Deltares

Can Precipitation-Related Incidents on the French Railway Network be Predicted in Advance?

Internship Project Report

Can Precipitation-Related Incidents on the French Railway Network be Predicted in Advance?
Internship Project Report

Author(s)
Peter Jansson

Can Precipitation-Related Incidents on the French Railway Network be Predicted in Advance?
Internship Project Report

Client	SNCF
Contact	Mark Cheetham
Reference	
Keywords	Railways, flooding, SNCF, France, Delft-FEWS, trains, earthworks, Roya, Villaine

Document control	
Version	0.1
Date	13-02-2024
Project nr.	-
Document ID	-
Pages	30
Classification	
Status	final

Author(s)		
	Peter Jansson	

The allowed use of this table is limited to check the correct order-performance by Deltares. Any other client-internal-use and any external distribution is not allowed.

Doc. version	**Author**	**Reviewer**	**Approver**
0.1	Peter Jansson	Jan Verkade	
		Jurian Beunk	

Summary

Rain and floods can cause large damages to railways, causing hazards to its trains and travelers. To improve flood mitigation, the French Railways (SNCF-Réseau) would like to use short-term forecasts to predict when a precipitation-induced incident would occur on the railway network. This study explored whether precipitation-induced incidents on the SNCF network between 2018-2022 can be predicted in advance. Precipitation accumulations from a gridded reanalysis dataset were matched with a log of incidents recorded by SNCF-Réseau. Local accumulation thresholds of a 10-year or 100-year event were tested to predict incidents. 24- and 72-hr precipitation accumulations were found to be weak – but not insignificant – predictors of the incidents provided (HSS<0.16). Filtering results based on antecedent moisture, using SPI-1, did little to improve results.

Two case studies were then conducted to explore the cause of incidents, given various preceding rainfall amounts. In the Villaine case study, incidents occurred despite normal precipitation accumulations and antecedent moisture (SPI<1). Photos of incidents indicate that they have likely been caused by short, localised precipitation events, some of which may not have registered in the precipitation data used. In the Roya case, all incidents were registered after one extreme precipitation event. This event was registered, but underestimated, by precipitation data. However, a hindcasting exercise - looking at past precipitation forecasts leading up to the event - showed that this event could have been forecasted with a lead time between 5 hours and 4 days, depending on the probability threshold used.

Results indicate that many of the incidents were triggered by very intense and localised precipitation events. These events were underestimated, or not captured, by the precipitation products we used. Further studies would also benefit from more contextual insights on past incidents, in order to better categorise each incident according to their causation pathway and hydro-meteorological predictors. To this end, we been recommend to explore predictors of past incidents and use high-resolution precipitation observations and nowcasts. While the former aspect relates to the entire SNCF network, analyses using high-resolution observations and forecasts should be tested to predicted incidents in small, steep catchments facing intense localised precipitation events. Also, recommendations have been made in further developing the risk management protocol to better gain insights on the hydro-meteorological context of each incident. Though results show a more nuanced approach is necessary to predict incidents across the SNCF network, recommendations have been made to set further research *on the right track*.

Contents

1 Introduction

Floods can wreak havoc on railways. As railways require complex and rigid infrastructure, their vulnerability to flooding is especially high (Doll et al., 2014; Thieken et al., 2016). Pluvial and fluvial flood events can, for instance, erode and inundate railway infrastructure (Ochsner et al., 2023). Another notable mechanism is scouring in rivers and streams, which can affect the structural integrity of both railway bridges and earthworks (Dikanski et al., 2018). Such incidents can lead to train services being delayed or cancelled, along with making costly maintenance and mitigation efforts a necessity (Cheetham et al., 2016). Demand for more resilient railways is on the rise. Train travel is being promoted for its relatively low greenhouse emissions (Bruckner et al., 2014) while their flood risks are escalating due to climate change (Bubeck et al., 2019). To cope, railway operators are increasingly required to better understand flood risk and provide mitigation measures to ensure improved resilience of their infrastructure (Ochsner et al., 2023).

The risk management process consists of the identification, assessment and treatment of risks (ISO 31000, see Hutchins, 2018). According to these principles, precipitation-induced flood risks are first identified and catalogued. Then, during risk assessment, stretches of railway are assessed for risk and vulnerability to floods using past data or future scenarios. Much of the literature is focused on this stage (Ochsner et al., 2023); studies often involve drawing up flood risk maps to identify vulnerable stretches of track where flood mitigation strategies could be directed (Bubeck et al., 2019; Petrova, 2020). However, few studies discuss flood risk management at an operational level, such as flood forecasting and monitoring (Ochsner et al., 2023). Potentially, flood forecasts can warn railway operators of possible flood events, giving operators time to carry out railway inspections, maintenance, or closures to mitigate damage and disruption (Kerin, 2020).

Risk is composed of hazard, exposure and vulnerability (UNISDR, 2016). Figure 1 illustrates that the presence of a hazard alone does not guarantee an incident. Railway infrastructure must also be exposed to the hazard and be vulnerable to it, resulting in damages. Incident occurrence is a result of a combination of these factors, occurring in a chain of events called a causation pathway. Understanding causal pathways and identifying relevant exposure and vulnerability variables is essential in linking precipitation with incidents.

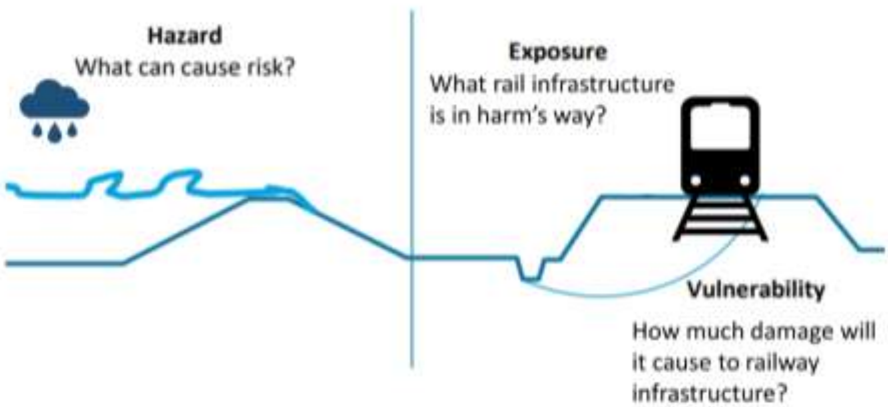

Figure 1 - Components of risk, as defined by UN International Strategy for Risk Reduction (UNISDR). Adapted from figure by Bles et al. (2023)*. Framework is analogous to RISK-VIP, as developed by* Cheetham et al. (2016)

The French national railways (Société Nationale des Chemins de fer Français – SNCF), through its SNCF-Réseau company, manages France's railway network. The network records multiple precipitation-induced incidents each year (SNCF Réseau, 2023, 2024). The network incurred flood damages averaging €100million/year between 1976-2005 (second in Europe to Germany), which are expected to rise to €400million/year between 2006-2100 (Bubeck et al., 2019). SNCF have implemented *Toutatis,* a system that monitors real-time precipitation observations from a nationwide radar network. Warnings are sent when high accumulations are observed. Preliminary results are encouraging – an event was confirmed on 2 of 3 instances where a warning was sent. SNCF-Réseau is considering the implementation of a system that can *forecast* precipitation-induced railway incidents. Such a system might improve their mitigation strategy through increasing the forecast reliability and extending the forecast lead time beyond that of *Toutatis* (5 minutes).

Delft-FEWS, developed by Deltares (Werner et al., 2013), has been considered as a possible system to forecast precipitation-induced railway incidents. Delft-FEWS, used operationally by forecasters worldwide, is able to run various hydrological models simultaneously using real-time data (Gijsbers et al., 2008). Despite its popularity at catchment, regional and national levels, integrating flood forecasting systems into transport infrastructure management is still in its infancy. Literature is limited to Kerin (2020), who incorporated Delft-FEWS into a management system of railway bridges in the Brandon catchment in Ireland. The system could forecast water levels and scour at the bridges, informing infrastructure managers on when to plan bridge maintenance, inspections, or even closures. Hypothetically, a similar system could be established for the SNCF railway network. However, due to the substantial investments involved in setting up such as system (Kerin, 2020), studying how past incidents on the railway network could be predicted in advance becomes are prudent step in justifying these investments.

