SUNRISE

Strategies and Technologies for **Un**ited and **R**esilient Critical **I**nfrastructures and Vital **S**ervices in Pandemic-Stricken **E**urope

D5.2 Demand Prediction and management Tool and training guide V1

Document Identification			
Status	Final	**Due Date**	26/09/2023
Version	1.0	**Submission Date**	29/09/2023

Related WP	WP5	**Document Reference**	D5.2
Related Deliverable(s)	D3.2, D5.1	**Dissemination Level (*)**	PU
Lead Participant	UPM	**Lead Author**	Javier Andión, UPM
Contributors	UPM, INT, SQD, XLB, HB, INS, IHS	**Reviewers**	Tom Sorell, UoW
			Andreja Markun, TS

Keywords:
Critical infrastructure, demand management, demand prediction, artificial intelligence

This document is issued within the frame and for the purpose of the SUNRISE project. This project has received funding from the European Union's Horizon Europe Programme under Grant Agreement No.101073821. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the European Commission.

(*) Dissemination level: **(PU)** Public, fully open, e.g. web (Deliverables flagged as public will be automatically published in CORDIS project's page). **(SEN)** Sensitive, limited under the conditions of the Grant Agreement. **(Classified EU-R)** EU RESTRICTED under the Commission Decision No2015/444. **(Classified EU-C)** EU CONFIDENTIAL under the Commission Decision No2015/444. **(Classified EU-S)** EU SECRET under the Commission Decision No2015/444.

Document Information

List of Contributors	
Name	Partner
Javier Andión	UPM
Daniel Sanz	UPM
Jose M. Del Alamo	UPM
David Rodriguez	UPM
Jakob Jenko	XLB
Dejan Štepec	XLB
Stavros Theocharis	SQD
Eftihis Thergiakis	SQD
Chirstos Vasilakis	SQD
Alessandro Calabrese	HB
Marianna Pedrazzi	HB
Tommaso Ruocco	HB
Christian Kimmich	IHS
Klaus Weyerstraß	IHS
Elisabeth Laa	IHS
Hannes Zenz	IHS
Ilias Nektarios Seitanidis	INT

Document History			
Version	Date	Change editors	Changes
0.1	19/07/2023	Javier Andión (UPM)	First draft with ToC.
0.2	24/07/2023	Jakob Jenko, Dejan Štepec (XLB)	Added demand prediction for the energy sector.
0.3	28/07/2023	Javier Andión (UPM)	Added demand prediction for the transportation sector.
0.4	03/08/2023	Stavros Theocharis (SQD)	Added demand prediction for water sector and tech parts in chapter 2
0.5	09/08/2023	Alessandro Calabrese (HB)	Added Management and What if analysis chapter 4
0.6	30/08/2023	Stavros Theocharis (SQD)	Added sections in chapters 2, 3
0.7	01/09/2023	Christian Kimmich (IHS)	Added section 4.4
0.8	02/09/2023	Ilias Nektarios Seitanidis (INT)	Added section 1.2.1 Front-end
0.9	05/09/2023	Javier Andión (UPM)	Integration of all contributions. Style unification. Conclusions, summary and introduction

0.10	07/09/2023	Christos Vasilakis, Stavros Theocharis (SQD)	Appendix I Training Guide
0.11	26/09/2023	Javier Andión, José M del Alamo (UPM)	
0.12	29/09/2023	Juan Alonso (ATS)	Quality Assessment
1.0	29/09/2023	Antonio Álvarez (ATS)	Final Version

Quality Control		
Role	Who (Partner short name)	Approval Date
Deliverable leader	Javier Andión (UPM)	26/09/2023
Quality manager	Juan Alonso (ATS)	29/09/2023
Project Coordinator	Antonio Álvarez (ATS)	29/09/2023

Table of Contents

SUNRISE

List of Tables

List of Figures

List of Acronyms

Abbreviation / acronym	Description
AI	Artificial Intelligence
ARIMA	Autoregressive integrated moving average
CI	Critical Infrastructure
DPM	Demand Prediction and Management
D5.2	Deliverable number 2 belonging to WP5
EC	European Commission
FR	Functional Requirement
GUI	Graphical User Interface
HPA	Horizontal Pod Autoscaler
IO	Inflow-Outflow
LP	Load profile
LSTM	Long short-term memory
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ML	Machine Learning
MLP	Multi-layer perceptron
MSE	Mean square error
NFR	Non-Functional Requirement
OD	Origin-Destination
PoC	Proof of Concept
RMSE	Root mean square error
RNN	Recurrent neural network
SLP	Standard Load Profile
SMAPE	Symmetric mean absolute percentage error
SSO	Single Sign-On
SotA	State of the Art
TRL	Technology Readiness Level
WP	Work Package

Executive Summary

This document describes the initial version of the Demand Prediction and Management Tool within the SUNRISE project. The Tool boasts a versatile graphical user interface and a service-based back-end engine that can be easily customized to accommodate the unique requirements of critical infrastructures in the different domains represented in the project, including energy, transport, health, and water. The general architecture and main technologies used to develop the Tool are described along with a guideline of the navigation through the first version of the graphical user interface. A training guide and user manual are also included in one of the appendices to better describe the user interface aspects.

For each domain a proof of concept has validated the feasibility and performance of the prediction models elaborated. Indeed, different prediction models have been trained on the historical data provided by the critical infrastructure providers participating in the work package, selecting those models that have demonstrated better performance, considering both their prediction ability and their computation time. The prediction models are complemented with a comprehensive "what-if" analyses, exploring the potential impacts of pandemic and climate change scenarios on critical infrastructures management.

These proofs of concept that have been tested in a lab environment, have supported achieving TRL5 for the Demand Prediction and Management Tool, thus contributing to the project milestone MS4 *Initial TOOLs [TRL5] and STRATEGY [TRL6].* Those proofs of concept are the Phase 0 of the pilot plans of the Demand Prediction and Management Tool presented on D5.1 Demand prediction and management conceptualization. Moving forward, the work package is well-prepared to transition into the next development phase, building upon the achievements and insights gained during this initial iteration.

1 Introduction

1.1 Purpose of the document

This report details the first version of the Demand Prediction and Management (DPM) Tool, including its overall design, deployment strategy, and verification and validation plan, as well as its different components with their initial validation results. It also describes the Proof of Concept (PoC) conducted to demonstrate the Tool's feasibility in the various domains.

1.2 Relation to other project work

While D5.1 Demand prediction and management conceptualization [1] focused on the DPM Tool conceptualization, the focus here is on demonstrating its feasibility in different domains through a set of PoCs to reach the TRL5 [1]. Therefore, in combination with the updated requirements and designs reported in the second version of the deliverable D3.1 Requirements and designs [2], and similar deliverables (DX.2) from the other technical Work Packages (WP4-WP7), D5.2 contributes to reaching the SUNRISE MS4 milestone: Initial Tools and strategy.

The results described in this document, particularly the validation results, as well as the updates from WP1-WP3 will guide the upgrade of the DPM Tool and its components in the forthcoming months. According to the SUNRISE work plan, the individual components will be upgraded by M16 and integrated by M20, which will be reported in the second version of this deliverable D5.2 Demand prediction and management Tool and training guide. By M23, the DPM Tool will be deployed in the environment of the Critical Infrastructure (CI) operators participating in WP5 to demonstrate its basic functionalities and bring it to TRL6. The results will be reported in the deliverable D5.3 Demand prediction and management pilot report by M23.

1.3 Structure of the document

The document is structured to deliver a presentation and detailed description of the Demand Prediction and Management Tool, and the first validation results.

Chapter 2 documents the most relevant decisions taken on the Tool development. It starts with an examination of the Tool's software architecture, describing its front-end and back-end components. It follows with the adopted deployment strategy and ends with the Tool's verification and validation.

Chapter 3 focuses on the demand prediction models elaborated for different domains. Each domain is analysed for its unique characteristics, and the prediction models together with their evaluation results are presented.

Chapter 4 presents the results of a What-If analysis that complements the prediction models to evaluate the impact of pandemics and climate change on the CIs.

Chapter 5 describes the status of the pilot trials plan, whose first phase of pilot trials has been already carried out.

Chapter 6 concludes the document with some reflexions on the work ahead.

Appendix I includes the first version of the training guide and user manual.

1.4 Glossary adopted in this document

Critical infrastructure (CI): Power distribution networks, transportation networks, and information and communication systems are all examples of critical infrastructure. The defense of these critical assets is very necessary to ensure the safety of the European Union and the wellbeing of its people. The electricity grid, the transportation network, and information and communication networks are examples of the so-called "critical infrastructures" that must be preserved if crucial society functions are to be kept operating normally. Natural disasters, acts of terrorism, and criminal activities all have the potential to cause damage to or destroy essential infrastructure, which may have serious repercussions for both the safety of EU residents and the complete EU.

Critical assets (CAs): are the significant resources that support both the social and business parts of an economy. If some of these assets fail, it will bring significant issues for business continuity. This does not mean that the likelihood of failing is high. For planning purposes, each business or organization must identify its critical assets and know the corresponding information about them.

2 The Demand Prediction and Management Tool

The Demand Prediction and Management (DPM) Tool, based on state-of-the-art technologies regarding demand prediction methods, like AI/ML-based models, has the general scope to provide accurate demand prediction for specific Critical Infrastructures and furthermore to efficiently facilitate resource management. More specifically, the DPM Tool aims to provide accurate predictions for the CI operators, in terms of their critical resources, mostly in a short-term period. Mid- and long-term predictions have also been examined and applied where applicable. In the extension of that, CIs can enhance the quality of their current predictions and their resource planning, making them also capable of more effective management of financial resources and able to address potential demand changes better.

An example that showcases a basic situation in which the tool can be used is showcased in the next few sentences. Supposing a scenario for a specific city, some decision-makers need to know the forecasts of the water consumption during the next period (some days ahead) in order to be prepared. The tool with the current models can provide such forecasts based on the historical data that they have been trained on. As next steps, simulations of how pandemics or other climate changes affect the forecasts are going to be researched. In addition, anomalies in the expected outcomes are also going to be experimented with and modelled.

2.1 General Context

The DPM Tool incorporates various approaches to achieve better forecasting and demand prediction. AI/ML-based models are applied to the historical data of the CIs, which are formed into a specialized type of dataset called time series. Through these methods, the Tool provides short-term forecasts of the different CIs resources. Additionally, a graph-based approach is used for the transportation sector to understand the flow (origin-destination) of passengers. On top of the above pandemic-related effects will also be considered to identify the possible demand changes in such periods.

For the Demand Prediction and Management Tool, a literature review was and will continue to be executed during the development of the Tool, to identify any new techniques from related works. Through the Tool, all the derived insights, suggestions and predictions will be shown in a user-friendly way through dashboards.

Work Package 5 includes different pilot countries (Spain, Slovenia and Italy). Each pilot country consists of different Critical Infrastructures in diverse sectors like energy, water, health, transportation and digital services. The Tool will be deployed for each one of the Critical infrastructures with personalized visuals. Furthermore, more services, like the “what-if analysis”, climate service, and socioeconomic service, will be added supplementary where it is applicable. To create a resilient system that complies with the requirements, there are specific needs for each sector as mentioned in D3.2 and D5.1.

2.2 Architecture

To reach the current structure of the architecture of the Tool, it was needed continuous cooperation between different partners and CIs of the Work Package. In that way it was easier to understand the specific needs of each partner and express them through the design of the Tool.

2.2.1 Front end

2.2.1.1 Introduction

Under the WP5 framework several modules will be developed for the various sectors. The system’s users will be able to interact with those modules through a shared Graphical User Interface (GUI). The following section’s purpose is to introduce an interface that will provide to the users a flowless experience, without confusing them, by exploiting UI/UX best practices. Based on these practices used by the industry for a variety of reasons, the back-end and the front-end should be separated. In our case this principle is adopted in order to

- develop in parallel the subsystem components.
- use the state-of-the-art technologies in each case, back-/front-end, without any restriction.
- re-usability, the front-end is not bounded to a back-end.
- deployment, in the case of a BE instance failure the FE will not fail as well.

The main issue that may arise for the UI is the facilitation of different modules under a shared GUI. For this reason, the creation of a Request for Comments (RFC) document that will describe the designs and functions implemented for the BE-FE communication is important. In this context, the data payload and format have to be agreed by all the involved parties. The enabling technology for the back-end front-end communication is through the widely used REST API communication. This approach offers an easy way to integrate different systems while maintaining a wide range of functionalities. In Figure 1 the main technologies used for the realization of the GUI are presented.

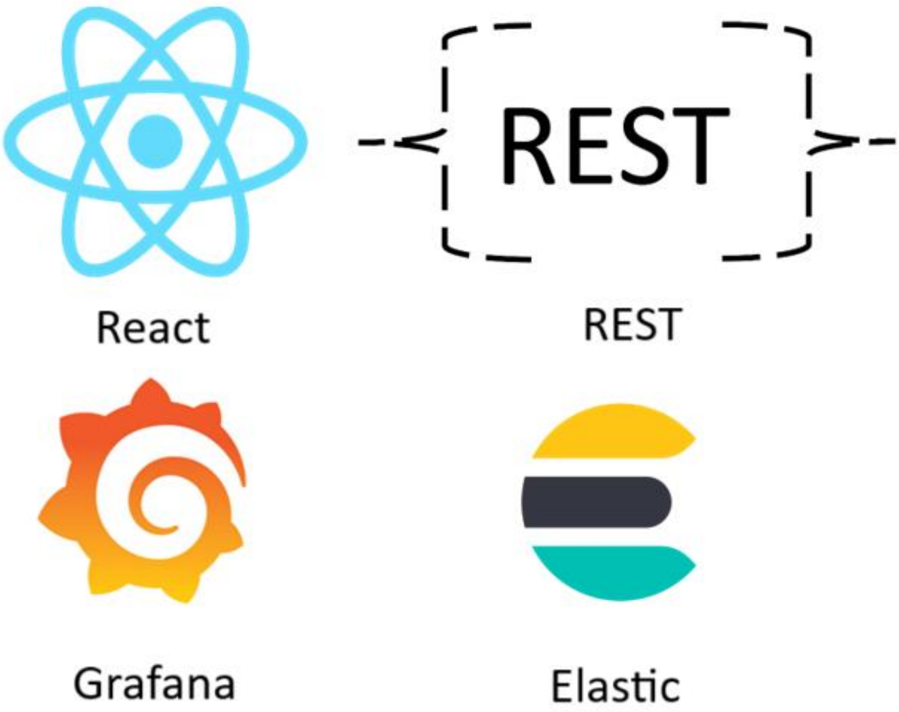

Figure 1. Main technologies used for the GUI.

2.2.1.2 Views

In this Section the initial Views of the GUI application will be presented. These views were produced based on the best practices of the UI/UX and the Tool's requirements as defined in the deliverable D3.1. It is foreseen that minor adjustments may apply as a result of the ongoing bilateral communication with the CIs in order to tailor the currently developed GUI for each specific involved CI.

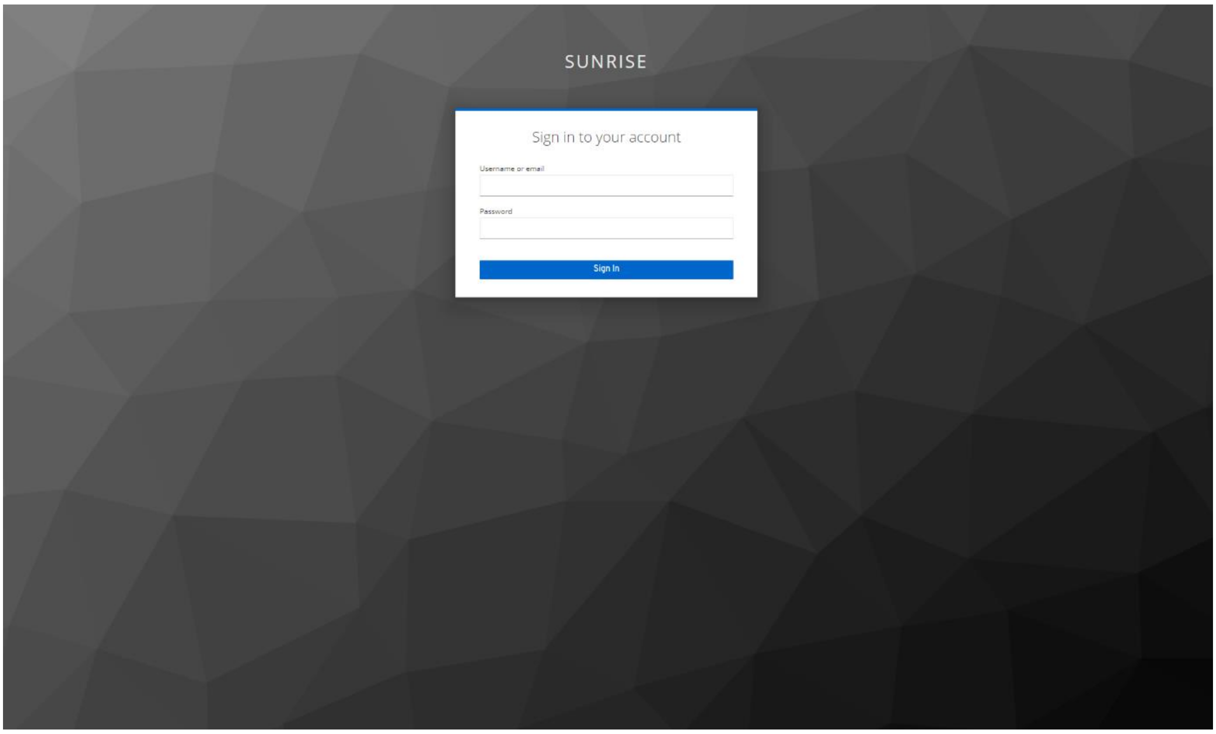

Figure 2. Login Page of the system.

In Figure 2, the landing page of the system is presented. This is the initial View that the users interact with. Upon a successful login the user is redirected to the main View of the application. The user authentication takes place by consuming a back-end resource which is responsible for the authentication (simple user/password login, SSO, etc.). After a successful user authentication, the user is landed on the Dashboard View. The purpose of this View is to facilitate all the required graphics (diagrams, charts, gauges, etc.) that will provide a clear view of the system's functions Figure 3 and Figure 4 demonstrate the various configurations of the graphics elements in the space. Another View origin from the Tool's requirements as defined in D3.1 is the Resource View. This View is responsible for demonstrating the latest used parameters for training and prediction. The user selects the desired functionality from the dropdown list (training or prediction) and the tableview is filled with the relevant variables. The user may alter the loaded values and upon pressing the button "Submit" the data values are sent to the back-end system.

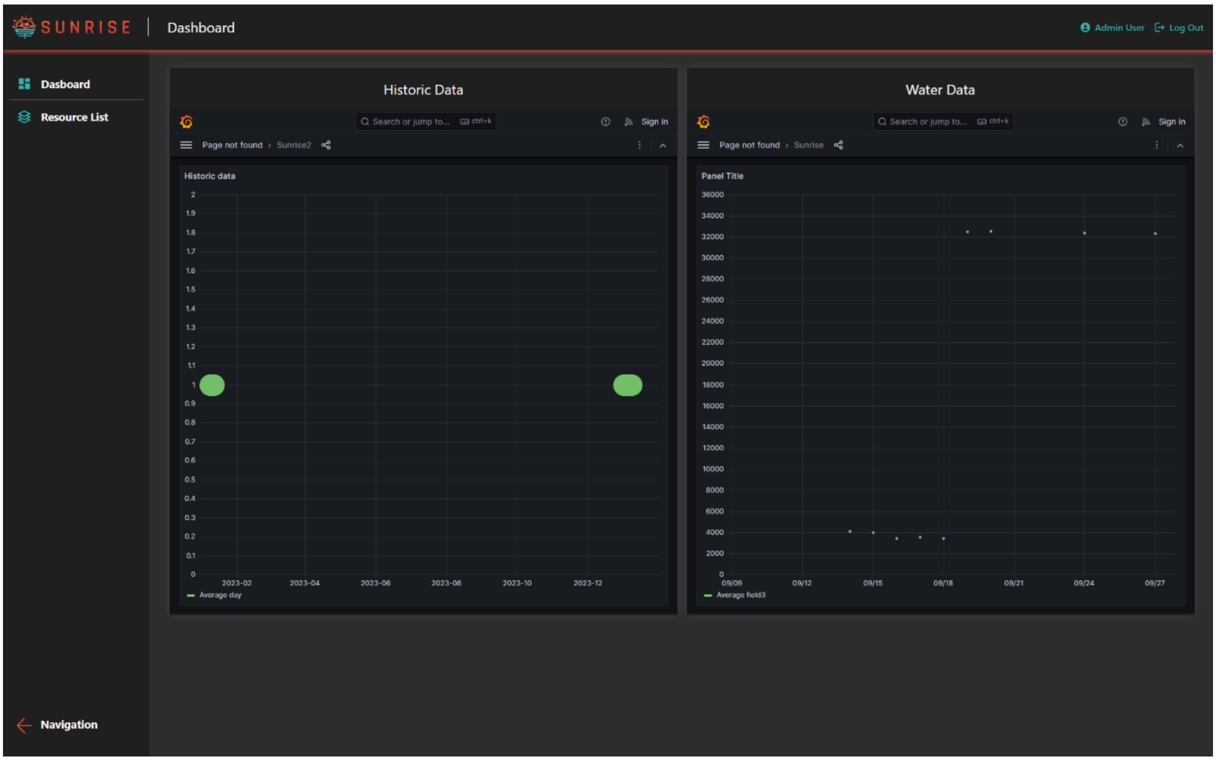

Figure 3. Dashboard View #1

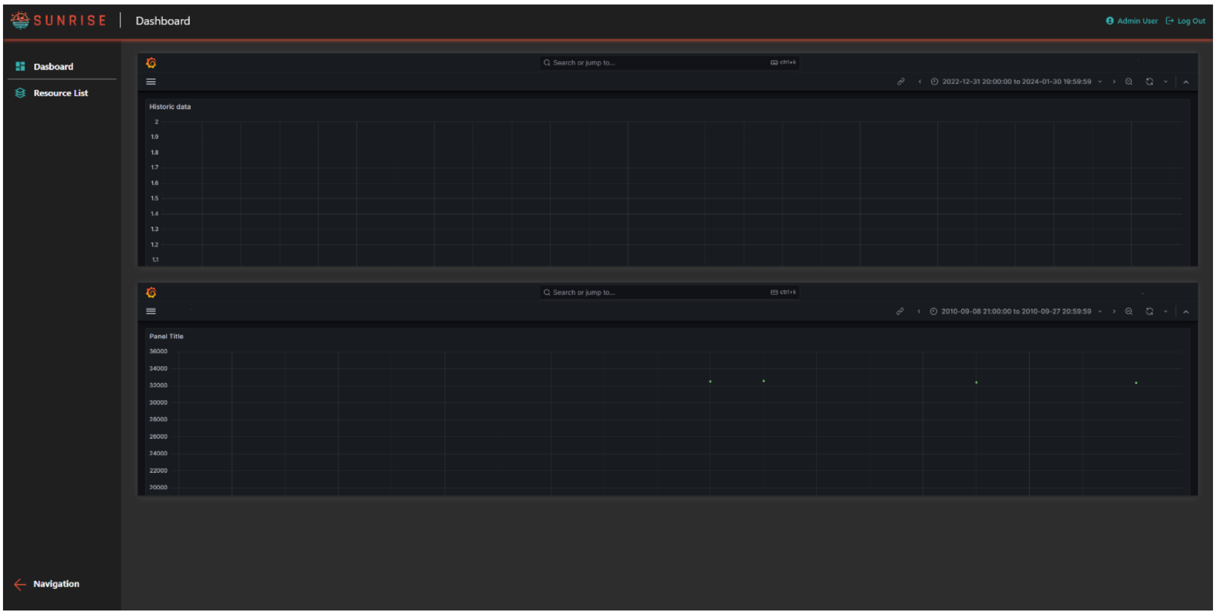

Figure 4. Dashboard View #2

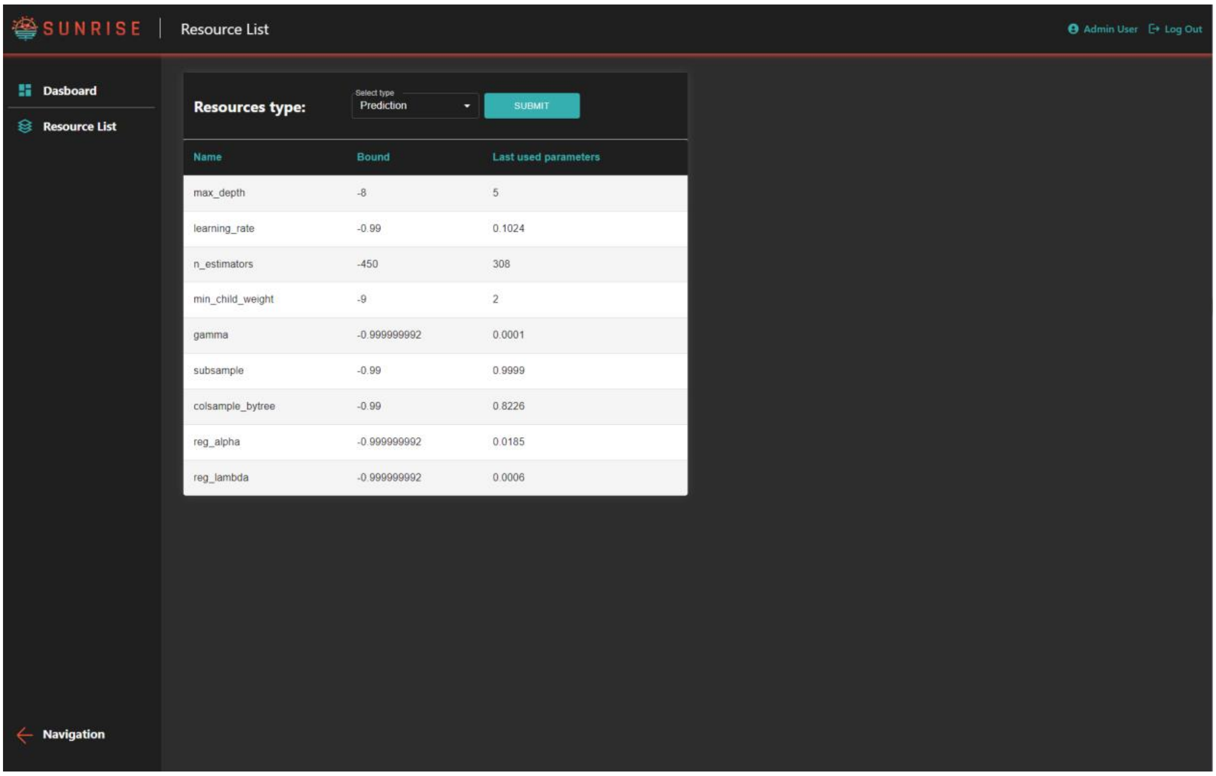

Figure 5. Resource View

2.2.1.3 Connection with back-end

As already stated, the front-end is separated from the back-end while the back-end front-end communication takes place through REST API endpoints. Below we describe the endpoints required for each of the Views presented in Section 2.2.1.2.

Login screen: The front-end sends the credentials from the User input to an Authentication Endpoint. This endpoint is responsible for authenticating that the user has access to the system. The most usual implementation of this kind of endpoint is using a Single Sign-On (SSO) mechanism like auth0 [4]. The response that the front-end awaits is a token, encrypted message, which contains all the necessary information regarding the user's permission to the system. Upon reception the FE checks the permissions and either proceeds to the next view or displays a "User's credentials error" message.

Resource Management: The front-end on the loading phase calls the relevant API endpoint that returns the last parameters used for the last model training/prediction. This endpoint is called using the GET HTTP verb. The values received are displayed at the tableview element of the page. The data of the tableview upon activating the "Send" action are sent to the back-end using a PUT verb to update those parameters. Based on the back-end's structure this functionality can be implemented in a single call API endpoint, with the execution of the PUT request for data transmission the retraining process may start. On the other hand, a second API call that will start the retraining/prediction process may be used.

Dashboards: The creation and maintenance of dashboards is a complicated procedure. In order to tackle this issue a unified mechanism is considered. The front-end is not directly connected with any of the partner's back-end infrastructure but is bound to a caching mechanism. The back-ends can send their data periodically through one of the various interfaces of the caching system or through a broker mechanism like Apache Kafka.

File Uploader: This view's main purpose is to upload binary data files to the partner's storage space. This will be achieved by sending the consumed data through a POST request as we create a new resource. Based on the back-end's functionality, a GET request may take place in order to retrieve and display the already stored data files.

2.2.2 Back end

2.2.2.1 Introduction

Each one of the sectors under WP5 can be identified as a different service in the whole architecture. All these services include all the corresponding pipelines implemented by the contributions of different partners under WP5., i.e., some main services are the energy service, the water service, the transportation service, the health service. There are also other support services such as the "what if analysis" service or the climate service.

