CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

Bachelor's Thesis

2023 **Vera Vavan**

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

Department of Water Resources and Environmental Modeling

The Role of SAP Analytics Cloud Smart Predict Time Series Feature in Evaluating Energy Consumption Trends and Environmental Impact of the SAP Metronom Business Center Office Location

Bachelor's Thesis

Academic Supervisor: prof. Ing. Martin Hanel, Ph.D.
Supervisor at the Company: Ing. Pavel Strnad
Author: Vera Vavan

2023

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

BACHELOR THESIS ASSIGNMENT

Vera Vavan

Environmental Data Science
Informatics

Thesis title

The Role of SAP Analytics Cloud Smart Predict Time Series Feature in Evaluating Energy Consumption Trends and Environmental Impact of the SAP Metronom Business Center Office Location

---

Objectives of thesis

- Investigating the usability and enablement of the SAP Analytics Cloud solution and the data democratization role in making data-driven decisions

-Measuring the individual impact of the Metronom office building in regards to the energy consumption of heating, cooling, electricity and water consumption as well as the corresponding CO2 production

- Using smart features with embedded AI to see whether the consumption trends will continue in the future after the catch-up effect caused by the COVID-19 office shutdown and seeing whether the SAP commitment to reaching carbon neutrality in 2023 is realistic

Methodology

SAP Analytics Cloud's tools used for the analysis are the Modeler for the data model creation, Story for the data visualizations and Predictive Scenario tool for the future forecast. The data used is the heating, cooling, water and electricity consumption, as well as the corresponding CO2 productions calculated using emission factors. The visualizations of historical and future data is then analyzed, in the context of the COVID-19 related office shutdown and the tools reliability.

The proposed extent of the thesis

40 pages

Keywords

SAC, Predictive model, Forecast, CO2 production, Office building, Data democratization, COVID-19

Recommended information sources

Banda, S., Chandra, S., & Gooi, C. (2022). SAP Business Technology Platform: An Introduction. Boston: Rheinwerk Publishing Inc.

Gole, V., & Shiralkar, S. (2020). Empower Decision Makers with SAP Analytics Cloud: Modernize BI with SAP's Single Platform for Analytics. Berkeley: Apress.

Ilinitch, A., Soderstrom, N., & E. Thomas, T. (1998). Measuring corporate environmental performance. Journal of Accounting and Public Policy, 383-408.

Lefebvre, H., Legner, C., & Fadler, M. (2021). Data democratization: toward a deeper understanding. International Conference on Information Systems (ICIS 2021).

Sidiq, A. (2022). SAP Analytics Cloud. Boston: Rheinwerk Publishing Inc.

Zhao, H. -x., & Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 3586-3592.

Expected date of thesis defence

2022/23 SS – FES

The Bachelor Thesis Supervisor

prof. Ing. Martin Hanel, Ph.D.

Supervising department

Department of Water Resources and Environmental Modeling

Electronic approval: 29. 3. 2023

prof. Ing. Martin Hanel, Ph.D.

Head of department

Electronic approval: 30. 3. 2023

prof. RNDr. Vladimír Bejček, CSc.

Dean

Prague on 30. 03. 2023

Restriction Note

The content of the present work must not be accessible in whole or in part to persons beyond the examination process and the evaluation procedure as long as no contrary approval of the workplace training provider exists.

SAP SE

Dietmar-Hopp-Alle 16

D 69190 Walldorf

Tel.: +49 (0) 6227 / 747474

Statutory Declaration

I hereby declare that I have independently elaborated the bachelor/final thesis with the topic of: "The Role of SAP Analytics Cloud Smart Predict Time Series Feature in Evaluating Energy Consumption Trends and Environmental Impact of the SAP Metronom Business Center Office Location", and that I have cited all of the information sources that I used in the thesis as listed at the end of the thesis in the list of used information sources.

I am aware that my bachelor/final thesis is subject to Act No. 121/2000 Coll., on copyright, on rights related to copyright and on amendments of certain acts, as amended by later regulations, particularly the provisions of Section 35(3) of the act on the use of the thesis.

I am aware that by submitting the bachelor/final thesis I agree with its publication under Act No. 111/1998 Coll., on universities and on the change and amendments of certain acts, as amended, regardless of the result of its defense.

With my own signature, I also declare that the electronic version is identical to the printed version and the data stated in the thesis has been processed in relation to the GDPR.

……………………………………….. ……………………………………

(City and Date) (VERA VAVAN)

Acknowledgement

I would like to thank Martin Hanel and Pavel Strnad for their support and contributions on this thesis, and for guiding me throughout my research. I would also like to thank all my colleagues from the SAP Enterprise Analytics Consumption Experience team who have always been willing to share their knowledge with me throughout the learning process and have provided me the right tools and authorizations for my research, which is also appreciated beyond the context of this thesis.

The Role of SAP Analytics Cloud Smart Predict Time Series Feature in Evaluating Energy Consumption Trends and Environmental Impact of the SAP Metronom Business Center Office Location

Abstract

The research constructed for this thesis aims to analyse the energy consumption and the corresponding CO_2 production trends and behaviours in one of the SAP Prague office locations, the Metronom Business Center. The date range includes the period from January 1, 2019, until March 31, 2022, and the analysis is performed using the SAP Analytics Cloud solution. The purpose of the analysis is to investigate the changes in office building consumption influenced by the COVID-19 disease emergence, as well as the possible future trends, and whether the solution enables corporate data democratization and to what extent.

Analysed data includes the cooling, heating, electricity and water consumption in the office building, from which the corresponding CO_2 productions were calculated using emission factors, and the tools from the SAP Analytics Cloud solution used for analysis are the Modeler for creating the data model, Story for data visualizations and Predictive Scenario for time series forecasting.

The findings of the data analysis are that the trend of energy consumption has been mostly negative since 2020, with a slight increase in 2021 as a result of office reopening after the COVID-19 shutdown and the catch—up effect. However, the overall decrease in CO_2 production is aligned with the SAP's commitment to reaching carbon neutrality in 2023, since the forecasted trends imply a steady decrease. The solution enablement empowers data democratization due to the simplicity and automation of the tools, nevertheless, the implied limitations include poor personalization options and low confidence of the results.

Research restrictions include missing data for electricity consumption for 2019 and 2020, and a restricted number of variables which influence CO_2 production due to data confidentiality. Further implementations and suggestions for the Smart Predict feature include building up on the textual explanations and data insights of the results and forecasts, as well as improving pattern recognition.

Keywords: SAC, Predictive model, Forecast, CO_2 production, Office building, Data democratization, COVID-19

Role funkce SAP Analytics Cloud Smart Predict Time Series při vyhodnocování trendů spotřeby energie a dopadu umístění kanceláří SAP Metronom Business Center na životní prostředí

Abstrakt

Výzkum vytvořený pro tuto diplomovou práci má za cíl analyzovat spotřebu energie a odpovídající trendy a chování produkce CO2 v jedné z pražských kanceláří SAP, Metronom Business Center. Období zahrnuje období od 1. ledna 2019 do 31. března 2022 a analýza se provádí pomocí cloudového řešení SAP Analytics Cloud. Účelem analýzy je prozkoumat změny ve spotřebě kancelářských budov ovlivněné výskytem onemocnění COVID-19 a také možné budoucí trendy a zda řešení umožňuje demokratizaci firemních dat a do jaké míry.

Analyzovaná data zahrnují spotřebu chlazení, vytápění, elektřiny a vody v kancelářské budově, ze které byly vypočteny odpovídající produkce CO2 pomocí emisních faktorů, a nástroje z řešení SAP Analytics Cloud používané pro analýzu jsou Modeler pro tvorbu datového modelu, Story pro vizualizace dat a Predictive Scenario pro prognózy časových řad.

Z analýzy dat vyplývá, že trend spotřeby energie je od roku 2020 převážně negativní, s mírným nárůstem v roce 2021 v důsledku znovuotevření kanceláře po odstávce COVID-19 a efektu dohánění. Celkový pokles produkce CO2 je však v souladu se závazkem SAP dosáhnout uhlíkové neutrality v roce 2023, protože předpokládané trendy naznačují trvalý pokles. Povolení řešení umožňuje demokratizaci dat díky jednoduchosti a automatizaci nástrojů, nicméně mezi předpokládaná omezení patří špatné možnosti personalizace a nízká spolehlivost výsledků.

Omezení výzkumu zahrnují chybějící údaje o spotřebě elektřiny za roky 2019 a 2020 a omezený počet proměnných, které ovlivňují produkci CO2 kvůli důvěrnosti údajů. Další implementace a návrhy pro funkci Smart Predict zahrnují zvýšení počtu textových vysvětlení a datových náhledů výsledků a prognóz, stejně jako zlepšení rozpoznávání vzorů.

Klíčová slova: SAC, Prediktivní model, Prognóza, Produkce CO2, Kancelářská budova, Demokratizace dat, COVID-19

Table of Contents

Abbreviations

AI	Artificial Intelligence
BI	Business Intelligence
BTP	Business Technology Platform
CSV	Comma Separated Values
DDEM	Data Democratization
ERP	Enterprise Resource Planning
GHG	GreenHouse Gas
HVAC	Heating, Ventilation and Air-Conditioning
KPI	Key Performance Indicator
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MBC	Metronom Business Center
PaaS	Platform as a Service
SaaS	Software as a Service
SAC	SAP Analytics Cloud
SAP	Systems Applications and Products
SE	Societas Europaea

1 Introduction

As the environmental issues are becoming more prominent and visible on the global scale, the need for sustainability and environmental impact monitoring in the corporate context has been increasing, as it is an important factor for all company stakeholders. Providing the past data and negative trends of the environmental impact of office buildings can change the outlook of consumers on the company's products, as customers have expressed a preference for greener companies, and the availability of information is increasing consumer's trust (Ilinitch, Soderstrom, & E. Thomas, 1998).

With the technological developments of data processing and visualization tools within software companies and the enablement of their usage, monitoring and predicting environmental data is becoming easier and democratized within all user groups, despite their knowledge of machine learning algorithms. When considering the data about energy consumption in office buildings, it is useful to analyse the past and current trends, as well as the resulting CO_2 productions, especially with the office shutdown related to the COVID-19 pandemic emergence in 2020. With data democratized tools, it is possible to perform such analysis and perform forecasts to investigate where the future data trends are heading towards. There is a general interest in the energy consumption changes and comparison of periods before and after the shutdown, as the employees and the stakeholders are wondering if the occupancy and environmental impact has changed after the offices were reopened and to what degree (Dacre, et al., 2021).

Analysis and visualization of data is enabled to almost anyone with access to the SAP Business Technology Platform (BTP), with the democratization of SAP Analytics Cloud (SAC), especially with the development of smart features which empower all consumers to use artificial intelligence (AI) tools for the analysis. However, the accuracy of automatized tools is brought into question, as an ill-informed user cannot assess the precision of the results. Smart features used and compared in this research are the SAC Predictive Scenarios tool and the time series chart automatic forecast within the SAC Story tool.

Considering the recent changes in the energy consumption of office buildings, a case study can be done for the location of SAP's Metronom Business Center (MBC), with the analysis of the past values and trends in cooling, heating, electricity and water consumption, as well as forecasting the possible future values and trends to see what the outlook is on energy consumption and CO_2 production in the future. With available information, SAP executives can make informed and strategic business decisions, taking into account employee preferences (Eldridge, 2022). As most companies already monitor their environmental impact and sustainability on a global level, including SAP SE, most employees are unaware of the detailed impact of their specific office location.

The proposed questions which the research aims to answer are whether the smart features are understandable to use and whether they produce realistic and accurate results which can be easily interpreted. Importance of the smart feature usability and accuracy comes from the fact that in order to use them, the user needs to invest time into learnings and trainings for the SAC software overall. Since the example of the tool usage is demonstrated using energy consumption data, the trends and analysis of past and forecasted future values aims to answer the question whether the consumption in MBC has been impacted by the COVID-19 shutdown and whether the values are returning to the state before the event. From an employee's perspective, the question imposes on whether their preference on their office location, whether it is in MBC or at home, has a major impact on the environment and overall CO_2 production of the company.

The structure and outline of the research firstly includes the detailed objectives of the research from an environmental and corporate perspective in *Section 2,* then the theoretical background about the data from MBC and how it is used further in the research, as well as the SAC tools introduction needed to understand the methodology of the research creation in *Section 3*. The location information of MBC from which the data was collected is also described, as to understand the layout and data collection. Secondly, it includes the methodology of the data model, chart visualizations and prediction creation using the SAC Modeler, Story and Predictive Scenario tools in *Section 4*. Afterwards, the outputs of the past and future visualizations are shown and analysed in *Section 5* in relation to the research questions proposed, while taking into consideration the many factors which influence energy consumption. Finally, the overall research and results are reflected and summarized in *Section 6,* along with the main contribution of the thesis and the suggestions for future research and implementations.

2 Objectives

The following sub-sections will describe the main objectives that the thesis targets, both from the SAP corporate and data analysis perspective.

2.1 SAP Objectives

2.1.1 Systems Applications and Products (SAP)

The goal of SAP SE as a software company is to provide Enterprise Resource Planning (ERP) system solutions used by businesses of all sizes, and ever since the establishment in 1972, it has set the global standard of such software in the industry (SAP SE, n.d.). Such systems are almost necessary to any enterprise, as it combines all needed software for business processes and functions, which enables the ease of management and accessibility in all business fields (Abd Elmonem, Nasr, & Geith, 2016). In case of SAP ERP solution, it is also provided via cloud where the user does not need to install or configure the system, and as such it is considered a Software as a Service (SaaS) (Abd Elmonem, Nasr, & Geith, 2016).

In addition, SAP BTP is a part of SAPs SaaS and Platform as a Service (PaaS) multi-cloud offering which provides a range of services to establish the technical foundation of a company and integrate with SAP ERP, including SAC as an all-inclusive analytics platform regarding the needs of the business (Banda, Chandra, & Gooi, 2022). SAC enables reporting, predicting and planning in a single cloud SaaS solution which is embedded in other SAP cloud-based applications and includes direct access to SAP data sources (Nazarov, Morozova, & Kokovikhin, SAP Analytic Cloud: a tool for the formation of professional competencies of business analyst, 2019).

As SAC contains many tools and functionalities within the solution, the opportunities for data analysis are vast, however, as the capabilities increase, the need for time and money investments into learnings and trainings for the solution's usage are also increasing (Abd Elmonem, Nasr, & Geith, 2016). Whether the outcomes are worth the investment depends on the objectives and the results it can produce. Hence, it is in the corporate interest to develop, maintain and enable the products with this in mind, and take the feedbacks into consideration.

