PREDICTIVE FAULT DETECTION AND EQUIPMENT MAINTENANCE USING AI-POWERED TOOL

T. ASWIIN

02/02/2024

Abstract

In today's fast-paced world, electricity consumption holds a vital position in fulfilling the energy needs of modern societies. As the demand for electricity keeps growing, optimizing energy usage becomes extremely important. Thankfully, advancements in technology have led to the emergence of machine learning, a potent tool that can predict electricity consumption with remarkable accuracy. By analysing vast historical data, including electricity usage, weather patterns and seasonal variations, machine learning algorithms can reveal trends, patterns, correlation in vast datasets and predict future energy consumption patterns. The objective of this project is to build an AI-driven tool that employs machine learning to help reduce energy costs, by decreasing waste, optimizing production, and anticipating malfunctions, the last of which helps workers keep up with equipment maintenance before damage occurs.

1.0 Introduction

India, one of the largest and fastest growing economies globally, with an average growth rate of approximately 10.4% in the year 2021, is also the world's third largest consumer of primary energy after China and the USA (*BP Statistical Review of World Energy,* 2022). Country's energy basket is dominated by fossil fuels like coal, oil, natural gas, and biofuels. While the total carbon emission due to intensive fossil fuels accounted for over 12.2% *(*BP Statistical Review of World Energy *2022).*

Australia	5.60	5.52	5.59	5.63	5.76	5.78	5.76	5.84	6.04	5.74	**5.72**	♦	0.2%	1.0%
Bangladesh	0.98	1.05	1.08	1.13	1.31	1.33	1.38	1.46	1.63	1.58	**1.65**	5.0%	5.4%	0.3%
China	112.80	117.43	121.85	125.41	127.02	129.15	133.60	138.88	143.92	147.58	**157.65**	7.1%	3.4%	26.5%
China Hong Kong SAR	1.19	1.14	1.17	1.14	1.18	1.21	1.29	1.31	1.24	0.93	**0.88**	-5.5%	-3.0%	0.1%
India	23.94	25.37	26.26	28.08	28.91	30.16	31.32	33.34	34.15	32.19	**35.43**	10.4%	4.0%	6.0%
Indonesia	6.79	7.03	6.78	6.86	7.00	7.07	7.39	8.19	8.75	8.10	**8.31**	2.9%	2.0%	1.4%
Japan	20.14	19.99	19.84	19.34	19.07	18.82	19.05	18.95	18.51	17.13	**17.74**	3.8%	-1.3%	3.0%
Malaysia	3.42	3.74	3.91	3.95	4.01	4.22	4.28	4.35	4.47	4.10	**4.19**	2.4%	2.1%	0.7%
New Zealand	0.85	0.86	0.86	0.90	0.91	0.92	0.93	0.93	0.95	0.86	**0.84**	-2.0%	-0.1%	0.1%
Pakistan	2.66	2.48	2.71	2.79	2.94	3.20	3.39	3.50	3.54	3.52	**3.86**	10.0%	3.8%	0.6%
Philippines	1.25	1.29	1.39	1.46	1.60	1.74	1.91	1.97	2.03	1.84	**1.96**	7.1%	4.6%	0.3%
Singapore	2.99	3.00	3.07	3.15	3.35	3.48	3.59	3.61	3.53	3.44	**3.46**	0.8%	1.5%	0.6%
South Korea	11.50	11.64	11.68	11.79	12.00	12.35	12.48	12.68	12.51	11.99	**12.58**	5.2%	0.9%	2.1%
Sri Lanka	0.27	0.28	0.29	0.31	0.34	0.37	0.37	0.38	0.39	0.37	**0.38**	3.1%	3.5%	0.1%
Taiwan	4.67	4.64	4.74	4.84	4.84	4.87	4.86	4.93	4.84	4.70	**4.98**	6.2%	0.6%	0.8%
Thailand	4.51	4.83	4.86	5.01	5.12	5.23	5.32	5.48	5.49	5.07	**5.11**	1.1%	1.2%	0.9%
Vietnam	2.15	2.25	2.38	2.63	3.00	3.24	3.48	3.91	4.34	4.22	**4.32**	2.6%	7.2%	0.7%
Other Asia Pacific	7.15	7.38	7.40	7.80	8.19	8.79	9.65	10.05	10.30	10.48	**10.65**	1.8%	4.1%	1.8%
Total Asia Pacific	207.66	214.59	220.48	226.61	230.59	235.62	243.28	252.74	259.51	256.69	**272.45**	6.4%	2.8%	45.8%
Total World	**520.90**	**528.18**	**537.56**	**543.52**	**548.14**	**555.91**	**566.66**	**582.38**	**587.43**	**564.01**	**595.15**	**5.8%**	**1.3%**	**100.0%**
of which: OECD	233.96	231.81	234.08	232.13	232.85	233.85	235.86	239.69	237.41	220.20	**229.89**	4.7%	-0.2%	38.6%
Non-OECD	286.94	296.36	303.48	311.39	315.29	322.05	330.81	342.69	350.01	343.82	**365.26**	6.5%	2.4%	61.4%
European Union	63.87	63.17	62.69	60.48	61.26	61.95	62.55	62.77	61.77	57.07	**60.11**	5.6%	-0.6%	10.1%

Figure 1. World Primary Energy Consumption 2022 (in Exajoules)

Australia	406.8	398.1	395.0	398.5	407.2	405.6	403.7	401.5	406.6	378.2	**369.4**	-2.1%	-1.0%	1.1%
Bangladesh	56.5	60.4	62.7	65.4	78.0	79.2	83.2	89.2	99.1	96.8	**100.9**	4.5%	6.0%	0.3%
China	8793.5	8978.7	9219.1	9256.7	9226.2	9234.4	9444.9	9676.0	9868.5	9974.3	**10523.0**	5.8%	1.8%	31.1%
China Hong Kong SAR	92.0	88.7	91.5	89.7	90.5	92.7	98.9	100.1	94.6	68.1	**64.6**	-5.0%	-3.5%	0.2%
India	1728.4	1861.4	1934.0	2090.7	2146.5	2241.5	2320.2	2442.6	2465.8	2281.2	**2552.8**	12.2%	4.0%	7.5%
Indonesia	470.6	489.5	460.6	469.0	489.0	487.8	514.6	565.6	613.2	560.8	**572.5**	2.4%	2.0%	1.7%
Japan	1207.5	1293.8	1282.2	1248.7	1209.1	1189.3	1182.7	1161.5	1121.7	1029.5	**1053.7**	2.6%	-1.4%	3.1%
Malaysia	209.8	227.7	234.1	243.1	247.1	252.6	241.2	251.1	256.9	240.6	**238.6**	-0.6%	1.3%	0.7%
New Zealand	34.4	36.0	35.6	35.6	36.2	35.5	37.6	37.4	38.5	33.4	**32.7**	-1.8%	-0.5%	0.1%
Pakistan	145.2	144.6	145.3	151.7	160.3	176.5	188.7	196.9	207.1	206.3	**226.4**	10.0%	4.5%	0.7%
Philippines	80.7	83.1	91.9	97.3	106.2	116.4	128.9	133.7	140.7	127.2	**136.8**	7.8%	5.4%	0.4%
Singapore	192.7	192.0	191.4	191.0	202.8	217.0	228.9	225.2	217.3	211.6	**215.7**	2.2%	1.1%	0.6%
South Korea	613.7	612.6	620.0	615.0	622.8	633.2	641.8	659.1	635.3	588.8	**603.8**	2.8%	-0.2%	1.8%
Sri Lanka	16.4	18.2	15.9	19.6	21.6	24.7	24.4	23.4	25.3	23.8	**22.8**	-3.9%	3.4%	0.1%
Taiwan	273.0	266.8	268.4	275.3	275.5	280.8	286.7	284.7	279.1	264.6	**279.2**	5.8%	0.2%	0.8%
Thailand	249.3	266.7	265.8	273.7	281.1	286.8	287.5	293.2	288.4	270.0	**269.4**	♦	0.8%	0.8%
Vietnam	132.1	129.7	134.5	151.0	184.2	198.7	200.4	241.9	292.6	276.3	**272.7**	-1.0%	7.5%	0.8%
Other Asia Pacific	431.5	443.3	422.2	447.7	476.1	519.6	597.9	613.1	641.0	630.4	**638.0**	1.5%	4.0%	1.9%
Total Asia Pacific	14813.5	15264.7	15555.4	15790.4	15910.2	16091.5	16486.7	16965.4	17249.9	16829.0	**17734.6**	5.7%	1.8%	52.3%
Total World	**31904.6**	**32241.1**	**32710.9**	**32820.2**	**32837.4**	**33020.6**	**33426.4**	**34148.5**	**34095.8**	**32078.5**	**33884.1**	**5.9%**	**0.6%**	**100.0%**
of which: OECD	12842.4	12655.7	12741.3	12518.7	12436.2	12348.5	12346.5	12459.4	12083.5	10744.7	**11292.5**	5.4%	-1.3%	33.3%
Non-OECD	19062.2	19585.4	19969.6	20301.4	20401.2	20672.1	21079.9	21689.2	22012.3	21333.9	**22591.5**	6.2%	1.7%	66.7%
European Union	3300.8	3218.7	3146.0	2980.8	3043.7	3075.7	3095.7	3069.0	2931.5	2564.2	**2728.2**	6.7%	-1.9%	8.1%

Figure 2. World $C0_2$ Emission 2022 (in Exajoules)

There are many reasons for energy waste: breakage, inefficiency, idling machinery drawing electricity, or bulk adjustments in a pharmaceutical plant when the equipment runs without manufacturing. In the past, India had attempted to improve the energy efficiency of the industrial sector through various voluntary measures that were not legally binding. The Energy Conservation Act of India was formulated in 2001 to devise government policies for energy conservation that would include penalties for non-compliance. The Act listed fifteen energy intensive industries as designated consumers (The Gazette of India, Ministry of Law, Justice and Company Affairs, 2001). The first phase of the PAT scheme assigned mandatory energy intensity reduction targets to DCs in eight out of fifteen industries.

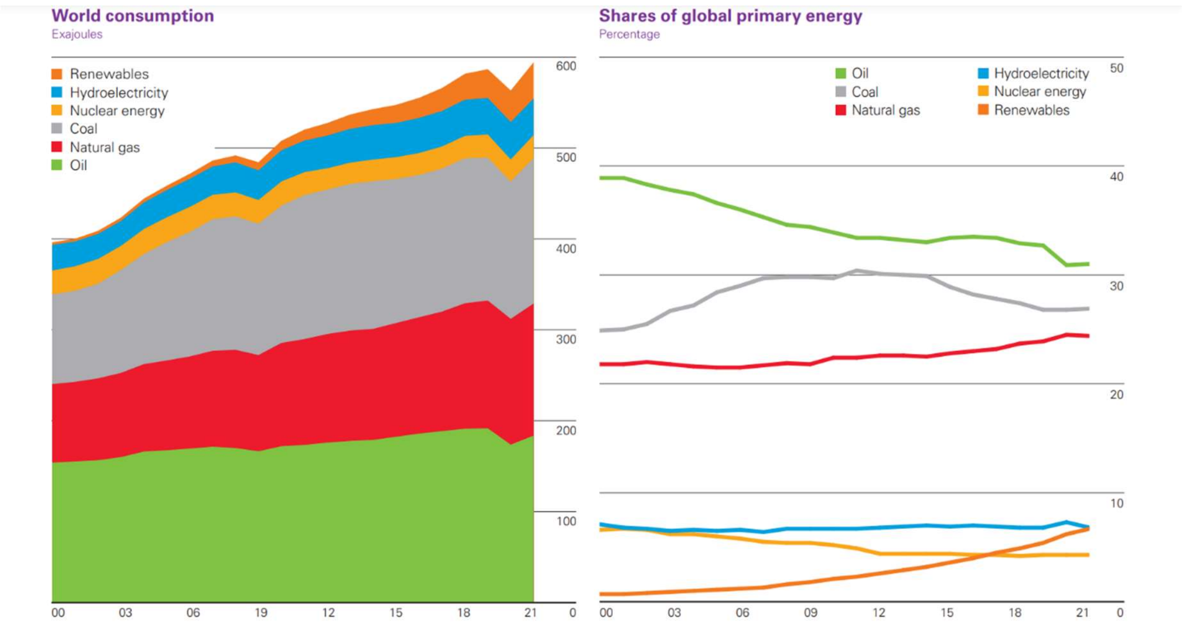


Figure 3. Area Chart and Line Graph representing World Energy Consumption 2022

Big businesses with resources in terms of capital and manpower are able to achieve zero-carbon efficiency updates to their business, making energy efficiency a big business in itself. The reality is that most Small and medium-sized enterprises (SMEs) simply do not have the resources on their own to pursue significant, lasting energy efficiency changes to their business. They do not have specialist energy efficiency departments or managers, do not have the spare capital to invest in the infrastructure needed to maximize energy savings, and simply do not have the time it takes to implement changes to their businesses while they are struggling with the day-to-day tasks required to keep their businesses afloat.

