

Ângelo Emanuel Gonçalves Morgado de Sousa

**SISTEMA DE MONITORIZAÇÃO DE CONSUMO E
PRODUÇÃO ELÉTRICA DA UNIVERSIDADE DO
ALGARVE**



UNIVERSIDADE DO ALGARVE
Instituto Superior de Engenharia
2024/2025

Ângelo Emanuel Gonçalves Morgado de Sousa

**SISTEMA DE MONITORIZAÇÃO DE CONSUMO E
PRODUÇÃO ELÉTRICA DA UNIVERSIDADE DO
ALGARVE**

Mestrado em Engenharia Eletrotécnica e de Computadores
(Especialidade em TIT – Tecnologia de Informação e Telecomunicações)

Trabalho realizado sob a orientação de:

- Prof. Dr. Jânio Monteiro
- Prof. Dr. Pedro Cardoso



UNIVERSIDADE DO ALGARVE
Instituto Superior de Engenharia
2024/2025

SISTEMA DE MONITORIZAÇÃO DE CONSUMO E PRODUÇÃO ELÉTRICA DA UNIVERSIDADE DO ALGARVE

Declaração de autoria da obra

Declaro ser o autor desta obra, que é original e inédita. Os autores e obras consultados são devidamente citados no texto e constam da listagem de referências incluídas.

(Nome Completo do Autor)

©2024/2025, Ângelo Sousa

A Universidade do Algarve reserva para si o direito, em conformidade com o disposto no Código do Direito de Autor e dos Direitos Conexos, de arquivar, reproduzir e publicar a obra, independentemente do meio utilizado, bem como de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição para fins meramente educacionais ou de investigação e não comerciais, conquanto seja dado o devido crédito ao autor e editor respetivos.

AGRADECIMENTOS

A concretização desta dissertação apenas foi possível graças ao apoio, incentivo e contributo de várias pessoas, às quais expresso a minha sincera gratidão.

Aos meus orientadores, Prof. Dr. Jânio Monteiro e Prof. Dr. Pedro Cardoso, pela orientação científica, pelo apoio sempre que necessário e pela confiança depositada ao longo de todo o desenvolvimento deste trabalho.

Ao Engenheiro André Pedro, dos Serviços Informáticos da Universidade do Algarve, pela ajuda fundamental numa fase inicial do projeto e pela colaboração contínua, em especial no que respeita à configuração da rede da UAlg para a integração dos dispositivos de monitorização.

Ao Engenheiro Sérgio Brito, colega e amigo, pelas inúmeras sessões de partilha de ideias, reflexões e procura de soluções.

Aos meus pais, Emanuel Sousa e Anabela Gonçalves, pelo apoio incondicional, pela motivação constante e pela dedicação diária que sempre me acompanharam.

À minha namorada, Ana Mendonça, pela paciência, pelo incentivo e pela presença diária, que me deram força para ultrapassar os desafios encontrados.

A todos os que, direta ou indiretamente, colaboraram neste projeto, manifesto o meu reconhecimento.

RESUMO

A necessidade de aumentar a eficiência energética e reduzir custos operacionais levou a Universidade do Algarve a reestruturar o seu sistema de monitorização elétrica. Embora já existisse um sistema, este apresentava falhas na recolha de dados, dificuldades na integração de novos medidores e incompatibilidades com versões recentes de software, tornando essencial a sua modernização.

Este projeto contemplou a migração da plataforma de monitorização para um servidor centralizado nos Serviços Informáticos, garantindo maior estabilidade, segurança e integração com a plataforma institucional UAlgNet, um dos objetivos centrais. Foram também integrados edifícios anteriormente não monitorizados, permitindo uma análise mais abrangente e detalhada.

O script responsável pela leitura dos medidores foi revisto e otimizado, melhorando a fiabilidade da aquisição de dados e a sua integração com as plataformas existentes. Como parte desta reformulação, foi criada uma base de dados, onde cada leitura é armazenada após ser enviada para o EmonCMS, garantindo um histórico detalhado das medições. Foi também desenvolvido um novo método de cálculo de custos energéticos, mais eficiente e compatível com versões atualizadas do software, assegurando maior precisão no cálculo tarifário tetra-horário.

Para garantir a robustez do sistema, foi concebida uma infraestrutura redundante, permitindo que a plataforma opere simultaneamente em dois locais distintos, assegurando a continuidade da monitorização elétrica em caso de falha de um dos locais.

O projeto integra ainda o desenvolvimento de modelos de *machine learning* para previsão de consumos e deteção de anomalias e garantia de continuidade dos dados perante falhas nos dispositivos de medição.

Com estas melhorias, a Universidade do Algarve passa a dispor de uma solução integrada e robusta para a gestão energética, promovendo a eficiência operacional, a redução de custos e um consumo energético mais sustentável.

Palavras-chave: Monitorização Energética, Gestão Energética, Eficiência Energética, EmonCMS, *Machine Learning*

ABSTRACT

The growing need to enhance energy efficiency and reduce operational costs led the University of Algarve to restructure its electrical monitoring system. Despite the existence of a previous system, it suffered from data collection failures, difficulties in integrating new meters, and incompatibility with updated software, making its modernization imperative.

This project involved migrating the monitoring platform to a centralized server hosted by the University's IT Services (SI), ensuring improved stability, security, and integration with UAlgNet—one of the key objectives. Additional buildings, previously unmonitored, were integrated into the system, enabling broader and more detailed energy analysis.

The meter reading script was redesigned and optimized to enhance data acquisition reliability and ensure seamless integration with the monitoring platforms. A database was implemented to store each measurement after being transmitted to EmonCMS, thus preserving a complete historical record. A new cost calculation method was also developed, compatible with updated software versions and capable of accurately applying the time-of-use energy tariff.

To ensure system resilience, a redundant infrastructure was deployed, allowing the platform to operate in parallel across two separate locations, thereby maintaining monitoring continuity in case of failure of one of the systems.

Furthermore, the project includes the development of machine learning models to forecast energy consumption, detect anomalies, and ensure data continuity in the event of meter malfunction.

With these advancements, the University of Algarve now possesses a robust and integrated energy monitoring solution that enhances operational efficiency, reduces costs, and supports more sustainable energy use.

Keywords: Energy Monitoring, Energy Management, Energy Efficiency, EmonCMS, Machine Learning

ÍNDICE

1	Introdução	1
2	Revisão Bibliográfica / Estado da Arte	5
2.1	Sistemas de Monitorização Energética	6
2.1.1	Arquitetura Geral	6
2.2	Comunicação e Transmissão de Dados	11
2.2.1	Modbus	11
2.2.2	API REST	13
2.2.3	MQTT	14
2.3	Medidores de Energia	15
2.3.1	Dispositivos da Marca Schneider Electric	16
2.3.2	Dispositivos da Marca Shelly	18
2.3.3	Outros Dispositivos de Medição	20
2.4	Grandezas Elétricas na Monitorização Energética	23
2.5	Plataformas de Monitorização Energética	24
2.5.1	EmonCMS – Plataforma <i>Open-Source</i> para Monitorização Energética	24
2.5.2	Grafana – Ferramenta Avançada de Visualização de Dados	30
2.5.3	Outras Plataformas IoT para Monitorização Energética	31
2.6	Bases de Dados para Armazenamento de Dados Energéticos	32
2.6.1	PostgreSQL – Base de Dados Relacional	32
2.6.2	InfluxDB – Base de Dados Temporal para Dados Energéticos	34
2.6.3	TimescaleDB – Expansão do PostgreSQL para Séries Temporais	34
2.6.4	SQLite	35
2.7	Sistemas Operativos e Infraestrutura Necessária para Implementação	35
2.7.1	Linux Ubuntu e a sua Aplicabilidade na Monitorização Energética	36
2.7.2	Gestão de Serviços e Processos no Ubuntu	36
2.7.3	Administração e Transferência de Ficheiros em Ambiente Linux	37
2.7.4	Automação de Tarefas e Monitorização Contínua	37
2.7.5	Segurança e Manutenção da Infraestrutura	38

2.7.6	Linguagem de Programação Python	39
2.8	Métodos de Cálculo de Custos Energéticos	40
2.9	<i>Machine Learning</i> Aplicado à Previsão e Detecção de Anomalias no Consumo Energético	45
2.9.1	ARIMA (<i>AutoRegressive Integrated Moving Average</i>).....	45
2.9.2	Prophet	46
2.9.3	<i>Isolation Forest</i>	47
2.9.4	<i>Local Outlier Factor</i>	48
2.9.5	Métricas de Avaliação da Previsão	49
2.9.6	Métricas de Avaliação em Detecção Não Supervisionada	52
2.10	Conclusão Crítica da Revisão Bibliográfica	53
3	Desenvolvimento do Trabalho Realizado	55
3.1	Arquitetura Inicial do Sistema	56
	Distribuição dos Medidores no Campus das Gambelas.....	59
	Distribuição dos Medidores no Campus da Penha.....	60
3.2	Arquitetura Intermédia do Sistema	64
3.2.1	Reestruturação e Otimização do Código.....	68
3.2.2	Aquisição de Dados via Modbus/TCP/IP	70
3.2.3	Aquisição de Dados via HTTP.....	72
3.2.4	Migração da Base de Dados PostgreSQL	73
3.2.5	Migração da Base de Dados SQLite	74
3.2.6	Integração Adicional de Edifícios Externos.....	75
3.3	Arquitetura Atual do Projeto.....	76
3.3.1	Reestruturação e Otimização do Código de Aquisição de Dados.....	78
3.3.2	Consolidação e Estruturação da Nova Base de Dados Energética.....	81
3.3.3	Custo Acumulado Mensal e Mecanismo de Reinicialização	83
3.3.4	Desenvolvimento e Planeamento de <i>Dashboards</i> Interativos.....	83
3.3.5	Particionamento de Medições no Interior dos Edifícios Monitorizados.....	85
3.3.6	Previsão Meteorológica Através de API.....	86
3.4	<i>Machine Learning</i> para Previsão Energética	88
3.4.1	Análise Preliminar dos Dados.....	89
3.4.2	Construção dos Conjuntos de Dados	89
3.4.3	Pré-Processamento e Limpeza	90

3.4.4	Estratégia de Treino e Otimização.....	94
3.4.5	Resultados do Treino e Otimização.....	98
3.4.6	Resultados da Previsão	99
3.4.7	Implementação Operacional do Modelo.....	102
3.5	<i>Machine Learning</i> para Detecção de Anomalias.....	103
3.5.1	Resíduos e Decisão por Persistência.....	103
3.5.2	Classificação Final com <i>Isolation Forest</i>	106
3.5.3	Detecção de Anomalias de Rede e Infraestrutura.....	109
4	Resultados Experimentais.....	110
4.1	Discussão	117
5	Conclusões e Trabalhos Futuros.....	119
5.1	Trabalhos Futuros	120
	Referências	122
	Apêndices	129
	A Tabela de Análise da Integridade dos Registos	130
	B Tabela do TOP 10 de combinações de hiperparâmetros.....	131

LISTA DE FIGURAS

Figura 1 – Comunicação Modbus [17]	11
Figura 2 – Estrutura de mensagens Modbus, RTU (cima) e TCP/IP (baixo) [19].....	12
Figura 3 – Fluxo de comunicação entre cliente e servidor numa API REST [22].....	13
Figura 4 – Estrutura de uma rede MQTT [24]	15
Figura 5 – Medidor PM710 [27].....	16
Figura 6 – Medidor PM5100 [29].....	17
Figura 7 – Medidor PM3250 [30].....	17
Figura 8 – Medidor Shelly EM com transformador [32].....	19
Figura 9 – Medidor Shelly 3EM [33]	19
Figura 10 – Medidor Shelly Pro 3EM [34].....	20
Figura 11 – Interface Modbus TRV00210.....	21
Figura 12 – Inversor Huawei Sun2000 [35]	22
Figura 13 – Equipamento de monitorização SolarLog [37].....	22
Figura 14 – Exemplo de dashboard no EmonCMS [44].....	30
Figura 15 – Exemplo de dashboard no Grafana [49].....	31
Figura 16 – Interface Valentina Studio.....	33
Figura 17 – Gráfico representativo do Ciclo Tetra-horário semanal [70]	42
Figura 18 – Exemplo genérico da estrutura da rede interna da UAlg.....	57
Figura 19 – Arquitetura inicial do sistema.....	62
Figura 20 – Arquitetura numa fase intermédia	65
Figura 21 – Conversão de registos Modbus.....	71
Figura 22 – Arquitetura na fase atual do sistema.....	78
Figura 23 – Interligação entre Grafana e PostgreSQL com dashboards interativos	85
Figura 24 – Integração entre sistema de monitorização e API meteorológica	88
Figura 25 – Mapa de calor da correlação de Spearman em função do lag	91
Figura 26 – Esquema de validação temporal do treino na primeira fase da otimização do Prophet.....	96
Figura 27 – Esquema de validação temporal do treino na segunda fase da otimização do Prophet.....	97
Figura 28 – Comparação da potência ativa real com a previsão em 24 horas do conjunto em holdout	100
Figura 29 – Evolução das métricas MAE e RMSE durante o período de teste (24h).....	101

Figura 30 – Potência ativa real e prevista no mês de outubro de 2025, com intervalos de previsão sombreados	102
Figura 31 – Séries temporais de potência ativa observada e previsões e limites de consumo do Edifício 1 no EmonCMS.....	103
Figura 32 – Detecção de eventos anômalos através de resíduos com persistência	105
Figura 33 – Potência ativa com detecções do algoritmo Isolation Forest, entre 4 de setembro e 26 de outubro de 2025	107
Figura 34 – Potência ativa no edifício 1, com os alarmes finais obtidos pela interseção entre resíduos persistentes e Isolation Forest, entre 4 de setembro e 26 de outubro de 2025	108
Figura 35 – Detalhe do período de 15 a 22 de setembro de 2025, com pormenor dos alarmes finais resultantes da interseção entre resíduos persistentes e Isolation Forest.....	108
Figura 36 – Visualizações do tipo gauge para monitorizar variáveis instantâneas.....	112
Figura 37 – Visualização em série temporal para variáveis relativas a consumo energético	113
Figura 38 – Visualização do custo instantâneo em série temporal e do acumulado em gráfico de barras.....	113
Figura 39 – Visualização da curva de consumo real e de previsão com limites de erro definidos.....	114
Figura 40 – Arquitetura final do sistema com fluxos de dados	116
Figura 41 – Proposta de esquema relacional da base de dados no futuro.....	120

LISTA DE TABELAS

Tabela 1 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.2 no Campus das Gambelas	59
Tabela 2 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.3 no Campus das Gambelas	59
Tabela 3 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.4 no Campus das Gambelas	59
Tabela 4 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.5 no Campus das Gambelas	60
Tabela 5 – Edifícios monitorizados pelo gateway com o IP: 10.18.22.12 no Campus da Penha	60
Tabela 6 – Edifícios monitorizados pelo gateway com o IP: 10.33.13.5 no Campus da Penha	60
Tabela 7 – Geração de edifícios monitorizados com IPs: 10.17.22.9, 10.17.22.10 e 10.17.22.11 no Campus da Penha.....	61
Tabela 8 – Geração do Edifício em U do ISE com IP:10.17.22.8 no Campus da Penha	61
Tabela 9 – Exemplo de três registos do dataset preparado para o modelo Prophet – ds (timestamp), y (potência ativa), radiação (W/m ²), temperatura (°C), humidade (%), pluviosidade (binária 0/1)	94
Tabela 10 – Espaço de procura dos hiperparâmetros e respetivas distribuições de amostragem	95
Tabela 11 – Métricas dos melhores hiperparâmetros nos 4 cutoffs.....	98
Tabela 12 – Métricas agregadas dos melhores hiperparâmetros nos 33 cutoffs.....	99
Tabela 13 - Métricas dos melhores hiperparâmetros no conjunto de teste (24h)	100
Tabela 14 – Métricas médias do modelo no mês de outubro de 2025.....	101

LISTA DE SIGLAS E ACRÓNIMOS

ACID – *Atomicity, Consistency, Isolation, Durability*
ANN – *Artificial Neural Network*
API – *Application Programming Interface*
ARIMA – *AutoRegressive Integrated Moving Average*
AVAC – Aquecimento, Ventilação e Ar Condicionado
BD – Base de Dados
BO – Otimização Bayesiana
CRC – *Cyclic Redundancy Check*
CRUD – *Create, Read, Update, Delete*
CSV – *Comma-Separated Values*
EDP – Energias de Portugal
EMS – *Energy Management System*
ERSE – Entidade Reguladora dos Serviços Energéticos
GAIA – *Green Awareness in Action*
GB – *Gigabyte*
GET – Método HTTP para leitura de dados
HTTP – *HyperText Transfer Protocol*
ID – Identificador
IEC – *International Electrotechnical Commission*
IEEE – *Institute of Electrical and Electronics Engineers*
IF – *Isolation Forest*
IP – *Internet Protocol*
ISE – Instituto Superior de Engenharia
ISO – *International Organization for Standardization*
JSON – *JavaScript Object Notation*
LAB160 – Laboratório 160
LAN – *Local Area Network*
LED – *Light Emitting Diode*
LOF – *Local Outlier Factor*
LSB – *Least Significant Byte*

LSTM – *Long Short-Term Memory*
LTS – *Long Term Support*
MA – *Moving Average*
MAD – *Median Absolute Deviation*
MAE – *Mean Absolute Error*
MASE – *Mean Absolute Scaled Error*
MBAP – *MODBUS Application Protocol*
ML – *Machine Learning*
MQTT – *Message Queuing Telemetry Transport*
MSB – *Most Significant Byte*
MVCC – *Multi-Version Concurrency Control*
PC – *Personal Computer*
PCA – *Principal Component Analysis*
PHP – *PHP Hypertext Preprocessor*
POST – *Método HTTP para envio de dados*
PUT – *Método HTTP para atualização de dados*
QoS – *Quality of Service*
RAM – *Random Access Memory*
REST – *Representational State Transfer*
RF – *Random Forest*
RMSE – *Root Mean Squared Error*
RS232 – *Recommended Standard 232*
RS485 – *Recommended Standard 485*
RTU – *Remote Terminal Unit*
SARIMA – *Seasonal AutoRegressive Integrated Moving Average*
SCADA – *Supervisory Control and Data Acquisition*
SCP – *Secure Copy Protocol*
SI – *Serviços Informáticos*
SQL – *Structured Query Language*
SSH – *Secure Shell*
SVM – *Support Vector Machine*
TCP – *Transmission Control Protocol*
THD – *Total Harmonic Distortion*
TIT – *Tecnologias de Informação e Telecomunicações*

TLS – Transport Layer Security
UAig – Universidade do Algarve
URL – *Uniform Resource Locator*
VLAN – *Virtual Local Area Network*
VPN – *Virtual Private Network*
WAN – *Wide Area Network*
WAPE – *Weighted Absolute Percentage Error*
XML – *Extensible Markup Language*

1 INTRODUÇÃO

A crescente necessidade de gestão eficiente dos recursos energéticos e a redução de custos operacionais tornaram-se prioridades para instituições de ensino superior, não apenas por razões económicas, mas também ambientais. Atenta a esta realidade, a Universidade do Algarve (UAlg) apostou na modernização do seu sistema de monitorização elétrica, com o objetivo de garantir maior precisão na recolha e análise de dados de consumo energético em todos os seus edifícios.

Apesar de já existir uma infraestrutura de monitorização, esta apresentava limitações significativas, tais como falhas na recolha de dados, dificuldades na integração de novos medidores, incompatibilidades com software atualizado e ausência de uma infraestrutura de armazenamento robusta. Estas limitações comprometiam a abrangência e a fiabilidade da análise energética, dificultando a implementação de estratégias eficazes de otimização. Face a este cenário, tornou-se essencial reformular o sistema, tornando-o mais robusto, escalável e adaptável às necessidades atuais e futuras da instituição. Este projeto surge precisamente como resposta a essas necessidades, focando-se na atualização da plataforma de monitorização, na expansão da cobertura dos medidores, na reformulação do processamento de dados e na implementação de soluções avançadas para previsão e deteção de anomalias no consumo energético.

1.1 Objetivos e Características do Projeto

O presente projeto tem como principal objetivo a reestruturação e expansão do sistema de monitorização energética da UAlg, garantindo a obtenção de dados consistentes que suportem a avaliação e gestão eficiente da infraestrutura energética.

Nesse contexto, foram estabelecidos os seguintes objetivos específicos:

- Migrar a plataforma de monitorização existente para um servidor centralizado nos Serviços Informáticos (S.I.), permitindo segurança e maior integração com outras infraestruturas institucionais;
- Expandir a cobertura da monitorização energética, integrando novos edifícios que até então não eram monitorizados, garantindo um controlo mais abrangente dos consumos energéticos da universidade;
- Rever e otimizar o código em Python responsável pela leitura dos medidores, tornando a aquisição de dados mais eficiente e integrando uma base de dados para armazenamento estruturado das medições;
- Implementar um novo método de cálculo de custos energéticos, assegurando uma maior precisão na aplicação do tarifário tetra-horário e facilitando a análise financeira do consumo energético da universidade;
- Criar um ambiente redundante, permitindo que o sistema opere simultaneamente nos Serviços Informáticos e num servidor do Lab160, garantindo continuidade e resiliência na monitorização elétrica;
- Desenvolver modelos de *Machine Learning*, que permitirão prever consumos, detetar anomalias e assegurar a continuidade dos dados mesmo em caso de falha dos medidores.

