



# A framework for now-casting and forecasting in augmented asset management

Jaya Kumari<sup>1</sup> · Ramin Karim<sup>1</sup> · Adithya Thaduri<sup>1</sup> · Pierre Dersin<sup>1</sup>

Received: 28 April 2022 / Accepted: 8 June 2022 / Published online: 6 July 2022  
© The Author(s) 2022

**Abstract** Asset Management of a complex technical system-of-systems needs cross-organizational operation and maintenance, asset data management and context-aware analytics. Emerging technologies such as AI and digitalisation can facilitate the augmentation of asset management (AAM), by providing data-driven and model-driven approaches to analytics, i.e., now-casting and forecasting. However, implementing context-aware now-casting and forecasting analytics in an operational environment with varying contexts such as for fleets and distributed infrastructure is challenging. The number of algorithms in such an implementation can be vast due to the large number of assets and operational contexts for the fleet. To reduce the complexity of the analytics, it is required to optimize the number of algorithms. This can be done by optimizing the number of operational contexts through a generalization and specialization approach based on both fleet behaviour and individual behaviour for improved analytics. This paper proposes a framework for context-aware now-casting and forecasting analytics for AAM based on a top-down, i.e., Fleet2Individual and bottom-up, i.e., Individual2Fleet approach. The proposed framework has been described and verified by applying it to the context of railway rolling stock in Sweden. The benefits of the proposed framework is to provide industries with a

tool that can be used to simplify the implementation of AI and digital technologies in now-casting and forecasting.

**Keywords** Now-casting · Forecasting · Asset management · Augmented asset management · Fleet management · Rolling stock

## 1 Introduction

An effective and efficient Asset Management (AM) is important to achieve operational excellence in asset intensive organizations such as railways. AM is defined as a coordinated activity of an organization to realize value from assets (ISO 2014). Currently, the AM framework includes the engineering and governance perspectives (Mardiasmo et al. 2008). From the engineering perspective, AM focuses on parameters related to cost and asset performance, such as availability, reliability, maintainability, and supportability parameters. From the governance perspective, AM focuses on aspects such as organization, ownership, responsibility, credibility, accountability.

Asset management is aimed at facilitating decision making and optimization (Tam and Price 2008). Therefore, managing complex technical assets in industry during their whole lifecycle is highly dependent on availability and accessibility of two main components: (a) The facts that provide data related to various features of the asset; and (b) The algorithms that extract information and knowledge from the facts. The developments in Information and Communication Technology (ICT) have increased the availability of data and accessibility to algorithms that were computationally expensive before. There have also been continuous advances in more efficient data structures and algorithms that require lower computational cost. This has enabled the

---

✉ Jaya Kumari  
jaya.kumari@ltu.se

Ramin Karim  
ramin.karim@ltu.se

Adithya Thaduri  
adithya.thaduri@ltu.se

Pierre Dersin  
pierre.dersin@ltu.se

<sup>1</sup> Luleå University of Technology, Luleå, Sweden

use of Artificial Intelligence (AI) as a potentially powerful tool to solve problems related to AM, operation, and maintenance (Rane and Narvel 2021; Tam and Price 2008).

The level of automation in AM can be enhanced by augmenting the decision making related to AM processes using AI tools and technologies. The concept of Augmented Asset Management(AAM) proposed by (Kumari et al. 2021), is based on the provision of a platform that aims to facilitate integration and transformation of data into information, knowledge and context models that are based on available sources of data, keeping the domain specific challenges at its center, and adapting the analytics to the needs of the end-user. AAM is expected to increase the efficiency and effectiveness of AM through the utilisation of AI and digital technologies.

AM of complex technical system-of-systems needs cross-organizational operation and maintenance, asset data management and context-aware analytics (Kumari et al. 2021). This creates the need for a common and distributed platform to integrate, analyze, and share information regarding overall asset health and performance. Emerging technologies such as AI and digitalization can facilitate the augmentation of AM by providing data-driven and model-driven approaches to analytics, i.e., now-casting and forecasting.

Now-casting deals with understanding the characteristics of the data and how it interacts with the Key Performance Indicators (KPI) identified for the system (Bragoli and Modugno 2017). Forecasting deals with predicting the future state of a system as accurately as possible based on historical data and the knowledge of future events that might impact the forecast (Hyndman and Athanasopoulos 2018).

Now-casting and forecasting analytics utilize the available data for a system and transform it into useful information based on the requirements of industrial contexts. These analytics aim to provide insight and knowledge about the relationships in data such as correlations and causalities of real-world phenomena. This often involves intensive processing of vast amounts of data. Factors such as lack of data, and insufficient data quality related to assets can reduce the accuracy of the now-casting and forecasting models.

To enable context-aware now-casting and forecasting, the provision of a concept to describe the industrial context is essential. One of the concepts for context-aware analytics suggested by (Karim et al 2016), describes four (4) phases of analysis, i.e.: (1) Descriptive Analytics – ‘What has happened?’; (2) Diagnostic Analytics – ‘Why something has happened?’; (3) Predictive Analytics – ‘What will happen in the future?’; and (4) Prescriptive analytics – ‘What needs to be done?’. From a knowledge discovery perspective, these four phases of analytics are interdependent and provide insight into aspects such as fault identification, fault isolation, and physics of failure. (Karim et al. 2016) These phases of analytics enhances industries’ capability in

‘Now-casting’ and ‘Forecasting’. The now-casting phase constitutes context-based descriptive and diagnostic analytics while the forecasting phase constitutes ‘predictive’ and ‘prescriptive’ analytics.

However, implementing context-aware now-casting and forecasting in an operational environment with vast number of contexts is challenging. If the analytics is adapted to each individual context, then the overall architecture and infrastructure of the AI solution becomes complex and complicated. This is due to a number of components in the solution and the interaction between these components. Simultaneously, the individual characteristics might not be well represented, when increasing the level of abstraction of the solution and developing generic algorithms to address multiple contexts, the individual characteristics of each context might not be well represented. This can reduce the accuracy of the analytics. Hence, optimizing the number of algorithms that can represent a system, becomes fundamental for the overall availability, reliability, maintainability, and supportability of the AI-solution.

Fleet management of railway rolling stock consists of a number of vehicles with different operational contexts such as the operating environment, the stakeholders that are responsible for ownership, operation and maintenance, and the end users. It is challenging to model the behavior of each individual in a fleet as it increases the complexity of the architecture and infrastructure of the AI solution, with requirement for data acquisition and model development for every individual operating in different contexts. The Individual2Fleet and Fleet2Individual approach proposed by (Kumari et al. 2021) can be used to simplify the infrastructure of the AI solution. The Fleet2Individual and Individual2Fleet approach is based on making generalisations for the fleet and identifying the individual behaviors that are outliers from the fleet. This approach can help to optimize the number of digital assets such as data acquisition and models that are required to represent the behavior of the fleet without compromising the unique behaviors of the individuals.

Hence, this paper proposes a framework for context-aware now-casting and forecasting analytics for AAM of railways based on the Fleet2Individual and Individual2Fleet approach. The proposed framework is described and verified by applying it to the case of fleet of railway rolling stock in Sweden. The practical implications of the proposed framework is to provide industries with a tool that can be used to simplify the implementation of AI and digital technologies in now-casting and forecasting.

## 2 Background

The introduction of AI empowered analytics to industry needs to be adapted to the industrial contexts. Industrial AI

(IAI), is part of the AI that is relevant and adapted to the specific characteristics of industrial contexts (Lee 2020). The four phases of the analysis used in context-aware analytics, suggested in the concept described by (Karim et al. 2016), is shown in Fig. 1.

Descriptive analytics gives a picture of the past and present condition of an asset with the use of narrative tools such as bar charts, scatter plots, pie charts etc. Diagnostic analytics gives an insight into the event that has happened. Additional data that might be related to the event, is needed for diagnostic analytics in order to explain the event. Data mining and data integration tools are needed for diagnostic analytics. Predictive analytics is based on the historical behavior of assets and knowledge of the future operating conditions, in order to predict failures in future. Techniques such as regression analysis, predictive modelling, machine learning, time series forecasting etc. are needed for predictive analytics. Prescriptive analytics provides insights on proactive strategies that can be implemented in order to prevent failures in the future based on the forecasting from predictive analytics. Optimization tools and simulations are needed for prescriptive analytics. (Famurewa et al. 2017).

These phases of analytics enhance industries' capability in now-casting and forecasting. Now-casting in industrial context is the process of providing the capability that aims to bring insight into what has happened and why has it happened. Now-casting uses data to recognize isolate and identify a real-world phenomenon. This process also tries to describe the underlying causes, which have led to the appearance of the identified phenomenon. Forecasting, on the other hand, refers to the process of providing capability that aims to predict what will happen in the future. Forecasting uses data and models to predict the appearance of a real-world phenomenon. Forecasting is essential to enable

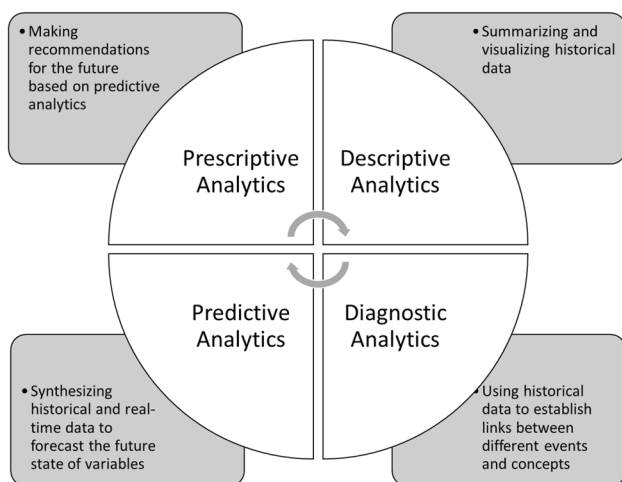
prescription that aims to prevent unwanted situations and failures in industry.