In order to have a robust integration between the different services, a specific and determined set of rules has to be followed. These rules will also help with an easy customization of the final demand management and prediction tool for each one of the involved CIs. The integration approach will be programming language agnostic. In order to be able to continue with the selected approach a communication of the infrastructure needs has to be set as next steps with the involved CIs.

For the backend part a REST API approach has been followed using modern web frameworks with high-performance capabilities.

Some of the features of the Sunrise Tool include secure authentication to prevent unauthorized use of the Tool, data validations to ensure the validity and integrity of information and easy-to-use data management functionalities. Additionally, users will have the ability to gain insights from predictive model analytics offering valuable demand forecasts and the option for model retraining to continuously enhance predictive accuracy.

2.2.2.2 Connections and endpoints

The full stack application of Sunrise is made up of several parts as listed below:

- Authentication Endpoints: Secure endpoints will be provided to authenticate users and prevent unauthorized use of the Tool.
- Data Validations: Various data validations will be implemented to aim the accuracy and integrity of the information.
- Data Management: The application will facilitate easy data management for the users to manage their data.
- Demand Forecasts (inference): The service can create forecasts based on the trained model. Users will have access to predictive analytics and demand forecasts to support decision-making.
- Historical timeline: The service can provide the historical values.
- Model Retraining: Occasional model retraining will be provided to continuously enhance predictive accuracy.
- Evaluation of a model: The service can calculate the evaluation metrics of the model it provides.
- Other functionality: Based on the sector other custom functionalities may be included.

All communications are done through HTTP calls to provide a UI framework agnostic functionality and in addition the API will be async and stateless to avoid keeping HTTP communications for long periods of time and provides better possible future scalability. All responses are in JSON format.

2.2.2.2.1 Example of API for uploading data and retraining a model

- Uploading Data:

When clients need to upload new data for the model to be trained, they can send a POST request to the appropriate API endpoint. This will be done through the platform UI for easier use. The backend servers will validate the data format and accuracy and will respond with a JSON response. After processing the data, the server will save them to a database for later use. In case of an error the server will respond with a JSON response and a useful error message.

- Model Retraining Pipeline:

When clients need to retrain the ML model and in scenarios that this is applicable, they need to send a POST request to another API endpoint. For example: */retrain_model*. The API server receives a retraining request and initiates the model retraining pipeline. This pipeline might involve fetching the uploaded data from the database, preprocessing it and then retraining the model based on the new data. During retraining the API server may run the training process in the background asynchronously, to avoid blocking other API requests. After the model retraining is complete the server stores the updated model or model parameters for future predictions.

- Responses in JSON Format:

Throughout the API flow all responses to client requests are formatted in JSON for consistency and ease of parsing. In case of errors or exceptions error responses are also provided in JSON format, including an error status code and a meaningful error message to assist in debugging.

- Model Prediction Endpoint:

After the model has been trained it is ready to provide predictions. The clients will send a POST request again through the UI platform and the server will initialise the prediction pipeline. The prediction pipeline might consist of fetching the new data from the database or proceed with the already trained model. The request payload should specify and narrow down as much as possible what they want to predict to make performance more efficient and quicker.

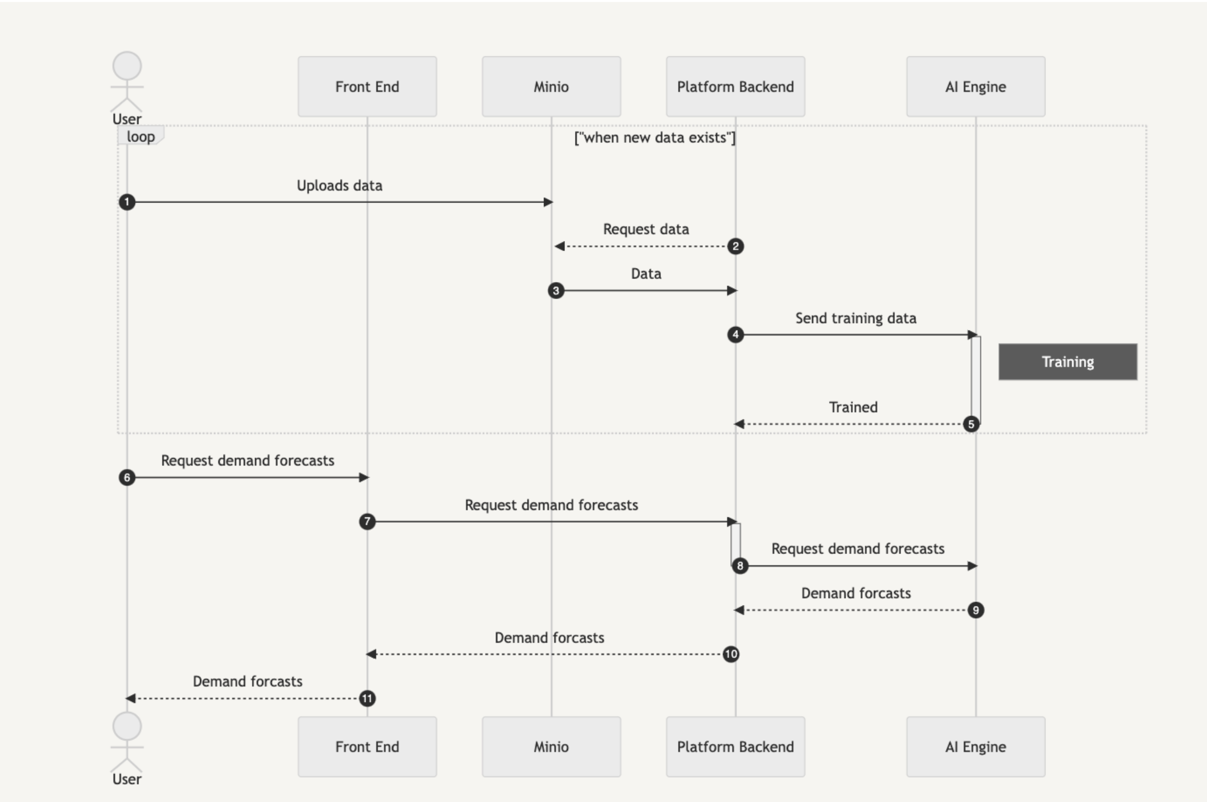

Figure 6: Back-end architecture diagram.

2.2.3 Deployment

Docker will be the tool of choice to scale the app efficiently. The way Docker works in that it allows for services and their dependences to be containerised hence avoid external dependences and achieve independence of the app. By containerisation Docker achieves an environment-agnostic performance and allows the app to run in any system regardless of its underlying environment. Scaling with Docker is done by replicating the service into multiple copies hence a higher load can be handled simultaneously. Docker also allows for horizontal scaling capabilities, meaning that developers can deploy and manage a cluster of containers that are eventually distributing the app's workload across these instances. This feature ensures a robust foundation is in place in order to handle higher loads and further manage with Kubernetes.

Kubernetes is built upon the containerisation Docker provides and offers advanced orchestration and management capabilities. Kubernetes offers an infrastructure for distributed and resilient manner to deploy and run Docker images. That infrastructure allows powerful auto-scaling features such as Horizontal Pod Autoscaler (HPA), meaning that it can detect increased incoming load hence automatically scale up the number of replicas (instances) of the app's pods to maintain performance. Similarly, during low-traffic periods Kubernetes can scale down the number of replicas to conserve resources. In addition, Kubernetes enables seamless rolling updates, making it possible to update the app to a new version without any downtime and thereby ensuring continuous availability during the scaling process. Another feature of Kubernetes is service discovery, load balancing and container health monitoring, all of which contribute to the app's efficient scaling and high availability. All the above are going to be applied on the corresponding premises of the CIs.

2.3 Tool validation and verification

In order to validate our Tool, we are going to track the intended requirements and check if the Tool fulfils them. One of the main factors is the accurate prediction of the Tool. So, based on the final evaluation results of the corresponding models, we will be able to understand the degree of accuracy of the outputs. The data sources have to be continuously validated because the quality of the data plays a significant role in training accurate models that produce accurate results. Some first iterations have been produced that give us high confidence in achieving appropriate results.

Most of the models that will be chosen have already been mentioned in the literature, and they seem to match the corresponding domains. Further experiments need to be done in order to have more choices between the possible implemented models and optimise them accordingly.

After having the first iteration of the Tool ready, several code reviews and quality assurance processes will take place in order to ensure that the coding standards are being followed. When the Tool is deployed on CIs premises, we will be able to get direct feedback from them about whether the Tool covers their needs and helps them with their decision-making processes by providing them with accurate insights.

3 Demand Prediction Methods & Validation

In this section the different domains are referenced with their corresponding details and progress over data exploration, experiments of modelling algorithms and next steps.

3.1 Energy domain

3.1.1 Introduction

Accurate prediction of energy demand is an important aspect of maintaining a stable energy grid. As the percentage of renewables in the energy grid continues to increase, the ability to accurately forecast the demand is crucial to minimise use of peaker plants which use fossil fuels to generate electricity and to reduce dependence on buying the energy on the costly spot markets. The demand follows a profile known as Standard Load Profile (SLP), as we can observe in Figure 7. The fact that the consumption follows a pattern, makes it possible to make accurate predictions. In simple terms, goal of the forecasting method is to match shape of it, best as possible. Events such as weather, holidays and other events have a big impact on the shape of the consumption curve. These factors induce additional uncertainty into predictions, and it's important to include them as an input to our model. The prediction uncertainty also changes when we predict on national level, the more we move toward more granular levels such as households, the harder it is to predict the demand accurately. This phenomenon is known as aggregation effect, and it makes it easier to predict consumption on higher levels such as national level.

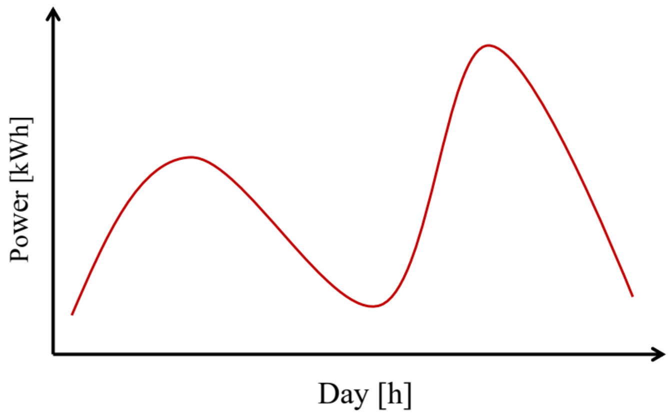

Figure 7: Standard load profile

The impact on the Load Profile (LP) can be divided into two parts. First are socio-economic factors such as holidays, major events and even pandemics. A good thing is that such events are usually predictable and are known in the future. In case of pandemics, restrictions are usually known for day-ahead forecasting only. For example, COVID-19 drastically reshaped the load profile curve, in some cases the morning peak was higher than evening peak.

The second part of the impact are weather related events such as storms, heatwaves, lighting storms. Exact weather conditions are generally not known in the future and are hard to accurately predict. However, state-of-the-art (SotA) weather models perform well for day-ahead forecasts, which can be used in demand prediction models.

To increase the modality of our model and improve predictions, we must include features carrying important information. In the case of pandemics, we will focus on incorporating SUNRISE data to increase modality. To achieve this, we will utilise mobility data obtained from Slovenian national telecommunication provider (Telekom Slovenije - TS). Mobility is measured by aggregating number of unique connections to each mobile cell (telecommunication grid is usually divided in grids of cells) or transmission tower.

Before this information is used for demand prediction in Slovenia, we will first evaluate the performance of timeseries forecasting models on public datasets. We will conduct studies on both classic and novel deep learning models to determine the best fit for our specific SUNRISE use-case.

The reason behind conducting this evaluation is that number of studies compared performance of novel transformer-based models with traditional models. For example, study [5] showed that simple one-layer linear models outperformed modern transformer-based models in long timeseries forecasting. On the other hand, publication [6] showed that transformers are effective for short timeseries forecasting. The main difference between the two papers is that the first evaluated on long timeseries (96+ hours), whereas the second focused on shorter periods (up to 24h). The publication [6], did not address the results for short-timeseries' forecasting, also known as day-ahead forecasting. This shows that it is not necessarily true that the latest models are the best fit. To find out, we will employ a variety of models and select the best fit for our use-case.

3.1.2 Methodology

3.1.2.1 Datasets

Datasets were described in depth in SUNRISE D5.1 (Section 3.4.2). While we have national level consumption for Slovenia and Serbia at our disposal, we will first focus on publicly available datasets described in section 3.4.2.2 of D5.1. Here, we will make a quick overview of the public datasets.

A detailed overview of the UCL dataset can be observed in Table 1. Where we can observe that all covariates, with exception of power usage were derived from time variable. By introducing such covariates, we provide additional insights on factors that influence power consumption patterns.

Table 1: dataset description for 370 power meters

	Power usage	t	Days From start	hour	day	Day Of week	month	Hours From start	Categorical Day of week	Categorical hour
count	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M	21.89 M
mean	616.45	29305.62	1220.59	11.50	15.43	3.01	4.65	29305.62	3.01	11.50
std	3625.14	1720.38	71.68	6.92	8.89	1.99	2.36	1720.38	1.99	6.92
min	0.00	26304.00	1096.00	0.00	1.00	0.00	1.00	26304.00	0.00	0.00
25%	51.51	27821.00	1159.00	6.00	8.00	1.00	3.00	27821.00	1.00	6.00
50%	120.81	29308.00	1221.00	12.00	15.00	3.00	5.00	29308.00	3.00	12.00
75%	307.46	30796.00	1283.00	18.00	23.00	5.00	7.00	30796.00	5.00	18.00
max	168100.00	32279.00	1344.00	23.00	31.00	6.00	9.00	32279.00	6.00	23.00

We will use all 370 datasets power meters. Overall, some of the power meters have up to 2 years of data, to make sure there is no missing data for any of the power meters we have selected a smaller interval seen in Table 2: Dataset train and test splits., which should suffice for means of validating and comparing the models. In Table 2: Dataset train and test splits. we can observe detailed description of train and test splits.

Table 2: Dataset train and test splits.

Meters = 370	N samples	from	to
train	1.8 M	01.01.2014	31.07.2014
test	0.3 M	01.08.2014	24.09.2014

Furthermore, we used a set of original datasets, with power meters measuring actual consumption. In the full dataset some metering instances have very little to no entropy in the given signal. Low entropy signals containing no consumption patterns, makes it hard for the model make predictions. To work with high entropy signals, containing actual consumption patterns, we filtered meters of which power consumption did not go over 500 W. This left us with 111 samples. While we could use more complex entropy filters, this simple filter turns out to be sufficient.

Table 3: Reduced dataset description for 111 power meters

	Power usage	t	Days From start	hour	day	Day Of week	month	Hours From start	Categorical Day of week
count	658584	658584	658584	658584	658584	658584	658584	658584	658584
mean	1822.79	29295.92	1220.18	11.50	15.41	3.01	4.64	29295.92	3.01
std	6448.97	1726.62	71.94	6.92	8.89	1.99	2.37	1726.62	1.99
min	0.00	26304.00	1096.00	0.00	1.00	0.00	1.00	26304.00	0.00
25%	335.70	27800.00	1158.00	5.75	8.00	1.00	3.00	27800.00	1.00
50%	595.99	29297.00	1220.00	11.50	15.00	3.00	5.00	29297.00	3.00
75%	1205.13	30794.00	1283.00	17.25	23.00	5.00	7.00	30794.00	5.00
max	168100.00	32279.00	1344.00	23.00	31.00	6.00	9.00	32279.00	6.00

Similarly, as before, in Table 4 we can observe that while train test time range is unchanged, number of samples decreased as we use less power metering instances.

Table 4: Reduced dataset train test split.

Meters = 111	N samples	from	to
train	98.48 K	01.01.2014	31.07.2014
test	16.41 K	01.08.2014	24.09.2014

3.1.2.2 Prediction models

The models that will be compared were described in depth in SUNRISE D5.1 (Section 3.3.2.3). In this study we will utilise models describe there - Arima, XGBoost, NHiTS, and TFT. Additionally, we will also utilise a simple linear regression as it is model that was compared against in the studies mentioned above. When fitting the models all will be given context length (window size) of 1 week and predict demand for the next 24 hours.

In this study we will use implementation of ARIMA, XGBoost and linear regression, TFT and NHITS in DARTS [7]. Additionally, we will use alternative implementation of NHITS and TFT in PyTorch forecasting [8].

3.1.2.2.2 Linear regression

Since linear regression was not described in D5.1, it will be described here. Linear regression works on basis of fitting following coefficients.

$$Y(t) = \beta_0 + \beta_1 \mathrm{X}(\mathrm{t}-1) + \beta_2 X(t-2) + \ldots + \beta_n * X(t-n)$$

Where:

- Y(t) represents the dependent variable at time t (the value we want to predict).
- X(t-n) represent the lagged values of the dependent variable at previous time points.
- β_n are the regression coefficients associated with each lagged variable, which determine the slope and height of the transform.

We will be performing a forecast for the next 24 hours which yields in total 4056 parameters to tune when performing gradient descent. Keeping in mind that we are working with 370 measurement stations which increases the number of total parameters to over 1.5M. Each parameter is written as 16-bit integer. In total the model is roughly 3 GB in size, which is more than complex recurrent neural networks (RNN) models. The size and time to train are parameters we will include in the study as they are important when developing new models with additional data. While an autoregressive approach, where we would have one model instead of 24, would reduce model size by 24, we must be consistent when building models, as others would be non-autoregressive models.

3.1.2.2.3 NHITS

The NHiTS [9] model employs multi signal sampling, which resamples input data to different frequencies to effectively capture temporal patterns. The multi-layer perceptron (MLP) stacks weights get adjusted to minimise the loss during training. The model aggregates different stack outputs to generate the final output. Contribution of each stack is also adjusted during training procedure. Each stack captures different aspect of a consumption pattern, such as trends, seasonality, and different small daily variations.

In Table 5 we can observe hyperparameters, we used for PyTorch forecasting implementation.

Table 5: Hyperparameters for NHiTS (PyTorch forecasting implementation).

Hyperparameter	Value
Context Length	168
Forecast Length	24
Hidden Size	441
Static Hidden size	100
Dropout	0.0213
Loss	Quantile loss
Down sample Frequencies	24, 8, 1
Number of Blocks	1, 1, 1
Number of Layers	2, 2, 2
Pooling Sizes	16, 4, 1

3.1.2.2.4 TFT

TFT [10] was already described in D5.1, here we will provide a short abstract. TFT, also known as temporal fusion transformer, uses past as well as future covariates to make prediction. In the training process all features go through variable selection network where their contribution is selected based on their importance. Next, long short-term memory network (LSTM) is used to make sense of time relations. Here past as well as future covariates are used. Static enrichment layer is useful when we are working with many classes, in our case 370 different power meters, enabling us to use a single model.

When comparing NHITS and TFT enables us to include future data to help the model make better predictions. In this case future variables will be date related and will be, day of the week, month, and time since start. In Table 6 we can observe hyperparameters, we used for PyTorch forecasting implementation.

Table 6: Hyperparameters for TFT (PyTorch forecasting implementation).

Hyperparameter	Value
Context Length	168
Forecast Length	24
Hidden Size	160
Hidden Continuous Size	6
Dropout	0.0210
Loss	Quantile loss
LSTM layers	1
Gradient Clip Val	5
Attention Head Size	21
Gradient Clip Val	0.01575

3.1.2.2.5 Comparison of different prediction models

Table 7: Table of model characteristics. Source: https://github.com/unit8co/darts#forecasting-models

Model	Univariate	Multivariate	Past-observed covariates	Future-known covariates
ARIMA	X			X
Linear Regression	X		X	X
XGBoost	X	X	X	X
NHiTS	X	X	X	
Transformer	X	X	X	X

3.1.2.3 Performance metrics

To measure the performance and the accuracy of forecasts we will utilise mean absolute error (MAE) and symmetric mean absolute percentage error (SMAPE). As we are working with different power meters measuring signals at different amplitudes, we must standardise contribution of each power meter accordingly. To achieve this, we normalised the signals using z-transform. In this case we wanted to know the MAE for each individual meter

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\text{z-score}(A_i) - \text{z-score}(F_i)|$$

The second metric we used was scaled mean absolute percentage error. In this case we do not perform a z-score normalisation but rather evaluate the performance on original signal

$$\text{SMAPE} = \frac{1}{\text{n}}\sum_{\text{i}=1}^{\text{n}} \frac{|\text{A}_\text{i} - F_i|}{(|A_i| + |F_i|)/2} \times 100$$

3.1.2.4 Evaluation

Results will be evaluated in two experimental pipelines. The first one will be evaluated on models implemented in DARTS. The hyperparameters were set to the same, except where explicitly specified otherwise.

3.1.3 Results

3.1.3.1 Results using DARTS implementation

Results were populated in following Table 8

Table 8: Results for all models (DARTS implementation)

Model	MAE	SMAPE	Training time (roughly)
ARIMA	0.824	31.71	48 H +
LINEAR REGRESSION	0.423	19.788	48 H +
XGBOOST	0.379	16.934	48 H +
NHITS	**0.352**	**15.99**	**1 H**
TFT[1] (smaller model)	0.45	19.25	5 H

NHiTS is a newer model, and it shows surprising performance compared to other models. While linear regression does not offer as much finetuning as for example XGBoost, where we could fine tune optimal tree depth to increase performance and reduce resource consumption. It was clear that NHiTS was better without a significant finetuning. This confirms the researc done by the authors.

While all models could converge faster, it is obvious that NHiTS is more compact compared to other models. Knowing the NHiTS outperformed SotA models we can still evaluate the TFT by extensively comparing it to NHiTS alone.

3.1.3.2 Results for PyTorch-forecasting implementation

To evaluate if TFT can compete with classical models we will evaluate it against NHiTS only. We noticed that full dataset exhibited issues. We have cleaned the dataset from meters that do not include a significant consumption pattern, which should let the models focus on actual patterns. Furthermore, we also used a different implementation of the two models. This time we are utilising PyTorch forecasting implementation of the two. In case of NHiTS, we kept the same hyperparameters. In this case we performed an extensive hyper parameter optimisation searching over all given hyperparameters. The optimal hyperparameters are the ones we mentioned in Table 5 and Table 6.

First thing we notice when comparing Table 8 and Table 9. This is because the two implementations are slightly different, and they treat classes differently. The PyTorch forecasting implementation seems to make more sense and is more efficient. The fact that we used less classes should not make a big impact on these two models.

Whereas we cannot directly and fairly compare the performances between the two implementations, we can compare the performance within the same implementation, as the whole data loading and evaluation pipeline is identical.

Table 9: Results for reduced dataset (PyTorch forecasting).

	MAE	sMAPE	Training time
NHITS	0.22	9.04	**8 minutes (35 batches/s)**
TFT	**0.1841**	**7.82**	4H (5.59 batches/s)

[1] TFT model used here was smaller than need due to implementation issues.

The first thing we noticed is that TFT slightly outperformed NHiTS. On the other hand, NHiTS converged to a solution a lot faster, and the model size was 6 times smaller. Even though the TFT is better, NHiTS is making the fast inference time competitive in this scenario. To further analyse the performance between the two, let's look at some actual forecasting examples.

3.1.4 Forecasting examples (qualitative analysis)

Examples show graphs of observed demand in blue and predicted in orange. More specifically we are plotting z-normalized consumption. In each example we plot NHiTS on the left and TFT on the right.

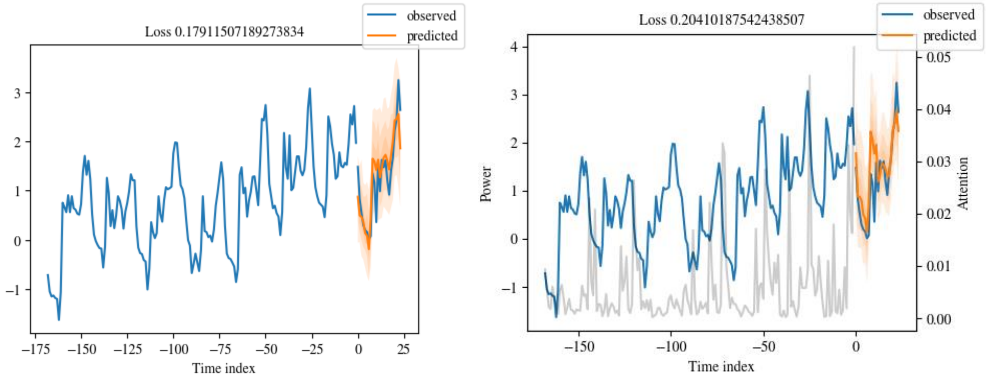

Figure 8: NHITS and TFT example 1.

When observing Figure 8 above, we can see that the right (TFT) figure includes attention for a specific time index. This specific example contains a consumption pattern that is hard to distinguish from noise. Non the less both models yielded accurate predictions. It seems as if TFT put attention to wrong part of the time index as its prediction is visibly worse compared to NHiTS.

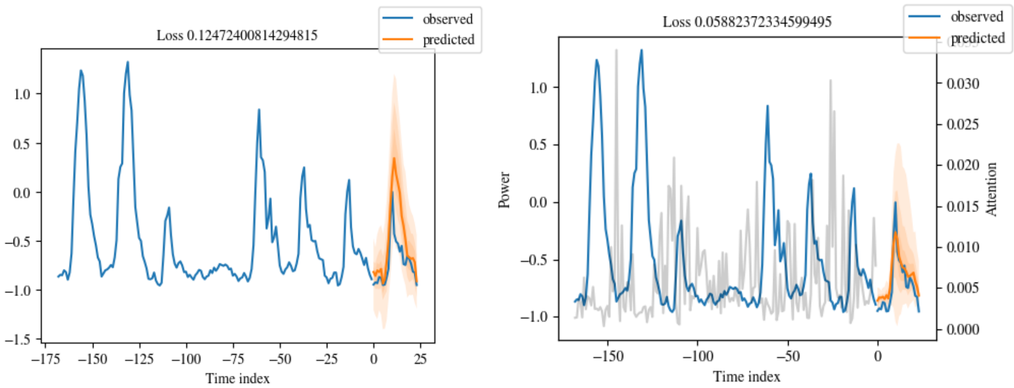

Figure 9 NHiTS and TFT example 2.

In the example in Figure 9 it is possible to see that the signal presents a much clearer consumption pattern. The exact source is unknown. It may be an office space as there is a significant peak in the morning with two-day break between -124 and -45 indexes, which may resemble a weekend. Overall, TFT managed to make a better prediction compared to NHiTS. As the two models are probabilistic, we can see that especially in evening hours TFT model is surer into its prediction compared to NHiTS which has the outer quantiles much more scattered.

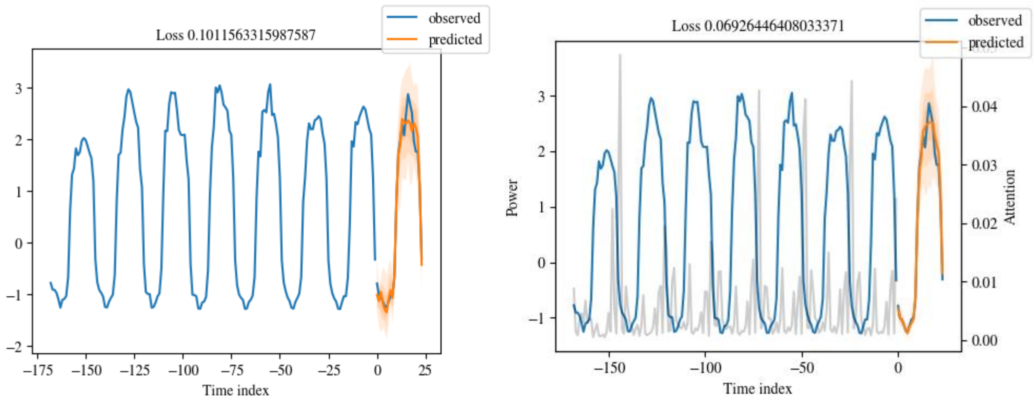

Figure 10: NHiTS and TFT example 3.

The last Figure 10 presents even clearer consumption pattern without any significant morning and evening peaks but rather having a peak during noon and minimum during the night. Again, TFT made a better prediction compared to NHiTS. While it did not predict the small dip in the morning, it was much surer into its prediction especially during the evening hours where its quantiles were much less scattered compared to NHiTS. We can also observe how last three days of the consumption and consumption from the same day last week have a big impact on final prediction, as that is where the attention is.

3.1.5 Discussion

Results show that novel models such as NHiTS and TFT manage to outperform existing models by a significant margin as was seen in Section 3.1.3.1. While simple linear regression performed surprisingly good, we must consider that it took a lot longer for linear model to converge, especially when comparing it to NHiTS.