Com estas melhorias todas concluídas, que no decorrer deste documento explicarei como foram implementadas, o sistema de monitorização passará a ser mais robusto, escalável e adaptado às necessidades futuras da instituição.

Neste contexto, o relatório do presente projeto encontra-se organizado em cinco capítulos. Após esta introdução, o segundo capítulo apresenta a revisão do estado da arte, abordando os principais conceitos relacionados com sistemas de monitorização energética, plataformas de recolha e visualização de dados, tecnologias de aquisição e técnicas de previsão baseadas em algoritmos de aprendizagem de máquina. O terceiro capítulo expõe a metodologia adotada, descrevendo em detalhe a evolução da arquitetura do sistema, desde a configuração inicial, passando por uma fase intermédia de transição, até à implementação da arquitetura atual. O quarto capítulo apresenta e discute os resultados experimentais, incluindo a análise e

interpretação dos dados recolhidos e a avaliação da eficácia das melhorias implementadas. Por fim, o quinto capítulo apresenta as conclusões e perspetivas futuras, destacando o impacto da solução proposta e potenciais desenvolvimentos no âmbito da monitorização e gestão energética.

2

REVISÃO BIBLIOGRÁFICA / ESTADO DA ARTE

A monitorização e optimização do consumo energético têm sido temas amplamente estudados, tanto no meio académico como na indústria, devido à crescente necessidade de eficiência energética, redução de custos operacionais e sustentabilidade ambiental. A evolução dos sistemas de medição, armazenamento e análise de dados tem permitido um controlo mais rigoroso e inteligente do consumo eléctrico em diversos contextos, como por exemplo, em instituições de ensino superior, empresas e infraestruturas industriais.

Este capítulo apresenta um levantamento detalhado do estado da arte nas áreas de sistemas de monitorização de energia, comunicação e transmissão de dados, métodos de cálculo de custos energéticos, bases de dados e abordagens baseadas em *Machine Learning* para previsão e optimização do consumo energético. O objetivo é enquadrar o presente projeto no contexto técnico e científico atual, identificar desafios frequentemente apontados em estudos anteriores e destacar oportunidades de melhoria.

A estrutura do capítulo está organizada da seguinte forma: A Secção 2.1 apresenta os sistemas de monitorização energética e as tecnologias de recolha e análise de dados. A Secção 2.2 descreve os principais protocolos de comunicação, com vantagens e limitações. A Secção 2.3 reúne exemplos de medidores de energia. A Secção 2.4 explica as grandezas eléctricas medidas e a sua relevância. A Secção 2.5 compara plataformas como EmonCMS, Grafana e OpenEMS. A Secção 2.6 discute bases de dados para séries temporais, incluindo PostgreSQL, InfluxDB e TimescaleDB. A Secção 2.7 indica requisitos de sistemas operativos e infraestrutura. A Secção 2.8 apresenta métodos de cálculo de custos e tarifários dinâmicos. A Secção 2.9 aborda modelos de previsão e deteção de anomalias, como ARIMA, Prophet, redes neuronais, Isolation Forest e LOF. A Secção 2.10 encerra com uma síntese crítica dos principais desafios e oportunidades.

2.1 SISTEMAS DE MONITORIZAÇÃO ENERGÉTICA

Os sistemas de monitorização energética têm evoluído de simples contadores mecânicos para plataformas digitais integradas, capazes de recolher, processar, armazenar e visualizar dados em tempo real. A sua adoção em instituições de ensino superior revela-se particularmente eficaz para promover a eficiência energética, reduzir custos operacionais e fomentar a consciencialização ambiental

2.1.1 ARQUITETURA GERAL

A arquitetura de um sistema de monitorização energética avançado integra essencialmente três camadas técnicas, como a seguir se detalha [1].

A camada de medição e aquisição de dados é composta por sensores de corrente, sensores de tensão e medidores multifuncionais que registam os parâmetros elétricos a taxas de amostragem pré-definidas, permitindo ajustar a resolução conforme a relevância da variável monitorizada e a dinâmica do sistema. Muitos destes dispositivos integram microcontroladores com conversores analógico-digitais de 24 bits, filtragem digital e capacidade de fornecer localmente variáveis energéticas derivadas, como a energia ativa, reativa e aparente, calculadas a partir das grandezas fundamentais de tensão e corrente.

A camada de concentração e pré-processamento é responsável pela agregação, validação e preparação dos dados recolhidos. Os *gateways*, enquanto nós intermédios, agregam medições de múltiplos dispositivos através de interfaces como o *Recommended Standard 485* (RS-485), assegurando a sincronização temporal das leituras, a compressão de dados, a filtragem de ruído e, em alguns casos, a realização de análises preliminares, como a verificação de consistência, antes de encaminharem a informação para o sistema central. A camada de processamento e visualização centraliza o armazenamento e a análise dos dados em servidores locais ou plataformas em nuvem, utilizando bases de dados otimizadas. Nesta fase podem ser aplicados algoritmos de análise estatística e energética, incluindo a identificação de padrões de consumo, tendências operacionais e cálculo de indicadores de desempenho energético. A informação pode ser posteriormente disponibilizada em *dashboards* interativos com funcionalidades avançadas de exploração em tempo real, geração automática de relatórios e emissão de alertas parametrizáveis perante desvios de comportamento.

A aplicação prática destas soluções tem demonstrado benefícios técnicos relevantes, especialmente em ambientes industriais, onde a implementação de estratégias automatizadas

de gestão de cargas contribui para uma utilização mais equilibrada da potência disponível, atenuando picos de consumo e melhorando a estabilidade da rede interna. Adicionalmente, a integração de mecanismos de compensação de energia reativa permite otimizar o fator de potência, reduzindo perdas energéticas e evitando penalizações tarifárias, o que se traduz numa operação mais eficiente e economicamente vantajosa.

2.1.2 Aplicações e Estratégias Digitais de Monitorização

Existem diversos exemplos de aplicações e estratégias digitais de monitorização energética que demonstram o seu impacto em contextos educacionais e institucionais. Um deles é o Projeto GAIA (*Green Awareness in Action*), desenvolvido entre 2016 e 2019 no âmbito do programa Horizon 2020 da União Europeia, foi implementado em 50 escolas europeias com o objetivo de monitorizar o consumo energético em tempo real e promover comportamentos sustentáveis [2], [3]. Através da integração de sensores IoT e *dashboards* interativos, foi possível identificar ineficiências, aplicar medidas corretivas e alcançar uma redução média de 20% no consumo energético. Os dados foram ainda utilizados em contexto educativo, contribuindo para o envolvimento ativo da comunidade escolar.

A norma ISO 50001, que estabelece diretrizes internacionais para a criação, implementação e melhoria contínua de sistemas de gestão de energia, tem vindo a ser adotada por diversas instituições de ensino superior com resultados promissores [4]. No Reino Unido, a Cranfield University encontra-se certificada segundo a ISO 50001, integrando a monitorização contínua e melhoria do desempenho energético no seu sistema de gestão [5]. A iniciativa incluiu a monitorização contínua do consumo energético em edifícios e infraestruturas técnicas através de sistemas automáticos de aquisição de dados em tempo real, permitindo a deteção precoce de ineficiências e o planeamento de ações corretivas. Foram igualmente instaladas infraestruturas de elevado desempenho, como sistemas de cogeração e painéis fotovoltaicos, e criada uma estrutura interna de controlo energético, com envolvimento ativo de docentes, técnicos e estudantes. A estratégia adotada privilegiou a gestão proativa de cargas, o feedback operacional imediato e a consciencialização da comunidade académica para práticas sustentáveis. Esta abordagem integrada conduziu a melhorias contínuas na eficiência energética e a uma redução significativa das emissões de carbono, comprovando a eficácia das orientações da norma ISO 50001 na racionalização do consumo e na consolidação de uma cultura institucional de sustentabilidade.

Na Universidade de León, em Espanha, foi desenvolvido um sistema de monitorização energética aplicado a vários edifícios do campus, baseado numa arquitetura em três camadas que combinou hardware heterogéneo e ferramentas de análise avançada [6].

Na camada de aquisição de dados, foram instalados contadores elétricos em cada edifício, comunicando através dos protocolos Modbus e TCP/IP com *gateways* que agregavam as leituras em tempo real. Estes *gateways* enviavam periodicamente tensões, correntes e potências para o servidor central.

Na camada intermédia - servidor, as medições eram armazenadas num banco de dados relacional e processadas por um motor de análise que aplicava técnicas estatísticas e de pré processamento de dados, permitindo detetar padrões de consumo, avaliar incoerências no uso energético habitual e até prever picos de carga. Esta camada incluía ainda funcionalidades de compressão de dados e filtragem de ruído, garantindo que apenas informações relevantes fossem mantidas para análises históricas e previsionais.

Por sua vez a camada cliente assentou numa interface web acessível remotamente, compatível com qualquer navegador, onde os utilizadores podiam visualizar *dashboards* interativos em tempo real, gerar relatórios automatizados e configurar alertas parametrizáveis para notificações instantâneas em caso de consumos anómalos.

A configuração técnica implementada permitiu uma gestão centralizada com elevada granularidade, resultando numa redução de aproximadamente 15 % no consumo energético global dos edifícios monitorizados. Estes resultados evidenciam a eficácia de uma solução escalável e interoperável em contexto universitário, especialmente quando se conjugam contadores digitais, protocolos normalizados (Modbus/TCP-IP) e algoritmos de análise avançada orientados para o apoio à decisão [6].

Com base nestas experiências, os Sistemas de Gestão de Energia (EMS – *Energy Management Systems*) assumem um papel essencial na análise [7], planeamento e otimização do desempenho energético de uma instalação. Estas soluções operam numa lógica estratégica e de supervisão, permitindo definir metas de eficiência, acompanhar indicadores de desempenho e ajustar o uso da energia com base em dados históricos, previsões de consumo e tarifas dinâmicas. Entre os seus principais módulos, destacam-se o planeamento tarifário, com integração de calendários de preços de energia para otimização financeira, a gestão de demandas, que permite programar cargas não críticas em horários mais vantajosos, e a análise de eficiência, com geração de relatórios sobre variáveis como o consumo específico (kWh/m²), o fator de potência e os indicadores normalizados de desempenho energético (EnPI), que relacionam o consumo de energia com variáveis de referência como área útil, produção, taxa

de ocupação ou condições climáticas, permitindo avaliar a eficiência energética de forma comparável ao longo do tempo e entre diferentes contextos [7].

Por contraste, os sistemas SCADA (Supervisory Control and Data Acquisition) apresentam uma orientação essencialmente operacional e em tempo real, focada na monitorização contínua de processos e na capacidade de atuar diretamente sobre os equipamentos do sistema energético. Ao contrário dos EMS, que operam a um nível mais agregador e analítico, os SCADA são desenhados para garantir a continuidade e segurança da operação, permitindo acionar comandos remotos, responder automaticamente a falhas e ajustar parâmetros críticos em tempo real. Esta capacidade de controlo granular e instantâneo torna-os indispensáveis em contextos industriais, redes elétricas e infraestruturas críticas, como demonstrado por Elias et al. [8], que propõem uma solução SCADA de baixo custo para uma instalação fotovoltaica, baseada na plataforma *open-source* EmonCMS e com funcionalidades locais de aquisição e armazenamento de dados [9].

Um outro exemplo de projeto é o ERIGrid – *European Research Infrastructure for Smart Grids*, financiado pela União Europeia no âmbito do programa Horizon 2020. Este projeto criou uma infraestrutura europeia colaborativa dedicada à validação do funcionamento de sistemas energéticos inteligentes. A componente central do projeto foi a integração funcional de sistemas SCADA com sistemas de gestão de energia (EMS) em ambientes de simulação e laboratório, permitindo testar estratégias avançadas de controlo energético, resposta à procura e integração de fontes renováveis, [10]. Através da combinação entre supervisão em tempo real e algoritmos de controlo, o projeto demonstrou como arquiteturas híbridas podem otimizar fluxos energéticos, detetar falhas antecipadamente e suportar operações em redes inteligentes com elevada presença de fontes de energia descentralizadas, como sistemas fotovoltaicos, eólicos ou de cogeração local, que exigem uma gestão coordenada e em tempo real.

Power Factors – Projeto Híbrido Solar e Armazenamento foi um dos maiores projetos híbridos desenvolvidos na região EMEA (Europa, Médio Oriente e África) [11], localizado em Abu Dhabi, Emirados Árabes Unidos, com 540 MW de produção de energia solar fotovoltaica e 225 MWh de capacidade de armazenamento em baterias. A plataforma SCADA, integrada com um sistema de gestão de energia (EMS), assegura o controlo contínuo da produção, do estado de carga e da injeção na rede elétrica nacional, otimizando a operação em função da procura e das condições ambientais.

Com sede em Espanha, a GreenPowerMonitor [12] desenvolveu uma plataforma modular que combina SCADA e EMS para gestão integrada de parques solares e eólicos. A solução permite a monitorização em tempo real, previsão meteorológica, controlo preditivo e resposta ao

mercado de energia, ajustando dinamicamente a produção, o armazenamento e a injeção na rede elétrica. Está implementada em várias centrais renováveis na Europa, otimizando o desempenho energético e económico dos ativos.

Estes casos evidenciam que a utilização de SCADA em conjunto com EMS não só permite um controlo operacional mais granular e uma resposta mais célere a eventos críticos, como também favorece a integração de variáveis ambientais, previsões meteorológicas e algoritmos de otimização em tempo real. Em cenários de elevada penetração de renováveis, como parques solares e eólicos, esta convergência é vital para garantir estabilidade, eficiência e rentabilidade da operação, promovendo uma visão integrada entre produção, armazenamento e consumo.

No entanto, importa salientar que, para além das soluções SCADA e EMS, existem outras abordagens de monitorização energética que, mesmo com menor complexidade, comprovam impactos significativos na eficiência e gestão de consumos. O estudo de Dat et al. [13] demonstra, através de um caso aplicado a uma grande empresa industrial no Vietname, que a simples implementação de um sistema de monitorização de energia permitiu uma redução de 9,6% nos custos energéticos anuais. Este trabalho reforça a importância da adoção de medições distribuídas, tal como recomendado na norma europeia EN 17267:2019 [14], que define diretrizes para a conceção e aplicação de planos de medição energética com vista à melhoria contínua do desempenho energético. A medição distribuída, ao permitir a recolha de dados detalhados por área, processo ou equipamento, possibilita a deteção de consumos anómalos, a identificação de perdas localizadas e a aplicação de medidas corretivas específicas. Esta granularidade na informação é fundamental para uma gestão energética mais informada, segmentada e orientada para a ação.

Outro exemplo relevante é o trabalho desenvolvido por Granadeiro [15], que concebeu um sistema de baixo custo para monitorização energética e comportamental num gabinete universitário em Portugal. O sistema permitia recolher dados de consumo elétrico e variáveis ambientais como temperatura e luminosidade, analisando a sua correlação com o comportamento adaptativo dos ocupantes. Os resultados obtidos evidenciaram a importância da monitorização detalhada para compreender padrões de utilização, identificar ineficiências e ajustar condições de conforto sem comprometer o desempenho energético do espaço.

Estes exemplos reforçam que os sistemas de monitorização energética, mesmo em configurações mais simples, podem contribuir decisivamente para a racionalização de consumos, melhoria de conforto e aumento da consciência energética dos utilizadores. A sua integração progressiva, adaptada à complexidade de cada contexto, constitui um pilar

fundamental para a transformação digital da gestão energética e para a concretização de objetivos de sustentabilidade.

2.2 COMUNICAÇÃO E TRANSMISSÃO DE DADOS

Nos sistemas de monitorização energética, a comunicação entre sensores, medidores e plataformas analíticas é essencial para garantir a fiabilidade dos dados obtidos e a eficiência operacional. Entre os protocolos de comunicação mais utilizados encontra-se o *Modbus* e API REST, como detalhado a seguir.

2.2.1 MODBUS

O protocolo *Modbus* foi desenvolvido pela Modicon [16] — atualmente parte integrante da Schneider Electric — e rapidamente se consolidou como um padrão industrial devido à sua simplicidade de implementação e à elevada compatibilidade com uma vasta gama de equipamentos. A comunicação assenta numa arquitetura mestre-escravo, em que um único dispositivo mestre (tipicamente um controlador ou sistema de monitorização) inicia todas as trocas de informação. Os dispositivos escravos apenas respondem quando solicitados, cada um com um identificador único que permite ao mestre direcionar os pedidos de leitura ou escrita para os respetivos registos. A comunicação decorre de forma sequencial, garantindo que apenas um escravo responde por ciclo de comunicação, o que assegura a integridade dos dados. A Figura 1 ilustra este modelo, evidenciando a direção dos pedidos (*requests*) enviados pelo mestre e das respostas (*responses*) devolvidas pelos escravos [16].

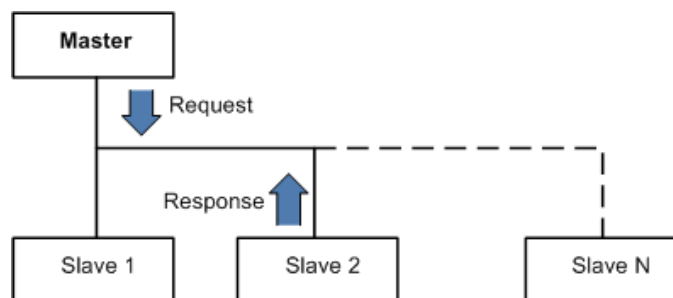


Figura 1 – Comunicação Modbus [17]

Atualmente, existem duas variantes principais do protocolo *Modbus*: RTU (*Remote Terminal Unit*) e TCP/IP (*Transmission Control Protocol/Internet Protocol*).

O *Modbus* RTU opera sobre comunicação serial, recorrendo tipicamente aos padrões RS485 ou RS232 (*Recommended Standard*) [18]. O RS485 é o mais utilizado em ambientes

industriais devido à sua robustez, capacidade de comunicação multiponto e maior imunidade a ruídos, permitindo que múltiplos dispositivos compartilhem um mesmo barramento de comunicação. O RS232, por outro lado, suporta apenas ligações ponto-a-ponto e apresenta limitações em termos de distância e velocidade. No *Modbus RTU*, os dados são transmitidos num formato binário compacto, o que resulta numa comunicação eficiente e com baixa latência. Cada trama *Modbus RTU* é composta por quatro elementos principais: o endereço do dispositivo escravo (Addr), o código de função (FCode), os dados transmitidos (Data) e o código de verificação de erro (CRC), que assegura a integridade da informação transmitida [16].

O *Modbus TCP/IP*, por sua vez, foi concebido para funcionar sobre redes Ethernet, permitindo uma comunicação mais rápida, escalável e flexível. As mensagens *Modbus* são encapsuladas em pacotes TCP/IP, o que possibilita a sua transmissão em redes locais (LAN), redes alargadas (WAN) e até através da Internet, eliminando as restrições físicas das ligações seriais. Nesta variante, o endereço IP do dispositivo substitui o conceito de endereço físico (Addr) utilizado no *Modbus RTU*, e a verificação de integridade da mensagem é assegurada pelos próprios mecanismos do protocolo TCP/IP, dispensando o uso de CRC no nível da aplicação [17].

Contudo, quando o *Modbus TCP/IP* é utilizado em conjunto com *gateways* que fazem a ponte para redes *Modbus RTU*, o campo *Unit Identifier*, presente no cabeçalho MBAP (*Modbus Application Protocol*), torna-se essencial [16]. Este campo permite ao *gateway* identificar qual o dispositivo escravo na rede serial (RS485) deve receber a mensagem, garantindo a interoperabilidade entre as duas variantes do protocolo. Assim, embora o *Unit Identifier* possa ser irrelevante em redes TCP/IP puras, a sua presença é crucial em arquiteturas híbridas, onde a comunicação entre dispositivos *Modbus TCP/IP* e RTU é necessária.

A Figura 2 apresenta, de forma comparativa, a estrutura de uma mensagem *Modbus RTU* e de uma mensagem *Modbus TCP/IP*, evidenciando as diferenças ao nível do encapsulamento, endereçamento e verificação de integridade.

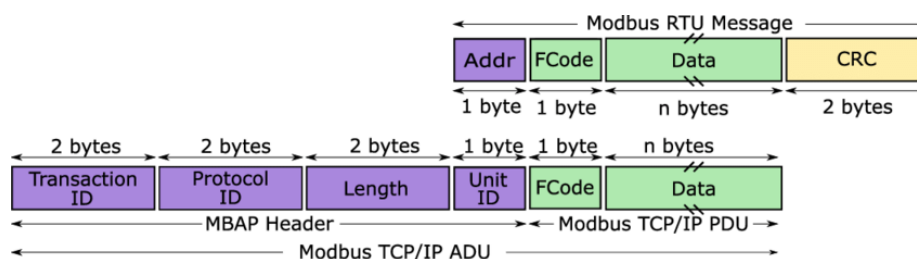


Figura 2 – Estrutura de mensagens *Modbus*, RTU (cima) e TCP/IP (baixo) [19]

2.2.2 API REST

Paralelamente aos protocolos tradicionais como o *Modbus*, as APIs (*Application Programming Interface*) de arquitetura REST (*Representational State Transfer*), constituem uma alternativa eficiente e moderna na comunicação de dados em sistemas de monitorização energética. Enquanto protocolos como o *Modbus* são essencialmente utilizados para comunicação direta entre dispositivos em redes locais, as API REST permitem a integração de sistemas através da internet e facilitam a troca de dados entre servidores, aplicações *web* e dispositivos IoT [20].

