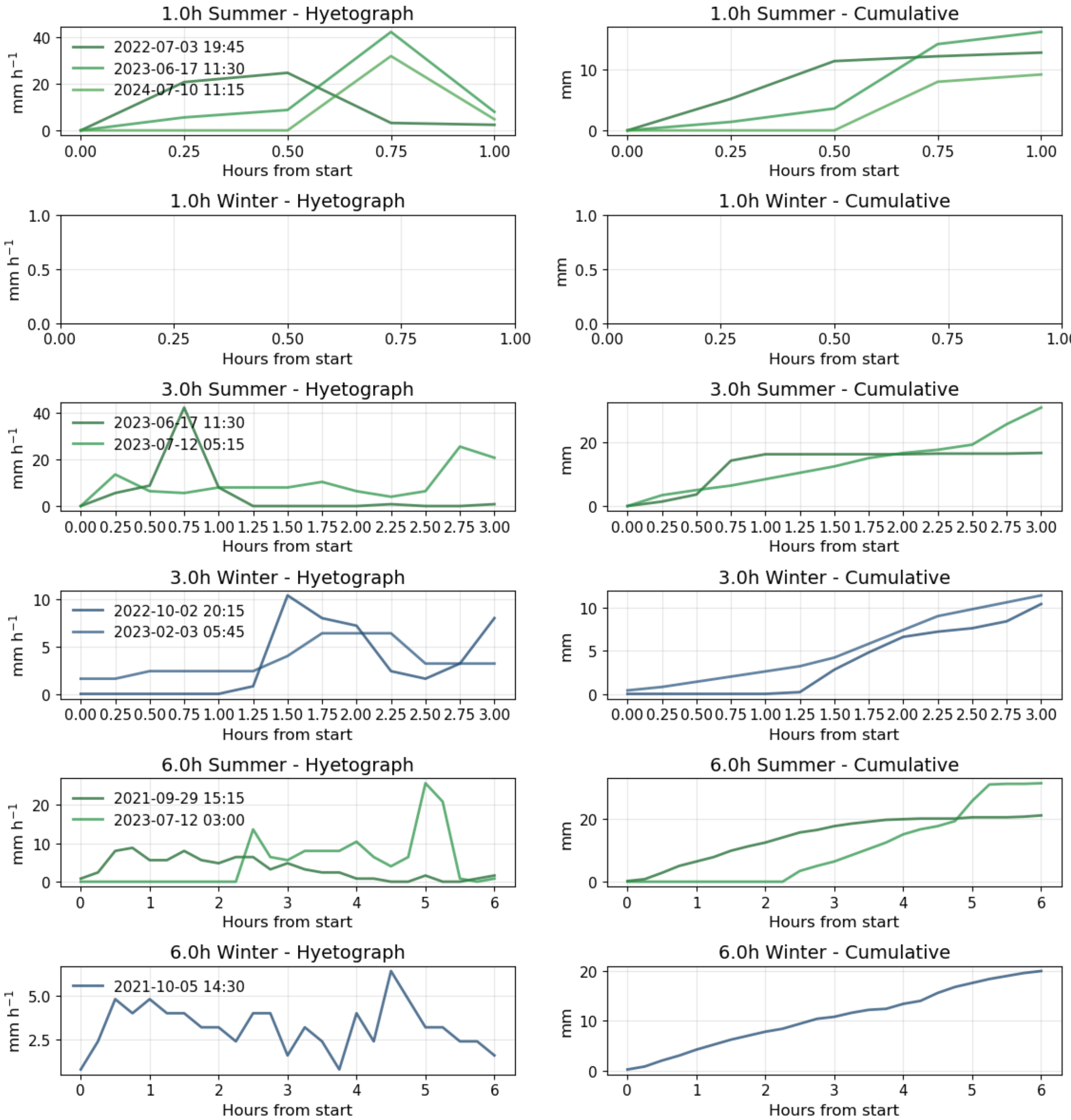

Figure A5: Hyetographs for events chosen from station Augustenborg.