To that end, this internship project explored the extent to which precipitation-induced incidents on the SNCF railway network can be predicted in advance. In the scope of this project, "precipitation-induced incidents" are ncidents triggered by precipitation or its induced runoff or streamflow, which can impact the structural integrity of the railway infrastructure. This study only considered past incidents impacting the railway earthworks that are hypothesised to have been caused by accumulated precipitation and resulting runoff processes in localised catchments (Cheetham, 2024).

2 Research Design

We first studied precipitation induced incidents across the entire SNCF network. When studying all incidents, we correlated incident occurrence with accumulated precipitation preceding the day of each event. Then, to study the relationship between precipitation and incidents on a local scale, two case studies were conducted. The datasets used in this study are summarised in table 4.

The study spanned the entire SNCF railway network, consisting of 35,115km of track across metropolitan France (figure 2). The incident log provided by SNCF includes 402 incidents that occurred between 2018 and 2022 and which were believed to be precipitation-induced. An incident is logged whenever abnormalities on the track are observed which may hinder the structural integrity of the track or underlying earthworks (table 1). These were categorized as either flooding, erosion or mudslide. This incident log was recorded by local railway workers, who record the location (line number and distance along track, i.e. chainage), time (day of recording) of each incident, as well as categorising the incident by type. Classification of flooding incidents is the responsibility of each local infrastructure maintenance team, meaning different types can be recorded for similar events, depending on experience. Also, when multiple incidents are observed close to each other (typically after a heavy storm) only some of the more serious incidents are logged. Due the sensitive nature of the log, only a selection of the incident database could be shared for this study.

Table 1 - Excerpt of incident log provided by SNCF. Incident location coordinates were obtained by georeferencing by line number and chainage.

Line	Chainage (m)	Date Obs	Type	X (WGS-84)	Y (WGS-84)
946000	81230	3-10-2020	EROSION	7.59030	44.10061
944000	6310	2-12-2019	MUDSLIDE	6.97147	43.56310
935000	852784	21-11-2018	FLOODING	5.15756	43.33850

2.1 All Incidents

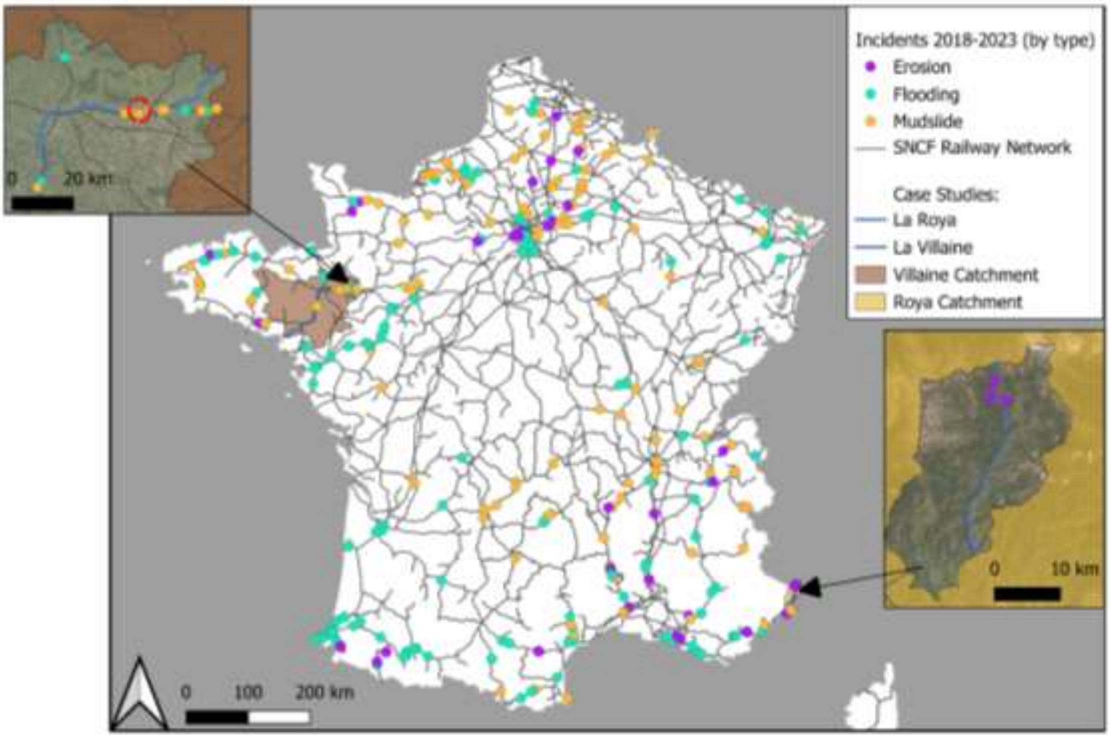

Figure 2 – Precipitation induced incidents logged between 2018-2022 on the SNCF railway network. The case study areas Villaine (circled) and Roya (entire catchment, in yellow) are highlighted.

Initially we ascertained the correlation between antecedent precipitation and incident occurrence. Daily accumulated precipitation observations were obtained from the EMO-1 reanalysis dataset (Gomes et al., 2020). This gridded dataset contains precipitation accumulations at a very high spatial resolution (1arcmin, roughly 1.8km, in both directions) necessary to capture extremely localised precipitation events (Thiemig et al., 2022). Such events could hypothetically lead to runoff processes on and near a certain stretch of track, leading to damages. Incidents were only logged by day of observation without noting the time at which an incident occurred. Hence, the shortest relevant time interval to accumulate precipitation is 24 hours preceding the incident date. The *pr24* precipitation variable was used, which comprises precipitation accumulated over 24 hours preceding 06:00am on the day after the incident. The precipitation grid was superimposed over the georeferenced incidents, yielding 314 grid cells (of 1x1 arcmin) where at least one incident was observed. Further analysis was limited to data from these 314 grid cells, which are termed 'incident locations'. This step saves on computation time, though ignores false alarms and misses that occur elsewhere in the railway network.

24- and 72-hour precipitation accumulations at each incident location were compared with those of local 10-year or 100-year events. These accumulation thresholds were available for all of France at 3km resolution from the SHYREG dataset (Arnaud et al., 2016). These amounts act as a threshold for precipitation severity, and hence used to predict incident occurrence. An incident is considered to have been predicted when if the preceding precipitation accumulation exceeded the threshold on a certain day. Differences between the precipitations during incidents and during all days can be compared statistically using the Kolmogorov-Smirnov (KS) 2-sample test (Kolmogorov, 1933; Smirnov, 1948). This test is widely used in comparing precipitation data, as comparisons are made on the empirical cumulative distribution functions without requiring the assumption of normality (Vlček & Huth, 2009).

As with the choice of precipitation data, precipitation accumulation thresholds of shorter durations were not used since the timing of the incidents are not known. For this analysis, observation data was preferred over forecast data, as the initial aim was to identify the statistical relationship between precipitation and incident occurrence, without having to consider forecast uncertainties.