These findings point to the opposite of the findings from paper [1], where authors stated that simple linear models outperform transformer-based model in long-timeseries forecasting. Our results are more in line with publication [6]. One thing to point out is that paper [5] did not include the TFT, NHiTS or other variations (NBEATS), which existed at the time and were specifically developed for long-timeseries forecasting. It's hard to make direct comparisons between the studies as the experimentation setup is not entirely the same. Overall, we can state that for our use-case of short-time series forecasting TFT and NHiTS outperformed linear models.

When comparing the performance between NHiTS and TFT, results in show that when TFT is complex enough, or in other words number of hidden layers is large enough, it significantly outperforms the NHiTS as was seen for PyTorch forecasting implementation in Section 3.1.3.2. Similar observations were made when we compared their performance on actual examples. On the other hand, NHiTS enables smaller models that can converge to a good solution very quickly. Due to these characteristics, we will utilise it in future work.

Overall PyTorch forecasting implementation was more stable and utilised generators which saved a significant amount of memory. This is one of the reasons we will utilise PyTorch forecasting in future implementations.

3.1.6 Conclusion

The results show that while transformer-based models perform better than traditional models, when comparing results of NHiTS and TFT with their adjacent publications [9] and [10], we can conclude that the experimentation pipeline is valid, as it yielded comparable results for both models. In future work we will focus on using both NHiTS and TFT models. NHiTS fast training time enables us to make quick iterations when developing and fine-tuning the models. TFTs option to include future covariates such as holidays or modelled weather may present a good asset for accurate demand prediction.

3.2 Transport domain [11]

3.2.1 Introduction

As a consequence of the large amount of data that transportation systems are currently capable of obtaining, different methods are emerging to use these data in a way that optimizes the means of transportation. One of the main use cases is to use these data to create models capable of predicting passenger behaviour, which usually follows rules and patterns that facilitate its extension in the time domain. Knowing passenger behaviour will make it possible to make the best use of resources and use them according to demand. In addition, this forecast will help to better manage the organization, allowing, among other things, the adjustment of schedules, the addition of services, the communication of information to passengers in response to conditions of increased demand, and the encouragement of passengers to postpone or alter their travel plans.

Therefore, the objective will be to create a model capable of generating samples and predicting their destination. This will be validated with data from line 12 of the Madrid metro. The main reason for choosing only this line is that it is one of the few in the Madrid metro system where the entry and exit records of each passenger are captured, so that the real destination of the users can be known, unlike what happens in the rest of the Madrid metro system.

3.2.2 Model

Before constructing the model, it was necessary to conduct research on the aforementioned field of study. Most studies aimed at passenger prediction in transportation systems focus on IO (inflow-outflow) prediction, which only predicts inflow and outflow. Few of them focus on OD (origin-destination) flow prediction, which not only predicts the inbound and outbound flow, but also how users move within the system. Therefore, our proposal has more value, as there are few studies in this area. Within the OD prediction studies, our proposal is framed in the static prediction, which is suitable for long-term transportation planning and design.

The construction of the model will be divided into several steps. First, a prior data processing will be performed to subsequently build the sample generation and prediction models that will allow the static OD prediction to be performed.

3.2.2.1 Data Processing

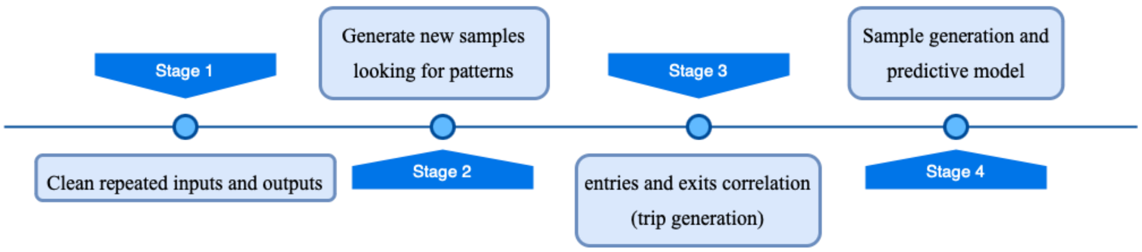

Figure 11: Pipeline of the data processing procedure followed.

The generation of the two models is divided into several steps, first, the reconstruction and elimination of the samples, followed by the correlation of each origin-destination pair. The previous figure summarizes the main steps followed, which will be further detailed below. This achieves a more correct modelling, since even though the records indicate some quantities of inputs and outputs, the reality may be somewhat different due to errors in the generation.

First, the samples are filtered and reconstructed. In order to do so, it will be necessary to know the different incidences that may occur with respect to these data. So, an important aspect to take into account is the repeated input and output samples, i.e., if a card records two or more input or output samples in a row at the same station, the system is generating records by error.

On the other hand, a reconstruction of samples not generated by the system is performed. Again, grouped by card, i.e., by user, there are several patterns that clearly, if they are met, indicate that there has been an error in the generation. So, the idea will be to search for these patterns in the dataset to generate these new samples. The principle for the regeneration is based on finding the pattern *Input 1-Output 1-Input 2-Output 2* of the same card, and where one of these four samples either input or output does not exist. In turn, either Input 1 must match Output 2 station, or Output 1 must match Input 2 station, depending on which sample is missing. Although it should also be noted that the search for the pattern will be somewhat different if the missing record is one of the two extremes, performing a check to show that the previous or next sample (outside the pattern) is an output, in the case of missing Input 1, or an input in the opposite case.

Once one of these four patterns is found, a new sample is generated, whose recording time is calculated as the difference in time between the two records (path time) correctly generated, adding or subtracting it from the remaining input or output record, respectively. Clearly, the standard route that a passenger usually follows is entry through one station, exit through a new one, and the return route.

Finally, each input and output pair are correlated, to create trips. For this purpose, queues of data grouped by card are used, where an input and an output are correlated as long as they are followed in the queue itself, their fields indicate that they are input and output, respectively, and the time between them is less than 2 hours. In all other cases, the items in the queue are discarded as incorrect data.

3.2.2.2 Sample Generation Model

Thanks to the above processing, the sample generation is carried out with the sample generation model. This model needs the probability density functions, which will generate the samples, as well as the total number of inputs per station.

The veracity of the model is based on the principle that the samples follow the same trend in the metro systems due to their weekly pattern (working week and weekend). The figures below show that the metro data follow the same pattern over the years, with only the total number of trips changing.

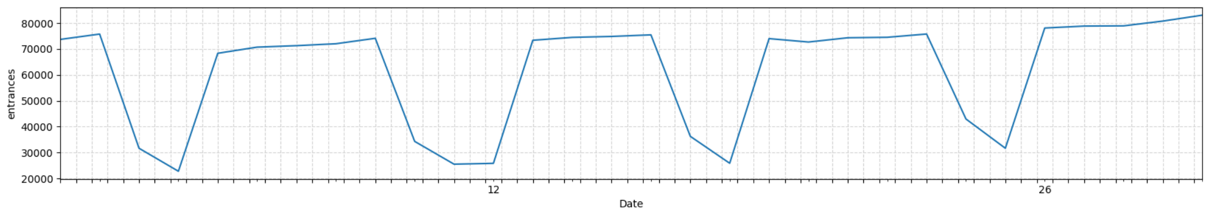

Figure 12: Madrid Metro passenger inflow October 2020

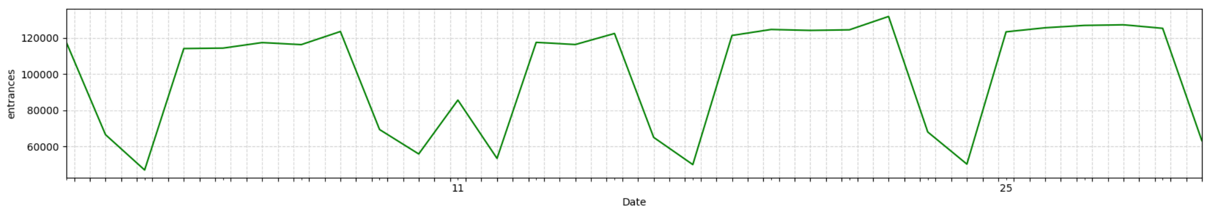

Figure 13: Madrid Metro passenger inflow October 2021

Furthermore, these graphs show the pandemic and post-pandemic effect, in which the passenger flow suffered large deviations. Therefore, this further verifies our hypothesis.

The pattern of the data is in the form of a plateau during the working week with a dip during the weekend. This pattern is followed with the exception of occasional events or events due to public holidays such as 12 October in 2020 or, in the case of 2021, 11 and 12 October.

This pattern is also particularly prominent in residential areas, where the origin-destination is generally composed of an origin O, which is a residential area, and a destination D, the work area, or vice versa, implying that the modelling of the samples keeps this trend constant over the years. All the above further verifies the validity of the model, given that our study area, line 12 of the Madrid metro, is a residential area.

As for the probability density functions, it will be created one per station and weekday following the previous discussion. The most important factor is setting a correct bandwidth, which will indicate how similar we want the generation to be to the initial model. Since looking for a very low bandwidth does not make sense because the samples will be generated in almost the same seconds, a slightly higher value is sought to reflect human behaviour, where, for example, a worker enters each station at very similar times, but not always the same.

3.2.2.3 Prediction model

The machine learning algorithm called XGBoost has been used to perform the prediction model. To make the destination prediction, our features will be the station of origin, the day of the week and the number of minutes elapsed since 5 a.m. of the day itself, since it is at this time that records begin to be stored in the system. Finally, a model that chooses the destination based on the probability weights of each destination has been used to make the prediction. This implies that, unlike other models, the most probable option is no longer selected, but one of them is chosen so that the overall prediction is adequately weighted. Otherwise, the most probable would always be selected even though other destinations may have a high percentage of occurrence.

3.2.2.4 General Process

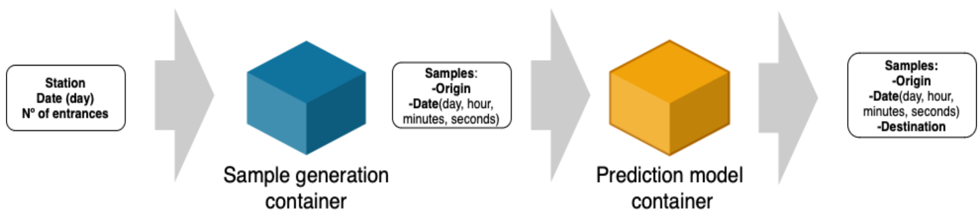

Figure 14: Box Chart

A box diagram of the logical procedure followed for the use of the global model is shown in the previous figure. This takes into account that the data processing has already been done and that the sample generation and prediction model have been generated.

As can be seen, in order to generate samples, the station, the number of inputs, and the day will need to be specified. This process will generate samples composed of date and station of origin, by means of probability density functions.

Obtaining the inflow is a problem that for metro systems would be reduced to time series prediction. However, this prediction has not been obtained in our work because of the lack of data. Therefore, our model needs to know the number of inputs per station in order to generate the samples. In short, it is a model that returns the metro usage. This can be used for both past and future prediction.

These origin samples will be the ones forecasted by the prediction model, thus obtaining the destination station for each one. Finally, it should be noted that each container can be executed separately, obtaining only the generation of samples or the target stations.

3.2.3 Validation and results

In order to validate the model, experiments have been carried out to test its functionality with data from line 12 of the Madrid subway. The data consist of one week of passenger records. Therefore, since our model is weekly, experiments can be carried out. Each record consists of the fields: card ID, station, hour, minutes, seconds, date and validation code, which indicates whether a record is an entry or an exit.

Consequently, a model has been trained looking for the best hyperparameters that minimize the error between each number of daily trips of each origin-destination group.

Once the prediction model was obtained, the dataset was divided into 80% for training and 20% for prediction, evenly distributed according to days of the week and hours. Due to the lack of data, it would be ideal to perform the training for one week and validate with another week, but this has not been possible. When using our prediction model as a function of the probabilistic weights, it loses a little of its great potential because all the weekly samples are necessary for a correct balancing of the final prediction. Therefore, to test this statement, another prediction is made with 100% of the dataset to train and the same 100% to validate.

Table 10: Experiment results

Metric	Score
MAE 80-20 %	5,42
MAPE 80-20%	16.91%
MAE 100%-100%	8.19
MAPE 100%-100%	5.07 %

The mean absolute error (MAE) and the mean absolute percentage error (MAPE) were used to measure the error.

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|M_i - \widehat{M}_i\right|$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{\left|M_i - \widehat{M}_i\right|}{M_i + \epsilon}$$

To conclude, the results obtained are particularly good, allowing an improvement in transportation planning. Further research will be needed to consolidate these results by taking into account other factors, such as changes in the distribution of probability density functions depending on seasons, weather, events, or other occurrences.

These factors are being considered in consequence of a new version of the dataset covering a longer time span. This extended dataset has passenger record data for three years, so the model will be able to take these factors into account. Finally, with this new set, the previous model has been verified, performing new experiments but this time training and validating with different weeks. The results obtained are a MAE = 14.56 and a MAPE = 10.05%. Thus, the validity of the model is reflected.

3.2.4 Conclusion

The model has been trained and tested with real-world data, and it has been validated showing a good performance. Data from the real world trend to be biased and frequently incomplete; in this case not all the travels registered have an origin and a destination; in some cases, one of the two points were missing. That implies that a perfect model is unfeasible, nevertheless, error margins are a very good approximation.

The next steps point towards understanding whether our results can be generalized to other transport infrastructures, including TT's. The method has been applied to a single metro line, but it does not require any specific characteristics of the network. This implies that it can be automatically applied to other transportation networks as long as information on individual travels is available, including origins, destinations, and timestamps.

Although the model has been tested with a grater dataset, the main development has been performed using one week of data. This could be a possible limitation, as this period cannot reflect the impact of exogenous events, such as holidays, weather events or event restrictions. The use of the extended dataset in combination with external sources will lead to better handling of these events.

3.3 Health domain

3.3.1 Introduction

Planning and model a hospital's needs during a pandemic such as Covid-19 has been very challenging. But the data collected during that time is very useful to predict future demands. This kind of data can help with short-term forecasts [11]. Short-term forecasts are very important to plan the hospital's needs when a pandemic period is triggered. In addition, long-term forecasts also play a critical role in order for the decision-makers to have time to be prepared and be able to plan accordingly their strategies. All the above can help with the planning of managing the staff, procuring drugs, or procuring personal protective equipment, which need some time to be allocated [12].

The possibility that Covid-19 has long-term effects raised the uncertainty of workforce planning. The continuous symptoms of the disease increased the fear of more medical treatments [13]. Moreover, the health staff had a greater chance of being infected, which indicated that the risks for Covid-related death or disability at work are especially significant in this type of job [14], which could potentially give rise to work impairment [15]. All these aspects can affect the total amount of staff and the corresponding demand.

Many studies have utilized ML algorithms in order to solve the problems discussed based on the Covid-19 period. A study proposed that it is very important to forecast the Covid-19 trajectory in China. This was proposed in order to estimate the size, length and ending time of Covid-19 in China. The method utilized an auto-encoder and clustering algorithms [16]. Being able to estimate the trajectory of a new pandemic seems helpful in order to predict other relevant healthcare factors.

A more straightforward method was proposed in another study utilizing some ARIMA models for specific countries like Spain, France, and Italy [17]. Other studies proposed logistic regression models, LSTM models [18] and support vector regression models [19].

It is important to mention that, during Covid-19 it was not feasible to provide what will take place in the near future for plenty of European countries and also for some states of the United States of America [20].

When it comes to supplying essential healthcare equipment, different models may result in very different outputs of the forecasted resources. This often results in an erroneous allocation of what is provided since the outputs of the models are inconsistent. Nevertheless, giving precise predictions of the peak demand for capacity in the healthcare sector is problematic owing to the lack of data or its inaccuracy since many times this data comes from manual work. This comes in addition to the difficulties involved with anticipating the results of the reported changes in mitigation actions [21].

The attempts that have been made up to estimate in a very precise way the course that any emergent pandemic would take over the next few days are limited because there are a variety of assumptions and factors. Consequently, using one forecasting approach will lead to wrong results of how the pandemic would develop.

3.3.2 Datasets

Datasets of the healthcare sector had a latency because of the manual work that was needed from the corresponding CIs side. Below some snapshots of the new updated data are depicted.

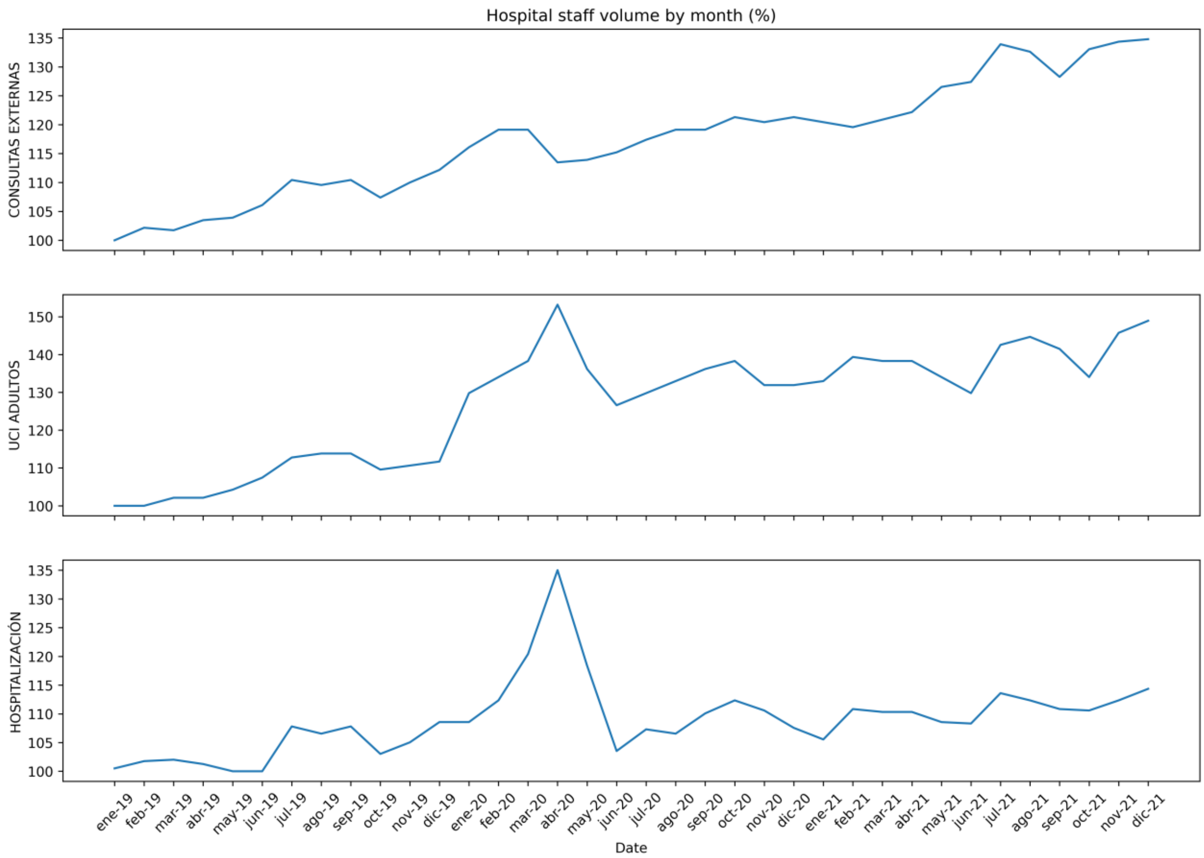

Figure 15: Sample of QS/HQM Staff requirement data compared with pre-covid situation.

Fecha compra	Descripción de línea		Cantidad
14/1/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	5
20/1/21	CICLOSPORINA 100 MG CAPSULAS MICROEMULSION	C/30	5
20/1/21	CICLOSPORINA 25MG CAPSULA	C/30	10
20/1/21	CICLOSPORINA 50 MG CAPSULAS	C/30	10
26/1/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	5
26/1/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	20
19/2/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	4
9/3/21	CICLOSPORINA 100 MG CAPSULAS MICROEMULSION	C/30	5
22/3/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	4
7/4/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	3
7/4/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	8
12/4/21	CICLOSPORINA 50 MG CAPSULAS	C/30	5
13/4/21	CICLOSPORINA 100 MG CAPSULAS MICROEMULSION	C/30	10
13/4/21	CICLOSPORINA 50 MG CAPSULAS	C/30	5
19/4/21	CICLOSPORINA 100 MG/ML SOLUCION ORAL 50 ML	UND	2
22/4/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	5
27/4/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	5
6/5/21	CICLOSPORINA 50 MG AMPOLLAS IV 1 ML	C/10	15
21/5/21	CICLOSPORINA 100 MG CAPSULAS MICROEMULSION	C/30	10
6/7/21	CICLOSPORINA 1 MG/ML COLIIRO EN EMULSION 0.3 ML.	C/30	5
26/7/21	CICLOSPORINA 1 MG/ML COLIIRO EN EMULSION 0.3 ML.	C/30	5
12/7/05	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	6
12/2/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	4
1/3/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	6
18/3/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	5
25/3/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	3
1/4/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	4
1/4/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	6
9/4/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	6
20/4/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	2
23/4/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	2
3/5/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	3
5/5/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	5
13/5/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	3
3/6/21	TOCILIZUMAB 200 MG VIAL PERFUSION IV 10 ML	UND	5

Figure 16 : Sample of QS/HQM Drugs requirement data

RESPIRADORES	UD.	FECHA	Nº INVENTARIO
RESPPIRADOR AVE COMPREHENSIVE	2	18/03/2020	2206709
		18/03/2020	2206710
RESPIRADOR iX5	1	18/03/2020	2206708
RESPIRADOR HAMILTON T1	2	16/03/2020	2206711
		16/03/2020	2206712
RESPIRADOR HAMILTON C3	1	16/03/2020	2206713
RESPIRADOR SERVO-AIR	3	01/04/2020	2206732
			2206733
			2206734

Figure 17 : Sample of QS/HQM Equipment requirement data

Ref.Prov/ Cod.N	Artículo	Descripción	Unidad	Enero	Febrero	Marzo	Abril	Mayo	Junio	Julio	Agosto	Septiembre	Octubre	Noviembre	Diciembre
144442	01-016128	BATA NO ESTERIL PROTECCION P/	C/100	7.000	8.700	12.200	8.200	8.400	9.900	6.900	7.500	7.000	10.400	8.800	8.800
132529	01-001505	CALZA TNT VERDE 16G	C/1000	12.000	15.000	15.000	12.000	12.000	12.000	12.000	6.000	16.000	18.000	17.000	12.000
128945	01-016147	MASCARILLA C/VALVULA AUTOF	C/6	102	84	84	24	54	72	72	108	0	90	42	66
128936	01-012528	MASCARILLA QUIRURGICA ALTO I	C/50	8.250	11.450	9.800	7.300	5.900	6.850	4.850	4.300	4.500	8.750	9.200	8.250
09030-1	01-016141	MASCARILLA PAPEL 1 CAPA C/GC	C/100	4.100	9.100	3.900	7.000	4.200	2.400	5.700	3.800	4.400	6.000	4.400	6.700
AP-DC03BV	01-016147	MASCARILLA C/VALVULA AUTOF	C/20	0	120	0	0	0	0	0	0	0	0	0	0
02090COVERSI	01-037234	GAFAS ESTANCAS	UND	0	0	7	10	0	0	0	0	0	0	0	0
0011030	01-016133	GORRO ENFERMERA VERDE	C/2000	4.000	8.000	8.000	4.000	8.000	4.000	4.000	4.000	8.000	8.000	8.000	8.000
687-826793845	01-009370	GUANTE EXAMEN NITRILO N/ESTEF	C/2000	2.000	0	0	0	0	0	0	0	0	0	0	0
942206	01-016074	GUANTE EXAMEN NITRILO N/ESTEF	C/200	22.000	22.000	28.000	18.000	28.000	24.000	28.000	12.000	20.000	28.000	22.000	24.000
992926	01-029594	BATA CIRUGIA ESTERIL REFORZA	C/28	1.260	1.820	1.624	476	420	140	364	252	392	504	364	308
992910	01-076294	BATA CIRUGIA STANDARD T-XL E	C/32	0	0	288	1.216	1.440	1.024	1.152	608	1.056	1.760	1.088	1.248
942207	01-009370	GUANTE EXAMEN NITRILO N/ESTEF	C/200	28.000	28.000	42.000	22.000	44.000	30.000	36.000	22.000	26.000	48.000	34.000	24.000
916304	01-023354	GEL HIDROALCOHOLICO 500ML	C/20	260	340	240	180	280	320	400	200	320	420	280	480
942208	01-012331	GUANTE EXAMEN NITRILO N/ESTEF	C/200	10.000	12.000	12.000	8.000	16.000	8.000	20.000	4.000	12.000	18.000	12.000	12.000
992909	01-001510	BATA CIRUGIA ESTERIL STANDAR	C/36	180	0	180	1.332	2.016	936	1.512	720	1.404	1.944	1.368	1.476
916302	01-031868	ANTISEPTICO MANOS SOLUCION 1	C/10	40	30	30	20	30	40	60	20	20	40	40	10

Figure 18 : Sample of QS/HQM Consumables requirement data

The given formats of the data are inappropriate to be used directly with the needs of the schemas for the processing and training pipelines and have to be redesigned and formatted appropriately.

3.3.3 Proposed methods

One of the main methods that is going to be followed is the multilayer LSTM network as proposed by Koç E. & Türkoğlu M. (2020) [22]. In order to build the model, the max-min normalization technique will be applied. After having normalized the data, which has already been split into train and test, they pass through the training pipeline in order to train the model. The test data is then passed through the prediction pipeline of the model, where the evaluation metrics are also being calculated.

Using the referenced LSTM model, the researchers achieved satisfactory results with an average of 3% of MAPE and 99% for R^2 when trying to predict the total equipment value. These results are significant and show us that such a model could be further experimented with and applied to our use cases. They refer to their model as highly accurate. ARIMA and SVM models also showed acceptable metrics, and we are also going to use such models in order to compare them. One limitation that has to be noted in our case is that many of the healthcare sector's data are on a monthly basis. However, the above models have been tested on daily data, and this is something that has to be considered.

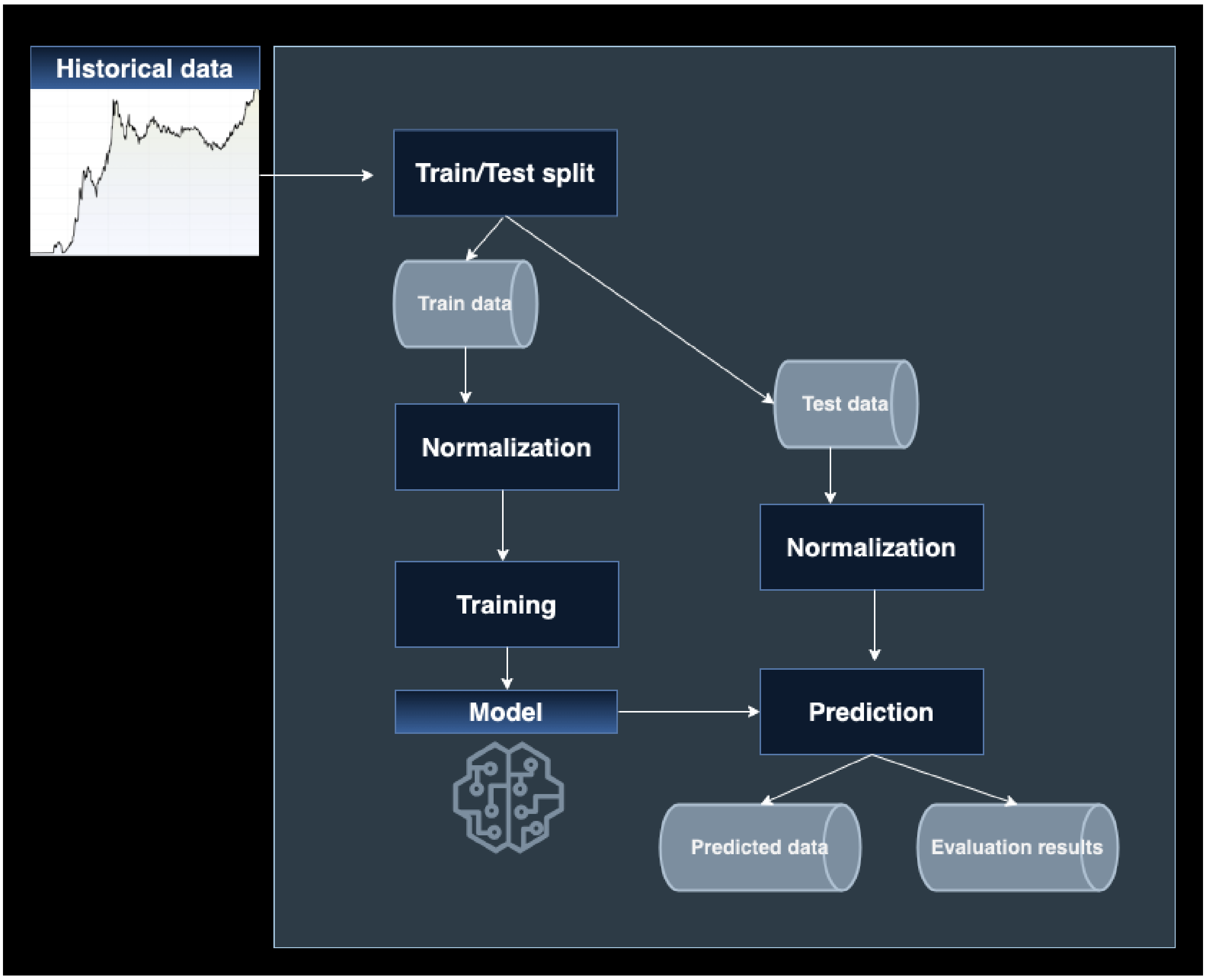

Figure 19: Simple modelling flow including the different phases

3.3.4 Conclusions and next steps

The demand for stuff, drugs, and equipment in healthcare environments is increasing due to infectious diseases and a lot more when a pandemic is happening. Based on the literature and the initial explorations of the given data, the LSTM networks can provide high-quality, accurate predictions.

