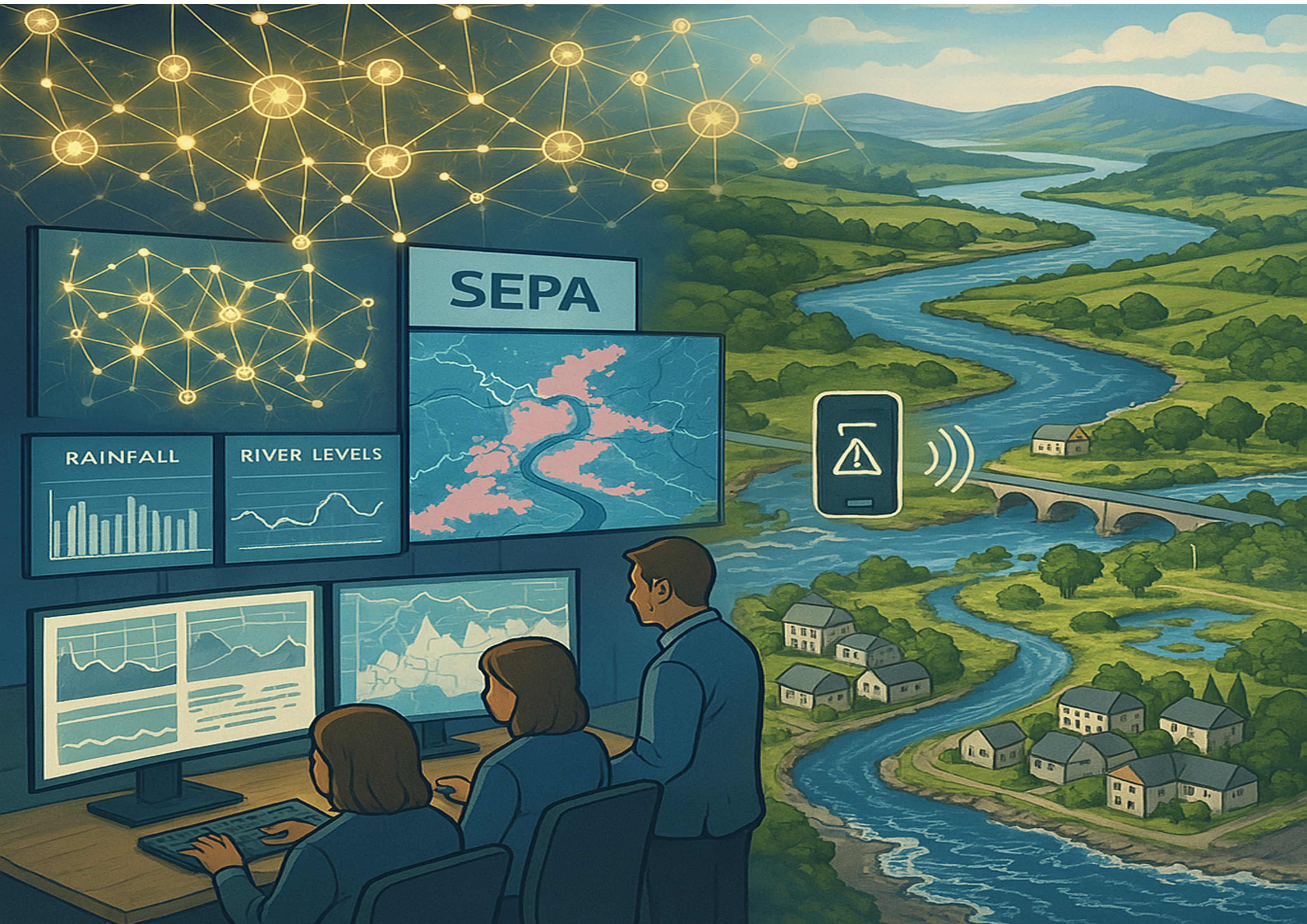
CREW CENTRE OF EXPERTISE FOR WATERS

# Exploring the use of Artificial Intelligence for flood forecasting in Scotland

**Christopher White, Douglas Bertram, Robert Atkinson, Muhammad Usman, Kamila Nieradzinska, Victoria Martí Barclay**




University of Strathclyde Glasgow

Scottish Government gov.scot

CENTRE OF EXPERTISE FOR WATERS

# Exploring the use of Artificial Intelligence for flood forecasting in Scotland

**Christopher White, Douglas Bertram, Robert Atkinson, Muhammad Usman, Kamila Nieradzinska, Victoria Martí Barclay**

University of Strathclyde Glasgow

Scottish Government gov.scot

CENTRE OF EXPERTISE FOR WATERS

**Published by CREW –** Scotland's Centre of Expertise for Waters. CREW connects research and policy, delivering objective and robust research and expert opinion to support the development and implementation of water policy in Scotland. CREW is a partnership between the James Hutton Institute and all Scottish Higher Education Institutes. The Centre is funded by the Scottish Government.

**This document was produced by:**
Christopher J. White, Douglas Bertram, Robert Atkinson, Muhammad Usman, Kamila Nieradzinska, Victoria Martí Barclay

University of Strathclyde, 16 Richmond Street, Glasgow G1 1XQ

**CREW Project Manager:** Rebekah Burman

**Please reference this report as follows:**
White, C.J., Bertram, D., Atkinson, R., Usman, M., Nieradzinska, K and Martí Barclay, V. (2025) *Exploring the use of Artificial Intelligence for flood forecasting in Scotland.* Main report. CSPF2025_01. Centre of Expertise for Waters (CREW).

Available online at: www.crew.ac.uk/publication/artificial-intelligence-for-flood-forecasting-in-Scotland

**ISBN:** 978-1-911706-41-0

**Dissemination status:** Unrestricted



**Acknowledgements:** The *DelugeAI* research team would like to thank the members of the Project Steering Group for their expertise, enthusiasm and guidance throughout the project: Michael Cranston, David Faichney and Darryl Baxter (all Scottish Environment Protection Agency), Neil Ryan (Environment Agency), Jonathan Sewell (Scottish Government) and Athena Zitrou (Scottish Water); the expert workshop participants, and the project management team at CREW: Rebekah Burman and Amy Cooper.

**Cover photograph courtesy of:** Michael Cranston. Image produced using Microsoft Copilot.

**Please contact enquiries@crew.ac.uk, to request report in an alternative format.**

# Contents

**List of figures**



**List of tables**

# Glossary

| | |
|---|---|
| AB | Abstract field in Web of Science search engine |
| AI | Artificial Intelligence |
| ANN | Artificial neural network (a machine learning model) |
| ARIMA | Autoregressive Integrated Moving Average (a statistical time series model) |
| CNN | Convolutional neural network (a deep learning model) |
| DEFRA | Department for Environment, Food & Rural Affairs |
| DL | Deep learning |
| DMI | Danish Meteorological Institute (Denmark) |
| DTU | Technical University of Denmark (Danmarks Tekniske Universitet) |
| EA | Environment Agency (England) |
| EW4All | Early Warning for All initiative |
| EWS | Early Warning System |
| GAN | Generative adversarial network (a deep learning model) |
| GNN | Graph neural network (a deep learning model) |
| IoT | Internet of Things |
| ITU | International Telecommunication Union |
| LLM | Large language model (an artificial intelligence model) |
| LSTM | Long Short-Term Memory (a deep learning model) |
| MCDA | Multi-Criteria Decision Analysis |
| ML | Machine learning |
| MLP | Multi-layer perceptron (a type of artificial neural network) |
| NIWA | National Institute of Water and Atmospheric Research (New Zealand) |
| NRW | National Resources Wales |
| RF | Random forest (a machine learning model) |
| RNN | Recurrent Neural Network (a category of deep learning models) |
| SEPA | Scottish Environment Protection Agency |
| SVM | Support vector machine (a machine learning model) |
| TI | Title field in the Web of Science search engine |
| TS | Topic field (including title, abstract, keywords and keywords Plus) in search |
| UKRI | UK Research and Innovation |
| UN | United Nations |
| U-Net | A type of convolutional neural network |
| UNDRR | UN Office for Disaster Risk Reduction |
| UNESCO | United Nations Educational, Scientific and Cultural Organization |
| WMO | World Meteorological Organisation |
| WoS | Web of Science |
| XAI | Explainable Artificial Intelligence |

# Executive Summary

## Purpose of research

The aim of this *DelugeAI* project, funded by Scotland's Centre of Expertise for Waters (CREW) and led by the University of Strathclyde, is to critically review the current state of Artificial Intelligence (AI) and Machine Learning (ML) technologies and methodologies in flood forecasting. The review is drawn from the latest research and assessments, evaluating AI's potential for flood forecasting and anticipatory actions in the Scottish context, assessing the feasibility of incorporating AI and ML within the Scottish Environment Protection Agency's (SEPA) current flood forecasting capabilities, and providing recommendations for future research, implementation and operationalisation.

To achieve this, *DelugeAI* poses the following four research questions:

1. Is there growing evidence in the literature that can be used to identify potential AI/ML methodologies and technologies which could be applied to flood forecasting?
2. Can experts in AI/ML and hydrological flood forecasting be engaged to facilitate discussions and provide state-of-the-art guidance on key challenges, opportunities and future directions?
3. Can the practicalities of AI/ML integration into SEPA's existing flood forecasting frameworks be assessed and quantified?
4. Can a plausible set of recommendations for advancing AI/ML-driven flood forecasting be developed outlining priority developments and implementation pathways?

## Background

SEPA serves Scotland as the national flood forecasting, flood warning, and strategic flood risk management authority. Their role is critical in mitigating flood risks, protecting communities and enhancing preparedness for flood events. As part of SEPA's Flood Warning Development Framework (2022-2028), SEPA outlined its strategic aim to upgrade its forecasting capabilities through targeted development and innovation, reflecting SEPA's commitment to adopt advanced technologies and methodologies.

One key area of emerging innovation that SEPA, together with Scottish Water and the Scottish Government, were keen to explore was the application of AI and ML in flood forecasting. AI/ML offers the potential to transform flood forecasting by analysing large datasets, recognising patterns, and providing real-time predictions. AI/ML tools and innovations have undergone rapid development in recent years, providing the potential to strengthen SEPA's ability to anticipate flood events, optimize response strategies and safeguard Scotland's environment and population

## Key findings

1. Academic research to date prioritizes replacing traditional models with end-to-end ML approaches, particularly using Long Short-Term Memory (LSTM) models and Convolutional Neural Networks (CNNs), while operational settings favour hybrid models that blend AI with physics-based simulations for greater interpretability and performance.
2. Grey literature emphasises the role of AI-enabled monitoring in underserved areas, often leveraging citizen science, though these solutions are typically localized and challenging to scale due to unclear methodologies and high resource requirements.
3. AI/ML applications in model calibration and enhanced input forecasting (e.g., precipitation) remain generally underexplored, particularly with respect to their integration into full flood forecasting workflows.
4. The use of AI/ML for decision support and issuing warnings is still limited, but expected to grow, especially following international initiatives such as the UN's Early Warnings for All (EW4All).
5. Expert insights confirm hybrid models improve forecast accuracy across event types and highlight the operational efficiency gains AI can provide, particularly in data assimilation and mapping, while underscoring the irreplaceable role of human judgment.
6. Despite global interest and momentum, most AI/ML implementations currently serve to support rather than supplant traditional flood forecasting systems, with ongoing challenges in transparency, data integrity, and skills development.
7. A feasibility study highlights the need for SEPA to adopt a phased AI/ML integration strategy, beginning with high-impact, low-effort applications such as early warnings and response and decision support, which will build confidence and capability for more complex future integrations.

## Recommendations

1. The *DelugeAI* project recommends SEPA to adopt a phased AI/ML integration strategy, beginning with high-impact, low-effort applications over the next 1-2 years such as early warnings and response and decision support.
2. In the 3–5-year horizon, efforts should focus on monitoring involving local communities to increase trust in AI/ML enhanced forecasts, model calibration and integration of enhanced inputs to complement existing approaches.
3. SEPA should trial AI/ML pilots that maintain human oversight to build confidence and demonstrate value.
4. Suggest SEPA invests in targeted training to equip forecasters with AI/ML fundamentals, best practices and model interpretation skills, ensuring human expertise remains central.
5. SEPA’s AI/ML model selection should be guided by clear problem definitions, data availability and quality, interpretability needs, and computational constraints.

# 1.0 Introduction

## 1.1 Background and scope

Flooding poses a significant threat in Scotland, exacerbated by a warming climate (Sniffer, 2021). Traditional flood forecasting approaches using physics-based rainfall-runoff flood models are well-established (Moor and Bell, 2001) but often struggle with complex hydrological processes and require extensive calibration. The rapid development and potential of Artificial Intelligence (AI) offers an emerging data-driven alternative, offering huge potential to complement existing capabilities that leverage large datasets to identify spatial patterns, widen the range of flood forecasting capabilities and predictive timescales (i.e., towards real-time predictions) and improve accuracy (e.g., Liu *et al.*, 2025). AI – in particular Machine Learning (ML) and deep learning (DL) (e.g., Xie *et al.*, 2021) – has the potential to transform SEPA's flood forecasting capabilities in Scotland by integrating diverse and novel hydrometeorological data sources, recognising nonlinear and multivariate relationships, and providing predictive insights for response optimisation. AI presents an incredible opportunity to enhance early warning systems (EWS), improve disaster preparedness, and support adaptive flood risk management (e.g., Ghaffarian *et al.*, 2023; UNU EHS, 2024; Cirri, 2023). However, challenges remain in model interpretability, data availability, and integration with existing flood forecasting approaches.

To address these challenges, the scope of the *DelugeAI* project was to critically review the current state of AI/ML technologies and methodologies in flood forecasting drawn from the latest research and assessments, evaluate its potential for flood forecasting and anticipatory actions in the Scottish context, and assess the feasibility of incorporating AI/ML with SEPA's current flood forecasting capabilities, providing a clear set of recommendations for integrating AI/ML with SEPA's existing flood forecasting capabilities in Scotland.

The multifaceted potential of AI/ML to enhance flood forecasting cannot be understated, spanning all aspects involved in producing, issuing and responding to a flood forecast, as well as the models, systems and processes that support them. This provides a complex picture of both the challenges and opportunities for using AI/ML to enhance flood forecasting. To structure this review and provide consistence and clarity, we adopted a framework that identifies seven phases of the flood forecasting process where we believe AI/ML can integrate with and/or enhance flood forecasting in a variety of ways, including monitoring, model development, and warning and response in real time (Figure 1). These seven phases were then used to guide and frame all aspects of the *DelugeAI* project, including the critical review, expert workshop, feasibility study and potential roadmap to ensure a consistent approach.

These seven phases are described in more detail as follows, including a rationale for their addition as an area where AI/ML may integrate with and/or enhance flood forecasting in Scotland as shown in Table 1.