The performances of local precipitation thresholds in predicting incidents were quantified in contingency tables (table 2). First, daily timeseries of precipitation accumulations spanning 2018-2022 were generated for each incident location. Each day was then categorised in a 2x2 contingency table based on whether an incident was reported, and whether the accumulated precipitation exceeded a local threshold (i.e. whether the incident was predicted). The resulting table consists of 4 categories, as seen in table 2. Various forecasting metrics were subsequently derived from each contingency table. The hit rate (equation 1) and false alarm rate (equation 2) respectively indicate the performance of the forecast in predicting the incidents and in generating false positives. These metrics are popularly used to communicate forecast performance (Wilks, 2011), as they capture the main requirements of a reliable forecast (capturing incidents while minimising false alarms) in an intuitive manner. An ideal forecast would maximise the hit rate (= 1) while minimising the false alarm rate (= 0).

$$Eq\ 1: Hit\ Rate = \frac{number\ of\ hits}{number\ of\ hits\ and\ misses}$$

$$Eq\ 2: False\ Alarm\ Rate = \frac{number\ of\ falses}{number\ of\ falses\ and\ quiets}$$

$$Eq\ 3: Hit\ Ratio = \frac{number\ of\ hits}{number\ of\ hits\ and\ falses}$$

$$Eq\ 4: Proportion\ Correct = \frac{number\ of\ hits\ and\ quiets}{sample\ size}$$

Equations 1-4: Performance metrics derived from the 2x2 contingency table.

Table 2 - Generalised contingency tables as used in this study.

Incident Reported? (Incident Observed?)	Threshold Exceeded? (Incident Predicted?)	
	True positive (Hit) ☺	False negative (Miss) ☹
	False positive (False) ☹	True Negative (Quiet) ☺

As the resulting contingency tables revealed predictions with low hit and false alarm rates, it became important to determine whether the predictor outperformed one that predicts incidents at random. The Heidke Skill Score (HSS) was used to compare the proportion correct (equation 4) of the prediction with that of a random (pure noise) predictor (equation 5). A predictor scoring HSS=1 is a perfect predictor, while one scoring HSS<0 underperforms a random forecast.

$$Eq\ 5: HSS = \frac{2(\#hits \times \#quiets\ -\ \#misses \times \#falses)}{(\#hits + \#misses)(\#misses + \#quiets) + (\#hits + \#falses)(\#falses + \#quiets)}$$

Equation 5 - The Heidke Skill Score

The occurrence of precipitation-induced incidents could also be linked to the antecedent wetness conditions in local area. To test this hypothesis, the standardised precipitation index (SPI-1) was calculated for each location, which also acts as an indicator for flood risk (Guerreiro et al., 2007). SPI-1 compares the precipitation accumulated in the preceding 30 days with the local 40-year climatology (1981-2020) (Edwards & McKee, 1997; McKee et al., 1993). SPI-1 hence indicates the relative antecedent precipitation levels in a normalised scale. In this study, SPI-1 was calculated using ERA-5 precipitation data (at 0.25^{O} resolution) (Hersbach et al., 2023). We distinguish days with wet conditions (SPI >1, see table 3), hypothesising that areas in wet conditions are more prone to flooding (Guerreiro et al., 2007; Seiler et al., 2002).

Table 3 - Definitions of wetness used in this study according to SPI-1 (SPI). Wet conditions are indicated by SPI>1.

Conditions	**Extremely Dry**	**Very Dry**	**Drier**	**Normal**	**Wetter**	**Very Wet**	**Extremely Wet**
SPI-1	**SPI < -2**	**-2< SPI ≤ -1.5**	**-1.5 <SPI ≤ 1**	**-1<SPI≤ 1**	**1<SPI≤1.5**	**1.5<SPI≤ 2**	**SPI> 2**

2.2 Case Study - Villaine

Figure 3 - Arial 3D impression of the stretch of line 420000, encompassing the Villaine study area. The river Villaine flows alongside the railway north-westwards (blue arrow), while the catchment area upstream of the tracks is highlighted in orange. Source: Google Maps (48.1049, -1.3417).

The choice of a case study location is somewhat arbitrary, but what sets the incident location in la Villaine (line 420000, chainage 342.430 to 342.790) apart from other locations was that more incidents were registered here (6) than in any other location in the network. The Villaine is also an SNCF priority catchment concerning fluvial floods and surface water runoff (Cheetham, 2024). The incident location, as well as the catchment, lies in the Western region of Brittainy, whose oceanic climate (Köppen-Geiger (Beck et al., 2018): Cfb) provides constant rainfall throughout the year. Daily precipitation accumulations were plotted for 2018-2022, allowing precipitation dynamics leading up to each incident date to be easily visualised. Similar visualisations can be made using SPI, as well as discharge data obtained from the measurement gauge of the Villaine at Chateaubourg. In addition, qualitative insights about the incidents were obtained from photos and excerpts from incident logs, which were communicated by Mark Cheetham (Cheetham, 2024).

2.3 Case Study – Roya

Figure 4 - Train passing through line 946000, also known as Train des Merveilles, through the Roya valley. Source: lifegate.it

The second case study looks into the incidents reported in the catchment La Roya following storm Alex in October 2nd 2020 (figure 4). The catchment lies in a valley in the region PACA (Provence Alpes Côte d'Azur), with a climate ranging between a drier Mediterranean climate in the main valley (Köppen-Geiger: Csb) and a subarctic climate (Köppen-Geiger: Dfc) at higher altitudes (Beck et al., 2018). Precipitation is punctuated by powerful convective systems with intense, localised precipitation, particularly during autumn. Following storm Alex, one such system, extensive damages were recorded on multiple sections of track, earthworks and bridges. Being able to predict and monitor precipitation events that lead to such extensive damages would be of great interest for multiple parties in SNCF-Réseau responsible for the different railway assets. As no other incidents in the catchment were recorded, we inspected the locations of the incidents, and using a 5m DEM, qualitatively explored the relevant hydrological processes.

Since it was known that this precipitation event was the cause of the railway incidents, a hindcasting exercise was also carried out. 15-day deterministic forecasts with 6-hourly precipitation accumulations generated by ECMWF were retrieved from their TIGGE repository. The Roya catchment spans neatly within 1 grid-cell at 0.5^{O} resolution (43.75-44.25^{O}N and 7.25-7.75^{O}E). Forecasts released at 00:00CET in the 15 days leading up to the event were analysed.

Hindcasting was also carried out with probabilistic forecasts. Uncertainties are inherent in meteorological and hydrological forecasting. These are not explicitly communicated by a deterministic forecast that represents the 'best guess' of the forecast model's conditions, states, structure and forcing values (e.g. Verkade & Werner, 2011). Hence, 50-member ensemble forecasts were also included to explicitly represent forecast uncertainties. It is important to note that that this case study represents an exercise on how an incident could have been forecasted. The exercise does not infer forecast reliability, especially since it involved forecasting a known incident, ruling out any chance for a false positive.

Table 4 - Overview of datasets used in this study.

Dataset	Source	Variable	Resolution	Type
EMO-1	(Gomes et al., 2020)	Accumulated precipitation (*pr24*) (mm/d)	1arcmin (1.8x1.8km); daily	Reanalysis
ERA-5 (for SPI-1)	(Hersbach et al., 2023)	Accumulated precipitation (mm/d)	0.25deg (21x21km); daily	Reanalysis
ECMWF	(ECMWF, 2024)	Accumulated precipitation (mm/6h)	0.5deg (42x42km); 6-hourly	Forecast (control and perturbed)
SHYREG	(Arnaud et al., 2016)	Accumulated precipitation (mm/d & mm/3d)	3kmx3km	Return-period precipitation thresholds
Villaine Discharge Gauge	(EauFrance, 2024) Gauge ID: J706 0620	Accumulated precipitation (mm/hr)	Point data; daily	Observations
Roya Precipitation Gauge	(EauFrance, 2024) Gauge ID: 0616 3007	Daily mean discharge (l/s)	Point data; hourly	