96% of the industrial units belong to small companies in the Indian economy. The small companies account for 40% of the nation's overall industrial production and 42% of all Indian exports. Small companies also offer various opportunities in the rural and urban areas of the country. The Indian economy generally experiences unemployment and small firms have helped increase the employment chances for the people.

Table 1. India's Urban and Rural Area Employment Distribution

MSME	Urban (in lakh)	Rural (in lakh)
Micro Businesses	306.43	324.09
Small Businesses	2.53	0.78
Medium Businesses	0.04	0.01

Cost-effective energy efficiency measures have the potential to reduce SMEs carbon emission by as much as Total Middle East annual emission combined.

The objective of this project is to build an AI-driven tool that employs machine learning to help reduce energy costs, by decreasing waste, optimizing production, and anticipating malfunctions, the last of which helps workers keep up with equipment maintenance before damage occurs.

2.0 Customer Needs Assessment

Table 2. Initial Customer Needs List Obtained from Interviews and Observations

1. Cost
2. Downtime/Disruption/Continuity
3. Lack of specialist contracting/consultancy resources
4. Workforce resistant to new technology
5. Lack of digital skills in workforce
6. Unsure how to improve efficiency/not understanding the technology
7. Not enough data to support business case

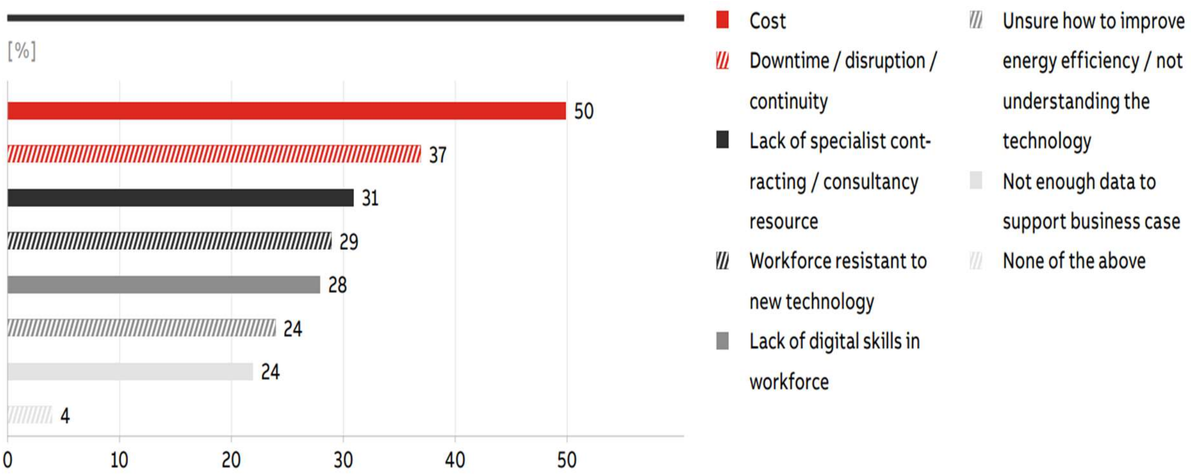
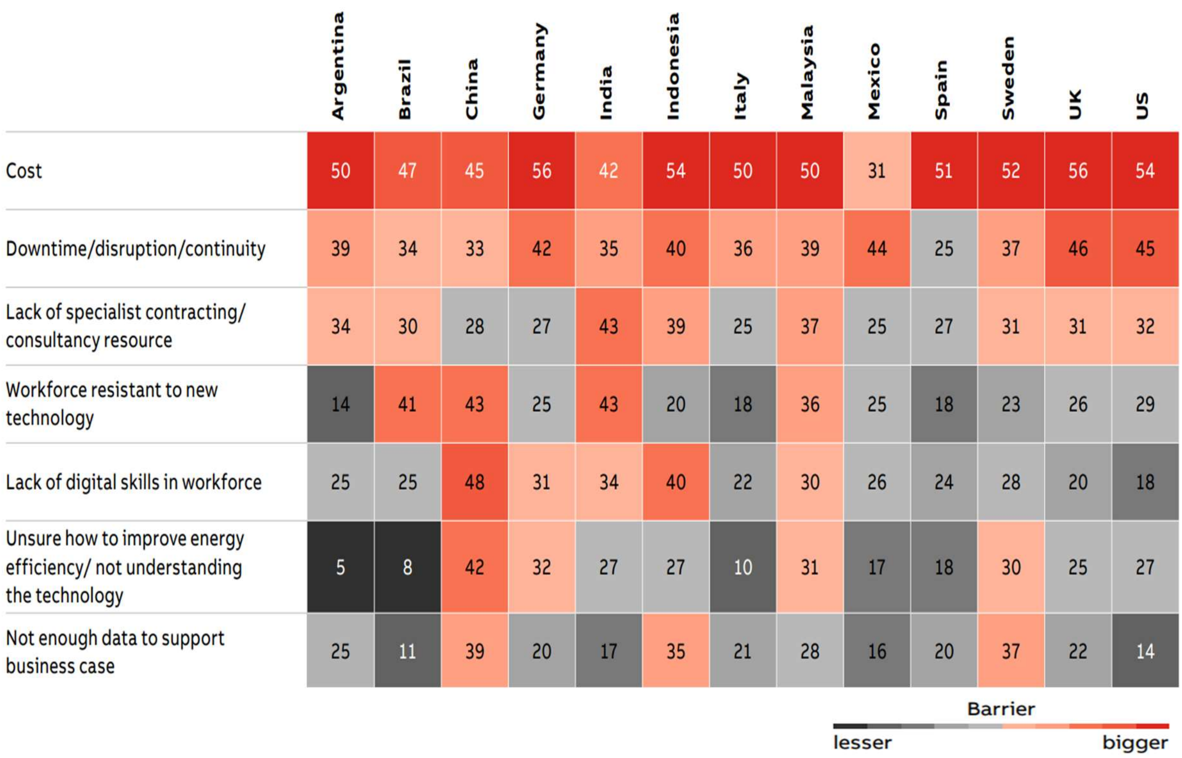


Figure 4. Graph representing Customer Needs



	Argentina	Brazil	China	Germany	India	Indonesia	Italy	Malaysia	Mexico	Spain	Sweden	UK	US
Cost	50	47	45	56	42	54	50	50	31	51	52	56	54
Downtime/disruption/continuity	39	34	33	42	35	40	36	39	44	25	37	46	45
Lack of specialist contracting/ consultancy resource	34	30	28	27	43	39	25	37	25	27	31	31	32
Workforce resistant to new technology	14	41	43	25	43	20	18	36	25	18	23	26	29
Lack of digital skills in workforce	25	25	48	31	34	40	22	30	26	24	28	20	18
Unsure how to improve energy efficiency/ not understanding the technology	5	8	42	32	27	27	10	31	17	18	30	25	27
Not enough data to support business case	25	11	39	20	17	35	21	28	16	20	37	22	14

Figure 5. Country Overview of Customer Needs

Table 3. Hierarchal Customer Needs List (With Weighting factors)

1. Cost
2. Downtime/Disruption/Continuity
3. Lack of specialist contracting/consultancy resources
4. Workforce resistant to new technology

4.1 Lack of digital skills in workforce

4.2 Unsure how to improve efficiency/not understanding the technology

5. Not enough data to support business case

2.1 Weighting of Customer Needs

Analytical Hierarchy Process is a decision-making method developed by Thomas L. Saaty that performs systematic structuring and analysing of complex decision problems.

Weighting criteria in AHP is a crucial step that plays a significant role in determining the relative importance of different criteria in decision-making by determining relative importance, ensuring consistency, analysing trade-offs, and facilitating objective and transparent decision-making. It enhances the robustness and reliability of the decision-making process, leading to more informed and effective decisions. **See Appendix A1**

Cat		Priority	Rank	(+)	(-)
1	Cost	49.1%	1	10.8%	10.8%
2	Downtime/Disruption/Continuity	22.9%	2	4.1%	4.1%
3	Lack of contracting/consultancy resources	12.6%	3	3.9%	3.9%
4	Workforce resistant to new technology	8.9%	4	2.3%	2.3%
5	Not enough data to support business case	6.5%	5	1.8%	1.8%

Number of comparisons = 10
Consistency Ratio CR = 2.9%

Figure 6. Weights for the Criteria based on AHP Pairwise Comparisons

	1	2	3	4	5
1	1	3.00	4.00	5.00	6.00
2	0.33	1	2.00	3.00	4.00
3	0.25	0.50	1	1.00	3.00
4	0.20	0.33	1.00	1	1.00
5	0.17	0.25	0.33	1.00	1

Principal eigen value = 5.129
Eigenvector solution: 4 iterations, delta = 2.5E-9

Figure 7. Weights based on Principal Eigen vector of Decision Matrix

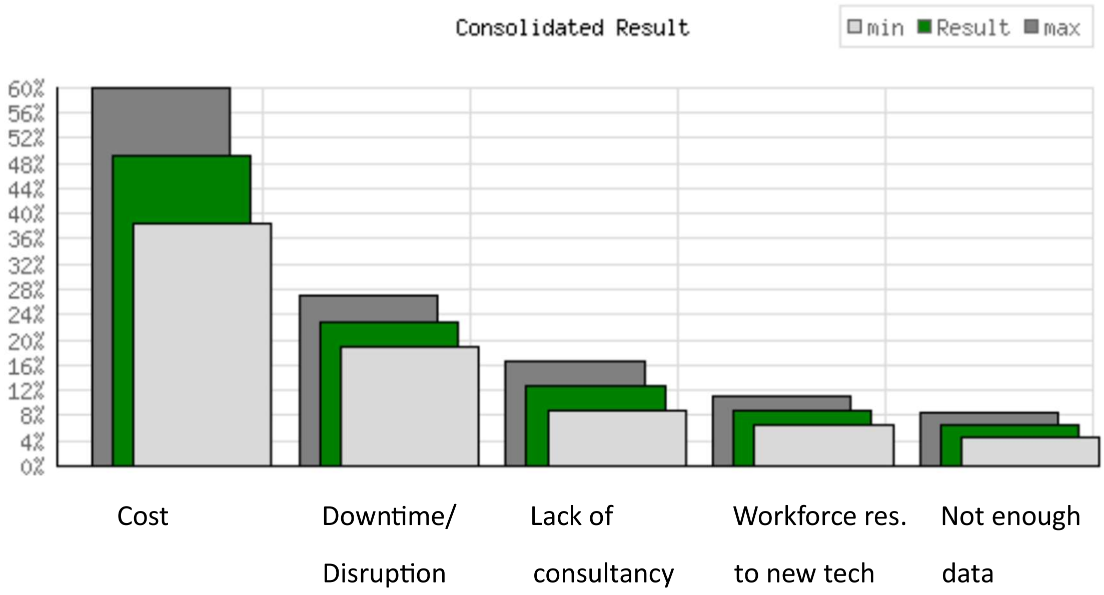


Figure 8. Consolidated Result of Priorities

3.0 Revised Needs Statement and Target Specifications

1. **Cost**
   The two most significant barriers to improving energy efficiency were cost (50 percent) and downtime (37 percent). Cost was by far the biggest barrier for those who have not and do not intend to invest in energy efficiency. However, an important upside to investing is that saving energy means saving money. Depending on the application, investments can have a quick payback time of 6 to 24 months while delivering long-term benefits for the environment. Without such investments, companies become vulnerable to fluctuations in energy prices. Rising energy costs can significantly impact operating expenses, profitability, opportunities to reduce improve operational efficiency and enhance overall financial performance.

2. **Downtime/Disruption/Continuity**
   Downtime and disruptions can have significant implications for energy efficiency in industries. The impact largely depends on the nature of the disruptions and the specific processes involved. When production processes are interrupted and then restarted, there is often a surge in energy consumption during the startup phase. This spike in energy usage can be inefficient and costly. Over time, frequent disruptions and downtime can contribute to wear and tear on machinery and equipment. Frequent disruptions can lead to production interruptions, and the need for backup power sources may result in less energy-efficient operations.

3. **Lack of contracting/consultancy resources**
   Many industries just lack the consultancy resources to adopt new changes. Energy consultants are well-versed in industry best practices for energy efficiency, local and international energy regulations. Without their input, industries may fail to implement proven strategies, technologies, and management practices that could enhance overall

efficiency. Companies lacking consultancy may choose projects with longer payback periods or lower returns on investment. This can make it more challenging to justify energy efficiency investments to management and stakeholders.

4. **Workforce resistance to new technology**
   More than two in five respondents mentioned workforce resistance to new technology as a significant barrier to improving energy efficiency. The introduction of new technology in the workforce has historically posed a threat to workers' sense of security, stability, and purpose. It will also enable them to adopt strategies that address people's concerns and reservations. Another big concern for industrial companies is lack of digital skills, with many naming it a substantial barrier to decision-making, improving energy efficiency. Expanded public-private partnerships may help address the gap between workforce skills and employers' needs.