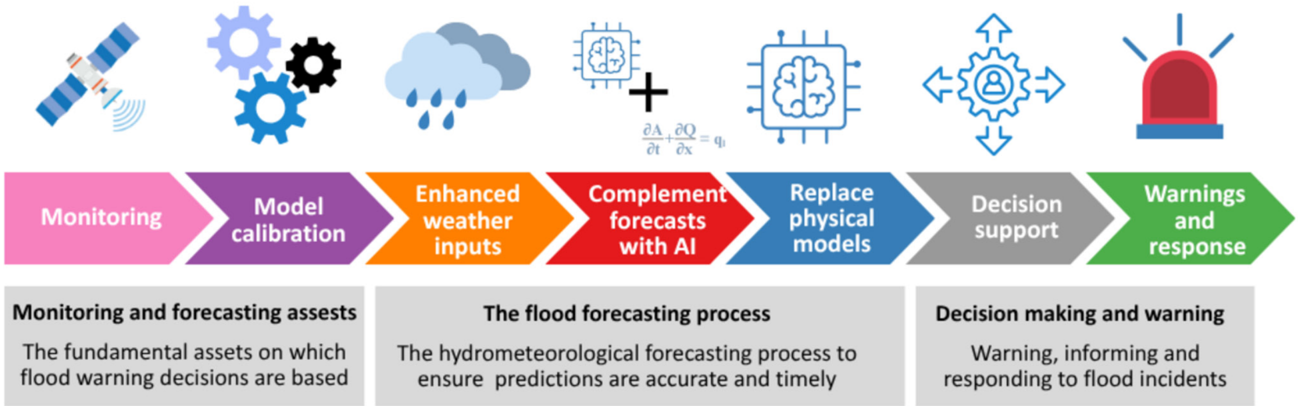



**Figure 1: Conceptual framework identifying seven flood forecasting phases where AI/ML can integrate with and/or enhance flood forecasting.**

| Table 1: The seven phases of the flood forecasting framework used in this study. | |
|---|---|
| **Phases** | **Description** |
| **Monitoring** | Not all the monitoring data available is being used to support flood forecasting. Combinations and gathering of data that can be incorporated in this phase include using ML (i.e., remote sensing, citizen science, cameras). This phase is explored to assess how AL/ML could help communities that are not supported by existing flood warning services. |
| **Model calibration** | Using ML to automatically recalibrate models (not in real time) and optimise existing tools could save demanding resource effort in the flood forecasting process. |
| **Integration of enhanced weather data** | Using ML to integrate flood forecasting with improved weather forecasts/data into existing models (e.g., improved precipitation forecasts). |
| **Complement existing flood forecasting approaches** | AI/ML could free up resources for high intensity/high impact events. 'Normal' conditions could potentially be replaced with AI/ML, but extreme events require the high-end models coupled with relevant flood forecasting expertise. This could also be extended to emulate complex processes with ML such as wave models, sea level rise, tidal surges, etc. |
| **Replacing flood forecasting with AI/ML** | This phase explores whether there is an opportunity to replace existing physically-based flood modelling assets using AI/ML techniques as well as making the calibration more efficient using AI/ML in real-time. |
| **Decision support** | This phase explores the potential of data mining from alternative libraries and repositories to assess flood impacts in real-time (or near real-time) such as real-time inundation, impact-based forecasting (e.g., forecasting how many houses a flood will affect). |
| **Warnings and response** | AI/ML automation of the response and warning system (e.g., road closures) is the final phase of the flood forecasting process. |

## 1.2 Project aim and objectives

The aim of the *DelugeAI* project, led by the University of Strathclyde, was to critically review the current state of AI technologies and methodologies in flood forecasting drawn from the latest research and assessments, evaluate its potential for flood forecasting and anticipatory actions in the Scottish context, assess the feasibility of incorporating AI/ML with SEPA's current flood forecasting capabilities, and provide recommendations for future research, implementation and operationalisation. *DelugeAI* achieved this aim through the following four core objectives and activities:

1. Undertake a systematic literature review analysing existing AI/ML methodologies and technologies applied to flood forecasting to identify gaps and opportunities;
2. Convene an online workshop with experts in AI/ML and hydrological flood forecasting to facilitate discussions on key challenges, opportunities and directions;
3. Undertake a feasibility study to assess the practicality of AI integration into SEPA's existing flood forecasting frameworks; and
4. Provide a set of recommendations for advancing AI/ML-driven flood forecasting, including a roadmap outlining potential priority developments and implementation pathways.

*DelugeAI* comprises an up-to-date evidence-based review that provides SEPA with a foundation that will inform of the potential of AI/ML and how it may be applied to flood forecasting to strengthen SEPA's safeguarding of Scotland's environment and population.

## 1.3 Structure of the report

This report is structured as follows: Section 2.0 provides an overview of the methods employed in the *DelugeAI* project; Section 3.0 summarizes the key findings from the critical literature review on the current state of AI/ML technologies and methodologies in flood forecasting, presents insights from an expert workshop evaluating their potential applications, and assesses the feasibility of integrating AI/ML into SEPA's existing flood forecasting systems; Section 4.0 discusses the key challenges and opportunities associated with the integration of AI/ML to improve flood forecasting, drawing from the experience and knowledge of other related sectors; Section 5.0 provides a set of recommendations together with a roadmap for their potential implementation and operationalisation; and Section 6.0 provides some concluding remarks.

# 2.0 Methods

## 2.1 Literature review

A structured literature search was conducted using Web of Science (WoS), a web-based bibliographic database. The aim was to identify publications at the intersection of flood forecasting and AI/ML. The search strategy was developed iteratively, refining and optimizing search terms for relevance and specificity. The core query targeted studies related to flood or inundation forecasting and AI methods, including both traditional ML, hybrid methods and deep learning architectures. The main search string combined terms for flooding (e.g., flood*, inundation) with forecasting-related terms (e.g., forecast*, predictive model*) and a comprehensive list of AI techniques and keywords (e.g., "machine learning", "deep learning", "neural networks", LSTM, CNN, "data-driven", "ensemble learning"). This was applied to both titles (TI) and abstracts (AB), and proximity operators (e.g., NEAR/15) were used to ensure conceptual relevance between terms. This query was applied across all publication years, restricted to the WoS Core Collection. The main query string is available in Appendix A.

The search was further refined using a combination of inclusion and exclusion filters by analysing the resulting keywords:

- **Inclusion:** Studies relevant to flood forecasting using AI, including both pure data-driven and hybrid approaches.
- **Exclusion:** Articles focused on precipitation, rainfall, sediment, runoff, reservoir operation, flood susceptibility, hazard mapping, and vulnerability assessments were removed using NOT filters in TI and topic (TS) fields as, although they can be related or used for flood forecasting, they are not strictly flood forecasting and do not fall within the seven identified phases. Additional unrelated applications (e.g., oil, spectroscopy, computer security, chemistry) were filtered out with key terms (e.g., oil, ship, $CO_2$, flood attack, fuel cell, etc.) from the TS and TI fields.

To explore the use of AI across the different phases of the flood forecasting chain, targeted sub-queries with relevant keywords and exclusion terms were developed and applied in combination with the main search. A list of the search queries is available in Appendix A. The topics covered for each theme include:

- **Monitoring:** using Internet of Things (IoT), citizen science, sensors, social media or remote sensing data. The search was performed in the Title AND Abstract.
- **Model calibration:** topics around offline learning, parameter estimation, error estimation or parameter tuning. The search was performed in the Title AND Abstract to reduce noise.
- **Integration of enhanced weather data:** incorporating weather, precipitation or atmospheric forecasts and bias correction, downscaling or ensemble forecasts. This search was conducted in the TS field.
- **Complement existing flood forecasting approaches:** focused on hybrid, surrogate and physics-informed models where ML is combined with traditional numerical models or where complex processes, such as storm surge or waves, are emulated. The search was performed in the TS OR TI fields, to ensure the key surrogate models were captured.
- **Replacing flood forecasting with AI/ML:** standalone data-driven models or real time calibration and learning. Surrogate, hybrid, statistical and physics-informed models were excluded. The search was constrained to Title AND Abstract to avoid returning the full query and having excessive noise.
- **Decision support:** based around flood impacts and knowledge discovery. The search was constrained to the title as often terms such as impacts or decision support are mentioned in passing as context or justification for flood-related research.
- **Warnings and response in real-time:** topics related to early warning systems (EWS), flood alerts and response strategies. This search was also constrained to the Title only.

It is important to note that records can belong to more than one category, either because they cover more than one phase or because they mention other opportunities in the abstract. This is particularly noticeable in the use of keywords such as "flood impact", "decision support" or "hydrological model" as context or justification. However, further overlap existed as the phase boundaries cannot be strictly defined. A particularly difficult term to classify was "hybrid model/forecast", as in the literature this term has many definitions: numerical modelling combined with ML; ML+ML, more than one type of ML model combined (e.g., CNN-LSTM model); statistical models used with numerical modelling; or statistical models combined with ML models. Although refinements were made to exclude data-driven non-AI models, it is possible some of these hybrid models were returned in the results.

Due to the large number of articles returned by the search and considering they span the entire flood forecasting workflow, it was not possible to further screen the results to establish further exclusion/inclusion criteria, and therefore there will be a large number of false positives. However, by tagging each model with a category, the results become more relevant and informative.

To complement the peer-reviewed literature, a targeted search for grey literature was conducted to capture government programmes, pilot projects, recent innovations and industry applications of AI in flood forecasting. The aim was to identify initiatives and developments not yet reflected in academic publications, but still highly relevant to real-world implementation. Grey literature was sourced from the following platforms:

- **Google and Google Scholar** – for broad discovery of technical reports, white papers, projects and presentations.
- **Government agencies** – including GOV.UK for reports related to UK Environment Agency and DEFRA, SEPA, NRW and other national hydrometeorological services.
- **International organisations** – such as UNESCO, WMO (World Meteorological Organization), ITU (International Telecommunication Union), World Bank Group, and UNDRR (UN Office of Disaster Risk Reduction).
- **Research and innovation platforms** – including UKRI, CORDIS (EU) and Horizon 2020/Europe project repositories.
- **Non-governmental and private sector sources** – including engineering consultancies (e.g., HR Wallingford, Jacobs) and technology firms (e.g., Microsoft, IBM).
- **Editorial and expert commentary platforms** – including Nature News and The Conversation

Limitations around the methodology mainly concern discoverability, language barriers and bias. Many relevant projects may not be well-documented online or lack standardized metadata, making them difficult to locate. In addition, many consultancies or agencies may be working behind the scenes on AI approaches but do not disclose this on their websites. Language barriers can also restrict access to local initiatives published in non-English sources. Furthermore, there is an inherent bias toward more prominent or well-funded projects with the resources to produce public-facing documentation, potentially overlooking smaller or community-led efforts that may be equally impactful. There is also more bias in the tagging of the forecasting stage as it depended on expert judgement rather than a systematic keyword match.

## 2.2 Expert workshop

A *DelugeAI* expert workshop titled 'Exploring the Use of Artificial Intelligence for Flood Forecasting in Scotland' was held on 23 April 2025 with support from CREW and SEPA and facilitated by the University of Strathclyde team led by Drs Chris White, Douglas Bertram and Robert Atkinson. As the second of four project stages, it followed an initial literature review and led the development of a feasibility study and a 5-year roadmap for AI/ML integration in flood forecasting.

The online workshop convened over 35 specialists from academia, government and industry. Their combined expertise in hydrology, hydraulics, meteorology, AI and data science ensured insights across the entire flood forecasting value chain. The objectives were to evaluate existing AI applications, establish research and operational priorities for Scotland, identify opportunities and discuss challenges.

Structured in three sessions; (1) the event opened with a contextual overview and keynote presentation, (2) progressed through a series of lightning talks accompanied by open discussion to surface opportunities and challenges, and (3) concluded with a horizon-scanning exercise that aligned potential AI applications with the development of a potential roadmap for implementation.

Appendix B provides additional details of the workshop including the agenda (B.1) and a list of attendees (B.2). B.3 provides details of the suggestions from the workshop attendees for AI developments in flood forecasting at the one, three and five-year horizons. Additional suggestions for areas of development were also offered.

## 2.3 Feasibility study

Building on the insights gained from the literature review and expert workshop, a feasibility study assessed the practical integration of AI into SEPA's operational flood forecasting system. While other phases of the DelugeAI project focused on conceptual understanding and stakeholder perspectives, this stage concentrated on evaluating the real-world viability of implementing AI solutions within SEPA's current and future workflows.

The feasibility assessment focused on the seven AI application areas that map directly to the flood forecasting phases introduced in Figure 1 and detailed in the introduction (Section 1.1). These phases reflect opportunities where AI and ML could meaningfully enhance forecasting performance – from initial data acquisition and model calibration to decision support and the delivery of real-time warnings. By aligning the feasibility study with this framework, the analysis maintains consistency across the project and ensures that recommendations are grounded in SEPA's operational context.

To ensure consistency with the project's wider framework and objectives, the feasibility study was structured around eight Multi-Criteria Decision Analysis (MCDA) steps. This method provided a systematic and transparent approach for evaluating each AI solution by combining expert judgment from SEPA's operational staff with weighted scoring aligned to strategic priorities. The use of MCDA ensured that complex trade-offs between technical feasibility, impact, and strategic alignment were addressed in a structured way. The next stages summarize the core implementation process, with further methodological detail available in Appendices C.1 to C.6.

Seven hypothetical AI solutions, one for each phase of SEPA's flood forecasting framework, were evaluated, spanning the full forecast chain from monitoring to warning response. Key stakeholders helped define the evaluation criteria and assign strategic weights to the five main evaluation categories. Each solution was scored by experts using a standardized 1-5 scoring scale, with results validated through peer review. Weighted scores were calculated to rank solutions, followed by an impact-effort analysis to assess implementation complexity. Final classifications identified priority solutions for potential adoption. Further detail can be found in Appendix C.

# 3.0 Results

## 3.1 Systematic literature review

A total of 1,845 papers across all years were identified in the final WoS dataset following the application of refined inclusion and exclusion criteria. The majority were research articles (1,470), including peer-reviewed journal articles (1,421), early access articles (25), and those also categorized as proceedings papers (15) or book chapters (6). Conference proceedings papers came to 318 entries. Additionally, there were 50 review papers and a small number of other types, such as editorial materials (5), corrections (2), and one retracted publication.

Annual publication trends revealed a rise in research activity from 2018, with exponential growth after 2020 (Figure 2), reflecting the increased interest in operational AI applications and climate adaptation technologies. Notably, 70% of the articles were published from 2020 onwards, highlighting a recent surge in academic focus on these topics. The rise of deep learning, especially the adoption of recurrent neural networks (RNNs) like the LSTM model, coincides with this period. LSTMs are especially effective at learning from sequential data because they include memory cells that help retain information over longer time periods, allowing them to remember previous inputs and capture long-term dependencies.

The majority of the articles focused on complementing or replacing the hydro-meteorological flood forecasts, whilst very little research has gone into model calibration, although the interest in this stage is more recent (Figure 2). The use of AI tools in combination with monitoring activities was the next most researched area, followed by decision support tools and using AI to integrate enhanced meteorological inputs. Focus on developing warning and response initiatives was low, but interest has risen in recent years, especially in 2024 (Figure 3).