3 Results

3.1 All Incidents

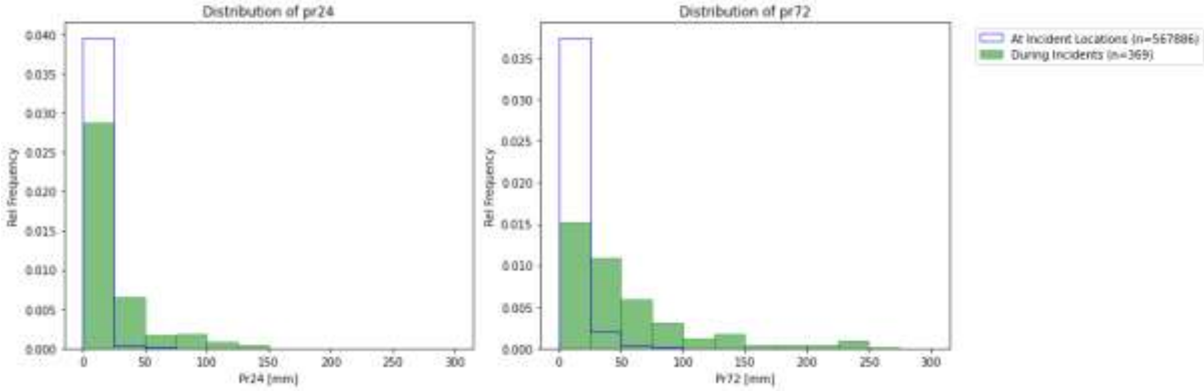

Figure 5 - Histograms of 24- and 72- hour precipitation accumulations. The plots compare accumulations preceding days with incidents with accumulations on all days.

At each incident location, the distribution of 24- and 72-hr precipitation accumulations (pr24 & pr72) preceding an incident were compared with distributions over the observation period (2018-2022). Figure 5 indicates that 72.5% of days with incidents are preceded by low (<25mm) 24-hour precipitation accumulations. This is substantially less than the proportion of low accumulations over the entire observation period (97.5%). This difference is starker when considering 72-hr precipitation accumulations, where only 37.5% of days with incidents are proceeded by <25mm accumulations, compared to 95% of all days. The skewedness in both plots indicate that days with incidents had higher antecedent precipitation accumulations compared to other days. Using the Kolmogorov-Smirnov 2-sample (KS) test, the distributions of pr24 and pr72 at incident locations were found to be statistically different from the distributions on all days ($p<0.01$).

However, higher precipitation accumulations only precede a minority of incidents. For instance, less than 10% of days with incidents registered more than 100mm in the preceding 72 hours. Though no single amount of precipitation accumulation can reliably predict incidents across all incident locations, there is a statistically significant relationship between precipitation at daily to monthly timescales (pr24, pr72, SPI-1) and incident occurrence.

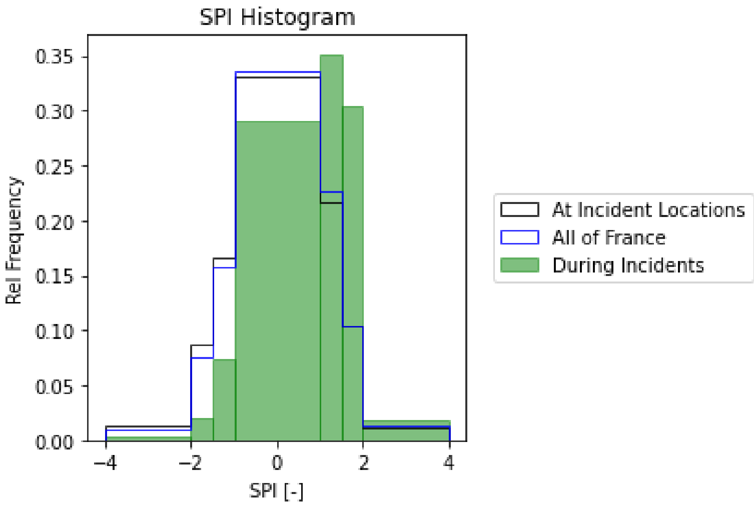

Figure 6 - Histograms of SPI-1 in all of France, at incident locations, and at incident locations during incidents.

The distribution of SPI-1 indices (figure 6) also reveals a tendency for incidents to occur in wetter conditions. Though all of France experienced on average slightly wetter weather than its climatology, this skewedness is more pronounced at incident locations, especially so on days with incidents. 37% of all days with incidents occurred during a wet period (SPI-1>1). Meanwhile, roughly only 20% of days at incident locations or in all of France experienced such wet conditions. However, t-tests run on the three distributions fail to reject the null hypothesis (H_0: SPI-1 ≤ 0), implying that the higher SPI-1 distribution during incidents does not exceed climatology with statistical significance ($p = 0.21$).

Contingency tables in figure 7 reveal that the local 10-year precipitation accumulations thresholds are poor predictors of incident occurrence. The best predictor of all incidents, pr72, only captures 14 of the 402 incidents, while generating 182 false positives. Pr24 captures fewer incidents (9), but also generates fewer false alarms (87). Isolating only instances when incident locations are in wet conditions leads to slight increases in hit rates, though coupled with increases in false alarms. The highest hit rate (0.125) is obtained when using pr72 to predict incidents during wet conditions, though this predictor also generates the highest false alarm rate (0.00065). Adjusting the threshold upwards (100-year events) or downwards (i.e. superimposing 10-year pr24 threshold on pr72 data) also lead to hit and false alarm rates shifting together, though never at a point where the hit rate reaches a practically useful level. For instance, the 10-year threshold on pr72 captures 12.5% of all incidents. However, this predictor generated over 4.14 times more false alarms than hits.

Predictor: $pr24_{obs} > pr24_{10yr}$

All incidents	9 Hits	393 Misses	Hit rate = 0.022 False rate = 0.00015 HSS = 0.03588
	87 Falses	567430 Quiets	
'Wet' incidents (SPI>1)	3 Hits	109 Misses	Hit rate = 0.027 False rate = 0.00037 HSS = -0.00078
	28 Falses	89023 Quiets	
'Not Wet' incidents (SPI<1)	6 Hits	284 Misses	Hit rate = 0.021 False rate = 0.00012 HSS = 0.03733
	59 Falses	478407 Quiets	

Predictor: $pr72_{obs} > pr72_{10yr}$

All incidents	32 Hits	370 Misses	Hit rate = 0.080 False rate = 0.00032 HSS = 0.10346
	182 Falses	567335 Quiets	
'Wet' incidents (SPI>1)	14 Hits	98 Misses	Hit rate = 0.125 False rate = 0.00065 HSS = 0.15134
	58 Falses	88993 Quiets	
'Not Wet' incidents (SPI<1)	18 Hits	272 Misses	Hit rate = 0.062 False rate = 0.00026 HSS = 0.08297
	124 Falses	478342 Quiets	

Figure 7 - Contingency tables generated when using local 10-year precipitation accumulation thresholds to predict incidents.

In terms of HSS, most predictors outperform one that guesses incidents by chance. 5 of the 6 predictors scored above 0, only the 24-hour threshold underperforms a forecast made by pure chance when predicting 'wet' incidents. Though the predictors are far from exhibiting a perfect forecast (HSS=1) their performances indicate some predictability. In terms of the skill scores, 72-hourly precipitation accumulations are relatively stronger indicators. Additionally, incidents occurring in 'wet' conditions are relatively more predictable.

3.2 Case study – La Villaine

The 6 incidents covered in this case study are not consistently linked to higher levels of precipitation, SPI-1 nor discharge (figure 8). The 6 incidents occurred on 5 days, spread between 2018 and 2021. Two incidents in June 2018 occurred during high levels of precipitation accumulation at both 24- and 72-hour timescales. One of the incidents followed 72 hours of precipitation that exceeded a 10-year event. This incident was the only one from this case study that was registered as a 'hit' in figure 2. Meanwhile, another peak event that occurred in 2021 did not coincide with an incident and was registered as a 'false'. All other incidents occurred during typical precipitation accumulations at the timescales studied, while a day with two incidents recorded no precipitation in the preceding 72 hours.