There have been some specific next steps and challenges identified for the healthcare sector. The provided data have been mostly produced manually, which raises concerns about their quality. In addition, it needs special treatment in order to reconstruct their schemas since they cannot be handled directly by the modelling algorithms, and also to apply some feature engineering pipelines in order to enhance them with relevant features. Possible Covid-19 data and weather data for capturing seasonality effects may be a restriction on the length of the timeseries data as it is limited to the pandemic period. One last part to be mentioned is the monthly aggregation of the data, which can lead to more inaccurate results. In order to bypass all these issues, several experiments will take place.

3.4 Water domain

3.4.1 Introduction

The prediction of how much water will be needed to be consumed in different areas and municipalities is considered a primary objective of those responsible for controlling water supply infrastructure and water resources. This field has attracted significant attention both from industry and academia with significant advancements due to the integration of modern machine learning and data science techniques. The demand for potable water is continuously increasing as a result of climate change.

The rising rates of urbanisation are also affecting this situation, in combination with industrialization and the overall population expansion. In order to effectively manage water resources, it is important to conduct demand projections over a variety of different time periods.

The corresponding systems that help in water demand forecasting allow, in general, the prediction of water use in both short-term and long-term perspectives, based on the objective of the observations being made. Long-term assessments (looking out 20–30 years into the future) are conducted with the intention of providing decision-making in matters pertaining to the design and development of water supply systems.

Short-term simulations, which often run hourly, daily, or weekly, are used to help with the more efficient amount of effort that pump stations put in. They also help reduce the amount of energy that is needed to be consumed as well as address existing operational challenges.

One main goal of such forecasting systems is to make certain that consumers will have access to an adequate quantity of water in the long-term. This is challenging, particularly in nations that have a dry environment. This situation is notably exacerbated in major metropolitan centres, where the number of inhabitants is always rising [23].

It is very important to use the methods that are used in long-term forecasting in order to be confident in evaluating the adequacy of the available water resources. General knowledge of the effects of the different variables that result in anomalies in water consumption from the water supply network is required for both short-term and long-term forecasting in order to produce accurate forecasts.

In order to make accurate short-term forecasts, a different mix of meteorological parameters must also be analysed in great detail. However, in order to make accurate long-term projections, demographic and economic factors also have to be considered.

In order to produce a water demand forecasting model of the finest possible standard, it is necessary to understand the many different weather elements. Failure-detecting systems that neglect to take into consideration the variable of water usage will provide a higher percentage of false forecasts if such important variables are ignored. It is thus of the utmost importance to include a combination of weather elements in the inputs that are used for water consumption forecasting.

3.4.2 Datasets

Datasets were described in depth in SUNRISE D5.1 (Section 3.4.4). We have two different countries for the water demand prediction in Italy and Spain, with the corresponding CI representatives CAF and ACO accordingly. Below we mention some new enhancements of the given datasets based on other dataset sources (i.e., Covid-19 data, holidays data, weather data, etc.).

An example of the seasonality, trend and residuals of the ACO's dataset is depicted in Figure 20.

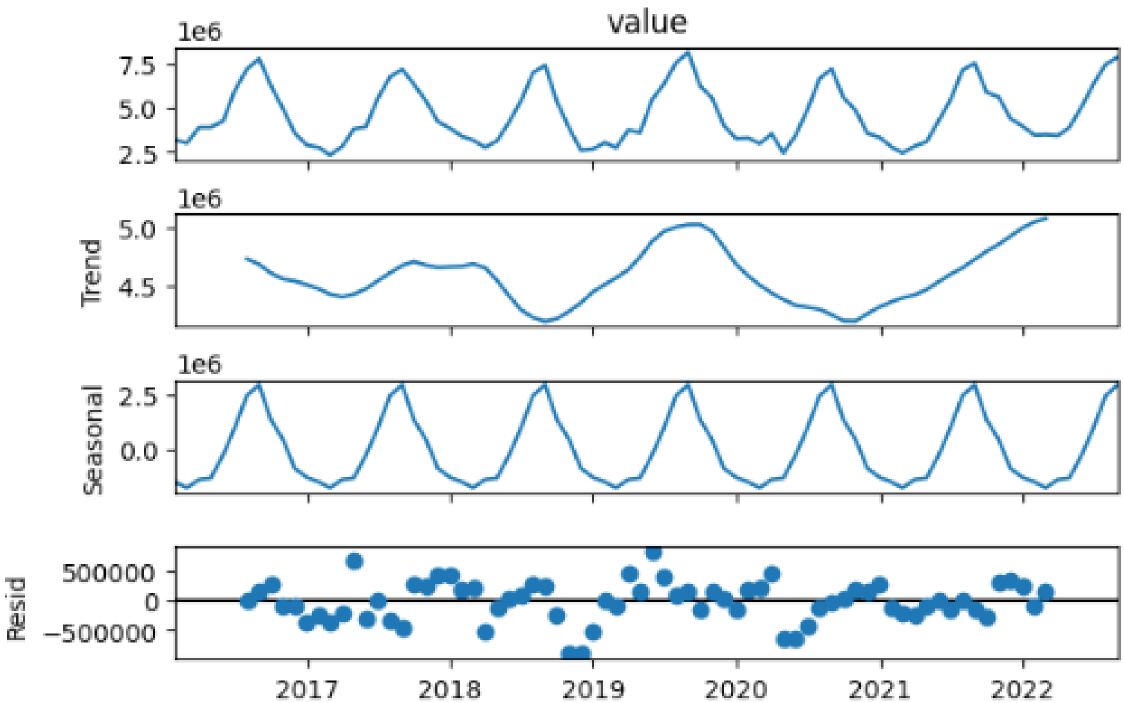

Figure 20: ACO's datasets seasonality, trend and residuals of water consumption variable

An example of the unified ACO dataset can be seen in Table 11.

Table 11: Example of the unified ACO's dataset

	date	value	Temp 2m max	Temp 2m min	Apparent Temperature max	Apparent Temperature min	Rain sum	Month index	Covid cat	holidays
0	2016-01-31	3191848	16.76	8.55	15.08	6.54	1.23	1	0	2
1	2016-02-29	3003755	17.04	8.31	14.26	5.49	0.69	2	0	0
2	2016-03-31	3910563	19.21	7.53	17.01	5.45	0.92	3	0	1
3	2016-04-30	3914040	21.19	11.38	20.08	9.97	2.55	4	0	0
4	2016-05-31	4261544	24.36	13.74	23.86	12.71	3.46	5	0	1
	...	...	...	...	...	...	...	...	...	...
75	2022-04-30	3907548	21.34	10.25	20.55	8.33	1.50	4	1	1
76	2022-05-31	5073066	29.56	16.05	30.16	15.70	0.23	5	1	0
77	2022-06-30	6372946	32.36	18.90	32.62	19.02	0.01	6	1	0
78	2022-07-31	7481473	37.78	22.65	38.89	23.54	0.00	7	1	0
79	2022-08-31	7909262	34.31	20.89	35.73	22.71	0.14	8	1	1

	place	entry_count	missing_values_sum	zero_values_sum
0	desaladora	80	1	10
1	etap	80	0	0
2	gibraltar	80	0	24
3	presa	80	0	0
4	rfuengirola	80	0	0
5	rguadalmansa	80	0	80
6	si4	80	0	78

Figure 21. ACO's dataset breakdown

As presented in Figure 21, two sources (si4 and rguadalmansa) have not been used so far. In detail, si4 is considered to be a possible resource for the Malaga city as well as the source of the river Guadalmansa. Moreover, the desalination plant "desaladora" has periods of maintenance or if the general water situation is very good, production could be paused as the running cost is high, that is why we also see 0 values in this source. Lastly, the water input from Gibraltar area is also only used, when necessary, as an extra resource. Based on these findings we have concluded in the following formula for the calculation of the total input inflow:

$$total_{inflow} = inflow_{etap} + inflow_{desaladora} + inflow_{rfuengirola} + inflow_{rguadalmansa} + inflow_{gibraltar} + inflow_{si4}$$

CAF also enhanced the previous given dataset with more inputs. Based on the noise that has been identified in the new dataset, different normalization and outlier detection methods are going to be followed. An example of the new series with these identifies issues can be seen in following figure.

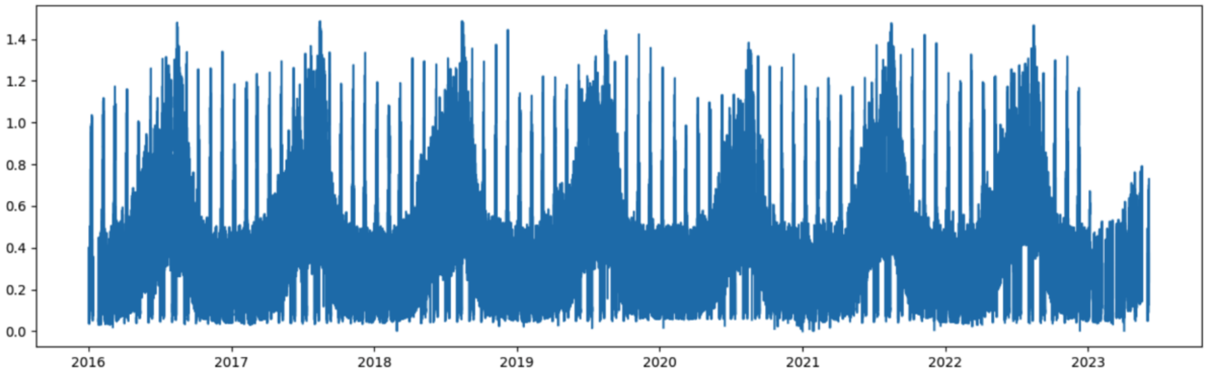

Figure 22: Example of the updated target timeseries from CAF

The different 4 new different districts are:

1. INTERMEZZO BIAUZZO-CROSERE-LIGNANO
2. INTERMEZZO BIAUZZO-LIGNANO (via Rivignano)
3. LATISANA
4. LIGNANO S.

3.4.3 Methodology

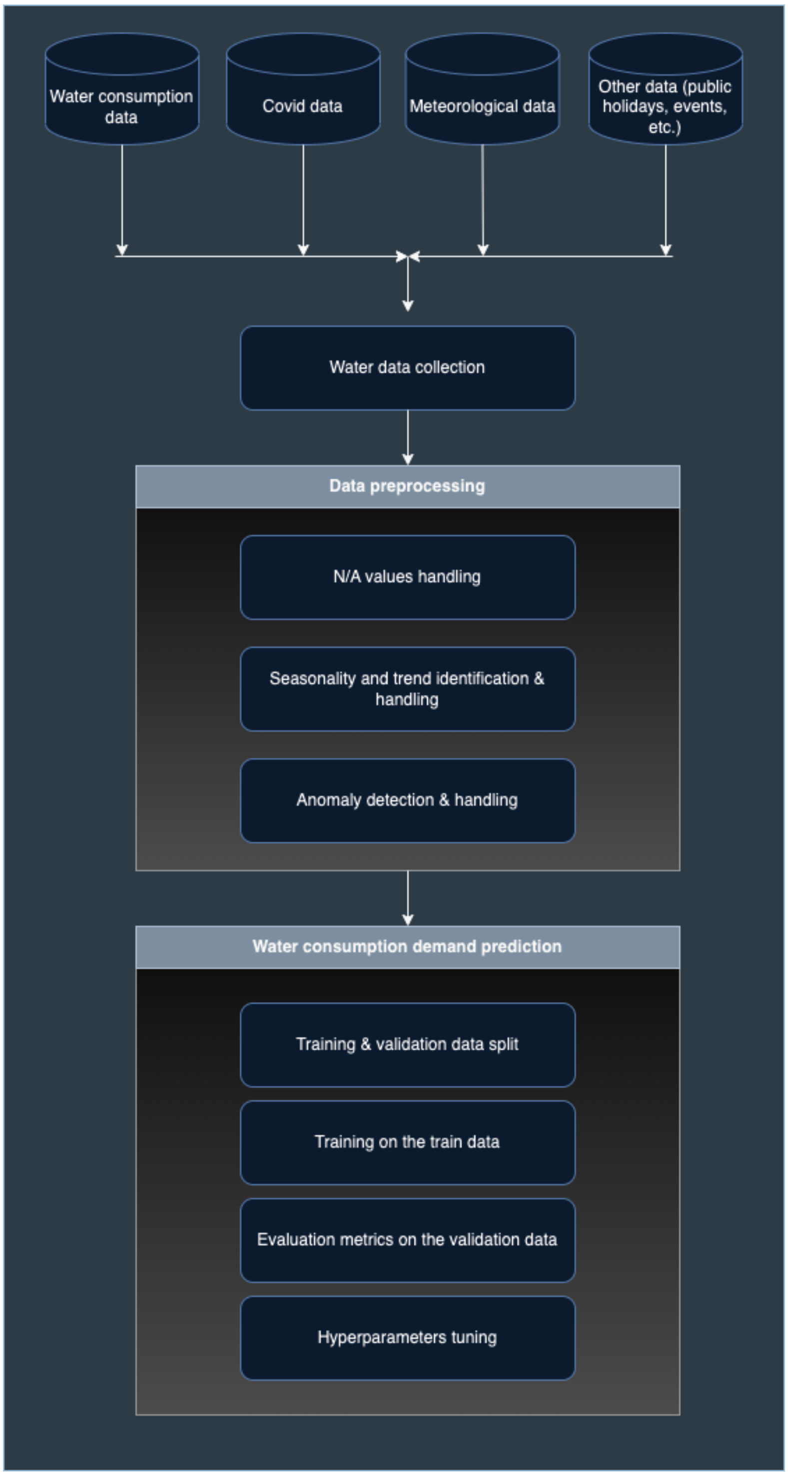

Figure 23: Water use case flow

The methodology we used for the first iteration of the modelling part is a common one for working with timeseries data. The only difference between the Water Cis is that in one case, the main water dataset is on an hourly level, while in the other, it is on a monthly level. The first part is to combine all the different sources into one unified dataset that will be used for the training of the Water models. So, we need a combination of the main water datasets with the corresponding meteorological data of the corresponding locations, the public holiday data, and also the Covid-19 data. Considering the Covid-19 data we are using in our experiments, the "severity" variables and also the "stringency index" and "lockdown periods". In our next iterations, we are also going to use some macroeconomic factors and test their possible effects and significance.

After we have the unified dataset, we need a preprocessing pipeline to run. This pipeline includes mainly the handling of NA values, the detection and handling of possible outliers, and the identification of possible trends and seasonality. At this step, normalisation also takes place where it is applicable.

As a final step, we split the data into training and validation sets, train the corresponding model, and then evaluate the results based on the preselected evaluation metrics. The hyperparameter tuning step is also included here.

As an in-between step, we do some feature engineering in order to produce needed variables (based on the model that we are going to train), like month or year index, specific lags of the targeted variable, etc.

In detail the training setup procedure used in the ACOSOL use case, which can also apply in the CAF use case, is described below.

On a first approach, we utilize Pycaret, which is a python machine-learning library for model development (training, tuning, evaluating). The library offers a very useful feature for model selection, which we use in our strategy. In this step, we used all available data, utilizing the built-in cross-validation feature of Pycaret, which in every fold holds out a validation set, and computes the metrics using this “unseen” data. Moreover, we have not used the COVID feature, as we believe that will not make a difference in the comparison of the models.

The metrics below are the average metrics for all k folds. The results in forecasting approach are ordered by MASE (Mean Absolute Scaled Error), which is a metric suitable for forecasting and in regression approach the models are ordered by R-squared. Our evaluation is based on all available metrics but mostly on R-squared, which gives an overall view about the correlation of our features to the target variable and MAPE (Mean Absolute Percentage Error), which gives us a comparable metric across all models. Unfortunately, our Pycaret setup does not include prophet, LSTMs, and ARIMA models in the available models, however we will also use it for forecasting, to set some baseline.

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
xgboost	Extreme Gradient Boosting	770068.1750	872770830336.0000	917862.4562	0.9021	0.1405	0.1123	1.2940
knn	K Neighbors Regressor	915408.8188	1172285528473.6001	1039739.8438	0.8741	0.1412	0.1252	1.2810
gbr	Gradient Boosting Regressor	874271.7155	1053092411841.4080	1002638.7173	0.8693	0.1478	0.1261	1.3360
rf	Random Forest Regressor	871495.3322	1268764824987.2483	1076052.0624	0.8251	0.1505	0.1223	1.3790
et	Extra Trees Regressor	1028820.2036	1655865622451.1519	1246062.0029	0.7887	0.1737	0.1413	1.3690
ada	AdaBoost Regressor	1035230.4869	1817388408370.6277	1282105.9359	0.7584	0.1606	0.1361	1.3000
dt	Decision Tree Regressor	1195257.7100	2080584997917.6631	1408547.3702	0.7192	0.2052	0.1701	1.2810
lightgbm	Light Gradient Boosting Machine	1939212.9684	6435502628026.3486	2489887.4376	0.2535	0.2905	0.2580	1.4350

Figure 24. Regression model selection

	Model	MASE	RMSSE	MAE	RMSE	MAPE	SMAPE	R2	TT (Sec)
snaive	Seasonal Naive Forecaster	1.1917	1.0900	1052739.2667	1284773.3132	0.1371	0.1370	0.8251	0.7500
stlf	STLF	1.2866	1.1889	1128756.4658	1394043.5126	0.1447	0.1398	0.7773	0.0333

Figure 25. Forecasting model-selection

Based on the above regression results, Figure 24, XGBoost and KNeighbors seem to have the best performance on our dataset. With R-squared more than 80% and MAPE less than 15%. To understand the comparison and evaluate the usefulness of the model, we have also included some baseline models for both forecasting and regression. KNeighbors can be used only for univariate timeseries so it is not going to be used further.

We have split our data into train (80%) and test (20%) set to evaluate our selected models on completely unseen data. For the hyperparameter tuning, due to small data sample, we used a different split (70% for training, 15% for validation and 15% for test).

In the regression approach, we use as features the year, month and the binary COVID feature, as explained earlier. We have also tested several parameters. More specifically, for the XGBoost, we tested the parameters below:

Table 12: Parameters of XGBoost

Name	Bound	Best Parameters
max_depth	1-9	5
learning_rate	0.01 - 1.0	0.1024
n_estimators	50-500	308
min_child_weight	1-10	2
gamma	0.000000008 - 1.0	0.0001
subsample	0.01 - 1.0	0.9999
colsample_bytree	0.01 - 1.0	0.8226
reg_alpha	0. 000000008 – 1.0	0.0185
reg_lambda	0. 000000008 – 1.0	0.0006

However, the default parameters of XGBoost (learning_rate : 0.1, max_depth: 3, n-estimators: 100, random_state: 42) led once again to better performance. Additionally, we believe that the lower metrics during the hyper-parameter tuning are due to the short data samples for training and validation, that in most cases do not include a complete seasonal cycle. In a different setup with more data samples, the tuning probably would have given us more fruitful results.

3.4.3.1 Prediction models

There are different models that have already been tested as described in the previous sub-section and were used as baselines. Here we reference some more that will be used due to their ability to lead to reliable results. These models are the following: SARIMAX, XGBoost, fbprophet and custom Long-Short-Term Memory models (LSTM). As next steps we are going to experiment new models and also optimize the results of the current models.

Some more details for the already tested models are referenced below:

SARIMAX: The main positive part of such models is their easy interpretation. The parameters of the model depict the different relationships between the past observations and the future forecasts. These models can also capture the possible seasonality effects and do not need many training data.

XGBoost: models are efficient in handling many input features. They can also adapt themselves in order to be used in cases with non-linear relationships between the variables. This kind of models can handle large amounts of data and need an easy preprocessing of the data in order to be used. One of the drawbacks is the seasonality capturing.

FBprophet: Is a library that can help in creating Prophet generalized additive models (GAM) with many components. In general, someone can include different seasonality like yearly, weekly, daily or also create custom monthly seasonality, etc. This model can also include trend and holiday effects. The Prophet model is probabilistic, and the model's equation is:

$$y_t = g(t) + s(t) + h(t)$$

Where t is the time index, y_t is the timeseries' targeted variable, g(t) is the trend function, s(t) is the different seasonalities and h(t) captures the effects of the holidays. Through such a model nonlinear relationships between features and a target can be modelled.

LSTM: This kind of models are actually the next generation of recurrent neural networks (RNN). They are often used in order to capture patterns of sequential data. This kind of model is able to remember long sequences which helps a lot in timeseries problems. For our current experiments we used an initial simple custom-made model in TensorFlow with 4 neurons in the first hidden layer. The input shape was one time step with nine features. The selected loss function was the MSE in combination with the efficient Adam version of stochastic gradient descent. We use this model on ACO's dataset which is not considered a large dataset since the data are in monthly level.

3.4.3.2 Evaluation metrics

As evaluation metrics we used the MAE, MAPE and RMSE. We also calculate the R square for each one of our models. The evaluation metrics are referenced below:

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\text{z-score}(A_i) - \text{z-score}(F_i)|$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}_t - y_t|}{y_t}\ 100\%$$

$$RMSE = \sqrt{\sum_{t=1}^{n}\frac{(\hat{y}_i - y_i)^2}{n}}$$

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2}$$

Where n is the size of the sample, $\hat{y}_t$ is the predicted value from the model on point t and y_t is the observed value at point t.

In order to conclude to safer results a combination of the metrics will be used. For example, there has been identified that some outliers exist in our data, and it is needed to increase a penalty to models that do not capture them. This is the reason the RMSE will be used as an indicator for such cases since it is more sensitive than MAE.

3.4.4 Results

The current results are depicted in Table 13: Results for water experimented models

Table 13: Results for water experimented models

Model/Metric	MAE	MAPE	RMSE	R^2
SARIMAX	373.295,00	0,086	450.352,00	0.912
FBprophet	217.603,00	0,04	256.614,00	0.98
XGBOOST	257.795,00	0,06	338.842,00	0.953
LSTM custom model	192.580,00	0,04	243.286,00	0.97

In the below visualisations and results, the target variable is normalised based on the MinMaxScaler of the "scikit-learn" package in Python. Using this technique, the numerical features are scaled between 0 and 1, and simultaneously, the relative relationships between the data points are preserved. Based on the results, this normalisation seems to be very useful. This is more important for the LSTM model we trained since such models are more sensitive to non-normalised values of the input variables. In Figure 26, Figure 27, Figure 28, Figure 29 we visualise the results of the different models.

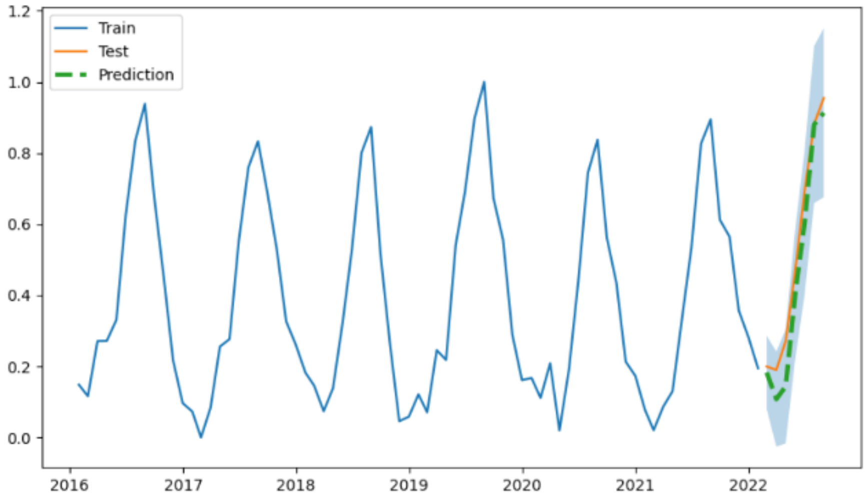

Figure 26: SARIMAX forecast.

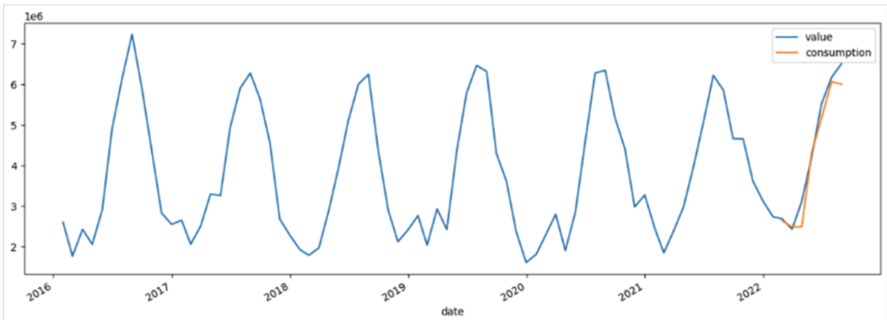

Figure 27: XGBoost forecast.

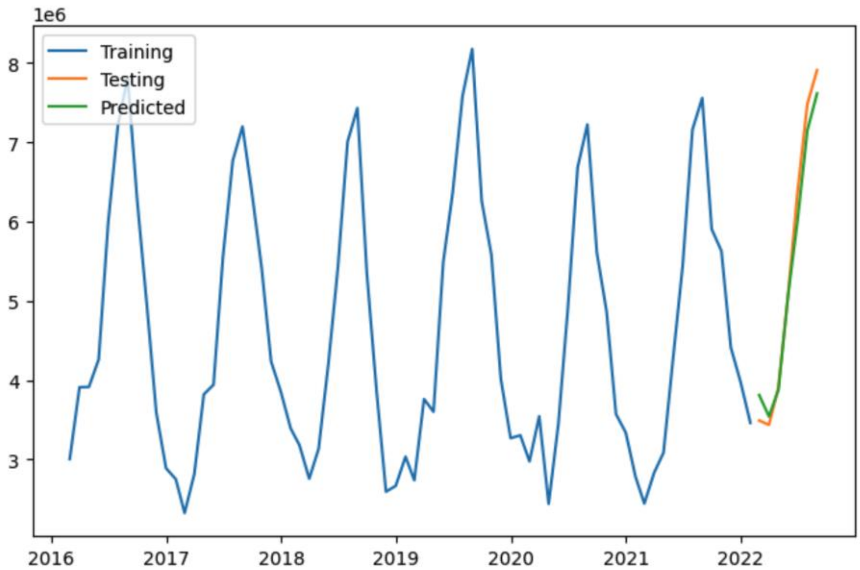

Figure 28: Custom LSTM forecast.

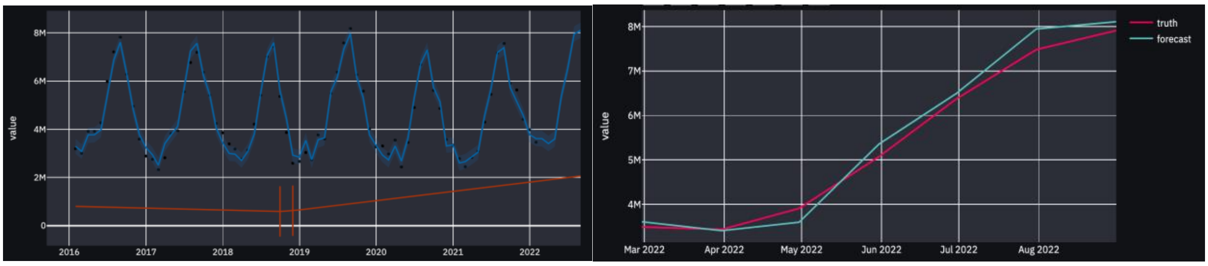

Figure 29: FBprophet forecast.