2.1.2 Data Democratization

Ever since the invention of the World Wide Web, the newly enabled access to data and knowledge has changed the way almost any business operates, and the technical developments have not stopped being created ever since. This has caused a large volume of data that is still accumulating, especially within businesses, but it is mostly accessible only to ones that have the research and data skills to analyse it, such as data scientists (Awasthi & George, 2020). Due to this method of data accessibility, data research can only be made by a small group of employees in a company on which the other employees depend on to make business decisions, which creates a large gap between the employees and time inefficiency issues.

However, data democratization (DDEM) is a concept of opening the access to corporate data to as many employees as possible to make data-driven decisions, despite their knowledge background, while still taking into account security and confidentiality, which enables their competitive advantage as a company (Awasthi & George, 2020). The way that DDEM is implemented in different companies varies, but in the case of SAP SE, it is manifested not only within the company, but outside as well. When using SAP ERP and BTP, DDEM can be easily implemented with enabling the authorizations for certain tools and data for certain user groups, which still ensures data security and confidentiality within the company.

Developing more functionalities within SAC which are easily learned and used is contributing to the solution's usage and empowers data-driven decision making within the company (Nazarov, Morozova, & Kokovikhin, SAP Analytic Cloud: a tool for the formation of professional competencies of business analyst, 2019). This is one of the main objectives of SAC, and especially of the planning features included, which aims to enable data analysis and predictions without using any programming languages (Sidiq, 2022). As AI has been implemented in the SAC smart features, data analysis has been made easier, with automatic forecasting and insights into the data. However, the ease of use does not automatically enable the accuracy and precision which is produced by data scientists and specialized experts, hence improving the reliability of the product must also be focused on while still maintaining the accessibility.

To which extent the accessibility and accuracy of SAC and its' functionalities has been maintained will be one of the objectives of this research, as the learning procedure of using the software is explained from a fresh beginning. Two of the SAC smart features are used and compared, time series chart forecasting and smart predict for a time series, in order to establish whether the results are easily obtainable, and whether the results are realistic and accurate.

2.1.3 Sustainability

SAP's approach to sustainability is to lead by example and provide solutions within their BTP which enable companies to follow their examples, while committing to reaching net zero emissions by 2030 and becoming carbon neutral in 2023 (SAP SE, 2021). Focusing on energy consumption, water usage and waste management, SAP aims to encourage other companies to follow their example using their solutions and enablement for tracking the consumption of the resources within the corporate system, as stakeholders urge companies for environmental performance indicators and consumers focus their attention to green companies and products (Ilinitch, Soderstrom, & E. Thomas, 1998).

Sustainability reporting in SAP is available via their integrated reports, however, the individual contribution of each office building is still not available, since there are office locations in more than 78 countries worldwide (SAP SE, 2021; SAP SE, 2022). By analysing the energy consumption in MBC, this thesis aims to assess how realistic the commitment to carbon neutrality in 2023 is in this location. However, the waste management portion of the analysis was excluded from the report, due to data confidentiality.

2.2 Data Analysis Objectives

Due to the COVID-19 outbreak in 2020, most office locations were shut down, including MBC, which created a large decrease in energy consumption and a false sense of increased sustainability which can increase once the offices are reopened (SAP SE, 2021). By taking into account the data from 2021 and the first quarter of 2022 when the offices were partially and afterwards fully opened, the changes can be analysed and predicted in order to see if the decrease in consumption is maintainable. Hence, this thesis aims to investigate the catch-up effect of consumption, comparing the three periods of 2019 data before the pandemic event, 2020 data when the offices were fully shut down, and the 2021 data when they were partially opened. By also considering the first quarter of 2022, it can be seen whether the consumption so far is potentially lower or higher than the consumption in the first quarters of the previous years. Besides the analysis of historical data, the research aims to predict the future trends with the smart features to see whether the changes are sustainable and realistic.

However, as with any research, there are many limitations which do not enable a full comprehensive analysis of the full environmental impact due to data confidentiality and

missing data, hence the main variables used are cooling, heating, electricity and water consumption in MBC.
While conducting the analysis and data model implementation, the user experience of the software usage is also a factor influencing the research, since the aim of DDEM is a smooth and understandable usage of SAP's solutions. The possibilities and limitations of the tools are also considered as for the future implementations.

3 Theoretical Background

3.1 Approach to Research

Understanding the current state and known information about the research topic is the most important first step for the analysis of the issue (Silverman & Marvasti, 2008). Due to the vast number of available literatures on such topic, a methodology for reviewing the information is needed. A systematic approach to researching the available sources is necessary in such a case, especially with the current availability of scientific and non-scientific online sources (Silverman & Marvasti, 2008). However, once the topic is initially grasped, a limitation to the subject must be set regarding the scope of the report in order to fully explore it and provide a comprehensive result (Flick, 2018).

To perceive the initial sense of the research subject, a basis level of knowledge was set with provided internal software trainings on the topic of SAC under the supervision of colleagues who can be seen as experts on the topic. Without understanding the overall competence of the software, it is hard to explore its individual features and their extent, which is also the case for Smart Predict, which is a generally unknown feature within SAC users.

Once the comfortability of using the software is established, the approach to research shifts to scientific resources which are specific to the research topic using academic search engines such as Web of Science, Google Scholar and ProQuest, as well as SAP internal and public resources (Silverman & Marvasti, 2008). In the case of SAP software, it is important to read official documentation and roadmap for the future, which is publicly available and up to date, as it provides general information regardless of the implementation of the product (SAP SE, 2016). The SAP Press books have been found to be a substantial source, as the content can be trusted to be correct and supportive in the research journey (Rheinwerk Publishing, 2022).

Finally, after establishing the full theoretical grasp on the general topic, the research itself on the particular subject can be performed under the chosen methodology while comparing and referencing other scientific work that explores similar ideas (Silverman & Marvasti, 2008). This method enables the author to provide only necessary theoretical background on the topic in a way that can be understood by anyone, while the author himself understands the overall scope down to the specific technical details which assures the provided information is correct (Flick, 2018).

3.2 SAP Analytics Cloud (SAC)

3.2.1 SAC Architecture

The first step of setting up an SAC instance is establishing a data connection. SAC supports a wide variety of data sources, some that are live data connections, but also import data connections of on-premise data sources (Sidiq, 2022). However, connections are created by system owners and admins and are outside of the scope of this report, as it focuses on file upload based on the local network as the data source, which also supports scheduling imports on a regular basis to maintain the data model (SAP SE, 2022).

Moreover, the data connections in an SAC system are tightly bound to the amount of data tenants that are connected to it, which is also based on the needs of a business (Sidiq, 2022). An example of an SAC landscape can be seen in *Figure 1*: Illustration of a Possible SAC Landscape Configuration using a BW live data connection (Sidiq, 2022). This figure can also be used to illustrate the landscape design used for this research, where the development tenant is transported to quality and only then to the production instance, which are all connected to their own instance of the data source. In the outlined landscape, the development tenants are used in the starting point of model and visualization construction, and the general guidance is that nothing is created from scratch in the productive and quality tenant where the files are consumed by other users, only transported to it. As the data model in this research is acquired from a file import and the data sources connected to the instance are not relevant, the SAC tenant used is the development tenant (S1V) with regard to the performance and purpose. The model and the visualizations are not meant to be consumed by other users; hence they are not transported.

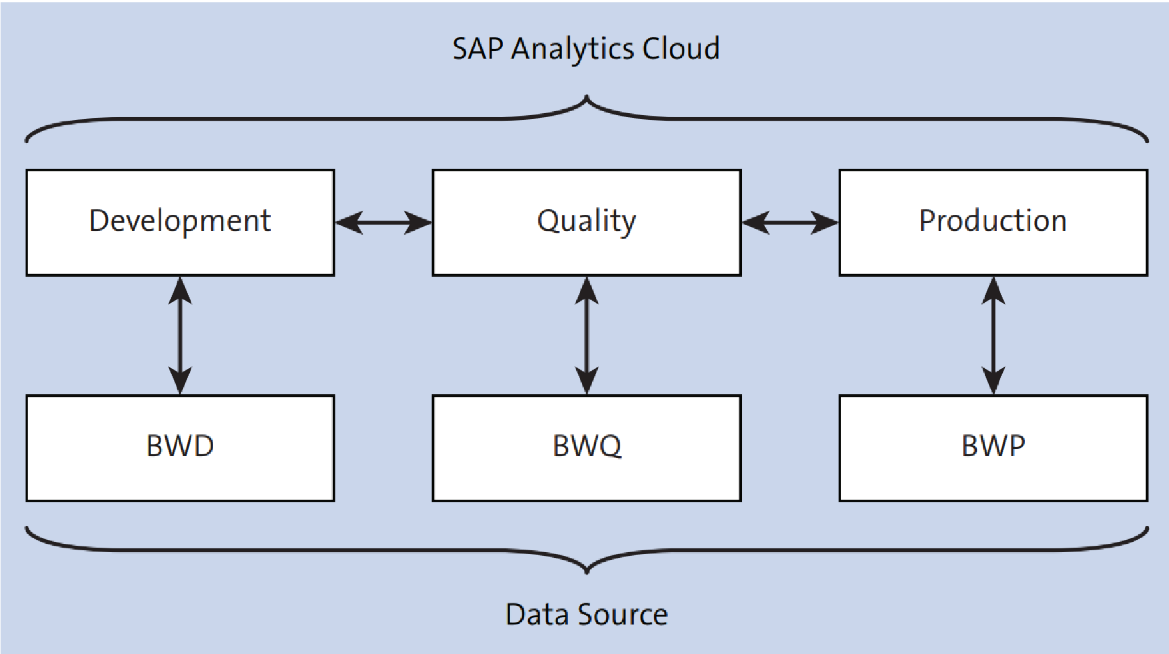


Figure 1: Illustration of a Possible SAC Landscape Configuration (Sidiq, 2022)

Finally, the SAC landscape can also be configured in terms of user authorizations and roles. SAC enables the system owner or admin to restrict user access to certain capabilities by assigning them roles with specific privileges enabled (SAP SE, 2022). This method ensures that data security and privacy in the tenant is maintained. For the purpose of this research, an admin role with full privileges and permissions was acquired to ensure that there are no limitations during the exploration, but otherwise users would need to explain the use case and needs for granting certain functionality permissions.

3.2.2 Modelling in SAC

Once the data source and tenant instance have been established, it is possible to use the Modeler tool in SAC to create a data model based on the source and initiate data wrangling. The first part of creating a model based on local files is data import where the user interface presents a data sample of 2000 rows in a table if the dataset has a larger size, but the tool automatically detects and displays the number of rows, columns, dimensions and measures, and performs a data quality check to ensure that the data does not have any anomalies which the user needs to address manually (SAP SE, 2022). However, any changes and calculations made in the data import screen are only applied to the sample rows, hence the quality check is not performed on the entire dataset and could cause problems in the further steps.

Once the data is imported and validated, the user can manipulate their data in the SAC Modeler tool, which enables a more detailed overview of the data with the full dataset and the data can be changed by removing or replacing rows and columns or the values themselves (Sidiq, 2022). Performing any calculations to add more measures and columns is also possible, but there is also the ability to create hierarchies, as well as time and geographical specific dimensions from the regular dimensions and see their data distribution (Sidiq, 2022). The formulas and calculations possibilities are quite limited for the user and do not use SQL functionalities, but a special set of available SAC formulas (SAP SE, 2022).

3.2.3 Visualization in SAC

The created data models can be then used to create multiple presentations of their data based on the use case. The prevalent form of data visualization is the story created using the SAC Story tool, which can also be called a dashboard if complex enough, or a report if it is lightly constructed (Sidiq, 2022). Stories have two views once they are connected to a data model, the data view enables quick visualizations of various types which are not automatically added to the story itself but are used for data exploration, while the story view enables direct story design with more functionalities (SAP SE, 2022).

As mentioned, stories can be as simple or as complex as the creator decides, since the capabilities of the tool include many functionalities. Some of the main functionalities are adding multiple pages to a story which are blank white canvases on top of which visualizations are displayed based on their common topic, and it is possible to apply filters for data dimension values which apply to either the entire story, just one page or just selected visualizations (SAP SE, 2022). The visualizations can also be of many types, with charts having the most variety, but also tables, geographical maps and R code visualizations (Bertram, et al., 2021).

Users also have the ability to calculate additional measures and dimensions on top of the data model, which are then only available within the story, and if another story were to be created using the same model, these calculations would not be available. The calculations within the SAC Story tool are more user friendly, with restrictions and filters being pre-set without the need to know any code functions, but just selecting the value members, such as restricting the date to a certain year by simply selecting the wanted year in a drop down menu (SAP SE, 2022).

With the main features of the SAC Landscape explained, the visual representation of the main system components can be seen in *Figure 2*: Visual Diagram of the SAC System Components and their Connectivity. Viewing from the bottom to the top, the first step is determining the data source, then preparing the data model and scheduling using the platform functionalities, then using business intelligence and augmented analytics within the SAC story capabilities.

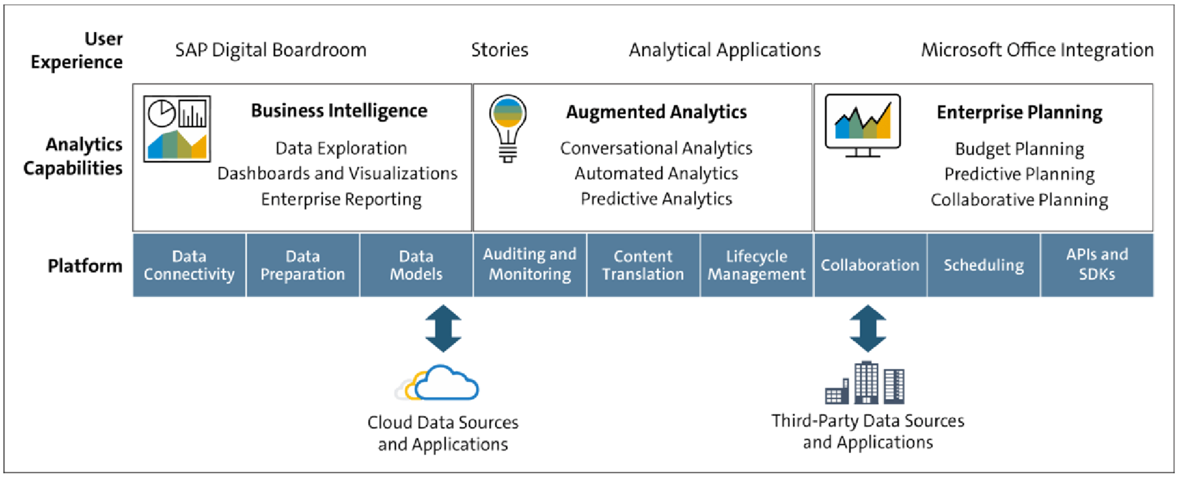


Figure 2: Visual Diagram of the SAC System Components and their Connectivity (Banda, Chandra, & Gooi, 2022)

3.2.4 Augmented Analytics (Smart Features)

Augmented analytics, often called predictive analytics, is a component of SAC that allows the use of smart features with machine learning embedment. It is divided into three groups based on the outcome of the feature, including smart insights, search to insights, smart discovery and smart predict, but each category enables the user to perform machine learning tasks without any technical knowledge besides the output analysis skills (SAP SE, 2022). The focus of this research is on the smart predict and forecast feature.