No incidents in the case study area occurred during 'wet' conditions. 5 of the 6 incidents occurred during normal wetness conditions (-1 < SPI-1 < 1), while another occurred during dry conditions (SPI < -1) combined with typical precipitation accumulations. Since none of the incidents in this case were caused by fluvial flooding, relationships between incident occurrence and high discharge could be attributed to a high moisture storage in the catchment (Kirchner, 2009). This state would likely coincide with peak discharge events, which in this case would include discharges above 5000l/s. 3 of the 6 incidents occurred during peak discharges. These results, combined with the patterns seen in SPI, imply that indicators for catchment moisture are also poor indicators of incident occurrence.

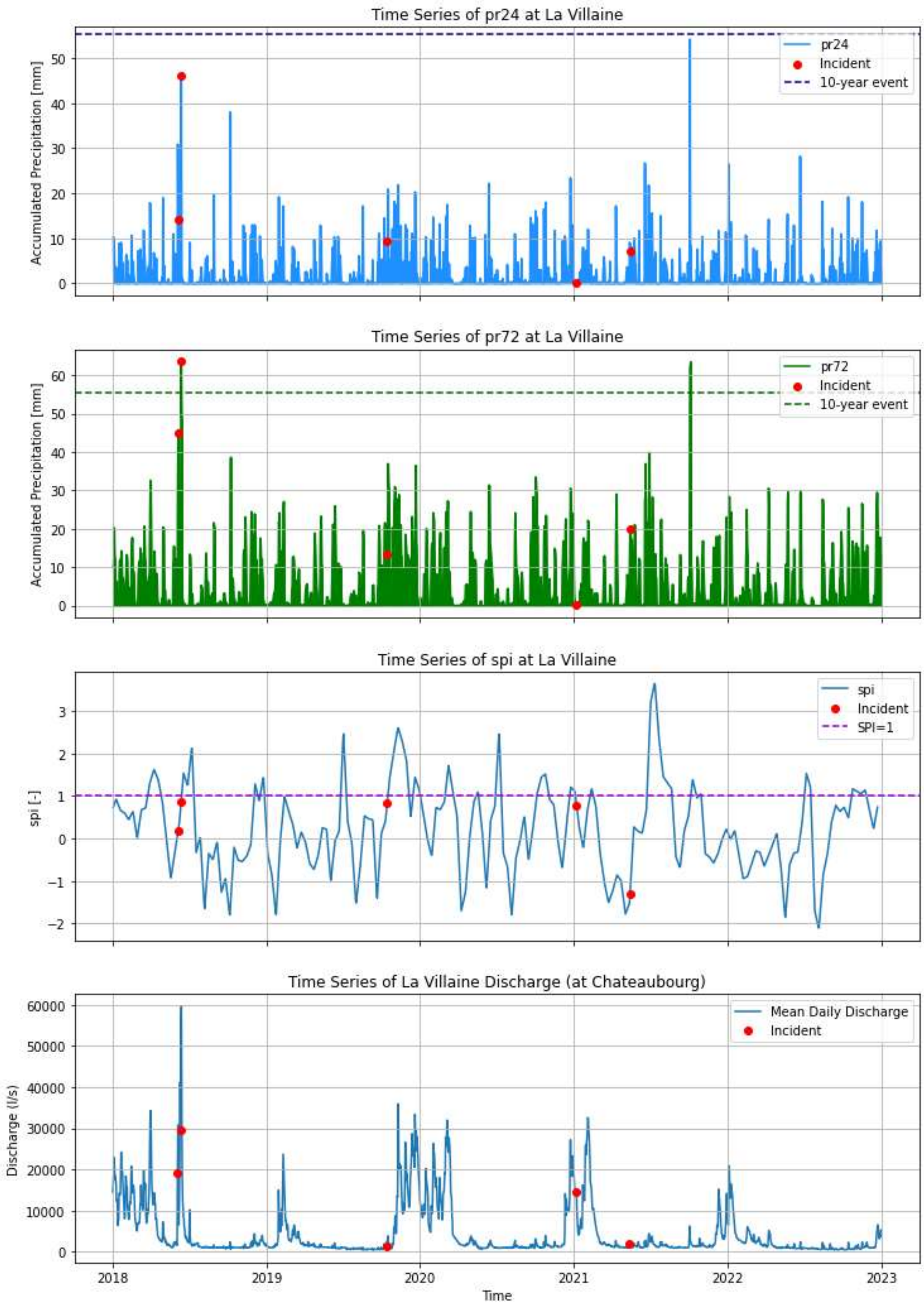

Figure 8 - Time series of 24-hour and 72-hour precipitation accumulations at the case study area in La Villaine, as well as of SPI-1 and discharge of La Villaine immediately downstream of the case study area.

The incidents occurred on a raised section of track on an embankment. Photos were requested of the incidents in the study area. These were not found, but we received photos of incidents in sections of track nearby (figure 9). Both showed evidence of small-scale rainfall-runoff processes. The left panel shows the railway located below a road; the erosion marks suggest runoff from an intense rainfall event moving towards the track. The right panel also appears to be the result of a precipitation-triggered runoff process. Here, runoff and sediment accumulated next to the railway embankment. The flood marks circled indicate that the water level was higher than the sediment deposited but remained below the tracks. Additionally, the arial photo in figure 3 shows that the area upstream of the railway consists of a plateau whose slopes converge towards the incident location, strengthening the case of surface water runoff being the culprit for these incidents. Despite the lack of correlations between hydro-meteorological variables and incidents, it remains likely that these incidents were caused by processes related to these variables.

Figure 9 - Photos of incidents on a stretch of track 42000 along the Villaine, near the study area.

3.3 Case study – Roya

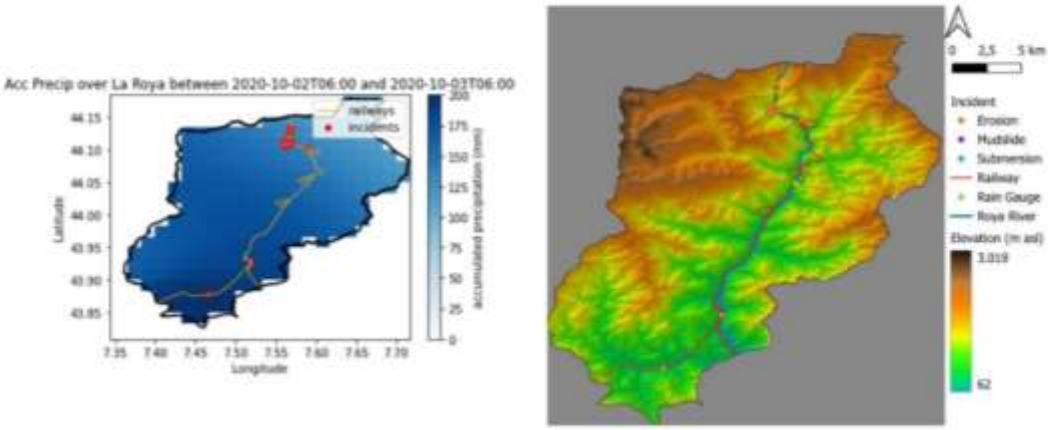

Figure 10 - Maps of the Roya catchment elevation (right) and precipitation accumulated during Storm Alex (EMO data) (right).

According to EMO data, the Roya catchment received on average 160.1mm in 24 hours leading up to 06:00 on 03/10/2020 (figure 10). This amount exceeds any other precipitation event in 2018-2022, though is shy of a 10-year event according to SHYREG data. The resulting runoff streamflow led to extensive damages along line 946000, classed in the incident log as landslides and erosion. The initial incident log received revealed 6 incidents in the area, though upon request for more details, 17 additional incidents were provided.

Figure 11 - Images of damages to line 968000 in the Roya catchment due to storm Alex.