The growth of end-to-end ML (replace) papers may reflect the accessibility of historical flood data, the maturity of DL frameworks, and the appeal of avoiding the complexity of physical modelling. However, hybrid models, where AI complements physics-based models, are also a rapidly expanding area (Figure 3), bridging the gap between operational practicality and data-driven innovation. These models stem from the need to improve the accuracy of physical models, make predictions faster and in real-time, and address the black-box nature and interpretability concerns of end-to-end

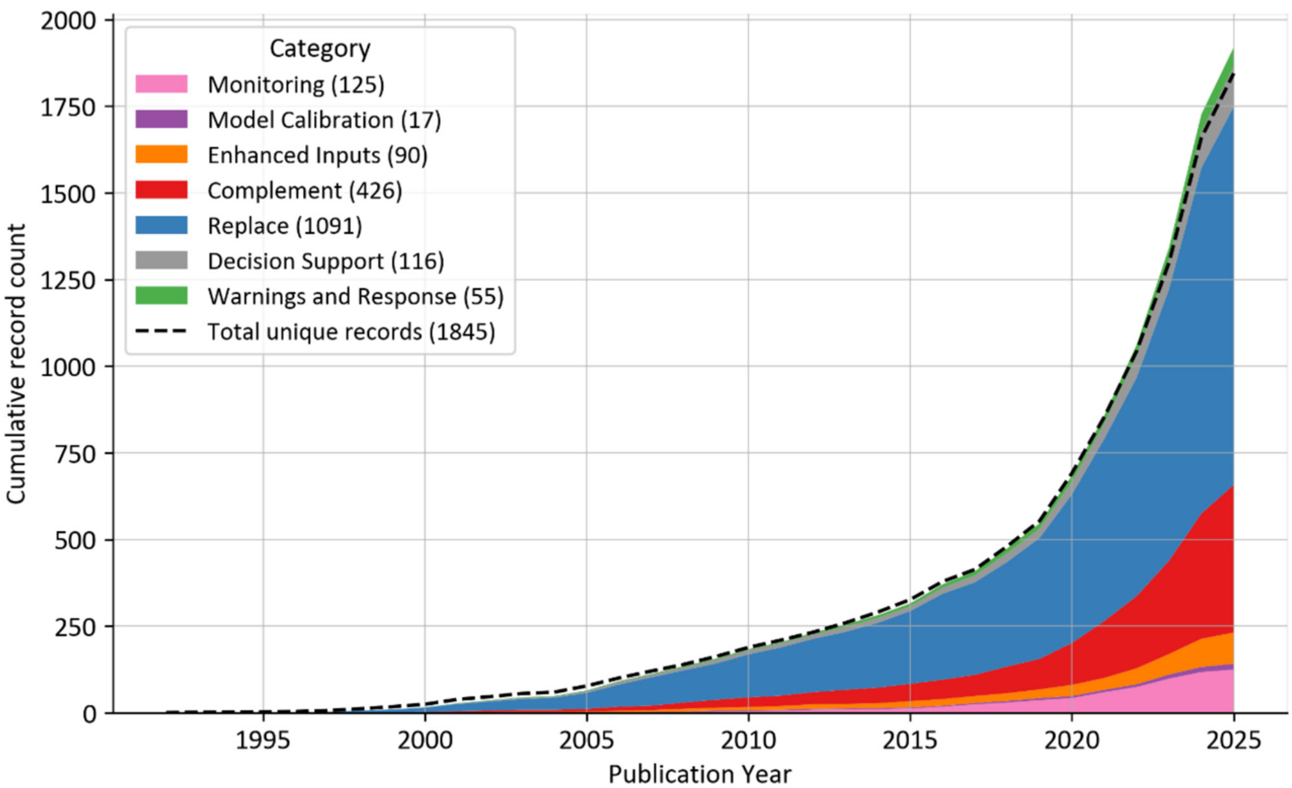



**Figure 2: Cumulative publication records over time stacked by forecasting stages (colours). Black dashed curve is total unique publications (no overlaps).**

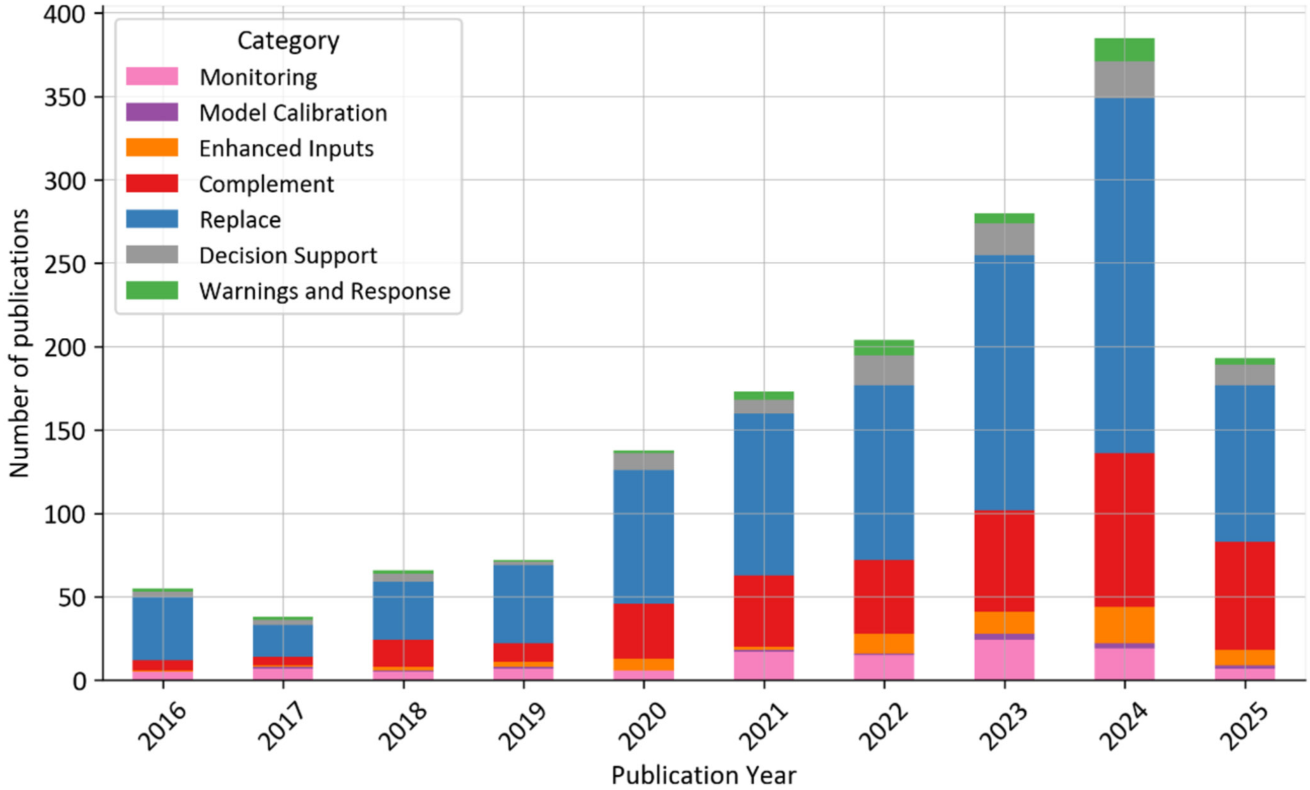



**Figure 3: Number of publications per year stacked by forecasting stage for the last 10 years. Note 2025 reflects only five months (Jan-May).**

ML models (Slater *et al*., 2023; Byaruhanga *et al*., 2024; Zhao *et al*., 2024).

A wide variety of ML models appeared in the results, from simple classical neural networks (e.g., multilayer perceptron (MLP)) to complex deep architectures with enhancements (e.g., STA-LSTM; Ding *et al*., 2020). To gain insights into model usage trends, a high-level keyword match was conducted across the title, abstract, and author keywords for the most frequently cited model types. These were grouped into five broad categories: traditional ML (e.g., random forests (RF) or support vector machines, SVM), deep learning (e.g., RNN, CNN, Transformers), ANN, generative models (e.g., generative adversarial networks (GANs)) and, finally, statistical and others (e.g., ARIMA (autoregressive integrated moving average) or fuzzy logic).

Earlier models were predominantly based on ANN structures, with traditional ML techniques appearing more frequently in the early 2000s, of which RF and SVM were the most used. ANNs remained the most common model type until the early 2020s. DL began to gain traction from around 2018 and, by 2024, had become the most widely used category. This trend aligns with the overall rise in publication numbers and reflects the growing adoption of DL techniques in flood forecasting. Lastly, although still a small proportion, the appearance of GANs from around 2021 highlights growing interest in advanced generative architectures for emulating flood dynamics. Within the deep learning category, the most used models are the LSTMs and CNNs, whilst the most common traditional models are SVMs and RFs. This analysis coincides with the results from comprehensive reviews on the use of AI for flood risk management (Liu *et al*., 2025) and for short-term flood forecasting (Asif *et al*., 2025).

The most common DL models are employed for different problems as they have different strengths and limitations (Asif *et al*., 2025). For image processing and spatial data, CNNs are a common tool (e.g., Guo *et al*., 2020). Time series or sequence problems are often addressed with recurrent networks, in particular LSTMs (e.g., Hunt *et al*., 2022), although GRUs are simpler and lighter models that are gaining traction in this area (Zhao *et al*., 2024). Generative models, such as GANs, are mostly employed to generate synthetic data (e.g., Weng *et al*., 2023) or fill data gaps, which is particularly valuable given the lack of abundance in extreme events like floods in many datasets. When reviewed and evaluated for short-term (<48 h) predictions, Asif *et al*. (2025) found that, overall, standalone models performed well, although the errors increased with lead times. Generally, LSTM performance was excellent from 1 to 12 hours while RF models performed best at 12 to 48 hours lead times. The best performing models across all lead times, however, were hybrid models that combined different ML algorithms, leveraging the strengths from each.

Combining AI with IoT and ground-based sensors (e.g., Mousavi *et al.*, 2021; Bande and Shete, 2017); cameras (e.g., Jafari *et al.*, 2021), or remote sensing (e.g., Lammers *et al.*, 2021) are common for monitoring before, during and after flooding events. These technologies are particularly useful for data scarce regions. However, drawbacks include the costs of implementation and maintenance of sensors, weather stations, cameras, etc.; the localized nature of these approaches; or the challenges around interoperability and data integration (Bukhari *et al.*, 2025).

Decision support tools increasingly incorporate ML algorithms to assess flood impacts – such as predicting damage to properties (e.g., Alipour *et al.*, 2020) or agricultural land (e.g., Jiang *et al.*, 2022). These models are often built using RFs, which are well-suited for classification and regression tasks. RFs are popular in this context due to their robust performance, low computational demand and relatively high interpretability compared to deep learning methods. Unlike many neural networks, RFs allow users to trace how decisions are made through feature importance metrics and tree structures, making them valuable for applications that require transparency and stakeholder trust.

Warning and response and EWS are the aim of many projects and research. However, they are not necessarily tackling the automation of warnings, but rather working towards accurate forecasting, faster predictions or data integration or aggregation from which a warning can be issued. Therefore, it is hard to distinguish in the literature search between the use of AI for informing the warning stage and for issuing warnings.

## 3.2 Grey literature review

In addition to the systematic literature review, a diverse set of flood forecasting-related projects, workgroups, and operational products were also identified in the grey (non-peer-reviewed) literature (ca. 50 projects) demonstrating how AI is rapidly (and increasingly) being applied to improve flood forecasting and management. Most initiatives focus on developing AI models to complement or accelerate traditional hydrodynamic approaches (Figure 4). There was also a notable emphasis on monitoring through sensor networks, satellite data integration, and enhancing forecasts for small or ungauged catchments where conventional data are limited.

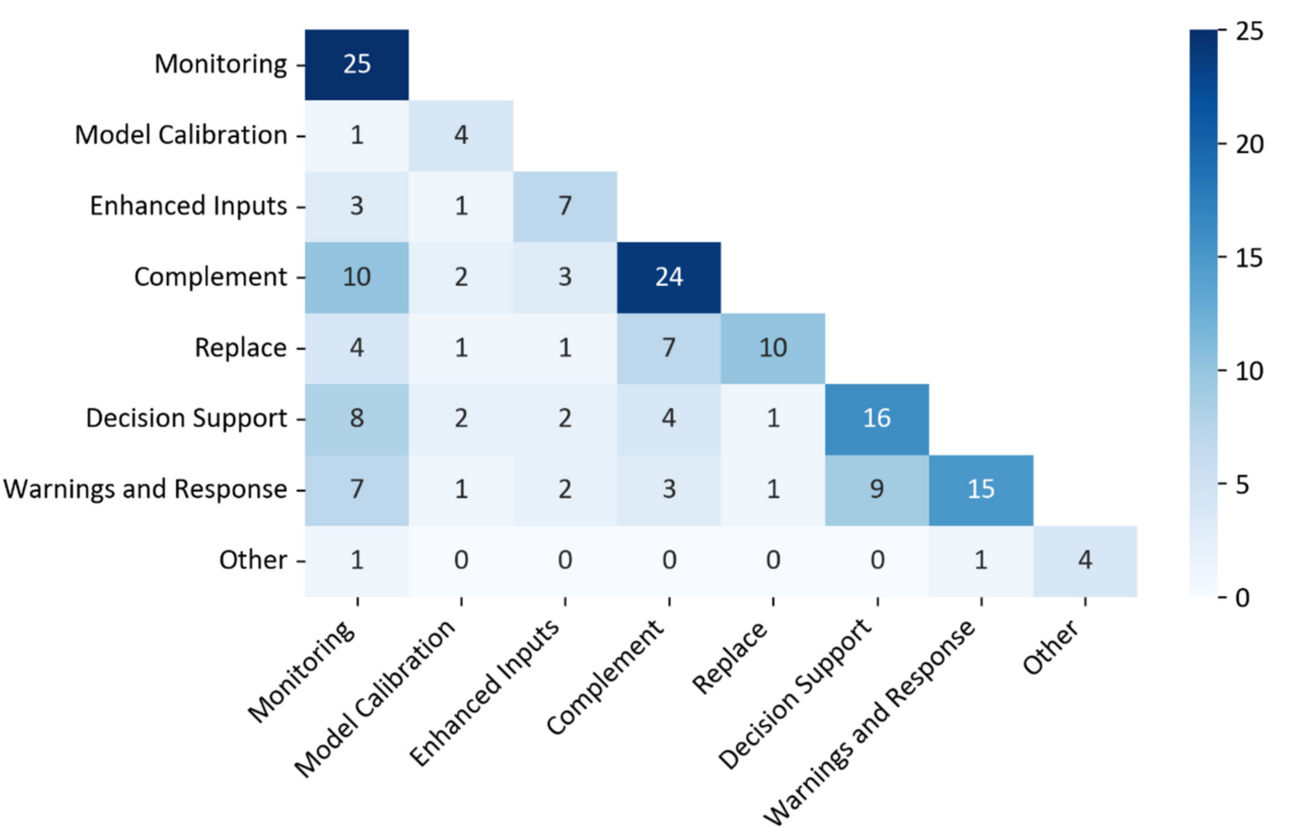



**Figure 4: Co-occurrence matrix of grey literature projects according to their forecasting stage. Numbers indicate the number of projects in each overlapping category. Total count per category is in the main diagonal.**

Early warning systems remain vital for protecting communities and building flood resilience and are being pushed by the United Nation's Early Warnings for All (EW4All) initiative. AI-driven technological advancements enable these systems to reach broader populations more efficiently. Consequently, growing interest is evident in AI applications for warnings and response, as well as decision support tools – particularly within the broader hazard, risk, and disaster mitigation frameworks that encompass flooding, for example under the UN's platform AI4Good or within the ITU AI for natural disaster management focus group (FG-AI4NDM).

AI-based decision support tools aim to translate flood forecasts into actionable insights. The AI-RiskAnalyzer (by FloodWaive) supports real-time flood risk evaluation through DL and localized data, helping assess impacts and respond effectively. The ISRV project (Aachen University) is developing an interactive tool for transport operators, combining AI forecasts with infrastructure risk to guide flood-related decisions.

There were very few projects specifically focused on model calibration or enhanced inputs within the flood forecasting domain (Figure 4). However, AI is actively used in related fields – such as weather forecasting – to develop models and forecasts that indirectly support flood forecasting, even when flood prediction is not their primary objective. These models are not covered by this review but future activity should examine developments in these areas as well, considering complementary activity, transferrable approaches, etc. that may benefit flood forecasting. Nonetheless, developments in these adjacent domains, particularly in improving atmospheric and precipitation forecasts using AI, may hold significant potential for future integration into flood forecasting workflows and are worth monitoring. For instance, the UK's Met Office is developing an experimental ML-based weather forecast, the FastNet, a model based on graph neural networks (GNNs). Additionally, the European Centre for Medium-Range Weather Forecasts (ECMWF) recently made the AIFS operational, the AI-driven version of their integrated forecasting system (IFS). These AI-generated outputs could serve as inputs in future flood forecasts.