Baseia-se em princípios como a separação entre cliente e servidor, a ausência de estado (*statelessness*) e a possibilidade de armazenamento em cache das respostas, com o objetivo de otimizar o desempenho e reduzir a carga no servidor. As operações mais comuns numa API REST, habitualmente implementada sobre o protocolo HTTP/TCP, incluem GET (leitura de dados), POST (envio de novos dados), PUT (atualização de dados existentes) e DELETE (remoção de informações). Estas operações seguem o padrão CRUD (*Create, Read, Update, Delete*), facilitando a interação com bases de dados e sistemas de gestão de energia [21].

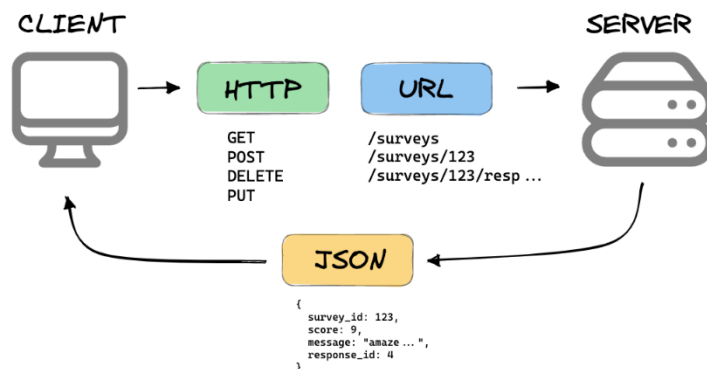


Figura 3 – Fluxo de comunicação entre cliente e servidor numa API REST [22]

A Figura 3 ilustra o fluxo de comunicação de uma API REST, evidenciando como um cliente — habitualmente um sensor, aplicação web ou plataforma IoT — comunica com um servidor através de requisições HTTP (*HyperText Transfer Protocol*) enviadas para URLs (*Uniform Resource Locator*) específicas. O servidor processa essas requisições e responde, por exemplo, em formato JSON (*JavaScript Object Notation*), com os dados solicitados ou com a confirmação da ação executada.

Este tipo de interface é amplamente utilizado em soluções baseadas na nuvem e em ambientes distribuídos, devido à sua flexibilidade, simplicidade e compatibilidade com diferentes sistemas e modelos de dispositivos. Comumente usada, a estrutura leve e legível do JSON contribuiu para a sua adoção como formato padrão, substituindo o XML (*Extensible Markup*

Language) em grande parte das aplicações modernas [21]. Além disso, muitas implementações recorrem a mecanismos de autenticação por *tokens*, como *API Keys*, para assegurar a segurança das comunicações e restringir o acesso a utilizadores autorizados.

Entre as principais vantagens deste modelo de comunicação destaca-se a facilidade de integração entre aplicações e dispositivos, independentemente de protocolos proprietários ou hardware específico. No contexto da monitorização energética, esta abordagem permite a recolha de dados em tempo real, o envio de comandos para equipamentos remotos e a integração com algoritmos de análise preditiva baseados em *machine learning*. Plataformas como o EmonCMS utilizam esta arquitetura para integrar medições de sensores e expandir continuamente o sistema, com a adição de novos dispositivos, serviços externos ou funcionalidades, sem necessidade de alterações estruturais [20].

2.2.3 MQTT

O MQTT (*Message Queuing Telemetry Transport*) é um protocolo de comunicação leve, baseado no modelo *publish/subscribe*, amplamente utilizado em sistemas *Internet of Things* (IoT) e em aplicações de monitorização energética distribuída [23]. Ao contrário de arquiteturas tradicionais ponto-a-ponto, em que um cliente comunica diretamente com um servidor, o MQTT introduz a figura de um *broker*, que atua como intermediário central da comunicação. O *broker* é responsável por receber as mensagens dos publicadores, processá-las e distribuí-las apenas aos subscritores interessados nos tópicos correspondentes, assegurando a eficiência, escalabilidade e desacoplamento entre emissores e recetores. Esta abordagem permite desacoplar os dispositivos emissores dos recetores, aumentando a escalabilidade e reduzindo a complexidade da rede.

O protocolo foi concebido para funcionar em redes com largura de banda limitada, elevada latência ou conectividade instável, sendo otimizado para ambientes onde a fiabilidade da comunicação é crítica e os recursos de hardware são restritos. O protocolo suporta diferentes níveis de Qualidade de Serviço (QoS), permitindo ajustar o compromisso entre desempenho e fiabilidade na entrega de mensagens. O nível QoS 0 corresponde ao envio sem confirmação de receção, adequado a aplicações onde a perda pontual de dados não compromete a operação. O nível QoS 1 assegura que a mensagem é entregue pelo menos uma vez, podendo ocorrer duplicações que devem ser tratadas pela aplicação. O nível QoS 2 garante a entrega única e confirmada de cada mensagem, sendo o mais seguro, mas também o mais exigente em termos de tráfego de rede e latência [24].

No contexto da monitorização energética, o MQTT é frequentemente utilizado para recolher medições de sensores ou medidores distribuídos e encaminhá-las para plataformas de análise ou armazenamento central. A sua leveza, combinada com a simplicidade de implementação, permite integrar rapidamente novos dispositivos e assegurar a transmissão contínua de dados em tempo real. Para além disso, o suporte a mecanismos de autenticação e encriptação, quando implementado sobre TLS (*Transport Layer Security*), reforça a segurança da comunicação e a proteção contra acessos indevidos. Graças a estas características, o MQTT consolidou-se como um protocolo de referência em arquiteturas IoT aplicadas à energia, oferecendo elevada flexibilidade, interoperabilidade e eficiência na transmissão de dados, em complementaridade com protocolos como o Modbus e interfaces baseadas em APIs REST. A Figura 4 ilustra o funcionamento do modelo *publish/subscribe* no protocolo MQTT, evidenciando o papel central do *broker* na receção de mensagens publicadas em diferentes tópicos e na sua distribuição subsequente aos subscritores [24].

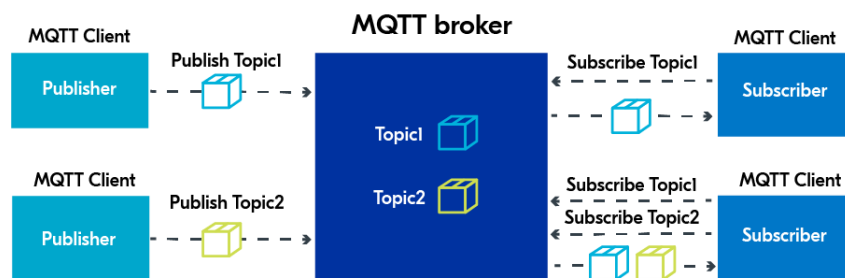


Figura 4 – Estrutura de uma rede MQTT [24]

2.3 MEDIDORES DE ENERGIA

A medição precisa e contínua de variáveis elétricas é um dos pilares essenciais da monitorização energética em edifícios e infraestruturas técnicas. O mercado disponibiliza uma vasta gama de medidores, com diferentes graus de precisão, funcionalidades, protocolos de comunicação e formatos físicos, adaptados a distintos contextos de aplicação — principalmente em edifícios públicos e industriais. Um dos parâmetros mais relevantes na escolha de um medidor de energia é a sua classe de precisão, normalmente definida pelas normas IEC 62053-21 [25] (para classe 1) e IEC 62053-22 [26] (para classe 0.5S ou 0.2S). Medidores de classe 1 possuem uma margem de erro até 1%, sendo geralmente utilizados em aplicações que não sejam críticas. Já os medidores de classe 0.5S ou superior (como 0.2S) oferecem maior precisão e fiabilidade, sendo recomendados para situações onde é necessário um controlo rigoroso, como em auditorias energéticas e faturação certificada.

No contexto deste projeto foram analisados diversos modelos de fabricantes distintos, sendo possível agrupá-los por marca. As subseções seguintes descrevem, de forma técnica e comparativa alguns dispositivos mais conhecidos e relevantes considerados.

2.3.1 DISPOSITIVOS DA MARCA SCHNEIDER ELECTRIC

A Schneider Electric é uma referência mundial no fabrico de equipamentos de monitorização e controlo energético. A linha PowerLogic destaca-se pela elevada precisão, robustez e integração facilitada em sistemas industriais. Todos os modelos descritos nesta secção comunicam através do protocolo Modbus TCP/IP, permitindo a recolha remota de grandezas elétricas em tempo real [27].

2.3.1.1 Dispositivo PowerLogic PM710

Medidor trifásico adequado para instalações comerciais e industriais, com classe de precisão 1. Fornece medições de tensões por fase, correntes, potências ativas e reativas, energia acumulada e frequência [28], conforme ilustrado na Figura 5.



Figura 5 – Medidor PM710 [27]

2.3.1.2 Dispositivo PowerLogic PM5100

Versão com maior rigor de medição (classe 0.5S), ideal para aplicações críticas. Para além das variáveis elétricas convencionais, permite a deteção de distorções harmónicas e a configuração de alarmes internos [29]. O aspeto visual do equipamento é ilustrado na Figura 6, comum a toda a série PM5100.

2.3.1.3 Dispositivo PowerLogic PM5110

Evolução da série PM5000, o PM5110 mantém o aspeto visual do seu antecessor, o PM5100, facilitando a substituição ou integração em painéis existentes. Inclui funcionalidades adicionais como registo de eventos, medição bidirecional e uma interface de configuração mais intuitiva. É particularmente adequado para instalações com produção renovável, como sistemas fotovoltaicos [29].



Figura 6 – Medidor PM5100 [29]

2.3.1.4 Dispositivo PowerLogic PM3250

Modelo de gama intermédia, concebido para quadros gerais de edifícios e aplicações de monitorização energética convencionais. Mede potência total, corrente por fase, tensões e energia acumulada [30], como mostrado na Figura 7.



Figura 7 – Medidor PM3250 [30]

2.3.2 DISPOSITIVOS DA MARCA SHELLEY

A marca Shelly, pertencente à empresa Allterco Robotics, disponibiliza uma gama inovadora de medidores de energia elétrica concebidos para aplicações domésticas, comerciais e industriais. Estes dispositivos distinguem-se pela sua arquitetura eletrónica baseada no microcontrolador ESP32, um sistema altamente integrado que conjuga capacidade de processamento local, conectividade Wi-Fi e funcionalidades de automação, proporcionando uma solução compacta e eficiente [31].

Um dos principais destaques tecnológicos dos dispositivos Shelly é a presença de um servidor web embutido, acessível diretamente através do endereço IP local do equipamento. Esta interface permite ao utilizador configurar o dispositivo, visualizar dados energéticos em tempo real, definir limiares de alarme, realizar atualizações de *firmware* e efetuar diagnósticos básicos, tudo a partir de um navegador web, sem necessidade de software adicional. Esta abordagem baseada em *webserver* local permite uma gestão descentralizada, intuitiva e multiplataforma.

Ao nível da instalação, os medidores Shelly adotam uma metodologia não invasiva, baseada no uso de transformadores de corrente externos, o que facilita a integração em quadros elétricos existentes sem interrupção dos circuitos. Estes transformadores estão disponíveis em diferentes gamas de medição e são adaptáveis às características elétricas do ponto de instalação.

2.3.2.1 Dispositivo Shelly EM

O Shelly EM é um dispositivo concebido para medições monofásicas ou bifásicas, conforme ilustrado na Figura 8. Está equipado com dois canais independentes, o que permite a monitorização paralela de dois circuitos distintos, como consumo geral e produção fotovoltaica. A sua configuração simples e comunicação via Wi-Fi tornam-no especialmente indicado para aplicações residenciais ou pequenas instalações comerciais, onde se pretende obter uma monitorização precisa com baixa complexidade de implementação. Apresenta precisão de $\pm 2\%$ na tensão e na corrente [33].



Figura 8 – Medidor Shelly EM com transformador [32]

2.3.2.2 Dispositivo Shelly 3EM

O Shelly 3EM é um dispositivo desenvolvido para monitorização trifásica, como mostrado na Figura 9, com capacidade para medir simultaneamente corrente, tensão, potência e energia elétrica em cada uma das três fases. Para além da medição bidirecional — útil na presença de sistemas de produção renovável — permite ainda o registo de valores médios, mínimos e máximos, facilitando a análise detalhada dos consumos. A comunicação é realizada via Wi-Fi, assegurando a integração com plataformas locais ou remotas de gestão energética. Pela sua versatilidade e facilidade de instalação, é amplamente utilizado em edifícios de média dimensão, por exemplo, pequenas unidades industriais ou comerciais, bem como em instalações fotovoltaicas. Apresenta precisão de $\pm 1\%$ na tensão e na corrente [33].



Figura 9 – Medidor Shelly 3EM [33]

2.3.2.3 Dispositivo Shelly Pro 3EM

O Shelly Pro 3EM é a versão profissional e avançada do modelo trifásico, ilustrado na Figura 10, concebida para ambientes mais exigentes ao nível da fiabilidade e da segurança na comunicação. Para além da ligação Wi-Fi, integra conectividade Ethernet, que garante maior estabilidade na transmissão de dados e permite operação em redes com requisitos técnicos mais rigorosos. Suporta também protocolo *Modbus* TCP/IP, permitindo a sua integração em soluções de monitorização mais específicas. A sua construção robusta e capacidade de comunicação segura tornam-no especialmente indicado para instalações institucionais, comerciais ou industriais, onde a continuidade e a fiabilidade da medição são essenciais. Apresenta precisão de $\pm 1\%$ na tensão e na corrente [34].



Figura 10 – Medidor Shelly Pro 3EM [34]

2.3.3 OUTROS DISPOSITIVOS DE MEDIÇÃO

Para além dos dispositivos destinados à monitorização do consumo energético, existem também soluções tecnológicas focadas na produção fotovoltaica, nomeadamente inversores híbridos e equipamentos dedicados à monitorização de sistemas solares. Estes dispositivos não se limitam apenas à conversão de energia, desempenham igualmente um papel relevante na recolha e disponibilização de dados elétricos de geração solar, permitindo a sua integração em sistemas de supervisão e gestão energética.

2.3.3.1 TRV00210

O TRV00210 é uma interface Modbus da gama Enerlin'X que liga as unidades Micrologic dos disjuntores Schneider Electric à rede de monitorização. Não mede por si, apenas expõe via Modbus as grandezas e estados fornecidos pela Micrologic. Está descontinuado e o substituto é o IFM LV434000, mantendo exatamente a função na arquitetura. A classe de precisão depende da unidade Micrologic instalada. O disjuntor com módulo de medição está certificado segundo a IEC 61557-12 em Classe 1 para potência e energia, sendo essa a classe efetiva herdada pelo canal Modbus do TRV00210. Na Figura 11 está representada a interface.



Figura 11 – Interface Modbus TRV00210

2.3.3.2 Dispositivo Huawei SUN2000

O Huawei SUN2000 é um inversor híbrido com capacidade de comunicação via *Modbus* TCP e compatibilidade com API REST, permitindo a sua integração com plataformas de monitorização e sistemas de supervisão externa. Na figura 12 encontra-se representado o dispositivo. Disponibiliza parâmetros como a potência de produção, tensão das *strings* fotovoltaicas, corrente e eficiência de conversão, sendo essencial para a monitorização contínua do desempenho do sistema solar [35].



Figura 12 – Inversor Huawei Sun2000 [35]

2.3.3.3 Dispositivo SolarLog

O SolarLog é um equipamento dedicado à monitorização de instalações fotovoltaicas, com capacidade para agregar dados provenientes de múltiplos inversores e disponibilizá-los através de API, como ilustrado na Figura 13. Este dispositivo permite acompanhar em tempo real o desempenho global do sistema solar, identificando eventuais falhas ou desvios de rendimento, e é frequentemente utilizado em contextos com múltiplos pontos de geração distribuída [36].



Figura 13 – Equipamento de monitorização SolarLog [37]

Estes dispositivos integram-se facilmente em plataformas de supervisão e gestão energética, fornecendo dados fíáveis e em tempo real. A sua compatibilidade com normas industriais e a flexibilidade de aplicação tornam-nos especialmente adequados para ambientes técnicos exigentes.

2.4 GRANDEZAS ELÉTRICAS NA MONITORIZAÇÃO ENERGÉTICA

A monitorização energética exige medir e interpretar de forma contínua as grandezas que descrevem o comportamento elétrico das instalações.

A tensão é a diferença de potencial que faz mover as cargas e deve manter-se dentro dos limites definidos. Em redes trifásicas é importante verificar a simetria entre fases para identificar desequilíbrios que afetam a qualidade da alimentação e podem reduzir a vida útil dos equipamentos [38].

A corrente é o fluxo de carga nos condutores e está ligada à energia utilizada. A sua medição permite avaliar o carregamento dos circuitos, identificar sobrecargas e perdas por efeito Joule e confirmar o dimensionamento de cabos e proteções. Em sistemas trifásicos é ainda necessário acompanhar desequilíbrios entre fases para evitar aquecimentos anómalos e degradação de componentes [39].

A potência ativa corresponde à parte que realiza trabalho útil e suporta o cálculo de consumos e da faturação. A sua monitorização ajuda a identificar padrões de utilização, a gerir cargas e a avaliar a eficiência dos equipamentos. A potência reativa está associada à energia temporariamente armazenada em campos elétricos e magnéticos de cargas indutivas ou capacitivas. Quando é excessiva pode originar penalizações, por isso deve ser medida para aplicar compensação e melhorar o fator de potência. A combinação das componentes ativa e reativa define a potência aparente, que expressa a exigência total colocada ao sistema e é usada no dimensionamento de transformadores, condutores e aparelhos de corte. O fator de potência resulta da razão entre ativa e aparente e indica a eficiência de utilização da energia. Valores baixos justificam correção com baterias de condensadores ou sistemas automáticos [40] [41]. A frequência indica o número de ciclos por segundo e deve manter-se estável. Em Portugal o valor típico é cinquenta hertz. Desvios de frequência costumam refletir desequilíbrios entre produção e consumo. Em sistemas isolados ou com produção renovável a sua monitorização é determinante para garantir estabilidade e evitar desconexões ou falhas de equipamentos [38] [42].

A análise de valores mínimos e máximos ao longo do tempo ajuda a detetar picos de consumo, quedas momentâneas de tensão e eventos transitórios. Esta informação é útil para diagnosticar falhas intermitentes, proteger equipamentos e planear ações preventivas [41].

A qualidade de energia é ainda avaliada pela distorção harmónica total em tensão e em corrente. Esta distorção resulta com frequência de cargas não lineares como fontes comutadas, variadores de velocidade e iluminação LED e pode causar aquecimentos e mau funcionamento

de equipamentos sensíveis. Em redes trifásicas deve também vigiar-se o desequilíbrio de fase expresso em percentagem, mantendo-o dentro de limites recomendados para evitar perdas adicionais e redução da eficiência de motores [43] [41].

Estas grandezas são os principais indicadores operacionais usados na aquisição, análise e gestão de energia. A sua interpretação cuidada permite identificar ineficiências, antecipar falhas e orientar estratégias de otimização contínua. Nem todas precisam de ser guardadas de forma permanente, já que algumas podem ser extraídas da análise de séries temporais, como por exemplo, os mínimos e máximos. A decisão sobre o que medir e armazenar deve ajustar-se à realidade de cada instalação, tendo em conta os objetivos da monitorização, os recursos disponíveis e a utilidade analítica de cada variável.

2.5 PLATAFORMAS DE MONITORIZAÇÃO ENERGÉTICA

A escolha da plataforma de monitorização energética adequada depende de vários fatores, tais como a facilidade de integração com sensores e bases de dados, capacidades de análise e visualização, compatibilidade com protocolos de comunicação e requisitos de infraestrutura. Atualmente, várias soluções estão disponíveis, tanto de código aberto como comerciais, cada uma oferecendo vantagens específicas consoante as necessidades dos sistemas.

Este capítulo analisa algumas das principais plataformas utilizadas para a monitorização e análise de consumo energético, como o EmonCMS, ThingSpeak, Arduino IoT Cloud e Home Assistant. Além disso, destaca ferramentas de visualização como o Grafana, que podem ser integradas a estas plataformas para criar *dashboards* interativos e personalizáveis.

2.5.1 EMONCMS – PLATAFORMA *OPEN-SOURCE* PARA MONITORIZAÇÃO ENERGÉTICA

O EmonCMS (*Energy Monitoring Content Management System*) é uma plataforma *open-source* desenvolvida no âmbito do projeto OpenEnergyMonitor [44], concebida especificamente para a monitorização e visualização de consumos energéticos. A sua natureza modular, flexível e gratuita torna-a particularmente atrativa para ambientes académicos e profissionais, permitindo a recolha, armazenamento e análise de dados energéticos em tempo real.