5. **Not enough data to support business case**
   Overall, twenty four percent of the respondents said they have struggled to gain meaningful information on energy efficiency from the government and third parties. This rises to over a quarter for those who have not invested and don't intend to, suggesting insufficient knowledge could be a barrier to investment. These concerns need to be addressed to make significant gains. By adopting energy efficient technology, industries can enjoy a fast return on investment while cutting CO_2 emissions. The bottom line is that energy efficiency is good for business, good for reputation and good for the environment.

4.0 External Search

ABB energy efficiency survey 2022

Sapio Research was commissioned by ABB to carry out a comprehensive global survey of 2,294 companies in 13 countries in the manufacturing, transportation, heavy industry, light industry, and energy industries. Respondent companies ranged from small enterprises with under 100 employees to those with a $5 billion turnover and over 5,000 workers. The survey, which aimed to understand these industries' current and future plans to invest in energy efficiency, took place in February 2022. The online survey was conducted via email invitation, with most of the responses coming from manufacturing (51 percent), transportation (20 percent), and heavy industry (11 percent). The majority of respondents (43 percent) were managers, 31 percent were specialists, and 26 percent were executives or business owners.

LEAP4SME Energy and Economic Mapping in Europe-D2.1-SME-ener

LEAP4SME aims to support Member States in establishing or improving effective policies for small and medium-sized enterprises (SMEs) to undergo energy audits and implement cost-effective, recommended energy-saving measures through identifying the barriers for unlocking energy efficiency measures, mobilising private stakeholders, and proposing effective solutions to realise both energy and non-energy benefits.

bp Statistical Review of World Energy 2022 | 71st edition

The Statistical Review of World Energy has a new custodian: the Energy Institute (EI), the chartered professional membership body for people who work in energy. The EI published the 72nd edition of the Statistical Review of World Energy on 26 June 2023, together with a webcast. The Statistical Review has a new home and a new look, but it offers the same timely, comprehensive, objective and free-to-access global energy data.

IBM Business Guide to Modern Analytics

This guide helped to perform the following actions:

- Navigate the modern predictive analytics landscape.
- Identify opportunities to grow and enhance your use of AI.
- Empower both data science teams and business stakeholders to deliver value, fast.

4.1 Applicable Regulations

1. Digital Personal Data Protection Act, 2023

2. Trade Secrecy
    - *Section 27 of the Contract Act* - Law that bound the parties not to disclose information contrary to the terms of the contract between the parties i.e., Non-Disclosure Agreements.
    - *Section 405-409 of the Indian Penal Code, 1860* - Deals with the cases when there is a Criminal Breach of trust.

3. The Environment (Protection) Act, 1986, amended 1991

4. Forest (Conservation) Act, 1980, amended 1988

5. Wildlife (Protection) Act, 1972

6. The Water (Prevention and Control of Pollution) Act, 1974, amended 1988

7. The Air (Prevention and Control of Pollution) Act, 1981, amended 1987

8. The E-Waste Management Rules of 2016 and Plastic Waste Management Amendment Rules, 2021

    The Ministry of Environment, Forest & Climate Change (MoEFCC) is the federal agency responsible for the implementation and oversight of environmental laws in India. The Central Pollution Control Board (CPCB) implements policies framed by the MoEFCC and provides technical services to the Ministry.

4.2 Applicable Constraints

Artificial intelligence (AI) has shown great potential in fault detection and prediction, but there are also several challenges and issues associated with its use. Some of these challenges and issues include:

a. **Data quality**
AI algorithms rely heavily on data to learn and make predictions. If the data used is of poor quality, the accuracy of the predictions will be affected. Therefore, it is important to ensure that the data used is accurate, complete, and representative of the system being monitored.

b. **Data quantity**
In order for AI algorithms to be effective, they require large amounts of data. This can be a challenge in situations where data is scarce or difficult to obtain.

c. **Model complexity**
AI models can be very complex, which can make it difficult to understand how they are making predictions. This can be a challenge when trying to diagnose faults or understand why a particular prediction was made.

d. **Model training**
AI models require training in order to learn from data. This can be a time-consuming and resource-intensive process, particularly if the data is complex or the model is large.

e. **Model validation**
Once an AI model has been trained, it is important to validate its performance on new data. This can be challenging, as it requires a large amount of data that is representative of the system being monitored.

f. **Interpretability**
AI models can be difficult to interpret, which can make it challenging to understand why a particular prediction was made. This can be a problem when trying to diagnose faults or understand how the system is behaving.

g. **False positives and False negatives**
AI models can produce false positives (predicting a fault when there is none) or false negatives (failing to predict a fault when there is one). This can be a problem if it leads to unnecessary maintenance or missed faults.

h. **Human Trust**
Finally, there is the issue of human trust. AI models can be seen as a "black box" by humans, which can make it difficult for them to trust the predictions being made. This can be a challenge when trying to convince operators or maintenance personnel to take action based on the predictions.

4.3 Business Opportunity

Predictive Maintenance Service can be thought of as a cost centre within a business. As new technologies are changing the game companies can turn predictive maintenance service into a revenue stream.

1. **Data as a raw material**
Every manufacturer has the potential to collect data from their customers and there are a number of options for monetizing anonymized data. There is a growing number of data markets that make it possible to monetize the data, such

as Datarade, Snowflake Data Exchange and AWS Data Exchange. Anonymizing the data and selling it to companies that are building out solutions, such as systems integrators, 3rd party software vendors, large customers that are developing their own internal solutions. Partner with other manufacturers that have complementary devices, pooling the data to create more value.

2. **Analytics**
   The next step in the monetisation chain is to use the data. Some of the most common uses are remote monitoring, condition monitoring, predictive maintenance for equipment. This can be done internally using a company's own development resources; bringing in a systems integrator to build out the analytics; or licensing pre-built solutions.

3. **Applications and solutions**
   Collecting data and understanding how to do analytics lays the foundation for building out applications and packaged solutions. Analytics could be device-specific, for example, monitoring a piece of equipment to calculate the remaining life, identifying anomalies, or getting a better idea of when to schedule the next service call. Applications are packaged solutions that can be sold on a repeatable basis to a wide range of customers. Building out applications that leverage the data is an effective way to:

a. Make customer relationships stickier.
b. Generate new sources of revenue.
c. Gain entry into accounts that are using competitive devices.
d. Monetize the investment in data collection.

4. **Managed service**
   Companies can combine the physical devices, connectivity, data collection, analytics, actions and maintenance as an end-to-end managed service. This business model works particularly well in industries that have consumables, where the rate of consumption can vary widely based on different variables. The managed service could deliver the entire system and reduce the overall operating costs for its customers, lock the customer in to using their brand, monitor their competition and grow market share.

   Pricing models will vary, but could be a subscription with a fixed monthly or annual payment, or variable pricing based on the actual output – a true outcome-as-a-service model. An example of this is Stihl hand tools. They have developed a power drill that measures the amount of torque that is needed. A customer could rent a drill for a weekend, and they would pay more if they were drilling through concrete, requiring extra torque, than if they are drilling through drywall.

5. **Platforms and ecosystems**
   Building a platform is a long-term strategy, a business model is developed such that it:
   a. Drives broad adoption of the platform either by becoming the standard or becoming part of someone else's ecosystem and have to comply with their standards.
   b. Generates revenues to:
      - "*train*" customers to see the added value they are getting from the technology.
      - fund the cost of sales and marketing to drive adoption.
      - fund continued development in new solutions.

c. Builds a critical mass of customers that makes the platform viable as it becomes attractive for third-parties (customers, systems integrators, other manufacturers to build additional apps and create a network effect).
d. Lays the foundation for accessing, aggregating and monetizing more data from more sources.
e. Creates a barrier to entry for competitors.

5.0 Concept Generation

TRIZ, The Theory of Inventive Problem Solving, is by definition a theory used for solving problems creatively and represents a new way of thinking about and approaching innovation based on the objective nature of technological analysis. TRIZ identifies and codifies these principles, using them to make the creative process more predictable. Part of solving problems creatively, then, is finding existing solutions that can be adapted to the specific problem.

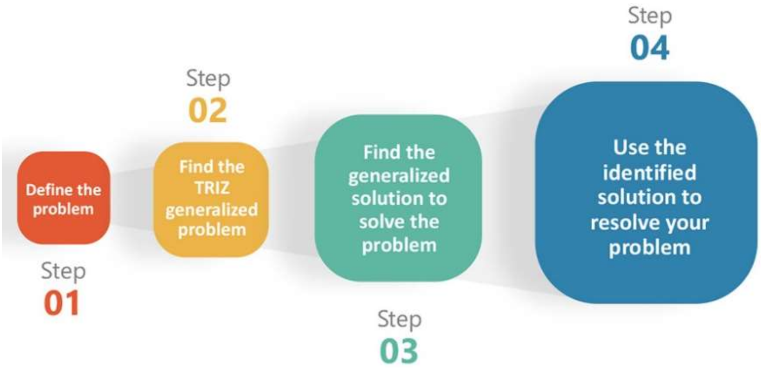


Figure 9. TRIZ for Problem Solving

1. **Define the Problem**
   The Problem that our project aims to solve is improving energy efficiency by predictive fault detection.

2. **TRIZ Generalized Problem**
   TRIZ has described the above problem as a contradiction that states "*To improve the Difficulty of detecting while preserving the Measurement accuracy.*"

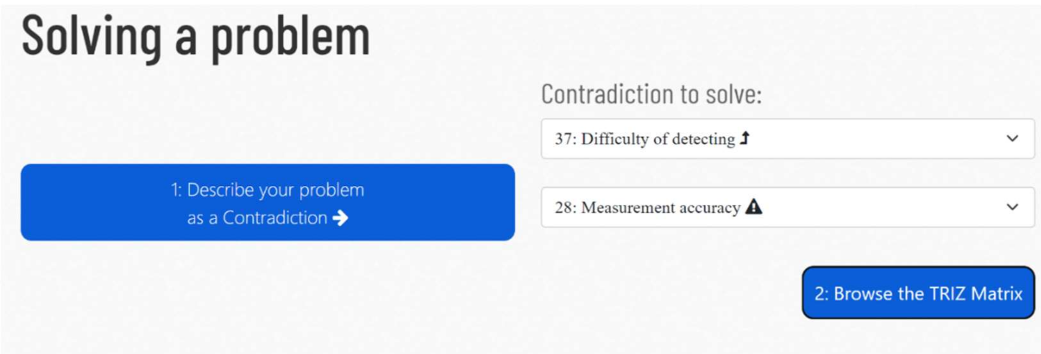


Figure 10. TRIZ Generalized Problem

3. **Generalized Solution**
   TRIZ proposes the following Principles to solve this contradiction:
   Copying
   - Instead of an unavailable, expensive, fragile object, use simpler and inexpensive copies.
   - Replace an object, or process with optical copies.
   - If visible optical copies are already used, move to infrared or ultraviolet copies.

   Intermediary
   - Use an intermediary carrier article or intermediary process.
   - Merge one object temporarily with another (which can be easily removed).

   Colour Changes
   - Change the colour of an object or its external environment.
   - Change the transparency of an object or its external environment.

   Mechanics Substitution
   - Replace a mechanical means with a sensory means.
   - Use electric, magnetic and electromagnetic fields to interact with the object.
   - Change from unstructured fields to those having structure.

4. **Applying Generalized Solution**
   The TRIZ process obtained generalised solution can be applied to our problem. This application is summarised in a brainstorming list and morphological chart.

   Brainstorming Lists
   1. **Copying**
      - Surveying from aerial images instead of on ground inspection.
      - Measuring an object by measuring the photograph.
      - Using Ultrasonic sound to detect flaws instead of direct testing.
      - Using Infrared images to detect heat sources.
   2. **Intermediary**
      - Employing many datasets to obtain accurate results.
   3. **Mechanics Substitution**
      - Antennas with very detailed structure of the pattern of radiation is preferred.

   Delighters
   Unique or unexpected features that could distinguish our product are:
   Mechanics Substitution
      - Instead of a mechanical or electrical sensor, a bad smelling compound like *t-butyl mercaptan* can be used in natural gas to alert users to leakage.

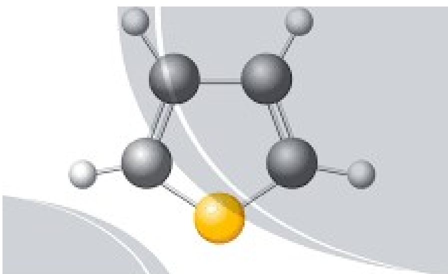


Figure 11. Compound to detect Leakage

Colour Changes

- Use photolithography to change transparent material to a solid mask.