21 of the 23 incidents were recorded in the most upstream section of track (figure 10), indicating a correlation between smaller catchment sizes (in this case 5 - 37km^2) and incident occurrence. Images (figure 11) and insights from experts (Cheetham, 2024) indicate that the incidents were generally triggered by higher flow velocities in the Roya weakening the railway embankments. The main 'mechanism' causing the incidents were hence the higher flow velocities in the Roya itself, as opposed to runoff coming immediately uphill from the damaged tracks. However, as no discharge gauge data was available, thresholds in flow dynamics leading to the incidents could not be obtained.

Focus instead shifted to a hindcasting exercise. As we know that these incidents were caused by extreme precipitation amounts over 24 hours, how far in advance was this forecasted using ECMWF forecasts? A threshold of 120mm/day was set as a predictor of incident occurrence, which just exceeds the second-highest precipitation accumulation in the study period (figure 12). The ECMWF control forecasts of the event were plotted in the 8 days leading up to it, to ascertain the possible forecast lead time (figures 13 and 14).

The deterministic (control) forecast underforcasted the (EMO) observed precipitation, though the 120mm/day threshold was exceeded on the 2/10/2020 00:00 forecast (figure 10). Comparison with the hourly gauged precipitation data implies a forecast lead time of 5 hours before the start of the precipitation event, and 10 hours before the peak hourly rainfall was reached. Meanwhile, the probabilistic forecasts paint a more optimistic picture of forecast lead time. 1 out of 50 ensemble members forecasted a threshold exceedance 3-5 days preceding the event. This amount increases to 9 members two days before the event, and 25 members on 2/10/2020. As a final note, there appears to be large discrepancies between data sources in the precipitation observed. The rain gauge recorded nearly double (312.1mm/d) as much precipitation as the EMO data (160.1mm/d), while ERA5 recorded close to 70mm/d.

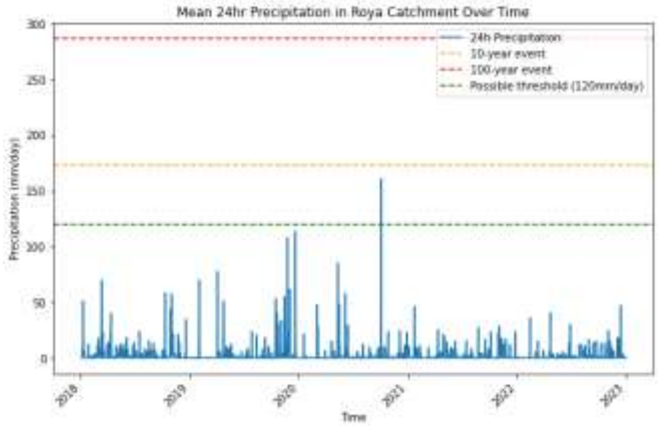

Figure 12 - Time series of 24-hourly accumulated precipitation in the Roya, with various thresholds.

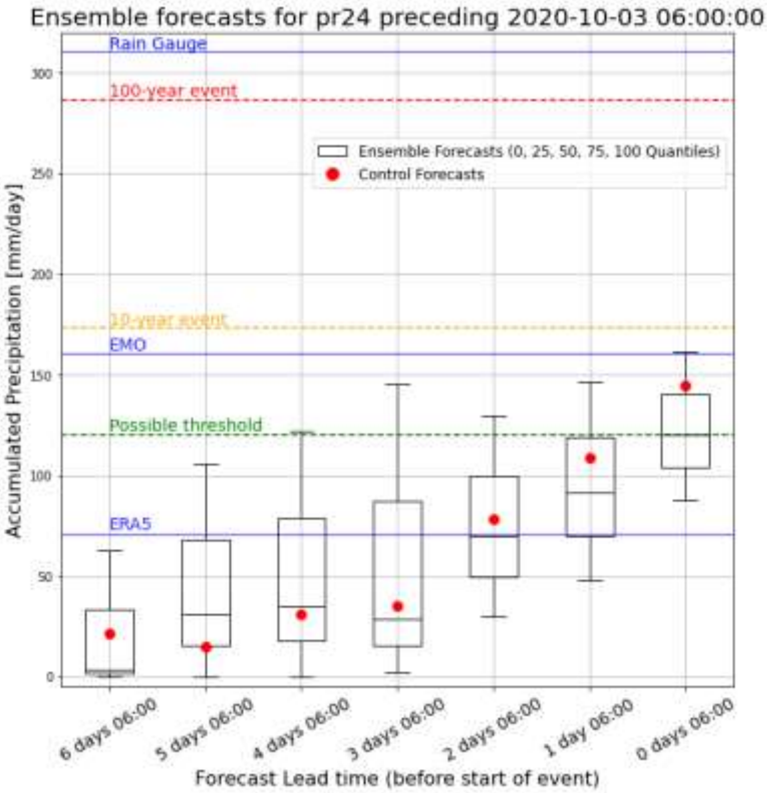

Figure 13 - Results from the hindcasting exercise, showing the (deterministically) forecasted precipitation accumulations in red, and the probabilistic forecasts as boxplots. Observed precipitation in grey.

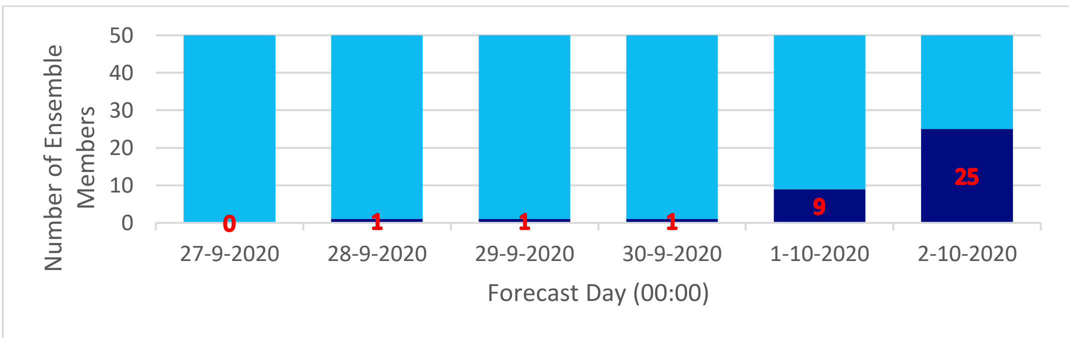

Figure 14 - Results from the probabilistic/ensemble forecasts. Number of ensemble members exceeding 120mm/day threshold in dark blue, labelled in red.

4 Discussion

The results indicate that 24–72-hour precipitation accumulations could not be used to predict incident occurrences with the necessary reliability. The study on all incidents showed that the majority of precipitation-induced incidents were preceded by ‘typical’ precipitation accumulations. Meanwhile, issuing alerts when precipitation accumulations are forecast to exceed a 10-year (or any less extreme) event would lead to a disproportionate amount of false positives.

However, results also point that forecasting railway incidents are far from a lost cause. The majority of predictors (10-year precipitation events with varying accumulation periods and SPI) outperformed a predictor based on pure chance, implying an (albeit weak) correlation between precipitation and incident occurrence. Also, the histograms of SPI and precipitation accumulations show that incidents tend to be preceded by wetter-than-average conditions, some at statistically significant levels. Lastly, the Roya hindcasting case indicated that the coarse ECMWF forecasts could already have forecasted the precipitation event up to 4 days in advance. As ECMWF also issues midday forecasts, the lead times of both probabilistic and deterministic forecasts could have been extended by another 12 hours.

Comparisons can be made with preliminary findings from SNCF’s *Toutatis* project (Cheetham, 2024). There, railway stretches in small catchments are monitored using real-time precipitation radar data at high spatial and temporal resolution. Thresholds were set based on local 5-minute precipitation accumulations at certain return periods. Track inspections were made after precipitation events exceeding thresholds, which were adjusted iteratively based on the extent of damages observed. So far, damages have been observed in roughly 2/3 of the time a warning was issued. This hit ratio (equation 3) is higher than any found in this study and indicates that incidents have a stronger correlation with precipitation events at shorter timescales than 24 hours. However, one should consider the risk of hit ratio being inflated by confirmation bias in that study. As maintenance workers were sent to inspect the track expecting to see damages, they may be more likely to observe and categorise abnormalities as incidents.