Esta plataforma pode ser instalada tanto em servidores locais como em dispositivos embebidos com recursos limitados, como o Raspberry Pi, o que a torna altamente versátil em contextos

de investigação, sistemas de automação doméstica ou infraestruturas energéticas de larga escala [44].

O EmonCMS baseia-se numa estrutura modular que inclui um núcleo e vários módulos opcionais que podem ser instalados consoante as funcionalidades pretendidas. A recolha dos dados pode ser realizada através de protocolos como MODBUS/TCP/IP, MQTT (*Message Queuing Telemetry Transport*) e HTTP (REST API), sendo posteriormente processados em inputs, que podem ser vistas como variáveis no estado bruto, podendo depois ser manipulados, transformados e encaminhados para *feeds*. Os dados resultantes deste processo encontram-se estruturados e organizados cronologicamente, estando preparados para armazenamento e posterior exploração. A plataforma disponibiliza dashboards personalizáveis, nos quais os feeds podem ser representados sob a forma de gráficos, widgets, indicadores ou imagens, potenciando a análise e a interpretação visual da informação [44].

A instalação do EmonCMS localmente num servidor Linux requer a presença de um conjunto de dependências essenciais, nomeadamente:

- Servidor Web – Responsável por disponibilizar a interface da plataforma e processar os ficheiros PHP (*Hypertext Preprocessor*) associados. A configuração mais comum recorre ao Apache2, dado o seu suporte nativo a `.htaccess` e a ampla compatibilidade com aplicações PHP.
- Sistema de gestão de bases de dados relacionais – Pode ser utilizado MySQL ou MariaDB, com a função de armazenar e gerir toda a estrutura lógica da aplicação.
- Ambiente de execução PHP, devidamente configurado com as extensões:
 - php-mysql – Necessária para a comunicação com a base de dados MySQL/MariaDB;
 - php-mbstring – Suporte a codificações *multibyte*, indispensável ao correto funcionamento da interface gráfica;
 - php-curl – Utilizada para efetuar operações de comunicação por HTTP com API externa ou interna.
- Serviço Redis – Recomendado para suporte a operações assíncronas e gestão eficiente da *cache* de dados, otimizando o desempenho geral da plataforma [43].

Para além da instalação das dependências, é importante compreender o papel funcional da base de dados relacional usada pelo EmonCMS, que corre sob MySQL ou MariaDB. Embora esta não seja utilizada para armazenar os dados de medição propriamente ditos, é responsável por suportar toda a estrutura lógica da aplicação. A base de dados organiza e relaciona

entidades fundamentais como utilizadores, *inputs*, *feeds*, *dashboards*, configurações do sistema, permissões de acesso e registos de eventos. Cada um destes elementos é representado em tabelas estruturadas, permitindo ao EmonCMS gerir o seu funcionamento interno de forma modular e eficiente [44].

O envio de dados para o EmonCMS é realizado através de pedidos HTTP simples, em que os dispositivos transmitem medições utilizando a API da plataforma. Os dados são normalmente enviados em formato JSON, juntamente com a identificação do nó correspondente e uma chave de autenticação designada por *API key*. Esta chave é única para cada utilizador e serve para validar o acesso ao sistema, garantindo que apenas fontes autorizadas possam enviar ou consultar dados. Um exemplo genérico de envio é:

```
http://<endereço_do_servidor>/input/post?node=nome_do_nó&json={variavel: valor}&apikey=sua_apikey
```

Após receção, os dados aparecem na secção *inputs* da interface gráfica. Para que sejam armazenados de forma persistente, é necessário associar esses *inputs* a *feeds* através da funcionalidade *log to feed*. É ainda possível aplicar operações adicionais no momento do registo, como acumulações, conversões ou cálculos, ajustando o tratamento dos dados de acordo com os objetivos do sistema de monitorização [44].

No que respeita à gestão da informação, o EmonCMS adota uma arquitetura híbrida. A base de dados MySQL é utilizada apenas para armazenar a estrutura interna da plataforma, incluindo configurações de utilizadores, *inputs*, *feeds* e *dashboards*. Já os dados históricos de séries temporais, pelo seu elevado volume e necessidade de escrita contínua, são guardados em ficheiros binários localizados em diretórios específicos do sistema.

O formato mais comum é o PHPFina, um método eficiente de armazenamento em intervalos regulares, ideal para séries temporais com cadência fixa. Este formato é altamente otimizado tanto em termos de espaço em disco como de desempenho nas operações de leitura e escrita.

Adicionalmente, o sistema suporta outros dois formatos:

- PHPFiwa, que permite a gestão de dados com resolução variável e cálculo de médias ao longo do tempo, embora atualmente se encontre obsoleto e em desuso nas versões mais recentes da plataforma.
- PHPTimeSeries, que é adequado para eventos com *timestamps* irregulares, como medições não periódicas [44].

Os dados históricos gerados pela monitorização são organizados em subdiretórios específicos localizados em `/var/opt/emoncms/`, nomeadamente: `phpfina/`, `phpfiwa/`, `phptimeseries/` e `backup/`. Cada um destes diretórios corresponde a um formato de armazenamento distinto, utilizado pelo EmonCMS para guardar séries temporais de forma eficiente. O diretório `backup/` é reservado para operações de salvaguarda, assegurando a preservação dos dados estruturais da plataforma.

Complementarmente, a estrutura principal da aplicação e os seus componentes modulares encontram-se no diretório `/opt/emoncms/`, que alberga o código-fonte do núcleo da aplicação web, bem como os módulos adicionais instalados. Estes módulos permitem estender a funcionalidade da plataforma, cobrindo áreas como visualização, transformação de dados, controlo de cargas, sincronização entre servidores e gestão de configuração [44].

A arquitetura modular do EmonCMS permite uma elevada adaptabilidade às necessidades do utilizador. Entre os módulos mais relevantes destacam-se:

- Graph e Dashboard, que permitem criar gráficos e painéis de visualização interativos;
- Apps, que disponibiliza aplicações analíticas pré-configuradas para análise de consumo, autoconsumo e desempenho energético;
- Postprocess, utilizado para aplicar cálculos retroativos sobre dados armazenados;
- Sync, concebido para sincronizar dados entre várias instâncias do EmonCMS. Esta funcionalidade é particularmente útil em sistemas com servidores a operar em paralelo ou em redundância, permitindo replicar automaticamente *feeds*, *inputs* e configurações entre plataformas;
- DemandShaper, que permite programar o acionamento de cargas com base em perfis tarifários, previsão de produção renovável ou horários definidos;
- Network, que disponibiliza uma visualização do estado de rede dos dispositivos conectados;
- Usefulscripts, que fornece um conjunto de scripts auxiliares para manutenção e diagnóstico;
- Backup, que oferece uma interface para exportar e importar configurações da plataforma [43]. Dentro deste módulo, encontram-se os diretórios `uploads/`, onde são colocados os ficheiros em formato `.tar.gz` para importação de instâncias anteriores, e `import/`, que contém os dados em fase de processamento. A importação pode ser feita tanto via interface gráfica como através do terminal, utilizando o script `emoncms-import.sh`, localizado habitualmente em `/opt/emoncms/modules/backup/`.

Para além destas funcionalidades, o EmonCMS permite também a personalização visual dos *dashboards* através da incorporação de imagens estáticas. Esta funcionalidade é útil para enriquecer a apresentação dos dados com esquemas elétricos, logótipos institucionais ou representações gráficas de edifícios e fluxos energéticos. As imagens devem ser colocadas manualmente no diretório `/var/www/emoncms/Modules/dashboard/Views/images`, tornando-se acessíveis à interface web da plataforma. Para que a imagem seja apresentada num *dashboard*, deve ser utilizado um bloco de texto (por exemplo, com o widget `Heading`), no qual se insere o URL da imagem correspondente ao seu caminho relativo. Por exemplo, para uma imagem denominada `image.png`, o campo de texto deverá conter o endereço `/Modules/dashboard/Views/images/image.png`. Ao carregar o *dashboard*, a imagem é automaticamente renderizada no local definido. Esta abordagem oferece elevada flexibilidade na criação de interfaces visuais adaptadas aos objetivos de cada projeto, permitindo representar zonas específicas, fluxos energéticos ou diagramas técnicos associados aos dados recolhidos em tempo real [44].

Para além das funcionalidades disponibilizadas através da interface gráfica, o EmonCMS permite também um nível adicional de personalização através da edição direta do ficheiro de configuração principal da aplicação.

A personalização das configurações globais da plataforma EmonCMS é realizada através do ficheiro `settings.ini`, localizado tipicamente em `/opt/emoncms/settings.ini` ou, em algumas distribuições, em `/var/www/emoncms/settings.ini`. Este ficheiro define diversos parâmetros de funcionamento do sistema e permite adaptar a plataforma a diferentes contextos de utilização.

Por exemplo, para permitir o registo de novos utilizadores diretamente através da interface web, deve ser incluída a linha `multiuser = true`. Esta linha deve ser colocada na secção `[interface]` do ficheiro `settings.ini`, respeitando a estrutura do mesmo. Após esta alteração, a página de login da plataforma passa a apresentar a opção registar novo utilizador, permitindo a criação de contas adicionais sem intervenção do administrador. Esta funcionalidade é particularmente útil em ambientes colaborativos, com múltiplos utilizadores a aceder à plataforma.

Também é possível ajustar o número máximo de *feeds* permitidos por utilizador. A linha `max_feeds = (limite pretendido)` pode ser adicionada na secção `[limits]` do ficheiro, permitindo aumentar ou diminuir a quantidade de *feeds* de cada utilizador. Esta configuração revela-se particularmente vantajosa em contextos onde o número total de

variáveis a monitorizar, distribuídas por vários edifícios, excede o limite predefinido de 32 *feeds* por utilizador na plataforma [44].

Após qualquer alteração ao ficheiro `settings.ini`, é necessário reiniciar o servidor web para que as novas configurações sejam carregadas. Esta operação é realizada através do terminal Linux com o comando `sudo systemctl restart apache2`. Esta etapa é essencial para garantir que a aplicação carrega corretamente as configurações atualizadas.

Outro componente relevante no ecossistema OpenEnergyMonitor é o EmonHub, utilizado sobretudo em configurações com sensores remotos ou dispositivos com comunicação série, como os EmonTx. Este serviço intermedeia a comunicação entre os dispositivos locais e o EmonCMS, recolhendo dados através de portas seriais ou rádio (RF), processando-os, validando-os e encaminhando-os para o servidor por HTTP ou MQTT. Embora não seja essencial em todas as implementações, o EmonHub é particularmente útil em sistemas descentralizados ou em contextos onde se pretende integrar sensores de baixo custo [44].

Vários estudos e aplicações têm reforçado o valor do EmonCMS, por exemplo, tem sido usado em conjunto com ferramentas de automação como o Node-RED, criando ecossistemas integrados para gestão energética [45]. Em contexto académico, projetos de monitorização demonstraram que a plataforma é adequada para recolha distribuída de dados e integração em soluções de *smart grids*, evidenciando a sua flexibilidade e relevância em investigação científica [46].

Em síntese, o EmonCMS constitui uma plataforma madura, extensível e bem documentada, com suporte ativo da comunidade open-source. A sua arquitetura modular, aliada à flexibilidade de integração com diferentes protocolos de comunicação e formatos de armazenamento, faz dele uma ferramenta particularmente adequada para projetos de monitorização energética em larga escala, como o caso da Universidade do Algarve. A possibilidade de instalação local, sem dependência de serviços externos, garante ainda a soberania dos dados e a personalização total da infraestrutura de monitorização [44]. A Figura 14 apresenta um exemplo de *dashboard* criado no EmonCMS, onde se combinam diferentes *widgets* para visualização de variáveis energéticas, ambientais e indicadores de desempenho, ilustrando a flexibilidade gráfica da plataforma na construção de interfaces adaptadas a múltiplos contextos de monitorização [44].

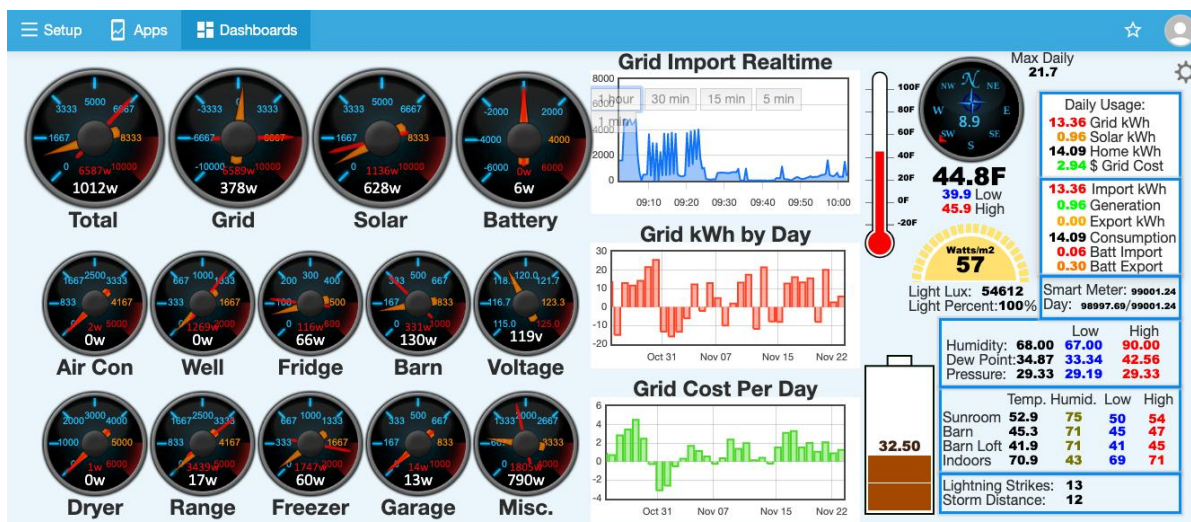


Figura 14 – Exemplo de dashboard no EmonCMS [44]

2.5.2 GRAFANA – FERRAMENTA AVANÇADA DE VISUALIZAÇÃO DE DADOS

O Grafana é uma plataforma amplamente utilizada para visualização e análise avançada de dados, com suporte para múltiplas fontes, incluindo bases de dados relacionais (como PostgreSQL e MySQL), bases de dados temporais, e API REST. Esta compatibilidade torna-o altamente flexível e adaptável a diversas aplicações, desde infraestruturas tecnológicas até sistemas industriais e de monitorização energética [47].

Embora não tenha sido desenvolvido especificamente para monitorização de energia, o Grafana destaca-se pela capacidade de criar *dashboards* interativos e altamente personalizáveis, o que o torna uma escolha popular para análise de dados em tempo real e históricos. A ferramenta permite consolidar informação proveniente de diferentes origens e representá-la através de gráficos dinâmicos e alertas configuráveis.

Além disso, o Grafana integra-se facilmente com bases de dados temporais especializadas, como o InfluxDB e o Prometheus, frequentemente utilizados para armazenar e consultar grandes volumes de séries temporais — os quais serão descritos em detalhe nas secções seguintes. Esta integração torna-o particularmente eficaz em sistemas que requerem análise histórica contínua e monitorização em tempo real.

O Grafana pode ser utilizado diretamente através da *cloud* oficial do Grafana Labs [48] ou instalado localmente num servidor, oferecendo flexibilidade consoante as exigências do projeto. Esta abordagem permite implementações em ambientes locais controlados ou em soluções centralizadas na nuvem, acessíveis remotamente. Importa, contudo, referir que o Grafana não recolhe dados diretamente dos sensores, dependendo de fontes externas para esse

efeito. Funciona como um complemento poderoso para análise e visualização, mas não substitui ferramentas especializadas na recolha e gestão de dados sensoriais, conforme ilustrado na Figura 15.



Figura 15 – Exemplo de dashboard no Grafana [49]

2.5.3 OUTRAS PLATAFORMAS IoT PARA MONITORIZAÇÃO ENERGÉTICA

Além do EmonCMS e do Grafana, existem outras plataformas IoT relevantes que oferecem suporte à recolha e monitorização de dados energéticos. Entre estas, destaca-se o ThingSpeak, uma plataforma baseada na *cloud* desenvolvida pela MathWorks, amplamente utilizada em cenários de monitorização remota e com capacidade de integração com MATLAB para análises avançadas [50]. A Arduino IoT Cloud que foi concebida para dispositivos Arduino e sensores compatíveis, fornecendo suporte nativo a API REST e ferramentas intuitivas de recolha e visualização de dados [51]. O Home Assistant que constitui uma solução *open-source* orientada à automação residencial, que permite integrar dispositivos inteligentes, medidores de consumo e assistentes de voz como Alexa e Google Assistant [52]. Estas plataformas podem ser utilizadas de forma complementar ou em alternativa, dependendo dos requisitos técnicos e funcionais do sistema de monitorização energética.

2.6 BASES DE DADOS PARA ARMAZENAMENTO DE DADOS ENERGÉTICOS

A eficiência do armazenamento de dados é um fator crítico na monitorização energética, dada a elevada frequência de recolha e processamento de informações provenientes de sensores e medidores. Para responder a estes desafios, são utilizadas diferentes bases de dados, cada uma com características específicas que as tornam mais adequadas para determinados casos de uso. Entre as mais relevantes para este tipo de aplicação encontram-se o PostgreSQL, InfluxDB, TimescaleDB e SQLite.

2.6.1 POSTGRESQL – BASE DE DADOS RELACIONAL

O PostgreSQL é um dos sistemas de gestão de bases de dados relacionais mais utilizados em contextos de monitorização energética, graças à sua robustez, desempenho e elevada conformidade com os padrões SQL (*Structured Query Language*). Suporta transações ACID (Atomicidade, Consistência, Isolamento e Durabilidade), assegurando integridade e fiabilidade dos dados mesmo em ambientes com múltiplas escritas simultâneas [53].

Entre as suas características mais relevantes, destaca-se o suporte nativo a tipo de dados JSON, permitindo o armazenamento eficiente de dados semiestruturados, como meta informação de sensores ou configurações específicas. Além disso, o PostgreSQL oferece indexação avançada, gestão de concorrência através de MVCC (*Multi-Version Concurrency Control*) e ampla compatibilidade com ferramentas de análise de dados, o que o torna particularmente adequado para sistemas de recolha e análise em tempo real.

Embora o MySQL continue a ser uma opção amplamente utilizada, especialmente pela sua simplicidade e desempenho em aplicações web, o PostgreSQL é frequentemente preferido em contextos mais exigentes, devido à sua maior flexibilidade, extensibilidade e capacidade de execução de consultas complexas. Enquanto o MySQL se foca na eficiência e facilidade de gestão, o PostgreSQL permite a personalização do sistema através de extensões como o TimescaleDB, concebido especificamente para otimização de séries temporais.

Esta combinação de desempenho, estrutura relacional avançada e suporte nativo a dados semiestruturados faz do PostgreSQL uma das soluções mais completas e escaláveis para sistemas de monitorização energética em grande escala [53].

Para facilitar a administração e consulta de dados armazenados no PostgreSQL, são frequentemente utilizadas ferramentas de gestão de bases de dados. O Valentina Studio é uma dessas soluções, oferecendo uma interface gráfica intuitiva para visualização, edição e análise rápida de dados. No contexto da monitorização energética, esta ferramenta permite executar consultas SQL de forma eficiente, visualizar tabelas e estruturar bases de dados sem necessidade de recorrer a comandos exclusivamente em linha de terminal. A sua compatibilidade com diferentes tipos de bases de dados torna-a um recurso valioso para a gestão e verificação rápida da integridade dos dados recolhidos pelos sensores, armazenados na base de dados ao longo do decorrer do projeto [54].

A Figura 16 ilustra um exemplo de utilização do Valentina Studio para a gestão de uma base de dados PostgreSQL, onde é possível visualizar a estrutura das tabelas e os campos armazenados. Esta interface simplifica a navegação e edição dos dados, permitindo uma administração mais eficiente e acessível da informação armazenada.