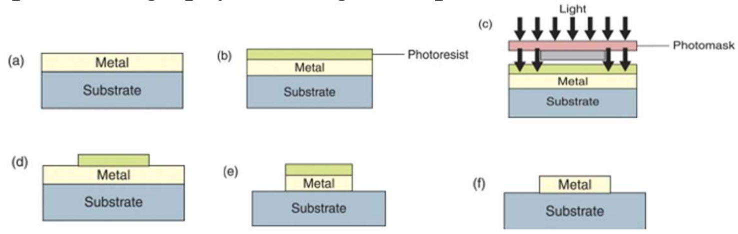
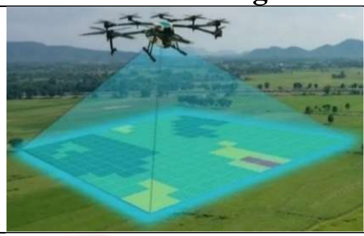
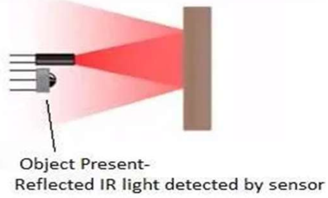
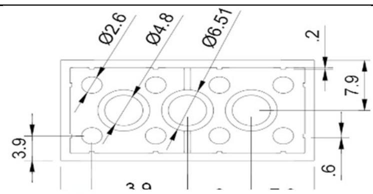
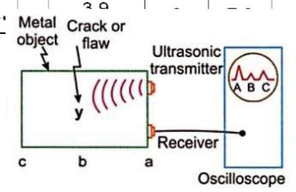
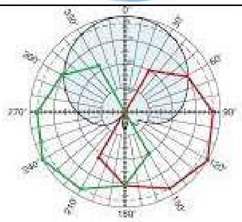


Figure 12. Photolithography

Table 4. Morphological Chart-I

Solution	**Application**		
	Contradiction	**Upgrade**	**Working**
Copying	Ground Inspection	Aerial Image Survey	
Copying	Ground Inspection	Infrared Images	
Copying	On-site Measurement	Image Dimensions	
Copying	Direct Testing	Ultrasonic Testing	
Inter-mediary	Few Datasets	Multiple datasets	
Mechanics Substitution	Open Broadcasting	Antenna Radiation	

6.0 Concept Development

Rather than how maintenance tasks are performed, the three models differ in terms of when they are performed.

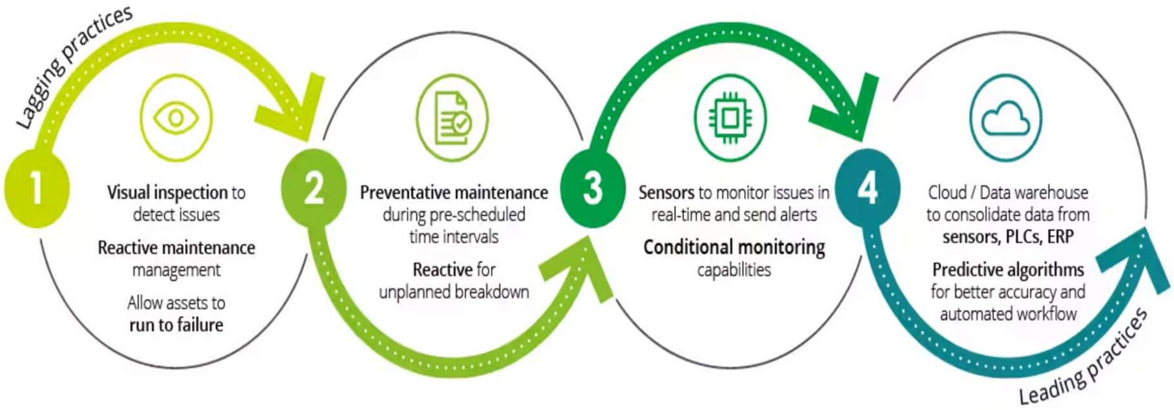


Figure 13. Types of Maintenance Models

Data analysis is what differentiates preventive maintenance from predictive maintenance. Indeed, three components make up predictive maintenance (PdM), which allow technicians to monitor the condition of equipment and alert them to upcoming failures:

1. Sensors installed in the machine generate real-time data on its health and performance.
2. The Internet of Things (IoT) connects machines, software, and cloud solutions, enabling the collection and analysis of huge amounts of data.
3. All the processed data are fed into predictive models, which can predict failures based on that information.

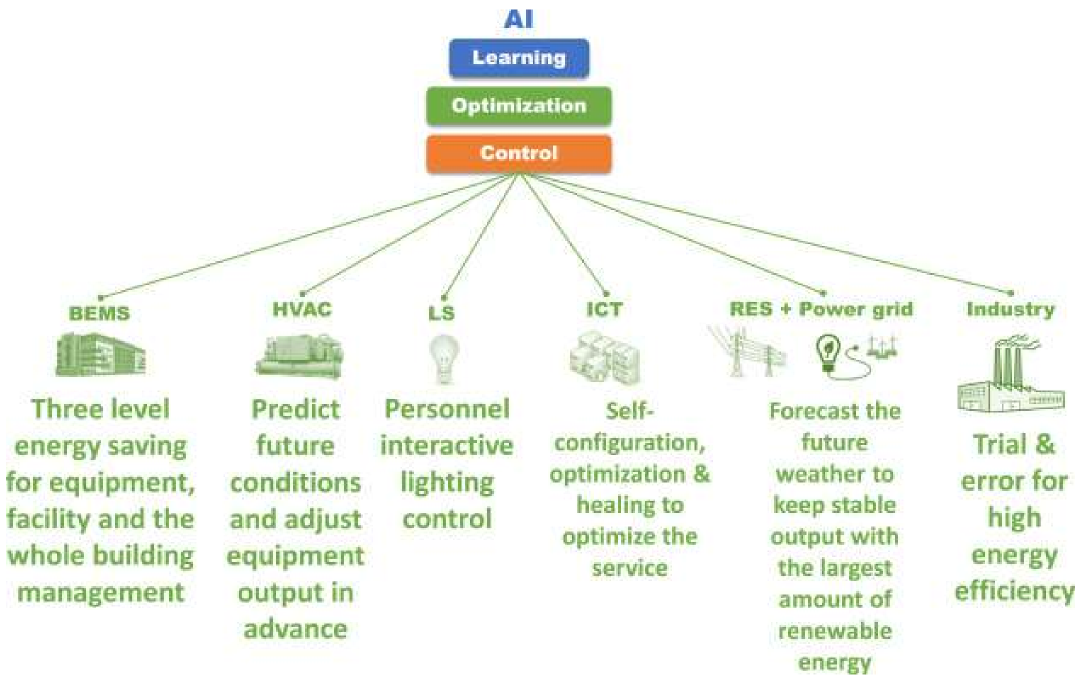


Figure 14. AI Workflow

7.0 Final Design

Our final product can perform Predictive Fault Detection and Equipment Maintenance to pursue efficient energy management. AI-based monitoring systems are coupled with drones to automate maintenance and system monitoring. First, the AI system can analyse weather and demand forecasts to generate drone flight schedules. Then, during its flight, the drone will capture the network's high-resolution images, which will be sent to a cloud-based AI application to identify the health of the network and classify it as "*functioning*", "*malfunctioning*" or "*sub-optimal functioning*". As an outcome, the AI application generates asset inspection reports that can be used for creating maintenance/repair work orders.

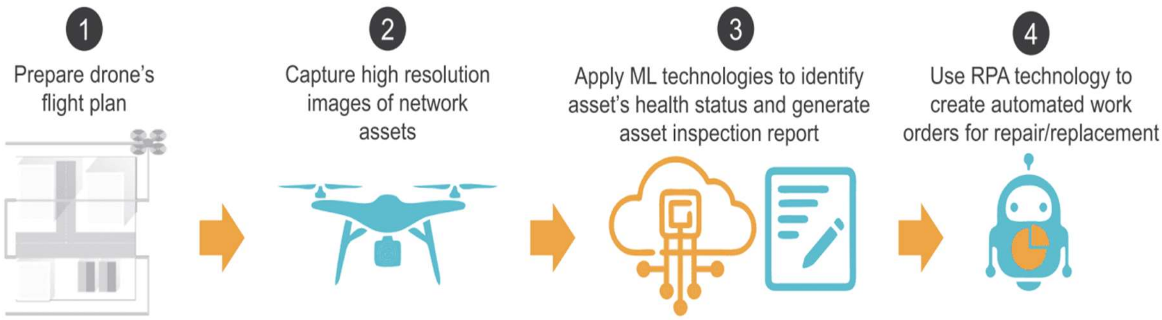


Figure 15. AI-drone based Predictive Equipment Maintenance

Long short-term memory (LSTM) networks, support vector machines (SVMs), and principal component analysis (PCA) are used to detect faults in power systems. Artificial neural network (ANN) based methods can illustrate the model's effectiveness when detecting the time and location of faults.

AI-powered Predictive Equipment Maintenance for equipment such as power grid/networks are economical as well as cost-effective. Such solutions help operators determine the overall system's health state to take proactive measures to prevent a catastrophe. Geothermal energy is an example of a rich amount of data used by predictive diagnostics to detect problems that can shut down power plants. Preventive measures such as chemical agent spray to avoid turbine shutdowns are optimized (quantity, composition, and timing) using the Internet of Things (IoT) and AI.

Figure 16. Drone deployed for Data Collection and Inspection

7.1 How does it work?

Figure 17. Predictive Maintenance Workflow

1. **Analysing the project's requirements**
   This step requires a thorough understanding of both the equipment and systems that are relevant to the project, as well as how they operate. It involves defining the physical quantities to be measured, selecting the sensors, and installing them if necessary. The next step involves defining a list of failure types that may occur.

2. **Understanding, collecting, and preparing data**
   Data collection
   Data can be collected and stored in databases using the sensors in the equipment. Drones equipped with advanced sensors and machine learning software can conduct efficient and thorough inspections, enabling utilities companies to detect and address potential risks promptly while ensuring the safety of personnel and the public.
   Understanding
   During this phase, the data to be analysed are determined, the quality of the data is identified, and its meaning is related to the data.
   Preparation
   In this sub-step, related data are selected; data are integrated by merging datasets; missing values are cleaned, outliers are checked and processed; updated data and enhanced features can be obtained by feature engineering. Data preparation takes a lot of time and work, typically between 70% and 90% of the project's total time.

3. **Data modelling**
   In data analysis, data modelling plays a pivotal role. The model receives input data (preparation), and output data are provided. Whether it is a classification problem, a regression problem, or a clustering problem, the first step is to select an algorithm to solve the problem. Different algorithms must be evaluated and parameterized in order to create a model.

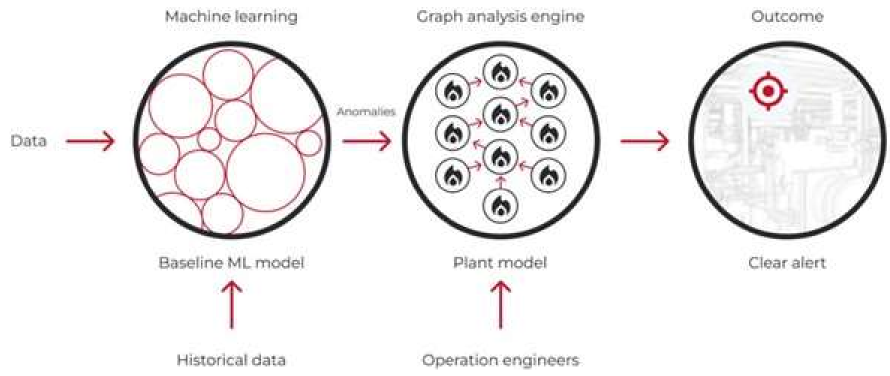


Figure 18. Data Modelling

4. **Evaluation and deployment**
The developed model should not be biased and should be both generalized and have good performance. During the deployment of the model, each step must be given proper attention, time, and energy.

5. **Decision-making**
Generally, operators use decision-making processes to decide how to resolve problems. Taking a step-by-step approach to making decisions can assist in making informed, thoughtful decisions that will make a positive difference in the short- and long-term.

7.2 Design, Manufacturing and Assembly

Design

a. **Central database** – Data are stored, processed and analysed either on-premise or on-cloud.
b. **Sensor** – The data-collecting sensors are installed in machine or physical product.
c. **Data communication** – A communication system that allows data to securely flow between the monitored asset and the central datastore.
d. **Predictive analytics** – Algorithms applied to the gathered data to identify patterns and generate insights in the form of alerts and dashboards.
e. **Root cause analysis** – It is a tool for data analysis that maintenance and process engineers use to study insights and plan corrective actions.