These four phases of analytics are data driven and model driven. To integrate data from multiple heterogeneous sources, for effective decision making, tools and technologies to acquire, integrate and visualize data are needed (Extract- 2015).

**Data acquisition**—One of the key performance drivers for asset management is data acquisition and data pre-processing for performing data analysis and decision support to improve the performance of engineering assets (Mathew et al. 2009). Furthermore, the utilization of the relevant engineering data needs to be transferred to information for improved reliability, availability, safety, efficiency, and sustainability of assets (Murphy and Chang 2009). The data usability also needs to be implemented for an efficient way of leveraging the data (Tretten and Karim 2014). Especially, in fleet management, there is an immense value of utilizing the data for information, since the data for each asset in the fleet needs to be analysed (Kinnunen et al. 2016). Hence, there is a need to investigate the quality of data before analyzing it (Aljumaili et al. 2016).

**Data integration**—The data gathered from various databases often have heterogeneity of datasets in terms of data types such as numeric, textual, image, audio, video, point cloud etc. However, all these disparate types of data need to be integrated to generate value for effective decision support for AM (Thaduri et al. 2015). It is also required to integrate asset data not only from computer maintenance management systems (CMMS) and online condition monitoring (CM) systems but also from supervisory control and data acquisition (SCADA) that are rarely used for asset diagnosis and prognosis (Galar et al. 2012). An architecture needs to be developed based on existing datasets to obtain a unified database that can be further exploited for data analysis. The advantages of data integration are to improve the data quality by decreasing inconsistencies, duplication, and conflicts (Beck et al. 2007; Rane and Narvel 2022). Integration can also be done by optimizing horizontal (crosscutting across various databases) and vertical (from higher level to lower level) databases for consistency (Grossmann et al. 2010). Automated data integration with expert judgement and learning-based approaches will be efficient in terms of utilization, processing, and usability for the data analysis (Dong and Rekatsinas 2018).

**Metaanalysis** – MetaAnalysis refers to an automated process of data pre-processing, feature selection, and selection of the suitable machine-learning algorithm based on performance evaluation of different algorithms on the given data set. The MetaAnalysis process is expected to expedite the initial analysis of newly received datasets (Kumari et al. 2022).



**Fig. 1** The four phases of analytics for context aware now-casting and forecasting analytics

From a knowledge discovery perspective, the four phases of analytics in Fig. 1 are inter-dependent and provide insight to aspects such as fault detection, fault isolation, and fault identification. Fault detection, isolation and identification can ensure the desired performance of a complex system-of-systems both in the presence and absence of faults (Hwang et al. 2010).

**Fault** – A fault is defined as a state of an item or system characterized by its inability to perform a required function, excluding the inability during preventive maintenance or other planned actions, or due to lack of external resources (CEN 2010). To improve the functioning of the system, fault diagnosis is performed to monitor the location, and identify the faults within the system (Wakiru et al. 2019). The goal for the early fault diagnosis is to get sufficient time for counter measures which include planning of maintenance actions, such as repair, replace, postponing of operations, etc. (Isermann 2005).

**Fault detection**—Fault detection involves detecting and reporting an anomalous condition. Generally, it is based on measurement data from various instrumentation systems and information provided by human operators. This process requires analytical knowledge and understanding of information provided by humans which is called heuristic knowledge. Hence, fault detection is often considered as a knowledge-based approach.

**Fault isolation**—Fault Isolation determines which component/item has a fault, such as, kind, shape, size, location, and time of fault by evaluating symptoms. For fault detection, analysis of a single measurement can be sufficient, while a set of measurements is normally necessary for performing fault isolation. To accomplish it, a set of structured attributes need to be designed. Each measurement is defined to be sensitive to a subset of faults, whilst remaining measurements are to be insensitive to other faults. The measurement set has the necessary sensitivity to an individual fault and is insensitive to other faults. This method is also known as residual-based method which relies on relationships and characteristics of all items of interest (Meskin and Khorasani 2011), (Rajeswari et al. 2014).

**Fault identification**—can be considered as an important objective of a fault diagnosis framework. Fault identification is the method for estimating information related to the fault after its detection. The identification outcome facilitates fault isolation through evaluating the expected measurement values with their normal values. It also determines the nature of the fault (Jardine et al. 2006).

Fault detection, isolation and identification is more challenging in fleet management due to the large amount of data collected from each individual in a fleet and the similarity and variability between the behaviors of the individuals in a fleet that require adaptive models (Zaccaria et al. 2018). These adaptive models can be derived based on the

similarity analysis between the individuals of the fleet and context adaptation based on the similarity analysis in situations, where the individuals of a fleet operate in different contexts. Similarity analysis is done by identifying a feature for which similarity/dissimilarity is measured, keeping the other features constant (Wang and Megalooikonomou 2008).

In this paper the above concepts have been applied to the augmented asset management in railways for the fleet management of rolling stock. There are studies for fleet management that aim to optimize aspects such as resource planning and management, logistics etc. (Penna et al. 2017, Bnouachir et al. 2020) for physical assets. In addition, there are several studies for forecasting in railways, for example, freight diesel locomotive (Lingaitis et al. 2014), railway traffic (Konstantinova et al. 2021), noise level (Orynach et al. 2021), train delays due to the condition of switches and crossings (Thaduri 2020), predicting train disruptions (Fink et al. 2013), data mining to draw up risk and control plans for trains (Kalathas and Papoutsidakis 2021), mathematical models for effective and efficient utilization of railway freight cars (Milenkovic and Bojovic 2019), and a simulation of railway fleet maintenance based on interactions of components such as rolling stock, infrastructure etc. (Bury et al. 2018). These studies have provided various optimization models for the fleet management of physical assets in railways.

The fleet management approach may also be used for optimization of digital assets such as a number of algorithms used in analytics. A similar framework for integration of models for each asset in a fleet of tools, to generate risk indices has been proposed by (Thomas et al. 2020). However, this model does not consider the optimization of the number of individual models based on the comparisons between the fleet behavior and the individual behavior. In context-aware now-casting and forecasting, the number of algorithms can be optimized by optimizing the number of operational contexts. The optimization might be done by a generalization and specialization approach based on fleet behavior and individual behavior. This has been introduced as a top down, i.e. Fleet2Individual, and bottom up, i.e., Individual2Fleet, approach to analytics (Kumari et al. 2021).

However, there is a need for an appropriate framework to enable the implementation of Fleet2Individual and Individual2Fleet approach in an industrial context of complex technical system-of-systems for now-casting and forecasting. A framework is defined as a meta-level model (a higher-level abstraction) through which a range of concepts, models, techniques, methodologies can either be clarified and/or integrated (Jayaratna 1994). The, proposed framework in this paper integrates the concepts discussed in this section and it consists of the categorization of components that have a process flow of interconnected steps. This proposed framework is applied and verified for the case of fleet of rolling stock in Sweden.

### 3 The proposed framework

The proposed framework for now-casting and forecasting consists of four (4) interconnected components i.e.: (a) context definition; (b) Data Extract/Transform/Load (ETL); and (c) Fault detection and identification, and 4) Now-casting and forecasting analytics. These components and their inherent items/steps are aimed to facilitate the analytics process in fleet asset management. Figure 2 illustrates the proposed framework and the relationships between its components and steps.