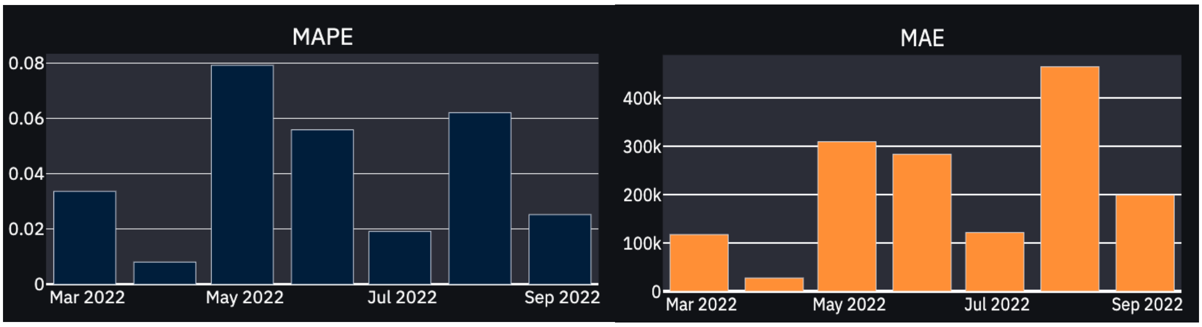

Figure 30: Examples FBprophet model metrics on month level.

As can be seen on the above table and the corresponding figures the FBprophet and the custom LSTM models had the best performance. Between these two models the custom LSTM model behaves a bit better.

3.4.5 Conclusions and next steps

We have seen that the LSTM and FBprophet indicate better results based on the selected evaluation metrics. As next steps, we are going to experiment more with these models and tune their hyperparameters. New timeseries models are also going to be experimented [24] with (i.e., Transformers for timeseries). Based on the experiments, we also identified that the different normalisation, NA handling, and outlier handling techniques have an impact on the results and evaluation metrics. So, we are going to experiment further with the different combinations. We are also going to enhance more of our feature engineering parts by identifying more possible features to be included in our training sets (more meteorological data, economic data, etc.) in order to improve our models. Finally, we want to explore the different forecasting results among the short-term and mid-term periods [25].

4 Management and What-If analysis

4.1 Effects of the Pandemic and Climate Change on Supply Chains

4.1.1 The Pandemic

One of the key aspects of globalization is the fragmentation of supply chains and the specialization of production at an international level. This has led to the concept of global supply chains, where the production of a single product involves several stages and components from different countries. Companies seek to maximize their profits by optimizing production costs using global supply chains to source product components or parts from countries where they can be manufactured at lower cost and higher quality.

This international division of labour allows companies to focus on their core competency and leverage global resources for efficiency and competitiveness. However, this model has a significant implication, namely an increased dependence on imports of components from other countries which can make national economies more vulnerable to supply-chains' disruption events: such as conflicts, natural disasters or pandemics [26].

Breaking down a key part of a supply chain can have a cascading effect (an unforeseen chain of events that occurs when an event in a system has a negative impact on other, related systems) throughout the manufacturing process. This situation occurred with the outbreak of the COVID-19 pandemic: in fact, until 2019 there was a well-functioning coordination in the concatenation of the supply chains and the delivery times of the goods were planned on a global scale. China's growing importance in the global manufacturing sector, for example, has led many companies and countries to be increasingly dependent on Beijing [27]. In many cases, Western companies have moved production in China to benefit from reduced labour costs and more available infrastructure, while at the same time reducing domestic production in other countries.

When the pandemic hit China in early 2020, there were disruptions in the production chains due to strict containment measures ("0 tolerance policies") which led to the closure of many factories and suspension of many services. This has caused significant delays in deliveries of key products and components for many European companies, resulting in reduced inventories and long supplies' times. For example, the interruption of the activities of the port of Shenzen (an important hub for the export of goods) and more specifically the closure of the Yantian terminal [28], which took place in June 2021 for almost a month due to health reasons related to the pandemic, has caused a heavy slowdown in exports given the congestion of international traffic and in the increment of shipping costs. Therefore, such situation has caused an indirect increase in costs for world consumers, especially Europeans, and a shortage of supplies of goods [29].

The aviation sector has been one of the hardest hits by the pandemic. With lockdowns and travel restrictions imposed by many European countries, air travel has dropped dramatically. Airlines have experienced a slump in demand and have had to cancel flights and reduce their fleets forcing them to refocus on air cargo. As passenger flights dwindled, many airlines began carrying goods in the cargo bays of their passenger aircraft [30] to compensate for lost revenue. Even if the earnings from the transport of goods have not been affected by the pandemic-related shutdowns, they were unable to compensate for the large loss of passenger revenues [31]. The public transport sector has also been heavily impacted by the pandemic. Social distancing measures and fear of contagion have led to a decrease in the number of passengers on trains, buses, trams and subways. Due to the infections to service personnel, hundreds of rides on the railway line and beyond have been cancelled, causing inconvenience mainly to commuter workers [32].

The reduction in the demand for the transport of goods has also led to a decrease in orders for transport by truck [33]. The increase in purchases on e-commerce sites by individual consumers has not been sufficient to compensate for the total decrease in the demand for freight transport [34]. In addition, many small and medium-sized transport companies have faced serious financial problems, with an increase in bankruptcies and closures, which has reduced the diversity and resilience of the sector [35]. The pandemic has also prompted many European governments to promote sustainable transport alternatives, such as cycling [36] and carpooling, to reduce the use of public transport and minimize the risk of contagion as well as to compensate for a reduced loading capacity of public transport required by law (50% load limit for public transport) [37].

The main problems regarding the lack of supplies have been registered in the health sector, where there has been a significant shortage of medicines and medical devices, from surgical masks, to swabs, up to the components for automatic ventilation systems, resulting in a dramatic increase of prices [38]. The pandemic has put a strain on the production capacity of medical devices and medications and, as a matter of fact, companies have found themselves faced with the need to increase production to meet growing demand, but this has not always been immediately possible, especially when raw materials primes had become scarce [39].

Strategies were implemented to facilitate the import and export of essential goods from alternative sources. However, the prevailing circumstances have underscored the imperative to enhance the robustness of supply chains within the healthcare sector. This entails diminishing reliance on external sources and fostering domestic production. This discernment has emerged as a crucial takeaway from the pandemic, essential for fortifying the healthcare system's ability to effectively manage forthcoming health crises.

To counter the shortages, numerous European States have endeavoured to amplify the local production of indispensable medical apparatus and pharmaceuticals. They have encouraged indigenous enterprises to reorient their manufacturing capacities, ensuring readiness to address unforeseen urgencies. In Italy, a notable 60% of the firms engaged in the manufacturing of specific medical devices originate from the textile and fashion domain. These entities were already equipped with the requisite machinery, workforce, and raw materials to fabricate personal protective gear like masks and gowns [40].

Illustratively, Prada stands as an exemplar, producing 80,000 lab coats and 110,000 masks in response to the Tuscany region's request to support regional healthcare personnel. In a similar vein, the Calzedonia textile group has undertaken the transformation of select facilities to manufacture masks and lab coats[2].

Regarding the water sector, potential risk incidents and hazards indirectly correlated with the pandemic emergency have been discerned. Primarily, these stem from diminished personnel resources within the sector and the constraints imposed by lockdown measures. Additionally, potential critical scenarios have been identified in conjunction with escalated consumption levels, coupled with inadequate aquifer replenishment owing to the extraordinary drought conditions that prevailed during the pandemic phase. These variables hold the potential to impose constraints on the water supply infrastructure and manpower availability, with attendant repercussions on public health.

Most European countries have legislated stringent regulations for water quality and have distribution and treatment networks that work efficiently to meet the needs of the population. While the pandemic has not caused a general water shortage, it may have caused some localized or temporary disruptions in certain areas, for example due to increased water demand in some areas or logistical problems in treatment and distribution operations. In particular, the lock-down policies, especially those involving movement restrictions, could have caused an impact on personnel involved in water treatment and distribution operations. Lack of staff or social distancing measures could have led to operational problems or delays in the maintenance and management of water systems. However, most infrastructures have been able to cope with the overall situation and with the increased demand [41].

The energy sector has experienced the most significant adversities as a result of the pandemic's impacts. The COVID-19 outbreak has had a significant effect on the energy supply chain around the world. The containment measures adopted by governments to limit the spread of the virus have led to export restrictions, factory closures and quarantines, which affected the production, transport and distribution of energy. Among the main impacts on the supply chain, the most relevant one led to the reduction in energy demand [42]. With the closure of many commercial and industrial activities, the demand for energy dropped significantly. Travel restrictions and remote working have reduced fuel consumption for transport, and industries that closed their activities have used less electricity. This has caused a temporary decrease in energy prices [43] and reduced revenues for energy companies. In particular, the sharp reduction in global oil demand has caused an unprecedented drop in prices. Price wars between major oil producers, such as Saudi Arabia and Russia, have further affected oil prices, driving them to extremely low levels. This has had a significant impact on the economies of countries dependent on oil exports, mainly due to the rebound in prices that occurred after the end of the lockdown's measures [44]. Many companies have suspended or delayed energy development projects due to economic uncertainty caused by the pandemic. Projects, such as those designed to build power stations, infrastructures and renewable energy systems, have been hit by permit delays, workforce shortages and funding problems [45].

As previously stated, the common risk factor for all the mentioned sectors was the lack of manpower. The pandemic had a significant impact on workforce availability due to various concomitant elements. The spread of the virus has led to a high number of infections among workers, forcing many people to isolate themselves or go into quarantine. This has reduced the number of workers available and compromised the continuity of operations in many sectors. Containment measures, such as lockdowns and social distancing, have led to the temporary or permanent closure of many businesses and made it difficult for workers to travel between regions or countries to get to their workplace [46].

Some job sectors, like the healthcare, have experienced increased demand for personnel but have struggled to recruit new employees due to restrictions and workplace safety concerns. The closure of schools and childcare facilities has presented difficulties for parents in organizing childcare, resulting in increased absenteeism or reduced working hours. This impact has been particularly pronounced for women, who have frequently assumed greater responsibilities in family caregiving [47]. These labour shortage challenges have highlighted the need to adopt policies and measures to protect and support the workforce during health crisis and strengthen the resilience of economies and supply chains.

4.1.2 Climate change

The increase in the frequency of extreme climate events, in recent years, has severely impacted supply chains, both globally and regionally, by disrupting the production and critical supplies. The interdependency of productive infrastructures implies that a sudden climatic event impacting on a single asset will affect every other connected asset, therefore disrupting business continuity [47].

Climate change has impacted global supply chains in multiple ways: firstly, extreme weather conditions cause a shortage of raw materials and an increase in their costs; secondly, damages to infrastructures such as roads and railroads causes delays in shipping and transportation; finally, water shortage and drought hinder manufacturing processes[3].

In this chapter, four main sectors will be analysed and related to each other in order to identify the critical aspects pertaining to the impacts of climate change. The four sectors are: healthcare, energy, water and transportation.

In 2020 the European Parliament published a report highlighting the shortage of active pharmaceutical ingredients (APIs), a great part of which are manufactured in distant countries such as India or China [48]. The increased frequency of extreme climate events, combined with their unpredictability

[3] Ibidem

determines the necessity to take into consideration the vulnerability of pharmaceutical companies to the disruption of the production or distribution of the finished product [49].

Furthermore, climate change impacts on both the medical device industry and the manufacturing sector, as it is associated with a rise in climate-related illnesses, consequently leading to an increased demand for medical devices specifically designed to address these conditions [50]. With such frequent and unpredictable weather events, such as floods or wildfires, there is a higher risk for the disruption of the global supply chains for these devices. Moreover, the temperature rise threatens the lifespan of medical devices whose components are particularly sensitive to heat and humidity variations, therefore compromising the functionality of such instruments[4]. This may require an update of testing protocols for the devices, as well as disposal procedures. Finally, global warming has an impact on sterilization processes since the change in humidity and air quality levels may hinder the effectiveness of the procedure by decreasing the germicidal properties of disinfectants and sterilant (such as chlorine dioxide and formaldehyde)[5].

Variations in climate conditions also indirectly affects the transportation industry, causing a loss in profit for the logistic sector. In 2021, Europe registered severe floods, followed by an unprecedented drought the following year. Flooding events caused a disruption in the supply chain of food, water, and medicine due to shipping delays, which required timely adaptation in logistics [51]. Similarly, under drought conditions inland ships are forced to reduce their loads, which determines an increased number of trips to transport the same amount of cargo, which in turn raises transportation's costs. The 2018 drought constitutes a primary example: on that instance, the Rhine River capacity lowered to a quarter of its normal freight capacity, becoming inaccessible to large cargo barges, and causing a major impact on industries that relied on the river. Traffic was redirected onto the major trade arteries, but it proved insufficient over time. Moreover, the reduced access to waterways led the operators to apply surcharges. The indirect costs of the lowered capacity of the river Rhine amounted to 5 billion euros loss in German industrial output for the second half of 2018 [52].

Another example of extreme weather affecting logistics is the 2022 heatwave in UK, which damaged the rail infrastructure, resulting in trains being cancelled, delaying shipments and deliveries, as well as an increased pressure on road freight. The heatwave also had a pronounced impact on the fresh produce delivery sector, primarily due to the inadequacy of appropriate equipment capable of withstanding elevated temperatures.

The inherent unpredictability and variability of weather patterns, coupled with fluctuations in agricultural production, necessitate adaptive measures within the transportation sector, which is among the most vulnerable industries. Climate change profoundly impacts freight supply chains, resulting in escalated costs for fuel and raw materials. Additionally, it exacerbates the vulnerability of small businesses within the logistics industry, leading to their collapse as they struggle to compete with larger companies [53].

According to the European Environment Agency (EEA), climate change is exerting notable effects on Europe's energy system, manifesting through impacts on the accessibility of energy sources, particularly renewable ones. Moreover, climate change is also influencing the distribution and storage of energy across the regions [54]. In particular, water scarcity impacts energy production by disrupting supply lines of gas and coal and causing the depletion of reservoirs for hydropower. Hydropower represents the main renewable energy source in Europe, accounting for almost half of all renewable power generation in the region. Changes in the hydrological cycle result in the change and intensification of precipitation patterns, making hydropower facilities especially vulnerable given their dependence on the streamflow. In particular, water scarcity in countries such as Italy, Portugal and Spain will determine a decreased streamflow, undermining energy production and increasing costs. On the other hand, North European countries such as Norway and Sweden are likely to experience

[4] Ibidem

[5] Ibidem

increased precipitation and therefore greater water availability. In the short term, this increases hydropower production, but it also puts pressure on hydropower's storage capacities, causing possible power outages. Moreover, glacier melt increases sediment transport, which accumulates reservoirs, and, over time, it causes turbine deterioration. Moreover, in the long term, temperature rise increases evaporation from the reservoir, impacting hydropower capacity [55]. Finally, variations in water supply inevitably impact electricity production costs, which are predicted to increase in southern Europe due to more frequent heatwaves, which will increase electricity demand [56].

4.2 Interdependency Among the Critical Infrastructures

As highlighted in the deliverable D1.1: "Local meetings with Critical Infrastructures Stakeholders" [57] the environment surrounding Critical Infrastructures is growing increasingly complex as a result of escalating natural disasters and the rise of pandemic dangers such as COVID-19. These risks are connected, compounding their detrimental impact. They no longer effect merely a single firm, a small community, or a specific region, but have global and cross-sectoral implications. European Critical Infrastructures rely on interdependence, which is strengthening over time. This increased interdependence emphasizes the importance of a cohesive management strategy, particularly during crises such as pandemics or catastrophic weather occurrences. Security and coordination among organizations, governments, and economic entities are critical, as is the ability to balance risk, costs, security, and resilience.

Previously, the emphasis was on protecting particular Critical Infrastructures (CIs) prior to the epidemic. However, due to the complex interdependence of several infrastructure sectors, these methods are insufficient for an effective crisis management. Following the pandemic, the focus has switched to managing cascading impacts and ensuring rapid recovery, emphasizing the critical importance of public-private partnerships and coordination among stakeholders at the national and international levels. These collaborations are critical for the correct operation of infrastructure and the management of key risks. Setting up a thorough structure to coordinate stakeholders is critical for effectively responding to risks and disasters.

The COVID-19 crisis has underlined the need for organizational resilience and the ability to keep key company processes running even when faced with adversity. Given the interconnectedness of society, understanding resilience on a larger scale, rather than on individual entities, is critical.

4.2.1 European Critical Infrastructure's Experience

Through the workshops organized within the SUNRISE Work Package 1 it has been possible to directly interrogate few Critical Infrastructures about the impact of the pandemic and climate hazard on their business continuity and supply chain systems, also highlighting the relevant interdependencies. In particular, these themes have been investigated during the first national workshop, summarized in the deliverable D1.1: "Local meetings with Critical Infrastructures Stakeholders" [57], and the first Cross-border collaboration workshop, summarized in the deliverable D1.2: "Pan-European meetings with CI stakeholders" [58]. In the next paragraph the results of these workshops will be presented to provide a comparison between the literature investigated in the previous chapter and the experience of the CIs interviewed.

The subsequent section places primary emphasis on the interdependence among distinct infrastructures and companies. It aims to provide a summarized overview, categorized by sectors, of the challenges faced and the commendable practices observed by Critical Infrastructures (CIs) throughout the pandemic, as elucidated in the workshops mentioned earlier.

Water Sector:

During the First National Workshop the participants from CIs belonging to the water sector underlined the complexity of the management of the COVID-19 pandemic due to the organizational structure of this kind of CIs. Indeed, a water infrastructure involves departments that require physical presence at the workplace and are connected to services that were disrupted during the pandemic, thus requiring the identification of suspended services. Another aspect concerned the communication system with end users, specifically citizens, alongside increased household water usage during lockdowns. This resulted in a surge of calls and unusual inquiries due to uncertainty about interacting with the organization for specific needs and concerns about water quality. For instance, an unusual request involved fears of infection from using tap water due to a perceived lack of water treatment chemicals [57].

At the same time, during the first cross-border collaboration workshop, many CIs have confirmed that their business continuity plans, and good practices allowed them to be self-sustained for up to 48 hours. Nevertheless, specific elements, like the distribution system, presented obstacles in maintaining uninterrupted functionalities and demanded extra backup systems. Several critical infrastructures had contingency strategies in place for securing the necessary water treatment chemicals, guaranteeing a two-week stock, although storage constraints brought about complications. The cooperation between supplier and client emerged as a vital factor in securing ongoing operations [58].

Energy Sector:

When it comes to the energy sector the participants of both workshops, regardless of the infrastructure they belong to, have highlighted the key role of this sector in the interdependencies theme. At the same time the members of the energy sector have outspoken how the functioning of energy suppliers relies on a long and complex supply chain that crosses several sectors [58].

Within this sector, a notable instance of heightened communication was observed with other sectors. Specifically, engagement occurred with the electricity distributor to facilitate the exchange of best practices, and collaborative efforts were undertaken in conjunction with European transmission network operators. Gas pipeline operators demonstrated exceptional international cooperation and solidarity during the pandemic. They formed a crisis planning group, conducted risk assessments, and executed agreed-upon plans effectively. Operational issues were minimal, and innovative practices like adjusted logistics, remote work, and isolated operations were adopted. Challenges arose in obtaining protective equipment and disinfectants. On the other hand, the electricity distribution sector lacked comprehensive pandemic plans. Measures were tailored to different pandemic phases and involved collaboration with entities such as the Ministry of Infrastructure, Civil Protection, and gas pipeline companies. Cross-border collaboration was also notable [57].

Healthcare Sector:

During the initial stages of the pandemic, one individual recounted that the healthcare and ICT sectors faced a challenge regarding external personnel provision. This was particularly relevant as numerous services were outsourced to external contractors. In response, a team of internal IT experts stepped up to establish essential ICT setups for employees, enabling remote working systems. It became evident that many employees were unfamiliar with basic conference systems and related technologies. Interestingly, a shortage of fundamental ICT equipment, such as cameras, was observed, and an inventive approach was taken to source this gear through social media networks. The intricacies of the public procurement system exacerbated the situation, given that the necessary equipment could solely be obtained from a particular supplier. However, the lengthy process of procurement posed a hindrance. As a result, ad-hoc decisions were made by the IT and management teams, providing flexibility in the face of regulatory constraints. The consideration of alternative solutions, such as approaching critical infrastructure operators to acquire equipment from their inventory, was deliberated upon. Regrettably, this option was contemplated belatedly.

Lastly, the individual noted disparities in how authorities collaborated with public and private healthcare entities during the pandemic. Notably, the lack of collaboration protocols between these sectors became evident, leading to challenges in patient transfers between public and private hospitals and securing free movement passes for essential workers during lockdown scenarios [57].

Similarly, in the inaugural cross-border collaboration workshop, the paramount importance of cooperation was conspicuously highlighted within the response efforts. Participants emphasized the value of productive alliances forged among hospitals, data providers, and communication networks, underscoring the pivotal role of effective communication in times of crises. They also pointed out the incorporation of weather predictions into operations, acknowledging the interconnectedness of critical infrastructures and the pivotal role of dependable information in shaping decisions [58].

Transport Sector:

During the first cross-border collaboration workshop participants have pointed out that Critical infrastructures (CIs) have exhibited a degree of autonomous operation, especially when it came to sustaining themselves, temporarily, in the events of disruptions, without relying on external links. Only the transportation sector has experienced a temporary disruption of its suppliers as a critical challenge due to the underscoring of the interdependence between different infrastructures [58].

Indeed, the transportation sector underwent significant changes during the pandemic. Freight transport witnessed remarkable success, achieving record business levels, largely attributed to the freeing up of transport routes that were no longer burdened by suspended passenger traffic. In light of these circumstances, substantial efforts were invested in maintaining the infrastructure. However, more extensive interventions in the infrastructure encountered challenges, given the need for multiple permits and intricate plans. Additionally, the acquisition of equipment followed public procurement protocols and involved tender processes, inherently consuming time. Swiftly organizing and executing various interventions was hindered by administrative barriers and limitations within the legal framework. Unfortunately, legislative procedures were not streamlined amidst the crisis.

In this context, there were no established protocols for continuous operations, prompting planning to be undertaken on a case-by-case basis, drawing upon past experiences with diverse service disruptions. Fortunately, operational services operated nearly seamlessly, thus circumventing any significant issues [57].

4.2.2 Conclusions

As economies have become more interconnected and specialized, the complexity of supply chains has increased, leaving them vulnerable to disruptions caused by conflicts, natural disasters, and pandemics like COVID-19. The pandemic laid bare various weaknesses, resulting in supply chain delays, shortages of essential goods, and significant cost escalations in numerous sectors [29].

Within the healthcare sector, the impact of the pandemic was especially pronounced, leading to severe shortages of medical supplies and equipment, as well as difficulties in the provision of external personnel for health and ICT services, and the lack of basic ICT equipment for remote work [38].

Critical infrastructures in Europe are deeply interconnected, necessitating a cohesive and integrated management approach. The pandemic underscored the significance of resilience and collaboration between organizations, governments, and economic entities during crises. Proactive planning and strategic cooperation were deemed crucial to mitigate potential disruptions and ensure the continuous functioning of critical infrastructures during challenging times [59].

The water sector confronted complex challenges during the pandemic. The organizational model of water infrastructures required physical presence for some departments, leading to difficulties when certain services were suspended. Communication with end-users, especially citizens, became crucial as domestic water consumption surged due to lockdowns [60].

The energy sector displayed resilience by maintaining collaboration with electricity distributors and European transmission network operators. Effective information exchange proved beneficial in managing operations. However, challenges were encountered in the procurement of protective equipment and disinfectants. The pandemic also highlighted the need for enhanced planning and coordination to tackle future crises effectively [42].

The transportation sector, particularly freight transport, demonstrated commendable performance during the pandemic, benefiting from the reduction in passenger traffic [45]. Challenges arose in intervening with infrastructure due to administrative obstacles and procurement delays. Collaboration among transport operators and other stakeholders played a crucial role in ensuring continuous operations.

Overall, the pandemic's impact on the workforce was a common challenge across various sectors. High infection rates among workers led to quarantines and isolation, reducing the number of available workers and disrupting operations [61].

Climate change further compounded supply chain disruptions. Extreme weather events caused water scarcity, disrupted transportation infrastructure, and hindered energy production. For instance, the hydrological cycle changes brought about by climate change affected hydropower facilities that heavily rely on streamflow. Water scarcity in certain countries led to reduced streamflow, impacting energy production and increasing costs [56].

Critical infrastructures in Europe are deeply interconnected, necessitating a cohesive and integrated management approach. The pandemic underscored the significance of resilience and collaboration between organizations, governments, and economic entities during crises. Proactive planning and strategic cooperation were deemed crucial to mitigate potential disruptions and ensure the continuous functioning of critical infrastructures during challenging times.

Cross-border collaboration workshops highlighted the significance of collaboration, interdependencies, and business continuity planning. The workshops emphasized the importance of reliable information, transparent communication, and coordination in critical sectors.

In conclusion, what is highlighted is the necessity for preparedness, collaboration, and resilience to mitigate the disruptive effects of the pandemic and climate change on critical infrastructures. Coordinated efforts, proactive measures, and strategic planning are essential to ensure the continuous functioning of vital services during crises and to build a secure and resilient future for Europe. By addressing vulnerabilities, embracing cooperation, and promoting adaptive strategies, European countries can foster a robust and interconnected network of critical infrastructures to withstand the challenges of the future.

4.3 Risk management Tool and What if analysis

The previous chapters highlighted the discrepancies between the impacts on the supply chain of a Critical Infrastructure as described in the theoretical literature and their real-life impact. Due to the heterogeneous nature of Supply Chains, for a substantively valuable outcome, the efficacy of the What if Analysis model necessitates reliance on authentic empirical data gathered directly from Organizations or CIs, rather than literary references. The main goal of Risk and What-If analysis model is to provide CIs Operators and their Supply Chain's Organization a reliable Tool for two main purposes:

1. Monitoring and forecasting of pandemic and climate emergencies' impacts;
2. Providing useful information for the business continuity decision making processes.

To actualize this objective, it is required that the Tool draws directly from data supplied by the appropriate critical infrastructure entities. The incorporation of risk scores pertaining to each Organization, CI or user is deemed to be particularly essential, as it represents the foundation for the effectiveness of the Tool.

The risks associated with the supply chain of critical infrastructure pertain to the organization's capacity to sustain uninterrupted business operations in the event of a compromise affecting a supplier. The What if Analysis requires to take into consideration every risk related to the supply chain: for this reason, it is mandatory to know the risk level of each supplier to cover the whole chain. The determination of the risk level requires the execution of a comprehensive risk analysis. Each individual organization is required to conduct this analysis, encompassing both its inherent risks and those arising from its suppliers. To facilitate this process, an effective analysis and risk management Tool is offered by Hermes Bay within the Sunrise project. However, in order to guarantee accuracy on the What if Analysis's result, a wide number of users should be reached in each geographical area. For this reason, within the SUNRISE project, a priority goal will be to obtain a high number of subscriptions to map most of the critical infrastructures and their suppliers, facilitating the attainment of a precise and nuanced What-if Analysis. In this context, the goal of Hermes Bay is to share and promote the Tool which will support the CI which will employ it. Ultimately, the extent to which various companies engage in the risk analysis via the risk management Tool will directly influence the breadth of information concerning distinct supply chains that will feed the Tool. This, in turn, will enable the Tool to provide a comprehensive "what if" analysis of heightened granularity.

4.3.1 Risk Management Tool

The first phase of the What if analysis is represented by Risk analysis. This assessment relies upon two primary metrics: one for potential threats and another for their associated impacts. The threat indicator originates from a list of open datasets, agreed among WP2 Sunrise partners, which quantifies the likelihood of a threat. For instance, the concept of probability must be understood as the probability of a threat to generate an impact on a CI or an organization. The Tool will calculate a regional probability value for each single threat. Meanwhile the indicator for the impacts is generated by data that are provided directly by the user through the Tool and it measures the impact score related to the category of impact that a threat can generate on a user.

Thus, risk analysis combines publicly available data with internally sourced information directly provided by users. The overarching objective of the risk management Tool is to compute an individual risk score for each organization, which will serve as the initial reference for the subsequent What-if analysis.

4.3.1.1 Threat Analysis

The risk management Tool offers two different modes for monitoring a threat: a real time one and a forecasting one. Each threat is characterized by two parameters: a real-time parameter and a forecasting parameter extending up to one week. The Tool can track both direct threats, such as pandemics and extreme climate events, and indirect threats, which are dependent on the direct ones.

The pandemics are described by three indicators: number of deaths, number of positive cases and number of hospitalizations. The first Demo of the Tool is provided with an historical dataset for COVID-19 where all the data for the previous indicators are collected and divided by regions for the following member states:

- France: the data were collected through the national health care system [62].
- Germany: the data were collected through the national health care system [63].
- Italy: the data were collected through the national health care system [64].
- Spain: the data were collected through the national health care system [65].

Conversely, extreme climate events threats are addressed by categorizing them based on specific types of climatic occurrences. The Tool is equipped with a dataset that aggregates historical data pertaining to these events within the designated states, categorized by distinct regions. The following table summarizes the extreme climate events considered and the open datasets that provide real time information.