Table 5. Sensors Required for Predictive Maintenance

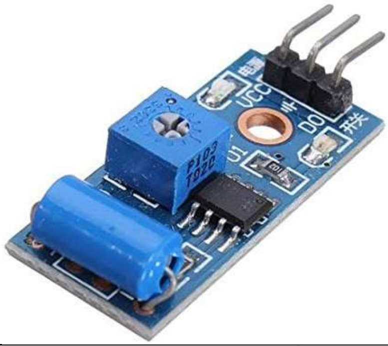
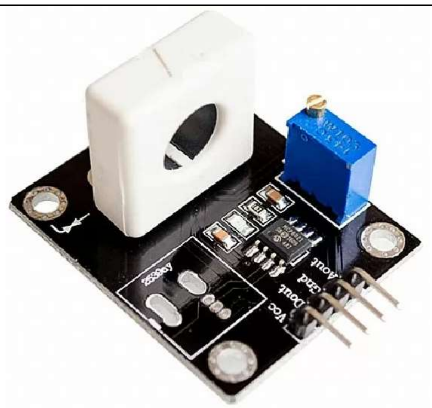
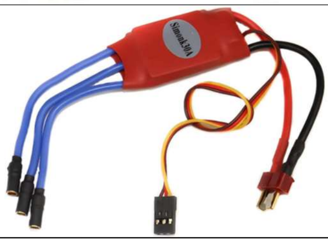
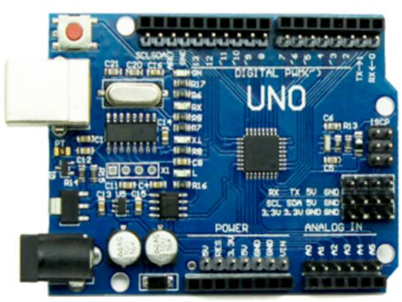
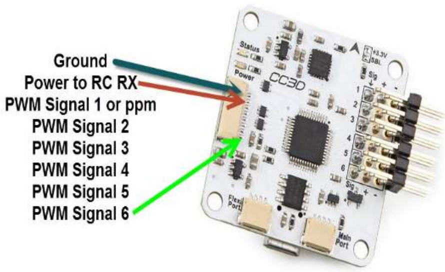
Component	Image	Function
Vibration sensor module SW-420		To detect: 1. Bearing wear and tear over or under lubrication 2. Shaft and coupling misalignment 3. Bent rotor 4. Broken rotor bar 5. Loose foundation
MLX90614 ESF Infrared Temperature Module		To detect: 1. Bearing heating 2. Coupling heating 3. Impeller heating 4. Fan heating 5. Rotor/Stator heating
US100 Ultrasonic Distance and Temperature Module		To detect: 1. Leakage detection 2. Pump Cavitation 3. Bearing friction and foreign particle detection 4. Structural faults, equipment housing cracks 5. Worn out, dirty leaking valves
WCS1700 Hall Current Sensor with over current protection		To detect: 1. Stator faults 2. Static/dynamic air gap irregularities 3. Broken rotor bar or cracked rotor end rings

Table 6. Components Required to build a Drone

Component	Image	Function
F450 Frame		Its sturdy frame has a robust construction and efficient design along with a diagonal wheel base of 450mm provide stability and durability for smooth flights.
LiPo Battery		The LiPo (Lithium Polymer) battery is a crucial component that powers the drone and determines its flight time.
Propellers		Propellers transform rotary motion into linear thrust. Drone propellers provide lift by spinning and creating an airflow.
BLDC Motors (Brushless DC)		BLDC Motors uses optical encoder to switch the fields at the correct time, thus allowing the motor to rotate in the same direction continuously.
Electronic Speed Controllers (ESC)		ESC allows the controllers to control and adjust the speed of drone's motor by raising or lowering the voltage to the motor as required.
Arduino UNO		It has 6 pins for PWM outputs, 6 pins for analog inputs, a 16 MHz ceramic resonator, a USB connection, a power jack, an ICSP header and a reset button.
CC3D Flight Controller		A flight controller has inbuilt gyro, accelerator controller. Arduino UNO sends PWM signals to microcontroller to control the individual BLDC motors.

Manufacturing

There is no requirement for manufacturing of these components as they are readily available in the market.

Assembling sensors and Arduino UNO

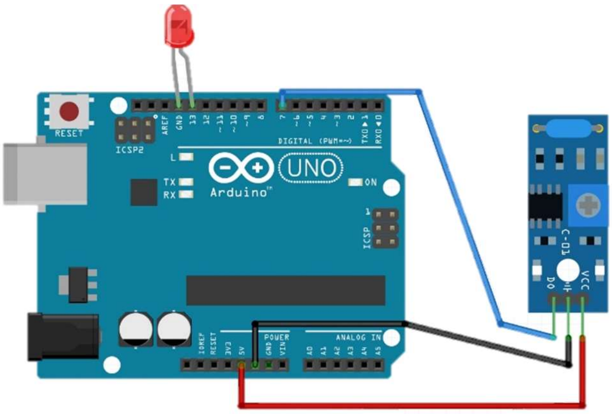


Figure 19. Interfacing Vibration Sensor module SW-420 with Arduino UNO

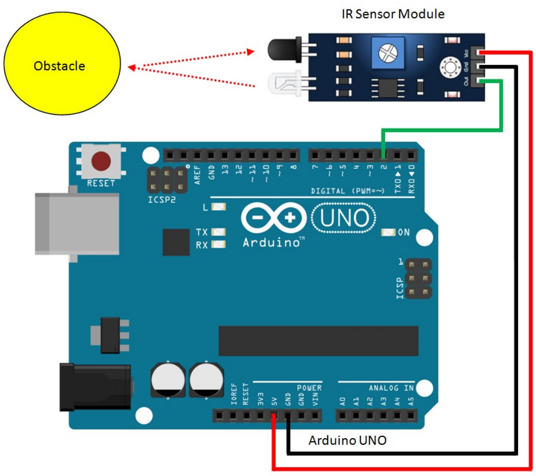


Figure 20. MLX90614 ESF Infrared Temperature Module with Arduino UNO

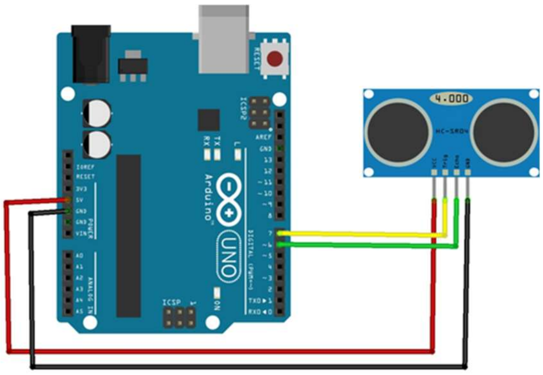


Figure 21. Interfacing US100 Ultrasonic Sensor Module with Arduino UNO

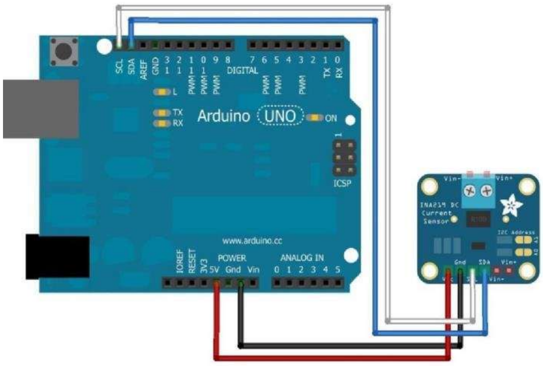


Figure 22. Interfacing Current Sensor with Arduino UNO

Assembling the components of Drone

1. **Controller Setup on Frame**

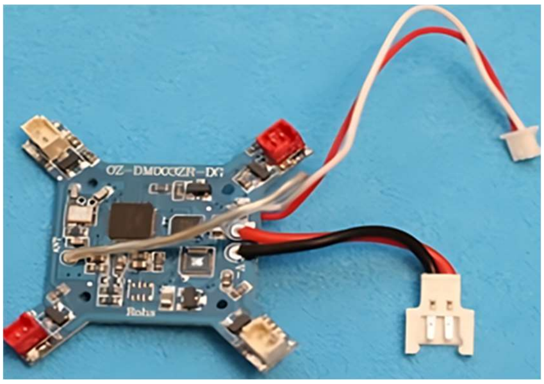


Figure 23. Controller Setup on Frame

Flight Controller is placed on the frame in proper direction for correct fitting. The Power cable is aligned with the cut on the frame and ensured that the holes match up with the frame's slots.

2. Connecting the main Microcontrollers

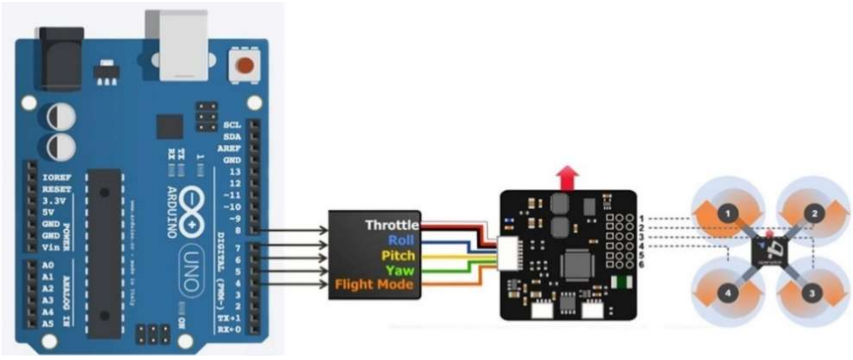


Figure 24. Arduino UNO interfaced with Flight Controller

The Arduino UNO is connected to the flight controller as shown in the above figure. Arduino UNO provides PWM signals input to controller.

3. Motor Installation

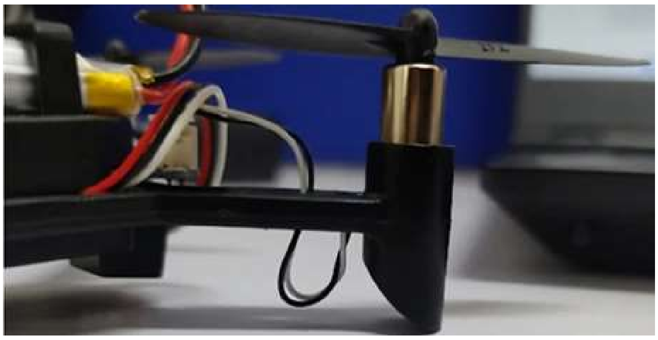


Figure 25. Motor Installation

The ESC motors are inserted into their corresponding arms based on the colours indicated on the circuit board. Then the wires are inserted into the corresponding sockets of the circuit board.

4. Battery Set-up

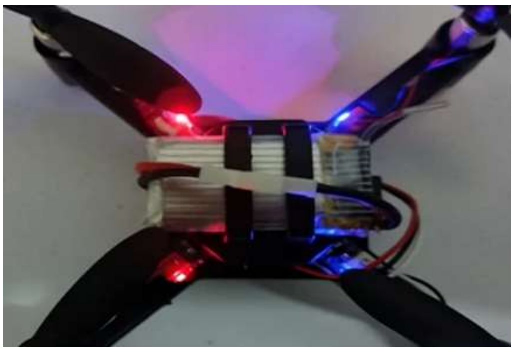


Figure 26. Battery Set-up

After connecting all the motors to the circuit board, it is important to attach the protective cover to the frame in proper direction. Once the battery is charged using the USB charger, it is inserted into the protective cover and connected to the circuit board.

5. Installing Propellers

Figure 27. Propellers

Propellers is connected to the motor. Different propellers are designed for clockwise and anticlockwise rotation. Propellers are connected with their matching motor wire colour. Its is made sure that propellers are pressed onto the motor firmly to secure them in place.

6. Overview of Transmitter

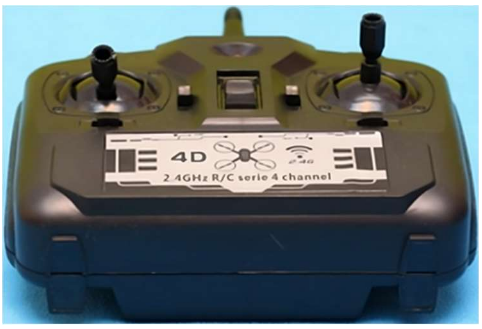


Figure 28. Transmitter

The transmitter has joysticks, buttons, and switches that correspond to various actions, such as taking off, landing and changing directions. There's a Power button to start the remote and an LED for indication.

Table 7. Bill of Materials

No.	Part	Qty.	Description	Wt. (gm)	Cost (INR)
1.	Jumper Wires	18	Female to Female, 100m	NA	30
2.	LED	2	3mm, Forward 1.6-2.03V, Wavelength 630-635 nm	NA	2.5
3.	Vibration sensor module	1	SW-420, 3.3-5V, 40x15x7 mm	3	38
4.	Infrared Temperature Module	1	MLX90614 ESF , 3-5V, 2 mA, -70 to 380 C, I2C, 11x17 mm	5	799
5.	US100 Ultrasonic Distance and Temperature Module	1	2.4-5V, 40000Hz, 2mA, 45x20 mm, 20 to 70 C, Sensing 15°, Diameter 16mm	9	185
6.	Hall Current Sensor (over current protected)	1	WCS1700, 5V, 5x5x4 mm, OP 0.4mA, Resolution 32mV/A	10	715
7.	Arduino UNO	1	ATmega328 SMD, 5V, Flash Memory 32Kb, SRAM 2Kb, EEPROM 1Kb, I/O Pin- 40mA, 3.3V Pin- 50mA, Clock Speed 16MHz, Supply 7-12V, Digital I/O Pins 14 (of which 6 provide PWM output)	25	285
8.	Transmitter	1	12V, 2 Channels, GFSK, 1024 dBm, 159x99x315 mm, DSC+Charging Port	365	1949
9.	Frame	1	450x55 mm, Motor mount bolt holes 16/19 mm	270	709
10.	Electronic Speed Controller	1	30A, 5V-20A, 2-3S LiPo, 1000/1400/2200kv motor, 45x24x11 mm	32	404
11.	Flight Controller with receiver board	1	ATmega644 PA, 4.8-6V, 6 Pin, 13x10x4.9 mm	20	4599
12.	LiPo Battery	1	11.1V, 2200mAh, 1016x34x23 mm, 1-3 C	100	1374
13.	Battery Charger	1	AC 110-240V, 3x800mA max	103	322
14.	Battery USB Charging Cable	1	20x30 mm, DC	50	223
15.	Coreless motors	4	2200Kv, 12 Magnet Poles, Stator length 13mm, Stator arms 12mm, Stator diameter 22mm, Rotor diameter 28mm, Shaft diameter 3.17mm	50	419
16.	Pair of leaf propellers	2	Length 10in, Pitch 4.5 in, Shaft diameter 7.7 mm	25	49

7.3 Design Validation

Dataset

This dataset is a daily time series of electricity demand, generation, and prices in Spain from 2014 to 2018. It is gathered from ESIOS a website managed by REE (Red Electrica Española) which is the Spanish TSO (Transmission System Operator). A TSO's main function is to operate the electrical system and to invest in new transmission (high voltage) infrastructure. As a system operator, REE forecasts electricity demand and offers and runs daily actions. As a result of daily actions, a PBF Plan Básico de funcionamiento) is yielded. This is a basic scheduling of energy production (upon it, several mechanisms are triggered to ensure supply).