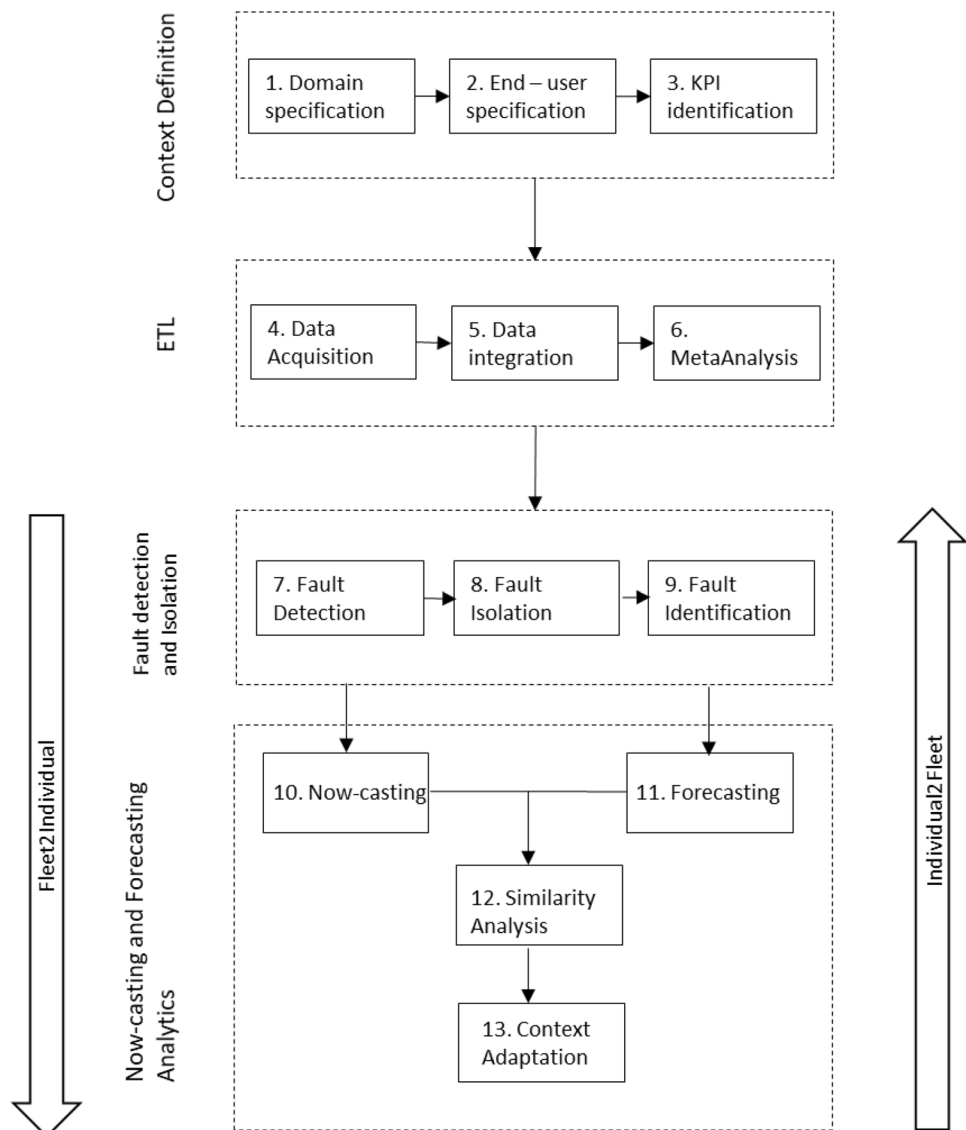
#### 3.1 Component: I. Context definition

*Step 1—Domain specification*—Industrial AI aims to apply AI enabled technologies to develop analytic services for real world challenges in industrial applications. To develop

an analytic service, it is important to define the industrial domain for which the service is being developed. An industrial domain definition can be on a higher abstraction level, such as ‘railways,’ or for a specific challenge within an industrial domain, such as ‘preventive maintenance of engines and compressors in railway vehicles in Norbotten.’ The subsequent steps of the framework are dependent on the industrial domain specified in Step 1.

*Step 2—End-user specification*—The analytic services developed for an industrial domain need to be further adapted to the context of the end user. For instance, the vehicle owners of a fleet of railway vehicles are interested in the overall performance of the fleet, while a maintenance engineer is interested in insights on failure and repair times of individual components used in the railway vehicles. The definition of the end user context leads to the next step of defining the KPIs.

**Fig. 2** The proposed framework for now-casting and forecasting analytics in augmented asset management with a fleet2individual and individual2fleet approach





*Step 3—KPI identification*—Now-casting analytics uses available data to calculate parameters of interest that are not known otherwise. These parameters of interest can be the KPIs for end users in an industrial domain. The end users might be interested in more than one KPI that contributes to factors such as cost, sustainability, short term goals and long term goals for their organization. However, in the first iteration of developing the analytic services, it is recommended to select a single KPI of interest. This can further contribute to complex analysis based on multiple KPIs the interactions and correlations between them. The efficiency and effectiveness of the analytic services will depend on the stepwise identification of KPIs in the domain either through literature survey or through interviews with domain experts.

### 3.2 Component: II. ETL

*Step 4—Data acquisition*—Data acquisition is dependent upon the nature of operational assets, nature of the organizations, and their strategy goals to achieve desired outcomes. Engineering asset data comprises of configuration and baseline data, asset condition data, event or incident data, environmental data and process data. Data quality of the acquired data is one of the key concerns from an application area point of view. Data quality has also a significant impact on decisions and their respective consequences since decisions with the help of data-drive algorithms.

*Step 5—Data integration*—In this step the asset data from multiple heterogenous sources is integrated into a unified database that is used for further analysis. This unified database consists of operation, maintenance, failure, and condition monitoring data of the fleet and individual assets of the fleet over different spatial and temporal operational contexts.

*Step 6—MetaAnalysis*—The MetaAnalysis of the integrated data automates the process of data pre-processing., the identification of important features that contribute to the KPIs, and the identification of suitable models for development of analytic services on the given data. This step, therefore, identifies and removes data quality issues such as missing, incomplete, and duplicate data and speeds up feature and model selection through automation.

*Fleet2Individual Individual2Fleet* – The fault detection, isolation and identification steps of the proposed framework are carried out in parallel (a) and (b) processes as shown in Fig. 2, for the whole fleet and for individual vehicles in the fleet, respectively. Establishing the two-way relationship between the fleet and the individual, helps to generalize on the fleet level that account for each individual of the fleet, as well as to identify individuals/groups of individuals that differ from those generalizations. The conditions encountered

by the individual vehicles are used to define the context of the fleet when the individual behavior is same as the fleet. Similarly, the context of the fleet specifies the expected behavior of an individual in that context, and therefore helps in identifying outliers.

### 3.3 Component: III. Fault detection and identification

*Step 7—Fault detection*—This step consists of (1) defining the normal behavior of the system under consideration (2) defining the fault i.e., the deviation from normal behavior, and (3) spotting the occurrence of the defined fault.

*Step 8—Fault isolation*— Fault isolation in the proposed framework, refers to the description of the characteristics of the fault such as, the type of fault, the location of occurrence and the time of occurrence. In this step, the fault is isolated by narrowing down the considered values of these characteristics, by one characteristic at a time. This is done to identify the set of characteristics that represent the majority of the faults occurring in the system.

*Step 9—Fault identification*—In this step the impact of the isolated fault on the overall system performance is identified. This is done by analyzing the isolated faults as a percentage of the total faults occurring in the system. This helps in identifying the extent of impact of the isolated fault on the system.

### 3.4 Component: IV. Now-casting and forecasting analytics

*Step 10—Now-casting*—The information about the system, that is extracted from fault detection, isolation and identification steps, is used for now-casting to answer the questions (1) What has happened and (2) Why it has happened. It is done by visualizing the historical and the current state of the KPIs and identifying trends or patterns in the data.

*Step 11—Forecasting*—The trend or pattern identified in the historical data is projected into the future under similar operational contexts, in order to forecast the future value of the KPIs. This can be done using various modelling techniques that involve a predicted variable and one/multiple predictor variables. The selection of now-casting and forecasting models depend upon multiple factors such as the intended outcome of the model, the availability of historical data, the desired model accuracy etc.

*Step 12—Similarity analysis*—The similarity analysis involves identifying the similarity and differences in behavior of the fleet and individuals/group of individuals. This is done by comparing the similarity/dissimilarity between

individuals and the fleet by comparing only one feature and meanwhile making other features constant. The features that remain constant for similarity analysis are narrowed down in the fault detection and isolation steps discussed previously. While narrowing down these features some data points from the entire dataset are lost as they do not represent these features. Therefore, it should be considered that the dataset after still has considerable number of data points so as to represent the behavior of the fleet and the individual. The similarity analysis helps to identify clusters of individuals that behave similarly to each other and to the fleet. This is done by comparing values of the features using visualization techniques, algorithmic comparisons. Similarity analysis when done with multiple variable features can also use machine learning algorithms such as clustering to identify similar groups in the data.

*Step 13—Context adaptation* – The clusters/groups identified in similarity analysis may be due to the individuals operating in similar context. The context adaptation step involves identifying this relationship between formation of clusters to contexts in order to make generalizations about groups of individuals in a fleet that have similar operational context. For scenarios in the future, when there is insufficient data available, the behavior of the individuals can be predicted based on the behavior of other individuals operating in similar context.

## 4 Case study description

The proposed framework in this paper has been verified on the case of railway rolling stock in Sweden. The Swedish Transport Administration (Trafikverket) registers all the deviations from scheduled arrival times for trains. For business management, causing a deviation means that a train is delayed by 5 min or more in the travel between two measuring points that follow directly after each other in the Swedish Transport Administration's system. Causing a deviation also means that a train becomes 5 min or more delayed compared to the timetable at the first measuring point.

The infrastructure manager's responsibility for deviations mainly covers additional delays caused by disruptions to infrastructure or operations management. The responsibility of railway undertakings or transport organizers for deviations includes mainly the railway vehicles and their driving. A quality fee needs to be paid by the organization to which the delay is assigned, for each minute of registered delay.

The additional delays are recorded with an associated reason code from a standardized code list. The time recorded for this additional delay is specifically attributed to the reason codes. The purpose of recording this reason code is to identify the cause of this delay and the party responsible for

addressing it. This reason code is associated with a three-level description. Level-1 (Nivå 1) describes who is the problem owner. Level-2 gives a description of the problem, in terms of where and what. Level-3 codes are entered only in specific cases. Train supervisor/train dispatcher must primarily indicate the first two levels while the third level can be completed later when the party that owns the delay requests for it. Railway companies and traffic organizers must pay a quality fee for Level-1 reason code as 'railway organizations' (Järnvägsföretag). These three level reason codes are used to decide the action that should be taken on these delays. Considering each additional delay record as a failure that impacts the system performance, the combination of these three level reason codes gives a specific failure mode. These delays with Level-1 reason code as 'railway organization' are referred to as 'registered additional delays' in this paper.