During the case studies, we investigated specific incidents in more detail. Judging from photographic evidence, incidents tended to occur during localised runoff events caused by intense precipitation in small and steep catchments. However, precipitation data either underestimated or failed to pick up many of these events. Also, the Villaine case showed that not all extreme precipitation and peak discharge events lead to incidents. The next challenge would hence be to better define the causation pathways between the incidents and their physical drivers.

Major limitations of these results could be linked to the incident data available. Due to legal limitations in this project, we received an incomplete list of incidents. The log included no information about the causes and severity of each incident. Incidents were only logged by their date of observation, which could be a day after the incident occurred. This limited us from exploring precipitation events at a finer timescale (i.e. hourly) where the *Toutatis* study indicated higher correlations with incident occurrence. Indeed, an hour of intense precipitation may be enough to trigger an incident but would not register as an abnormally high 24-hour precipitation accumulation. That being said, the EMO dataset accumulates 24-hour precipitations every 6 hours, making it possible to investigate 6-hour accumulations. Given the same data, further research could correlate incident occurrence with the maximum 6-hour accumulation that occurred at the day and location of the incident. We would expect this predictor to generate more hits and fewer false alarms, as intense localised precipitation events are more likely to be captured at this finer temporal scale.

Questions could also be raised on whether the incidents provided were all of a hydro-meteorological cause. The categorisation of incidents may be somewhat subjective and are made based on the observed damages to the track infrastructure without inferring about its causes. For instance, if an unsafe amount of water was observed on the tracks, the incident would be categorised as a flood, without considering whether the source of water was from rain, river flooding, or from a burst water pipe. Such subjectivities in incident categorisation may lead to the number of misses being exaggerated, while other precipitation-induced incidents may not have been included in the incident log we received.

Uncertainties were also evident in the precipitation data we received. Results from the Roya hindcasting exercise show large disagreements between different data sources. Differences are inherently linked to differences in how each dataset was generated (think point measurements versus interpolated grids). The EMO dataset tends to underestimate short and intense precipitation events (Thiemig et al., 2022). As results hint that such precipitation events are the ones most likely to trigger incidents, future studies may benefit from using higher resolution observations and forecasts. Opting for high-resolution data often entails the trade-off of shorter lead times, though this issue can be negated using blended forecasts (Imhoff et al., 2022). These forecasts combine numerical weather predictions with radar forecasts. If similar precipitation data were to be used, the methodology could be refined by incorporating return-period estimates generated from precipitation observations from the same dataset. This method would ensure stronger comparability between precipitation data and the return-period thresholds.

Further research could explore nowcasting, defined as forecasts for the next few hours, (World Meteorological Organization, 2017)) One example is operational nowcasts provided by MeteoFrance (Bouttier & Marchal, 2024). This dataset integrates radar and satellite data from metropolitan France as inputs for using the AROME model (lead time ≤ 6hours, with 15min accumulations) . This dataset benefits from having a higher spatial resolution compared to the EMO data used in this study (1.3km compared to 1.8km). The AROME model has been developed to better capture intense localised convective precipitation events, such as Storm Alex in the Roya case (Brousseau et al., 2016). These two characteristics indicate that the AROME nowcasts may perform better than the EMO data. However, preliminary findings indicate that AROME can reliably capture localised precipitation with return-periods of up to only several years (Bouttier & Marchal, 2024). Future research could utilise this (recently published) dataset to better understand the precipitation dynamics preceding the incidents used in this study.

Future work should also focus on categorising the main types of precipitation-induced incidents that occur in the SNCF network. It is known that, even for floods, the strongest hydrological indicator for incident occurrence can depend on the catchment or railway infrastructure type (Kellermann et al., 2016). Referring to the risk framework from figure 1, this study explored precipitation as the main hazard, and touched upon SPI and catchment size as forms of exposure. We argue that intense rainfall could be the 'main' driver of incident occurrence, whereas the SPI and multi-day precipitation accumulations inform us about the state of the hydrological system and its exposure to incident occurrence. Future studies could explore other forms of exposure (catchment steepness and flow velocities) and vulnerabilities (e.g. time since last inspection) linked to railway incidents (Kellermann et al., 2015).

5 Conclusion

This study explored whether precipitation-induced incidents on the SNCF network can be predicted in advance. , 24- and 72-hr accumulated precipitation was found to be a weak – but not insignificant – predictor for the incidents provided. Filtering results based on antecedent moisture, using SPI-1, did little to improve results. Two case studies were then made to explore the cause of incidents, given various preceding rainfall amounts. In the Villaine case, incidents occurred despite normal precipitation accumulations and antecedent moisture (SPI<1). Photos of incidents indicate that they could have been caused by short, localised precipitation events, which may not have registered in the precipitation data used. In the Roya case, all incidents were registered after one extreme precipitation event. This event was registered (but underestimated) by EMO precipitation data, while a hindcasting exercise showed that the event could have been forecasted with a lead time between 5 hours and 4 days, depending on the probability threshold used.

Major limitations of the study were linked to the spatial resolution of precipitation data. The relatively coarse data has limited ability to capture intense, localised precipitation events. Such events appear to be an important predictor of precipitation-induced incidents. Further studies would also benefit from more contextual insights on past incidents, to better categorise the incident and identify their hydrological predictors. Though results show a more nuanced approach is necessary to predict incidents across the SNCF network, important insights have been found to set further research *on the right track*.

6 Recommendations

The main results and insights gained from this study were distilled into the following recommendations for SNCF-Réseau (ordered in terms of importance and feasibility). Some recommendations (notably 4 & 5) may not follow directly from the results, but from conversations and other activities during the project. Lastly, it is noted that SNCF-Réseau are already actively focusing on many of the following recommendations; these recommendations are still listed here to encourage their continued progress to the benefit of predicting and forecasting precipitation induced incidents.

1. **Refine incident logging protocol:**
   Formal definitions could be made for incident types. Allow for railway workers to log contextual information succinctly, such as the state of track, how it was damaged and what the source of damage was. Such insights are useful to better identify which (hydrological) variables are needed to predict the incident. Insights from statistical risk analyses and from experts could be crucial inputs to this task (see recommendations 4 and 5). We recommend SNCF to account for this issue when devising improvements to their risk management protocol.

2. **Pursue radar-based nowcasting and forecasting products in vulnerable catchments:**
   Results indicate that incidents occurring due to localised runoff caused by intense precipitation events in small, steep catchments may be those most predictable using precipitation data. This study recommends exploring other nowcasting and forecasting products (such as AROME) when monitoring these catchments and investigate their ability to predict and forecast incidents in these catchments.

3. **Expand network monitoring data to include track vulnerability:**
   Results indicate that precipitation data may not be enough to reliably predict incident occurrence. This study agrees with ongoing efforts within SNCF-Réseau to monitor and collect data on the state of their railway assets. Data on railway infrastructure vulnerability could be analysed, coupled with more contextual insights on each incident (including zoomed-out photos of railway damages). These efforts can reveal trends in infrastructure vulnerability which can predict incident occurrence.

4. **Engage with those on the ground:**
   We encourage SNCF-Réseau to further facilitate active engagement between their central office and their agents on the ground. Regional officers who regularly observe the state of the railway assets have an intimate understanding of their local railway infrastructure that may be hard to quantify. Insights from the ground can guide further research in identifying which variables to focus on when monitoring the hydrological state of the catchment, as well as the state of the railway assets themselves.

5. **Identify main incident predictors across SNCF network:**
   Overlay incidents with other geospatial information to identify main variables that predict incidents. Important indicators can include topographical derivatives, soil types, land-use, return-period thresholds for precipitation or runoff. These variables can also be adapted to reflect future scenarios to anticipate changes in climate and land-use. Results, possibly represented as a risk map, will aid identifying important predictors for incident occurrence where monitoring and mitigation measures should be prioritised. SNCF is currently undertaking such an analysis.