Code

Importing Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import skew, kurtosis, shapiro
```

Reading the Dataset

```
path = "spain_energy_market.csv"
data = pd.read_csv(path, sep=",", parse_dates=["datetime"])
data = data[data["name"] == "Demanda programada PBF total"]
data["date"] = pd.to_datetime(data["datetime"], format="mixed")
data["date"] = data["date"].dt.date
data.set_index("date", inplace=True)
data = data[["value"]]
data = data.asfreq("D")
data = data.rename(columns={"value": "energy"})
data.info()
print(data[:5])
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 1825 entries, 2014-01-01 to 2018-12-30
Freq: D
Data columns (total 1 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   energy  1825 non-null   float64
dtypes: float64(1)
memory usage: 28.5 KB
```

```
              energy
date
2014-01-01  620107.7
2014-01-02  615641.4
2014-01-03  598826.9
2014-01-04  653093.8
2014-01-05  535445.0
```

```
data.plot(title="Energy Demand")
plt.ylabel("MWh")
plt.show()
```

Output

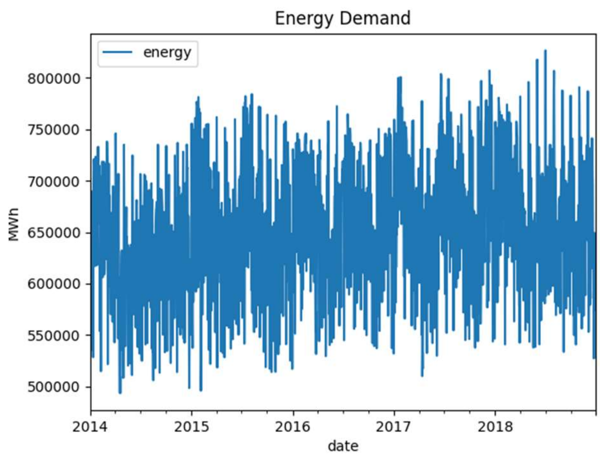


Figure 29. Energy Demand

```
print(len(pd.date_range(start="2014-01-01", end="2018-12-31")))
```

Output

```
1826
```

There are no missing values in the dataset, and we have a four-year span of data to work with.

```
data["year"] = data.index.year
data["qtr"] = data.index.quarter
data["mon"] = data.index.month
data["ix"] = range(0, len(data))
data[["movave_7", "movstd_7"]] = data.energy.rolling(7).agg([np.mean, np.std])
data[["movave_30", "movstd_30"]] = data.energy.rolling(30).agg([np.mean, np.std])
data[["movave_90", "movstd_90"]] = data.energy.rolling(90).agg([np.mean, np.std])
data[["movave_365", "movstd_365"]] = data.energy.rolling(365).agg([np.mean, np.std])
plt.figure(figsize=(20, 16))
data[["energy", "movave_7"]].plot(title="Daily Energy Demand in Spain (MWh)")
plt.ylabel("(MWh)")
plt.show()
```

Output

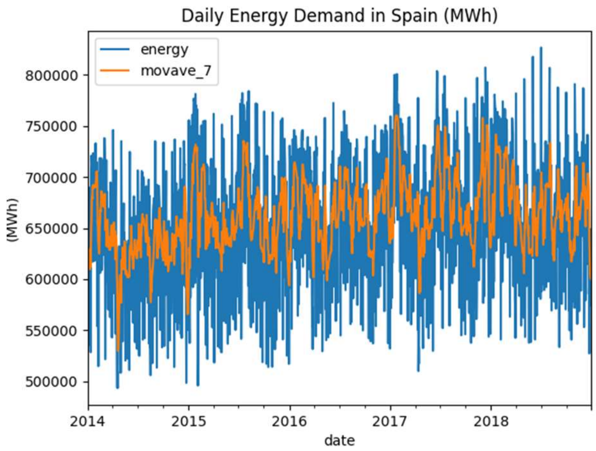


Figure 30. Daily Energy Demand in Spain

Exploratory Data Analysis (EDA)

Analysing the target variable involves studying its seasonality and trend. Our aim is to visually understand the patterns and fluctuations in the time series data without heavily relying on statistical techniques such as decomposition. By graphically examining the data, we can gain insights into the underlying patterns and trends that may exist.

Target Analysis (Normality)

```
mean = np.mean(data.energy.values)
std = np.std(data.energy.values)
skew = skew(data.energy.values)
ex_kurt = kurtosis(data.energy)
print("Skewness: {} \nKurtosis: {}".format(skew, ex_kurt+3))
```

Output

```
Skewness: -0.2555279252628293
Kurtosis: 2.605260697245368
```

In terms of data distribution, negative skewness indicates that the data is not perfectly symmetrical and has a longer left tail. Additionally, the kurtosis value below 3 suggests that the tails of the distribution are slightly thinner compared to a normal distribution. This characteristic is known as platykurtic, indicating that the likelihood of encountering extreme values is lower than in a normal distribution.

```
def shapiro_test(data, alpha=0.05):
    stat, pval = shapiro(data)
    print("H0: Data was drawn from a Normal Distribution")
    if pval < alpha:
        print("pval {} is lower than significance level: {}, therefore null hypothesis is 
rejected".format(pval, alpha))
    else:
        print("pval {} is higher than significance level: {}, therefore null hypothesis cannot be 
rejected".format(pval, alpha))
shapiro_test(data.energy, alpha=0.05)
```

Output

```
H0: Data was drawn from a Normal Distribution
pval 2.546819185197872e-10 is lower than significance level: 0.05, therefore null hypothesis is rejected
```

```
sns.distplot(data.energy)
plt.title("Target Analysis")
plt.xticks(rotation=45)
plt.xlabel("(MWh)")
plt.axvline(x=mean, color='r', linestyle='-', label="\mu: {0:.2f}%".format(mean))
plt.axvline(x=mean+2*std, color='orange', linestyle='-')
plt.axvline(x=mean-2*std, color='orange', linestyle='-')
plt.show()
```

Output

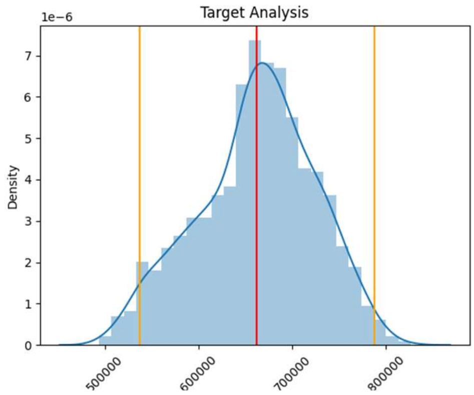


Figure 31. Target Analysis

In general, the data does not exhibit a normal distribution as it displays a smaller left tail and a reduced likelihood of observing extreme values compared to normally distributed data.

Inserting the rolling quantiles to the monthly returns

```
data_rolling = data.energy.rolling(window=90)
data['q10'] = data_rolling.quantile(0.1).to_frame("q10")
data['q50'] = data_rolling.quantile(0.5).to_frame("q50")
data['q90'] = data_rolling.quantile(0.9).to_frame("q90")
data[["q10", "q50", "q90"]].plot(title="Volatility Analysis: 90-rolling percentiles")
plt.ylabel("(MWh)")
plt.show()
```

Output

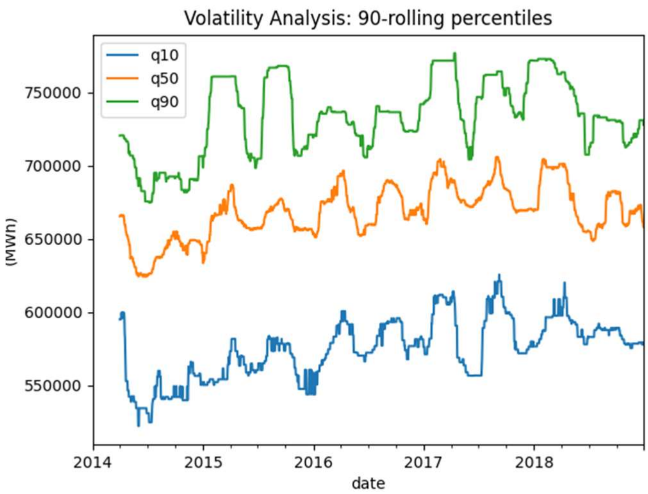


Figure 32. Volatility Analysis

```
data.groupby("qtr")["energy"].std().divide(data.groupby("qtr")["energy"].mean()).plot(kind="bar")
plt.title("Coefficient of Variation (CV) by qtr")
plt.show()
data.groupby("mon")["energy"].std().divide(data.groupby("mon")["energy"].mean()).plot(kind="bar")
plt.title("Coefficient of Variation (CV) by month")
plt.show()
```

Output

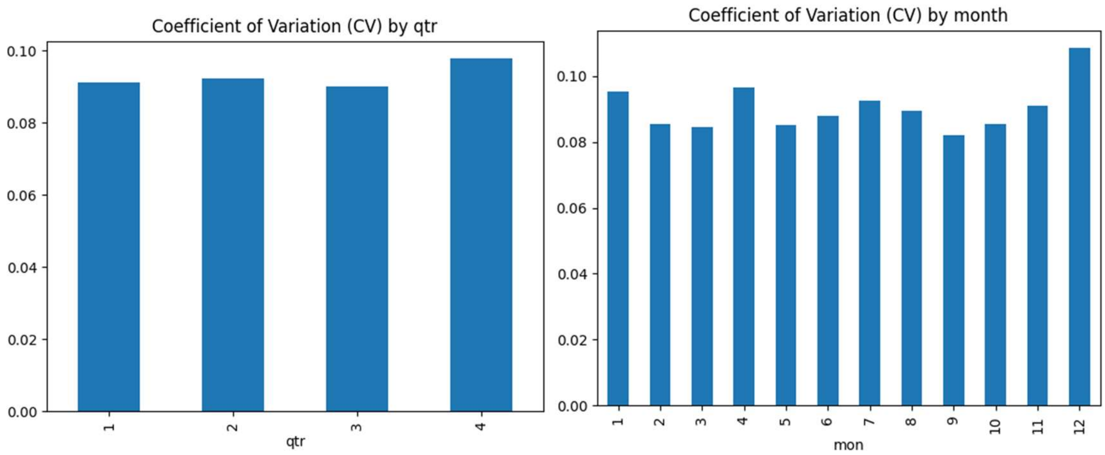


Figure 33. Coefficient of variation by qtr. and month

```
data[["movstd_30", "movstd_365"]].plot(title="Heteroscedasticity analysis")
plt.ylabel("(MWh)")
plt.show()
```

Output

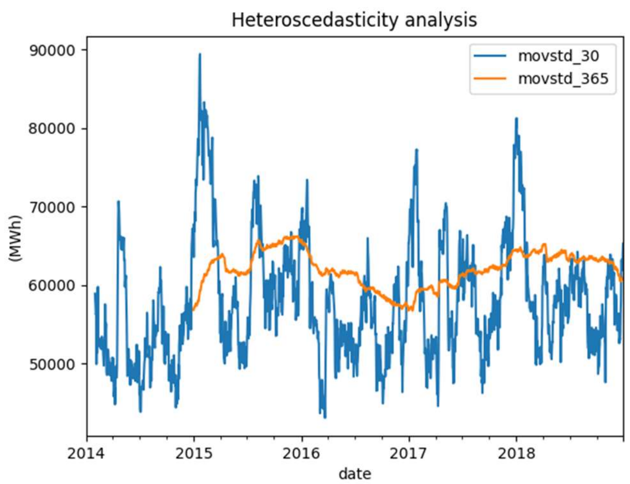


Figure 34. Heteroscedasticity Analysis

When considering shorter time periods such as quarters and months, volatility tends to vary, but over the long term (in a yearly window), it remains relatively stable. As a result, potential predictors need to account for the seasonal pattern in variance.