The railway organization to which these delays are assigned owns a fleet of 12 railway vehicles. Apart from Level-1, 2, and 3 reason codes, a delay record also contains information about the vehicle number in which the delay was reported, the route on which the vehicle was operating, and the place and time of the occurrence of the delay. Since the railway organization has to pay for each minute of registered additional delay assigned to them, this delay is an important Key Performance Indicator (KPI) to assess the health of the fleet and individual vehicles in the fleet.

## 5 Verification of the proposed framework

### 5.1 Component: I. Context definition

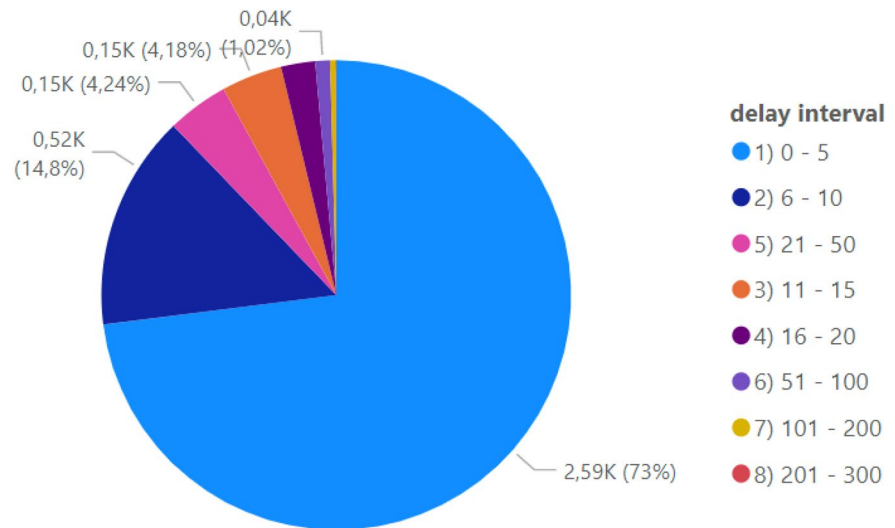
*Step 1—Domain specification*—This paper applies the context based now-casting and forecasting framework for asset management in railways, to the fleet management of railway rolling stock in Sweden. The considered case in the paper is a fleet of 12 passenger railway vehicles operating in Sweden.

*Step 2—End-user specification*—The end users for this case are the vehicle owners of the fleet of railway vehicles operating in Sweden (Fig. 3).

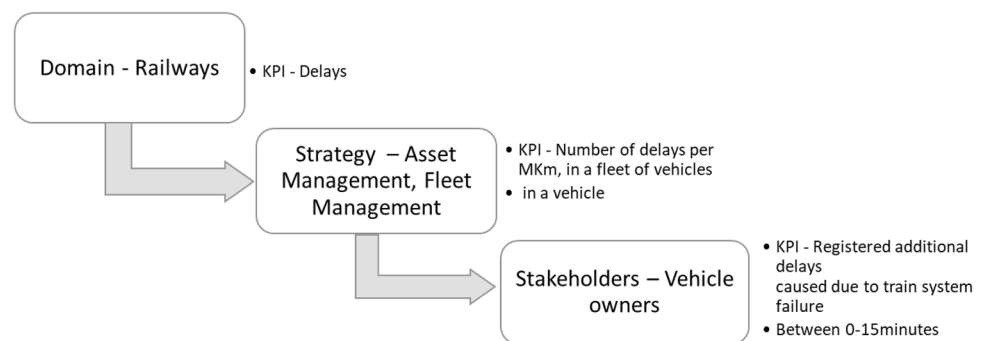
*Step 3—KPI identification*—Once the industrial domain and the end users are identified, the related KPI is identified systematically as shown in Fig. 4. The first step, specifies the domain as railways and identifies arrival delays in trains as a KPI. In the next step, after the strategy for management of railway vehicles is known, two KPIs are defined as the count of delays per Million Km of performance for a (1) fleet of vehicles and for (2) each individual vehicle. Further, the vehicle owners/operators are only responsible for delays

**Fig. 3** The count of delays observed in the fleet of 12 railway vehicles with the delay intervals

**Count of registered additional delays by delay interval for records with no value in level 3 reason code**



**Fig. 4** Systematic identification of KPIs based on the industrial domain, the end-users



that have ‘railway organization’ listed as the Level-1 reason codes for additional delays. Hence, the KPI of concern is only the registered additional delays that are caused due to Railway organizations.

As shown in Fig. 3, over 90% of the delay count comprises of the delays between 0 and 15 min. Due to high frequency of occurrence, registered additional delays between 0 and 15 min are the identified KPI for now-casting and forecasting analytics in order to explain, predict, and reduce these uncertainties.

## 5.2 Component: II. ETL

**Step 4—Data acquisition** – Data from different stakeholders was acquired to carry out analytics for the fleet management of rolling stock. For the analytics in this paper, history of vehicle performance and the history of damage in vehicles was acquired from vehicle owners. The data on arrival delays in vehicles and registered additional delays was acquired by

the Swedish Transport Administration (Trafikverket). The weather data was acquired from open source API based on the time and location of the train operation. The acquired datasets were for a fleet of 12 railway vehicles operating in Sweden. This data is from Jan – 2018 until Dec 2020. All this data were received from different stakeholders in separate csv files.

**Step 5—Data integration** – The data acquired from multiple stakeholders was integrated to obtain a holistic picture of the vehicles in the fleet. The data sources were integrated based on the date associated to the records, the vehicle numbers associated to the records and the place of operation associated to the records.

**Step 6—MetaAnalysis** – This step was done to identify the important features in the data that had an impact on the KPI, i.e. registered additional delays. The parameters used for this analysis were the reason code for the delays, place of



occurrence of the delays, month/quarter/year of occurrence of the delays, and weather parameters such as temperature, humidity, and weather condition associated to the delays. After conducting a regression analysis through an automated MetaAnalyser platform, it was observed that the 'reason code' was the prominent feature that had impact on the KPI. Therefore, reason codes were the selected features for the now-casting and forecasting analytics.

### 5.3 Component: III. Fault detection and identification

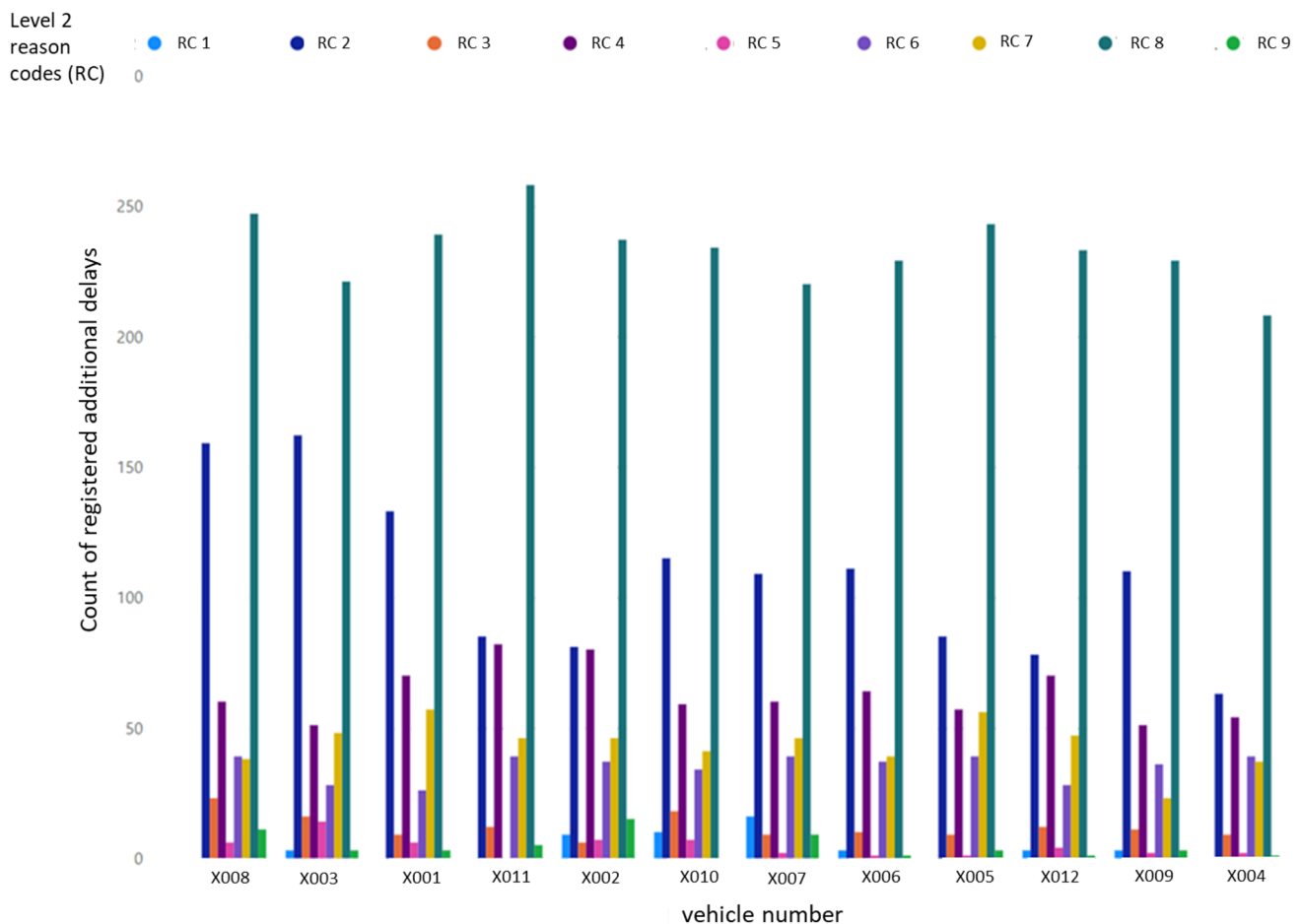
*Step 7—Fault detection*—A registered additional delay is considered as fault in the context of this paper, if it satisfies all the 3 criteria.