Smart predict manifests itself in two ways, one is within an SAC story on a time series or line chart and one as a separate functionality called Predictive Scenarios. Time series forecast is the simplified version of a predictive scenario, where the user can automatically forecast a chart with a time dimension and measure, but the options for choosing a prediction model are limited to linear regression and triple exponential smoothing, with the option of adding multiple measures (SAP SE, 2022). Such actions are as simple as pressing two buttons on a chart and can be used by any user, regardless of their knowledge, especially the automatic forecast feature where the prediction model is chosen automatically by the software.

On the other hand, Predictive Scenarios are a separate SAC functionality outside of the story itself and can become more complex to use and interpret, as there needs to be research done in order to use it. Upon creating a scenario, a choice must be made based on the data type and use case, you can choose the classification, regression and time series scenario, however, due to the nature of the dataset used in the research, time series predictive scenario will be the focus of the report (SAP SE, 2022). The choice was made based on the wanted result, as the aim is to predict future consumption values and a date dimension is present.

The user responsibility in the creation of a time series scenario is to connect it to the SAC data model, select a target variable which should be the key performance indicator (KPI) of your interest, a date dimension, an optional entity if you wish to separate the KPI into groups and the number of forecast periods which is the number of hypothetical predictive forecasts that will be generated (Vahlkamp & Vahlkamp, 2020). Once the options are selected, the software automatically trains the model by dividing it into two datasets, the training set which is 75% of the inputted data used to generate multiple hypothetical predictive models and a validation set which is the other 25% of the data, and compares it to the training set in order to determine the best counterpart model as a result (SAP SE, 2022). A schema of the partition strategy can be seen in *Figure* 3.

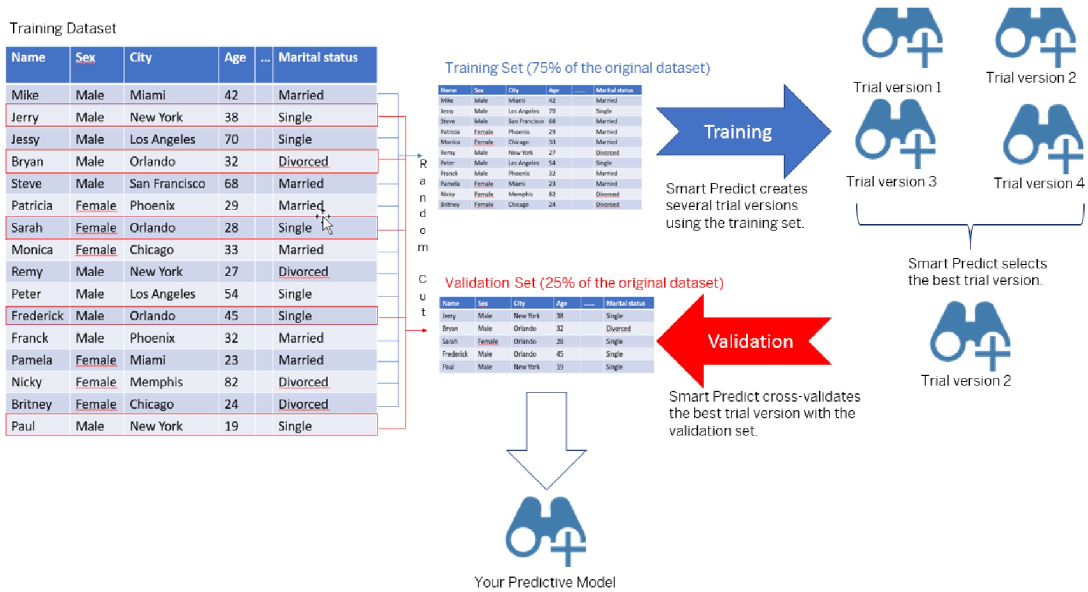


Figure 3: Partition Strategy of SAC Predictive Scenarios Schema (SAP SE, 2022)

The user does not have any insights or decisions on the selection of the best fit model; however, the user is able to produce multiple scenarios to compare and is able to provide the number of forecast periods to be generated in a single scenario. The methodology for selecting the best model fit is based on horizon-wide mean absolute error (MAE) in a time series predictive scenario, where MAE is found for each forecast period and then the mean is taken as the horizon-wide MAE (SAP SE, 2022).

MAE is a KPI found by first calculating the absolute error by subtracting the forecasted value from the real value and then finding the mean of the difference for all of the observations, as it can be seen in *Equation 1*, therefore this method does not take into account the scale and does not indicate whether the result is good or bad, but just the magnitude of the error (Foss & Modderman, 2019). The output of the time series predictive scenarios shows a mean absolute percentage error (MAPE) as a measure of accuracy, which is similar to MAE, but the absolute difference between the values is also divided by the real value and the result is a percentage where the higher value indicates worse accuracy if the forecasted values are much higher than the real values, and it can be seen in *Equation 2* (Foss & Modderman, 2019). The outputs of the time series predictive scenarios will be analysed additionally.

$$MAE = \frac{\sum |real\ value - forecast\ value|}{n} \tag{1}$$

$$MAPE = \frac{\sum \frac{|real\ value - forecast\ value|}{real\ value}}{n} * 100\% \tag{2}$$

3.3 Office Sustainability

3.3.1 Metronom Business Center

Currently, there are three SAP office locations in the Czech Republic, two of which are in Prague with the largest location being the Metronom Business Center (SAP SE, n.d.). The building is divided into three interconnected towers, which are referred to as building A, B and C regardless of the connectivity, which adds up to 31 199 m^2 of office space divided into 250-4 500 m^2 per floor depending on the location, and the layout can be seen in *Figure 4* (White Star Real Estate, n.d.). The center was built in 2015 by HB Reavis located at Bucharova Street Nos. 2817/9-13, Nové Butovice, Prague 5 and it is owned by METRONOM BC s.r.o. at the current time (REICO , n.d.). It comprises of 7-9 floors depending on the building and not including the ground floor where SAP offices are not located except the entrance and elevator.

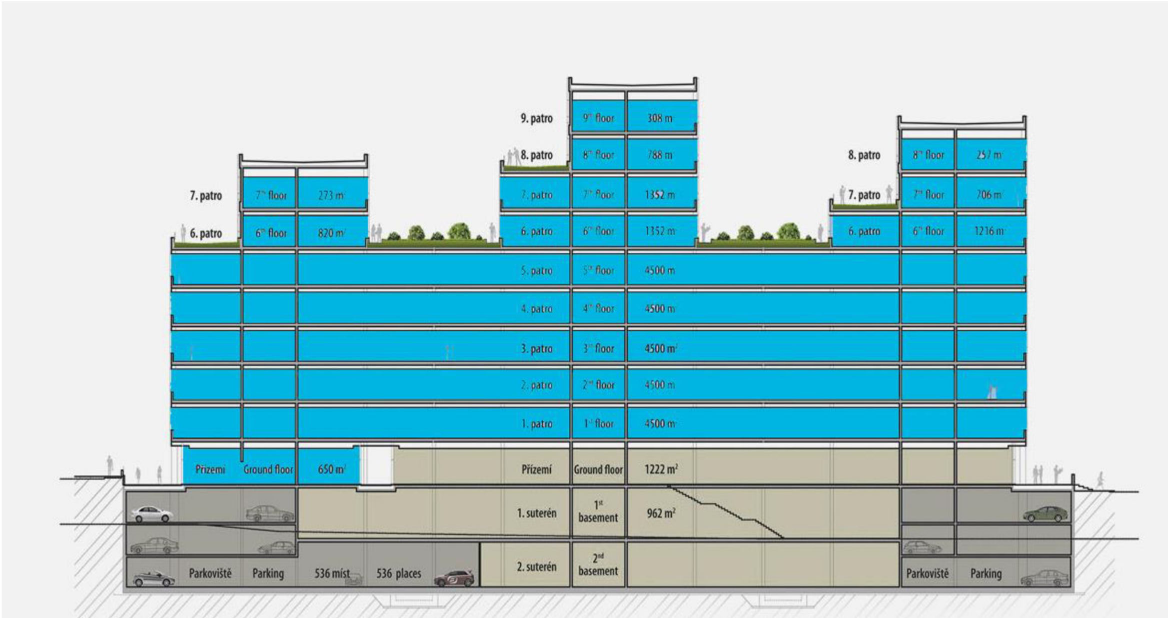


Figure 4: Schema of the Metronom Business Center (White Star Real Estate, n.d.)

As it can be seen when considering both *Figure 4* and *Appendix A*, building A includes 8 floors, building B includes 9 floors and building C only 7 floors. The energy consumption in the buildings is measured 1:1 for each tenant, hence the consumption data provided to SAP does not include the unoccupied ground floor space (White Star Real Estate, n.d.).

Heating, ventilation and air conditioning (HVAC) has separate regulations for each office and it has received the Excellent grade on the Building Research Establishment Environmental Assessment Methodology (BREEAM) certificate, meaning that it received a score of 70.7% in 2018 (BRE Group, 2018). The BREEAM grade specific distribution for some of the categories was:

- energy considering the operational energy of the building and CO2 emissions: 78
- water consumption and efficiency: 56
- land use & ecology in terms of site and building footprint: 70
- transport in terms of related CO2 and location-related factors: 100
- water and air pollution: 50 (White Star Real Estate, n.d.; BRE, 2020)

3.3.2 The Environmental Impact of Energy Consumption

As the consequences of the anthropogenic impact on the climate become more manifested and researched on the global scale, the urge to decrease the global warming increase rate, in hopes of not reaching the 1.5°C temperature increase above pre-industrial level in the following years, is becoming stronger in most business sectors (Ara Begum, et al., 2022). Even in the optimistic case of 1.5°C increase, the possible consequences are severe, such as natural system collapse, water and food resources restrictions, extreme weather events, mortality, economic damages, and many more depending on the location and area (O'Neill, et al., 2022).

Considering greenhouse gasses (GHGs) as one of the main drivers of climate change, of which carbon dioxide (CO_2) takes up 72%, the need to reduce emissions from coal, oil and natural gas combustion that mostly produces it is increasing (Olivier & Peters, 2020). As energy is consumed from these sources, CO_2 emissions are increasing, as well as the need for monitoring and policy implementation on the way that energy consumption is demonstrated globally (Worrell, Bernstein, Roy, Price, & Harnisch, 2009). However, some studies propose that the increase in CO_2 emissions is having a negative effect on energy consumption, suggesting that these measures are having a positive effect on turning to renewable energy sources (Acheampong, 2018).

The SAP's commitment to net zero emissions explains that their part in helping the future temperatures not reach a 1.5°C temperature increase is by releasing less GHGs into the atmosphere, so that the outgoing emissions can be balanced out to the gas removals from the atmosphere and reach a net zero balance (SAP SE, 2021). This is possible in cooperation with other major companies, which is reflected in SAP's lead by example philosophy. Carbon neutrality differs from the net zero commitments as it focuses specifically on CO_2 gas removal from the atmosphere, and it can be accomplished on a company level by investing in many environmental innovations within the buildings and data centers which are carbon neutral (SAP SE, 2021). Another way to achieve both commitments is investing funds and efforts into turning to renewable energy sources, especially in relation to carbon emissions.

3.3.3 Energy Consumption of Office Buildings

Considering the pressure on companies to reduce their environmental impact, the need for monitoring and forecast implementation on the energy consumption practices in office buildings has been extensive (Ilinitch, Soderstrom, & E. Thomas, 1998). Since buildings account for even 40% of total energy use, they also produce 36% of total CO_2 emissions, and to measure individual impact of an office building of a company the information about energy types consumed is needed in order to obtain the carbon intensity factor and determine the CO_2 production (Zhao & Magoulès, 2012). Most companies that monitor their environmental impact globally already have specified emission factors for energy consumption.

Some of the main components of building energy consumption are the HVAC systems, controlling and measuring separately the cooling, heating and energy supply (Zhao & Magoulès, 2012). As most modern buildings are equipped with building automation systems (BASs), extracting and optimizing the energy consumption is available using various energy sub-meter sensors for measuring each variable on separate floors, offices or buildings, as it is in the case of MBC (Wang & Ma, 2008). In MBC, employees can control the heating and cooling systems per their needs (SAP SE, 2023).

Cooling, heating, electricity and water consumption data is measured using electronic sub sensors which provide the energy consumption usually in different units and locations, and depending on their placement and type they can provide the information of energy used for electricity by sensing the values of current and voltage in watt-hours (Wh), for heating and cooling units by using thermal sensors measuring the energy used to heat/cool them in giga-joules (GJ), and for water with flow sensors which measure the amount of water released into the systems for distribution in m^3 (Ahmad, Mourshed, Mundow, Sisinni, & Rezgui, 2016). Such data is usually maintained by the landlord of the building and can be provided to the tenant on request.

Once the energy consumption is measured, extracting the CO_2 production is done by determining the emission factor based on the business type, energy source, geographical location and data variable, and multiplying the values (World Resources Institute, 2015). Since the emission factors vary based on the variable, the total CO_2 production of the system cannot be found from the total energy consumption, but from the individual CO_2 productions of each variable summed up to the total. The methods and products mentioned are used in the case of MBC, however, the emission factors were predetermined by SAP SE in regards to the business type and scope. There are a lot of other factors influencing the CO_2 production of a building, including the indirect sources such as commuting, data centers and waste management, but they are outside of the scope for this research and are protected due to data confidentiality.

Recent trends in the corporate industry also showed a lot of change on the energy consumption in the physical offices with the rising popularity of teleworking as a consequence of the COVID-19 pandemic event which continues to be the main method of working even after the pandemic (Hook, Court, Sovacool, & Sorrel, 2020). In the case of the MBC, SAP shut down the office entry for almost all employees on March 17, 2020, but reopened the entrance in the start of 2021 to some employees who have been vaccinated or tested negative for the disease, and reopened the entrance to all employees and office areas on February 18, 2022 (SAP SE, 2023). However, most of the employees have still not returned to the office as there is no internal policy in place which requires mandatory office attendance and SAP reported a building utilization level of ~6% on Fridays, during which the energy consumption services are reduced to only ~20% of the building (SAP SE, 2023). Hence, the dataset used for forecasting of future values needs to include data from all periods, before the event, during and after the event in order to realistically represent the trends and behaviours.