Table 14: Climate events datasets sources

Extreme Climate Event	Real Time Dataset	Forecasting Dataset
Cold Spell	https://www.efas.eu/en	https://open-meteo.com/en/docs
Droughts	https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1111	https://edo.jrc.ec.europa.eu/edov2/php/index.php?id=1111
Earthquake	https://www.seismicportal.eu/	https://www.seismicportal.eu/
Extreme Rain	https://www.efas.eu/en	https://www.efas.eu/en
Floods	https://www.efas.eu/en	https://www.efas.eu/en
Forest Fire	https://effis.jrc.ec.europa.eu/	https://effis.jrc.ec.europa.eu/
Heat Waves	https://www.efas.eu/en	https://eswd.eu/cgi-bin/eswd.cgi

Finally, each direct threat indicator will be normalized with a threat score from 1 to 5. As far as the pandemic threat is concerned, the score is generated regionally by the weighted average of the three indicators, while the threat score of the extreme climate events is based on a 1 to 5 score normalization of real time single indicators. These scores represent the real time probability of the direct threats to generate an impact on the CI and its supply chains.

On the other hand, the term “indirect threat” refers to all the threats generated by the occurrence of a direct threat. Presented below are some examples of indirect threats and their description:

- **Forced remote working and acceleration of digital transition:** The adoption of this measure due to an extreme climate event and/or a pandemic caused the retraining of employees, the adoption of remote-working technologies and the reorganization of workspaces.
- **Shortage of medicines and medical devices**: Due to an extreme climate event and/or a pandemic event, a shortage of medicine and medical devices was registered due to either problems in the supply chain or an unexpected surge in demand.
- **Mandatory quarantine for tested positive**: Due to an extreme climate event and/or a pandemic, infected people were forced to quarantine, and they couldn’t leave their houses.
- **Increase of cyber-attacks**: The adoption of technologies and solutions for remote working on a large scale increased the number of cyber-attacks.

Due to the large amount and the heterogeneity of the literature about the precise probability of occurrence of indirect threats, a qualitative study has been conducted. Through an ontology, each indirect threat is related to the direct ones, or other indirect ones, which could trigger its occurrence. Indeed, indirect threats are not solely contingent upon direct ones; they can also arise from the occurrence of other indirect threats. To elucidate this concept, consider the following example: due to a pandemic as COVID-19 the hospitals are overcrowded and congested (indirect threat: layer 1); in order to overcome this problem a lockdown is implemented (indirect threats: layer 2); at this moment all the other measures to contain the impact of a lockdown are implemented: remote working, smart education, closure of non-essential sectors (indirect threats: layer 3) and so on.

Indirect threats are likewise assigned a corresponding threat score, which mirrors that of the direct threat they stem from, albeit with a one-week delay: if an indirect threat is generated by more than one direct threat, the highest “threat’s score” will be selected. Ultimately, each direct and indirect threat will be correlated with its corresponding threat score, resulting in the compilation of a list enumerating the threats and their respective scores. The forecasting model is organized in the same way as the real time model. To forecast the pandemics, a probabilistic model has been used, which considers the transformation of the pandemic’s indicators depending on the changing of meteorological parameters and the implemented national countermeasures registered on historical evidence. Instead, the forecasting model for the extreme climate events is based on the datasets showed in the previous table, that already provide forecasting data on these events.

4.3.1.2 Impact Analysis

In order to guarantee a meaningful risk analysis's result, critical infrastructures or organizations intending to utilize the Tool, hereafter referred to as "users", must manually input organizational information and data into the Tool, alongside requisite definitions.

As previously explained, the Tool's optimal functionality is contingent upon the user furnishing this information, which will then be integrated with data generated during the threat analysis, thereby affording a comprehensive risk analysis to each respective user.

The user must provide three kinds of information:

- The location of his regional branches,
- Impact assessment metrics,
- The regional impact value for each threat.

Location of the regional branches: The user must provide the regional location of his branches. Indeed, the probability of the threats is calculated on a regional base, as demonstrated in the previous paragraph. Moreover, the kind of threat to which a branch is susceptible can vary due to the location of the branch itself.

Impact assessment metrics: The user must define his global impact severity metric scoring based on an economic perspective. Five metrics are presented to the user that must fill them up with the right monetary amount. The categories are the following:

1. Very Low/NA Impact Scenario - From 0 € to ... €
2. Low Impact Scenario - From [max of Very low/NA] € to ... €
3. Medium Impact Scenario - From [max of Low] € to ... €
4. High Impact Scenario - From [max of Medium] € to ... €
5. Critical Impact Scenario - Major of [max of High] €

For instance, if an expenditure falls within the range of €1,000 to €3,000 and is considered commonplace by the user, this range will correspond to the "Very Low / NA" impact scenario for that user. Only numerical inputs are admitted into this interface.

Regional impact value for each threat: The user must define the impact value and the impact categories triggered by each threat, previously selected, for each branch. The impact should be assessed on the whole organization. This requirement arises from the recognition that, as previously indicated, these values may vary based on the geographical location of the branch. The impact categories are:

- **Economic impact:** The main impact registered after the occurrence of the selected threat is due to economic or financial losses. It should be considered the loss of profits and the actual damage caused by the selected threat.
- **Operational impact**: The main impact registered after the occurrence of the selected threat is due to organizational issues related to inefficacy/ disruption of internal procedures, lack of equipment, working teams' organization and management.
- **Manpower (HR) impact**: The main impact registered after the occurrence of the selected threat is due to lack personnel, sick-leaves, HR management problems…
- **Reputational impact:** The main impact registered after the occurrence of the selected threat is due to an "image" damage on the organization detected on social media, newspaper and social network. Because of the registered impact the following topics should be considered: costumer reduction, drop in demand for service, discontent of clients.

- **Impact on service quality/service level:** The main impact registered after the occurrence of the selected threat is due to a reduction or disruption of service level that could lead to a contractual non-compliance, penalties or sanctions and lowering of service quality.

Table 15: Impacts assessment table example

Threat	Economic Impact	Operational Impact	Manpower (HR) Impact	Reputational Impact	Impact on service quality/service level
Threat 1	- Region 1: [select 1;2;3;4;5] - Region 2: [select 1;2;3;4;5] - Region 3: [select 1;2;3;4;5]				
Threat 2			- Region 1: [select 1;2;3;4;5] - Region 2: [select 1;2;3;4;5] - Region 3: [select 1;2;3;4;5]		
Threat 3		- Region 1: [select 1;2;3;4;5] - Region 2: [select 1;2;3;4;5] - Region 3: [select 1;2;3;4;5]			

After the user has provided all the requested information, a list of threats for each region and their impact scores are given.

Finally, at the end of the impact assessment process, each user's region will show 3 scores: a threat probability score, an impact score and the risk score.

The regional risk score of a threat is calculated as the product of the probability score of the threat and the impact value generated by each threat. In case a threat is generating more than an impact, the highest value has to be selected among the impact categories. Also, the risk results are normalized with a score from 1 to 5. Through this process a risk value is obtained for each direct and indirect threat selected by the user. The user's global risk value will be the weighted average of all the threats' risk score selected by the user.

4.3.1.3 Interdependencies

Now that each user has its own global risk score, the subsequent step involves assessing the interdependencies among these users.

The first step to structure interdependencies consists in the selection, for each region, of the NACE code of the sectors on which the user considers itself dependent.

When a user designates a sector as their provider for a specific region, they are effectively selecting all users who have similarly chosen to be affiliated with that sector during their subscription. Subsequently, a notification is dispatched to these selected users.

Thereafter the user must declare for each supplier the dependency level, giving a score from 1 to 5. The dependency metric is defined by the user's the maximum tolerable period of disruption (MTPD) after the unavailability of a supplier sector:

1. very low dependency - in case of supplier disruption business continuity is guaranteed for more than 1 month.
2. low dependency - in case of supplier disruption business continuity is guaranteed from 2 weeks to 1 month.
3. medium dependency - in case of supplier disruption business continuity is guaranteed from 5 days to 2 weeks.
4. high dependency - in case of supplier disruption business continuity is guaranteed from 1 day to 5 days.
5. critical dependency level - in case of supplier disruption business continuity is guaranteed for less than 1 day.

Meanwhile, the supplier retains the discretion to either accept or decline their role as the user's supplier. In the event of an affirmative response, communication will be initiated between the user and the supplier. At that point, the user will be able to see the supplier that has just been placed within its supplier list instead of a selected NACE code. The NACE code is visible only if the supplier doesn't accept the invitation or if no company belonging to that sector is using the Tool. The sector risk score is calculated as the average risk score of all the risk scores of the users affiliated to that sector.

4.3.2 What if analysis

Once the global risk score and the extent of dependency have been determined, it becomes feasible to compute the What-if index.

This index is generated by the product between the dependency level, previously defined by the user, and the risk score of the supplier, normalized with a score from 1 to 5.

In case the infrastructure does not have a connection with a specific supplier for a sector, the risk score of the supplier is calculated as the average of the risk indices of the users belonging to that sector in that region. On the other hand, for the suppliers with whom the user has made a contact, it is possible to see who their suppliers (suppliers of suppliers) are.

This process is not limited solely to companies reaching out to suppliers; it is also applicable in the reverse scenario, where suppliers can initiate connections with companies. Once a supplier has received and confirmed a connection request from a company, they will gain access to information detailing the extent of dependency the specific customer places on the supplier.

The following image shows an example of the What-If Cascading effect tree. With this function the user will be able to analyse its own dependencies by viewing direct suppliers (where a contact is established) or supplier sectors. Each of them will provide, where available, their risk score and will be organized depending on the respective Dependency Level. Furthermore, the user will be able to overview a second level of the supply chain showing the suppliers of its suppliers. This function is available only if a contact is established. Finally, if the user accepted any contact from its clients as a provider, it will be able to see also his direct clients' dependency tree.

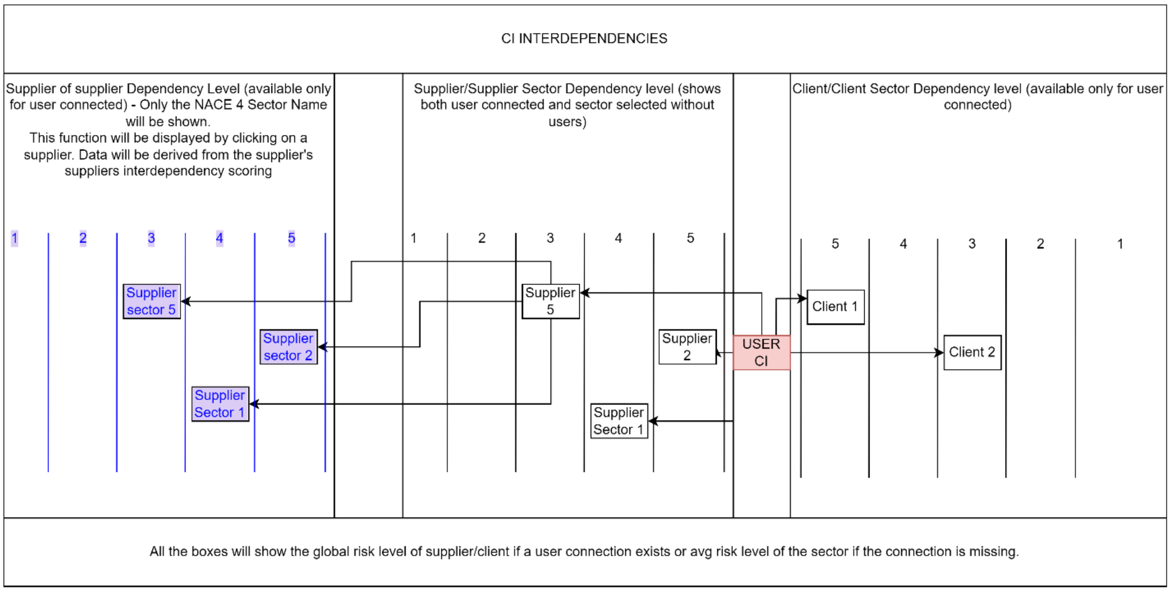

Figure 31 CI interdependencies

In case the risk score of an entire sector or a supplier is missing, but the user is aware of it, it is given the possibility to enter this value, scoring it manually from 1 to 5.

4.4 Macroeconomic scenarios

In principle, all CI providers, including water, energy, transport, and health sectors, are confronted with a dynamic cost-benefit optimization problem concerning the infrastructure capital that is necessary to cope with anticipated future demand. The prediction of future demand needs to account both for economic growth and structural changes in the economy. For example, the electricity sector needs to cope with future demand due to the growth of e-mobility and the increased use of heat pumps.

In addition, the capital stock needs to be sufficient to account for demand shocks due to extreme events. In the case of the Covid-19 pandemic, the health care sector had to be large enough to cope with the pandemic. In the case of climate change effects such as droughts, the water sector and its infrastructures need to meet increased demand from all water-demanding sectors; furthermore, the timely distribution of water has to be secured.

This second category of precautionary infrastructure investments [66] needs to account for the frequency and severity of extreme events, including infrastructure capacity, robustness, and resilience to different kinds of shocks. The costs of building up and maintaining infrastructure need to be weighed against the costs of under-provision of critical infrastructure in the case of extreme events. In many cases, the CI operator costs, and societal costs may diverge. For example, the CI costs of under-provision in the case of a disaster may be lower than the costs of building up and maintaining infrastructure, while societal costs of under-provision may be significantly higher. In other cases, the costs of under-provision could also be dramatic for the CI itself.

In addition to maintaining an optimal infrastructure capital stock, the speed of building up sufficient infrastructure needs to be considered. In the case of the Covid-19 pandemic, the rapid scaling and re-purposing of hospital beds is an illustrative example. If the critical infrastructure stock can be increased rapidly, then the need to make precautionary infrastructure investments is lower than in a case where the infrastructure cannot be built up or repurposed quickly. Usually, it takes time to install new capital. This is particularly relevant in the case of distribution grids such as for water or electricity supply. In the case of the electricity grid, not only the construction time has to be considered, but time-consuming approval procedures can also play a role here.

Finally, not only infrastructure capital, but also inventories need to be large enough to cope with shocks. In the case of production outages, for example, stocks help to prevent interruption in supply chains. Like in the case of infrastructure, the costs of keeping inventories should be weighed against the costs of not having sufficient inventories in the case of extreme events.

Our (IHS) sector-disaggregated agent-based macroeconomic model can account for critical infrastructures and related capital stocks as sectors at the NACE level of the Figaro input-output tables (Full international and global accounts for research in input-output Analysis) provided by Eurostat. Production of each sector is modelled as a Leontief production function with labour, capital, and intermediate inputs, which implies that there is no substitution between these inputs. However, currently, only the sum of intermediate inputs as an aggregate is a limiting factor. The firm agents' production functions thus need to be adapted to sector-specific characteristics regarding critical inputs such as energy, given that a critical sector input usually cannot be easily replaced by another, if at all. Electricity input in an industry sector mostly cannot be replaced by other forms of energy or other goods, for example. Therefore, a critical input constraint for production will lead to reduced macroeconomic output. This assumption is a plausible approximation for many cases of critical infrastructures.

For the purpose of planning their production and consumption, respectively, companies and households generate forecasts of general economic growth. To do so, the economic agents are currently forming expectations based on autoregressive models of lag order one (AR1) [67]. This means that they base their forecasts on past observations of the relevant variable (i.e., GDP growth and inflation), only taking the previous period into account. The forecasting period for the macroeconomic model is divided into quarterly timesteps and is up to three years, after which prediction uncertainty becomes too large. Hence, the agent-based model applied here does not contain an explicit long-term equilibrium such as computable general equilibrium (CGE) models. On the other hand, ABMs are much more detailed since they contain a very large number of individual firms and households.

Our macroeconomic simulations include a baseline scenario with simple agents without forward-looking expectations where capital stock is not increased, which will be contrasted with a scenario where critical infrastructures are forward-looking and have built up sufficient capital stock to cope with a shock. The assumption is that all CI providers will have adopted an adequate demand prediction and management Tool. In both scenarios, the economy will be shocked with an extreme event, such as an event that leads to a rapid demand increase. We derive macroeconomic indicators, including GDP and unemployment rate, among others, for each scenario. The difference between the baseline scenario and the anticipated shock scenarios will help to assess the overall macroeconomic benefits of having implemented a respective demand prediction Tool in critical infrastructure sectors and we contrast these with the costs of not having implemented such a Tool, which could be derived from the results of the impact assessment for selected CIs carried out with the Risk Management Tool displayed in Section 4.3 above and scaled to the sectoral level for macroeconomic modelling purposes. In addition to the extreme scenarios, we will also test intermediate scenarios to assess the up-scaling impact of the demand prediction Tool. These scenarios can potentially also relate back to the five global impact severity metric scores used in the Risk Management Tool, ranging from very low/NA impact scenarios to critical impact scenarios.

5 Pilot trials execution (feasibility analysis)

Deliverable D5.1 Demand prediction and management [1] described the pilot trials execution planning. This planning is made of three phases, starting with Phase 0 that is concluded on M12. The results of this phase have been described in Section 3. The intended proofs of concepts have been carried out by domain and they show the prediction capabilities of the proposed methods.

5.1 Description of piloting activities

5.1.1 Energy

For the energy demand prediction pilot, we have first constructed and fine-tuned the models to match the performance of the state-of-the-art (pilot 0 - TRL5) on existing public datasets. Extensive details on the performance of each model are available in Energy domain.

As for the piloting plan described in D5.1 [1], the next steps in the pilot's execution will be:

- We will apply mobility data into the pipeline, where we will assess its impact on performance and ability to predict demand in cases of rapid changes in consumption and compare it on existing production datasets collected in an offline manner. This data will be collected and provided by ELES and EKC (pilot 1 –TRL6, M23).
- We will refine the developed approaches and integrate it with pilot's legacy systems, as well as explore its effectiveness when deployed into a production-like environment (pilot 2 – TRL7, M34).

5.1.2 Transport

For the transport demand pilot, we have designed a method that enables the creation of a model able to predict the passenger's arrival to the transportation network and the trip distribution between stations. Details on the method and the results obtained validating it with CIs dataset are available in Section 3.2.

As for the piloting plan described in D5.1 [1], the next steps in the pilot's execution will be:

- Phase 1 (M23): The Tool will be deployed in transport CIs to demonstrate the basic functionalities.
- Phase 2 (M34): CIs will pilot the Tool in their operational environment.

5.1.3 Health

For the health sector pilot, we have proposed algorithms that will be used to predict the impact of a pandemic event or its countermeasures (e.g.: vaccines, quarantines) on the effort needed, in terms of staff or human workforce, and their resources in order to effectively provided its services to the regional healthcare ecosystem. Details of this analysis are available in Section 3.3.

As for the piloting plan described in D5.1 [1], the next steps in the pilot's execution will be:

- Phase 1 (M23): The Tool will be deployed in CI to demonstrate the basic functionalities, in terms of prediction.
- Phase 2 (M34): CIs will pilot the Tool in their operational environment.

5.1.4 Water

For the water distribution sector, we have studied the different solutions in the state-of-the-art for water consumption forecasting, compare them and select the most favourable for our scenario. The comparation of the different models and the study of the water consumption distribution during time are available in Section 3.4.

As for the piloting plan described in D5.1 [1], the next steps in the pilot's execution will be:

- Phase 1 (M23): The Tool will be deployed in water CIs to demonstrate the basic functionalities.
- Phase 2 (M34): Water CIs will pilot the Tool in their operational environment.

6 Conclusions

This document has presented the status of the Demand Prediction and Management Tool after the first development cycle to reach TRL5. The overall architecture of the tool has been presented, as well as the main components it integrates. The models for demand prediction in the different domains have been detailed as well, together with the initial proofs of concept that have been carried out to validate their performance, thus demonstrating the feasibility of the approaches in a lab environment.

The Tool architecture presents a flexible solution that will be adapted to meet the requirements of the individual Critical Infrastructure domains. Those domains have each own requirements and needs of forecasting unique variables at different speeds. Both the back- and front-ends have been designed with those requirements as a guideline. The back-end is ready to handle the prediction models lifecycle, from training to prediction. Also, front-end is prepared to show different views customized to each specific domain.

The approaches followed to train the prediction models vary among domains, as each shows different features:

- **Energy**: the focus in this sector is forecasting energy demand, validated with data from Slovenia's network. To accomplish this, a range of state-of-the-art methods has been put to the test, and among them, transformer-based models have emerged as the top performers, offering promising results in energy demand prediction.
- **Transportation**: the objective is to predict how travellers utilize the transportation network. This entails understanding the profiles of passenger arrivals at stations over time and constructing a model to estimate the origin-destination matrix, which reveals travel patterns between different locations. This model possesses adaptability, making it easily applicable to new infrastructure provided that the necessary data is available.
- **Health**: the primary prediction objective are the physical resources such as disposable protective equipment, medical supplies, drugs, and the workforce categorized by roles such as doctors, nurses, administrators, and more. Multiple data sources contribute to this endeavour, with the data often requiring manual input due to its diverse origins and formats, so they need thorough data preparation. In this context, LSTM networks are suggested as a potential solution, building upon their successful application in other sectors for accurate predictions.
- **Water**: the goal is to forecast water demand at the distribution network level. This prediction task is notably characterized by strong seasonality patterns, although it has been altered by the behavioural changes during and after the pandemic. To address this challenge, various predictive models have been put through rigorous testing, and it has been determined that LSTM models consistently deliver the most accurate results for water demand forecasting.

In addition to the prediction models and methods, a what-if analysis tool has been introduced, which takes various factors into consideration. These factors encompass pandemics, where the main risk to critical infrastructures lies in a shortage of manpower, with certain sectors being more severely affected. Furthermore, climate change introduces a heightened frequency of extreme weather events, significantly influencing sectors like logistics. Another critical aspect is between domains, as critical infrastructures form a network of interconnected dependencies; thus, failures in one sector have ripple effects on others. These risks undergo evaluation across diverse domains and regions to conduct a comprehensive risk analysis. The resulting risk indexes are then employed in conducting what-if analyses, revealing potential cascading effects among various critical infrastructures.

The second development cycle starting in M13 will upgrade the individual components by M16 and integrate them in the Tool by M20. This will allow deploying and demonstrating the tool in the partners (Critical Infrastructure) premises by M23, thus achieving TRL6.

References

[1] SUNRISE, “D5.1 Demand prediction and management,” 2023.

[2] European Union, “https://ec.europa.eu/,” [Online]. Available: https://ec.europa.eu/research/participants/data/ref/h2020/wp/2014_2015/annexes/h2020-wp1415-annex-g-trl_en.pdf.

[3] G. Tsakirakis, “SUNRISE. D3.1 - Requirements and designs,” 2023.

[4] “Auth0,” [Online]. Available: https://auth0.com/. [Accessed 22 09 2023].

[5] A. Z. a. M. C. a. L. Z. a. Q. Xu, “Are Transformers Effective for Time Series Forecasting?,” 2022.

[6] “Web Page Name (i.e. Deliverables),” [Online]. Available: web url (i.e. http://www.witdom.eu/deliverables). [Accessed day acceessed (i.e. 06) month accessed (i.e. 11) year accessed (i.e. 2017)].

[7] J. H. a. F. L. a. S. G. P. a. T. N. a. L. T. a. G. R. a. T. V. P. a. M. P. a. A. S. a. N. H. a. M. D. a. J. K. a. D. Ba, “{Darts: User-Friendly Modern Machine Learning for Time Series,” Journal of Machine Learning Research, vol. 23, pp. 1-6, 2022.

[8] J. Beitner, “https://towardsdatascience.com/introducing-pytorch-forecasting-64de99b9ef46,” Medium, 12 September 2020. [Online]. Available: https://towardsdatascience.com/introducing-pytorch-forecasting-64de99b9ef46. [Accessed 6 July 2023].

[9] C. O. K. G. O. B. N. G. R. F. M. C. M. &. D. A. Challu, “NHITS: Neural Hierarchical Interpolation for Time Series Forecasting,” in Proceedings of 37th AAAI Conference on Artificial Intelligence, Washington , 2023.

[10] S. Ö. A. N. L. T. P. Bryan Lim, “Temporal Fusion Transformers for interpretable multi-horizon time series forecasting,,” International Journal of Forecasting, , vol. 37, no. 4, pp. 1748-1764, 2021.

[11] D. S. Sobrino, J. Andión, J. C. Dueñas, J. M. d. Álamo and F. Cuadrado, “Model Generation and Origin - Destination(OD) Matrix Forecasting for metro systems,” in 2023 7th International Young Engineers Forum (YEF-ECE), Caparica / Lisbon, Portugal, 2023.

[12] D. G. Manuel, C. van Walraven and A. J. Forster, “A commentary on the value of hospital data for COVID-19 pandemic surveillance and planning,” Int J Popul Data Sci, vol. 5, p. 1393, March 2021.

[13] N. Shaukat, D. M. Ali and J. Razzak, “Physical and mental health impacts of COVID-19 on healthcare workers: a scoping review,” Int J Emerg Med, vol. 13, p. 40, July 2020.

[14] Z. Al-Aly, Y. Xie and B. Bowe, “High-dimensional characterization of post-acute sequelae of COVID-19,” Nature, vol. 594, pp. 259-264, 2021.

[15] H. Poustchi and et al., “SARS-CoV-2 antibody seroprevalence in the general population and high-risk occupational groups across 18 cities in Iran: a population-based cross-sectional study,” The Lancet Infectious Diseases, vol. 21, pp. 473-481, 2021.

[16] C. L. Hodgson and et al., “The impact of COVID-19 critical illness on new disability, functional outcomes and return to work at 6 months: a prospective cohort study,” Critical Care, vol. 25, 2021.

[17] Z. Hu, Q. Ge, S. Li, L. Jin and M. Xiong, “Artificial Intelligence Forecasting of COVID-19 in China,” ArXiv, 2020.

[18] Z. Ceylan, “Estimation of COVID-19 prevalence in Italy, Spain, and France,” Sci. Total Environ., vol. 138817, 2020.

[19] O. Torrealba-Rodriguez, R. A. Conde-Gutiérrez and A. L. Hernández-Javier, “Modeling and prediction of COVID-19 in Mexico applying mathematical and computational models,” Chaos Solitons Fract., p. 109946, 2020.

[20] D. Parbat and M. Chakraborty, “A Python based support vector regression model for prediction of Covid19 cases in India,” Chaos Solitons Fract., p. 109942, 2020.

[21] I. C. O. V. I. D.-1. health service utilization forecasting Team and C. J. Murray, “Forecasting the impact of the first wave of the COVID-19 pandemic on hospital demand and deaths for the USA and European economic area countries,” medRxiv, 2020.

[22] I. F. Miller and et al., “Disease and healthcare burden of COVID-19 in the United States,” Nature Medicine, vol. 26, pp. 1212-1217, 2020.

[23] E. Koç and M. Türkoğlu, “Forecasting of medical equipment demand and outbreak spreading based on deep long short-term memory network: the COVID-19 pandemic in Turkey,” Signal Image Video Process., vol. 16, pp. 613-621, 2022.

[24] M. Ghiassi, F. Fa'Al and A. Abrishamchi, “Large metropolitan water demand forecasting using DAN2, FTDNN, and KNN models: A case study of the city of Tehran, Iran,” Urban Water Journal, vol. 14, p. 655–659, 2017.

[25] L. Deng, X. Chang and P. Wang, “Daily Water Demand Prediction Driven by Multi-source Data,” Procedia Computer Science, vol. 208, pp. 128-135, 2022.

[26] A. Niknam, H. K. Zare, H. Hosseininasab, A. Mostafaeipour and M. Herrera, “A Critical Review of Short-Term Water Demand Forecasting Tools—What Method Should I Use?,” Sustainability, vol. 14, p. 5412, 2022.

[27] T. A. Peterson K. Ozili, Spillover of COVID-19: Impact on the Global Economy, IGI Global, 2023.

[28] “Supply chain disruptions and the effects on the global economy.,” [Online]. Available: https://www.ecb.europa.eu/pub/economic-bulletin/focus/2022/html/ecb.ebbox202108_01~e8ceebe51f.en.html. [Accessed 22 09 2023].

[29] “Coronavirus controls pushing China’s southern container ports to limit,” [Online]. Available: https://www.scmp.com/economy/china-economy/article/3136629/coronavirus-controls-pushing-chinas-southern-container-ports. [Accessed 22 09 2023].

[30] “La catena di approvvigionamento globale si è rotta. Le conseguenze per imprese e consumatori,” [Online]. Available: https://www.aduc.it/articolo/catena+approvvigionamento+globale+si+rotta_34283.php. [Accessed 22 09 2023].

[31] X. Sun, S. Wandelt and A. Zhang, “Aviation under COVID-19 Pandemic: Status Quo and How to Proceed Further?,” SSRN, p. 31, 2022.

[32] "COVID-19 has been an unprecedented shock," [Online]. Available: https://www.iata.org/en/iata-repository/publications/economic-reports/covid-19-has-been-an-unprecedented-shock/. [Accessed 22 09 2023].