1. delays that are > 3 min
2. the Level-1 reason code is assigned as 'railway organization' (Järnvägsföretag)
3. delays that are < 15 min

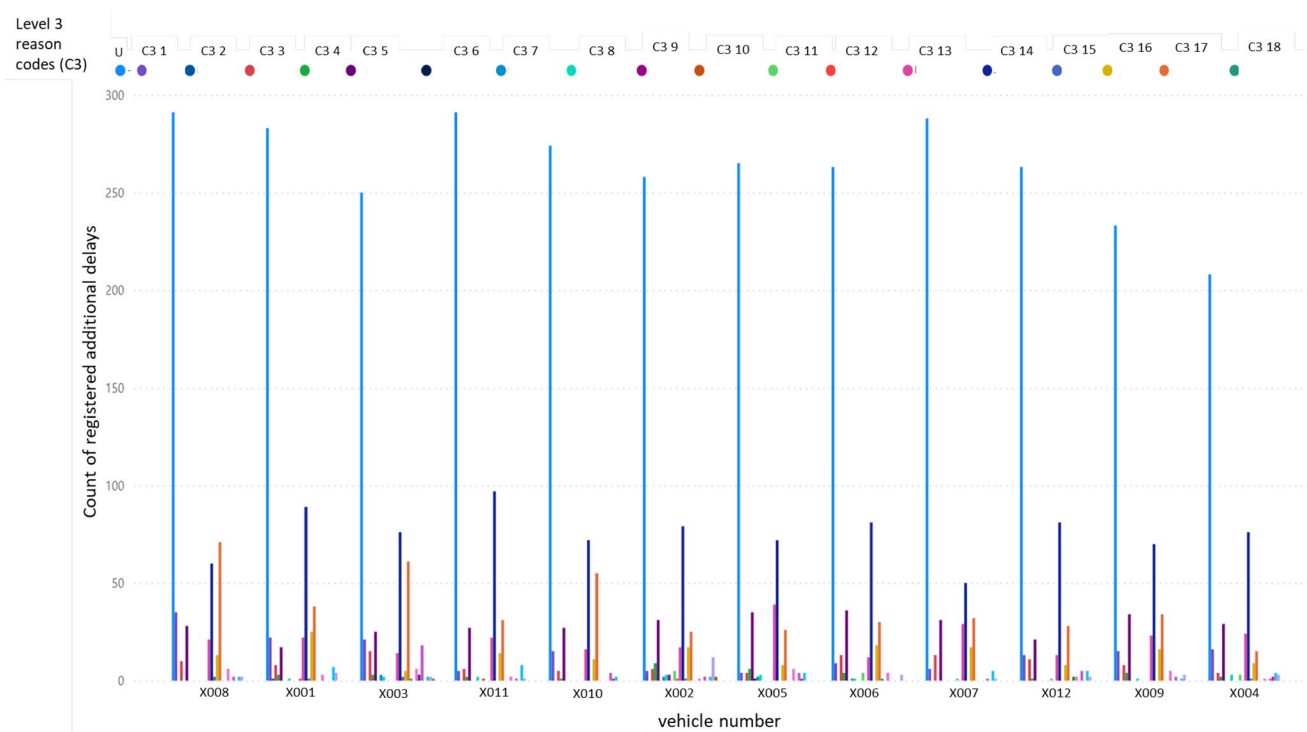
It was observed that there were a total of 6090 registered additional delays in total for the fleet of 12 railway vehicles, that fulfilled the above criteria.

*Step 8 – Fault isolation* -The registered additional delays are assigned with 3 level reason codes. The Level-1 reason code describes the problem owner which is 'railway organization' in the context of this case study. The level 2 and Level-3 reason codes give a broad and narrow description respectively, of the reason of occurrence of the delay (Figs. 5 and 6).

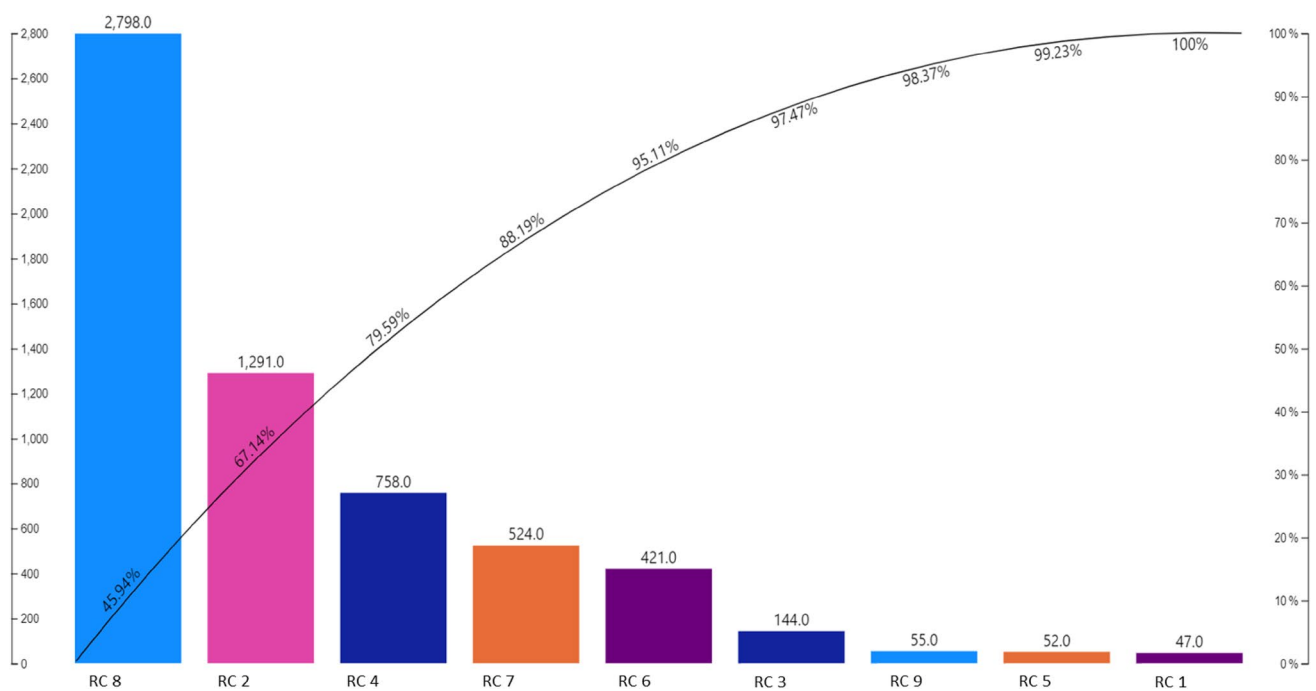
Figure 7 shows a relative distribution of registered additional delays by Level-2 reason codes for the fleet of 12 vehicles. The most frequently attributed Level-2 reason code to the KPI considered in this case is RC8, followed by RC2, RC4, RC7 and RC6. Figure 5 shows a relative distribution of registered additional delays by Level-2 reason codes for each individual vehicle in the fleet. It is observed in Fig. 5 as well that The most frequently attributed Level-2 reason code to the KPI considered in this case



**Fig. 5** The relative distribution of registered additional delays by Level-2 reason codes for individual vehicles



**Fig. 6** The relative distribution of registered additional delays for top 5 Level-2 reason codes by Level-3 reason codes for individual vehicles