```
data[["movave_30", "movave_90"]].plot(title="Seasonal Analysis: Moving Averages")
plt.ylabel("(MWh)")
plt.show()
```

Output

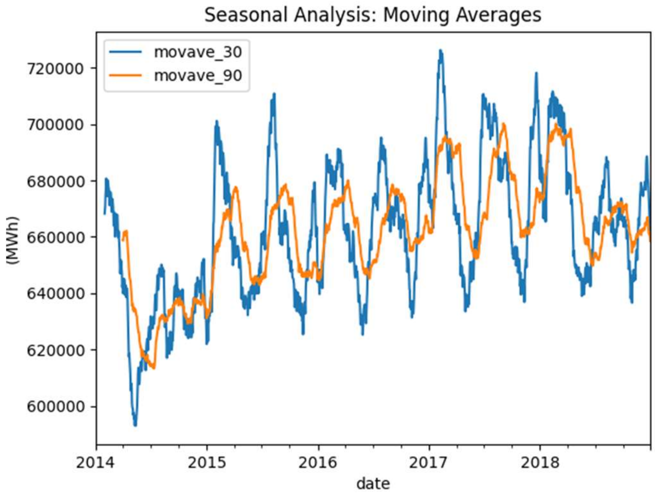


Figure 35. Seasonal Analysis: Moving Averages

```
sns.boxplot(data=data, x="qtr", y="energy")
plt.title("Seasonality analysis: Distribution over quarters")
plt.ylabel("(MWh)")
plt.show()
```

Output

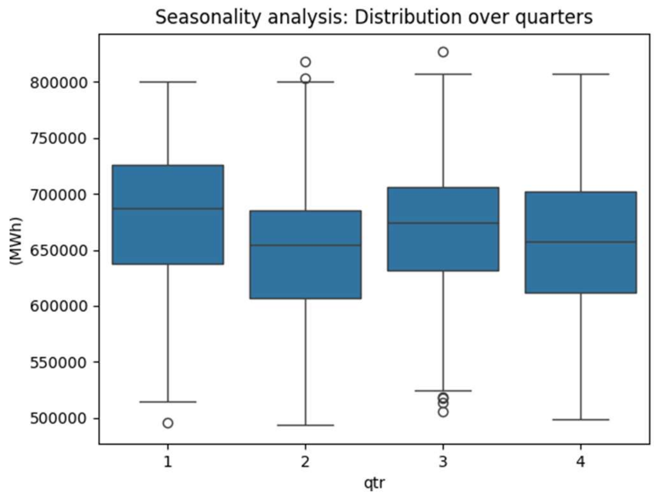


Figure 36. Seasonal Analysis: Quarterly distributed

As anticipated, there are distinct seasonal patterns observed in the data when considering quarters.

```
data_mon = data. energy.resample("M").agg(sum).to_frame("energy")
data_mon["ix"] = range(0, len(data_mon))
print(data_mon[:5])
```

Output

date	energy	ix
2014-01-31	20706385.1	0
2014-02-28	18626996.7	1
2014-03-31	19956417.7	2
2014-04-30	18024449.2	3
2014-05-31	18982544.9	4

```
sns.regplot(data=data_mon,x="ix", y="energy")
plt.title("Trend analysis: Regression")
plt.ylabel("(MWh)")
plt.xlabel("")
plt.show()
```

Output

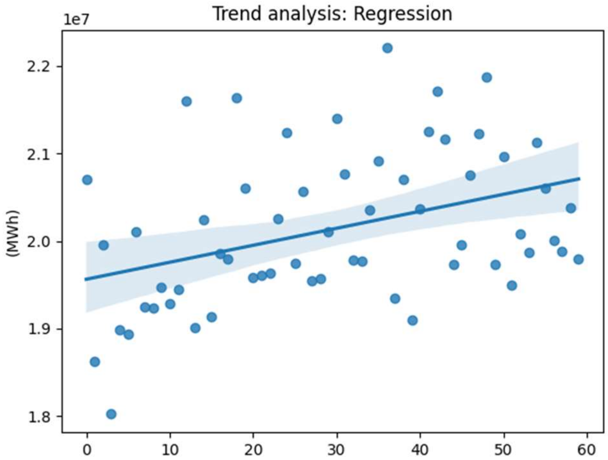


Figure 37. Trend Analysis: Regression

```
sns.boxplot(data=data["2014":"2017"], x="year", y="energy")
plt.title("Trend Analysis: Annual Box-plot Distribution")
plt.ylabel("(MWh)")
plt.show()
```

Output

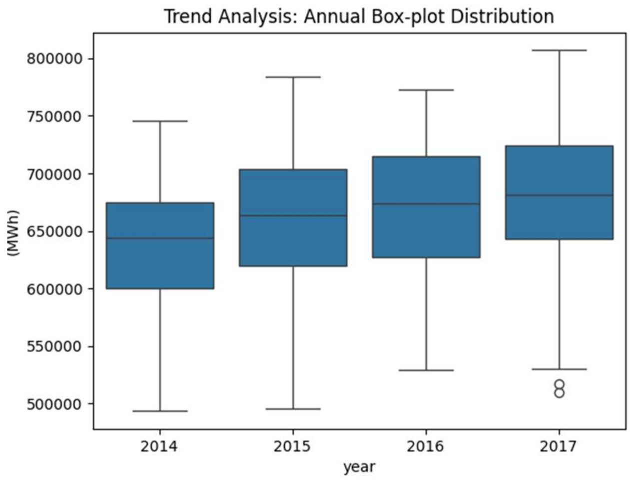


Figure 38. Trend Analysis: Annual Box-plot Distribution

The energy demand shows a positive linear trend, or a slightly damped trend, which can be attributed to the steady economic growth resulting from the recovery from a previous recession.

Feature Engineering

Standardizing the data is a necessary step to enable the application of models that are sensitive to scale, such as neural networks or support vector machines (SVM). By standardizing the data, we ensure that the distribution shape remains unchanged while only altering the first and second moments, namely the mean and standard deviation. This process allows for more accurate and effective modelling of the data using these particular machine learning algorithms.

```
data["target"] = data.energy.add(-mean).div(std)
sns.distplot(data["target"])
plt.show()

features = []
corr_features = []
targets = []
tau = 30
```

Output

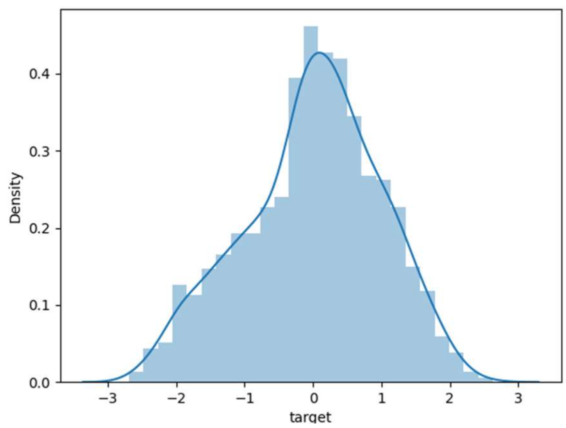


Figure 39. Target-Density

Forecasting periods

```
for t in range(1, tau + 1):
    data["target_t" + str(t)] = data.target.shift(-t)
    targets.append("target_t" + str(t))

for t in range(1, 31):
    data["feat_ar" + str(t)] = data.target.shift(t)
    # data["feat_ar" + str(t) + "_lag1y"] = data.target.shift(350)
    features.append("feat_ar" + str(t))
    # corr_features.append("feat_ar" + str(t))
    # features.append("feat_ar" + str(t) + "_lag1y")

for t in [7, 14, 30]:
    data[["feat_movave" + str(t), "feat_movstd" + str(t), "feat_movmin" + str(t),
          "feat_movmax" + str(t)]] = data.energy.rolling(t).agg([np.mean, np.std, np.max, np.min])
    features.append("feat_movave" + str(t))
    # corr_features.append("feat_movave" + str(t))
    features.append("feat_movstd" + str(t))
    features.append("feat_movmin" + str(t))
```

```
    features.append("feat_movmax" + str(t))

months = pd.get_dummies(data.mon,
                        prefix="mon",
                        drop_first=True)
months.index = data.index
data = pd.concat([data, months], axis=1)

features = features + months.columns.values.tolist()
corr_features = ["feat_ar1", "feat_ar2", "feat_ar3", "feat_ar4", "feat_ar5", "feat_ar6", "feat_ar7", 
"feat_movave7",
                "feat_movave14", "feat_movave30"]
```

Calculating correlation matrix

```
corr = data[["target_t1"] + corr_features].corr()
top5_mostCorrFeats = 
corr["target_t1"].apply(abs).sort_values(ascending=False).index.values[:6]
```

Plotting Heat Map of correlation matrix

```
sns.heatmap(corr, annot=True)
plt.title("Pearson Correlation with 1 period target")
plt.yticks(rotation=0)
plt.xticks(rotation=90)  # fix tick label directions
plt.tight_layout()  # fits plot area to the plot, "tightly"
plt.show()
```

Output

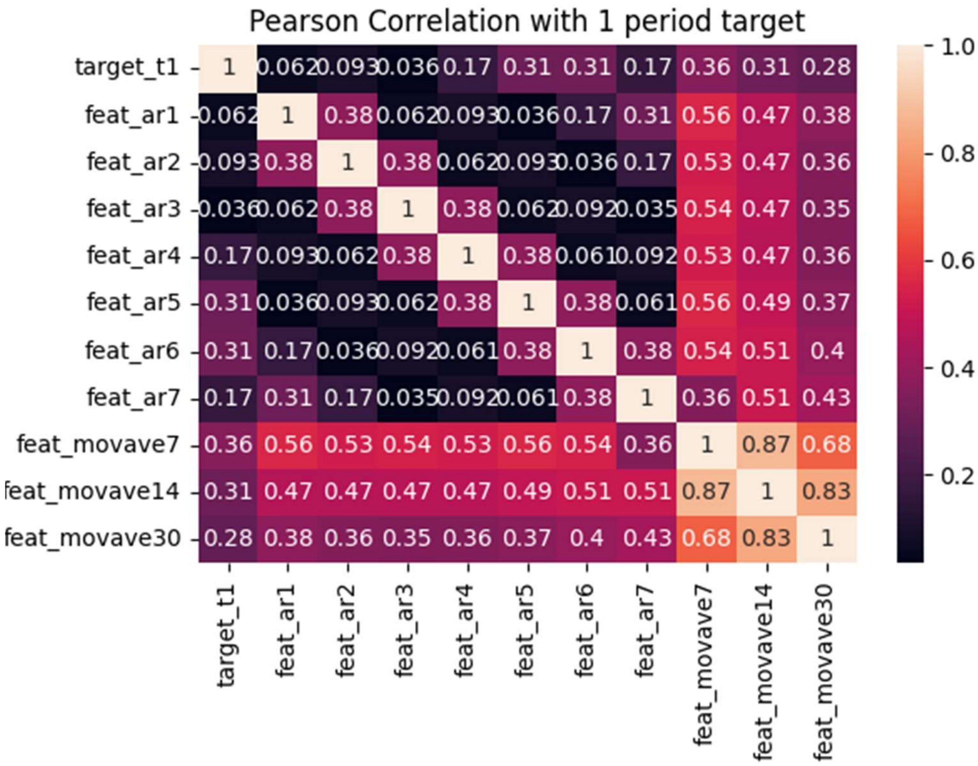


Figure 40. Heat Map of Correlation Matrix