[33] L. Ecke, M. Magdolen, B. Chlond and P. Vortisch, "How the COVID-19 pandemic changes daily commuting routines – Insights from the German Mobility Panel," Case Studies on Transport Policy, vol. 10, no. 4, pp. 2175-2182, 2022.

[34] "COVID-19 Impact on the Road Transport Industry," [Online]. Available: https://www.itf-oecd.org/sites/default/files/docs/covid-19_impact_on_the_road_transport_industry_-_june_2021.pdf. [Accessed 22 09 2023].

[35] "COVID-19 Impact on the," [Online]. Available: https://www.itf-oecd.org/sites/default/files/docs/_iru_report.pdf. [Accessed 22 09 2023].

[36] "Halesowen haulage firm collapse blamed on fuel costs," [Online]. Available: https://www.bbc.com/news/uk-england-birmingham-61744651. [Accessed 22 09 2023].

[37] "The impact of the COVID-19 pandemic on the behaviour of bike sharing users," [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2210670722003237. [Accessed 22 09 2023].

[38] "DECRETO DEL PRESIDENTE DEL CONSIGLIO DEI MINISTRI 3 novembre 2020," [Online]. Available: https://www.gazzettaufficiale.it/eli/id/2020/11/04/20A06109/sg. [Accessed 22 09 2023].

[39] "How COVID-19 Medical Supply Shortages Led to Extraordinary Trade and Industrial Policy," [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8441910/. [Accessed 22 09 2023].

[40] "Medicine shortages in the EU: causes and solutions," [Online]. Available: https://www.europarl.europa.eu/news/en/headlines/society/20200709STO83006/medicine-shortages-in-the-eu-causes-and-solutions. [Accessed 22 09 2023].

[41] "Prada the latest fashion brand to make medical face masks," [Online]. Available: https://www.theguardian.com/fashion/2020/mar/24/prada-the-latest-fashion-brand-to-make-medical-face-masks. [Accessed 22 09 2023].

[42] ISS - Istituto Superiore di Sanità, "Rapporto ISS COVID-19 n. 10/2020 - Indicazioni ad interim su acqua e servizi igienici in relazione alla diffusione del virus SARS-CoV-2," 2020.

[43] "Covid-19 impact on electricity," [Online]. Available: https://www.iea.org/reports/covid-19-impact-on-electricity. [Accessed 22 09 2023].

[44] "Prezzi medi annuali dei carburanti e combustibili," [Online]. Available: https://dgsaie.mise.gov.it/prezzi-annuali-carburanti. [Accessed 22 09 2023].

[45] "The Russia-Saudi Arabia oil price war during the COVID-19 pandemic," [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8652835/. [Accessed 22 09 2023].

[46] "Covid-19 impact on renewable energy growth," [Online]. Available: https://www.iea.org/reports/renewable-energy-market-update/covid-19-impact-on-renewable-energy-growth. [Accessed 22 09 2023].

[47] "Lavoro, Istat: crollo occupazione da inizio pandemia, persi 945mila posti," [Online]. Available: https://www.ilsole24ore.com/art/lavoro-istat-crollo-occupazione-inizio-pandemia-persi-945mila-posti-AD7ePfVB. [Accessed 22 09 2023].

[48] "Is Childcare Affordable? POLICY BRIEF ON EMPLOYMENT, LABOUR AND SOCIAL AFFAIRS," [Online]. Available: https://www.oecd.org/els/family/OECD-Is-Childcare-Affordable.pdf. [Accessed 22 09 2023].

[49] "Public information on medicine shortages," [Online]. Available: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/availability-medicines/public-information-medicine-shortages. [Accessed 22 09 2023].

[50] "Pharmaceutical companies need to act quickly on their climate risk," [Online]. Available: https://cervest.earth/news/pharmaceutical-companies-need-to-act-quickly-on-their-climate-risk. [Accessed 22 09 2023].

[51] "The Effect of Climate Change on Medical Device Manufacturing," [Online]. Available: https://www.mddionline.com/manufacturing/effect-climate-change-medical-device-manufacturing. [Accessed 22 09 2023].

[52] "How can climate change impact the logistics of fresh produce?," [Online]. Available: https://www.girteka.eu/how-can-climate-change-impact-the-logistics-of-fresh-produce/. [Accessed 22 09 2023].

[53] "Will climate change dry up our supply chains?," [Online]. Available: https://biggerpicture.ft.com/climate/article/will-climate-change-dry-our-supply-chains. [Accessed 22 09 2023].

[54] "Climate Impacts on Logistics and Supply Chain," [Online]. Available: https://supplychaingamechanger.com/climate-impacts-on-logistics-and-supply-chain/. [Accessed 22 09 2023].

[55] "Hydropower in Europe: How climate change will impact the region's leading renewable energy source," [Online]. Available: https://www.nsenergybusiness.com/features/hydropower-europe/. [Accessed 22 09 2023].

[56] "European insights: hydropower, climate change and EU law," [Online]. Available: https://www.waterpowermagazine.com/features/featureeuropean-insights-hydropower-climate-change-and-eu-law-7502217/. [Accessed 22 09 2023].

[57] "Climate change and electricity production," [Online]. Available: https://joint-research-centre.ec.europa.eu/system/files/2020-09/04_pesetaiv_energy_supply_sc_august2020_en.pdf. [Accessed 22 09 2023].

[58] J. M. d. Alamo, "SUNRISE; D1.1 Local meetings with Critical Infrastructures," 2022.

[59] M. Tarman, "SUNRISE; D1.2 Pan-European meetings with CI stakeholders V1," 2023.

[60] "Supply chain disruptions and the effects on the global economy," [Online]. Available: https://www.ecb.europa.eu//pub/economic-bulletin/focus/2022/html/ecb.ebbox202108_01~e8ceebe51f.en.html. [Accessed 22 09 2023].

[61] “Rapporto ISS COVID-19 n. 10/2020 - Indicazioni ad interim su acqua e servizi igienici in relazione alla diffusione del virus SARS-CoV-2. Versione del 7 aprile 2020,” [Online]. Available: https://www.iss.it/rapporti-covid-19/-/asset_publisher/btw1J82wtYzH/content/rapporto-iss-covid-19-n.-10-2020-indicazioni-ad-interim-su-acqua-e-servizi-igienici-in-relazione-alla-diffusione-del-virus-sars-cov-2.-versione-del-7-aprile-2020. [Accessed 22 09 2023].

[62] R. Bardhan, T. Byrd and J. Boyd, “Workforce Management during the Time of COVID-19—Lessons Learned and Future Measures,” COVID, vol. 3, pp. 1-27, 2023.

[63] “Synthèse des indicateurs de suivi de l’épidémie COVID-19,” [Online]. Available: https://www.data.gouv.fr/fr/datasets/synthese-des-indicateurs-de-suivi-de-lepidemie-covid-19/. [Accessed 22 09 2023].

[64] “COVID-19 Tracking Germany,” [Online]. Available: https://www.kaggle.com/datasets/headsortails/covid19-tracking-germany?resource=download&select=covid_de.csv. [Accessed 22 09 2023].

[65] “COVID-19 dati regioni,” [Online]. Available: https://github.com/pcm-dpc/COVID-19/tree/master/dati-regioni. [Accessed 22 09 2023].

[66] “COVID 19 ccaa_covid19_datos_sanidad_nueva_serie.csv,” [Online]. Available: https://github.com/datadista/datasets/blob/master/COVID%2019/ccaa_covid19_datos_sanidad_nueva_serie.csv. [Accessed 22 09 2023].

[67] S. Fankhauser, J. B. Smith and R. S. J. Tol, “Weathering Climate Change: Some Simple Rules to Guide Adaptation Decisions,” Ecological Economics, vol. 30, pp. 67-78, 1999.

[68] S. Poledna, M. G. Miess, C. Hommes and K. Rabitsch, “Economic Forecasting with an Agent-Based Model,” European Economic Review, vol. 151, p. 104306, 2022.

Annex I: Training Guide and User Manual for DPM POC

A.1. Introduction

The Demand Prediction and Management (DPM) tool is going to be used in different sectors, as has been defined. These sectors are water, energy, health, transportation, and digital. This tool will incorporate state-of-the-art technologies and will use AI/ML and graph-based models where applicable in order to produce accurate forecasts and valuable insights for the involved CIs.

The forecasts and insights can be customised and parametrized in order to help with both short-term and long-term forecasts. The variables that are being forecasted are actually the critical resources of each of the involved CIs. In order to produce the outputs that will be visualised, the historical data of each of the CIs will be used. In addition, specific events will be used in cases where they are applicable, like Covid-19, lockdowns, climate changes, etc.

In this manual developed by SQD a first guide of the POC is showcased. In order to cover a large part of the identified requirements, a more generic POC has been developed since many of the requirements are common between the different sectors. The next iterations of the development will use this POC as a reference to build the remaining functionalities and logic. Based on that, this guide will be helpful for the CIs to understand the possibilities of the tool at this step.

A.2. Getting Started

It is considered that the POC is deployed for experimentation purposes and QA sessions. It can be accessed through the corresponding URL. In the example in Figure 32 the tool is running locally.

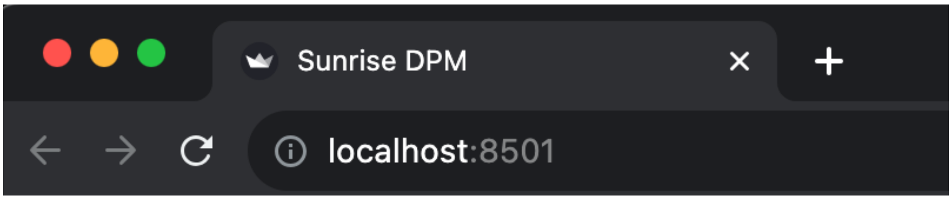

Figure 32: Local hosting of the POC app

No setup is required since the POC tool is already deployed, and no other actions for installation are required by the users. The only parts that are needed are the possible customizations, as they are described below.

The authentication mechanism has not been integrated yet, so for POC purposes, no credentials are needed in order to log in to the tool. When opening the app, the user is able to see some basic information about it (Figure 34).

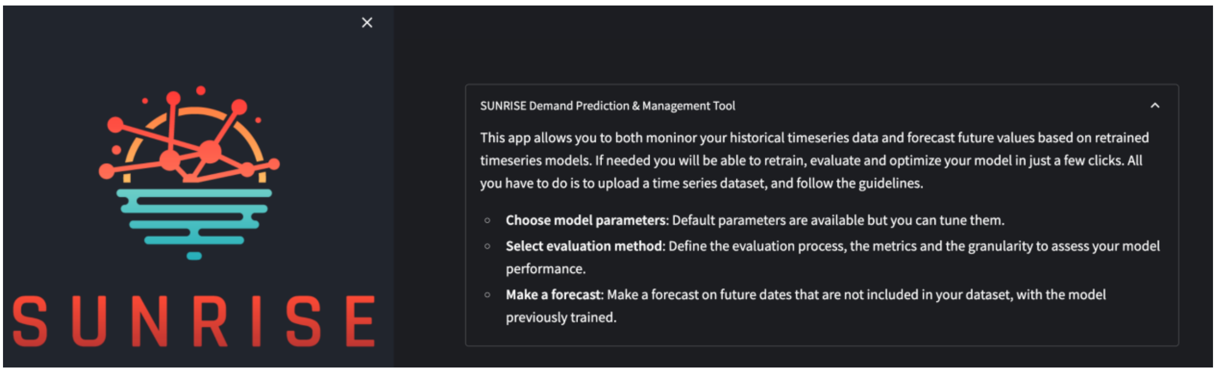

Figure 33: Initial screen of the POC app

A.3. User Interface (UI) Overview

In the UI there are different layouts and parts, as mentioned in this subsection. We can split it into two different parts. The sidebar, where the user can set their own parameters, has a part of the control and the main monitoring part of the app where the user is able to see the produced forecasts, outputs, and insights.

In the sidebar, four main tables are included: data, modelling, evaluation, and forecast.

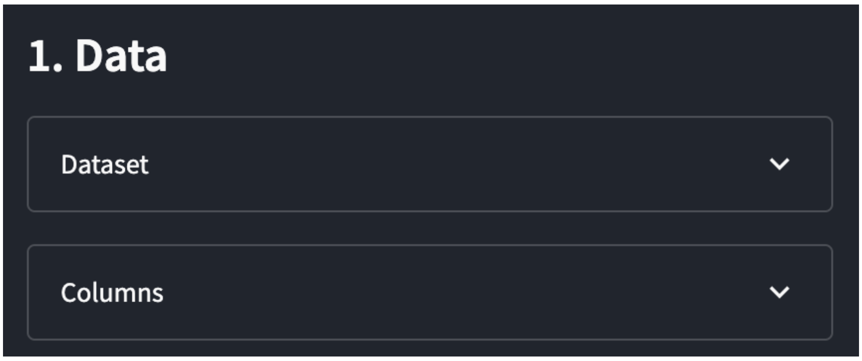

Figure 34: Data table in sidebar

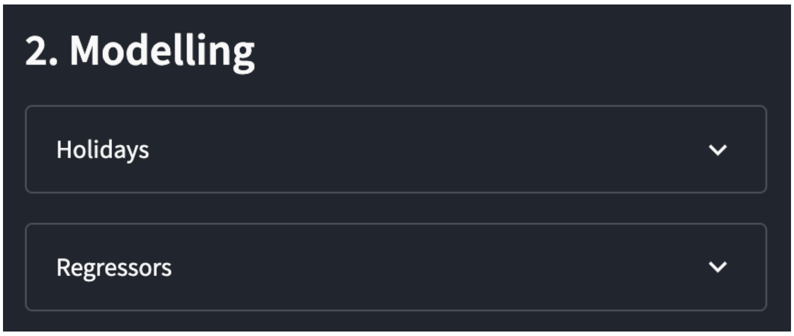

Figure 35: Modelling table in sidebar

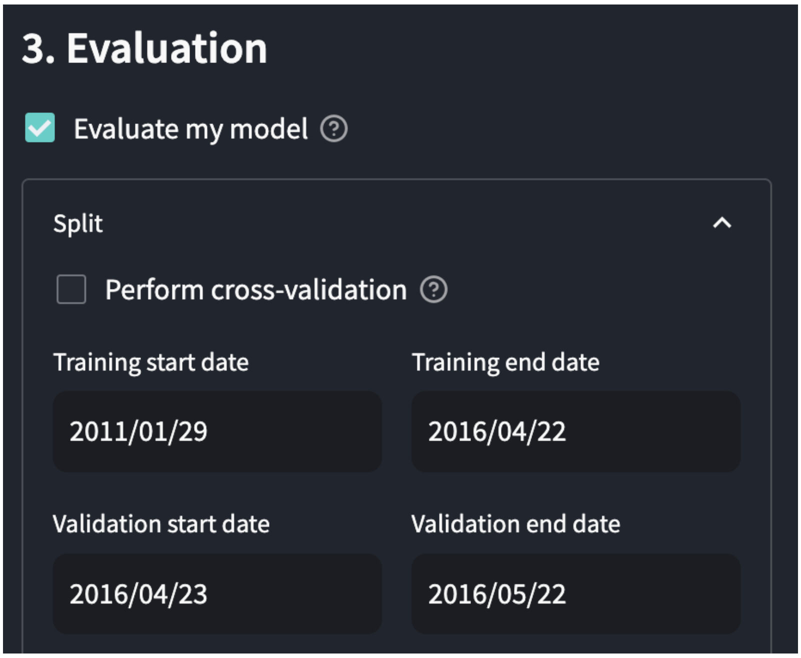

Figure 36: Evaluation table in sidebar

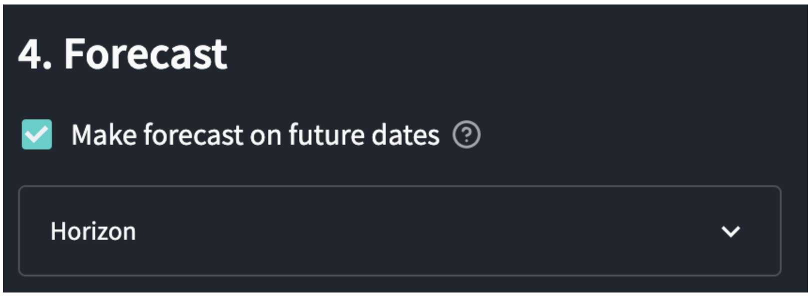

Figure 37: Forecast table in sidebar

On the monitoring part, two tabs are included: "Overview" and "Forecast".

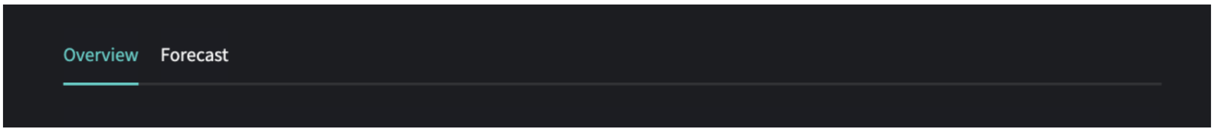

Figure 38: Overview and Forecast tabs in monitoring part

A.4. Basic Features in Sidebar

In this section, all the basic features are explained.

In the Data table of the sidebar, the user is able to select the data that needs to be loaded in order to use it for the forecasting session. This depends on the data that has already been included in the database. In the next iteration, data upload will also be effective.

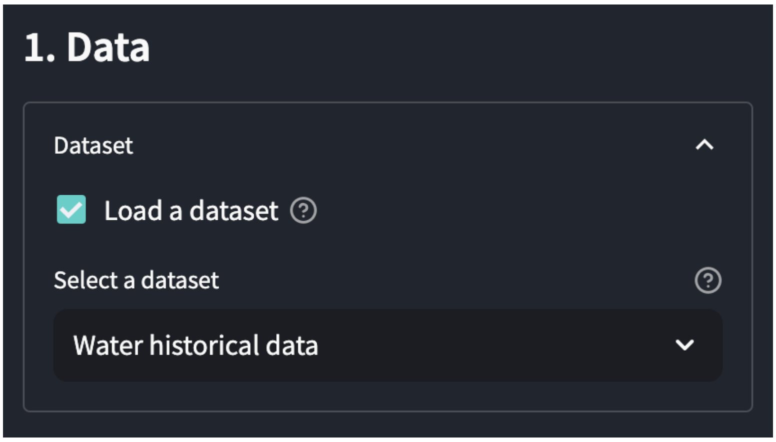

Figure 39: Loading dataset in Data table

The “columns” field in Data table helps user to validate that the correct main columns are being used by the system.

Figure 40: Validating selected columns from the dataset in Data table

The modelling table includes two main parts: the "Holidays" and "Regressors" tabs. In the Holidays tab, the country and the specific holidays of the country can be selected to be used to enhance the dataset. In addition, lockdown events can be selected to be used in the forecasting.

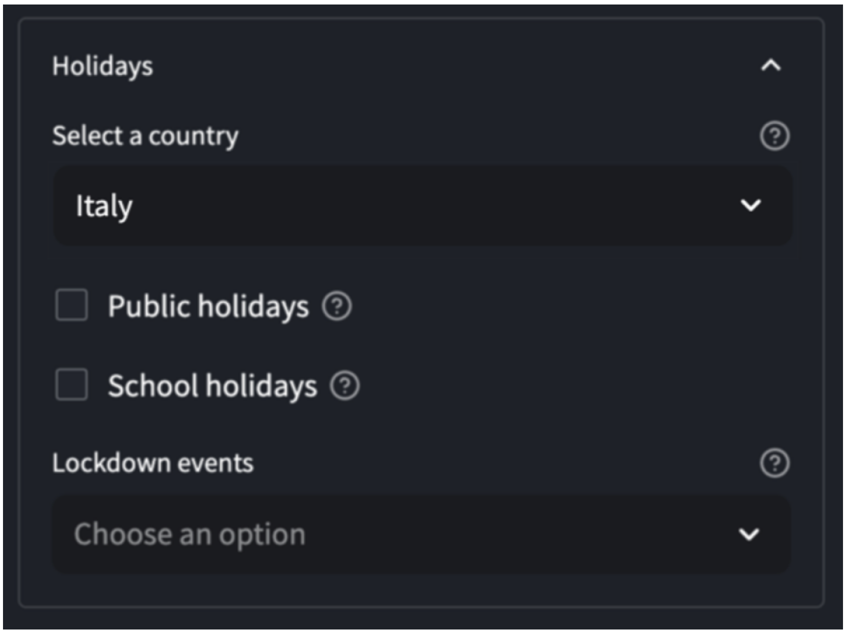

Figure 41: Holidays and lockdown events selection in Holidays tab in Modelling table

In the Regressors tab, other regressors can be used based on those saved in the database.

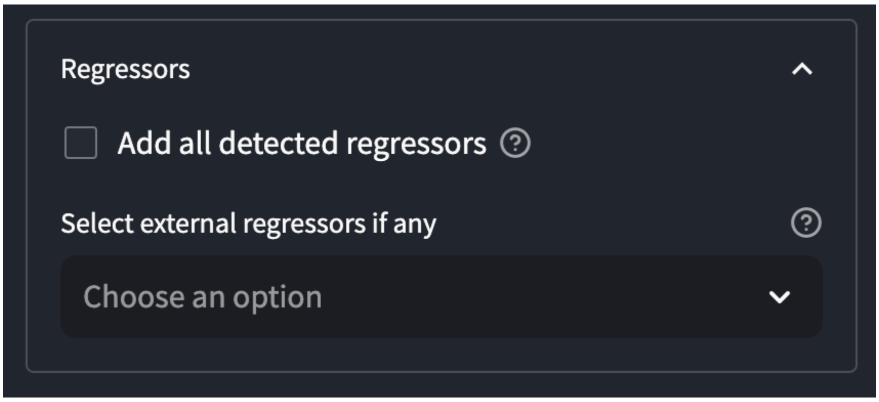

Figure 42: Other Regressors selection in Modelling table

In the “Evaluation” table, the user is able to select the desired metrics to be visualised and also set the scope.

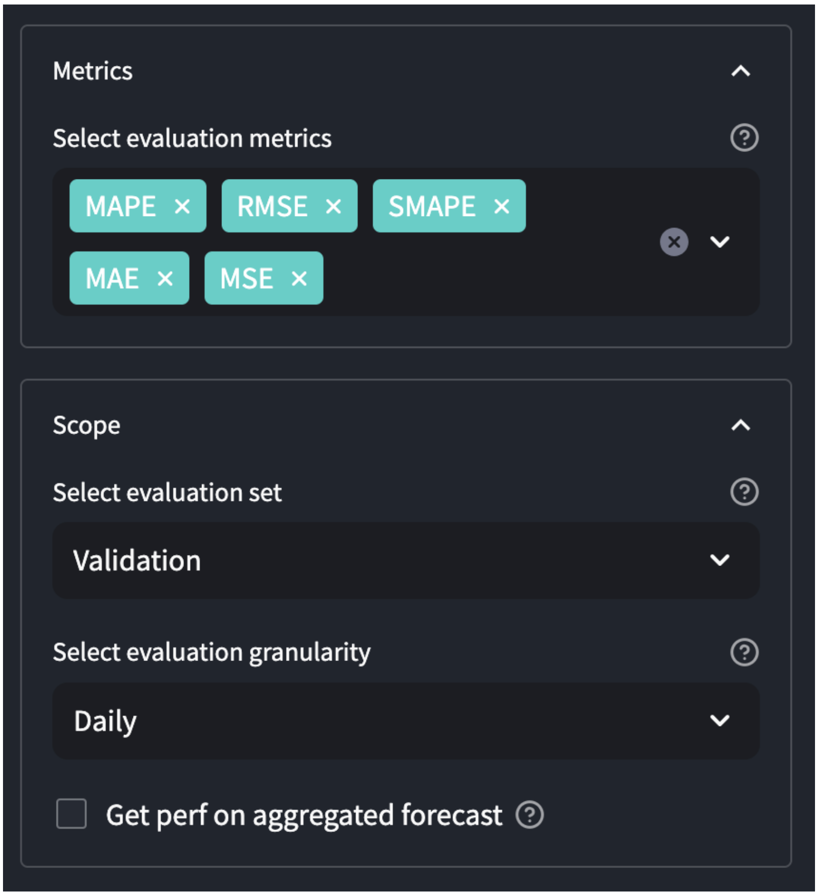

Figure 43: Metrics and scope in Evaluation table

In the evaluation set, the user can choose whether to evaluate the model on training data or validation data. You should look at validation data to assess model performance, but evaluation on training data can also be useful to detect overfitting.

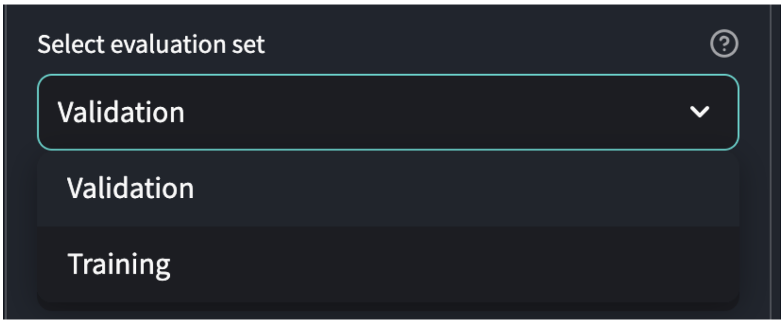

Figure 44: Select evaluation set in Evaluation table

In the field of evaluation granularity selection, the user can set the granularity at which predictions on the evaluation set will be averaged. If 'Global' is selected, the user can compute performance on the whole evaluation set.

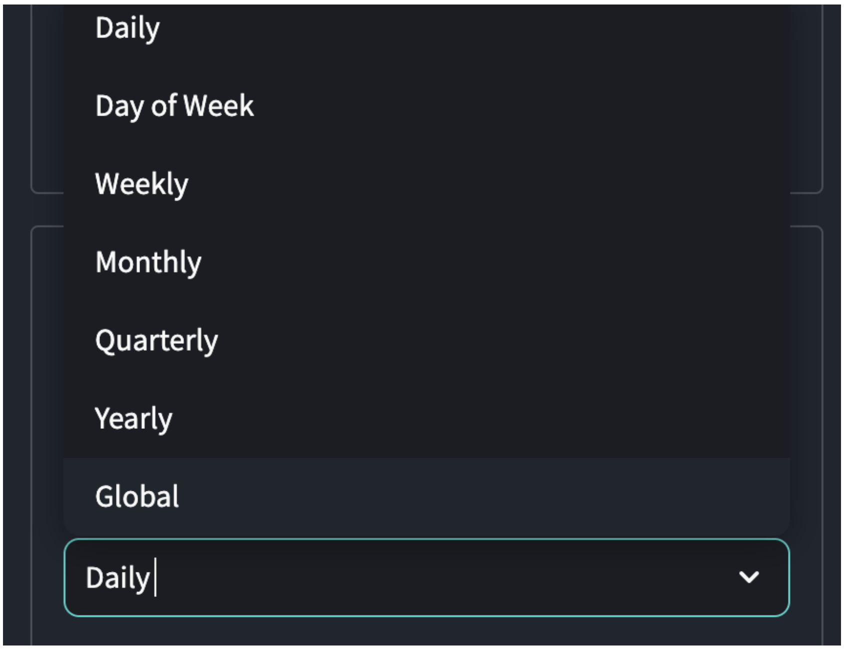

Figure 45: Select evaluation granularity in Evaluation table

In the “Forecast” table the user is able to select the horizon of the forecast. The horizon is the length of the period to forecast after the last date available in the input dataset.

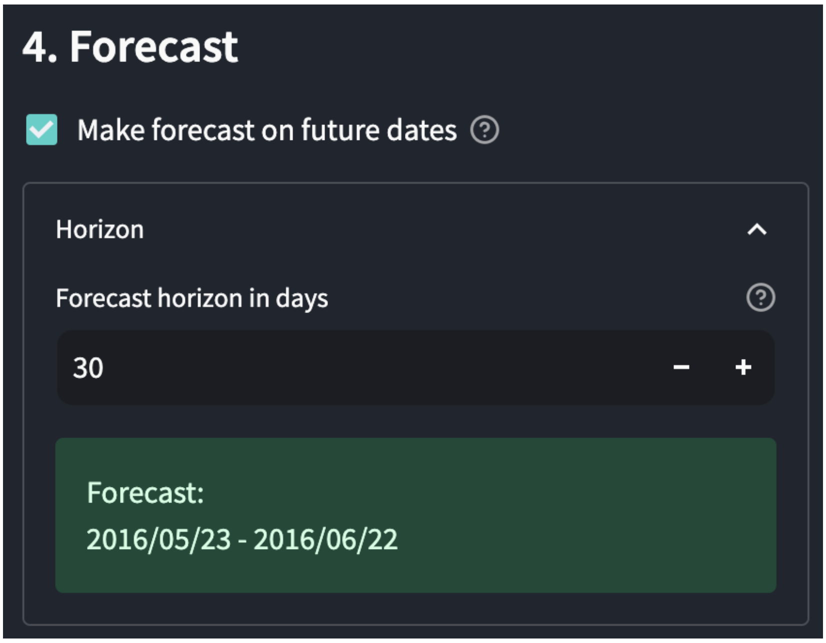

Figure 46: Forecast session in Forecast table

A.5. Basic Features in Monitoring part

When the “Overview” tab is selected in the monitoring part and the loaded dataset is chosen in the sidebar, the historical timeseries data is visualised.