4 Methodology

4.1 Data Wrangling

The first step to creating the forecast scenario is to extract and clean the data itself, from which the model is being created (Silverman & Marvasti, 2008). This step is important since there could be many consequences and difficulties if the SAC software cannot detect the format and recognize the data when creating the model, since the actual user functionalities are limited in the SAC modeler tool compared to using R or Python to load data.

4.1.1 Data

The energy consumption data for the MBC was provided by the landlord of the building, which is the White Star Real Estate company. It includes the information about the sensor identifier, sensor type, unit of the measured variable, floor and building that it is placed in, as well as the street location of the building itself and the timestamp of the measured consumption value. The building location is the same throughout the entire dataset, as the landlord provided the information specifically for the MBC and for the space occupied by SAP SE only. Data file provided is in the comma separated values (CSV) format with 1 473 701 rows and 8 columns, with the first row used as a header, and the timestamp values include the date and the time of measurement in one column, which is every hour. In total, the dataset includes seven dimensions and one measure, which all add up to 123 117 KB in file size, and a sample of 34 rows can be seen in *Figure 5*.

The sensor column includes the unique identifier of the sensor, indicating the building and the floor on which it is located, the sensor type and serial number. The available sensor types are the electricity meter (Wh), calorimeter cooling (GJ), calorimeter heating (GJ) and the water meter (m^3). As mentioned before, the building dimension can take the values A, B or C which are shown in *Figure 4,* and the floor dimension is written in the format '[number].NP' where 'NP' stands for 'nadzemní podlaží' in the Czech language, indicating that the floor numeration starts from the ground floor with the number 1, meaning that the total number of floors in MBC is 10. However, the ground floor which is labelled '1.NP' does not have any energy consumption monitoring, hence the number of floors with data assigned is 9 starting from '2.NP' until '10.NP'. There are 141 individual sensors spread across the floors, with each floor in each building having at least one sensor for each variable, except for 2.NP and 10.NP floor in building B which do not have any water sensors.

The date range also depends on the time that the sensors were installed in the building, hence different variables can have different starting dates. To reduce the differences between starting dates, the dataset was restricted to start on January 1, 2019 and end on March 31, 2022 at 23:00. However, the electricity monitoring in buildings A and B has not started in 2019, in building A the first electricity sensor measured data on April 23, 2020 at 22:00 and in building B on November 11, 2020 at 22:00. This is due to the installation of new sensors, which replaced the previous ones, but the previous data was not provided. This is an important note to keep in mind in further analysis, since the COVID-19 event has already started at this point and the dataset does not include the electricity consumption prior to it, hence there will be a sudden apparent increase in electricity consumption in 2020. Timestamps in the provided data are in the format 'MM/DD/YYYY HH:MM'.

```
Sensor,Sensor Type,Unit,Floor,Building,Location,TimeStampNew,Consumption
2RSC.D_PW_23  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 0:00,768
2RSC_PW1_21  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 0:00,225
2RSC_PW2_22  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 0:00,608
4RSC_PW1_41  ,electricity meter  ,Wh  ,4.NP,C,Bucharova 2817/11,1/1/2020 0:00,110
4RSC_PW2_42  ,electricity meter  ,Wh  ,4.NP,C,Bucharova 2817/11,1/1/2020 0:00,2496
5RSC.D_PW_53  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 0:00,130
5RSC_PW1_51  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 0:00,106
5RSC_PW2_52  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 0:00,2384
6RSC_PW1_61  ,electricity meter  ,Wh  ,6.NP,C,Bucharova 2817/11,1/1/2020 0:00,111
6RSC_PW2_62  ,electricity meter  ,Wh  ,6.NP,C,Bucharova 2817/11,1/1/2020 0:00,2064
7RSC_PW1_71  ,electricity meter  ,Wh  ,7.NP,C,Bucharova 2817/11,1/1/2020 0:00,112
7RSC_PW2_72  ,electricity meter  ,Wh  ,7.NP,C,Bucharova 2817/11,1/1/2020 0:00,1808
8RSC.D_PW_83  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 0:00,458
8RSC_PW1_81  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 0:00,584
8RSC_PW2_82  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 0:00,364
2RSC.D_PW_23  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 1:00,764
2RSC_PW1_21  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 1:00,226
2RSC_PW2_22  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 1:00,616
4RSC_PW1_41  ,electricity meter  ,Wh  ,4.NP,C,Bucharova 2817/11,1/1/2020 1:00,109
4RSC_PW2_42  ,electricity meter  ,Wh  ,4.NP,C,Bucharova 2817/11,1/1/2020 1:00,2496
5RSC.D_PW_53  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 1:00,130
5RSC_PW1_51  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 1:00,107
5RSC_PW2_52  ,electricity meter  ,Wh  ,5.NP,C,Bucharova 2817/11,1/1/2020 1:00,2448
6RSC_PW1_61  ,electricity meter  ,Wh  ,6.NP,C,Bucharova 2817/11,1/1/2020 1:00,110
6RSC_PW2_62  ,electricity meter  ,Wh  ,6.NP,C,Bucharova 2817/11,1/1/2020 1:00,2320
7RSC_PW1_71  ,electricity meter  ,Wh  ,7.NP,C,Bucharova 2817/11,1/1/2020 1:00,112
7RSC_PW2_72  ,electricity meter  ,Wh  ,7.NP,C,Bucharova 2817/11,1/1/2020 1:00,1800
8RSC.D_PW_83  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 1:00,460
8RSC_PW1_81  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 1:00,754
8RSC_PW2_82  ,electricity meter  ,Wh  ,8.NP,C,Bucharova 2817/11,1/1/2020 1:00,356
2RSC.D_PW_23  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 2:00,844
2RSC_PW1_21  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 2:00,224
2RSC_PW2_22  ,electricity meter  ,Wh  ,2.NP,C,Bucharova 2817/11,1/1/2020 2:00,616
4RSC_PW1_41  ,electricity meter  ,Wh  ,4.NP,C,Bucharova 2817/11,1/1/2020 2:00,109
```

Figure 5: Energy Consumption Data Sample in the CSV Format

As mentioned in the previous section, the CO_2 production can be calculated from energy consumption using emission factors which are provided by SAP SE separately. For electricity the emission factor is 0.382, for heating and cooling it is 0.055 and for water it is 0.00038.

The following two sections are dedicated to the SAC Modeler tool, hence consult *Appendix B* and *Appendix C* for the user interface design of the data import and the SAC Modeler tool.

4.1.2 Data Import

After obtaining the data file, the data was imported to the SAC Validation tenant using the data import functionality in the SAC Modeler tool and the option of inserting data from a CSV file from the local system, with using the first row as headers and the CSV delimiter comma. As expected, the modeler tool presented a sample size of only 2 000 rows from the 1 473 700 original ones with regards to performance.

The issues with the software interpretation of the data were visible from the initial overview presented where the model requirements were not satisfied. The warning indicated that the timestamp column contained data quality issues, and upon opening the details of the column it could be seen that the date format registered by the software was 'DD/MM/YYYY' which does not match the actual date format in the data, hence it interpreted the issue of months having a value higher than 12. The date format of the column was replaced with 'MM/DD/YYYY' and the data quality issues were fixed, therefore the software data validation was initiated. Data validation results showed that there were no issues were found and planning was enabled for the model, which is required for the creation of predictive scenarios.

Data was inspected manually as well, however, the sample size did not allow a full inspection of the data distribution, as all the rows were sampled from the electricity meters in building C only, hence some data quality issues might occur once the changes are applied to the full dataset. However, based on prior experience, a calculation was made for each column where the values were split on the last space symbol using the transform functionality. The formulas used for the splitting can be seen in *Equation 3*, where the integer used was either 1 for columns which have a string with no spaces, or 2 for the columns which have a space in the values. If the column values had a space, *Equation 4* was also used to merge the split data values back together, leaving a third blank column of whitespaces.

$$Split\ \text{[column name]}\ on\ \ repeat\ \text{"[integer]"} \tag{3}$$

$$Concatenate\ \text{[column name 1], [column name 2]}\ using\ \text{" "} \tag{4}$$

The reason for this action is that the SAC software detects a difference when there is a whitespace at the end of the value in the column, resulting in multiple value groups for the same value and issues in further calculations. Once the split was performed, the secondary columns containing only the whitespaces were deleted.

4.1.3 Data Model

Once the data import was done, the model was created under the name 'METRONOM_CONSUMPTION_thesis' in private user files. The model was opened in the SAC Modeler tool and the model structure was inspected to ensure that the previous split action has been done successfully. All of the dimension member counts were displayed and inspected to be correct, with the count having one additional value for the unassigned members. The additional unassigned member does not indicate that there are actual values not assigned to any member in the dimension, it is an automatic addition performed by the software (Sidiq, 2022). To assure that there are no discrepancies in the data, it was inspected manually with the full row count where the consumption was added to the table and then each dimension individually. If there were any unassigned values, they would be shown beside the usual members, but there were none.

Upon the inspection of the dataset, it was noticed that the default hierarchy of the timestamp column was set as ‘Year Quarter Month Day’, but for the purpose of the research it was changed to ‘Year Month Day’ and the column label name was changed from ‘TimeStampNew’ to ‘Date’ for easier readability. Another issue is that the model creation has emitted the actual timestamp values and has kept only the date values without time. The reason for this behaviour is that the planning capability of the SAC model does not support the timestamp data format, only the date format and the data was aggregated to a daily time step (SAP SE, 2022).

Once the manual quality inspection was complete, new measures were calculated as the energy consumption was not divided into groups corresponding to the sensor type. The new measures for consumption were Cooling Consumption, Heating Consumption, Electricity Consumption and Water Consumption which were created using *Equation 2*. Once that the consumption values were found for each variable, the individual CO_2 production was also calculated using *Equation 3*, and the measures Cooling CO_2 Production, Heating CO_2 Production, Electricity CO_2 Production and Water CO_2 Production were created. The reason for multiplication with 1000 is for unit conversion from gram to kilogram (kg), so that all the resulting CO_2 production values are unified in kg, and so the only measure where the value was divided by 1000 is the electricity consumption to convert it to kWh and then kg. In the end, the total CO_2 production was found by aggregating the individual consumption measures, since they are all in kg, and named Total CO_2 Production.

$$RESTRICT\ ([Consumption], [d/Sensor\_Type] = \text{"[value]"}) \tag{5}$$

$$[\text{Consumption measure}] * [Emission\ factor] * 1000 \tag{6}$$

With the added calculations, the model for energy consumption is finished and it is ready to be used in the SAC Story tool. As a general suggestion, SAC models are advised to be as simple as possible and should only contain the main calculations and all other more specific calculations which apply to a specific story should be made in the SAC Story tool so that the model can be reused for other purposes and is not crowded with too many measures.

4.2 Visualization

4.2.1 Story

The SAC story was created with the responsive option where the proportions of the layout are adapted to the screen resolution, and the model made in the previous section was linked to it. Prior to the creation of charts, the data was explored in the data exploration view to see which points and dimensions can be used and make sense to present. Afterwards, two pages were created, one called Energy Consumption and the other CO_2 Production, and they can be seen in *Figure 6* and *Figure 7*.

Both pages are similar in design and layout, with numerical charts displaying the total number of consumption or production per variable, and a corresponding time series chart displaying the consumption or production trend per building over time with different coloured lines. The Energy Consumption page contains only four pair of charts for cooling, heating, electricity and water consumption, while the CO_2 Production page also includes total production of CO_2 summed up. The reason for not including the total energy consumption is that the values cannot be aggregated as the units of the values vary between the variables. The pages were designed with the purpose of analysing the past trends and differences in mind.

Additional calculations were performed to find the consumption and production per variable for each year by creating a restricted measure and using the Date dimension to limit the data to 2019, 2020, 2021 and 2022, afterwards these measures were added to the numerical charts as secondary values to display the consumption and production breakdown per year. The names of the new variables were constructed using the first letter of the variable (C, H, W, E and T), the first letter of the measure (C or P) and the year for which it is calculated.

Variance for these measures was added to the numerical charts by comparing the year for which the variance is being calculated to the corresponding previous year value and displaying the percentage variance value in brackets next to the consumption or prediction number with the green colour indicating negative values and the red colour indicating positive values. The percent change in the values for the previous year could not be calculated for the year 2019, as there is no previous year data to compare it to. Moreover, the change from 2021 to 2022 has to be considered in regards to the fact that the 2022 data is recorded until March 31 only and displays a lower value.

The time series charts are displayed with three lines for each building differentiated by the colours blue for A, orange for B and green for C, with the legend displayed above the chart and below the chart title. The X axis on all charts is identical, with the date range including all of the dataset starting from January 1, 2019 until March 31, 2022, while the lines for A and B electricity start at different points. The Y axis of the charts differs for each chart, depending on the range of values included in the variable. The unit for the values is also displayed either in the chart title, or in the top left corner, depending on the chart type.

4.2.2 Forecast

Once the SAC story was created, the automatic forecast functionality was applied only to the CO_2 Production page to each time series chart, which can be seen in *Figure 8 and Appendix D* can be consulted for a closer look. The reason for only forecasting the one story page is that the consumption and production are directly related, and the prediction scenario can only be applied to the original consumption variable, and not the calculated measures. The forecast periods for all charts were 92 which is the maximum option, and the date range predicted is from April 1, 2022 until July 1, 2022, which could not be manipulated. It contains three prediction lines and three confidence intervals coloured according to the legend.

The charts X axis still begins on January 1, 2019, however, the screen captures shown in *Figure 8* have the charts zoomed in for the purpose of value investigation and start on July 1, 2021. In the SAC story, there is a slider which enables the user to see the full date range from the beginning by scrolling.

4.3 Predictive Scenario

The SAC Predictive Scenario tool was used to create one time series forecasting predictive model using the model created in *Section 4.1.3*. The predictive goal was selected to be the Consumption measure and the Date dimension was used for the date. The number of forecast periods selected was 92 to match the chart forecast and the entity on which the consumption was split was only the Sensor Type dimension since the building consumption differences are not as high and would require a deeper analysis with 12 different results.

The predictive model was trained using all available observations and the negative forecast values were converted to zero, since consumption cannot be negative. Training and creation of the model results took less than 5 minutes, and the forecasted period is from April 1, 2022 until July 1, 2022, same as for the time series chart forecast. The result of the scenario also provides the individual forecasted values, as well as the minimum and maximum errors.

5 Results and Discussion

5.1 SAC Story Pages

The story pages created in the previous section can be seen in *Figure 6* and *Figure 7* below.