**Fig. 7** The relative distribution of registered additional delays by Level-2 reason codes for the fleet

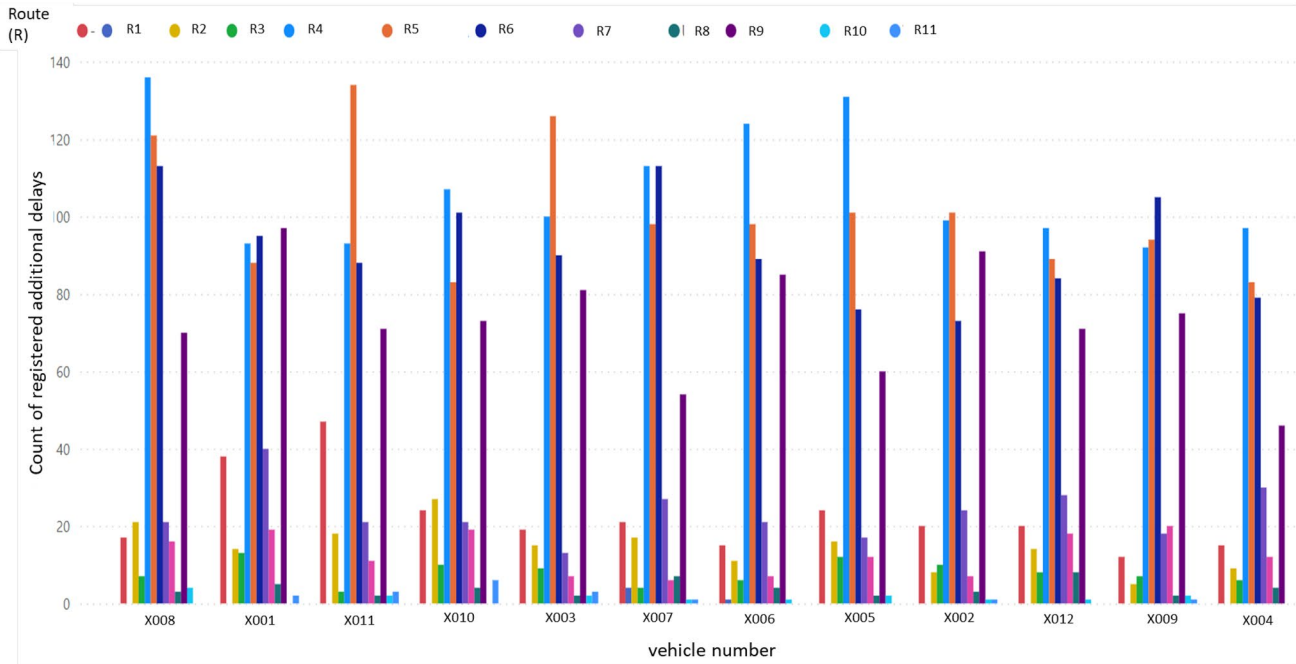
is RC8, followed by RC2, RC4, RC7 and RC6. Based on this above observation, in the next step only the registered additional delays with most frequently attributed Level-2

reason codes (RC8, RC2, RC4, RC7 and RC6) are considered for further analysis.

In Fig. 6, a relative distribution of registered additional delays by Level-3 reason codes with filtered Level-2 reason codes in the previous step, for individual vehicles is shown. Delays with most frequently attributed Level-3 reason codes are identified in this step. It can be seen in Fig. 6, that for most of the recorded delays, the Level-3 reason code is unknown (denoted by U in Fig. 6).

A similar relative distribution of registered additional delays by Level-3 reason codes with filtered Level-2 reason

codes for the fleet of 12 vehicles showed that the fleet behavior is similar to the behavior of individual vehicles in terms of occurrence Level-3 reason codes with 3 most frequently attributed Level-3 reason codes being, ‘U’, ‘C3 14’ and ‘C3 17’. However, the distribution registered additional delays that were attributed to other Level-3 reason codes was observed to be different for the fleet than for individual vehicles. Therefore, in order to represent the fleet behavior as well as the variation in behavior of the individual



**Fig. 8** The relative distribution of registered additional delays for top 5 Level-2 and Level-3 reason codes by train routes for individual vehicles

**Table 1** The percentage of delays that comprises the isolated delays

Vehicle number	Count of total registered additional delays considered as KPIs	Count of registered additional delays in the isolated scenario	Isolated delays as percentage of total delays %
X001	543	373	69
X002	518	364	70
X003	546	397	73
X004	413	305	74
X005	493	368	75
X006	495	396	80
X007	510	378	74
X008	583	440	75
X009	468	366	78
X010	518	364	70
X011	527	386	73
X012	476	341	72
Fleet	6090	4478	74

vehicles, registered additional delays attributed to 8 most frequently attributed Level-3 reason codes were selected for further analysis.

In the next step, a distribution of registered additional delays by their route of operation was observed for individual vehicles as shown in Fig. 8. It was observed that most of the registered additional delays for individual vehicles occurred in routes R4, R5, R6 and R9. A distribution of registered additional delays by route of operation for the fleet showed similar result. Therefore, R4, R5, R6 and R9 were identified as the routes with highest frequency of occurrence of registered additional delays.

**Step 9—Fault identification**—Table 1 shows the isolated delays from step 8 as percentage of the total registered additional delays that were identified as the KPIs for this case. The first column in Table 1 shows the vehicle number. The second column shows the total number of delays that were considered as KPI according to the criteria in step 3 for each vehicle and the fleet. The third column shows the count of delays that were isolated by Level-2 reason codes, Level-3 reason codes and routes of operation in step 8 for each vehicle and the fleet. The last column shows the isolated delays in column 3 as a percentage of total delays in column 2. It

is observed that the isolated delays comprise an average of 70% of the delays registered for each individual vehicle and the entire fleet.

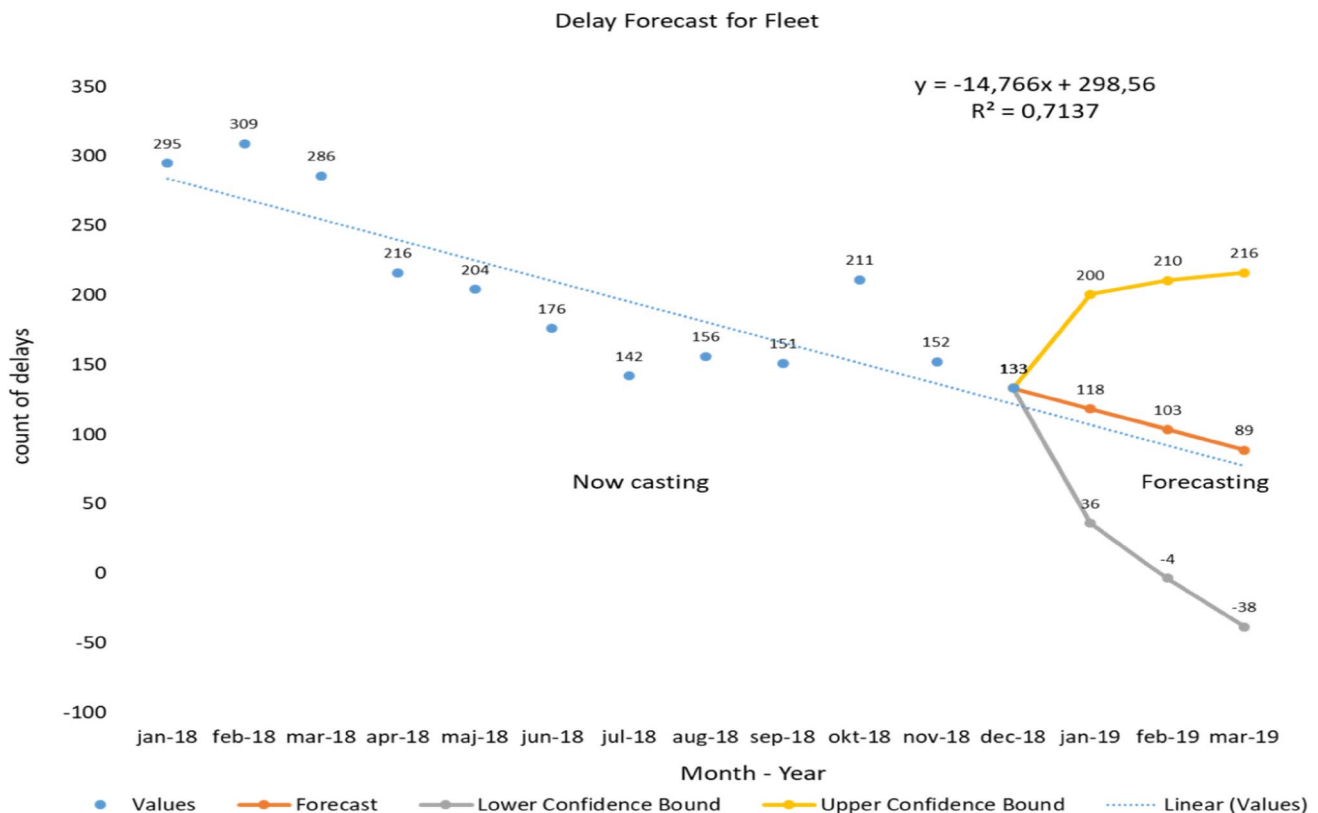
#### 5.4 Component: IV. Now-casting and forecasting analytics

**Step 10, 11—Now-casting and forecasting**—This step demonstrates the now-casting and forecasting for the isolated faults in previous steps. On the left side of Fig. 9 the count of delays for each of the 12 months in the year 2018, is shown which is assumed to be the current/historical state of the system. The historical data for 12 months in 2018 is considered in order to forecast for the next 3 months in 2019, as shown on the right side of Fig. 9. The now-casting and forecasting were performed using a simple linear regression model.

The equation of the trendline as shown in Fig. 9 was observed to be:

$$y = -14.766x + 298.56$$

The R2 value has been used for the evaluation of the linear model.



**Fig. 9** The now-casting and forecasting of the count of delays in each month based identification of the trend in monthly delay count for 12 consecutive

**Table 2** Linear regression model parameters for individual vehicles and for the fleet

Vehicle	Slope	Intercept	R2 score
X001	−0,8396	22,048	0,3237
X002	−1,2643	25,181	0,5775
X003	−2,0571	31,857	0,436
X004	−1,2179	26,143	0,5424
X005	−1,0214	23,105	0,242
X006	−0,6286	20,629	0,1542
X007	−0,775	20,933	0,2049
X008	−1,5964	30,838	0,4002
X009	−1,1143	25,548	0,332
X010	−1,3357	25,819	0,5577
X011	−0,719	19,81	0,2492
X012	−0,6706	20,2	0,3058
fleet	−1,6682	26,995	0,8375

**Table 3** Similarity analysis and consequent grouping of individuals based on the slope value of their individual linear regression models

Slope value range	Vehicles with slope in the give range				
0.6–0.8	X001	X006	X007	X011	X012
1.0–1.3	X002	X004	X005	X009	X010
2.05	X003				
1.5–1.6	X008	fleet			

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}$$

where

$$SS_{res} = \sum_i (y_i - f_i)^2$$

$$SS_{tot} = \sum_i (y_i - \bar{y})^2$$

where  $x$  = independent variable,  $y$  = dependent variable,  $f$  = predicted values of  $y$ .