Figure 47: Historical timeseries data visualisation

The user can zoom in, zoom out, autoscale, or select a specific time frame to visualise.

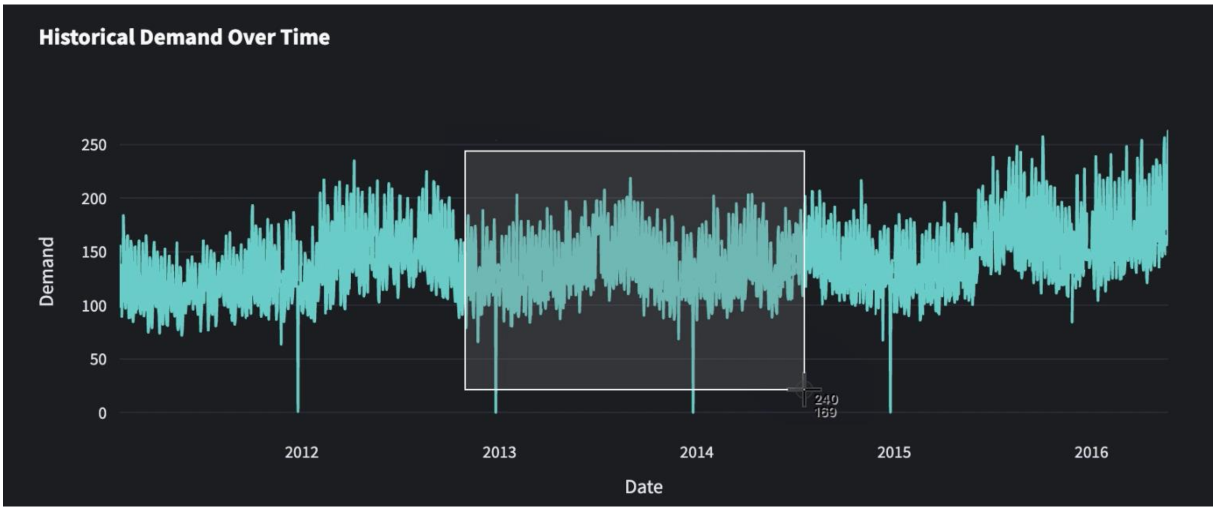

Figure 48: Timeframe selection in timeseries data visualisation

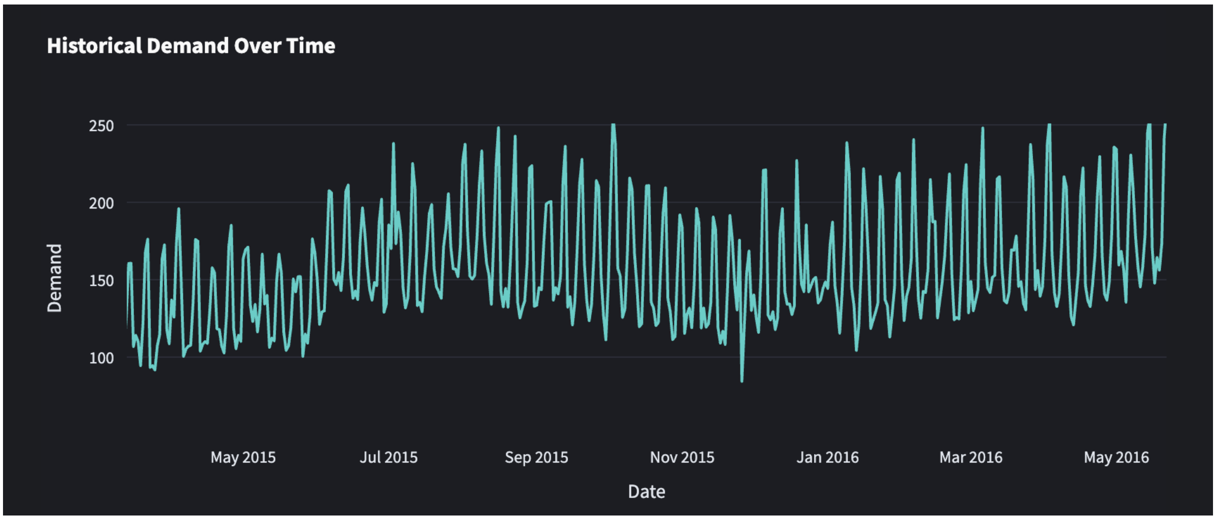

Figure 49: Timeframe selection and zoom-in in timeseries data visualisation

In the "Forecast" tab, all the corresponding results regarding the forecasting session are depicted. The first element that is depicted is the forecasted timeseries. The user is able to select different granularity levels according to the needed analysis. The cyan line shows the predictions made by the model on both training and validation periods. The blue shade around is an 80% uncertainty interval. The yellow points depict the actual values of the target during the training period. The red line is the trend estimated by the model, and the vertical lines show the changepoints at which this trend evolves.

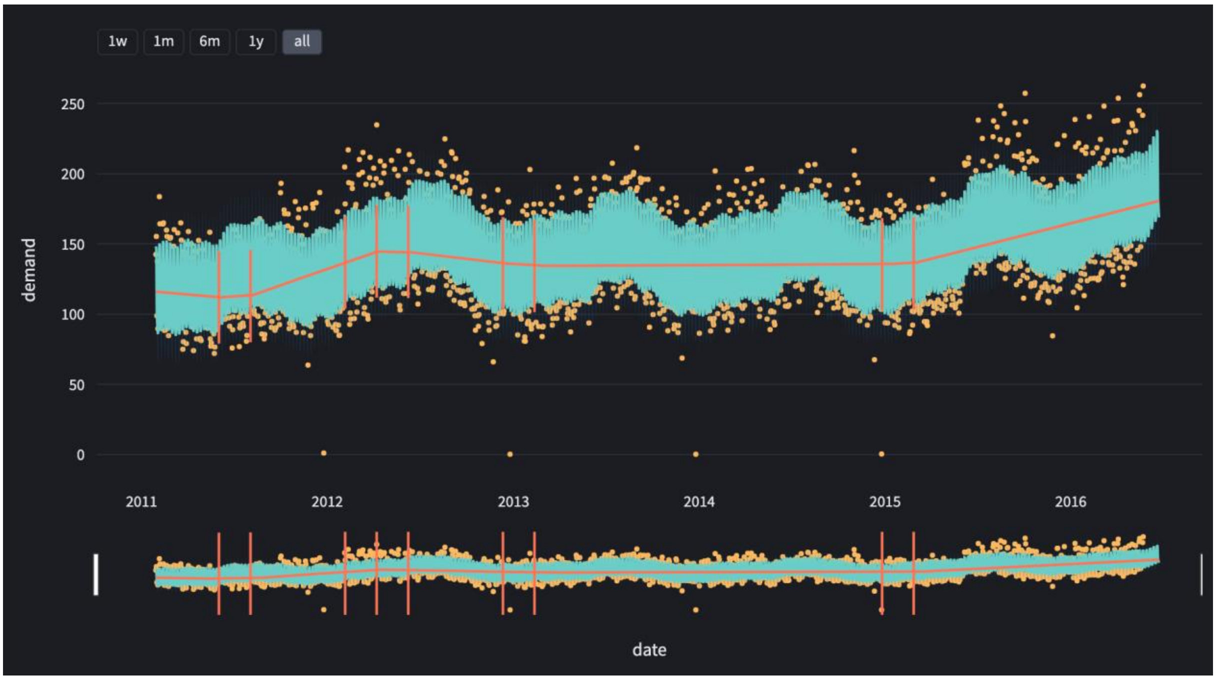

Figure 50: Forecasted timeseries (visualisation of the whole timeseries)

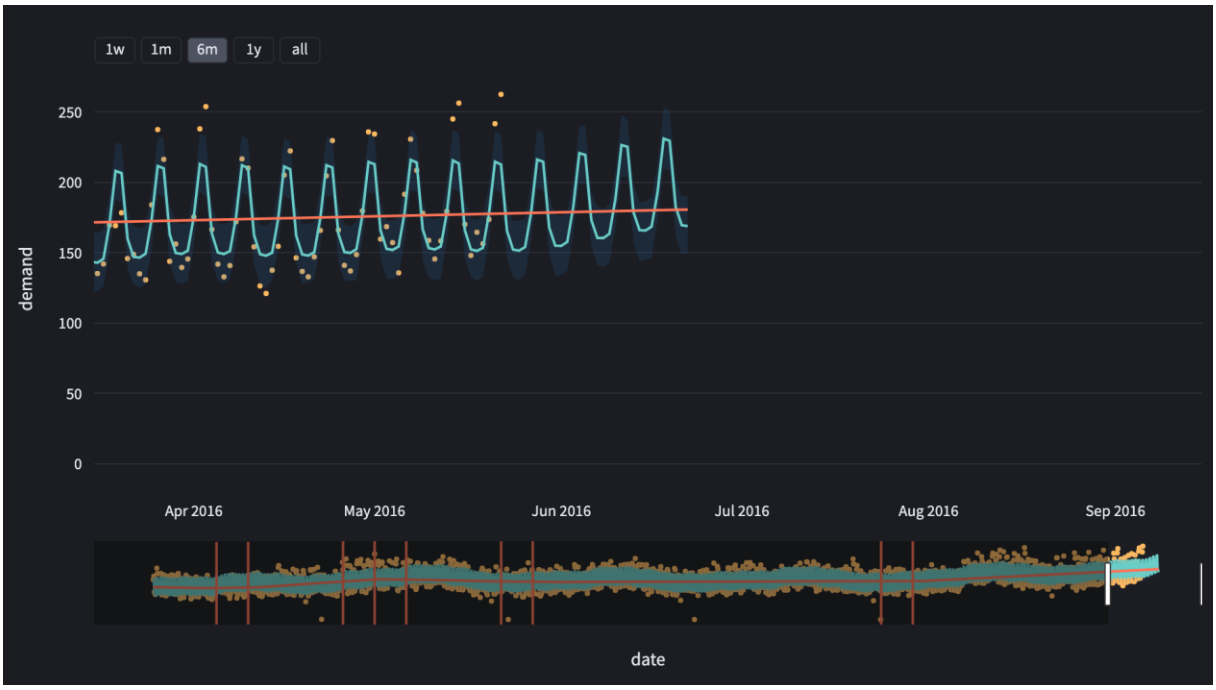

Figure 51: Forecasted timeseries (visualisation of the 6 months granularity)

The second element is the evaluation on the validation set. This part includes the performance metrics of the model, the error analysis, and the distribution of the errors. The meaning of the used metrics is also included in the app for quick reference.

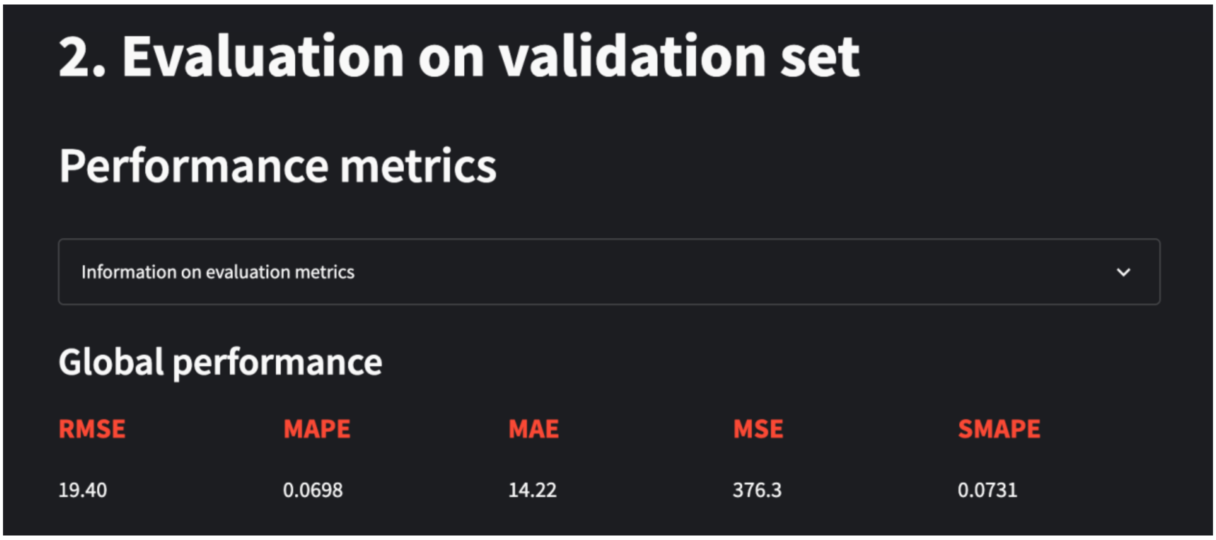

Figure 52: Evaluation on validation set

On the detailed evaluation metrics, the user is able to understand in which periods each one of the evaluation metrics produced better results and the opposite.

Figure 53: Detailed evaluation metrics

In the error analysis element, the user can see the depicted differences between the forecasted and the truth values of the model on the validation set.

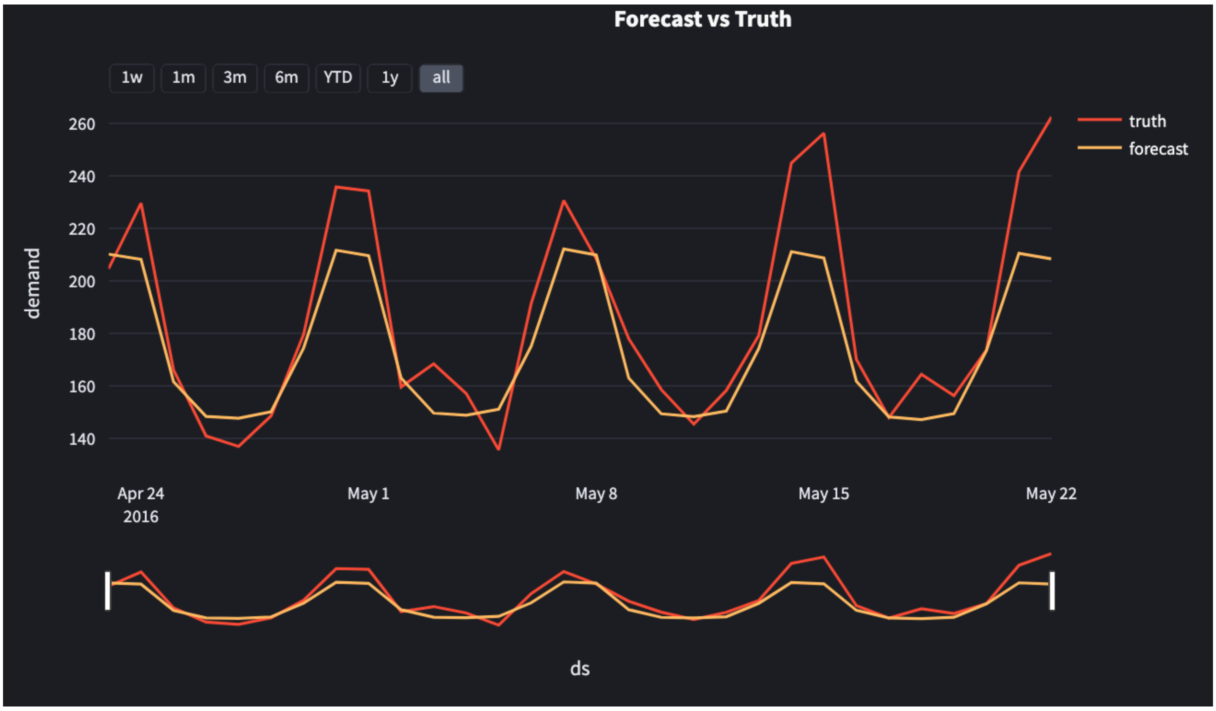

Figure 54: Truth vs forecasted values

In the third element of the “Forecast” the user is able to understand the different impact that each one of the used features (seasonality, trend, other factors such as pandemic, climate change, etc.) had on the forecasted values. In the “Local Impact” part, the user is able to choose more specific time periods in order to see the averaged contributions.

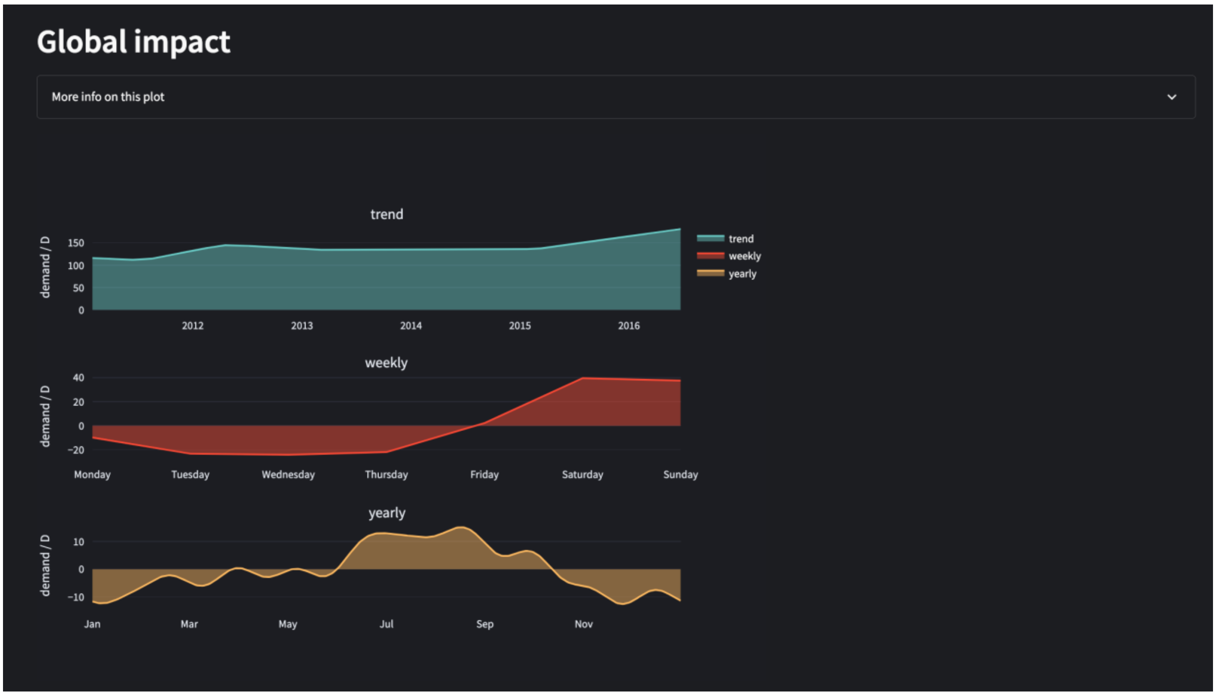

Figure 55: Global Impact

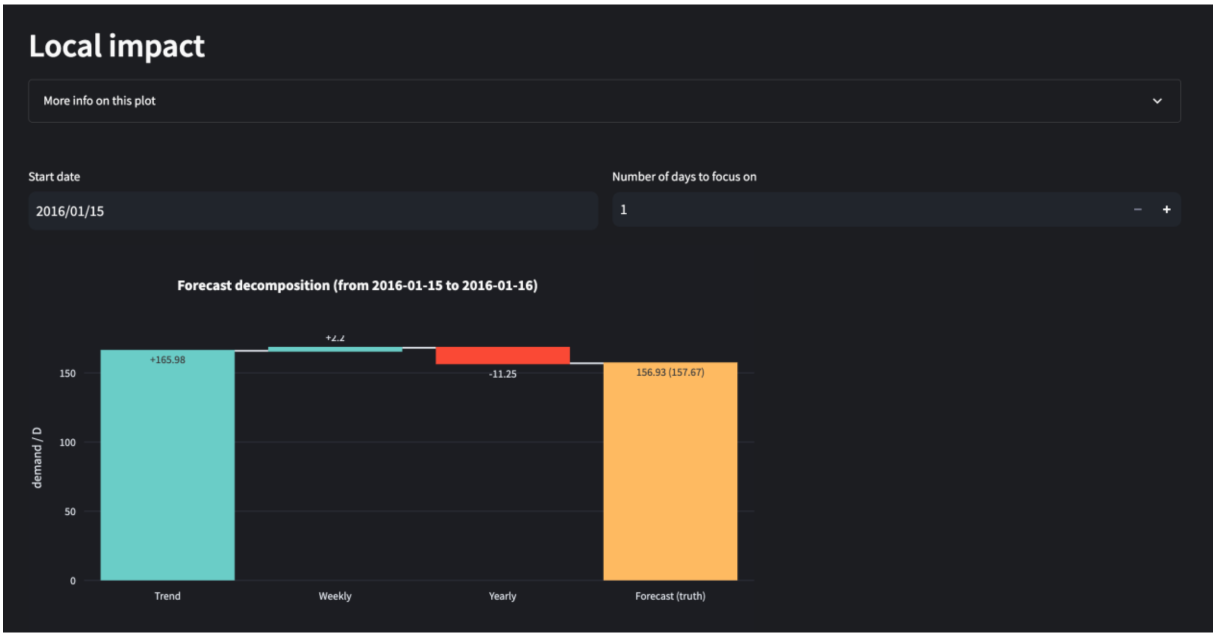

Figure 56: Local impact

In the fourth element, the user is able to see the depicted future forecast if the check box for future forecasts is selected in the sidebar. Through this, the user can see the future forecasts that are further from the validation set. The cyan line shows the forecasts made by the model for the period to be forecast. The blue shade is an 80% uncertainty interval. The red line is the trend estimated by the model.

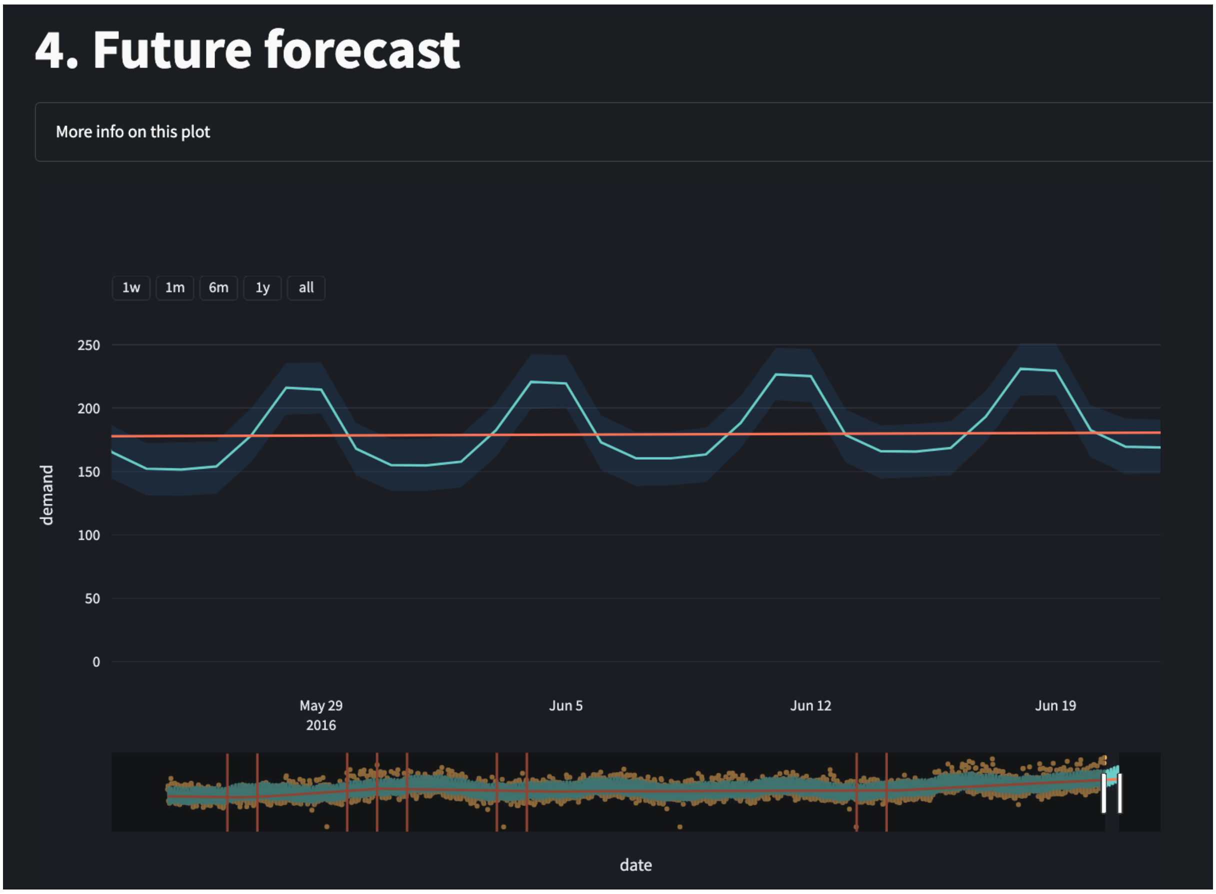

Figure 57: Future forecast

A.6. Example scenario flows

Here are represented two example scenario flows:

Scenario 1: Basic functionality of the tool

Steps

- **Demo:** Demand Prediction & Management Tool
- **Scenario 1:** Basic functionality of the tool
- **Step 1:** The tool loads the historical data
- **Step 2:** User can reads some basic guidelines of the tool
- **Step 3:** **U**ser selects and checks basic fields of the data to be used
- **Step 4:** User is able to see the visualization of the historical timeseries data
- **Step 5:** User is able to see the visualization of the forecasted timeseries data (on validation data)
- **Step 6:** The tool evaluates the used model and visualizes the performance metrics
- **Step 7:** The tool visualizes the Truth vs the Forecasted values
- **Step 8:** The tool visualizes the different impact that each one of the used features has on the predictions

Flow description

- **Data Loading:** Users start by loading historical data into the tool, the very foundation of their forecasting flow.
- **Guiding Insights:** Users are provided with valuable guidelines to help them navigate the tool effectively.
- **Data Selection:** Selecting relevant data fields ensures that forecasting is based on the most pertinent information.
- **Historical Data Visualization:** Users can explore past trends through intuitive visualizations.
- **Forecast Visualization**: The tool offers a glimpse into the future, helping users plan and strategize.
- **Model Evaluation**: Ensuring the reliability of the forecasting model through key performance metrics.
- **Truth vs. Forecast:** Measuring the accuracy of predictions against actual outcomes.
- **Feature Impact:** Understanding the influence of individual features on demand forecasts, aiding in data-driven decision-making.

Scenario 2: Forecasting functionality of the tool

Steps

- **Demo:** Demand Prediction & Management Tool
- **Scenario:** Forecasting Functionality of the tool
- **Step 1:** The tool loads the historical data
- **Step 2:** User selects the "future forecasts" functionality of the tool
- **Step 3:** User is able to select the horizon of the forecasts
- **Step 4:** User is able to see the visualization of the forecasted timeseries data based on the selected horizon

Flow description

- **Data Loading**: The tool loads historical data, laying the groundwork for accurate forecasting.
- **Future Forecasts Selection**: Users unlock the forecasting feature, granting them access to the part of business insights.
- **Horizon Selection:** Users can customize their view of the future by selecting the time horizon for their forecasts, making predictions highly tailored.
- **Visualization of Forecasted Data**: The tool transforms data into actionable insights by providing users with compelling visualizations of forecasted time-series data, helping them make informed decisions about the future.

A.7. Requirements covered by POC

Table 16: FRs and NFRs being fulfilled from the POC

NFR.WP5.03	The system MUST provide accurate forecasts for consumption and the needs required from the CI operators (from very short periods of time to medium term)
FR.WP5.01	The data MUST have defined formats and schemas for CI topology, including connections or geographical information, and their supplied resources
FR.WP5.02	The Dataset MUST be complete in order to proceed with modelling
FR.WP5.03	The component MUST make available an API for importing resource/network data
FR.WP5.05	The system MUST find demand patterns, bottlenecks, and root causes for resource demand stress by applying both AI-based and Graph-based technologies
NFR.WP5.04	The module SHOULD provide a dashboard for training and inference control
NFR.WP5.05	The module MUST provide a dashboard for exploring the results and the several models' forecasts on demand prediction. An administrator needs a way to inspect logs and assigned metrics
FR.WP5.07	The tool MUST integrate its result into a CI oriented tool
FR.WP5.09	The tool MUST offer the users a dashboard to enable the parametrization of the specific CI and show information of interest to the users.
NFR.WP5.07	The tool MUST provide a dashboard that visualizes the performance metrics
NFR.WP5.08	The tool MUST provide a dashboard that visualizes the forecasted vs the true values