5.1.1 Energy Consumption

The Energy Consumption page of the SAC story shows clear yearly trends in the time series charts, with the most prominent cycles being for cooling and heating consumption as the cooling devices are used mostly during summer and the heating devices during winter months. Another overall trend is visible in the decrease of consumption levels in 2020 and a slight increase in 2021 compared to the previous year, but compared to 2019 the levels are still lower.

Figure 6: SAC Story Page Called Energy Consumption Showing the Consumption per Variable and Building Over Time

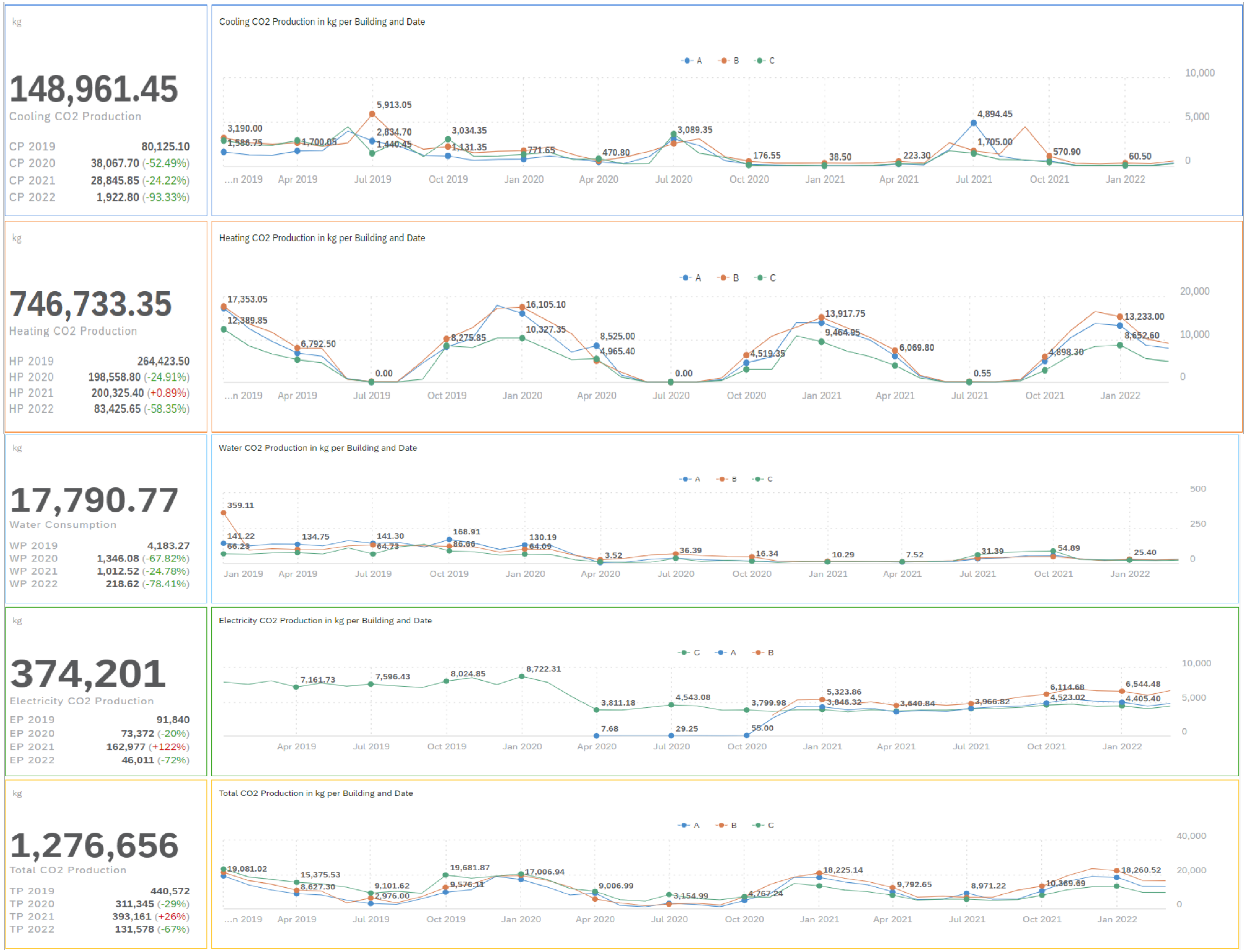


Figure 7: SAC Story Page Called CO2 Production Showing the CO2 Production per Variable and Building Over Time

Cooling consumption has experienced a large and steady decrease, with the consumption steadily decreasing over the years. Comparing the total value of 2 708.39GJ over more than three years to the 2019 value of 1 456.82GJ shows that more than half of the total cooling consumption was produced before the COVID-19 event and it has still not increased after the return of employees to the office. When examining the cooling consumption over the winter period in 2019 and beginning of 2020, it can be seen that the cooling units were even relatively highly used during the colder months. Overall, most cooling units are consumed in building B, followed by building C, which are generally more used. Comparing the Y axis scale and overall values to the heating consumption in the same units, the cooling consumption has the lowest values but still high fluctuations. Three anomalies can be recognized from the chart, with building B experiencing a large value compared to the other buildings in July 2019 and September 2021, and the building A experiencing the highest value and a very high difference from other values in July 2021. Such events could be accredited to a conference, or an all-hands meeting performed in only one building which is aggregating the employees in one location.

On the other hand, heating comparison has more than two times bigger Y axis scale, and the total consumption is more than five times bigger than the total cooling consumption. The consumption follows an opposite trend from the previous chart, as the heating increases during the winter months, but the fluctuation between them is much higher with the lowest values reaching zero for longer periods of time. However, the winter periods are longer in length than the summer periods, with the employees using the heating units even in May and September months. In contrast to the cooling consumption, heating consumption in years after the COVID-19 event had less of a decrease, with 2021 even experiencing a slight increase of 0.89% compared to 2020, attributed to the catch-up effect, which still keeps this value lower than the 2019 value. Even though the 2022 data is incomplete, the decrease compared to 2021 is only 58.35%, which has potential to become an increase in the following months since it is the lowest decrease in this period. Overall, the highest heating consumption was in building B, followed by building A, which can explain the overlap of both the heating and cooling units being used simultaneously in different buildings.

The water consumption does not follow such a clear seasonal trend; however, it does show a yearly trend experiencing an increase in consumption over the period of June until October, perhaps due to the higher water consumption of employees and a larger need of plant watering to be performed in this period. However, the trend does not occur before the COVID-19 event, as the consumption has drastically decreased in general, while beforehand it was quite high and steady. The consumption in 2020 has decreased by even 67.82% compared to 2019, which is the highest decrease of this comparison, and it has continued the decreasing trend in the following years. One anomaly can be seen at the very beginning of the dataset in January 2019 in building B, of which the source is unknown. Different buildings have slight variations in water consumption, with the highest being building B with 6 725.53 GJ consumed in total, followed by building A with 6 494.57 GJ, and then the highest difference is with building C which only consumed 4 570.67 GJ.

Analysis of the electricity consumption is quite limited due to the difference in dataset starting time, but the overall trend has been quite steady over the years, with a decrease in the beginning of the COVID-19 event. The trend is experiencing a smaller increase after the event., however, it is hard to compare the values and the changes since the consumption data has experienced a large increase as more sensors were implement in the buildings. Focusing only on building C which has a complete dataset, the decrease has been steady kept over the two years with values being half of the values in 2019. Also considering just the data from November 2020 until the end of the dataset, building B has the highest electricity consumption, followed by building A.

The differences between the consumption in different buildings comes from their general usage, where building B seems to be the most used in general, as there are also kitchens located in this area. This is also the building which connects A and C, hence the movements between it are increased.

5.1.2 CO_2 Production

The CO_2 production calculation is directly based on the consumption; hence the values follow the same trends and the same percentual differences with a different scale and range of values. However, the units of all the charts are the same (kg), hence the actual impact of each variable can be examined. By looking at the Y axis scales of the charts, it is perceived that the water consumption has the lowest CO_2 production, cooling and electricity have similar values, while the heating consumption has the highest impact on CO_2 emissions taking up more than 50% of the total CO_2 production.

Even though the cooling and heating consumption have the same emission factor, the heating production is overall three times higher compared to cooling and it has the highest total value overall. CO_2 production from electricity consumption has the second highest total value, even with large parts of the dataset values missing. Overall, water consumption has the lowest impact on CO_2 production as expected.

The trend of total CO_2 production is highly impacted by the seasonal changes of the heating consumption, as the water and electricity trends are quite steady in comparison. The overall change over the years is also impacted by the missing data from the electricity consumption, indicating that the total consumption and CO_2 production has increased 26% in 2021 compared to 2020, which is when additional sensors for electricity were installed. However, building C has no missing data, hence the change of the production and consumption in 2021 compared to 2020 is -15%, indicating that the decrease has been maintained, contrary to the overall data.

Overall, the energy consumption and CO_2 production in MBC has been kept on a decreasing trend starting during the COVID-19 event, with only a slight increase in some cases, which shows that the employees have not been returning to the office space with the same intensity as before. On the other hand, the changes could have been caused by the SAP efforts for sustainability practices and employees have been mindful of when and how they are using the appliances.

5.2 Time Series CO2 Production Charts Forecasts

The time series automatic forecasts for CO_2 production of each variable can be seen in *Figure 8*, however, the X axis scale is different from the charts before, as the forecast date range is quite small in comparison. Contrary to the analysis in the previous section, the SAC automatic forecast option on the time series charts has not detected the clearest seasonal cycle in cooling and heating fluctuations, mentioning that the trends cannot be recognized from the dataset.

The forecast quality for cooling, heating and water variables was rated 1/5, alerting that the confidence of the forecast is very low due to the pattern. Forecast quality for electricity was rated 5/5 with very high confidence, and for the total variable it was rated 4/5 with high confidence.

Without the data semantics surrounding the cooling and heating consumption, as well as the consumption overall, the forecast does not take into consideration trends and the Y axis limitations, and sometimes predicts negative values which are not possible in this scenario. The cooling consumption is meant to increase in the following months as the summer months approach, and the values are forecasted to be almost double than the summer months in 2021, and there is no difference in predicted values for each month and buildings which is quite unrealistic. It is unclear why the predicted values are starting the increase in 2022, when ever since 2019 the CO_2 production data for cooling has been decreasing considerably.

A similar scenario is shown in the heating forecast, although the decrease trend during the summer months has been overestimated with the predicted values even being negative for building B. The values for each month are decreasing and eventually reaching below zero, despite the fact that the values in the same months of 2021 are consistently zero, hence the shown forecast is unrealistic and quite impossible.

The forecast for water CO_2 production follows a fluctuation caused by the low water usage over the weekends, however, the negative values are still predicted for the future, which is again impossible. Overall, the trend is steady, but the confidence intervals and the values seem much more realistic and accurate for building A than for the other buildings. Perhaps if the negative values were converted to zeroes, the trend and prediction would be more reliable. Considering the fact that the SAC software issued the warning with the confidence rating of 1/5 for the three charts mentioned, the reliability of the forecasts is expectably low and cannot be considered as potential future trends.

On the other hand, the confidence rating for the electricity forecast is 5/5, displaying a clear trend in weekly fluctuations and no negative values, as the negative values for electricity consumption in a software company office are quite rare. The missing data in 2019 and 2020 in buildings A and B seem to have not influenced the forecast as it shows no variation in the building consumption. The range of predicted values seem to mostly follow the trend of previous six months, and not the months beforehand, which is believable due to the increase of overall electricity consumption in 2021 compared to the period before. The differences in consumptions of different buildings are clearly shown, with the building B having the highest values in comparison, and A having the lowest, which is consistent with the overall historical data.

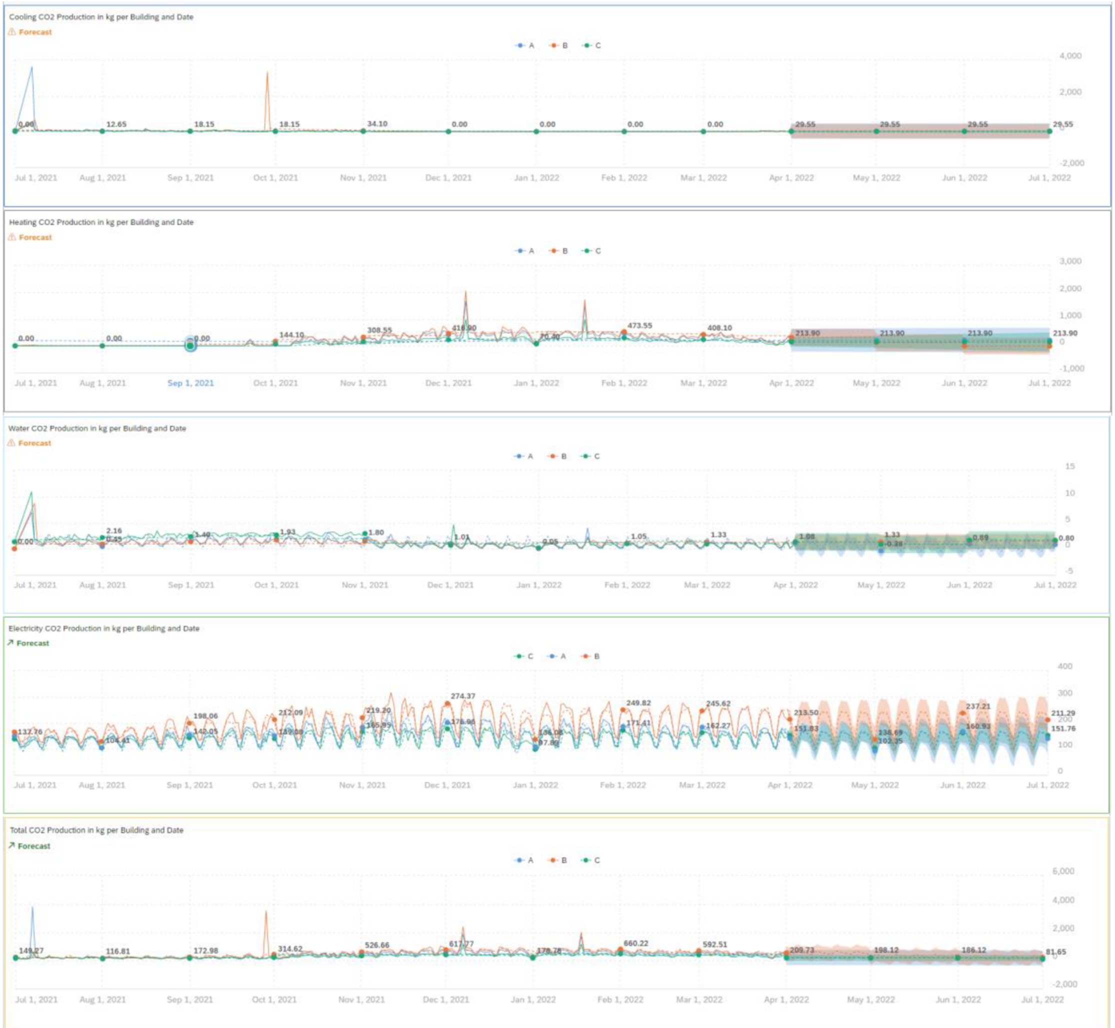


Figure 8: Charts from the CO2 Production SAC Story Page Showing the Forecasted CO2 Production per Variable and Building Over Time

The total CO_2 production forecast has quite a steady trend, with different values for each month which are decreasing compared to the previous winter months, as shown in the true trend influenced highly by the heating consumption. The confidence intervals are including negative values, but the actual predicted value is mostly positive. Prediction in general seems to correspond to the confidence rating of 4/5, except the negative values and the fact that the confidence interval for building A is not showing any variation compared to the other buildings. The building A confidence interval is strange in the total production forecast, as it is the only building where the weekly trends were recognized in both the water and electricity forecasts before.