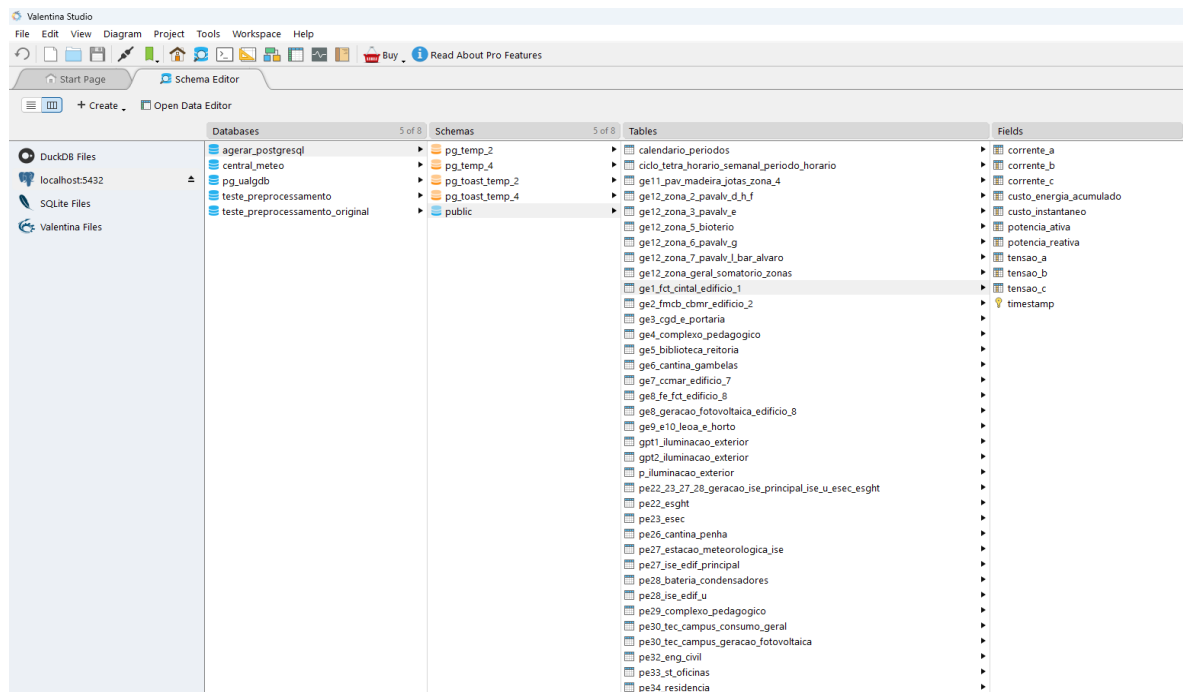


Figura 16 – Interface Valentina Studio

2.6.2 INFLUXDB – BASE DE DADOS TEMPORAL PARA DADOS ENERGÉTICOS

O InfluxDB é uma base de dados otimizada para séries temporais, sendo amplamente utilizada para armazenar medições energéticas com elevada granularidade. A sua arquitetura foi desenvolvida especificamente para lidar com grandes volumes de dados de sensores em tempo real, proporcionando consultas altamente eficientes e baixo consumo de recursos [55].

Entre as principais vantagens do InfluxDB encontram-se:

- A alta taxa de ingestão de dados, permitindo registos contínuos de medições energéticas com baixa latência.
- A otimização para séries temporais, oferecendo um modelo de armazenamento eficiente que reduz o espaço necessário para arquivar grandes quantidades de dados.
- Consulta de dados rápida e eficiente, utilizando um motor de processamento desenhado para análise de séries temporais, permitindo a visualização de padrões de consumo energético e deteção de anomalias em tempo real.
- Integração nativa com ferramentas de monitorização, como o Grafana, facilitando a análise e visualização dos dados [56].

No entanto, o InfluxDB não oferece as mesmas garantias transacionais que bases de dados relacionais como o PostgreSQL. Por este motivo, é frequentemente utilizado para análise de dados recentes e em tempo real, enquanto bases de dados relacionais continuam a ser a escolha ideal para armazenamento de longo prazo e integração com outros sistemas [57].

2.6.3 TIMESCALEDB – EXPANSÃO DO POSTGRESQL PARA SÉRIES TEMPORAIS

O TimescaleDB é uma extensão do PostgreSQL desenvolvida especificamente para lidar com séries temporais. Esta extensão adiciona funcionalidades avançadas para o armazenamento e análise de medições contínuas, tornando-a numa solução robusta para aplicações de monitorização energética de grande escala [58].

Entre as principais características do TimescaleDB destacam-se:

- O particionamento automático em *hypertables*, melhora significativamente o desempenho ao armazenar e consultar grandes volumes de dados temporais.
- A compressão eficiente de dados, reduzindo o espaço de armazenamento necessário sem comprometer a velocidade de acesso.
- A compatibilidade total com PostgreSQL, permitindo integração com sistemas já baseados nesta tecnologia, facilitando migração e implementação.

- O desempenho otimizado para consultas em séries temporais, tornando-o ideal para análises históricas, previsões energéticas e detecção de padrões de consumo.

O TimescaleDB é vantajoso para sistemas que necessitam de armazenamento estruturado, alta escalabilidade e consultas eficientes sobre grandes volumes de dados históricos. Desta forma representa uma solução intermédia entre o PostgreSQL puro e bases de dados especializadas em séries temporais, como o InfluxDB.

2.6.4 SQLITE

O SQLite é um sistema de gestão de bases de dados leve e incorporável, frequentemente utilizado em aplicações locais e sistemas embebidos. A sua principal vantagem é a simplicidade, dispensando a necessidade de configuração de um servidor de base de dados e permitindo armazenar informações num único ficheiro [59]. No entanto, no contexto da monitorização energética, o SQLite apresenta várias limitações que justificam a sua substituição por soluções mais robustas, como o PostgreSQL:

- Escalabilidade limitada – O SQLite não é adequado para gerir grandes volumes de dados provenientes de sensores de monitorização energética, uma vez que não suporta múltiplos acessos concorrentes de forma eficiente.
- Ausência de otimização para séries temporais – Diferente do InfluxDB e do TimescaleDB, o SQLite não possui funcionalidades específicas para processamento de séries temporais, o que compromete a análise de padrões de consumo energético.
- Menor desempenho em consultas complexas – Em sistemas que requerem análises avançadas e cruzamento de grandes volumes de dados, o SQLite tende a ter um desempenho inferior comparado com bases de dados relacionais robustas, como o PostgreSQL.

Dada a necessidade de um sistema mais eficiente, escalável e otimizado para análise energética, o SQLite foi descartado no projeto atual em favor do PostgreSQL, que oferece maior desempenho, integridade transacional e suporte para integrações avançadas.

2.7 SISTEMAS OPERATIVOS E INFRAESTRUTURA NECESSÁRIA PARA IMPLEMENTAÇÃO

A escolha do sistema operativo e da infraestrutura necessária desempenha um papel crucial na implementação de sistemas de monitorização energética. Estes sistemas requerem estabilidade, eficiência no processamento de dados e capacidade de automação para garantir a

recolha contínua e a análise de grandes volumes de informação proveniente de sensores e medidores. Para tal, o Linux Ubuntu tem-se demonstrado uma escolha altamente fiável, tendo sido usado neste projeto desde a versão 16.04 LTS, até à versão 24.04 LTS.

2.7.1 LINUX UBUNTU E A SUA APLICABILIDADE NA MONITORIZAÇÃO ENERGÉTICA

O Ubuntu é amplamente utilizado na implementação de plataformas de monitorização energética devido à sua estabilidade, segurança e compatibilidade com softwares *open-source* essenciais, como PostgreSQL, EmonCMS e Grafana [60]. Além disso, permite ajustes específicos que melhoram a eficiência dos sistemas de monitorização, como a otimização do consumo energético através de ferramentas especializadas, a configuração avançada de serviços automatizados e a personalização de permissões e segurança para garantir a integridade dos dados.

As versões do Ubuntu LTS (*Long Term Support*) apresentam vantagens significativas, ao garantirem suporte prolongado e atualizações de segurança, promovendo a longevidade e fiabilidade do sistema. As atualizações nestas versões priorizam a compatibilidade e segurança, garantindo que não introduzem mudanças disruptivas, tornando-as ideais para infraestruturas [61].

2.7.2 GESTÃO DE SERVIÇOS E PROCESSOS NO UBUNTU

Nos sistemas de monitorização energética, a correta gestão dos serviços é essencial para garantir que os processos de recolha, armazenamento e visualização dos dados funcionam de forma interrupta. O Ubuntu utiliza o Systemd, um sistema de inicialização e gestão de serviços que permite configurar processos críticos para serem executados automaticamente, assegurando que componentes essenciais, como a base de dados (e.g., PostgreSQL), servidor web (e.g., Apache2, sistemas de cache (e.g., Redis) e outras aplicações necessárias para o funcionamento de plataformas de recolha de dados, estejam sempre disponíveis [62].

A utilização de serviços automatizados garante que falhas inesperadas não resultem em perdas de dados, permitindo reinícios automáticos dos processos em caso de erro. Além disso, o Systemd permite o agendamento, monitorização e gestão de serviços, possibilitando a configuração de aplicações para que sejam iniciadas automaticamente no arranque da máquina. Esta funcionalidade garante que os processos de recolha de dados sejam ativados conforme necessário, assegurando a continuidade da monitorização sem falhas.

Por fim, a integração do Systemd com ferramentas de monitorização de *logs* de diagnóstico permite obter uma análise do desempenho dos serviços, ajudando a identificar anomalias no sistema e a otimizar a sua operação.

2.7.3 ADMINISTRAÇÃO E TRANSFERÊNCIA DE FICHEIROS EM AMBIENTE LINUX

A operação e manutenção de sistemas de monitorização como o EmonCMS, quando instalados em servidores com sistemas baseados em Linux, requer frequentemente a utilização de ferramentas de linha de comandos para garantir segurança, controlo remoto e transferência eficiente de ficheiros. Entre os comandos mais relevantes neste contexto destacam-se o SSH (*Secure Shell*) e o SCP (*Secure Copy Protocol*) [63].

O comando SSH permite estabelecer uma ligação segura e encriptada entre dois sistemas através da rede, funcionando como uma ferramenta essencial para a administração remota de servidores e outros dispositivos. Através desta ligação, é possível aceder ao terminal da máquina remota, executar comandos, editar configurações ou monitorizar serviços, tudo com elevado nível de segurança.

Neste comando, o utilizador é o nome do *user* remoto e IP do servidor corresponde ao endereço IP da máquina a aceder.

O comando SCP baseia-se no mesmo protocolo seguro do SSH, sendo utilizado para transferência de ficheiros entre máquinas. É especialmente útil em cenários de migração de sistemas, sincronização de backups ou transferência de dados históricos.

Ambos os comandos são amplamente utilizados em ambientes técnicos e académicos pela sua fiabilidade, simplicidade e segurança, sendo ferramentas indispensáveis para administradores de sistemas e técnicos responsáveis por plataformas de monitorização energética baseadas em Linux.

2.7.4 AUTOMAÇÃO DE TAREFAS E MONITORIZAÇÃO CONTÍNUA

Para assegurar um funcionamento eficiente e autónomo dos sistemas de monitorização energética, é fundamental a implementação de mecanismos de automação que garantam a recolha, processamento e armazenamento de dados de forma sistemática. No Ubuntu, a automação de tarefas pode ser realizada através de ferramentas como o *crontab*, que permite agendar a execução periódica de scripts responsáveis pela recolha e processamento de dados. O *crontab* pode ser configurado para executar tarefas em qualquer periodicidade de tempo,

desde minutos até intervalos mais longos como semanas ou meses, conforme as necessidades do sistema, assegurando que a monitorização não depende de intervenções manuais e ocorre de forma contínua [64]. Estes scripts podem ser utilizados para diversas finalidades, como a comunicação entre o servidor e os medidores de energia para recolha de dados em intervalos regulares, a consulta e envio de medições em tempo real para análise, a deteção de falhas de comunicação entre sensores e servidor, o armazenamento dos dados recolhidos numa base de dados, a realização de cálculos necessários para derivar variáveis adicionais ou até a reconfiguração automática de dispositivos remotos em resposta a condições específicas do sistema.

2.7.5 SEGURANÇA E MANUTENÇÃO DA INFRAESTRUTURA

Sistemas de monitorização energética envolvem dados críticos que devem ser protegidos contra falhas e acessos não autorizados. O Ubuntu oferece diversas camadas de segurança, incluindo atualizações regulares, permissões avançadas de utilizador e suporte a *firewalls* e encriptação de dados. Manter o sistema atualizado é uma prática essencial para garantir que vulnerabilidades conhecidas são corrigidas e que o sistema por de trás da plataforma mantém um elevado nível de segurança.

A implementação de políticas de *backup* periódicas é essencial para assegurar a integridade dos dados armazenados e permitir uma rápida recuperação em caso de falha. Para garantir a continuidade operacional, este processo pode ser automatizado através de *scripts* programados, que realizam cópias dos dados recolhidos e da estrutura da base de dados em períodos específicos [65]. Estes *backups* podem ser configurados para incluir tanto os dados históricos armazenados como as tabelas, relações e configurações essenciais do sistema, garantindo uma recuperação eficiente e minimizando o risco de perda de informação. Além dos *backups* da base de dados, é igualmente fundamental realizar *backups* completos do servidor, garantindo que todas as configurações do sistema, serviços e aplicações críticas são preservadas. A realização regular de cópias do sistema permite que, em caso de falha grave, o servidor possa ser restaurado rapidamente para um estado funcional anterior, minimizando tempos de inatividade e prevenindo perdas de dados irreversíveis. Além disso, a manutenção preventiva do servidor, incluindo a verificação periódica da integridade da base de dados e a análise de registos de erro, contribui para a longevidade e fiabilidade do sistema.

2.7.6 LINGUAGEM DE PROGRAMAÇÃO PYTHON

A linguagem de programação Python tem assumido um papel central no desenvolvimento de sistemas de aquisição e análise de dados, especialmente no domínio da monitorização energética [66], [67]. A sua sintaxe clara, associada a uma vasta comunidade de utilizadores e ao suporte de um conjunto alargado de bibliotecas, torna-a particularmente adequada para tarefas de recolha, tratamento, armazenamento e visualização de dados provenientes de sensores. Em comparação com linguagens como C++ ou Java, o Python distingue-se pela simplicidade de implementação e pela legibilidade do código, o que permite acelerar o desenvolvimento de soluções técnicas em ambientes académicos e de investigação [66].

Os códigos programados em Python são executados de forma sequencial por um interpretador que traduz, em tempo real, o código-fonte em instruções compreendidas pelo processador. Este modelo contrasta com o das linguagens compiladas, que exigem uma conversão prévia para código executável. Apesar de interpretada, a linguagem pode tirar partido de bibliotecas com componentes otimizados em linguagens de baixo nível, o que permite ultrapassar algumas limitações de desempenho em operações computacionalmente exigentes [67].

As bibliotecas Python consistem em coleções de funções e módulos que ampliam as capacidades da linguagem, evitando a necessidade de reimplementação de funcionalidades comuns. A sua utilização permite aceder a métodos específicos para a comunicação com dispositivos, manipulação e análise de dados, ligação a bases de dados ou geração de relatórios. Cada biblioteca é geralmente concebida para resolver um problema específico, incorporando boas práticas e soluções otimizadas [66], [67].

No contexto da monitorização energética, são frequentemente utilizadas bibliotecas como `pyModbusTCP`, que permite a comunicação com dispositivos através do protocolo Modbus; `requests`, utilizada para o envio de dados por HTTP; `psycopg2`, que estabelece ligação com bases de dados PostgreSQL; `json`, para tratamento de dados estruturados; `numpy`, dedicada ao processamento numérico e `datetime`, para a gestão de datas e horas com precisão [66], [67], [68]. Estas bibliotecas são utilizadas de forma integrada num fluxo de trabalho que se estende desde a recolha de dados até ao seu armazenamento e análise.

Destaca-se ainda a biblioteca `struct`, fundamental para a interpretação correta dos dados binários recebidos de dispositivos *Modbus*. Em muitos casos, as medições recolhidas (como potência, corrente ou tensão) são transmitidas em blocos de dois registos *Modbus* de 16 bits, que juntos representam um número em ponto flutuante de 32 bits. A biblioteca `struct` permite, através das funções `pack` e `unpack`, realizar esta conversão de forma segura e

eficiente, garantindo que os valores interpretados correspondem às medições reais. Esta funcionalidade é essencial para sistemas que interagem com dispositivos industriais, onde os dados são transmitidos em formatos binários normalizados e requerem tradução para tipos de dados compreensíveis pelo sistema de monitorização [65].

Uma biblioteca adicional relevante neste contexto é a `emon-tools`, desenvolvida para interagir com a plataforma EmonCMS. Esta biblioteca disponibiliza funcionalidades para automatizar tarefas como o envio de dados, a gestão de `feeds` e a leitura de medições. Para além disso, permite interpretar os ficheiros binários utilizados pela plataforma, geralmente nos formatos PHPFina, e convertê-los para estruturas tabulares, como CSV (*Comma-Separated Values*). O formato PHPFina recorre a registos de valor fixo armazenados como números em ponto flutuante em intervalos temporais regulares, definidos por um carimbo temporal inicial e um intervalo de amostragem constante. Esta organização permite armazenar séries temporais de forma compacta e eficiente, embora não seja diretamente legível por ferramentas de análise convencionais. A `emon-tools` interpreta a estrutura binária dos ficheiros `.dat` e `.meta`, reconstruindo a série temporal completa e permitindo a exportação dos dados para formatos mais versáteis, facilitando a integração com plataformas externas ou a realização de análises estatísticas offline [69].

A linguagem de programação Python, em articulação com um vasto ecossistema de bibliotecas — desde ferramentas especializadas para comunicação com dispositivos até módulos reutilizáveis para tratamento e análise de dados —, constitui uma base sólida para a construção de sistemas de monitorização energética robustos e escaláveis. Este ecossistema, mantido e continuamente expandido por uma comunidade ativa de programadores e investigadores, confere à linguagem uma versatilidade que permite adaptar a solução a diferentes necessidades técnicas, assegurando modularidade, clareza e facilidade de manutenção ao longo do tempo.

2.8 MÉTODOS DE CÁLCULO DE CUSTOS ENERGÉTICOS

A determinação dos custos energéticos é um fator fundamental na gestão eficiente do consumo de eletricidade. O método utilizado para calcular os custos associados à utilização da energia elétrica depende diretamente do tipo de tarifário contratado no local. Entre os diferentes modelos tarifários, alguns ajustam os preços ao longo do tempo, permitindo uma melhor adequação do consumo às condições do mercado e da rede elétrica [70].

2.8.1.1 Estrutura dos Tarifários Elétricos

Os tarifários de eletricidade podem ser classificados em dois grupos:

- Tarifário fixo: Mantem um preço constante por kWh independentemente da hora do dia ou do nível de consumo.
- Tarifários dinâmicos: Variam ao longo do tempo, refletindo fatores como horários de pico, oferta de energia renovável e variações de mercado.

Entre os tarifários dinâmicos mais utilizados destacam-se [70]:

1. Tarifário Bi-Horário – Divide o dia em dois períodos:
 - Vazio – preço reduzido, normalmente durante a noite e madrugada;
 - Fora do vazio – preço mais elevado, normalmente durante o dia.
2. Tarifário Tri-Horário – Adiciona um terceiro período:
 - Vazio – preço mais baixo, geralmente de madrugada;
 - Cheia – preço reduzido;
 - Ponta – preço mais elevado, correspondendo a períodos de maior procura.
3. Tarifário Tetra-Horário – Modelo mais detalhado utilizado em indústrias e grandes consumidores, com quatro períodos distintos [70]:
 - Super Vazio – tarifa mais reduzida, geralmente de madrugada, onde o consumo elétrico é mais baixo;
 - Vazio – tarifa baixa, geralmente aplicável durante a manhã e noite, quando o consumo ainda não atingiu os picos diários;
 - Cheia – tarifa intermédia, aplicada em períodos de maior procura energética ao longo do dia;
 - Ponta – tarifa mais elevada, refletindo os momentos de maior carga na rede elétrica, onde a procura atinge o seu pico.

A implementação destes tarifários permite que os utilizadores ajustem os seus padrões de utilização de energia, transferindo consumos para períodos de menor custo e contribuindo para a redução da fatura elétrica. A Figura 17 ilustra a estrutura do tarifário tetra-horário semanal aplicado em Portugal Continental, evidenciando a divisão horária ao longo da semana e as variações entre os períodos de inverno e verão. Esta estrutura foi publicada pela ERSE (Entidade Reguladora dos Serviços Energéticos) em 2023 e está atualmente em vigor porque não surgiram alterações nos anos seguintes [70]. Esta diferenciação possibilita uma gestão otimizada da energia, especialmente em contextos institucionais e empresariais, como na Universidade do Algarve, onde este tarifário está atualmente em vigor. A aplicação de estratégias de gestão adaptadas a esta estrutura contribui para uma utilização mais eficiente

dos recursos energéticos. Importa referir, contudo, que o preço por kWh em cada período não é fixo, sendo revisto anualmente com base nas dinâmicas do mercado e nas decisões das fornecedoras de energia, exigindo uma adaptação constante para garantir a sustentabilidade económica do sistema [71].