Most examples lack clear technical detail, and terms like AI or ML are mentioned without further elaboration – particularly in cases where AI is applied mainly for data processing or fusion, rather than for predictive modelling. Deep Learning techniques appear more frequently in the forecasting stages, where LSTMs (e.g., used in HydroForecast or HydroSphereAI products or by the AI4Flood team) or convolutional networks (e.g., in projects by NIWA or JBA) are employed. These models are used to either accelerate and complement traditional models (e.g., AI-FloodCast), or to replace them entirely, as in the models developed through the FruítPunch AI "AI for Inland Flood Prediction" challenge. CNNs are mainly used for flood extent mapping and spatial pattern recognition; for instance, FloodSENS is based on a U-Net architecture. Additionally, several cases showcase how using large language models (LLMs) can serve as warning and response mechanisms, such as UNESCO's flood awareness chatbot.

A wide variety of data sources are used across the projects, although technical details are often limited. Satellite data, such as Sentinel imagery, are employed for flood mapping, real-time monitoring, or as inputs into AI models (e.g., the CAMEO project in Ireland or DTU's Wet Index tool). Sensor networks and in-situ measurements, including river gauges and IoT devices, are also widely used – particularly for real-time monitoring in small catchments (e.g., KI-HopE-De project in Germany or FloodAI in England). AI is also applied to transform citizen science inputs or crowdsourced data into usable hydrological information that can feed into flood models, as is done in CrowdWater, OpenSafe Fusion or BluPix. However, the integration of multi-source datasets is usually described in broad terms, with little information on how the data is processed, cleaned or fused with AI. Projects like Wet Index (developed by DTU), HüPros (by Aachen University) or FAST (from Intellialert Technologies) aim to integrate diverse geospatial and hydrological data – ranging from soil moisture to seawater levels – into unified platforms that lead to flood forecasts and warnings. Detailed documentation of their data processing workflows remains scarce, making it difficult to assess scalability and robustness.

Several initiatives apply digital twin concepts to flood management. Projects like DestinE (Destination Earth) aim to build large-scale digital twins of the Earth system, providing enhanced weather input data with flood-related applications. More localized efforts such as PYRAMID and FLOODTWIN use digital twins to simulate urban flooding scenarios by integrating live sensor data, weather forecasts, and physical models. AI can be integrated into digital twins in many ways, for example, the PYRAMID project used ML to identify floating debris from floods and establish risk. However, this same project, led by researchers at Newcastle and Loughborough Universities, remarked that the technical challenges around developing a digital twin are considerable. In this case, they are particularly constrained by the quality of the driving precipitation datasets.

AI-based flood forecasting initiatives range from hyper-local small catchment solutions to global platforms or products designed for scalability and transferability. Local solutions often emerge from immediate needs to increase community flood resilience, are tailored to the specific infrastructure and hydrological characteristics of an area and often rely on the monitoring assets to inform early warning systems. On the other hand, larger complementary forecasting tools (e.g., HydroForecast, HydroSphereAI, FloodMapp, etc.) are usually products that rely on weather forecasts or satellite data, and are transferable across regions with some model calibration or tuning for local relevance.

These diverse applications reflect the different actors involved. At the local level, council-led initiatives often partner with environmental consultancies or research institutions to address specific urban or regional challenges. For example, Auckland Council, in collaboration with Mott Macdonald, are working to develop an ML solution for predicting real-time surface flooding. Other initiatives, like CENTAUR, are driven by university-led research with a focus on enhancing decision support at the urban scale. National agencies and government research bodies, such as NIWA (New Zealand) or DMI (Denmark), operate within more formal governance structures and are responsible for operational flood models and services. Meanwhile, environmental consultancies and tech companies often lead the development of proprietary forecasting products, either independently or through public-private partnerships. On a global scale, large consortia coordinated by organisations such as WMO or UNDRR aim to create generalisable tools, platforms or working groups. Additionally, organisations such as the Red Cross Red Crescent Climate Centre or UNESCO focus on early warning systems and decision support in under-resourced or high-risk regions, often with a focus on capacity building and humanitarian aid response.

## 3.3 Expert workshop

The workshop underscored that hybrid AI-hydrology approaches reliably enhance flood forecasts. Combining physical rainfall-runoff or hydraulic simulations with ML post-processing produced more accurate river-flow and extreme event predictions, demonstrating clear benefits over standalone models. Participants agreed that AI can obviously accelerate forecasting workflows whether by fusing remote sensing observations with gauge data, post-processing ensemble outputs or using convolutional networks to reconstruct flood extents under cloud cover while emphasising that human expertise must remain central to interpretation and warning issuance.

A systematic review revealed a rapidly expanding global landscape of AI in flood forecasting since 2018, ranging from large-scale platforms (for example, Google's Flood Hub) to regional projects such as FloodCast, FloodAI and FloodWaive. In virtually every case, AI complemented rather than supplanted existing forecasting systems, delivering decision support at multiple junctures of the forecast chain. Expert discussions highlighted common concerns around "black box" models, data quality and the need to keep a human in the loop to maintain trust, particularly in high stakes scenarios.

Participants reached a clear set of findings that highlight AI's potential while emphasising the need for robust scientific and ethical practices. Everyone agreed that hybrid methods, which combine traditional hydrological models with ML, produce noticeably better forecasts for both everyday flows and extreme events. The experts pointed out that AI can speed up tasks such as gathering remote sensing data and creating flood maps, but they also stressed that human judgement remains essential for interpreting results and issuing warnings. The group of experts noted that since 2018 there has been a rapid increase in AI tools that support rather than replace existing forecasting systems, and they identified common challenges around data quality, model transparency and staff training.

The workshop's sessions highlighted several essential discussion points aimed at shaping the future role of AI in Scotland's flood forecasting. Participants explored how to translate the insights gained into practical steps, emphasising the need for balance between technological innovation and human expertise. These discussion points reflect the collective priorities and concerns raised by experts throughout the event:

- Emphasise low risk, quick win AI pilots that maintain human oversight to build confidence and demonstrate value.
- Strengthen data quality through accurate QA/QC procedures, expanded sensor networks and shared repositories to support reliable AI forecasting.

- Design transparent, impact-based warning systems that clearly communicate risks and tailor messages according to user.
- Invest in targeted training to equip forecasters with AI fundamentals, ethical best practices and model interpretation skills, ensuring human expertise remains central.
- Establish a multi-stakeholder group to design ethical frameworks and best practices, addressing transparency, equity, legal considerations and human oversight.

Appendix B.3 provides details of workshop attendees suggestions for AI developments in Flood Forecasting at the one, three and five-year horizons. Additional suggestions for areas of development were also offered and recorded in Appendix B.3.

## 3.4 Feasibility study

The MCDA evaluation produced comprehensive scoring profiles across all seven AI solutions, revealing significant variation in performance across the five assessment criteria. Solutions demonstrated diverse strengths, with some excelling in technical capabilities while others showed superior deployment feasibility or cost-effectiveness. The analysis identified clear patterns that informed the prioritisation of solutions for SEPA's implementation strategy.

Following consensus building with the stakeholder group, a criteria weighting for the project was adopted as shown below in Table 2. It is noted that decisions in this exercise are dependent on values chosen that require both a degree of consensus building and an understanding of sensitivity. Repeating the exercise at regular intervals and with different stakeholders' representations would be beneficial.

The aggregated weighted scores revealed a clear hierarchy of implementation priorities, as detailed in Table 2. High-priority solutions, highlighted in green (scoring above 3.0) emerged as those offering the optimal balance of technical capability, practical feasibility, and strategic value for SEPA's operational context.

The MCDA analysis reveals a clear strategic pathway for SEPA's AI adoption in flood forecasting. The emergence of Warnings and Response and Decision Support as high-scoring, easily deployable solutions provides an opportunity for immediate impact while building organisational capability and confidence. These solutions can serve as stepping stones toward more ambitious technical implementations like comprehensive Weather Input integration and eventual replacement of forecasting with AI initiatives. A breakdown of the feasibility study results is available in Appendix C.7 and Appendix C.8.

The analysis demonstrates that technical excellence alone does not guarantee high priority status. Solutions must balance technical capability with practical considerations of deployment feasibility, cost management, and alignment with organisational objectives. This balanced approach ensures sustainable AI adoption that delivers tangible benefits while building toward more transformative long-term capabilities.

Moving forward, SEPA should consider a phased implementation approach that begins with high-priority, low-effort solutions while developing the infrastructure and capabilities necessary

**Table 2: MCDA scoring matrix for AI flood forecasting solutions. This table presents the detailed scoring breakdown for each AI solution across the five assessment criteria, along with the calculated overall weighted scores. Scores range from 1 (Very low feasibility) to 5 (Very high feasibility) for each criterion. Solutions are ranked by their overall weighted scores, clearly distinguishing high-priority solutions (≥3.0) from medium-priority options. Green highlighting indicates high-priority solutions whereas amber highlighting indicates medium-priority solutions. For further information on weighting and scoring see Appendices C.4 and C.5.**

| Category | Weight (%) | Monitor | Model Calibration | Enhanced weather inputs | Forecast (Complement) | Forecast (Replace) | Decision Support | Warnings and Response |
|---|---|---|---|---|---|---|---|---|
| Technical | 30% | 3 | 3 | 4 | 4 | 4 | 3 | 2 |
| Deployment | 15% | 1 | 4 | 3 | 1 | 2 | 4 | 4 |
| Improving Flood Resilience | 25% | 4 | 4 | 3 | 4 | 4 | 2 | 4 |
| Cost/ Resource | 20% | 3 | 2 | 3 | 2 | 2 | 4 | 4 |
| Sustainability & Ethics | 10% | 3 | 2 | 3 | 1 | 2 | 4 | 4 |
| **Total** | **100%** | **2.95** | **3.1** | **3.3** | **2.85** | **3.1** | **3.2** | **3.4** |

for more complex AI integrations. This strategy maximizes both short-term impact and long-term transformation potential in flood forecasting capabilities.

### 3.5 Summary

The development of AI/ML in flood forecasting reflects a dynamic interplay between academic innovation and real-world implementation. The academic literature exhibits a strong focus on advancing model architectures, pushing boundaries and exploring the theoretical potential of AI to enhance or, in some cases, fully replace traditional hydrometeorological forecasting systems. The exponential rise in publications since 2018 – particularly those employing deep learning methods such as LSTM and CNN – coincides with broader interest in AI for climate resilience and adaptation. Academic efforts frequently target performance metrics, benchmark comparisons and algorithmic novelty, often just very slightly improving a previous model with a more complex alternative.

In contrast, the grey literature and the expert workshop reveal a more pragmatic approach. Projects led by environmental agencies, engineering consultancies, and humanitarian organisations demonstrate a preference for using AI to complement and strengthen existing forecasting tools rather than displace them. These implementations are solution-driven and prioritize feasibility, speed, usability, system integration, and interpretability. This was echoed by the experts during the workshop, highlighting the improvement of both routine and extreme predictions with hybrid modelling.

Hybrid models, where AI complements traditional models, as well as aiding interpretability (Slater *et al*., 2023; Zhao *et al*., 2024), have the ability to integrate large datasets and various sources of data at once, including outputs from ensemble forecasts, helping reduce and address uncertainties (Slater *et al*., 2023; Byaruhanga *et al*., 2024), an overlooked point in much of the literature on end-to-end ML models. Therefore, ensemble flood forecasts are also easier to produce with coupled models (Byaruhanga *et al*., 2024). However, as highlighted in the workshop, AI tools cannot yet replace forecaster judgement.

Both in academic and grey literature, LSTM networks emerge amongst the most frequently applied deep learning models, showing influence between research priorities and practical feasibility. The widespread adoption of LSTMs reflects their balance between performance and interpretability, as well as their ability to handle sequential data. In addition, they are mature enough that many toolkits and tutorials exist, lowering the barrier to entry for both researchers and practitioners. This popularity is echoed in grey literature applications where reliability and clarity are critical for decision-making in operational settings. CNNs were also noted in the workshop for their utility in reconstructing flood events under cloud cover, supporting rapid post-event assessments and large-scale inundation mapping.

Recent academic work, however, is increasingly characterized by the exploration of more complex and data-intensive architectures such as Transformers, GNNs, and hybrid ensemble approaches. While these models often achieve marginal gains in performance benchmarks, they typically require larger datasets, higher computational power, and more intricate training procedures. As a result, their practical applicability in real-world, resource-constrained environments remains limited for now, with potential for future improvements as better systems or more resources become available. The type of model used depends on the problem that is being solved. Fluvial flooding, for example, is a simpler time series problem that can be addressed with classical ML (Liu *et al*., 2025); whilst pluvial or coastal flooding present non-linear complex interactions between hydrometeorological parameters, requiring a model that can handle spatio-temporal patterns, usually deep learning (Liu *et al*., 2025).

As well as defining the problem, it is important to consider what data is available, both in terms of quantity and quality (Al-Rawas *et al*., 2024; Asif *et al*., 2025, Liu *et al*., 2025). These questions are vital to address model accuracy and overfitting (where the model learns the training data very well but cannot generalize to unseen data), impacting viability of the model outputs and the use of AI/ML. The answers to these questions should determine the complexity of the system (Byaruhanga *et al*., 2024). Other noted challenges around data include ensuring data integrity, quality control and diversity of sources, which also tackles reducing bias (Liu *et al*., 2025). Maintaining transparency and accountability in AI outputs and throughout all stages is still a concern for researchers and experts.

Monitoring is theoretically one of the most accessible and impactful entry points for integrating AI into flood forecasting systems. The frequency of digital twins, IoT-based infrastructures, and Big

Data platforms in the grey and academic literature demonstrate the use of AI to aggregate vast volumes of sensor, camera, or citizen-sourced data to detect patterns, anomalies, or early signs of risk. These tools are relatively low-risk to implement and can be especially valuable in underserved or at-risk areas, where moving from a no warning system to one based on tangible, real-time data can significantly improve outcomes and build public trust in AI-driven systems. AI-enhanced monitoring also supports community engagement, as it can incorporate citizen science and local knowledge, raising awareness and empowering populations to play an active role in risk reduction, and thus contributing to democratizing AI (Bukhari *et al*., 2024; Liu *et al*., 2025). However, implementation is not one-size-fits-all; local hydrological conditions, infrastructure maturity, and institutional capacity shape the potential benefits (Bukhari *et al*., 2025). This is also reflected in the feasibility study scoring. Where historical data exists but remains underutilized, AI can unlock value by training models to extend or improve existing observational networks, but this requires considerable time, effort and monetary investment. Solutions must balance technical capability with practical considerations of deployment feasibility, cost management, and alignment with organisational objectives.

Despite the growing breadth of research and application, both the academic and grey literature reveal important gaps in current AI-driven flood forecasting practice. One of the most notable is the limited attention to model calibration. There could be several reasons for this. For example, model calibration could be considered a hybrid model where AI is used to enhance a numerical model and therefore would be tagged under "complement". On these lines, there could also be a discovery gap as language around this task is not very specific and targeted, making it difficult to create a search string that captures everything.

Regarding the warning and response stage, although comparatively little research and emphasis has been given to this, we expect it to receive more focus as the UN EW4All initiative pushes innovation in the coming years, particularly for those who currently do not receive early warnings. Furthermore, the feasibility study found that these solutions score highly as they are deployable to provide impact and opportunities. The workshop stressed the importance of starting with low-risk AI pilots that maintain human-in-the-loop control to build institutional confidence. This aligns with the feasibility study recommendation of a phased AI adoption strategy. Starting with low-effort, high-impact solutions in decision support and warnings and response can deliver early wins while laying the groundwork for more advanced capabilities. This balanced pathway helps bridge technical potential with institutional readiness.

Projects in the grey literature typically involve government agencies, research institutions, and private sector technology providers. Multi-sector collaborations are common, facilitating data sharing and practical deployment. Knowledge sharing, open-source data and interdisciplinary collaboration have also been highlighted in the literature as a key necessity for fast application of AI in operational settings (Al-Rawas *et al*., 2024; Byaruhanga *et al*., 2024, Liu *et al*., 2025). In addition, although some AI implementations are simple, the more complex DL models require specific skills. The literature remarked that there will need to be changes in education and preparation of the workforce (Slater *et al*., 2025) and the experts highlighted the need for upskilling staff to interpret and apply AI driven insights.