The  $R^2$  value for the model was found to be 0.7137. Figure 9 shows the now-casting and forecasting for the fleet. Similar individual regression models were developed for each of the 12 vehicles in the fleet. These models are compared for similarity analysis in step 11.

**Step 11—Similarity analysis-** In this step, the slope of the regression models for individual vehicles have been compared with the slope of the regression model for the fleet for similarity analysis. Table 2 shows the linear regression model parameters for individual vehicles and for the fleet. It

is observed that all the vehicles have a negative slope indicating decrease in the number of delays with time.

For similarity analysis, the slope values were divided into 4 intervals as shown in column 1 of. The vehicles with slope values in the same intervals form a group. individual behavior. Based on the groups observed in, the number of linear regression models for now-casting and forecasting of the count of delays for the isolated scenario in step 8,, can be reduced from 13 (fleet and 12 individual vehicles) to 4 (Table 3).

**Step 12—Context adaptation –** The now-casting and forecasting analytics in this paper has been verified for the fleet of rolling stock in Sweden. This fleet consists of passenger trains. Further the context of the considered KPI in this case, i.e. registered additional delays assigned to railway organizations was narrowed down in the fault detection, isolation and identification steps. This has been performed by filtering the delays that were between 0 and 15 min and were assigned to specific Level-2 reason codes, Level-3 reason codes and routes. Further, the now-casting and forecasting model was developed using historical data for the year 2018 to predict the count of delays in the first quarter of 2019. These factors were the considered context for delays in this paper. However, there can be more operational contexts such as place of operation, weather conditions, type of faults that causes the delays, the time to repair the faults etc. that can be considered.

## 6 Discussion

Based on interviews with domain experts in the railway vehicle organizations it was found that the delays that were considered as KPIs in this case are those delays for which the reason of occurrence remains largely unknown. This can be due to the reason that Level-1 and level 2 codes are entered for each registered additional delay, but Level-3 reason codes are provided only when requested by the railway organizations, as explained in the case study description section. This lack of information about the reason of occurrence of the delays can cause uncertainty in the system. Therefore, the now-casting and forecasting of such delays is of importance for the railway organizations.

The fleet of railway vehicles in Sweden is owned operated and maintained by multiple parties such as vehicle owner, local transport authority, vehicle operator and maintenance service provider. The vehicle owners feel the responsibility and need, that all parties handling the vehicles should base their operation and maintenance decisions on a common image of the fleet status and the vehicle status. They are interested in analytics that helps to visualize the overall health of the fleet in terms of KPIs calculated on the fleet



level and on individual vehicle level. For example, after the integration of data, the performance in kilometers, recorded damages, registered delays, failures in components and weather conditions for the vehicles could be retrieved for each vehicle and for the fleet on a specified date, place or interval in time. This provides a common picture of the fleet health to all the stakeholders.

In Sect. 5, the registered additional delays were filtered by Level-2 reason codes, Level-3 reason codes and routes. It is possible to further isolate the delays by identifying contexts such as the place and time of occurrence of the delays. In this case study, it was observed that the distribution of delays in the places Sundsvalls C and Sundsvalls Västra were similar for the fleet and for each individual vehicles in fleet. However, for other places such as Östersunds c, Storien, Ånge, Umeå, the distribution of delays in individual vehicles could not be represented by the fleet distribution. Additionally, the distribution of delays by quarter of the year, where Q1 – Jan–Mar, Q2 – Apr–Jun, Q3 – Jul–Sep, Q4 – Oct–Dec, was observed to be, similar for the fleet and most of the individual vehicles with maximum delays observed in Q1. However, the context for fault isolation in this case study was only considered for Level-2 reason codes, Level-3 reason codes and routes of operation. This was due to the observation that isolation of delays by context such as places and quarter of the year, led to significant reduction in number of data points. This may lead to overfitted models that may not represent the general behavior of the vehicle or the fleet.

It was observed in ‘step 9 – fault identification’ of Sect. 5, that, the isolated delays in step 8 comprise of 70% of the total delays considered as KPI in this case. This implies that the isolated delays represent a considerable fraction of the registered additional delays. The diagnostics performed on the set of characteristics of these isolated delays in terms of reason of occurrence and route of operation, can contribute to the reduction of a substantial number of delays that were recorded for the fleet.

The linear regression model for now-casting and forecasting of registered additional delays in this case study was developed by considering the count of delays for 12 months in the year 2018 as historical data, to predict the count of delays in the first quarter of 2019. This combination was established based on observation of the available data for the count of delays for each month in this case. It was observed that the forecast, was less accurate when the historical data for more than 12 months was considered and when the forecast was made for a longer period than 3 to 4 months in the future. R<sup>2</sup> value was chosen as the performance metric for linear regression model in this case. R<sup>2</sup> value is the proportion of variance in the dependent variable that can be explained by the independent variable. The obtained R<sup>2</sup> value of 0.7137 is considered as a moderate effect of the

independent variable on the dependent variable. The linear model has been chosen for this case study due to simplicity in order to explain the framework. However, other regression models, e.g. exponential, logarithmic, or polynomial, may also be applied for now-casting and forecasting of delays. Furthermore, for the given dataset in this case study, more complex model evaluation should be conducted in order to predict using historical data for 2–3 years and to be able to forecast over more than 3–4 months.

The slope of the trendline of the linear regression model has been considered for similarity analysis in this case study. The slope represents the shape of the trendline. The comparison of the slope is done to compare the footprint of the behavior of the shape between the fleet and individual vehicles. The intercept of a linear model is based on the values of the data, while the slope tells the rate of change in the predicted variable based on the predictor variable.

The similarity analysis is done to optimize the number of algorithms for a system with a generalization approach for the fleet as well as a specialization approach for individual vehicles. The representation of the vehicles in the fleet with fewer models reduces the complexity of the architecture and infrastructure of the AI solution. However, when very few models are representing the fleet, then specialised behavior of some of the individual vehicles might also be lost. Therefore, it is important to find an optimised number of models that can represent the general fleet behavior as well as specialized individual behaviors. When the models developed for a specific context, need to be adapted to a different context, the features of the context should be varied one by one. This should be done in order to identify the parameters that are similar and different in different contexts.

Within this study, a framework for context-aware now-casting and forecasting has been proposed. The proposed framework contains of a number of components and steps, aimed to facilitate and simplify the process of enhanced context-aware analytics in industrial asset management. The framework is proposing a fleet management approach to reduce the management of number of contexts that the analytic models need to consider and adapt to. It is assumed that by increasing the abstraction level of the algorithms for now-casting and forecasting, the universality of individual model can be increased. For verifying and validation the level of universality of analytics, a similarity check has been included in the framework. The similarity check tries to identify context individuals that are representatives for more than one context, which are considered as a fleet of contexts. In the other words, the proposed framework can be used as a tool to optimize the number of context-based algorithms based on common behavior of a group of individuals operating in different contexts and specialization of the specific characteristics of an individual in a given context.

The practical implication of this work is that it will facilitate the implementation AI-based solutions in industrial context by reducing the complexity related to development, implementation, and maintenance of algorithms during the system-of-interest's whole lifecycle.

The work with development of the proposed framework have been carried out by conceptual modelling of the framework. The conceptual model of the framework has then been verified with the case of railway rolling stock. It is believed that the proposed framework can be applied to similar industrial contexts for context-aware now-casting and forecasting in augmented asset management for mining and construction industries. However, additional research effort might be necessary to verify the universality of the proposed framework over a system's whole lifecycle. This have not been possible to conduct due to practical limitations, since many of the complex technical systems in industry have a long lifetime, e.g. 10–50 years.

## 7 Conclusions

The purpose of this paper has been to explore if AI empowered now-casting and forecasting can be simplified, and to develop and propose an appropriate framework for context-aware analytics in various industrial contexts. Implementation of now-casting and forecasting to an industrial context with a large number of operational contexts is challenging. To overcome this challenge, the fleet2individual and individual2fleet approach was used. This approach is used to optimize the number of algorithms by optimizing the number of operational contexts. The proposed framework consists of 4 components i.e. context definition, data extract transform and load, fault detection and identification, and now-casting and forecasting analytics. The fault detection and isolation, and now-casting and forecasting analytics have been conducted for the top-down i.e. fleet2individual and bottom up i.e. individual2fleet approach followed by similarity analysis between individuals and fleet and proposing the idea of adaptation of the analytics to the context. The proposed framework was described and verified by using the case of railway rolling stock in Sweden.

Based on the findings from the conducted research activity, it can be concluded that the proposed framework can be utilized as a handrail, by industries dealing with complex technical systems to facilitate the implementation of AI and digital technologies for context-aware now-casting and forecasting analytics. The future work may be focusing on the investigation and identification of appropriate universal technologies, methodologies, and tools within the individual components and steps of the proposed framework. This is to increase the universality of the framework and its inherent components.