The overall experience of using the automatic forecast in SAC time series charts is quite simple and straightforward, however, the prediction values seem to not take into account the semantic of the data, and while the weekly trends seem to be recognized in some cases, the trends on a yearly basis are not taken into account. Negative values have severely impacted most of the charts, where the predictions interpret steady zero values as possible negative values in the future decreasing trend, and an option to exclude them would be highly useful. However, the forecast feature does issue a warning about the prediction confidence, so the user can know that the prediction values cannot be trusted, and the ratings seem to be relatively accurate. But there are no options and suggestions on improving the accuracy of the forecasts, hence the predictions with the bad confidence ratings should just not be used in that case. Improving user interactivity with the forecast options, beyond selecting the forecast model, could improve user experience and the accuracy of the forecasts.

However, the overall trends of the forecast still indicate in which direction the future values are heading. If interpreting the forecasted negative values as possible zeros, it can be said that the future trend is negative and the overall decreasing in consumption will be kept in the future as well, based on the historical data.

5.3 Predictive Scenario

The first page of the predictive model results shows the top entities based on the MAPE KPI, as well as the median expected value of 193.15% and the average of 163.36%. The breakdown of individual MAPE values is 12.13% for electricity, 131.15% for water, 254.79% for cooling and 255.01% for heating.

Similar to the time series chart forecast, the confidence in the forecast for the cooling and heating consumption is quite low, with the MAPE values reaching more than 200%, and the water consumption over 100%, which are not acceptable errors. The electricity consumption has an acceptable error where the values can be considered with a lower percentage in mind, but the overall prediction confidence of consumption is decreased by the other variables. Therefore, the individual forecast charts, which can be seen in *Figure 9*, and values cannot be accepted for variables other than electricity.

The predicted chart for cooling is similar to the time series chart forecast, but the predicted values are not negative, only the floor of the confidence interval which is still realistic. Yearly seasonal trends in cooling consumption were not recognized as the prediction line is linear, and the values seem to follow a downward trend in the beginning of the summer months, with values being lower than 0.5, besides the fact that the usual values are usually higher than that, and the confidence intervals are quite large compared to the usual values, as the MAPE value indicates. Differences between the actual past data and the predicted model for the past are quite big, with the forecast being more steady and lower than the real values. In comparison to the forecast chart from *Figure 8,* the trend is more believable due to only positive values, but the confidence is quite low for the actual data points. However, the decreasing trend might still be considered, just not in the same intensity.

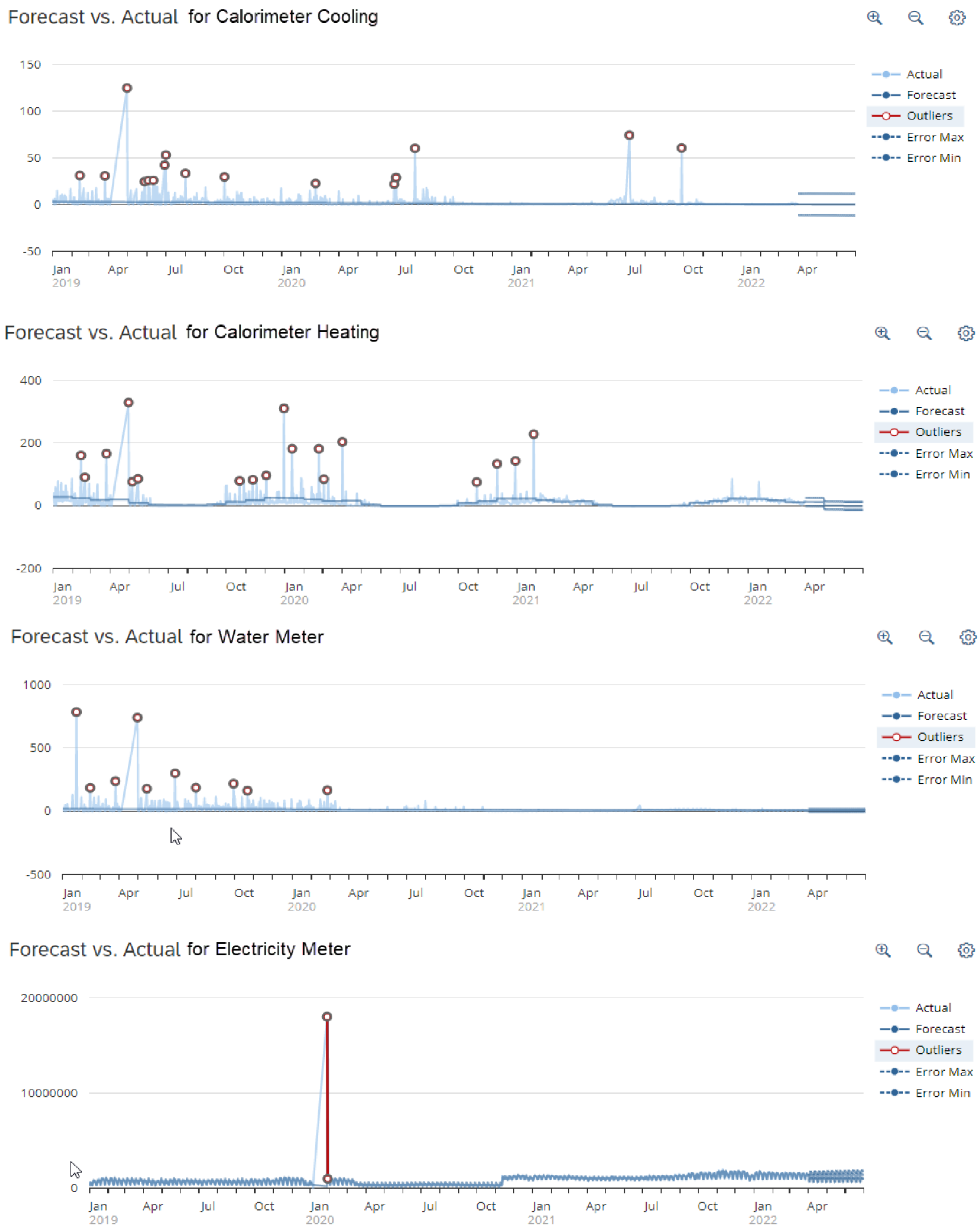


Figure 9: SAC Predictive Scenario Forecast Charts for Cooling, Heating, Electricity and Water Consumption

On the other hand, the heating consumption forecast shows a downward trend, settling at zero in the end of May 2022, which is quite realistic for summer months. However, the confidence intervals are quite large, and the MAPE value is higher compared to the cooling consumption. The actual past data contains large amounts of outliers, but the general predicted trend for the past fits well with the data, except the daily variations in the data, and the future prediction follows the consumption trend. Yearly seasonal variations were only recognized for the heating consumption and can be seen in *Figure 10.* The overall impact of the yearly cycles is recognized to be only 26.86%, but the actual impact per month is quite well presented in the chart with the highest consumption being in December and January. Other weekly trends are not noticed; hence the predicted past values are not directly fitted to the values, only on the overall trend increases and decreases. Therefore, the forecasted data points cannot be taken into consideration, but the overall trend predictions are quite well summarized on a seasonal level and the future trend can be viewed. Unfortunately, the predicted period does not include winter months, which would be valuable to see whether the consumption is actually decreasing overall.

The prediction of the values for electricity consumption seem to recognize the weekly pattern of the data in the trend lines, but the software did not recognize it separately as it did for the seasonal consumption of heating. Confidence intervals are still quite large, but the overall values are kept decreasing in comparison to 2021 data, and the trend seems to start to decrease in the start of the summer months, which was the case in the historical data. The fitting of the actual data values and the forecasted values for the past is quite similar, corresponding to the MAPE value which was most likely kept the lowest due to the weekly trend recognition. When excluding the period when only building C was monitored, the historical and predicted trend is decreasing after the catch-up effect in the end of 2021. However, a very high outlier was recognized in the forecast on January 28, 2020, with a very high value of 18 033 349 Wh consumed in just one day. This date was found in the original dataset and it seems that 15 sensors in building C recorded extremely high values, for which the source is unknown.

Water consumption weekly trend was not recognized by the prediction model, where the forecasted trend is mostly linear. Besides the overall shape, the future trend seems to be decreasing as the historical data indicates. The result seems similar to the prediction of cooling and heating consumption, but the overall MAPE value is much lower in comparison, which is still quite high. Another trend was recognized for the water consumption, which can be seen in *Figure 10,* where the water consumption seems to increase in the second half of the year compared to the first, which was recognized in the water consumption charts as well for the period between June and October. However, the overall impact seems to be quite low with only 8.81% and a slight difference between the year periods. The overall downward trend seems to be recognized, but the prominent weekly trend was not, and the forecast does not include the future increase period in the second half of the year due to the limited end date. The fitting of the predicted data to the past data is quite poor due to the linear values, which has probably influenced the high MAPE value.

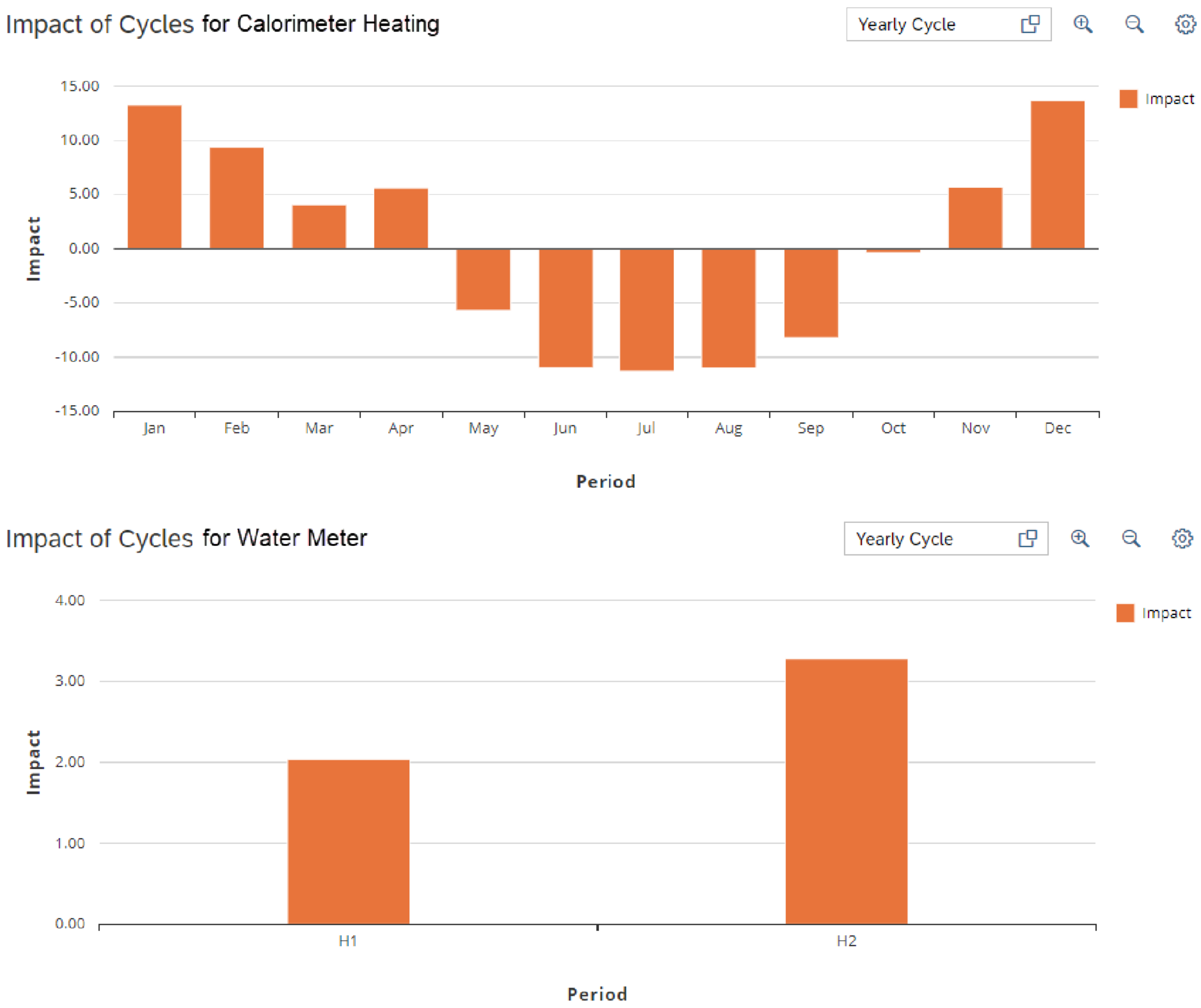


Figure 10: SAC Predictive Scenario Charts for the Heating and Water Consumption Impact of Yearly Cycles

Overall, it seems that the prediction trends have improved in comparison to the time series chart forecasts, due to the fact that all negative forecast values were converted to zero and not included. However, the pattern recognition did not recognize the seasonal trends in cooling consumption which were recognized for heating, perhaps due to the cooling unit usage during winter months. The weekly trends of usage were exclusively recognized in electricity consumption only, where the forecast was produced with the highest confidence in both tools, despite the missing data. The yearly trend for water consumption was recognized, but the impact is not as high and the weekly trends are not present, making the reliability of the forecast poor. Since most predicted values are linear, except for electricity consumption, the actual values seem unlikely to occur in the future. By inspecting the MAPE values, it can be seen that only electricity consumption scenario can be taken into consideration when forecasting the future consumption data points. The trends also seem to not be influenced by the past values in 2019 before the COVID-19 event, with no drastic increases in the future.