```
sns.pairplot(data=data[top5_mostCorrFeats].dropna(), kind="reg")
plt.title("Most important features Matrix Scatter Plot")
plt.show()
sys.exit(1)
```

Output

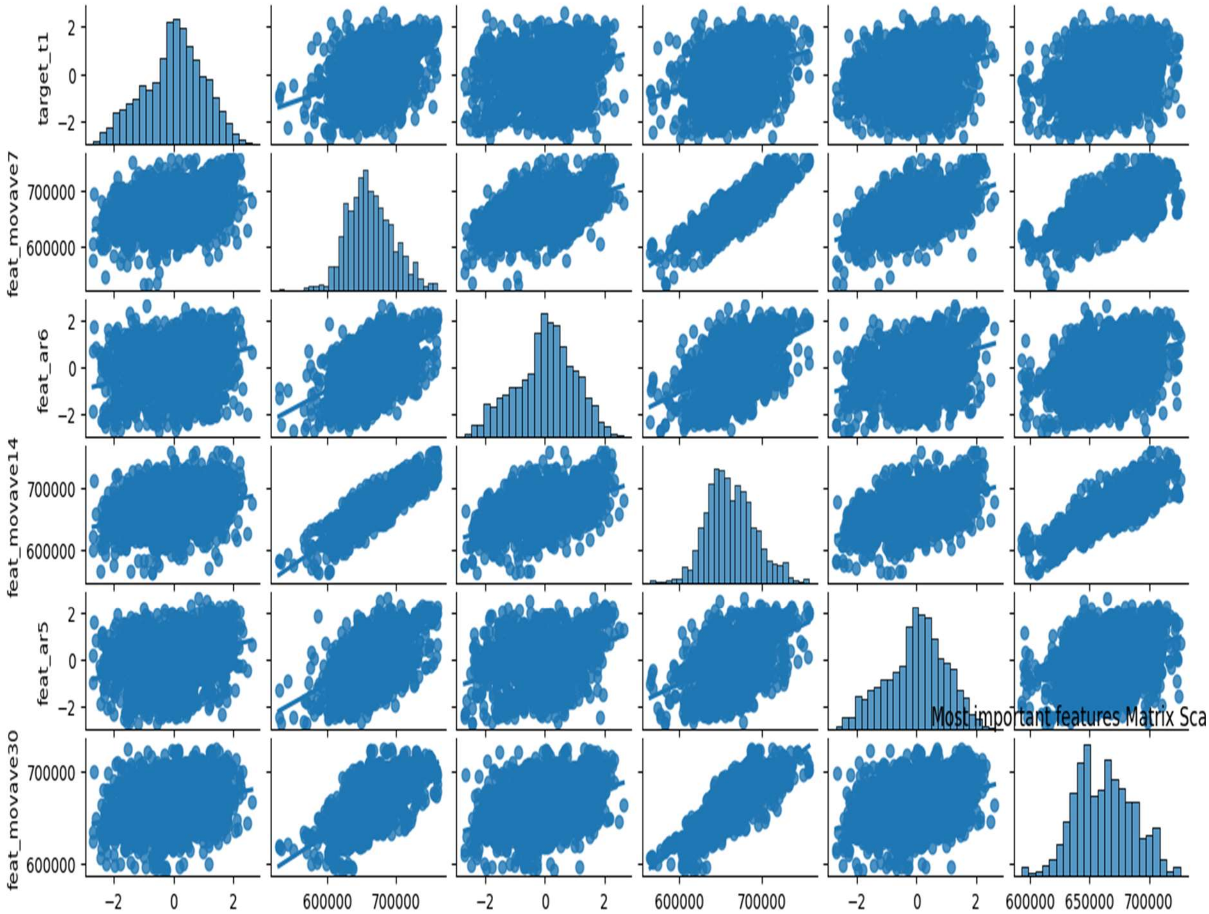


Figure 41. Matrix Scatter Plot

Model Building

In this step, we have built two candidate models using a convenient feature in Scikit-Learn called MultiOutput Regression. This feature allows us to efficiently and automatically fit models that can predict multiple target variables simultaneously. By leveraging this framework, we can train our models to predict several target variables in a streamlined manner. This not only simplifies the modelling process but also enables us to evaluate the models' performance across multiple targets effectively.

First, we will fit a baseline model using linear regression and compare it to a more advanced model, such as Random Forest. The linear regression model does not require extensive hyperparameter tuning and provides a solid foundation for our analysis. However, there are several considerations to keep in mind:

1. **Non-Normal Distribution and Varied Variance:** The target variable does not follow a perfect normal distribution and exhibits varying levels of variance. This can affect the assumptions of linear regression, which assumes normality and constant variance. We need to be cautious of potential deviations from these assumptions.

2. **Multicollinearity Among Predictors:** There is a high degree of multicollinearity among the predictor variables, meaning that some predictors are highly correlated with

each other. This can introduce challenges in interpreting the individual effects of these predictors on the target variable and may impact the model's performance.

3. **Non-Independence of Observations:** The observations in our dataset may not be independent, which violates one of the key assumptions of linear regression. Non-independence can arise from various factors, such as temporal dependencies or clustering within the data. We need to consider this when interpreting the model results and evaluating its accuracy.

   Scikit-Learn provides TimeSeries Split to achieve optimal performance. This technique allows us to perform GridSearch in a time-aware manner by preserving the temporal order of the data. It splits the data into sequential time-based folds, ensuring that each fold respects the chronological order of the observations.

   By using TimeSeries Split, we can iteratively train and evaluate our Random Forest model with different combinations of hyperparameters. This approach enables us to find the best set of hyperparameters that maximizes the model's performance on unseen future data points.

```
data_feateng = data[features + targets].dropna()
nobs= len(data_feateng)
print("Number of observations: ", nobs)
```

Output

```
Number of observations:  1765
```

Splitting the Data

To ensure an unbiased evaluation of our model's performance and conduct thorough residual analysis, we reserve the data points from the year 2018 as a separate holdout dataset. This means that we keep this data untouched during the model development process.

```
X_train = data_feateng.loc["2014":"2017"][features]
y_train = data_feateng.loc["2014":"2017"][targets]

X_test = data_feateng.loc["2018"][features]
y_test = data_feateng.loc["2018"][targets]

n, k = X_train.shape
print("Total number of observations: ", nobs)
print("Train: {}{}, \nTest: {}{}".format(X_train.shape, y_train.shape,X_test.shape, y_test.shape))

plt.plot(y_train.index, y_train.target_t1.values, label="train")
plt.plot(y_test.index, y_test.target_t1.values, label="test")
plt.title("Train/Test split")
plt.xticks(rotation=45)
plt.show()
```

Output

```
Total number of observations:  1765
Train: (1431, 53)(1431, 30),
Test: (334, 53)(334, 30)
```

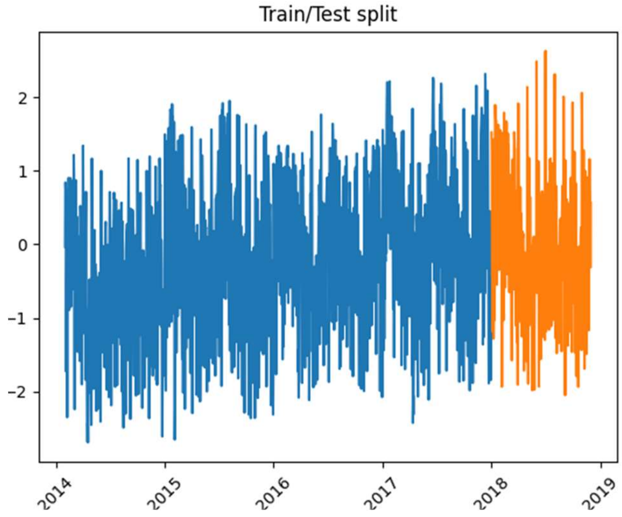


Figure 42. Train Split

Baseline Model: Linear Regression

```
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error

reg = LinearRegression().fit(X_train, y_train["target_t1"])
p_train = reg.predict(X_train)
p_test = reg.predict(X_test)

RMSE_train = np.sqrt(mean_squared_error(y_train["target_t1"], p_train))
RMSE_test = np.sqrt(mean_squared_error(y_test["target_t1"], p_test))

print("Train RMSE: {}\nTest RMSE: {}".format(RMSE_train, RMSE_test))
```

Output

```
Train RMSE: 0.7829033680347028
Test RMSE: 0.7954296844838591
```

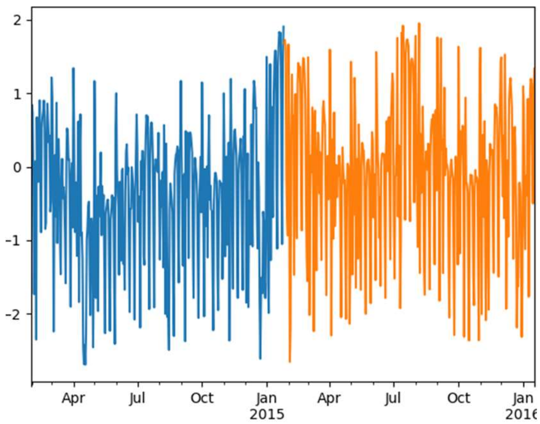
Training a Random Forest with Time Series Split to tune Hyperparameters

In this particular example, we illustrate the use of the TimeSeriesSplit framework. With this approach, each fold of the data is constructed in such a way that the training data is closer to the beginning of the forecasting period.

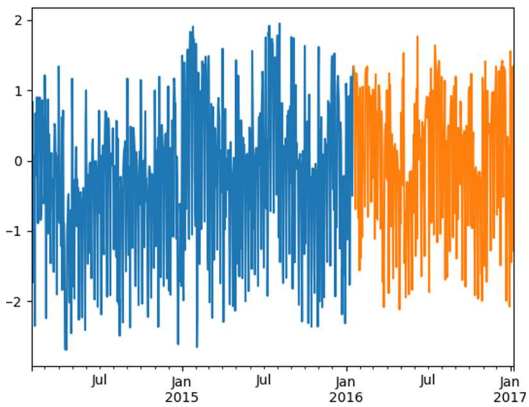
```
from sklearn.model_selection import TimeSeriesSplit, ParameterGrid

splits = TimeSeriesSplit(n_splits=3, max_train_size=365 * 2)
for train_index, val_index in splits.split(X_train):
print("TRAIN:", len(train_index), "TEST:", len(val_index))
y_train["target_t1"][train_index].plot()
y_train["target_t1"][val_index].plot()
plt.show()
```

Output

```
TRAIN: 360 TEST: 357
```




```
TRAIN: 717 TEST: 357
```




```
TRAIN: 730 TEST: 357
```

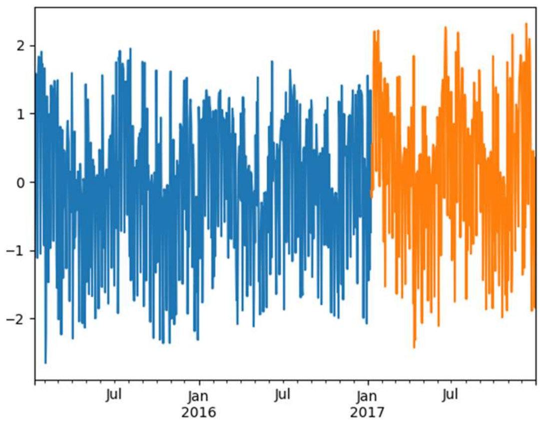


Figure 43. Trained Sets

8.0 Conclusion

Thanks to the Predictive Maintenance, the occurrence of corrective maintenance interventions is greatly reduced, as well as their environmental repercussions. However, strategies where the service intervals or the used time of the components are prolonged without imposing on safety levels will likely have a positive effect on both environmental and cost performance. Furthermore, a product consists of several components and there can be a variety of maintenance plans for one product. It is a system of systems where changes in one area will most likely affect other areas. This can be addressed by seeing it as a part integrated into a larger system, and by implementing new maintenance plans that is evaluated for the overall system.

This project clearly reveals that integration of artificial intelligence in Predictive Maintenance and Fault Detection brings numerous benefits to industries. By harnessing the power of AI, organizations can detect faults in advance, optimize maintenance operations, improve accuracy, enhance safety, cost savings and improve overall operational efficiency. Despite the numerous benefits, challenges related to data quality, algorithm selection, interpretability, integration, overall maintenance plans need to be addressed. Overcoming these challenges and investing in robust AI solutions, businesses can stay ahead of potential faults, reduce downtime, and achieve significant competitive advantages in today's fast-paced and technology-driven landscape.

9.0 References

[1]A Review of Machine Learning Models for Forecasting Electricity Consumption

[2]Jui-Sheng Chou, Duc-Son Tran,Forecasting energy consumption time series usingmachine learning techniques based on usage patterns of residentialhouseholders,Energy,Volume 165, Part B,2018,Pages709-726,ISSN0360-5442

[3]A. González-Briones, G. Hernández, J. M. Corchado, S. Omatu and M. S. Mohamad,"Machine Learning Models for Electricity Consumption Forecasting: A Review,"2019 2nd International Conference on Computer Applications & Information Security(ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6

[4]A 2019 study by Alsmadi, Al-Madi, and Al-Turjman. a short-term powerconsumption forecasting method based on machine learning. IEEE Access, 7,141329–141338.

[5]In 2019, Wang, X., Li, Y., and Wang, K. an innovative extreme learning machinealgorithm-based prediction model for power use. Energy 12, page 3283.

[6]Zhang, J., Gao, F., Liu, L., and Zou, Y. (2021). A thorough examination of short-termload prediction utilising machine learning methods. 14, no. 10 (energy), 2958.

[7]In 2020, Xu, X., Yang, Y., Jiang, L., and Wang, W. Forecasting electricity use basedon upgraded extreme learning machine and meteorological data. 51162-51172. IEEEAccess, 8.

[8]Almutairi, A. H., Alqahtani, A. M., & Alqahtani, M. A. (2020). Alotaibi, M. R. Aoverview of machine learning-based methods for forecasting power use. SustainableDevelopment, 12(8), 3381.

[9]Hu, M., Huang, and Chen, Y. (2020). a hybrid strategy combining machine learningand wavelet transform for estimating power usage. 20632-20644. IEEE Access, 8.

[10] 2019: Yin, J., Li, H., Zhang, X., and Liu, X. a method based on deep learning forpredicting power usage in a smart grid. 12, 703 Energy

10.0 Appendix

A1- AHP Scale

1- Equal Importance,3- Moderate importance, 5- Strong importance, 7- Very strong importance, 9- Extreme importance (2,4,6,8 values in-between)