7 References

Arnaud, P., Cantet, P., & Odry, J. (2016). SHYREG, a national database of flood frequency estimation. *3rd European Conference on Flood Risk Management (FLOODrisk 2016)*, *7*, 7.

Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, *5*(1), 1–12.

Bles, T., van Marle, M., Boonstra, H., van Muiswinkel, K., & de Bel, M. (2023). Multi-hazard resilience assessment and adaptation planning for the Dutch highway network. *Transportation Research Procedia*, *72*, 3801–3808. https://doi.org/https://doi.org/10.1016/j.trpro.2023.11.508

Bouttier, F., & Marchal, H. (2024). Probabilistic short-range forecasts of high precipitation events: optimal decision thresholds and predictability limits [PREPRINT]. *EGUsphere*, *2024*, 1–30.

Brousseau, P., Seity, Y., Ricard, D., & Léger, J. (2016). Improvement of the forecast of convective activity from the AROME-France system. *Quarterly Journal of the Royal Meteorological Society*, *142*(699), 2231–2243.

Bruckner, T., Bashmakov, I. A., Mulugetta, Y., Chum, H., & de la Vega Navarro, A. (2014). *Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*.

Bubeck, P., Dillenardt, L., Alfieri, L., Feyen, L., Thieken, A. H., & Kellermann, P. (2019). Global warming to increase flood risk on European railways. *Climatic Change*, *155*(1), 19–36. https://doi.org/10.1007/s10584-019-02434-5

Cheetham, M. (2024). *Personal Communication*.

Cheetham, M., Chirouze, F., & Bredier, L. (2016). RISK VIP: Evaluation of Flood Risk on the French Railway Network Using an Innovative GIS Approach. *E3S Web of Conferences: FLOODrisk 2016 - 3rd European Conference on Flood Risk Management*, *7*(10004).

Dikanski, H., Imam, B., & Hagen-Zanker, A. (2018). Effects of uncertain asset stock data on the assessment of climate change risks: A case study of bridge scour in the UK. *Structural Safety*, *71*, 1–12.

Doll, C., Trinks, C., Sedlacek, N., Pelikan, V., Comes, T., & Schultmann, F. (2014). Adapting rail and road networks to weather extremes: case studies for southern Germany and Austria. *Natural Hazards*, *72*, 63–85.

EauFrance. (2024). *Hydroportail*. https://www.hydro.eaufrance.fr/

ECMWF. (2024). *ECMWF TIGGE Archive*. apps.ecmwf.int/datasets/data/tigge/

Edwards, D. C., & McKee, T. B. (1997). *Characteristics of 20th century drought in the United States at multiple time scales* (Vol. 97). Colorado State University Fort Collins.

Gijsbers, P., Werner, M. G. F., & Schellekens, J. (2008). *Delft FEWS: A proven infrastructure to bring data, sensors and models together*.

Gomes, G., Thiemig, V., Skøien, J. O., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J., Kolbe, C., Pichon, D., Schweim, C., & Salamon, P. (2020). *EMO: A high-resolution multi-variable gridded meteorological data set for Europe*. https://doi.org/10.2905/0BD84BE4-CEC8-4180-97A6-8B3ADAAC4D26

Guerreiro, M. J., Lajinha, T., & Abreu, I. (2007). *Flood analysis with the standardized precipitation index (SPI)*.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., & Thépaut, J.-N. (2023). *ERA5 hourly data on single levels from 1940 to present*. https://doi.org/10.24381/cds.adbb2d47

Hutchins, G. (2018). *ISO 31000: 2018 enterprise risk management*. Greg Hutchins.

Imhoff, R. O., Brauer, C. C., van Heeringen, K.-J., Uijlenhoet, R., & Weerts, A. H. (2022). Large-sample evaluation of radar rainfall nowcasting for flood early warning. *Water Resources Research*, *58*(3), e2021WR031591.

Kellermann, P., Schöbel, A., Kundela, G., & Thieken, A. H. (2015). Estimating flood damage to railway infrastructure - The case study of the March River flood in 2006 at the Austrian Northern Railway. *Natural Hazards and Earth System Sciences*, *15*(11), 2485–2496. https://doi.org/10.5194/nhess-15-2485-2015

Kellermann, P., Schönberger, C., & Thieken, A. H. (2016). Large-scale application of the flood damage model RAilway Infrastructure Loss (RAIL). *Natural Hazards and Earth System Sciences*, *16*(11), 2357–2371.

Kerin, I. (2020). *The development of a bridge management system involving standardised scour inspection procedures and flood forecasting* [Doctoral Thesis, University College Cork, Ireland]. https://hdl.handle.net/10468/11414

Kirchner, J. W. (2009). Catchments as simple dynamical systems: Catchment characterization, rainfall-runoff modeling, and doing hydrology backward. *Water Resources Research*, *45*(2).

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge didistribuzione. *Giorn Dell'inst Ital Degli Att*, *4*, 89–91.

McKee, T. B., Doesken, N. J., & Kleist, J. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, *17*(22), 179–183.

Ochsner, M., Palmqvist, C.-W., Olsson, N. O. E., & Winslott Hiselius, L. (2023). The effects of flooding on railway infrastructure: A literature review. *Transport Research Procedia* , *72*, 1786–1791.

Petrova, E. (2020). Natural hazard impacts on transport infrastructure in Russia. *Natural Hazards and Earth System Sciences*, *20*(7), 1969–1983.

Seiler, R. A., Hayes, M., & Bressan, L. (2002). Using the standardized precipitation index for flood risk monitoring. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *22*(11), 1365–1376.

Smirnov, N. (1948). Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, *19*(2), 279–281.

SNCF Réseau. (2023). *Liste des ponts rail*.

SNCF Réseau. (2024). *Formes des lignes du RFN*. https://ressources.data.sncf.com/explore/dataset/fichier-de-formes-des-voies-du-reseau-ferre-national/information/

Thieken, A. H., Bessel, T., Kienzler, S., Kreibich, H., Müller, M., Pisi, S., & Schröter, K. (2016). The flood of June 2013 in Germany: how much do we know about its impacts? *Natural Hazards and Earth System Sciences*, *16*(6), 1519–1540.

Thiemig, V., Gomes, G. N., Skøien, J. O., Ziese, M., Rauthe-Schöch, A., Rustemeier, E., Rehfeldt, K., Walawender, J. P., Kolbe, C., & Pichon, D. (2022). EMO-5: a high-resolution multi-variable gridded meteorological dataset for Europe. *Earth System Science Data*, *14*(7), 3249–3272.

UNISDR. (2016). *Report of the Open-Ended Intergovernmental Expert Working Group on Indicators and Terminology Relating to Disaster Risk Reduction* (Issue A/71/644).

Verkade, J. S., & Werner, M. G. F. (2011). Estimating the benefits of single value and probability forecasting for flood warning. *Hydrology and Earth System Sciences*, *15*(12), 3751–3765.

Vlček, O., & Huth, R. (2009). Is daily precipitation Gamma-distributed?: Adverse effects of an incorrect use of the Kolmogorov–Smirnov test. *Atmospheric Research*, *93*(4), 759–766. https://doi.org/https://doi.org/10.1016/j.atmosres.2009.03.005

Werner, M., Schellekens, J., Gijsbers, P., van Dijk, M., van den Akker, O., & Heynert, K. (2013). The Delft-FEWS flow forecasting system. *Environmental Modelling & Software*, *40*, 65–77.

Wilks, D. S. (2011). *Statistical methods in the atmospheric sciences* (Vol. 100). Academic press.
World Meteorological Organization. (2017). *Guidelines for Nowcasting Techniques* (2017th ed.).

Deltares is an independent institute for applied research in the field of water and subsurface. Throughout the world, we work on smart solutions for people, environment and society.

Deltares

www.deltares.nl