Ciclo BTE tetra-horário ciclo semanal para todos os fornecimentos em Portugal Continental



Figura 17 – Gráfico representativo do Ciclo Tetra-horário semanal [70]

2.8.1.2 Cálculo do Custo Energético

Nos sistemas de monitorização energética, o custo da energia pode ser calculado em tempo real com base na potência instantânea medida e no tarifário vigente no período da medição [72]. A equação que expressa este cálculo é dada por:

$$Ci = \frac{Pi \times \Delta t \times p(ti)}{60}; \quad (1)$$

onde:

- Ci = custo da medição em euros;
- Pi = potência ativa instantânea em kW na medição i ;
- Δt = duração do intervalo em minutos;
- $p(ti)$ = preço em €/kWh aplicável no instante ti em função do período horário aplicado;

Além do custo instantâneo, o cálculo do custo acumulado ao longo do tempo é essencial para a análise de faturas e previsão de gastos futuros. O custo total pode ser determinado através da seguinte expressão:

$$C_{total} = \sum_{t=1}^n \left[\frac{P(t) \times \Delta t \times p(t)}{60} \right]; \quad (2)$$

onde:

- C_{total} = custo total acumulado em euros;
- $P(t)$ = potência instantânea em kW;
- Δt = duração do intervalo em minutos;
- $p(t)$ = preço da energia em €/kWh

O processamento dos dados de consumo energético é realizado em intervalos de um minuto, assumindo que a potência medida no início de cada intervalo se mantém constante até à leitura seguinte. Desta forma, o sistema obtém uma aproximação ao perfil real de consumo a partir das sucessivas medições de potência

O cálculo rigoroso da energia consumida num intervalo de tempo $[t_0, t_f]$ exige a integração contínua da potência instantânea:

$$E = \int_{t_0}^{t_f} P(t) dt ; \quad (3)$$

Contudo, nos sistemas de monitorização com medições discretas, esta integral é aproximada por uma soma de Riemann:

$$E \approx \sum_{i=1}^n P(t_i) \times \Delta t ; \quad (4)$$

em que $P(t_i)$ representa a potência medida no instante t_i e Δt corresponde ao intervalo de amostragem [72]. Esta aproximação implica assumir que a potência se mantém constante durante cada intervalo, o que raramente corresponde à realidade. Assim, em determinados instantes, a potência real pode ser superior ou inferior ao valor considerado, originando um erro de aproximação no cálculo da energia desse intervalo.

No entanto, a cada nova medição, o valor atualizado de $P(t)$ pode introduzir um desvio oposto ao anterior, compensando parcialmente o erro acumulado. Embora este comportamento seja imprevisível para intervalos individuais, quando se consideram longas séries temporais o conjunto dos desvios tende a anular-se estatisticamente. Deste modo, a soma discreta resulta, em média, numa estimativa global do consumo energético bastante próxima do valor real, justificando a validade do método para análise de faturas e projeções de consumo.

2.9 MACHINE LEARNING APLICADO À PREVISÃO E DETECÇÃO DE ANOMALIAS NO CONSUMO ENERGÉTICO

A aplicação de técnicas de aprendizagem automática na gestão energética tem sido amplamente estudada, com resultados promissores na previsão de consumo e na deteção de comportamentos anómalos. Estudos recentes demonstram que algoritmos como Redes Neurais Artificiais (ANN), *Support Vector Machines* (SVM) e modelos baseados em séries temporais (LSTM, ARIMA) apresentam elevada precisão na previsão de consumo [73]. Entre os modelos de séries temporais mais utilizados destacam-se o ARIMA, conhecido pela sua capacidade de modelar padrões estatísticos, e o Prophet, desenvolvido pelo Facebook, que se adapta melhor a tendências sazonais e dados esparsos [74].

Além da previsão, a deteção de anomalias em padrões de consumo é uma ferramenta essencial para identificar falhas em medidores ou desperdícios energéticos. Algoritmos como o *Isolation Forest* [75], [76] e o *Local Outlier Factor* [77] são amplamente utilizados para reconhecer padrões atípicos, permitindo otimizações na gestão de energia.

A combinação de previsão e deteção de anomalias pode contribuir significativamente para a melhoria da eficiência energética e redução de custos operacionais, permitindo uma gestão mais proativa e reativa perante desvios inesperados no comportamento energético.

Em seguida, são aprofundados alguns dos algoritmos que suportam estas abordagens.

2.9.1 ARIMA (AUTOREGRESSIVE INTEGRATED MOVING AVERAGE)

O ARIMA é um modelo estatístico clássico para previsão de séries temporais, que combina três componentes principais [78]:

- AR (*AutoRegressive*) – Usa valores passados da série para prever valores futuros, assumindo que há uma relação linear entre os dados históricos e as previsões.
- I (*Integrated*) – Aplica uma diferenciação aos dados para remover tendências e tornar a série estacionária (sem crescimento ou declínio sistemático).
- MA (*Moving Average*) – Modela a relação entre erros passados e o valor atual da previsão, ajudando a capturar flutuações de curto prazo nos dados.

Deste modo, o ARIMA identifica padrões históricos de consumo e ajusta equações matemáticas para prever a procura futura. Estudos recentes confirmam a sua aplicação em diferentes contextos, como na previsão da produção de energia eólica na Polónia, onde foi

comparado com modelos SARIMA e de aprendizagem automática [79], ou na análise do consumo de edifícios públicos na China, evidenciando o seu desempenho face a abordagens baseadas em *machine learning* [73]. No entanto, o modelo assume que a série temporal é linear e estacionária, o que pode ser uma limitação quando há padrões sazonais recorrentes ou mudanças abruptas nos dados devido a fatores externos, como variações climáticas, políticas energéticas ou eventos inesperados [78].

Para lidar com variações sazonais bem definidas, utiliza-se uma extensão chamada SARIMA (Seasonal ARIMA), que inclui parâmetros adicionais para capturar padrões cíclicos que ocorrem em períodos fixos, como flutuações diárias, semanais ou sazonais no consumo energético [78].

2.9.2 PROPHET

O Prophet é um modelo de previsão de séries temporais desenvolvido pelo Facebook [80], concebido para lidar com grandes volumes de dados e capturar padrões complexos, como tendências não lineares e variações sazonais. Ao contrário de modelos clássicos como o ARIMA, que exigem séries estacionárias, o Prophet adapta-se bem a séries com componentes sazonais de curto e longo prazo, podendo ainda incorporar efeitos específicos como feriados e eventos extraordinários.

O funcionamento do modelo baseia-se na decomposição da série temporal em três componentes principais: a tendência (que pode ser linear ou logística), a sazonalidade (que pode incluir padrões diários, semanais e anuais) e os efeitos externos (tais como feriados ou eventos específicos que afetem o comportamento da variável em estudo). Esta estrutura modular permite ajustar o modelo de forma intuitiva e interpretar separadamente as influências de cada componente sobre a previsão final. Uma das grandes vantagens do Prophet é a sua robustez na presença de dados esparsos ou com lacunas temporais, sendo particularmente útil em domínios como a monitorização energética, onde podem existir períodos sem registos.

Estudos recentes evidenciam a eficácia do Prophet em aplicações reais de previsão de consumo e produção energética. Por exemplo, nos Estados Unidos, estudos com dados da *American Electric Power* mostraram que a afinação de parâmetros do Prophet pode reduzir o erro percentual absoluto médio significativamente, sobretudo nos períodos de treino, embora os ganhos nem sempre se mantenham nos dados de teste [81]. No Canadá, estudos comparativos entre o Prophet e modelos baseados em redes neuronais, como as LSTM, demonstraram que o Prophet apresenta boa capacidade para identificar

padrões sazonais e de tendência em séries temporais de elevada granularidade, embora os modelos híbridos ou neuronais revelem, em alguns casos, maior precisão na captura de flutuações rápidas e voláteis [83]. Além disso, o modelo suporta a inclusão de variáveis exógenas, ou seja, fatores externos ao sistema que influenciam diretamente a variável alvo, o que o torna flexível para aplicações mais realistas e complexas, como a previsão de consumo energético em sistemas sujeitos a variações dinâmicas [80].

O Prophet, contudo, apresenta também algumas limitações. Pode ter dificuldades em capturar alterações abruptas no comportamento dos dados, sobretudo quando essas mudanças não seguem padrões temporais consistentes. Nestes casos, o ajuste à realidade pode ser impreciso ou demorar vários ciclos de atualização. Apesar da sua facilidade de utilização e interpretação, o desempenho do Prophet tende a ser inferior ao de modelos mais sofisticados baseados em redes neuronais profundas ou em abordagens híbridas, especialmente quando a quantidade de dados históricos disponíveis é reduzida ou quando existe uma elevada variabilidade estrutural no sistema monitorizado [80].

2.9.3 ISOLATION FOREST

Além da previsão, a deteção de anomalias em padrões de consumo energético constitui uma ferramenta essencial para identificar falhas em medidores, desvios operacionais ou potenciais situações de desperdício. Entre os métodos mais utilizados para este fim destaca-se o algoritmo *Isolation Forest (IF)*, que se baseia em árvores de decisão para isolar rapidamente observações que se desviam da norma [75].

O funcionamento do *Isolation Forest* assenta na criação de múltiplas árvores de decisão construídas aleatoriamente, onde os dados são sucessivamente particionados em subconjuntos mais pequenos. O princípio subjacente é que observações anómalas podem ser isoladas com um número reduzido de divisões, ao contrário das observações normais, que tendem a exigir mais partições para serem separadas [75]. Assim, quanto menor o número médio de divisões necessárias para isolar uma determinada instância, maior a probabilidade de esta ser classificada como uma anomalia. Esta abordagem revela-se particularmente eficaz na deteção de padrões atípicos em grandes volumes de dados, mesmo em contextos com elevada dimensionalidade, sendo também computacionalmente eficiente para aplicações em tempo real.

Apesar destas vantagens, o *Isolation Forest* apresenta algumas limitações. O seu desempenho pode ser influenciado pela densidade local dos dados, podendo levar à classificação incorreta

de observações legítimas que simplesmente pertencem a regiões menos densas da distribuição. Esta sensibilidade pode resultar em falsos positivos, especialmente em séries com variabilidade natural acentuada. Por outro lado, o algoritmo não fornece explicações detalhadas sobre a natureza da anomalia detetada, o que implica a necessidade de análises complementares para identificar a causa subjacente do comportamento atípico [75].

A aplicação prática do *Isolation Forest* em contextos energéticos tem sido explorada em vários estudos recentes. Um exemplo relevante é o trabalho de Mendes et al. [76], desenvolvido na Universidade do Algarve, que analisou o consumo energético em unidades hoteleiras, comparando o desempenho do *Isolation Forest* com o de um *Autoencoder* variacional. Os resultados evidenciaram elevada eficácia do algoritmo na deteção de consumos anómalos em dados reais de monitorização. Outros trabalhos têm igualmente explorado o uso do *Isolation Forest* em dados de redes elétricas reais, como no estudo de Labura et al. [83], onde o algoritmo foi aplicado a dados de medidores inteligentes no domínio do tempo e da frequência para detetar tanto anomalias pontuais como contextuais em redes de baixa tensão. Este tipo de aplicação evidencia a versatilidade e a eficiência deste algoritmo em múltiplos domínios da gestão energética.

2.9.4 LOCAL OUTLIER FACTOR

O *Local Outlier Factor* (LOF) é um algoritmo de deteção de anomalias baseado na densidade local dos dados, que permite identificar pontos atípicos ao comparar a densidade de um ponto com a dos seus vizinhos mais próximos [76]. Este método fundamenta-se na ideia de que anomalias são observações cuja densidade local é significativamente inferior à densidade média dos dados circundantes.

A principal característica do LOF é a sua capacidade de medir a densidade relativa de uma observação em relação ao seu ambiente local, o que o torna particularmente eficaz em contextos onde a densidade dos dados varia de forma acentuada ao longo do espaço de características. Ao contrário de abordagens globais como o *Isolation Forest*, o LOF considera a estrutura local dos dados, permitindo uma análise mais sensível a padrões específicos de vizinhança [76]. Este comportamento é vantajoso em situações onde grupos distintos de pontos apresentam diferentes concentrações, pois permite identificar anomalias que seriam ignoradas por métodos que assumem uma distribuição uniforme.

No entanto, o desempenho do LOF depende fortemente da escolha do número de vizinhos utilizados no cálculo da densidade relativa. A definição inadequada deste parâmetro pode

comprometer a detecção eficaz de anomalias, resultando em classificações errôneas. Além disso, devido à sua natureza local e à complexidade dos cálculos envolvidos, o LOF tende a ser menos eficiente quando aplicado a grandes volumes de dados, sendo mais apropriado para conjuntos de dados de dimensão moderada e com uma estrutura relativamente bem definida [76].

A aplicação do algoritmo LOF à detecção de anomalias em dados de consumo energético tem sido explorada em diversos contextos reais. Um exemplo é o estudo de Zangrando et al. [84], desenvolvido em colaboração entre instituições de Itália e Espanha, que comparou o desempenho do LOF com outros algoritmos de detecção de anomalias em séries temporais de consumo energético. Os resultados evidenciaram que métodos baseados em densidade, como o LOF, podem superar abordagens neurais em determinados cenários, mostrando elevada eficácia na generalização entre diferentes equipamentos. Estes casos reforçam a aplicabilidade do LOF na análise de dados energéticos, contribuindo para a detecção precoce de falhas operacionais e desperdícios.

2.9.5 MÉTRICAS DE AVALIAÇÃO DA PREVISÃO

A documentação existente sobre previsão em séries temporais aplicada à energia converge para um quadro de avaliação com três eixos complementares. Métricas determinísticas ponto a ponto quantificam a discrepância entre valores observados e estimados, métricas escala-livre asseguram comparabilidade entre séries e horizontes distintos e métricas probabilísticas que se aplicam quando os modelos produzem quantis ou intervalos de previsão permitindo avaliar simultaneamente a calibração e a largura dos intervalos. Com este quadro é possível avaliar o desempenho dos algoritmos, comparar métodos em séries heterogêneas e mitigar o efeito de valores extremos.

Em primeiro lugar apresentam-se as métricas determinísticas, que quantificam o desvio ponto a ponto entre observações reais e previsões em unidades físicas da variável.

O **MAE** (*Mean Absolute Error*) [85] quantifica o desvio médio, em valor absoluto, entre as observações e as previsões, expresso nas mesmas unidades da variável. Para uma sequência de n pares $(y_t, \hat{y}_t)$ define-se:

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (5)$$

Valores mais baixos indicam melhor desempenho. Por usar desvios absolutos, a MAE é menos sensível a valores extremos do que o RMSE (*Root Mean Squared Error*).

O **RMSE** [85] mede a raiz da média dos erros ao quadrado, nas mesmas unidades da variável.

Para um conjunto de avaliação com n pares $(y_t, \hat{y}_t)$, define-se:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (6)$$

Onde em (5) e (6),

- y_t é o valor observado e $\hat{y}_t$ a previsão no instante t ;
- n é o número de pontos usados na avaliação.

Valores mais baixos indicam melhor desempenho. Por elevar os erros ao quadrado, o RMSE atribui maior peso a desvios de grande magnitude do que o MAE.

Em seguida, introduz-se uma métrica de escala livre que permite comparar desempenhos entre séries e horizontes distintos. O MASE (*Mean Absolute Scaled Error*) [85] divide o erro absoluto médio do modelo por um erro de referência calculado na própria série através de uma estratégia de persistência, usando o último valor observado ou, o valor observado m períodos antes, quando há sazonalidade. Define-se por:

$$MASE = \frac{\frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|}{\frac{1}{T-m} \sum_{t=m+1}^T |y_t - y_{t-m}|} \quad (7)$$

Onde em (7),

- y_t é o valor observado e $\hat{y}_t$ a previsão no instante t ;
- n é o número de pontos usados na avaliação;
- T é o número de pontos usados para calcular o erro de referência;
- m é o período sazonal, se não houver sazonalidade usa-se $m = 1$.

O **WAPE** (*Weighted Absolute Percentage Error*) [86] representa a razão entre a soma dos desvios absolutos e a soma dos valores observados, expressa em percentagem. Define-se por:

$$WAP E_{\%} = 100 \times \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{\sum_{t=1}^n |y_t|} \quad (8)$$

Onde em (8),

- y_t é o valor observado e $\hat{y}_t$ a previsão no instante t ;
- n é o número de pontos usados na avaliação.

Quando os modelos produzem quantis ou intervalos de previsão, a avaliação deve considerar propriedades probabilísticas. Para previsões quantílicas, a perda *pinball* para o quantil τ (nível de probabilidade τ) mede a qualidade do quantil previsto através de uma função de custo assimétrica:

$$L_{\tau}(y_t, \hat{q}_{\tau,t}) = \begin{cases} \tau (y_t - \hat{q}_{\tau,t}), & \text{se } y_t \geq \hat{q}_{\tau,t} \\ (\tau - 1) (y_t - \hat{q}_{\tau,t}), & \text{se } y_t < \hat{q}_{\tau,t} \end{cases} \quad (9)$$

Onde em (9),

- y_t é o valor observado no instante t ;
- $\hat{q}_{\tau,t}$ é o quantil previsto ao nível de probabilidade τ .

Quando os modelos fornecem intervalos de previsão, recorrem-se a duas medidas amplamente usadas neste tipo de modelos. A **coverage** estima a proporção de observações que ficam dentro do intervalo ao nível $1 - \alpha$ [87]:

$$\widehat{Coverage}_{1-\alpha} = \frac{1}{n} \sum_{t=1}^n 1 \{L_t \leq y_t \leq U_t\} \quad (10)$$

A largura média dos intervalos, designada **sharpness**, quantifica a sua amplitude para o mesmo nível [88]:

$$Sharpness_{1-\alpha} = \frac{1}{n} \sum_{t=1}^n (U_t - L_t) \quad (11)$$

Onde em (10) e (11),

- L_t e U_t (limites inferior e superior ao nível $1 - \alpha$);
- n é o número de pontos usados na avaliação.

2.9.6 MÉTRICAS DE AVALIAÇÃO EM DETECÇÃO NÃO SUPERVISIONADA

Em contexto não supervisionado não existem rótulos que permitam calcular métricas como precisão ou sensibilidade. A avaliação deve caracterizar a carga operacional de alarmes, verificar a estabilidade do detetor entre janelas temporais e identificar derivas na distribuição dos scores, tomando como base um limiar definido por critério que assegure uma taxa de alarmes coerente com os requisitos operacionais.

A **FAR** (*False Alarm Rate*) taxa de alarme por 24h mede a carga operacional média de alarmes quando se assume uma janela sem incidentes conhecidos [89]. Define-se por:

$$FAR_{24h} = \frac{N_{alarmes}}{T_{horas\ observadas}} \times 24h \quad (12)$$

A **ARL** (*Average Run Length*) representa o tempo médio até ocorrer um alarme quando não há anomalias, quanto maior melhor [89]. Define-se por:

$$ARL_0 = \mathbb{E}[\tau_{FA}] \quad (13)$$

Onde em (13),

- ARL_0 é o comprimento médio até falso alarme em regime sem anomalias;
- τ_{FA} é o tempo até ao primeiro falso alarme, medido em unidades coerentes com a aplicação.

O **PSI** (*Population Stability Index*) quantifica a mudança da distribuição dos *scores* entre uma janela de referência e a janela corrente. Valores mais altos indicam maior deriva. O índice é calculado como a soma, sobre *bins* (intervalos) comuns às duas janelas, das diferenças ponderadas entre as proporções observadas [90]. Define-se por:

$$PSI = \sum_{b=1}^B (p_b - q_b) \ln \left(\frac{p_b}{q_b} \right) \quad (14)$$

Onde em (14),

- B representa o número de *bins* (intervalos);
- p_b representa a proporção de *scores* no *bin* b na janela corrente;
- q_b representa a proporção de *scores* no *bin* b na janela de referência.

O **Quantil-alvo** representa o percentil superior da distribuição dos *scores* s usado para fixar o limiar τ , permitindo controlar diretamente a fração esperada de observações em alarme [91]. Define-se por:

$$\tau = Q_{1-\alpha}(s) \quad (15)$$

Onde em (15),

- s representa os scores de anomalia na janela de calibração;
- $Q_{1-\alpha}(s)$ representa o quantil $(1 - \alpha)$ de s ;
- α representa a fração alvo de observações acima do limiar (fração esperada de alarmes).

2.10 CONCLUSÃO CRÍTICA DA REVISÃO BIBLIOGRÁFICA

A revisão bibliográfica evidencia que a monitorização e otimização do consumo energético têm sido amplamente estudadas, com soluções como o EmonCMS a desempenharem um papel relevante na gestão e visualização de dados de consumo energético. No entanto, a integração de medidores de diferentes fabricantes continua a representar um desafio, especialmente devido à diversidade de protocolos de comunicação e padrões de dados adotados pelas empresas do setor. A continuidade da monitorização em caso de falhas na infraestrutura depende de soluções específicas, como redundância na recolha de dados e armazenamento distribuído.

Os métodos de cálculo de custos energéticos apresentam variações significativas conforme o tarifário adotado e a granularidade das medições disponíveis. Os tarifários dinâmicos têm sido amplamente recomendados para a otimização dos custos energéticos, permitindo que consumidores ajustem os seus padrões de consumo de acordo com períodos de menor custo. No entanto, a necessidade de atualização manual dos tarifários continua a ser um fator limitante em muitos sistemas, uma vez que os valores de energia variam periodicamente com base na distribuidora energética. A implementação de mecanismos automatizados de atualização dos preços, idealmente através da disponibilização de uma API que fornecesse os preços atualizados por parte das fornecedoras de energia, poderia representar um avanço significativo na eficiência dos sistemas de monitorização.