# 4.0 Discussion

From the expert workshop, it was evident that recurrent ML models, particularly LSTM networks, have become widely used in flood forecasting due to their ability to model temporal dependencies within time series data. This is also reflected in the literature review conducted (Section 3.0). Unlike other AI models that treat each input independently, LSTMs consider both current and past inputs, making them well suited for capturing patterns in sequences (time sampled data) such as rainfall, river discharge, or soil moisture over time. This capability enables them to forecast future events with improved temporal awareness.

Beyond temporal modelling, some researchers have employed two-dimensional (2D) CNNs to incorporate spatial context into flood predictions. These models can process spatially distributed data such as satellite imagery or gridded precipitation fields, enhancing the ability to detect regional patterns relevant to flood events. Again, the interest in these models is well reflected in the literature review.

Defining the problem clearly is therefore essential, particularly in spatial terms. Does the model need to provide a highly localized forecast based on timeseries data, cover multiple catchments, or operate at a national scale? Equally important to this problem definition is the question of data availability: what type, resolution, and quality of data are accessible to train the model? The quality of the input data will ultimately determine the quality of the model outputs – if the training data are poor, the predictions will be as well. In terms of potential implementation, understanding the technical constraints must be placed in consideration alongside any potential policy or financial constraints as well. This activity would likely form a key part of future phased consideration of AI adoption and implementation.

There was also broad agreement between the views expressed in the workshop and use-case as exposed by the literature review: AI models are often used in tandem with traditional hydrological models, resulting in hybrid frameworks that combine the strengths of data-driven learning with physics-based simulation. Common integration strategies often take the form of:

- AI2Hydro: an AI model processes raw input data, such as meteorological observations or remote sensing imagery and its outputs serve as inputs to a hydrological model. For example, an AI model might estimate soil moisture or rainfall intensity from satellite data before passing those estimates to a rainfall-runoff model. AI models can also be used to automate data preprocessing pipelines by identifying errors or gaps in raw readings from sensors.
- Hydro2AI: the hydrological model performs initial simulation, and its outputs are fed into an AI model that refines the forecast. This setup allows the AI model to correct for systematic errors or biases in the hydrological model, improving predictive accuracy.
- TunedHydro: here, the AI model is used to optimize the parameters of the hydrological model itself, such as infiltration rates or routing coefficients, leading to better calibration and overall model performance.
- ModelOfModel: in this configuration, an AI model is trained to replicate the outputs of a hydrological model rather than the physical system directly. Once trained, the AI surrogate can produce results much faster than the original model, making it suitable for real-time forecasting or large-scale scenario testing.

The workshop participants expressed interest in the use of generative AI in the broader business context:

- Intelligent decision support: in this context, using AI models to summarize forecasts, support managerial approval processes, and explain decisions (e.g., why a warning was not issued); also includes using LLMs to enable natural (human) language queries of forecast data.
- Media Monitoring: using AI models to gain an overview of citizen perspectives via social media monitoring.
- Digitisation: the use of AI models as part of digital transformation processes such as the conversion of paper records to digital (machine-readable) formats.
- Productivity: the use of LLMs to summarize complex information and produce the initial draft of technical reports.

In operational settings, both AI and hydrological models often require periodic retraining to maintain performance. This is particularly significant in environments where the underlying hydrological characteristics – such as land use, saturation levels, or drainage patterns – evolve over time and is critical in the context of a changing climate. While such retraining practices are sometimes informal, they are crucial for ensuring that the model continues to reliably reflect current conditions and remains reliable. AI models offer a key advantage, they can be retrained more quickly and frequently than traditional models, enabling more responsive adaptation to evolving data. In addition, they can leverage transfer learning, where knowledge gained from one region or dataset is reused to improve model performance in another, often with limited additional training. This allows AI systems to generalize across basins or regions more efficiently than traditional hydrological models, which typically require extensive recalibration for each new setting. Such flexibility makes AI particularly attractive for operational use in data-scarce or rapidly changing environments.

There is broad acknowledgement that for AI models to be utilized to their fullest capability, significant upskilling is required within the pre-existing staff within the workforce. A workforce that has the necessary AI skill set should be able to advise which computing platform(s) should be used, and so upskilling may solve this particular issue. SEPA have recently adopted Microsoft Copilot for general use across the organisation; a more AI-skilled workforce will likely develop as uptake and experience increases. The organisation has also adopted a 'cloud-first' approach to computing and data management. Deeper examination of workforce upskilling for digital transformations would likely feature as a key initiative for SEPA in the near future.

An area of interest is the application of explainable AI (XAI) to flood forecasting; however, for whom the explanations provided was less clear. Philosophically, the explanation for any event is shaped by the perspective, or specialism, of the observer, each of whom emphasizes a different causal factor. This idea is succinctly captured by Hanson (1958), who stated, "There are as many causes of x as there are explanations of x." Consider, for example, a fatal car accident at a hazardous junction. A medic might attribute the death to multiple haemorrhages; a lawyer may view it as a case of driver negligence; a road planner might highlight poor road design, such as obscuring shrubbery as the root cause. None of these explanations is inherently more correct than the others; each is valid within its own disciplinary lens. The corollary is that different stakeholders may seek different types of explanations. In flood or drought forecasting, emergency planners may seek actionable insights on timing and severity; hydrologists may focus on model sensitivity to input variables; policymakers may be interested in broader socioeconomic implications.

The challenge is especially pronounced with deep learning models.  These systems develop highly opaque internal representations of the data they are trained on. To address this opacity, researchers have developed various strategies. One common approach is sensitivity analysis, where input variables are systematically varied to observe how changes affect the output. For instance, in a flood forecasting model, one might alter rainfall inputs to see how it influences predicted water levels. More advanced techniques like SHAP values or saliency maps are also used to gain insight into model decision-making. Another strategy is to use inherently interpretable models, such as decision trees or simple random forests, which allow for more straightforward explanation but often at the cost of predictive performance. While these simpler models may be easier to understand, they may not match the accuracy of deeper architectures, particularly when dealing with complex or high-volume data. Finally, coupling AI models with hydrological models, whether that is AI2Hydro, Hydro2AI, TunedHydro or ModelOfModel, has also proven to be an effective way of increasing trust, as here the outputs are, at some point, bound by physical constraints.

Deploying ML in industrial settings also raises issues of human trust. Workers may fear that automation will render their roles obsolete. However, in many cases, ML models handle repetitive or low-level analytical tasks, freeing human experts to focus on higher-order strategic or interpretive work. In flood risk management, for example, a model might continuously monitor weather data (and other data sources) and predict flood patterns, allowing the domain specialists to concentrate on response coordination rather than real-time data crunching.

Another barrier is the reluctance of domain experts to trust models whose internal logic they cannot scrutinize. This scepticism often fades as the model demonstrates reliability and improves decision-making (forecasting) over time. As with any innovation, trust in AI systems tends to grow through iterative use, validation, and integration into established workflows. Transition and overlap periods during implementation of AI, supported by broad engagement approaches, will promote better outcomes for SEPA and other stakeholders.

Therefore, defining the required level of performance and accuracy for each problem, and asking how explainable the model needs to be and for whom, will be key. Together with defining the spatial scale and data availability, this will not only guide the choice of model but also help ensure that the models used are both robust and trusted by experts and the public.

In summary, it can be seen there is considerable agreement between the views and interest expressed during the expert workshop and the literature review on the classes of model used (e.g., LSTM, CNN, etc.) and the specific application (e.g., AI2Hydro, Hydro2AI). However, the workshop did expose some interests not well represented in the literature such as media monitoring, digitisation and productivity. It should be noted that these are more generic interest, not confined to flood forecasting, and will be represented in broader business domains and addressed by tool vendors (e.g., Microsoft), making implementation of these types of tools at the organisation level easier, as reflected by the feasibility study.

# 5.0 Potential roadmap and recommendations

## 5.1 Policy-aligned AI development

Based on the results presented in this report, SEPA's first steps should be leveraging AI to enhance its flood forecasting and warning and response capabilities – in this area its impact on the public and stakeholders can be immediate and apparent. This initiative aligns with the broader goals of the Scottish Government's current AI strategy, which emphasizes trustworthy, ethical, and inclusive AI applications (Scottish Government, 2021a), and note that a future update to the AI Strategy is anticipated later this year. By integrating AI, SEPA can improve accuracy and timeliness of flood predictions, thereby enhancing community preparedness and resilience against flood events. This aligns with the Scottish Government vision for AI in Scotland (Scottish Government, 2021a) and our world leadership in the use of trustworthy, ethical and inclusive AI, and for building resilient communities. For flood forecasting, a focus on trustworthy AI will be key to successful warn and inform activity, as well as robust decision-making support.

The recent and timely release of the Scottish Government Programme for Government confirms investment in the Digital Programme, with funding already made available to the public sector. SEPA also focus on delivering digital transformation in their current Annual Operating Plans (SEPA, 2024), seeking to drive operational efficiencies, be an enabler for an agile organisation, and support partnership collaboration. SEPA's AI strategy should focus on developing and analysing large datasets, recognising patterns and providing real-time predictions through complimentary or hybrid AI/ML systems. This approach is consistent with the Scottish Government's vision of using Digital Systems to drive innovation and improve public services (Scottish Government, 2021b). The Flood Warning Development Framework (2022–2028) outlines SEPA's commitment to adopting digital technologies to support its flood forecasting efforts, maintaining and improving existing services while upgrading capabilities through development and innovation, and delivering enhanced communications (SEPA, 2022).

The UK National AI Strategy underscores the importance of AI in increasing resilience, productivity, and innovation across various sectors (UK Government, 2021). SEPA's AI-driven flood forecasting aligns with this strategy by contributing to the UK's goal of becoming a global leader in AI technology and Digital Developments (UK Government, 2021; Department for Digital, Culture, Media & Sport, 2022). The integration of AI in flood forecasting not only supports wider, overarching strategies but also positions SEPA as a forward-thinking agency committed to leveraging cutting-edge technology for public safety and environmental protection. The parallels to the UK Met Office programmes developing AI for forecasting (Met Office, no date) are clear, and each agency can learn much from the other in terms of AI use, and in partnership working for AI delivery.

Beyond the UK, European partners are looking at digital developments that can transform activities, including the integration of AI/ML in catchment early-warning systems (e.g., the AI-based flood early-warning platform for small rivers and catchments project funded by Enterprise Europe Network), next-generation weather forecasting (e.g., DMI), flood forecasting (e.g., EO4FLOOD — Earth observation data for advancing flood forecasting by the European Space Agency) and more, while further afield others like the NIWA in New Zealand are developing AI and deep learning approaches to enhance flood mapping. This initiative by SEPA that led to this *DelugeAI* report positions them at the cutting edge of exploring AI/ML for flood forecasting.

## 5.2 Recommendations and potential roadmap

To enhance flood forecasting, the *DelugeAI* project recommends SEPA adopt a phased approach (Figure 5). Start with high-priority, low-effort solutions while building the infrastructure for advanced AI/ML integrations. This ensures immediate benefits and sets the stage for long-term improvements.

This *DelugeAI* review, coupled with international expert input, suggests starting with low-risk, quick-win AI pilots that maintain human oversight to build confidence and demonstrate value, aligning with national AI visions. The MCDA analysis outlines a strategic pathway for SEPA's AI adoption in flood forecasting, targeting a **1–2-year horizon**. High-scoring solutions like Warnings and Response and Decision Support (Table 2) offer impact and build organisational capability, paving the way for ambitious projects like Weather Prediction integration and Forecasting Replacement initiatives. This would build the community of practice amongst stakeholders and develop SEPA's AI literacy. Developing clear, impact-focused warning systems that communicate risks and tailor messages based on the users will enhance the AI offering. This approach will build resilient communities and foster trust in AI solutions, meeting Government vision. In parallel, SEPA should identify what data is currently available and what gaps exist, and to establish if these gaps can be covered to allow for (or to not limit) future AI/ML implementations.

In parallel, SEPA should identify what data is currently available and what gaps exist, and to establish if these gaps can be covered to allow for (or to not limit) future AI/ML implementations.

More complex AI/ML solutions would require longer **3–5-year horizons**. Monitoring, particularly emphasized in grey literature, involves local communities to increase trust in AI/ML enhanced forecasts but requires significant investment and is

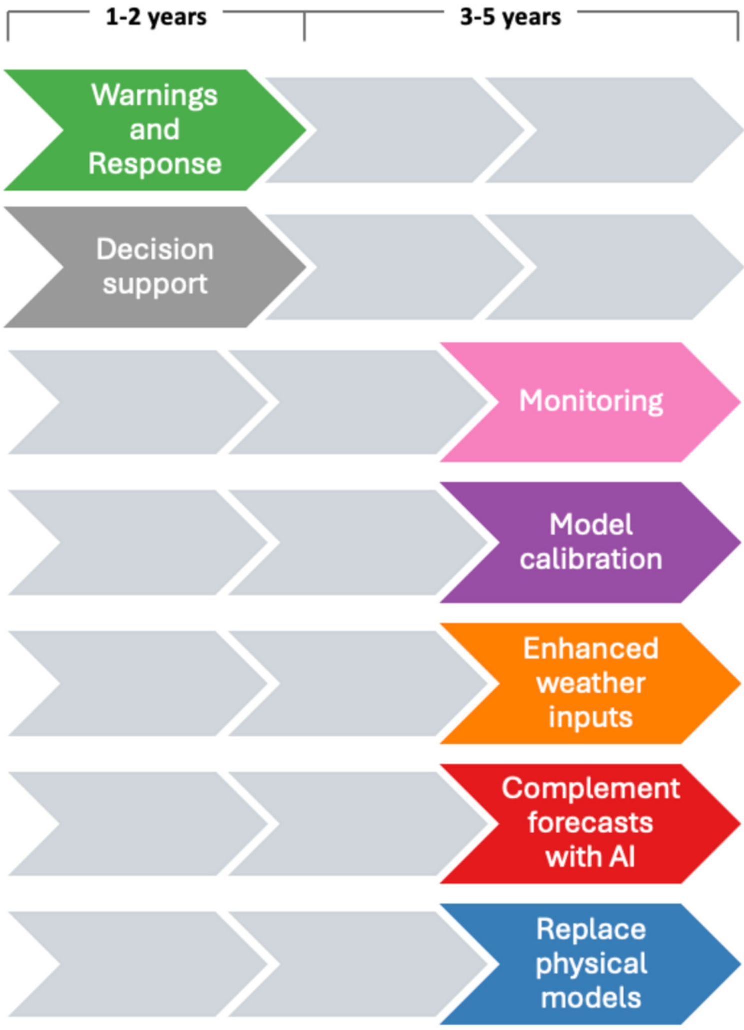



**Figure 5: *DelugeAI* recommended 5-year potential roadmap for using AI for flood forecasting in Scotland.**

often localised. Model calibration and integration of enhanced inputs are underexplored but could complement existing approaches. Academic research focuses on replacing forecasts with end-to-end ML models, while grey literature and operational settings favour hybrid models (ML + physics) for their combined benefits, including ensemble forecasting and reduced uncertainties.

Staff training and upskilling and development activities will form a key part of the AI and digital transitions. Investing in targeted training to equip flood forecasters with AI fundamentals, ethical best practices, and model interpretation skills, ensuring human expertise remains central. Additionally, establishing a multi-stakeholder group to design ethical frameworks and best practices, addressing transparency, equity, legal considerations, and human oversight is a solid first step for SEPA.