Besides the actual forecasted values and weekly trends, the overall trends are generally recognized on a higher monthly scale and the predicted lines show that there is a decreasing trend even after the catch-up effect. The poor confidence in results generally comes from the missed weekly pattern recognition, hence the predicted values are not fitted properly, but on a monthly scale. Since the overall trends seem to fit the historical patterns, it can be said that all measures are heading towards a general decrease in consumption.

Compared to the carbon emissions of all sources on a global SAP scale, the catch-up effect and decrease trend seems to be similar, as the trend follows a large decrease in 2020 followed by a slight increase in 2021 which is still lower than the production before COVID-19 (SAP SE, 2021). However, in comparison to the research in MBC, the data is more extensive and includes many more unavailable factors. The energy consumption trend is also similar, with MBC having a better general decrease trend in consumption (SAP SE, 2021). The research also implies that the policy to enable employees to stay in the remote working model is helping in the SAP's sustainability targets.

The usability of the software is quite accessible and straightforward, with not many user options and a lot of official documentation that describes each step in detail. However, the initial investment in understanding the general architecture and setup of SAC is quite important, as the user can feel overwhelmed with the interface. As the user gets more comfortable, the additional functionalities are easier to understand and use correctly. The limitations which come with the ease of usage are that the results can be affected by the limited options of personalization, especially with the time series chart forecasts. If there were more customization options, especially with the negative value conversion and data semantic, perhaps the results would be more accurate and realistic, as it is the case for electricity consumption.

On the other hand, the predictive scenario brings more options to be included in the forecast, but both tools have their limitations. While the time series chart can forecast any measure put into the chart, the predictive scenario only allows the prediction of original measures from the dataset, and no calculations. However, it provides more detailed analysis, with the MAPE accuracy measures and cycle recognition, which is not as accurate as expected. Weekly trends can be recognized, but in the case of MBC dataset, it was not noticed even for the electricity measure which was forecasted with the weekly variations. Yearly seasonal trend was only noticed for heating, while the cooling measure was not recognized due to a slightly smoother trend. The poor MAPE values are speculated to come from poor weekly pattern recognition, but if the user could select the future scale of the forecast, such as only monthly values for fitting and trends, the confidence of the results would be much higher.

The user is notified of the inaccuracy of the predictions, therefore a scenario where an ill-informed user does not recognize the bad accuracy of the forecast seems unlikely. Even without the warnings, the user might notice the large confidence intervals for the predicted values where the predicted value can take almost any possible value and the prediction makes no sense with such a high error. Unfortunately, there are no textual descriptions and insights in the predictive scenarios results page, just the charts and values, hence the

SAC Help Page must be consulted for the interpretation of the results, which most users do not consider. Adding short descriptions of the prediction results could incredibly improve the user experience, as well as adding additional information on why the forecast has such poor results and how that can be improved.

5.4 Discussion

Based on the results in the previous section, it can be seen that there are several trends relating to energy consumption, and consequently to the respective CO_2 production as well. With regards to the environmental impact indicators, energy consumption and CO2 production seem to be good corporate indicators (Herva, Franco, Carrasco, & Roca, 2011). In the case of cooling and heating consumption, it can be hard to differentiate whether the changes in consumption are related to office employee attendance, or perhaps the changes in local temperature. The electricity consumption yearly increases also cannot be taken into consideration, as the missing sensor data is heavily affecting the results. Overall, it seems that the consumption is starting to decrease after the catch-up effect, and the general usage has a definite decrease in comparison to 2019. Other research also suggests that the catch-up effect has been generally more severe in colder climates, which Czech Republic can be classified as, as it is directly dependent on heating consumptions (Duarte & Cortiços, 2022).

The COVID-19 event has greatly decreased the energy consumption of the SAP office in this period, however, studies suggest that the actual change in CO_2 emissions as a result of teleworking resulted in a modest decrease or even an increase as the employees switch to a home office environment (Santos & Azhari, 2022; Hook, Court, Sovacool, & Sorrel, 2020). Unfortunately, the more extensive data is hard to collect for SAP employees and it is currently unavailable. Another data limitation is the fact that some variables which can provide great insights into the indirect CO_2 production, such as commuting, office building materials, waste management, and data center emissions, are not allowed to be analysed and shared in detail due to confidentiality. Nevertheless, mitigating and controlling energy usage in office buildings based on occupancy seems to have a high impact on energy consumption, which is one of the newer policies in MBC (O'Brien & Yazdani Aliabadi, 2020). Monitoring the actual occupancy and implementing tools for office space booking can help SAP reach their targets regarding the environmental impact in time.

In comparison to other research, the automatic forecast feature in time series charts seems to be quite generalized on a larger scale as well, as it predicts more linear trends and relatively high confidence intervals (Nazarov, Kovtun, & Reichert, 2020). Even though the other research does not include weekly and seasonal trends, the situation compared to the predictions performed in this research is similar with poor confidence in the predictions. However, compared to similar research which takes place outside of SAP's BTP, there are many more possibilities and functionalities that can be implemented, especially with the forecasted periods that can be chosen by the creator of the forecast (Meng & Noman, 2022). This would highly influence the prediction trends in the case of energy consumption, especially seeing the heating consumption forecasted trends in the winter months. On the other hand, it would not contribute to the DDEM empowerment and enablement in a corporate setting, which seems to have a large impact on business decisions in a corporate context (Gole & Shiralkar, 2020). Nevertheless, the availability of data and software solutions seems to empower and motivate a large range of employees in different job fields and increase general knowledge (Lefebvre, Legner, & Fadler, 2021).

Perhaps it would be a better investment to contribute the learning time to studying a programming language, such as R, and create your own machine learning models with the R visualizations functionality in SAC, which can increase the accuracy and reliability of the results. Unfortunately, Python integration in SAC is not available, however, the integration is possible in the SAP BTP solution with other tools.

6 Conclusion

Findings from the energy consumption historical data research can be summarized by the catch-up effect due to the COVID-19 related office shutdowns. Similar to the SAP's global energy consumption trends, there was a large decrease in 2020 in comparison to 2019 data. However, the negative trend was not kept in the same magnitude in 2021 and 2022 as the employees eventually started going back to the office, but the consumption and corresponding CO_2 production has still not returned to pre-pandemic level. The prediction trends indicate that the catch-up effect will soon stabilize or potentially decrease in the near future, as the employees keep their preference for the remote working model. Following the predicted trends, SAP has the opportunity to reach their emission and energy consumption targets in time, if the policies implemented align with their goals.

In the case of MBC, the most concerning issues come from the heating consumption, which is kept relatively high and has the highest contribution to CO_2 total production. Investing in new innovations and limiting energy usage based on office occupancy seems to be a favorable implementation on reducing the high emissions. However, the research is highly limited, especially in case of electricity consumption, due to data unavailability of sensors and measures which also indirectly influence CO_2 emissions. The future trends also have low confidence level in most measures of energy consumption; hence the future data points must be considered with skepticism.

One of the main influences of the data trends in energy consumption have been the seasonal and weekly cycles in data. With the predictions only forecasting the following three months, it is not possible to effectively analyze the future trends of both heating and cooling consumption and CO_2 production, as the winter and summer months cannot both be taken into consideration. The weekly patterns have also been unrecognized by the predictive models, making the MAPE values and the overall confidence in the predictions very low. Improving the prediction model pattern recognition, or perhaps creating multiple predictive scenarios for the same data, could heavily improve the accuracy of the results.

Comparing the results of the Predictive Scenario tool and the time series chart forecast feature, one of the main differences is in the negative values restriction. Even though the time series forecast does not require almost any background knowledge to interpret, the negative values prediction in place of a steady zero consumption trend, makes the results of the forecast completely unrealistic. On the other hand, the predictive model training is limited in the personalization of the predictions by insisting on forecasting only the values included in the original dataset and no calculations to be used.

Considering the poor accuracy and overgeneralized predictions of SAC's smart features, it is hard to say whether the DDEM has negatively or positively influenced data-driven decision making in the corporate environments. Even though the users do not explicitly need to have any knowledge about machine learning algorithms to use them, the user should be aware and educated on the interpretation of the results and he must be able to assess the reliability of the provided forecast. The time invested in learning to use SAP's BTP is heavily affected by learning the underlying architecture and integrations. By providing more insights into the results of the predictive models and forecasts, the user experience can be heavily enhanced. Nevertheless, the access to data and tools for analysis on the unified SAC platform has enabled many employees in all job fields to educate themselves and investigate data on their own which has empowered them to make more educated decisions.

7 References

Rheinwerk Publishing. (2022). *Rheinwerk Publishing -- Publisher of SAP PRESS*. Retrieved from SAP PRESS: https://www.sap-press.com/the-publisher/

Abd Elmonem, M., Nasr, E., & Geith, M. (2016). Benefits and challenges of cloud ERP systems – A systematic literature review. *Future Computing and Informatics Journal*, 1-9.

Acheampong, A. O. (2018). Economic growth, CO2 emissions and energy consumption: What causes what and where? *Energy Economics*.

Ahmad, M., Mourshed, M., Mundow, D., Sisinni, M., & Rezgui, Y. (2016). Building energy metering and environmental monitoring – A state-of-the-art review and directions for future research. *Energy and Buildings*, 85-102.

Ara Begum, R., Lempert, R., Ali, E., Benjaminsen, T., Bernauer, T., Cramer, W., . . . Wester, P. (2022). *Point of Departure and Key Concepts. In: Climate Change 2022: Impacts, Adaptation.* Cambridge, UK and New York: Cambridge University Press.

Awasthi, P., & George, J. (2020). A case for Data Democratization. *AMCIS*, 23.

Banda, S., Chandra, S., & Gooi, C. (2022). *SAP Business Technology Platform: An Introduction.* Boston: Rheinwerk Publishing Inc.

Bertram, E., Charlton, J., Hollender, N., Holzapfel, M., Licht, N., & Paduraru, C. (2021). *Designing Dashboards with SAP Analytics Cloud.* Boston: Rheinwerk Publishing Inc.

BRE. (2020). *BREEAM Rating System | 5-minute guide*. Retrieved from Ongreening: https://ongreening.com/breeam-rating-system-5-minute-guide/

BRE Group. (2018). *How BREEAM works*. Retrieved from BRE: https://bregroup.com/products/breeam/how-breeam-works/#:~:text=The%20BREEAM%20ratings%20range%20from,assessment%20and%20its%20BREEAM%20rating.

Dacre, H., Western, L., Say, D., O'Doherty, S., Arnold, T., Rennick, C., & Hawkins, E. (2021). Detectability of COVID-19 global emissions reductions in local CO2 concentration measurements. *Environmental Research Letters, 16*(9).

Duarte, C., & Cortiços, N. (2022). The Energy Efficiency Post-COVID-19 in China's Office Buildings. *Clean Technologies, 4*(1).

Eldridge, T. (2022, December 7). *Measuring the Carbon Footprint of Employees Can Help You Make More Informed and Strategic Business Decisions*. Retrieved from SAP Community: https://blogs.sap.com/2022/12/07/measuring-the-carbon-footprint-of-employees-can-help-you-make-more-informed-and-strategic-business-decisions/

Flick, U. (2018). *Designing Qualitative Research.* London: SAGE Publications.

Foss, G., & Modderman, P. (2019). *Practical Data Science with SAP.* Sebastopol: O'Reilly Media, Inc.

Gole, V., & Shiralkar, S. (2020). *Empower Decision Makers with SAP Analytics Cloud: Modernize BI with SAP's Single Platform for Analytics.* Berkeley: Apress.

Herva, M., Franco, A., Carrasco, E., & Roca, E. (2011). Review of corporate environmental indicators. *Journal of Cleaner Production, 19*(15).

Hook, A., Court, V., Sovacool, B., & Sorrel, S. (2020). A systematic review of the energy and climate impacts of teleworking. *Environmental Research Letters, 15*(9).

Ilinitch, A., Soderstrom, N., & E. Thomas, T. (1998). Measuring corporate environmental performance. *Journal of Accounting and Public Policy*, 383-408.

Lefebvre, H., Legner, C., & Fadler, M. (2021). Data democratization: toward a deeper understanding. *International Conference on Information Systems (ICIS 2021).*

Meng, Y., & Noman, H. (2022). Predicting CO2 Emission Footprint Using AI through Machine Learning. *Atmosphere, 13*(11).

Nazarov, D., Kovtun, D., & Reichert, T. (2020). SAP Analytics Cloud: intellectual analysis of small and medium-sized business activities in Russia in the context of COVID-19. *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*, (pp. 1-6).

Nazarov, D., Morozova, A., & Kokovikhin, A. (2019). SAP Analytic Cloud: a tool for the formation of professional competencies of business analyst. *International Conference of Information Systems and Design.* Moscow: CEUR Workshop Proceedings.

O'Brien, W., & Yazdani Aliabadi, F. (2020). Does telecommuting save energy? A critical review of quantitative studies and their research methods. *Energy and Buildings, 225.*

Olivier, J., & Peters, J. (2020). *Trends in global CO2 and total greenhouse gas emissions: 2020 Report.* The Hague: PBL Netherlands Environmental Assessment Agency.

O'Neill, B., van Aalst, M., Zaiton Ibrahim, Z., Berrang Ford, L., Bhadwal, S., Buhaug, H., . . . Warren, R. (2022). *Key Risks Across Sectors and Regions. In: Climate Change 2022: Impacts, Adaptation and Vulnerability.* Cambridge, UK and New York: Cambridge University Press.

REICO . (n.d.). *Metronom Business Center*. Retrieved from REICO: https://www.reico.cz/en/cs-nemovitostni-fond/real-estate-portfolio/metronom-praha-cr

Santos, G., & Azhari, R. (2022). Can we save GHG emissions by working from home? *Environmental Research Communications, 4*(3).

SAP SE. (2016, July 25). *Terms of Use for SAP Websites | About SAP*. Retrieved from SAP: https://www.sap.com/corporate/en/legal/terms-of-use.html

SAP SE. (2021). *SAP Integrated Report.* Retrieved from SAP: https://www.sap.com/integrated-reports/2022/en.html

SAP SE. (2022). *Augmented Analytics (Smart Features).* Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/cd897576c3344475a208c2f7a52f151e.html

SAP SE. (2022). *Getting Started with SAP Analytics Cloud Stories*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/fa31c0480f1f4c39af0692cd9a80f95b.html

SAP SE. (2022). *Import and Prepare Fact Data for a Model with Measures*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f08100

2fd3691d86a7/75683f452594426bb05cb49dc4c5af99.html#loioaa7a3a58a4d3497b9d4218d7a1181167

SAP SE. (2022). *Running a Forecast in a Time Series or Line Chart*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/8f9665dcf53f4f3f90e920995cdd5f05.html

SAP SE. (2022). *Smart Predict – Using Predictive Scenarios*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/37db2128dab44d15b46e1918829c1ff1.html

SAP SE. (2022). *Training a Predictive Model*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/4ac765768eb84806aabee0874c36809c.html

SAP SE. (2022). *Understand Licenses, Roles, and Permissions*. Retrieved from SAP Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/72ae65446c7943be80905c1d83a57a4a.html

SAP SE. (2022). *Update and Schedule Models*. Retrieved from SAP Help Portal : https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/6aa6a99d4eb5467896c4fd47bc9c7642.html

SAP SE. (2022). *Variable Data Types*. Retrieved from SAC Help Portal: https://help.sap.com/docs/SAP_ANALYTICS_CLOUD/00f68c2e08b941f081002fd3691d86a7/f1b2f80a12604f08bb358ce0a04ce1d6.html

SAP SE. (2022). *Worldwide Directory*. Retrieved from SAP: https://www.sap.com/cz/about/company/office-locations.html

SAP SE. (2023). Internal SAP Source.