**Acknowledgements** First, we would like to convey our appreciation to Sweden's Innovation Agency Vinnova, JVTC (Luleå Railway Research Center), Trafikverket, Alstom, Tågforetagen, Norrtåg, Infranord, Trasnitio, Bombardier, Sweco, Omicold, and Damill and partners for their financial support to carry out this work within the AI Factory project. Secondly, we would like to express our particular gratitude to Eva Sandberg (Transitio) and Håkan Jarl (Norrtåg) for their enlightening feedback and valuable domain knowledge during this research work. We would also like to thank Miguel Castano for providing valuable reviews to the paper.

**Funding** Open access funding provided by Lulea University of Technology.

## Declarations

**Conflict of interest** There are no financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

**Funding** This research is part of the project funded by JVTC (Luleå Railway Research Center).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aljumaili M, Karim R, Tretten P (2016) Metadata-based data quality assessment. *VINE J Inf Knowl Manage Syst*.
- Beck AR, Fu G, Cohn AG, Bennett B, Stell JG (2007) A framework for utility data integration in the UK. In: *Proceedings of UDMS Symposium*, pp 261–276.
- Bnouachir H, Chergui M, Machkour N, Zegrari M, Chakir A, Deshayes L, Semmar A, Medromi H (2020) Intelligent fleet management system for open pit mine. *Int J Adv Comput Sci Appl* 11(5):327–332
- Bragoli D, Modugno M (2017) A now-casting model for Canada: Do US variables matter? *International Journal of Forecasting* 33(4):786–800. <https://doi.org/10.1016/j.ijforecast.2017.03.002>
- Bury H, Spieckermann S, Wortmann D, Hübler F (2018) A case study on simulation of railway fleet maintenance. In: *2018 Winter Simulation Conference (WSC)*, pp 2851–2860. IEEE.
- Cen EN (2010) 13306: maintenance-maintenance terminology. Brussels.
- Dong XL, Rekatsinas T (2018) Data integration and machine learning: a natural synergy. In: *Proceedings of the 2018 international conference on management of data*, pp 1645–1650.
- Extract-, A. S. (2015). Integrating big data :
- Famurewa SM, Zhang L, Asplund M (2017) Maintenance analytics for railway infrastructure decision support. *J Qual Maint Eng* 23(3):310–325. <https://doi.org/10.1108/JQME-11-2016-0059>

- Fink O, Zio E, Weidmann U (2013) Predicting time series of railway speed restrictions with time-dependent machine learning techniques. *Expert Syst Appl* 40(15):6033–6040
- Galar D, Palo M, Van Horenbeek A, Pintelon L (2012) Integration of disparate data sources to perform maintenance prognosis and optimal decision making. *Insight-Non-Destruct Test Cond Monit* 54(8):440–445
- Grossmann G, Stumptner M, Mayer W, Barlow M (2010) A service oriented architecture for data integration in asset management. In *Engineering Asset Lifecycle Management*. Springer, London, pp 785–795
- Hwang I, Kim S, Kim Y, Seah CE (2010) A survey of fault detection, isolation, and reconfiguration methods. *IEEE Trans Control Syst Technol* 18(3):636–653. <https://doi.org/10.1109/TCST.2009.2026285>
- Hyndman RJ, Athanasopoulos G (2018) *Forecasting: principles and practice*. OTexts, Melbourne
- Isermann R (2005) *Fault-diagnosis systems: an introduction from fault detection to fault tolerance*. Springer, Heidelberg
- ISO. (2014). *ISO 55000: Asset Management*.
- Jardine AKS, Lin D, Banjevic D (2006) A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mech Syst Signal Process* 20(7):1483–1510
- Jayaratna N (1994) *Understanding and evaluating methodologies: NIMSAD, a systematic framework*. McGraw-Hill Inc, New York
- Kalathas I, Papoutsidakis M (2021) Predictive maintenance using machine learning and data mining: a pioneer method implemented to Greek railways. *Designs* 5(1):5
- Karim R, Westerberg J, Galar D, Kumar U (2016) Maintenance analytics – the new know in maintenance. *IFAC-PapersOnLine* 49(28):214–219. <https://doi.org/10.1016/j.ifacol.2016.11.037>
- Kinnunen S-K, Marttonen-Arola S, Kärri T (2016) Value of fleet information in asset management. In: 2016 Proceedings of the 6th International Conference on Maintenance Performance Measurement and Management (MPMM 2016). Luleå, Sweden. November, vol 28, pp 76–80.
- Konstantinova MV, Olentsevich AA, Konyukhov VY, Guseva EA, Olentsevich VA (2021) Automation of failure forecasting on the subsystems of the railway transport complex in order to optimize the transportation process as a whole. *IOP Conf Ser Mater Sci Eng* 1064(1):12020
- Kumari J, Karim R, Thaduri A, Castano M (2021) Augmented asset management in railways—Issues and challenges in rolling stock. *Proc Inst Mech Eng Part f: J Rail Rapid Transit*. <https://doi.org/10.1177/095440972111045782>
- Kumari J, Karim R, Karim K, Arenbro M (2022) MetaAnalyser - a concept and toolkit for enablement of digital twin. In: 14th IFAC workshop on intelligent manufacturing systems. Tel Aviv.
- Lee J (2020) *Industrial AI. Applications with sustainable performance*. Springer, Singapore
- Lingaitis LP, Lebedevas S, Liudvinavičius L (2014) Evaluation of the operational reliability and forecasting of the operating life of the power train of the freight diesel locomotive fleet. *Eksplotacija i Niezawodność* 16(1)
- Mardiasmo D, Tywoniak S, Brown K, Burgess K (2008) Asset management and governance—An analysis of fleet management process issues in an asset-intensive organization. In: 2008 First international conference on infrastructure systems and services: building networks for a brighter future (INFRA), pp 1–6. <https://doi.org/10.1109/INFRA.2008.5439593>
- Mathew A, Ma L, Hargreaves D (2009) Understanding data management in asset management: a survey. In: *Proceedings of the 3rd world congress on engineering asset management and intelligent maintenance systems conference (WCEAM-IMS 2008)*, pp 1096–1107. Springer-Verlag London Ltd.
- Meskin N, Khorasani K (2011) *Fault detection and isolation: multi-vehicle unmanned systems*. Springer, Heidelberg
- Milenkovic M, Bojovic N (2019) *Optimization models for rail car fleet management*. Elsevier, Amsterdam
- Murphy G, Chang A (2009) A capability maturity model for data acquisition and utilisation. In: *ICOMS asset management conference proceedings: sustain your business through good asset management*, pp 1–7. ICOMS.
- Orynychak M, Melnyk M, Havran V (2021) Methods for forecasting the noise level of rail vehicles. In: 2021 IEEE 26th international seminar/workshop on direct and inverse problems of electromagnetic and acoustic wave theory (DIPED), pp 253–256. IEEE.
- Penna M, Shivashankar Arjun B, Goutham KR, Madhaw LN, Sanjay, KG (2017) Smart fleet monitoring system using Internet of Things(IoT). In: 2017 2nd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT), pp 1232–1236. <https://doi.org/10.1109/RTEICT.2017.8256795>
- Rajeswari C, Sathiyabhama B, Devendiran S, Manivannan K (2014) A gear fault identification using wavelet transform, rough set based GA, ANN and C4.5 algorithm. *Proced Eng* 97:1831–1841. <https://doi.org/10.1016/J.PROENG.2014.12.337>
- Rane SB, Narvel YAM (2021) Leveraging the industry 4.0 technologies for improving agility of project procurement management processes. *Int J Syst Assur Eng Manag* 12(6):1146–1172
- Rane SB, Narvel YAM (2022) Data-driven decision making with Blockchain-IoT integrated architecture: a project resource management agility perspective of industry 4.0. *Int J Syst Assur Eng Manag* 13(2):1005–1023
- Tam ASB, Price JWH (2008) A generic asset management framework for optimising maintenance investment decision. *Prod Plan Control* 19(4):287–300. <https://doi.org/10.1080/09537280802034042>
- Thaduri A (2020) Nowcast models for train delays based on the railway network status. *Int J Syst Assur Eng Manag* 11(2):184–195
- Thaduri A, Galar D, Kumar U (2015) Railway assets: a potential domain for big data analytics. *Proced - Proced Comput Sci* 53:457–467. <https://doi.org/10.1016/j.procs.2015.07.323>
- Thomas S, Dubey A, Viassolo DE, Zanette M (2020) Digital fleet management: a scalable cloud framework based on data-driven prediction models. *Ann Conf PHM Soc* 12(1):8
- Tretten P, Karim R (2014) Enhancing the usability of maintenance data management systems. *J Q Mainten Eng* 20(3):290–303
- Wakiru J, Muchiri PN, Pintelon L, Chemweno P (2019) A cost-based failure prioritization approach for selecting maintenance strategies for thermal power plants: a case study context of developing countries. *Int J Syst Assur Eng Manag* 10(5):1369–1387
- Wang Q, Megalooikonomou V (2008) A dimensionality reduction technique for efficient time series similarity analysis. *Inf Syst* 33(1):115–132. <https://doi.org/10.1016/j.is.2007.07.002>
- Zaccaria V, Stenfelt M, Aslanidou I, Kyprianidis KG. 2018 Fleet monitoring and diagnostics framework based on digital twin of aero-engines. In: *Turbo expo: power for land, sea, and air*, vol 51128, p V006T05A021. American Society of Mechanical Engineers

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.