As SEPA explores the use of AI/ML, it will be important for the organisation to consider developing new tools as well as adopting and integrating with existing ones. Several AI-driven flood forecasting models already exist, such as Google's open-access Flood Hub, or commercial platforms like HydroForecast and HydroSphereAI. These vary in accessibility, from open access services to licensed products requiring formal partnerships. Future work should assess which external tools are suitable for Scotland's flood forecasting needs and resources and how SEPA might engage with them, whether by adapting, integrating, or collaborating. By upskilling staff, it will also support innovation and making use of readily available open-source code (e.g., NeuralHydrology or ML4Floods python packages or EdgeImpulse's river level predictor tool). The Feasibility Study can be repeated and/or adapted for future developments in AI and digital transformation to support this approach. SEPA are recommended to do so at regular intervals through the planning and implementation phases of activity, noting the fast pace and dynamic nature of the AI/ML landscape may see rapid and currently unforeseen changes in capability and opportunities.

# 6.0 Conclusions

This *DelugeAI* project critically reviewed the current state of AI/ML technologies and methodologies in flood forecasting to inform SEPA's future development plans. The report highlights that research to date into AI/ML for flood forecasting predominantly centres around two categories: replacement and complement. Academic literature focuses on replacing traditional models using ML, particularly with time series and spatial data through LSTMs and CNNs. In contrast, the grey literature and operational contexts favour hybrid approaches that combine ML with physics-based models, leveraging the speed and data integration of AI with the interpretability and constraints of traditional methods. Monitoring solutions, often involving local communities, are emphasized in grey literature but are typically localized and resource-intensive, with limited clarity on replicability more widely.

Insights from a workshop, which brought together international experts spanning hydrology, forecasting and AI/ML, reinforce the value of hybrid modelling in improving flood predictions, especially for both routine and extreme events. AI/ML applications show promise in operational support, such as accelerating data processes and enhancing satellite image interpretation through CNNs. While global AI/ML adoption in flood forecasting is expanding, most efforts are aimed at augmenting rather than replacing traditional methods. Notable challenges persist, including the need for transparency, data quality assurance, and workforce training to interpret AI-generated insights responsibly.

Finally, a feasibility study and potential roadmap, produced in collaboration with SEPA and the Scottish Government, suggests a pragmatic path for adopting AI/ML for flood forecasting, starting with low-effort, high-impact applications like early warnings and response and decision support. These offer immediate value and help build internal capabilities, paving the way for more sophisticated implementations such as full weather prediction integration and eventual model replacement. The analysis stresses that technical prowess must align with deployment feasibility and strategic priorities, recommending phased approaches of development and implementation to ensure sustainable and effective AI/ML integration into flood forecasting in Scotland.

# References


Al-Rawas, G., Nikoo, M. R., Al-Wardy, M., and Etri, T. (2024) 'A Critical Review of Emerging Technologies for Flash Flood Prediction: Examining Artificial Intelligence, Machine Learning, Internet of Things, Cloud Computing, and Robotics Techniques', *Water*, 16(14), 2069. Available at: https://doi.org/10.3390/w16142069.

Alipour, A., Ahmadalipour, A., Abbaszadeh, P. and Moradkhani, H. (2020) 'Leveraging machine learning for predicting flash flood damage in the Southeast US', *Environmental Research Letters*, 15(2), p. 024011. Available at: https://doi.org/10.1088/1748-9326/ab6edd.

Asif, M., Kuglitsch, M.M., Pelivan, I. and Albano, R. (2025) 'Review and Intercomparison of Machine Learning Applications for Short-term Flood Forecasting' *Water Resour Manage*, 39, 1971–1991. Available at: https://doi.org/10.1007/s11269-025-04093-x.

Bande, S. and Shete, V.V. (2017) 'Smart flood disaster prediction system using IoT & neural networks', in *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*. Bangalore: IEEE, pp. 189–194. Available at: https://doi.org/10.1109/SmartTechCon.2017.8358367.

Bukhari, S.A.S., Shafi, I., Ahmad, J., Gracia Villar, S., García Villena, E., Khurshaid, T. and Ashraf, I. (2025) 'Review of flood monitoring and prevention approaches: a data analytic perspective', *Natural Hazards*, 121(5), pp. 5103–5128. Available at: https://doi.org/10.1007/s11069-024-07050-w.

Byaruhanga, N., Kibirige, D., Gokool, S., and Mkhonta, G. (2024) 'Evolution of Flood Prediction and Forecasting Models for Flood Early Warning Systems: A Scoping Review', *Water,* 16(13), 1763. Available at: https://doi.org/10.3390/w16131763.

Cirri, G. (2023) 'AI for Climate Change: Managing floods using AI Early Warning Systems', *Oxford Insights*, 18 December. Available at: https://oxfordinsights.com/insights/ai-for-climate-change-managing-floods-using-ai-early-warning-systems/ (Accessed: 4 June 2025).

Department for Digital, Culture, Media & Sport (2022) UK Digital Strategy. Policy Paper. UK Government. Available at: https://www.gov.uk/government/publications/uks-digital-strategy/uk-digital-strategy (Accessed: 4 June 2025).

Ding, Y., Zhu, Y., Feng, J., Zhang, P. and Cheng, Z. (2020) 'Interpretable spatio-temporal attention LSTM model for flood forecasting', *Neurocomputing*, 403, pp. 348–359. Available at: https://doi.org/10.1016/j.neucom.2020.04.110.

Ghaffarian, S., Taghikhah, F.R. and Maier, H.R. (2023) 'Explainable artificial intelligence in disaster risk management: Achievements and prospective futures', *International Journal of Disaster Risk Reduction*, 98, p. 104123. Available at: https://doi.org/10.1016/j.ijdrr.2023.104123.

Guo, Z., Leitão, J.P., Simões, N.E. and Moosavi, V. (2021) 'Data-driven flood emulation: Speeding up urban flood predictions by deep convolutional neural networks', *Journal of Flood Risk Management*, 14(1), p. e12684. Available at: https://doi.org/10.1111/jfr3.12684.

Hanson, N.R. (1958) *Patterns of Discovery: An Inquiry into the Conceptual Foundations of Science*. New York: Press Syndicate of the University of Cambridge.

Hunt, K.M.R., Matthews, G.R., Pappenberger, F., and Prudhomme, C. (2022) 'Using a long short-term memory (LSTM) neural network to boost river streamflow forecasts over the western United States', *Hydrology and Earth System Sciences*, 26(21), pp. 5449–5472. Available at: https://doi.org/10.5194/hess-26-5449-2022.

Jafari, N.H., Chen, Q., Le, C.-X., Betzer, L.P. and Liang, Y. (2021) 'Real-time water level monitoring using live cameras and computer vision techniques', *Computers & Geosciences*, 147, p. 104642. Available at: https://doi.org/10.1016/j.cageo.2020.104642.

Jiang, Z., Yang, S., Liu, Z., Xu, Y., Xiong, Y., Qi, S., Pang, Q., Xu, J., Liu, F. and Xu, T. (2022) 'Coupling machine learning and weather forecast to predict farmland flood disaster: A case study in Yangtze River basin', *Environmental Modelling & Software*, 155, p. 105436. Available at: https://doi.org/10.1016/j.envsoft.2022.105436.

Lammers, R., Li, A., Nag, S. and Ravindra, V. (2021) 'Prediction models for urban flood evolution for satellite remote sensing', *Journal of Hydrology*, 603, p. 127175. Available at: https://doi.org/10.1016/j.jhydrol.2021.127175.

Liu, Z., Coleman, N., Patrascu, F.I., Yin, K., Li, X., and Mostafavi, A. (2025) 'Artificial intelligence for flood risk management: A comprehensive state-of-the-art review and future directions', *International Journal of Disaster Risk Reduction*, 117, p. 105110. Available at: https://doi.org/10.1016/j.ijdrr.2024.105110.

Met Office (no date) *Artificial Intelligence for Numerical Weather Prediction*, Met Office. Available at: https://www.metoffice.gov.uk/research/approach/collaboration/artificial-intelligence-for-numerical-weather-prediction (Accessed: 4 June 2025).

Moore, R.J. and Bell, V.A. (2001) 'Part 1: literature review of models', in *Comparison of rainfall-runoff models for flood forecasting*. Bristol: Environment Agency. Available at: https://assets.publishing.service.gov.uk/media/5a7cb3c440f0b65b3de0aa3d/strw241-e-e.pdf (Accessed: 4 June 2025).

Mousavi, F.S., Yousefi, S., Abghari, H. and Ghasemzadeh, A. (2021) 'Design of an IoT-based Flood Early Detection System using Machine Learning', in 2021 26th *International Computer Conference, Computer Society of Iran (CSICC)*, pp. 1–7. Available at: https://doi.org/10.1109/CSICC52343.2021.9420594.

Scottish Government (2021a) *Scotland's Artificial Intelligence Strategy - Trustworthy, Ethical and Inclusive.* Edinburgh: The Scottish Government. Available at: https://www.gov.scot/binaries/content/documents/govscot/publications/strategy-plan/2021/03/scotlands-ai-strategy-trustworthy-ethical-inclusive/documents/scotlands-artificial-intelligence-strategy-trustworthy-ethical-inclusive/scotlands-artificial-intelligence-strategy-trustworthy-ethical-inclusive/govscot%3Adocument/scotlands-artificial-intelligence-strategy-trustworthy-ethical-inclusive.pdf (Accessed: 4 June 2025).

Scottish Government (2021b) *A changing nation: how Scotland will thrive in a digital world*. Available at: https://www.gov.scot/publications/a-changing-nation-how-scotland-will-thrive-in-a-digital-world/pages/part-one-people-and-place/ (Accessed: 4 June 2025).

SEPA (2022) *Flood Warning Development Framework 2022-28*. Stirling: Scottish Environment Protection Agency. Available at: https://www.sepa.org.uk/media/594489/sepa-flood-warning-dev-framework-2022-28.pdf (Accessed: 4 June 2025).

SEPA (2024) *Our Annual Operating Plan 2024-2025*. Scottish Environment Protection Agency. Available at: https://beta.sepa.scot/media/svqbokvx/report-annual-operating-plan.pdf (Accessed: 20 June 2025).

Slater, L.J., Arnal, L., Boucher, M.-A., Chang, A.Y.-Y., Moulds, S., Murphy, C., Nearing, G., Shalev, G., Shen, C., Speight, L., Villarini, G., Wilby, R.L., Wood, A. and Zappa, M. (2023) 'Hybrid forecasting: blending climate predictions with AI models', *Hydrology and Earth System Sciences*, 27(9), pp. 1865–1889. Available at: https://doi.org/10.5194/hess-27-1865-2023.

Sniffer (2021) *Evidence for the third UK Climate Change Risk Assessment (CCRA3): Summary for Scotland*. Available at: https://www.ukclimaterisk.org/wp-content/uploads/2021/06/CCRA-Evidence-Report-Scotland-Summary-Final-1.pdf (Accessed: 4 June 2025).

UNU EHS (2024) *5 Ways AI Can Strengthen Early Warning Systems, United Nations University Environment and Human Security*. Available at: https://unu.edu/ehs/series/5-ways-ai-can-strengthen-early-warning-systems (Accessed: 4 June 2025).

UK Government (2021) *National AI Strategy*. Command paper 512. Office for Artificial Intelligence. Available at: https://assets.publishing.service.gov.uk/media/614db4d1e90e077a2cbdf3c4/National_AI_Strategy_-_PDF_version.pdf (Accessed: 4 June 2025).

Weng, P., Tian, Y., Liu, Y. and Zheng, Y. (2023) 'Time-series generative adversarial networks for flood forecasting', *Journal of Hydrology*, 622, p. 129702. Available at: https://doi.org/10.1016/j.jhydrol.2023.129702.

Xie, K., Liu, P., Zhang, J., Han, D., Wang, G. and Shen, C. (2021) 'Physics-guided deep learning for rainfall-runoff modeling by considering extreme events and monotonic relationships', *Journal of Hydrology*, 603, p. 127043. Available at: https://doi.org/10.1016/j.jhydrol.2021.127043.

Zhao, X., Wang, H., Bai, M., Xu, Y., Dong, S., Rao, H. and Ming, W. (2024) 'A Comprehensive Review of Methods for Hydrological Forecasting Based on Deep Learning', *Water*, 16(10), p. 1407. Available at: https://doi.org/10.3390/w16101407.

# Appendix A: Literature review

## A.1 Literature review strings

Table 3: Search strings used in WoS for each forecasting stage. The search was performed in WoS core database with no time constraints, i.e., spanning all years. Search date is date of results download for analysis. The first search is the main query, and the second search is the main query refined with the exclusion criteria. The subsequent searches were based on the combination of queries 1 and 2 with each stage string.

| # | String | Results | Search date | Forecasting stage |
|---|---|---|---|---|
| 1 | TI=((flood* OR inundation) AND (forecast* OR predict* OR "predictive model*" OR "time*series forecast*") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR neural network OR "artificial neural network" OR ANN OR "ensemble learning" OR hybrid model OR "data-driven" OR AI OR LSTM OR "long short-term memory" OR comput* intelligence OR soft comput* OR CNN OR "convolutional network" OR "convolutional model" OR recurrent)) OR AB=((flood* OR inundation) NEAR/15 ("forecast*" OR "predict*" OR "predictive model*" OR "time*series forecast*") AND ("artificial intelligence" OR "machine learning" OR "deep learning" OR neural network OR "artificial neural network" OR ANN OR "ensemble learning" OR hybrid model OR "data-driven model" OR AI OR LSTM OR "long short-term memory" OR comput* intelligence OR soft comput* OR CNN OR "convolutional network" OR "convolutional model")) NOT TI=(precipitation OR rainfall OR runoff OR sediment OR "flood mapping" OR "inundation mapping" OR (reservoir OR dam) *flow OR reservoir operation OR dam operation) NOT TS=(flood susceptibility OR hazard mapping OR vulnerability) | 1,995 | 22/05/2025 | Main |
| 2 | #1 NOT TS=(oil OR carbon dioxide OR nitrogen OR nitrification OR ammonium OR zooplankton OR distillation OR arsenic OR methane OR chlorophyll OR column OR spacecraft OR "flood* attack*" OR "computer security" OR "intrusion detection" OR pathology OR enzyme OR sparrow OR opioid OR ship OR spectroscopy OR equity OR abrasion OR "heat pipe" OR aerosol OR porous OR salt marsh OR rock) NOT TI=(CO2 OR earthworm OR amphibian OR mosquito OR wood OR eucalyptus OR oxygen OR fuel cell OR supply chain OR tool wear OR scour OR dike OR "lightning prediction" OR "track prediction" OR fish) | 1,845 | 22/05/2025 | Main with exclusions |
| 3 | #2 AND<br>(TI=(monitor* OR "real-time monitoring" OR "remote sensing" OR sensor OR WSN OR "wireless sensor network*" OR IoT OR internet of things OR "citizen science" OR crowdsourcing OR "social media" OR camera OR video OR CCTV OR drone OR UAV OR "data fusion" OR "flood detection" OR "community-based" OR "ungauged basin" OR "data-sparse" OR "underrepresented communit*" OR "participatory sensing" OR gauge OR Twitter OR radar)<br><br>AND AB=(monitor* OR "real-time monitoring" OR "remote sensing" OR sensor OR "WSN" OR "wireless sensor network*" OR IoT OR internet of things OR "citizen science" OR crowdsourcing OR "social media" OR camera OR video OR CCTV OR drone OR UAV OR "data fusion" OR "early warning" OR "flood detection" OR "community-based" OR "ungauged basin" OR "data-sparse" OR "underrepresented communit*" OR "participatory sensing" OR "gauge" OR "Twitter" OR radar)) | 125 | 23/05/2025 | Monitoring |
| 4 | #2 AND TI=(recalibrat* OR calibrat* OR "parameter optim*" OR "parameter tuning" OR "model optim*" OR "model updat*" OR "parameter estim*" OR "offline learning" OR "data-driven optim*" OR "efficiency improvement" OR "model reduction" OR "computational saving" OR "workflow autom*" OR "automatic model*" OR "error updat*" OR "error correction" OR "error analysis" OR "intelligent correction" OR "parameter prediction" OR "pre-training")<br><br>AND AB=(recalibrat* OR calibrat* OR "parameter optim*" OR "parameter tuning" OR "model optim*" OR "model updat*" OR "parameter estim*" OR "offline learning" OR "surrogate model" OR "hybrid model" OR "data-driven optim*" OR "efficiency improvement" OR "model reduction" OR "computational saving" OR "workflow autom*" OR "automatic model*" OR "manual calibration" OR "reduce effort" OR "numerical model improvement" OR "error updat*" OR "error correction" OR "error analysis" OR "intelligent correction" OR "parameter prediction" OR "pre-training") | 17 | 22/05/2025 | Model Calibration |