SAP SE. (n.d.). *Prague | Office Locations*. Retrieved from SAP Careers: https://www.sap.cn/about/careers/joining/locations/prague.html

SAP SE. (n.d.). *What is SAP? | Definition and Meaning*. Retrieved from SAP: https://www.sap.com/about/company/what-is-sap.html

Sidiq, A. (2022). *SAP Analytics Cloud.* Boston: Rheinwerk Publishing Inc.

Silverman, D., & Marvasti, A. (2008). *Doing Qualitative Research: A Comprehensive Guide.* London: SAGE Publications.

Vahlkamp, D., & Vahlkamp, L. (2020). *Smart Predict in SAP Analytics Cloud.* Gleichen: Espresso Tutorials.

Wang, S., & Ma, Z. (2008). Supervisory and Optimal Control of Building HVAC Systems: A Review. *HVAC&R Research, 14*(1).

White Star Real Estate. (n.d.). *Metronom - offices for rent in Prague 5*. Retrieved from Metronom: https://www.metronombc.cz/en/

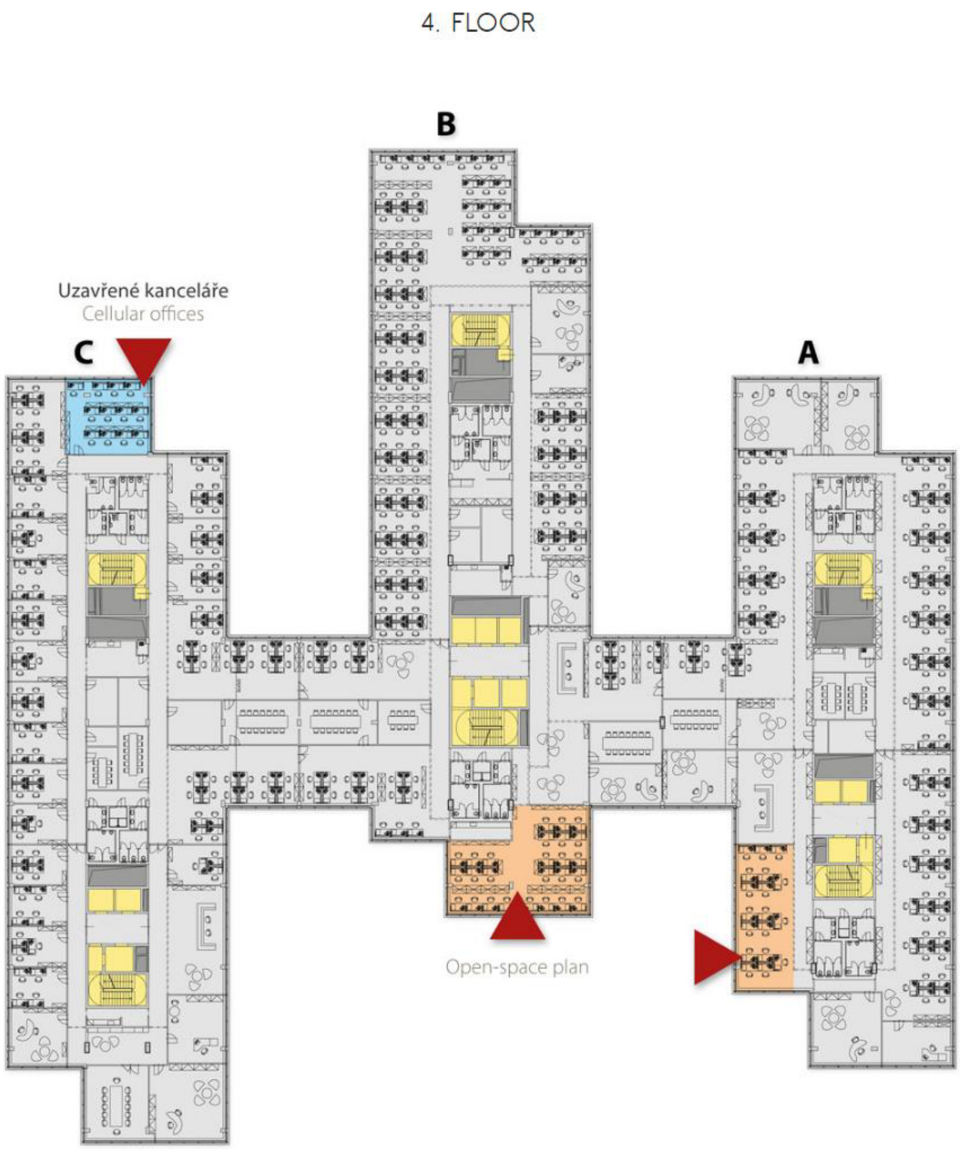
World Resources Institute. (2015). *Corporate Standard.* Retrieved from Greenhouse Gas Protocol: https://ghgprotocol.org/corporate-standard

Worrell, E., Bernstein, L., Roy, J., Price, L., & Harnisch, J. (2009). Industrial energy efficiency and climate change mitigation. *Energy Efficiency*, 109-123.

Zhao, H.-x., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, 3586-3592.

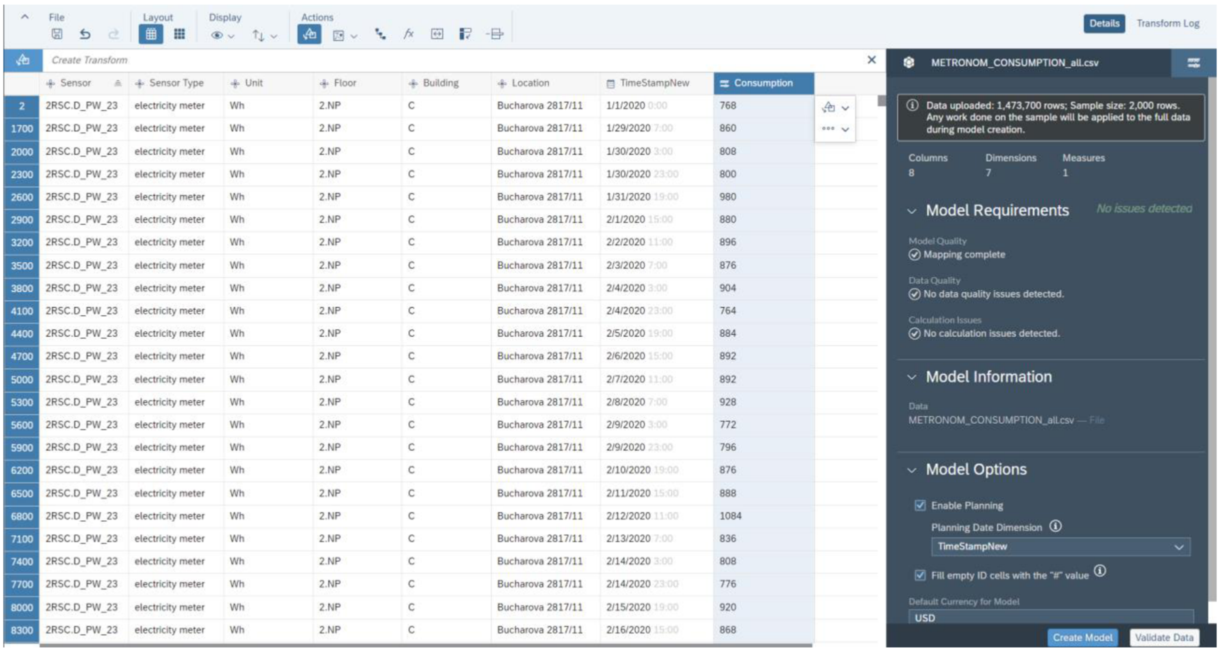
8 List of Figures

Appendix A



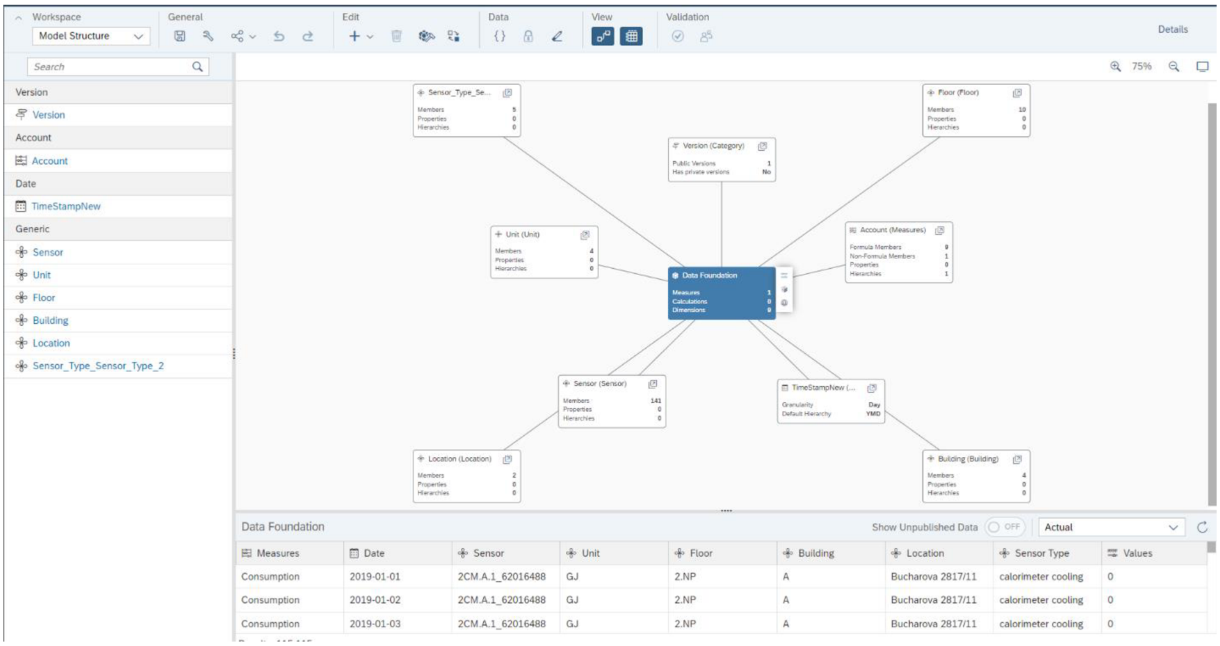
Appendix A: Schema of the Fourth Floor Plan in the Metronom Business Center (White Star Real Estate, n.d.)

Appendix B



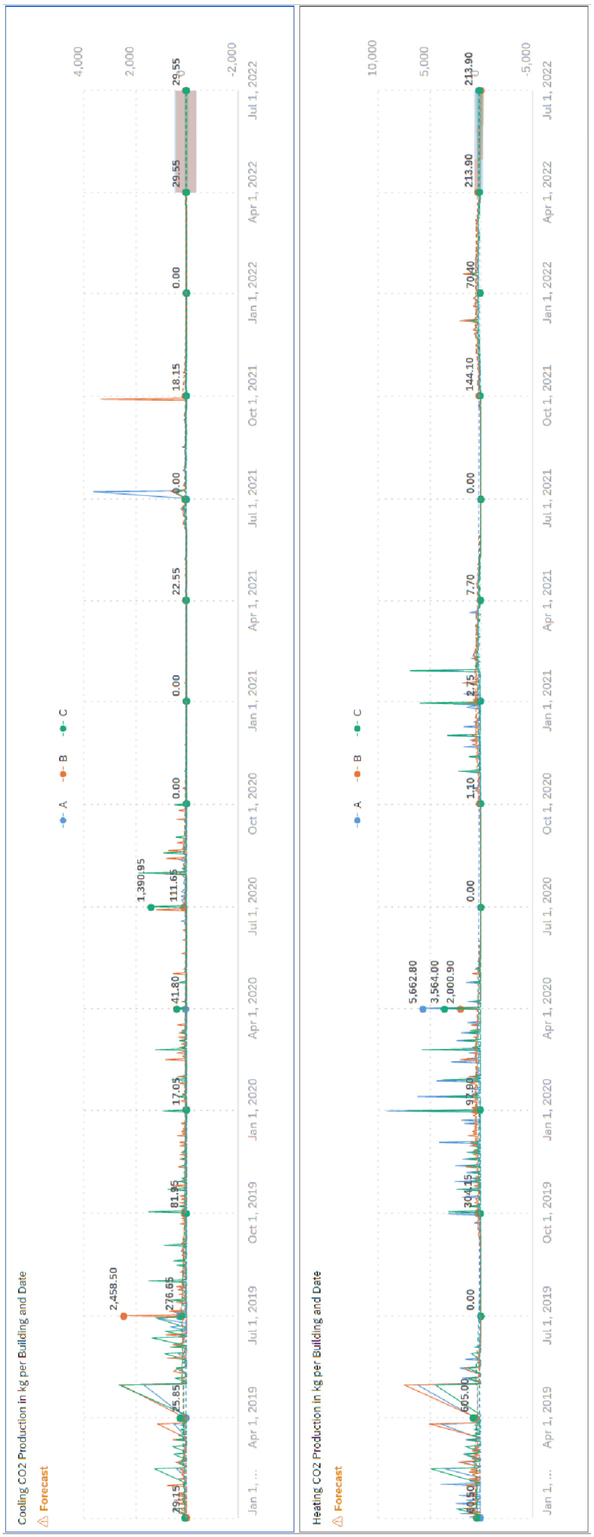
Appendix B: Screen Capture of the Data Import Screen in the SAC Modeler with the Data Sample Size

Appendix C



Appendix C: Screen Capture of the SAC Modeler Tool with the Model Structure Diagram of the METRONOM_CONSUMPTION_thesis Model

Appendix D



Appendix D.1: Cooling and Heating Charts from the CO2 Production SAC Story Page Showing the Forecasted CO2 Production per Variable and Building Over Time

Appendix E.2: Water, Electricity and Total Charts from the CO2 Production SAC Story Page Showing the Forecasted CO2 Production per Variable and Building Over Time