No domínio da previsão de consumos e deteção de anomalias, diversos algoritmos têm sido estudados, incluindo ARIMA, Prophet, *Isolation Forest* e LOF. Cada um destes modelos apresenta vantagens e limitações consoante a estrutura dos dados e os objetivos da análise. Enquanto o ARIMA e o Prophet são amplamente utilizados para previsão de consumo

energético, o *Isolation Forest* e o LOF são mais indicados para a detecção de padrões anómalos, podendo identificar falhas em medidores ou consumos atípicos que indicam desperdícios energéticos. A escolha do modelo mais adequado depende não apenas da qualidade dos dados recolhidos, mas também do contexto da aplicação, sendo essencial uma fase de validação para aferir a eficácia da abordagem escolhida.

Desta forma, a análise realizada permitiu compreender os avanços existentes nesta área, bem como os desafios técnicos e operacionais que ainda persistem. A necessidade de maior integração entre plataformas, a automatização de processos como a atualização dos tarifários e a adoção de técnicas de *machine learning* são alguns dos aspetos que podem contribuir para a melhoria da monitorização, do cálculo de custos e da otimização do consumo energético. O presente projeto assenta nestas premissas, explorando soluções que possibilitem maior precisão na análise energética e maior capacidade de adaptação a diferentes cenários de consumo.

3

DESENVOLVIMENTO DO TRABALHO REALIZADO

O presente projeto tem como objetivo a otimização e estabilização de um sistema de monitorização energética já existente na Universidade do Algarve. O sistema foi inicialmente desenvolvido para recolha e análise de consumos energéticos dos edifícios da universidade, utilizando comunicação via MODBUS/TCP/IP e armazenamento dos dados numa base de dados local.

Com o tempo, verificaram-se problemas de manutenção, limitações na escalabilidade e falhas na recolha de dados, evidenciando a necessidade de modernizar e melhorar a sua infraestrutura. Para solucionar estes problemas, foram implementadas melhorias que garantem maior fiabilidade, otimização da recolha, armazenamento e análise dos dados.

Este capítulo descreve as diferentes fases do desenvolvimento do projeto, desde a arquitetura inicial do sistema até à sua configuração atual, incluindo a reestruturação da recolha de dados, armazenamento e aumento de desempenho.

Este capítulo organiza-se em cinco secções que introduzem o sistema de forma progressiva. A Secção 3.1 apresenta a arquitetura inicial e a sua evolução até à configuração atual, destacando as opções técnicas que definiram a infraestrutura. A Secção 3.2 descreve a aquisição dos dados com leitura das variáveis energéticas dos medidores através de MODBUS sobre TCP/IP e de pedidos HTTP. A Secção 3.3 trata do processamento e da conversão dos dados com transformação dos valores recebidos para formatos utilizáveis, cálculo de custos e estruturação consistente das medições. A Secção 3.4 foca o armazenamento e o envio dos dados para o EmonCMS e para a base de dados PostgreSQL garantindo histórico organizado e pronto para análise. A Secção 3.5 apresenta os mecanismos de tolerância a falhas e as otimizações implementadas para assegurar fiabilidade e estabilidade do sistema.

3.1 ARQUITETURA INICIAL DO SISTEMA

A arquitetura inicial do sistema de monitorização energética da Universidade do Algarve foi desenvolvida em 2019, com o objetivo de implementar uma plataforma simples e eficaz para monitorizar os consumos energéticos através do protocolo Modbus/TCP/IP, integrando os medidores já instalados nos quadros gerais elétricos de grande parte dos edifícios da universidade. A solução foi concebida para funcionar de forma autónoma e fiável, recorrendo a software livre, *scripts* desenvolvidos em Python e servidores dedicados.

O sistema foi desenvolvido com base na infraestrutura de rede existente na Universidade do Algarve, a qual se encontra organizada, desde há vários anos, em múltiplas redes locais virtuais, VLANs (*Virtual Local Area Networks*). Estas VLANs, criadas sobre uma infraestrutura física comum, estão separadas por edifícios, departamentos ou funções institucionais. Embora esta segmentação seja essencial para a segurança e a gestão dos serviços, representa um desafio adicional para o sistema de monitorização, que precisa de aceder a medidores localizados em diferentes segmentos da rede.

A Figura 18 apresenta um esquema representativo e genérico desta organização em VLANs, ilustrando como diferentes segmentos da rede interna podem conter equipamentos de medição e como estes precisam de comunicar com os servidores centrais para garantir a recolha de dados. Embora não represente a configuração real da rede da Universidade, este exemplo permite compreender a lógica subjacente à infraestrutura existente e o contexto em que foi desenvolvido o sistema.

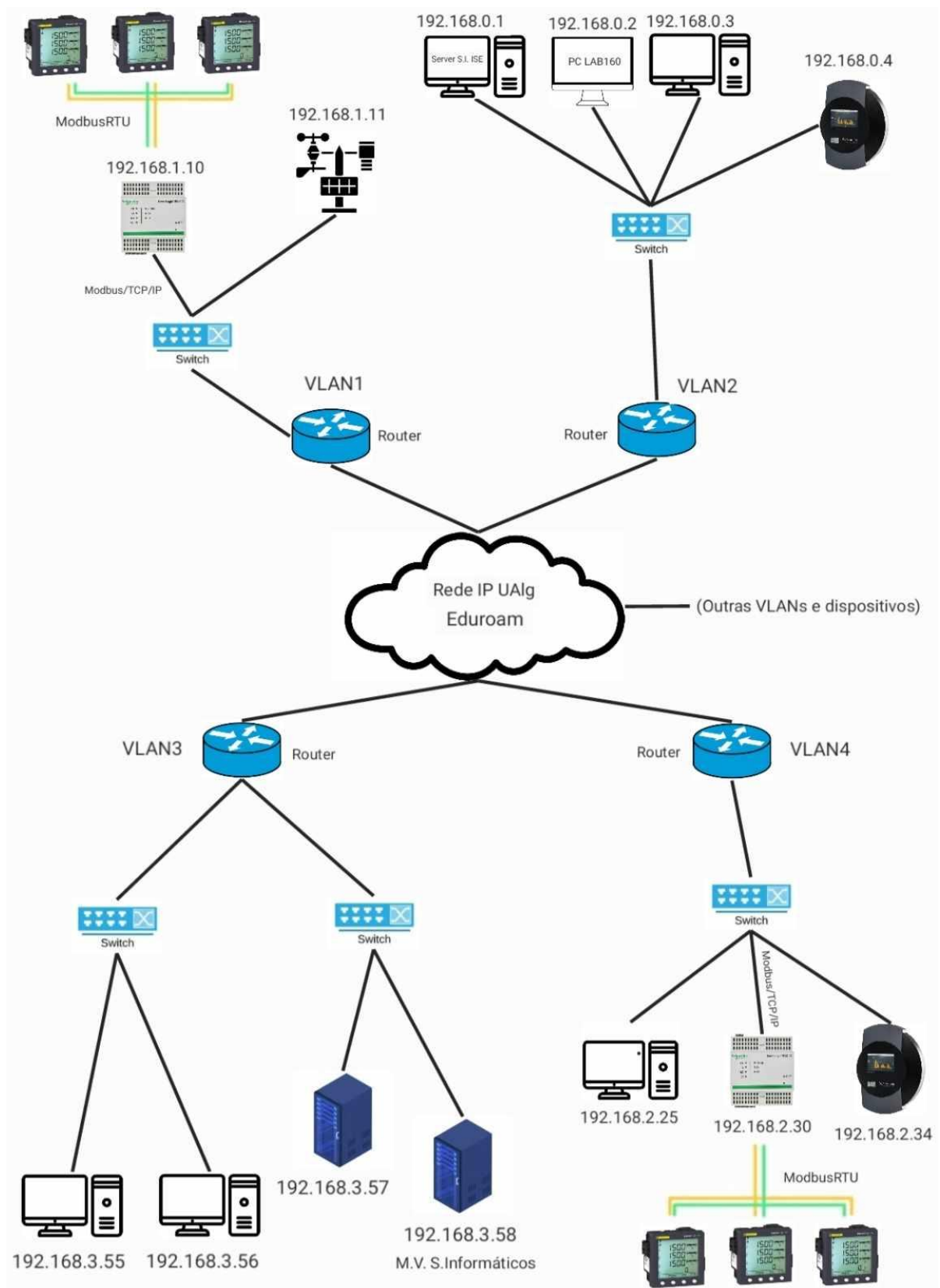


Figura 18 – Exemplo genérico da estrutura da rede interna da UAlg

A fase inicial do sistema era composta por dois componentes principais, ambos a operar sob o sistema Linux Ubuntu 16.04 LTS. O primeiro era um servidor central, localizado nos Serviços Informáticos do edifício principal do ISE, onde se encontrava alojada a plataforma EmonCMS. Este servidor dispunha de um processador com dois núcleos a 2.90 GHz, 4 GB de memória RAM e 500 GB de armazenamento interno. O segundo componente correspondia a um servidor físico instalado no laboratório 160, frequentemente referido como PC LAB160, que era responsável, à data, pela execução de três scripts distintos desenvolvidos em Python. Este servidor apresentava uma configuração semelhante, com dois núcleos a 3.40 GHz, 4 GB de RAM e igualmente 500 GB de armazenamento interno. A comunicação com os medidores era realizada através de *gateways Modbus*, que recebiam as requisições do PC LAB160 e transmitiam os valores das medições. O sistema utilizava um conjunto diversificado de medidores distribuídos pelos diferentes edifícios dos Campus de Gambelas e da Penha.

Importa referir que, embora a organização atual dos pontos de medição esteja documentada nas Tabelas 1 até à 8, nem todos os edifícios listados estavam incluídos na fase inicial do sistema. O edifício Tec Campus, por exemplo, foi apenas integrado mais tarde, já no decurso do presente projeto.

À data da implementação original, não existia qualquer uniformidade na nomeação dos pontos de medição, o que dificultava a gestão, análise e manutenção da infraestrutura. Essa limitação foi superada neste projeto, através da definição de uma convenção padronizada de nomenclatura, que associa a cada medidor uma designação clara e estruturada, baseada na inicial do campus e no número oficial do edifício na instituição. Esta uniformização revelou-se essencial para melhorar a organização da informação e facilitar a escalabilidade futura do sistema.

Durante este processo, foi ainda necessário resolver a ambiguidade associada a medições identificadas genericamente como as Zonas, numeradas de 2 a 5, que até então não apresentavam correspondência clara com os nomes dos edifícios da universidade. Para ultrapassar essa limitação, foi realizada uma ação conjunta com o Engenheiro João Mascarenhas, dos Serviços Técnicos da Universidade do Algarve, na qual, aproveitando um período de menor ocupação dos campus, foram efetuados desligamentos controlados para observar os efeitos nas medições. Esta abordagem permitiu mapear com precisão essas zonas aos respetivos edifícios ou agrupamentos de pavilhões.

Como referido, as tabelas 1 a 8 resumem os pontos de medição organizados por *gateway* e endereço IP, já com os nomes dos edifícios devidamente identificados e a nomenclatura padronizada aplicada. Neste conjunto as tabelas 1 até à 4 resumem a distribuição dos medidores no Campus das Gambelas e as tabelas 5 até à 8 a distribuição dos medidores no Campus da Penha.

DISTRIBUIÇÃO DOS MEDIDORES NO CAMPUS DAS GAMBELAS

Tabela 1 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.2 no Campus das Gambelas

ID	Nome no Sistema	Edifício / Localização	Medidor
11	ge1_fct_cintal_edificio_1	Edifício 1 (FCT/CINTAL)	PM710
12	gpt1_iluminacao_exterior	Iluminação Exterior GPT1	PM710
13	ge9_e10_Leoa_e_horto	Edifícios 9 e 10 (Leoa e Horto)	PM710
14	ge2_fmcb_cbmr_edificio_2	Edifício 2 (FMCB/CBMR)	PM710

Tabela 2 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.3 no Campus das Gambelas

ID	Nome no Sistema	Edifício / Localização	Medidor
21	ge3_cgd_e_portaria	Edifício 3 (CGD e Portaria)	TRV00210
22	gpt2_iluminacao_exterior	Iluminação Exterior GPT2	TRV00210
23	ge6_cantina_gambelas	Cantina de Gambelas	TRV00210
24	ge5_biblioteca_reitoria	Biblioteca e Reitoria	TRV00210
25	ge4_complexo_pedagogico	Complexo Pedagógico	PM710

Tabela 3 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.4 no Campus das Gambelas

ID	Nome no Sistema	Edifício / Localização	Medidor
5	ge7_ccmar_edificio_7	Edifício 7 (CCMAR)	PM5100
6	ge8_fe_fct_edificio_8	Edifício 8 (FE e FCT)	PM5100
33	ge8_geracao_fotovoltaica	Geração Fotovoltaica Edifício 8	PM5100

Tabela 4 – Edifícios monitorizados pelo gateway com o IP: 10.17.22.5 no Campus das Gambelas

ID	Nome no Sistema	Edifício / Localização	Medidor
41	ge12_zona_geral_somatorio_zonas	Edifício 12 (Zona Geral – somatório das zonas)	PM710
42	ge12_zona_2_pavalv_d_h_f	Edifício 12 Zona 2 (Pavilhões D, H e F)	PM710
43	ge12_zona_3_pavalv_e	Edifício 12 Zona 3 (Pavilhão E)	PM710
44	ge11_pav_madeira_jotas_zona_4	Edifício 11 (Pavilhão Madeira - Jotas Zona 4)	PM710
45	ge12_zona_5_bioterio	Edifício 12 Zona 5 (Biotério)	PM710
46	ge12_zona_6_pavalv_g	Edifício 12 Zona 6 (Pavilhão G)	PM710
47	ge12_zona_7_pavalv_l_bar_alvaro	Edifício 12 Zona 7 (Pavilhão L e Bar Álvaro)	PM710

DISTRIBUIÇÃO DOS MEDIDORES NO CAMPUS DA PENHA

Tabela 5 – Edifícios monitorizados pelo gateway com o IP: 10.18.22.12 no Campus da Penha

ID	Nome no Sistema	Edifício / Localização	Medidor
1	pe28_ise_edif_u	ISE Edifício U	PM5100
2	pe27_ise_edif_principal	ISE Edifício Principal	PM5100
3	pe29_complexo_pedagogico	Complexo Pedagógico Penha	PM5100
4	pe28_bateria_condensadores	Bateria de Condensadores <i>(não monitorizado inicialmente)</i>	PM5100
5	pe23_esec	ESEC	PM5100
6	pe26_cantina_penha	Cantina Penha	PM5100
7	pe22_esght	ESGHT	PM5100
8	pe34_residencia	Residência Penha	PM5100
9	pe33_st_oficinas	Oficinas	PM5100
10	pe32_eng_civil	Engenharia Civil	PM5100

Tabela 6 – Edifícios monitorizados pelo gateway com o IP: 10.33.13.5 no Campus da Penha

ID	Nome no Sistema	Edifício / Localização	Medidor
12	pe30_tec_campus_geracao	Tec Campus – Geração <i>(não monitorizado inicialmente)</i>	PM5110
13	pe30_tec_campus_consumo_geral	Tec Campus – Consumo Geral <i>(não monitorizado inicialmente)</i>	PM3250

Tabela 7 – Geração de edifícios monitorizados com IPs: 10.17.22.9, 10.17.22.10 e 10.17.22.11 no Campus da Penha

Nome no Sistema	Edifício / Localização	Medidor
pe23_esec_geracao	Geração Solar – ESEC	SUN2000
pe27_ise_edif_principal_geracao	Geração Solar – ISE Principal	SUN2000
pe22_esght_geracao	Geração Solar – ESGHT	SUN2000

Tabela 8 – Geração do Edifício em U do ISE com IP: 10.17.22.8 no Campus da Penha

Nome no Sistema	Edifício / Localização	Medidor
pe28_ise_edif_u_geracao	Geração Solar – ISE U	SOLARLOG

Após a identificação e padronização dos pontos de medição, importa agora descrever o funcionamento do sistema de recolha e armazenamento de dados na arquitetura inicial.

Os três *scripts* em execução no PC LAB160 tinham como função recolher dados de grupos distintos de medidores e transmitir os valores obtidos para o servidor central com EmonCMS, onde eram posteriormente armazenados e analisados. Sendo que, dois desses scripts, desenvolvidos na fase inicial do projeto, eram responsáveis pela maioria das leituras dos medidores localizados nos diferentes campus. Posteriormente, foi adicionado um terceiro script com o objetivo de integrar novos medidores instalados em edifícios não contemplados na implementação original, incluindo dispositivos de geração solar SolarLog, que comunicam através de API REST.

Esta abordagem, assente em múltiplos scripts independentes desenvolvidos em momentos distintos, originou uma estrutura fragmentada, que dificultava a gestão, manutenção e uniformização do sistema de recolha.

Paralelamente, os dados recolhidos eram armazenados localmente em duas bases de dados distintas (SQLite e PostgreSQL), ambas com falhas recorrentes. Estas falhas resultavam em lacunas nos registos e períodos prolongados sem dados, provocadas por erros nos *scripts*, interrupções de comunicação com os medidores e por uma arquitetura de armazenamento desajustada às necessidades de uma aplicação de monitorização contínua.

Todos os *scripts* estavam configurados para execução automática minuto a minuto via *crontab*, o que permitia uma recolha constante.

A Figura 19 representa de forma detalhada a arquitetura inicial do sistema, antes das intervenções realizadas no âmbito deste projeto, evidenciando o servidor central com EmonCMS (localizado nos Serviços Informáticos do ISE), o computador físico do Laboratório 160, responsável pela execução dos *scripts* e pelo alojamento das bases de dados locais e a comunicação com os medidores de energia. Para garantir a comunicação entre todos os dispositivos envolvidos, era essencial assegurar a visibilidade mútua a nível de IP dentro da rede.

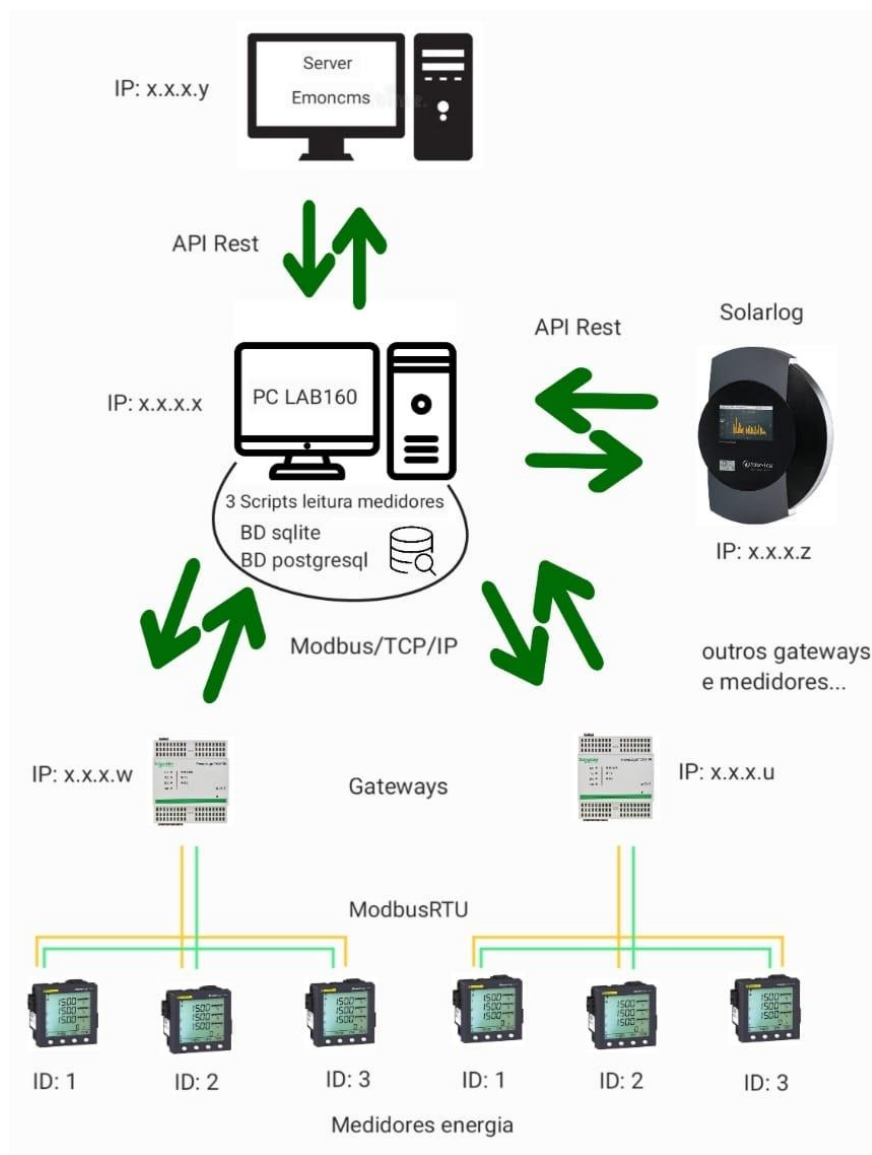


Figura 19 – Arquitetura inicial do sistema

A arquitetura inicial, embora funcional, revelou ao longo do tempo várias limitações que comprometeram a fiabilidade, escalabilidade e sustentabilidade do sistema de monitorização energética. Entre os principais problemas identificados destacam-se:

- Dependência exclusiva de um único computador físico localizado no laboratório 160, responsável pela execução dos scripts de recolha de dados. Esta configuração introduzia um ponto único de falha, sem qualquer mecanismo de redundância ou recuperação automática. Existia um segundo servidor, mas este encontrava-se dedicado a outras funções, não tendo capacidade para assegurar as tarefas em caso de falha do sistema principal.
- Fragmentação dos scripts de recolha, inicialmente divididos em três ficheiros independentes, o que aumentava consideravelmente a complexidade operacional. Esta divisão dificultava a manutenção e a atualização do sistema, aumentando o risco de erros e incoerências no processo de recolha.
- Ausência de uma estrutura centralizada e organizada para lidar com a adição ou alteração de medidores. Sempre que era necessário integrar um novo equipamento, o processo exigia modificações manuais dispersas no código, sem documentação uniforme, o que tornava a tarefa propensa a erros e fortemente dependente do conhecimento individual de cada programador que, em momentos distintos, contribuiu de forma isolada para o desenvolvimento deste sistema.
- Frequência elevada de falhas na comunicação entre o servidor e as *gateways* Modbus dos medidores. Estas falhas não eram automaticamente detetadas ou tratadas, o que comprometia a continuidade da recolha de dados. No caso de ocorrer um erro no envio de determinada variável por parte de um medidor, o sistema atribuía-lhe o valor *None*, provocando um bloqueio momentâneo da execução do script por alguns segundos antes de prosseguir. Embora o script não terminasse abruptamente, esta paragem afetava negativamente a sua eficiência, uma vez que o processo completo de recolha tinha de ser concluído dentro de um minuto, respeitando o intervalo entre medições definido para o sistema.
- Presença de lacunas significativas nas bases de dados utilizadas (SQLite e PostgreSQL), não apenas pela ocorrência de perdas de dados, mas também devido a uma estrutura inadequada, desorganizada em termos de nomenclatura de colunas e tabelas, e tecnicamente desatualizada, o que afetava a fiabilidade e a coerência do armazenamento histórico.
- Ambos os servidores que suportavam o sistema apresentavam sistemas operativos desatualizados, operando ainda com o Ubuntu 16.04 LTS, cuja manutenção de longo prazo já tinha expirado. Esta desatualização comprometia a segurança e a compatibilidade com bibliotecas e serviços mais recentes. Para além disso, a instalação do EmonCMS em

funcionamento encontrava-se na versão 10.2 — uma versão significativamente desatualizada — sem suporte a vários módulos importantes introduzidos em atualizações posteriores, nomeadamente funcionalidades de sincronização de *feeds*, gestão de utilizadores mais segura, e melhorias ao nível da API e da interface de *dashboards*. Esta obsolescência técnica limitava a evolução do sistema e a sua integração com novas ferramentas ou funcionalidades pretendidas para fases futuras do projeto.