| Table 3: Search strings used in WoS for each forecasting stage. The search was performed in WoS core database with no time constraints, i.e., spanning all years. Search date is date of results download for analysis. The first search is the main query, and the second search is the main query refined with the exclusion criteria. The subsequent searches were based on the combination of queries 1 and 2 with each stage string. | | | | |
|---|---|---|---|---|
| # | String | Results | Search date | Forecasting stage |
| 5 | #2 AND TS=((precipitation forecast* OR rainfall forecast* OR "weather forecast*" OR "atmospheric forecast*" OR "meteorological forecast*" OR "meteorological model*" OR "precipitation model*" OR WRF OR "numerical weather prediction" OR NWP)<br>AND ("bias correction" OR "post processing" OR "post-processing" OR downscaling OR "ensemble post-processing" OR "statistical correction" OR "forecast fusion" OR "forecast integration" OR "input enhancement" OR "ensemble forecast*" OR coupl*)) | 90 | 23/05/2025 | Enhanced weather inputs |
| 6 | #2 AND (TS=((hybrid OR "hybrid model*" OR "AI-assisted" OR "hybrid hydrological model*" OR "physics-informed" OR "physics-guided" OR surrogate OR "emulate*" OR augment* OR complement* OR aid OR combin* OR "support tool*" OR "coupled model*" OR "coupled system*" OR "data assimilation" OR "digital twin" OR "proxy model*" OR "approximat* model*" OR "approximat*" OR ROM OR "reduced-order model" OR ARIMA OR SARIMA) AND ("hydrological model*" OR "hydraulic model*" OR "hydrodynamic model*" OR "hydrodynamic simulation*" OR "flood* model*" OR "numerical model*" OR "numerical simulation" OR "physical model*" OR process-based OR " conventional hydro* model* " OR "statistic* model*" OR wave model* OR "storm surge forecast*" OR "storm surge model*" OR "sea level rise" OR "tsunami forecast*" OR MIKE* OR SWAT OR LISFLOOD OR TELEMAC OR Delft3D OR "HEC-RAS" OR MODFLOW OR CAMA-FLOOD)) OR TI=(surrogate* OR "surrogate* model*" OR physics-informed OR physics-guided OR hybrid)) | 426 | 23/05/2025 | Complement forecasting |
| 7 | #2 AND (TI=(*LSTM* OR "long short-term memory" OR *GRU* OR RNN OR *CNN* OR convolutional OR SVM OR "support vector" OR MLP OR "multi-layer perceptron" OR *NARX* OR ANN OR "artificial neural network*" OR "neuro-fuzzy" OR "fuzzy logic" OR "backpropagation" OR "BP neural network*" OR "back-propagation" OR transformer OR "sequence-to-sequence" OR feedforward OR FFNN OR "genetic progrm*" OR "genetic algorithm" OR recurrent OR "deep learning" OR "extreme learning" OR "data-driven" OR "AI-based forecast*" OR "time-series" OR "time series" OR "end-to-end model" OR "real-time calibration" OR "online learning" OR "online training" OR "adaptive model" OR "model-free") OR AB=(LSTM* OR bi*LSTM OR "long short-term memory" OR GRU* OR RNN OR CNN* OR convolutional OR SVM OR "support vector" OR MLP OR "multi-layer perceptron" OR NARX* OR ANN OR "artificial neural network*" OR "neuro-fuzzy" OR "fuzzy logic" OR "backpropagation" OR "back propagation neural network" OR "back-propagation" OR transformer OR "sequence-to-sequence" OR feedforward OR FFNN OR "BP neural network*" OR "back-propagation" OR recurrent OR "genetic progrm*" OR "genetic algorithm" OR "deep learning" OR "extreme learning" OR "data-driven" OR "AI-based forecast*" OR "time-series" OR "time series" OR "end-to-end model" OR "real-time calibration" OR "online learning" OR "online training" OR "adaptive model" OR "model-free"))<br>NOT TI=("surrogate" OR ARIMA OR SARIMA OR "physics-informed" OR "hybrid hydro*" OR hybrid)<br>NOT AB=("surrogate" OR ARIMA OR SARIMA OR "physics-informed" OR "hybrid hydro*" OR "hydro* model*") | 1,093 | 23/05/2025 | Replace forecasting |
| 8 | #2 TI=(impact OR consequence OR damage OR disaster OR destruction OR "flood loss" OR "inundation impact" OR "impact forecast*" OR "impact assessment*" OR "real-time impact*" OR "decision support" OR "risk-informed decision*" OR "decision making" OR "decision-making" OR "emergency plan*" OR "impact-based forecast*" OR "situational awareness" OR "operational platform" OR (data mining AND (impact OR support)) OR "knowledge discovery" OR road OR electricity OR house OR building OR rail OR train OR infrastructure OR transport) | 116 | 22/05/2025 | Decision support |
| 9 | #2 AND TI=(warning OR "early warning system*" OR "early warning*" OR EWS OR "real-time alert" OR "real-time warning" OR alert OR "flood alert" OR "response system" OR "emergency response" OR "flood response" OR "risk warning" OR "automated decision" OR "automated response" OR "actionable intelligence" OR "trigger-based response" OR "smart warning" OR "intelligent warning" OR "response strategy" OR "hazard warning") | 55 | 22/05/2025 | Warnings and response |

# Appendix B: Expert workshop

## B.1 Workshop agenda and handout

**Workshop: Exploring the Use of Artificial Intelligence for Flood Forecasting in Scotland**

Date: Wednesday 23 April 2025

Workshop time: 10:00-15:15 BST

Location: Teams (online)

**Background, status and workshop objectives**

**Background**

Flooding is one of the most significant climate-related threats facing Scotland, with risks expected to intensify due to a warming climate (CCRA3, 2021). Traditional flood forecasting models, such as physics-based rainfall-runoff models (Environment Agency, 2001), have been widely used to predict flood events. However, these models often struggle to account for the complex hydrological processes involved in flood events and require extensive calibration to maintain precision.

The rapid advancement of Artificial Intelligence (AI) presents a promising data-driven alternative to conventional floods forecasting methods. AI has the potential to complement existing forecasting capabilities by leveraging large datasets to identify spatial patterns, expand predictive timescales (towards real-time predictions), and improve accuracy (e.g., Liu *et al.*, 2025). In particular, Machine Learning (ML) and deep learning approaches (e.g., Xie *et al.*, 2021) could significantly enhance SEPA's flood forecasting capabilities in Scotland by integrating diverse and novel hydrometeorological data sources, identifying nonlinear and multivariate relationships, and delivering predictive insights to enhance response optimisation.

AI presents a major opportunity to improve early warning systems, strengthen disaster preparedness, and support adaptive flood risk management (e.g., Ghaffarian *et al.*, 2023; United Nations UNU EHS, 2024; Oxford Insights, 2025). However, despite its potential, challenges remain in terms of model interpretability, data availability, and the seamless integration of AI-driven insights into existing forecasting frameworks.

**Current status of AI in flood forecasting**

AI is increasingly being integrated into flood forecasting, providing improved predictive capabilities and real-time monitoring. AI-driven approaches offer the ability to process vast amounts of hydrometeorological data, enhance forecasting accuracy, and contribute to more effective early warning systems. These advancements allow for a transition from hazard-based forecasting towards impact-based approaches, which provide better insights into how flooding events will affect communities and infrastructure (United Nations UNU EHS, 2024). While AI holds great promise in advancing early warning systems (EWS), several challenges must be addressed to ensure its effective and responsible use. Known challenges to achieving this include:

- Data availability and quality: AI models rely on high-quality, extensive datasets, yet many flood-prone regions lack sufficient historical and real-time hydrometeorological data (United Nations UNU EHS, 2024).
- Integration with existing forecasting systems: Traditional physics-based flood models remain widely used. AI must be effectively integrated with these models rather than replacing them, ensuring compatibility with current forecasting frameworks (Oxford Insights, 2025).
- Resource constraints: Implementing AI-driven flood forecasting systems requires significant computational resources, infrastructure, and technical expertise, which may not be available in all regions (Liu *et al.*, 2025).

## Workshop objectives

The Scottish Environment Protection Agency (SEPA) is the national authority for flood forecasting in Scotland, playing a critical role in mitigating flood risks, protecting communities and enhancing preparedness for flood events. SEPA seeks to explore how AI and ML can enhance flood forecasting in Scotland.

This workshop, led by Dr Chris White from the University of Strathclyde as part of the '*DelugeAI*: A review of the emerging opportunities of using artificial intelligence for flood forecasting in Scotland' (funded by CREW, Scotland's Centre of Expertise for Water; CSPF2025_01), will explore the key questions related to AI-driven flood forecasting as an alternative to traditional rainfall-runoff modelling identifying both opportunities and challenges in its adoption.

The workshop aims to:

- Demonstrate and discuss findings from AI-driven flood forecasting research, gathering expert insights on its relevance and practical applications.
- Identify current applications of AI in flood forecasting within existing forecasting frameworks.
- Explore the challenges and opportunities of AI-based flood forecasting, distinguishing realistic applications from theoretical possibilities.
- Provide recommendations for integration AI into Scotland's flood forecasting systems.

The workshop will include a presentation of current AI applications (both in flood forecasting and more general) and facilitate a broad discussion with participants from across disciplines. The goal is to identify where AI can be realistically integrated into flood forecasting in Scotland, distinguish feasible applications from more aspirational ideas, and establish a shared vision for AI's role in flood forecasting over the next 3-5 years.

## Expert workshop agenda

| Workshop: Wednesday 23 April 2025 | | |
|---|---|---|
| **Part 1: Setting the scene** | | |
| 10:00 – 10:20 | Welcome, introductions & workshop objectives | Chris White, University of Strathclyde and Michael Cranston, SEPA |
| 10:20 – 10:50 | Keynote: From Continental Models to Local Insights: Advancing Hydrological Prediction in Europe with AI | Ilias Pechlivanidis, SMHI (Sweden) |
| 10:50 – 11:00 | Break | |
| **Part 2: Understanding the use of AI for flood forecasting** | | |
| 11:00 – 11:30 | Expert insights on AI applications – Lightning talks: Michael Butts, Massimiliano Zappa and Jonathan Frame | Robert Atkinson, University of Strathclyde |
| 11:30 – 11:50 | Initial findings from a systematic review of flood forecasting and AI | Vicky Marti, University of Strathclyde |
| 11:50 – 12:30 | Open discussion: Expert insights on AI applications | Robert Atkinson, University of Strathclyde |
| 12:30 – 13:45 | Break | |
| **Part 3: Looking forward** | | |
| 13:45 – 14:15 | Expert insights on the future of AI for flood forecasting – Lightning talks: Maria Luisa Taccari, Steven Ramsdale and Jan Verkade | Doug Bertram, University of Strathclyde |
| 14:15 – 14:45 | Open discussion: Identifying realistic applications of AI in flood forecasting | Doug Bertram, University of Strathclyde |
| 14:45 – 15:00 | Next steps: Building a collaborative output | Chris White, University of Strathclyde |
| 15:00 – 15:15 | Closing remarks | Chris White, University of Strathclyde |
| 15:15 | Workshop close | |

### Invited speakers

This section highlights our invited speakers, their areas of expertise, and links to their professional profiles.

| Table 4: List of invited speakers at the expert workshop. | |
|---|---|
| **Ilias Pechlivanidis** | Lead Scientist at SMHI and Chair of the HEPEX initiative – Profile |
| **Michael Butts** | Chief Consultant for Flooding at DMI, Adjunct Professor of Water Resources Engineering (TVRL), Lund University – Profile |
| **Massimiliano Zappa** | Senior Scientist at the Swiss Federal Institute for Forest, Snow and Landscape Research (WSL) – Profile |
| **Jonathan Frame** | Assistant Professor of Machine Learning and Artificial Intelligence in Geological Sciences at the University of Alabam – Profile |
| **Maria Luisa Taccari** | Scientist for Data-driven Hydrological Forecasting at the European Centre for Medium-Range Weather Forecasts (ECMWF) – Profile |
| **Steven Ramsdale** | Chief meteorologist at the UK Met Office – Profile |
| **Jan Verkade** | Senior Hydrometeorologist at Deltares – Profile |

## B.2 List of expert workshop participants

Table 5: List of expert workshop participants and their organisation in alphabetical (surname) order.

| Name | Affiliation |
|---|---|
| **Phillip Aarestrup** | Technical University of Denmark |
| **Robert Atkinson** | University of Strathclyde |
| **Douglas Bertram** | University of Strathclyde |
| **Michael Brian Butts** | Danish Meteorological Institute |
| **Rebekah Burman** | CREW |
| **Annie Chang** | MeteoSwiss |
| **Deidre Cleland** | National Institute of Water and Atmospheric Research |
| **Steven Cole** | UK Centre for Ecology & Hydrology |
| **Amy Cooper** | CREW |
| **Michael Cranston** | Scottish Environment Protection Agency |
| **Richard Crowder** | Jacobs |
| **Robert Cowling** | Environment Agency |
| **Anthony Duke** | UK Met Office |
| **Matthew Fry** | UK Centre for Ecology & Hydrology |
| **Jonathan Frame** | University of Alabama |
| **Andrew How** | Natural Resources Wales |
| **Victoria Martí Barclay** | University of Strathclyde |
| **Simon Moulds** | University of Edinburgh |
| **Kamila Nieradzinska** | University of Strathclyde |
| **Adam Parkes** | Jacobs |
| **Ilias Pechlivanidis** | Swedish Meteorological and Hydrological Institute |
| **Charles Pilling** | UK Met Office |
| **Malcolm Price** | The Data Lab |
| **Marc Roper** | University of Strathclyde |
| **Jenny Roberts** | JBM Consulting |
| **Steven Ramsdale** | Met Office |
| **Neil Ryan** | Environment Agency |
| **Jonathan Sewell** | Scottish Government |
| **Rachel Skinner** | Natural Resources Wales |
| **Bryony Smith** | Jacobs |
| **Saleh Seyedzadeh** | The Data Lab |

| Table 5: List of expert workshop participants and their organisation in alphabetical (surname) order. | |
|---|---|
| **Name** | **Affiliation** |
| **Maria Luisa Taccari** | European Centre for Medium-Range Weather Forecasts |
| **Muhammad Usman** | University of Strathclyde |
| **Lucile Verrot** | Scottish Environment Protection Agency |
| **Christopher White** | University of Strathclyde |
| **Richard Weston** | Natural Resources Wales |
| **Jan Verkade** | Deltares |
| **Massimiliano Zappa** | Swiss Federal Institute for Forest, Snow & Landscape Research |

## B.3 Identified development areas, aligned to the road map horizons

During the expert workshop, several suggested areas or lines of development were identified across the one, three, and five-year development horizons, which are detailed here from Table 6 to Table 9.

| Table 6: One-year horizon. | |
|---|---|
| **Application/Suggestion/Comment** | **Relevant Phase Identified** |
| Improved flood model inputs (e.g., snow, rain, temperature, soil moisture) 1+ years. | Monitoring |
| Using LSTM to post-process (bias correction) the output from physics-based models. | Complement existing flood forecasting approaches |
| Formal model intercomparison and benchmark datasets to quantify potential benefits from AI. | Complement existing flood forecasting approaches |
| Use of AI / ML to help reduce model run times especially as we look towards more of a probabilistic approach which will increase run times and processing power requirements. Can also hopefully improve model outputs within the bounds of more frequent floods where we have data. | Complement existing flood forecasting approaches |
| Data assimilation, forecasting Potential to improve current flood model outputs using AI/ML assimilation and updating approaches. 1 year plus. | Complement existing flood forecasting approaches |
| Use AI for any kind of post-processing. | Complement existing flood forecasting approaches |
| Use of multi model ensemble to capture the skill of different ML algorithms, but the trade off is how many models should we run while keeping the system computationally reasonable. | Complement existing flood forecasting approaches |
| Can AI be used to reduce uncertainty in near time weather forecasting. | Complement existing flood forecasting approaches |
| Inclusion of LSTM rainfall runoff models alongside existing legacy models. | Complement existing flood forecasting approaches |
| Emulating physics-based models, which are too slow to run in real-time. | Replacing flood forecasting with AI/ML |
| Having decisions as to why you have or haven't issued a warning from a forecast being fed directly back into the system rather than through a human paper trail. Support the development of junior forecasters and comprehensive log of actions. | Decision Support |
| Scraping real-time data from social media, news outlets and incoming phone calls, summarised by AI and presented to operational staff during events to validate in real-time and inform. | Decision Support |
| Local decision-making. Can we learn when road barriers should be closed by learning from past decisions. | Decision Support |
| Using AI as direct decision support - e.g., Summarising forecasts, required procedures and/or asking approval from the human interaction to proceed to next stage (e.g., issue warning). | Decision Support |
| I see an increasing use of AI to flag areas that we need to prioritise on. Warnings from AI to drive more attention/activity. | Warnings and response in real-time: |
| Impact Forecasting Potential to use this quickly using existing model/mapping datasets. 1 year +. | Warnings and response in real-time: |
| Use AI to help draft our internal documents, e.g., project plans. | Other |
| Validation of flood maps using xAi to compare multiple sources e.g., EO-derived flood outlines or geo-referenced social media posts. | Other |

| Table 7: Three-year horizon. | |
|---|---|
| **Application/Suggestion/Comment** | **Relevant Phase Identified** |
| Physical processes. Could be used to model poorly modelled/understood processes (where appropriate data available). E.g., reservoirs. | Model calibration |
| Some of the modelling concepts (e.g., LSTM) should be near-ready for implementation. | Complement existing flood forecasting approaches |
| Feedback to prioritise model development. Identify issues in current process models to prioritise developments. 1 year plus. | Complement existing flood forecasting approaches |
| As confidence in AI solutions grow so will the confidence in the super complex processes of multi input challenges. | Complement existing flood forecasting approaches |
| Hybrid approaches (conceptual/physical/ML/AI) – particularly for small and ungauged catchments and river flow/level forecasts. 3 years? Benefits – improved forecasts, still some physical understanding (e.g., snow) Issues/Risks – ML only models for ungauged area or models encountering processes not in the "training" dataset. | Complement existing flood forecasting approaches |
| AI can surely efficiently support in impact-based decisions to prioritize action | Decision Support |

| Table 8: Five-year horizon. | |
|---|---|
| **Application/Suggestion/Comment** | **Relevant Phase Identified** |
| Real time inundation/impact modelling to inform forecasting/warning service. | Replacing flood forecasting with AI/ML |
| I'm not convinced we will replace physical models so much as shift our trust to AI and significantly reduce reliance on trad models. | Replacing flood forecasting with AI/ML |
| Towards foundation models that design/perform analyses that respond to natural language queries (e.g., will it flood in the next five days?). | Warnings and response in real-time: |
| Forecasting/warning service that uses impact/inundation mapping to deliver variable warning areas based on forecast AI/ML used to reduce model run times. | Warnings and response in real-time: |

| Table 9: Other aspects to consider. | |
|---|---|
| **Application/Suggestion/Comment** | **Relevant Phase Identified** |
| Please avoid any try to replace monitoring with AI. Just use AI for support in errors detection and gap filling. | Monitoring |
| Using AI to optimise data collection by prioritising locations of new hydrometric sensors; also digitisation of paper records to extend data record and therefore amount of training data. | Monitoring |
| Impact of data quality. Is this amplified by AI? Should some AI 'funding' be diverted to data verification? | Monitoring |
| Use of AI to process alternative monitoring data - e.g., satellite, drone, camera data to capture hydrometric information. | Monitoring |
| Explore levels of predictability at various scales to ensure valuable use of effort. | Complement existing flood forecasting approaches |
| Platforms for transparent development and assessment of AI (and conceptual) forecasting models. Opportunities – define objective (e.g., flood extent, flood levels), use emerging data/compute platforms (e.g., FDRI, https://fdri.org.uk/), reproducible, benchmarking standards/initiatives (e.g., flood hydrology roadmap). Issues – require compute and data platforms, with appropriate security (e.g., Trusted Research Environments), access to data (e.g., high res satellite). 3-5 years. | Complement existing flood forecasting approaches |
| Research AI decision support tools in other sectors - there will be several sectors who have this operational already which could apply to all the peripheral forecasting needs aside from actual FF and weather models. | Decision Support |
| Something very basic to consider before setting up an AI system: what infrastructure should/can an agency run its AI model operationally on, e.g., access to GPU? | Other |
| Need to continue traditional modelling/research/understanding to ensure accurate data exists to train future models. If we end solely AI we lose our understanding. | Other |
| Future workforce skills needed and how to upskill or complement existing staff. | Other |
| Base training in data science/AI. | Other |
| Are we killing the world trying to save it with AI? Is there any way of making AI for FF less impactful on the env? | Other |

# Appendix C: Feasibility study

## C.1 AI Solution definition and scope

Seven AI solutions (Table 10) were defined based on SEPA's flood forecasting framework, spanning the complete operational spectrum from initial monitoring through to warning response:

Table 10: Solutions identified for feasibility analysis for each stage of the flood forecasting process.

| Solution | Primary Function |
|---|---|
| Monitoring | Small or ungauged catchments – river flow/level forecasts where not captured elsewhere |
| Model Calibration | Maintaining and improving accuracy of existing SEPA model assets through automated recalibration |
| Weather Prediction | Improved flood model inputs – enhanced representation of rainfall, snow, and temperature forecasts |
| Forecasting (Complement) | Hybrid modelling – combining AI with physics-based models to support routine forecasting |
| Forecasting (Replace) | Full model replacement with ML |
| Decision Support | Reducing bias in decision-making for more consistent and informed future responses |
| Warnings and Response | Warning communication optimization |

## C.2 Feasibility study process

Step-by-step MCDA process implementation:

1. **Solution identification:** seven AI solutions were identified from SEPA's flood forecasting framework, covering the complete spectrum from monitoring to warning response.
2. **Stakeholder engagement:** key stakeholders were consulted to define evaluation criteria and assign category relative importance, i.e., weightings.
3. **Criteria development:** five main evaluation categories were established with detailed sub-criteria definitions.
4. **Weighting assignment:** percentage weights were allocated to each category based on strategic importance to SEPA's objectives.
5. **Scoring matrix development:** a standardised 1-5 scoring scale was developed with clear performance descriptors.
6. **Expert evaluation:** each AI solution was independently scored against all criteria by subject matter experts.
7. **Score validation:** scores were reviewed and validated through peer review and stakeholder feedback.
8. **Weighted calculation:** overall scores were calculated using the weighted sum method.

## C.3 Evaluation criteria development

Five primary evaluation categories were established through stakeholder consultation, with each category containing multiple sub-criteria to ensure comprehensive assessment. The criteria development process involved:

1. **Literature review:** best practices in AI implementation for flood forecasting were reviewed.
2. **Stakeholder workshops:** technical and management stakeholders identified key success factors.
3. **Criteria validation:** draft criteria were tested against known AI implementations.
4. **Final refinement:** criteria were refined based on applicability and measurability.

## C.4 Weight assignment process

Category weights were assigned through a structured stakeholder consultation process:

1. **Initial weight proposal:** draft weights were proposed based on SEPA's strategic priorities.
2. **Stakeholder review:** technical and management stakeholders reviewed and provided feedback on proposed weights.
3. **Consensus building:** final weights were agreed through facilitated discussion sessions.
4. **Validation:** weights were tested against hypothetical scenarios to ensure logical outcomes.

Following consensus building and validation, final weightings are indicated in Table 11 below.

**Table 11: Category weightings, sub-criteria, and rationale used in the MCDA process to evaluate AI solutions for flood forecasting. The table outlines the five evaluation categories, their respective percentage weights, detailed sub-criteria, and the rationale for each weighting based on strategic alignment with SEPA's objectives.**

| Category | Weight (%) | Sub-criteria | Weight rationale |
|---|---|---|---|
| Technical | 30 | Technology, accuracy, scalability, transparency, integration capability, real-time performance | Assigned highest weight (30%) due to critical importance of technical performance in flood forecasting accuracy and reliability |
| Improving Flood Resilience | 25 | Creating flood resilient communities, supporting good practice in flood-resilient placemaking, supporting multi-partner flood resilience delivery | Second highest weight (25%) reflecting core mission alignment with SEPA's flood resilience objectives and community protection goals |
| Cost/Resource | 20 | Infrastructure, development, staff training, resource efficiency, data availability, licensing & ongoing maintenance | Significant weight (20%) recognizing budget constraints and resource optimization requirements in public sector context |
| Deployment | 15 | Near-term wins build credibility for longer-term projects | Moderate weight (15%) emphasizing achievable implementation and building organizational confidence in AI solutions |
| Sustainability & Ethics | 10 | Environmental sustainability, energy efficiency, and alignment with Scottish Government AI policies | Lowest weight (10%) but essential for alignment with environmental stewardship and ethical AI deployment principles |

## C.5 Scoring scale development and application

The scoring process was conducted through the following systematic approach:

1. **Evaluator selection:** Subject matter experts were selected for each category based on relevant expertise.
2. **Independent evaluation:** Each evaluator independently scored all solutions against their assigned criteria.
3. **Calibration session:** Initial scores were reviewed in calibration sessions to ensure consistency.
4. **Score revision:** Scores were revised based on feedback from SEPA.
5. **Final validation:** Final scores were validated through peer review process.

## C.6 Mathematical calculation methodology

The overall scores were calculated using the weighted sum method, implemented through the following mathematical process:

$$\text{Overall score} = \sum \textit{(category score x category weight)}$$

Where:

- Category score = individual score for each of the five categories (1-5 scale), with 1 being very low feasibility and 5 being very high feasibility
- Category weight = percentage weight assigned to each category
- $\Sigma$ = summation across all five categories

Example calculation (warnings and response solution):

- Technical score: $2×0.30=0.60$
- Deployment score: $4×0.15=0.60$
- Flood resilience score: $4×0.25=1.00$
- Cost/Resource score: $4×0.20=0.80$
- Sustainability score: $4×0.10=0.40$
- Overall score: $0.60+0.60+1.00+0.80+0.40=3.40$

## C.7 MCDA results

The aggregated weighted scores revealed a clear hierarchy of implementation priorities, as shown in Figure 6. High-priority solutions (scoring above 3.0) emerged as those offering the optimal balance of technical capability, practical feasibility, and strategic value for SEPA's operational context.

## C.8 Results and categorisation

Based on these results, the seven AI solutions were categorized into distinct priority groups that reflect their readiness for implementation and potential impact. The analysis identified five high-priority solutions that scored above the 3.0 threshold, each offering unique advantages for immediate deployment:

- **Warnings and response (3.4)** emerged as the highest-scoring solution, achieving exceptional performance in deployment feasibility, cost-resource efficiency, and sustainability. While its technical performance was moderate, the solution's strength lay in practical implementation and operational efficiency.
- **Integrating weather predictions (3.3)** demonstrated strong technical capabilities and excellent flood resilience contributions through the adoption of improved ML inputs such as snow, rain and temperature data. The moderate deployment score reflected the complexity of integrating advanced weather prediction models, while reasonable cost considerations made this solution attractive for enhancing the foundational meteorological inputs essential for accurate flood forecasting.
- **Decision support (3.2)** excelled in cost-effectiveness and deployment feasibility, making it highly attractive for immediate implementation. This AI-driven system leveraged historical data to reduce decision-making bias, promising more consistent and informed future responses despite moderate technical performance scores.
- **Forecasting replacement (3.1)** represented the most ambitious technological advancement, demonstrating strong technical capabilities and good flood resilience potential through comprehensive replacement of physical models with ML. However, significant deployment challenges and cost concerns reflected the substantial commitment required for fundamental system transformation.
- **Model calibration (3.1)** focused on using ML to automatically recalibrate physically based flood forecasting models. Strong deployment characteristics were offset by technical limitations and cost concerns, positioning this as an incremental improvement rather than transformational change.

Two solutions were classified as medium priority, falling below the threshold:

- **Monitoring (2.95)** addressed the enhancement of monitoring for small or ungauged catchments, showing good flood resilience benefits but suffering from poor deployment scores and moderate technical performance. The solution required significant infrastructure development but provided essential data foundation for other AI applications.
- **Forecasting complement (2.85)** utilized a hybrid modelling approach with strong technical performance and good flood resilience potential but faced significant deployment challenges and cost considerations that limited immediate viability.

The MCDA analysis highlights a phased, strategic approach to AI adoption in flood forecasting, prioritising high-impact, easily deployable solutions that balance innovation with feasibility — laying the groundwork for sustainable, long-term transformation.

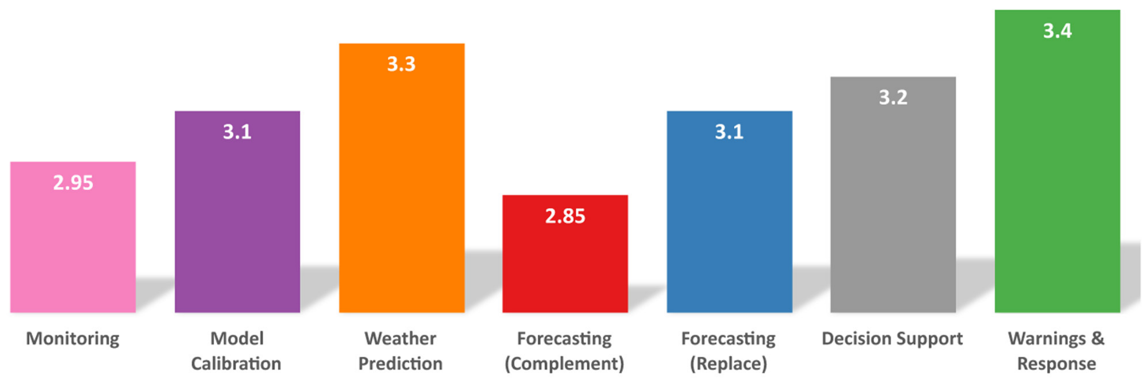



**Figure 6: Overall Multi-Criteria Decision Analysis (MCDA) scores for the evaluated solutions, highlighting performance across key categories. Warnings & Response achieved the highest score (3.4), while Forecasting (Complement) scored the lowest (2.85).**

CENTRE OF EXPERTISE FOR WATERS

## Centre of Expertise for Waters

**James Hutton Institute**
**Craigiebuckler**
**Aberdeen AB15 8QH**
**Scotland UK**

**www.crew.ac.uk**

CREW publications can be accessed here:
www.crew.ac.uk/publications

@crew_waters

CREW-scotland

Bluesky
@crew-waters

CREW is a partnership between the James Hutton Institute
and all Scottish Higher Education Institutes and Research Institutes.
The Centre is funded by the Scottish Government.

Scottish Government
Riaghaltas na h-Alba
gov.scot