A plataforma EmonCMS encontrava-se alojada num servidor acessível externamente através da internet, permitindo que qualquer utilizador pudesse aceder à interface de login a partir de qualquer localização. Esta abertura ao exterior tinha como objetivo inicial permitir que outros projetos da Universidade do Algarve, a operar fora da rede local, pudessem enviar dados de sensores para a plataforma. No entanto, tendo em consideração que a plataforma passou a conter dados críticos relacionados com o consumo energético de toda a universidade, esta configuração representava um risco significativo de segurança, comprometendo a integridade e a confidencialidade da informação. Tornou-se, assim, evidente a necessidade de restringir o acesso à rede interna da instituição.

3.2 ARQUITETURA INTERMÉDIA DO SISTEMA

Com o objetivo de superar as limitações identificadas na fase inicial, foram implementadas melhorias técnicas e operacionais durante a primeira fase deste projeto. As alterações realizadas visaram essencialmente a modernização e robustecimento do sistema.

Entre as principais ações realizadas, destaca-se a migração do sistema para uma máquina virtual nos Serviços Informáticos (S.I.) no campus de Gambelas. Todo o sistema anteriormente operacional nos dois servidores iniciais foi integralmente transferido para esta nova infraestrutura, alojada nos S.I. da Universidade do Algarve, o que permitiu centralizar a recolha e o armazenamento de dados num ambiente mais robusto e seguro, reduzindo significativamente a dependência do servidor físico localizado no laboratório 160. A máquina virtual atribuída foi configurada com especificações adequadas a uma operação contínua e expansível, nomeadamente: 4 núcleos de processamento, 8 GB de RAM, 512 GB de armazenamento em disco e sistema operativo Linux Ubuntu 22.04 LTS.

A Figura 20 ilustra a organização do sistema durante a fase intermédia, na qual toda a recolha, tratamento e armazenamento de dados passou a estar centralizada numa única máquina virtual, alojada nos Serviços Informáticos da Universidade do Algarve. Esta abordagem contrastava com a estrutura anterior, distribuída entre múltiplos servidores e com diferentes níveis de

fiabilidade. A comunicação com os *gateways* é realizada via Modbus/TCP/IP, enquanto alguns dados de sistemas fotovoltaicos são obtidos através de API REST. Toda a informação recolhida é enviada para a plataforma EmonCMS, instalada localmente na mesma máquina, e simultaneamente registada nas bases de dados PostgreSQL e SQLite, assegurando redundância, persistência e simplificação da gestão técnica. Esta configuração permitiu estabilizar o sistema e consolidar uma base sólida para a fase seguinte de desenvolvimento.

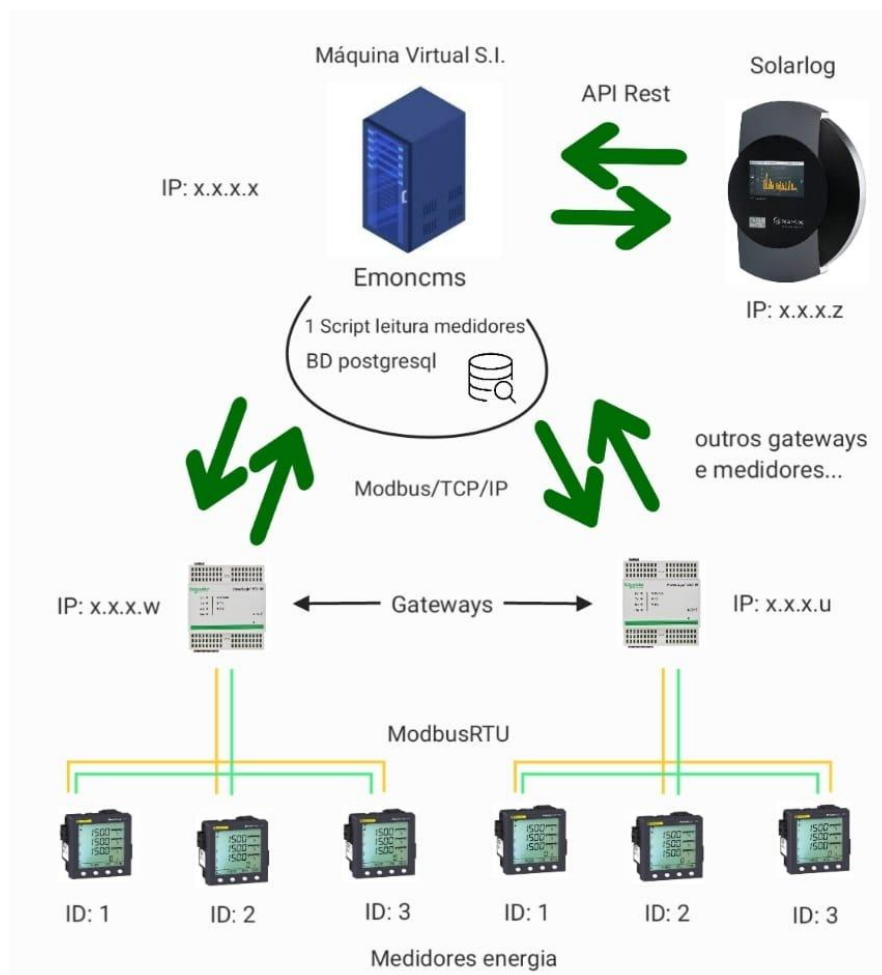


Figura 20 – Arquitetura numa fase intermédia

Esta migração exigiu igualmente uma reconfiguração da rede interna, levada a cabo em estreita colaboração com os Serviços de Informática da universidade, com o objetivo de assegurar uma comunicação estável e contínua entre a nova máquina virtual e os diversos medidores instalados nos edifícios dos diferentes campus.

Um dos aspetos mais críticos deste processo foi a necessidade de garantir que as VLAN onde se encontram fisicamente os *gateways* Modbus — normalmente isoladas para fins de segurança e segmentação de tráfego — estivessem devidamente roteadas até à sub-rede onde a máquina virtual foi instalada. Esta configuração envolveu a atualização das tabelas de

encaminhamento, a criação de rotas estáticas e o ajustamento das regras de *firewall*, de modo a permitir o tráfego entre os diferentes segmentos da rede interna de forma segura e sem colisões.

Durante a análise das condições da instalação anterior, verificou-se que os dados energéticos da Universidade do Algarve se encontravam acessíveis a partir do exterior da rede institucional, uma vez que a plataforma EmonCMS estava exposta à internet pública e permitia o acesso à sua interface de autenticação a qualquer utilizador. Esta configuração, embora inicialmente útil para permitir a integração com projetos externos à rede da universidade, passou a representar um risco relevante em termos de segurança, dado o carácter sensível e estratégico da informação recolhida.

Com a nova instalação sobre a máquina virtual dos Serviços Informáticos, o acesso à plataforma passou a estar limitado exclusivamente à rede interna. Esta medida visa mitigar potenciais ameaças externas e garantir que apenas utilizadores autorizados, dentro da rede ou através de VPN à rede interna, possam aceder aos dados energéticos da instituição. Um dos desafios mais significativos durante esta fase foi a transferência dos dados históricos armazenados na instalação antiga para a nova. A versão antiga em utilização não possuía um módulo incorporado para realizar automaticamente *backups* completos dos dados históricos. Ao contrário da nova versão instalada na máquina virtual, que já integra esta funcionalidade, a antiga exigiu uma abordagem manual.

Foi necessário identificar manualmente as pastas essenciais da instalação local antiga, especificamente as pastas e a estrutura da BD:

- `phpfina` - dados históricos de alta frequência;
- `phpfiwa` - dados históricos com média granularidade;
- `phptimeseries` - dados temporais estruturados;
- Base de dados SQL com estrutura essencial (utilizadores, *feeds*, *inputs* e *dashboards* gráficos)

Após a identificação e agrupamento das pastas essenciais, procedeu-se à criação manual de um backup integral no formato `.tar.gz`, por ser o único suportado pelo módulo de importação da versão atual do EmonCMS. Para tal, foi utilizado o comando `tar -czvf`, que permite comprimir múltiplos diretórios num único ficheiro compactado no formato pretendido. A

transferência deste ficheiro da máquina original para a nova instalação foi realizada através do comando `SCP`.

Dado o elevado número de dados históricos e o número significativo de variáveis envolvidas, tornou-se inviável realizar a importação diretamente através da interface gráfica do EmonCMS, mais propriamente utilizando o módulo *backup* disponível, porque este processo frequentemente resultava em erros ou interrupções devido a problemas de permissões e limites de tempo de execução – *timeouts* - impostos pelo servidor *web* Apache.

Perante esta limitação, optou-se por uma solução mais robusta, recorrer à execução direta do script `emoncms-import.sh`, disponibilizado pelo EmonCMS e localizado por omissão em `/opt/emoncms/modules/backup/`. Para que este método funcionasse corretamente, foi necessário colocar manualmente o ficheiro de backup `.tar.gz` na pasta `uploads/`, situada em `/var/opt/emoncms/backup/uploads/`.

A execução do *script bash* através do terminal Linux permitiu contornar as limitações impostas pela interface gráfica do EmonCMS e realizar uma importação integral e bem-sucedida dos dados. Esta abordagem revelou-se especialmente eficaz face ao volume significativo de informação acumulada ao longo de vários anos de operação e à multiplicidade de *feeds* provenientes de dezenas de medidores distintos. Apesar do tempo de execução prolongado, que se estendeu por várias horas, o processo assegurou a replicação fiel de toda a instalação anterior. A nova instância do EmonCMS passou assim a refletir integralmente a configuração original, mantendo todos os *feeds*, *inputs*, *dashboards*, utilizadores e definições existentes, sem perdas ou discrepâncias. Esta fidelidade na transição foi essencial para garantir a continuidade funcional da plataforma e preservar o histórico energético.

Após a migração, foi implementada uma rotina automatizada de backup diário, de forma a garantir a segurança e integridade contínua dos dados. Esta rotina baseia-se na execução periódica do script existente na instalação do Emoncms, `emoncms-export.sh`, configurado através do `crontab` do sistema Linux, assegurando assim cópias de segurança diárias da base de dados e das configurações essenciais da plataforma.

Ainda no âmbito desta fase de reestruturação técnica, foi também iniciada a recolha sistemática da potência reativa medida diretamente pelos medidores instalados em todos os edifícios monitorizados. Esta variável, embora já estivesse disponível nos equipamentos, não era anteriormente armazenada, limitando as possibilidades de análise técnica. A sua inclusão

permite efetuar análises energéticas mais completas, nomeadamente através do cálculo e avaliação do fator de potência, contribuindo para um melhor diagnóstico dos consumos e alinhando-se com os requisitos previamente identificados pelos Serviços Técnicos da Universidade.

3.2.1 REESTRUTURAÇÃO E OTIMIZAÇÃO DO CÓDIGO

Paralelamente à migração, foi realizada uma reestruturação integral e atualização dos *scripts* utilizados para a recolha dos dados energéticos. Como referido, o sistema utilizava três *scripts* separados, cada um responsável pela leitura de medições específicas associadas a diferentes grupos de edifícios. Procedeu-se à integração destes *scripts* num único *script* centralizado, simplificando significativamente a sua manutenção e tornando o processo de recolha mais eficiente e menos propenso a falhas.

Devido à escolha da versão mais recente com suporte de longo prazo (LTS) do sistema operativo Ubuntu (22.04 LTS), selecionada para ser utilizada na nova máquina virtual, e à consequente adoção da versão atualizada do Python (Python 3.12), surgiram incompatibilidades técnicas com a estrutura do código original, nomeadamente com listas aninhadas de comprimentos diferentes — uma estrutura anteriormente utilizada para representar os tarifários horários. Este tipo de organização, além de ser difícil de manter, deixou de ser suportado diretamente nas versões mais recentes do Python, forçando uma reestruturação profunda da lógica de cálculo dos custos energéticos.

O tarifário em vigor na Universidade do Algarve é o tetra-horário, que divide cada dia em quatro períodos distintos — Ponta, Cheia, Vazio e Super Vazio — cuja distribuição varia entre verão e inverno. A definição detalhada dos períodos e a sua calendarização encontra-se descrita na Figura 17.

Na versão inicial do código, o cálculo do custo energético baseava-se numa matriz complexa com listas aninhadas, onde cada elemento representava blocos horários específicos da semana, diferenciados por estação do ano. Essas listas continham sublistas com diferentes comprimentos, correspondentes a intervalos de tempo variáveis e, a cada execução, o script realizava cálculos dinâmicos intensivos, como geração de *arrays* em tempo real e conversões de índice, para determinar o período tarifário correspondente a cada instante. Com a nova abordagem, o método foi simplificado e otimizado. O código foi reestruturado

para utilizar listas fixas de 96 elementos, representando os 96 blocos de quinze minutos que compõem as 24 horas do dia. Foram criadas listas distintas para dias úteis, sábados e domingos, bem como para as estações verão e inverno. Cada posição dessas listas contém um valor numérico representando o período tarifário correspondente (1 – Super Vazio, 2 – Vazio, 3 – Cheia, 4 – Ponta).

Caso os períodos tarifários venham a ser alterados em anos futuros, será necessário criar vetores de 96 elementos com base nas novas definições horárias atualizadas. Estes vetores devem depois substituir as listas existentes no script principal do sistema, garantindo que o cálculo do custo energético continua alinhado com os tarifários oficialmente em vigor. A escolha da lista apropriada é feita automaticamente com base no dia da semana e na verificação da estação do ano, que são obtidos do sistema. Importa, contudo, notar que, apesar de os cálculos históricos permanecerem armazenados na base de dados e não serem afetados por alterações posteriores, uma eventual necessidade de recalcular consumos passados exigiria a preservação das tabelas tarifárias originais, de modo a assegurar a reprodutibilidade e a consistência dos resultados.

Após esta seleção, identifica-se o índice do bloco de quinze minutos correspondente ao momento atual utilizando a fórmula:

$$\text{Índice} = (\text{hora atual} \times 4) + \left\lfloor \frac{\text{minuto atual}}{15} \right\rfloor \quad (16)$$

Ou seja, em (3) a hora é multiplicada por 4 (correspondente aos quatro blocos de quinze minutos por hora) e ao resultado soma-se o quociente inteiro da divisão dos minutos por 15, ignorando os decimais. Este índice permite aceder à posição correta dentro da lista já selecionada, indicando o bloco de quinze minutos correspondente ao momento atual. O valor obtido nesse índice representa o código do período tarifário, de 1 a 4, que por sua vez determina o custo por kWh previamente definido. Este novo método permite determinar de forma imediata e eficiente o período tarifário e o respetivo preço por kWh, sem recorrer a estruturas aninhadas complexas ou cálculos dinâmicos *on-the-fly*, o que torna o código mais claro, mais leve computacionalmente e facilmente adaptável a futuros tarifários. Para o caso

em concreto aqui considerado, em que os tarifários apresentam períodos fixos e repetitivos definidos pela ERSE, este cálculo mostra-se plenamente adequado. Importa, no entanto, reconhecer que em cenários de tarifários dinâmicos, nos quais o preço da energia varia diariamente ou em tempo real, a abordagem teria de ser adaptada através da integração de tabelas de preços atualizadas ou da ligação direta a plataformas de mercado, assegurando a correta determinação do custo horário da energia em contextos mais voláteis.

3.2.2 AQUISIÇÃO DE DADOS VIA MODBUS/TCP/IP

A aquisição de dados é realizada por comunicação direta com os medidores instalados nos diferentes edifícios da universidade, utilizando o protocolo Modbus/TCP/IP. Esta comunicação é implementada no script principal de aquisição através da biblioteca `pyModbusTCP`, uma biblioteca em Python concebida para facilitar a implementação do protocolo Modbus sobre conexões TCP/IP. Esta biblioteca permite o estabelecimento de sessões com os dispositivos remotos, a definição do identificador de escravo, *slave ID*, e a leitura direta dos registos de forma simplificada.

No código desenvolvido, cada variável medida corresponde a um pedido Modbus independente, ou seja, para cada conjunto de registos, por exemplo, tensões, correntes ou potências, é feita uma chamada individual à função `read_input_registers`. Esta abordagem, apesar de resultar num maior número de pedidos por medidor, permite isolar falhas de leitura e simplificar o tratamento de exceções.

Cada pedido especifica o endereço IP da *gateway* associada ao equipamento e o respetivo *slave ID* do medidor. A função de leitura abre automaticamente a ligação TCP, executa o pedido dos registos pretendidos e encerra a ligação após a resposta. Este mecanismo garante uma gestão eficiente das conexões, evitando a manutenção de sessões persistentes desnecessárias e permitindo uma melhor tolerância a falhas temporárias na rede.

Os dispositivos utilizados — de diferentes modelos (PM710, PM5100, TRV00210) — disponibilizam os seus dados através de registos Modbus, cuja organização e codificação variam consoante o fabricante. Embora estes equipamentos permitam o acesso a um vasto conjunto de variáveis elétricas — incluindo tensões fase-fase e fase-neutro, correntes por fase, potências por fase, fatores de potência, frequência, energias acumuladas, bem como valores

mínimos, máximos e por vezes médios — foi adotada uma abordagem seletiva. Com o objetivo de simplificar o processamento e centrar a análise nas grandezas de maior relevância energética, optou-se por recolher apenas um subconjunto de variáveis fundamentais: tensões fase-neutro (A, B e C), correntes por fase (A, B e C), potência ativa e potência reativa.

As variáveis são lidas sob a forma de registos Modbus de 16 bits (2 bytes). Para representar grandezas superiores, como potências ou energias, os dispositivos recorrem à junção de dois registos consecutivos (totalizando 32 bits), normalmente codificados em ordem *little endian*, LSB (*Least Significant Bit*) seguido de MSB (*Most Significant Bit*). A conversão destes pares de registos binários para valores físicos é efetuada no próprio *script* de aquisição, utilizando a biblioteca `struct`, que permite interpretar corretamente a ordem dos bytes e extrair os valores em formato de ponto flutuante (*float*). Quando necessário, é também aplicado um fator de escala próprio do fabricante, que ajusta o valor bruto à unidade física adequada (por exemplo, amperes, volts ou watts).

Esta lógica modular e robusta de leitura permite uma integração eficaz e fiável dos medidores no sistema de monitorização energética, garantindo a coerência e qualidade dos dados recolhidos.

Exemplo de conversão de dois registos com fator de escala para energia em kWh

Suponhamos que um medidor fornece os seguintes registos consecutivos:

- Registo LSB = 0xC350
- Registo MSB = 0x0003

A concatenação (*little endian*) resulta em:

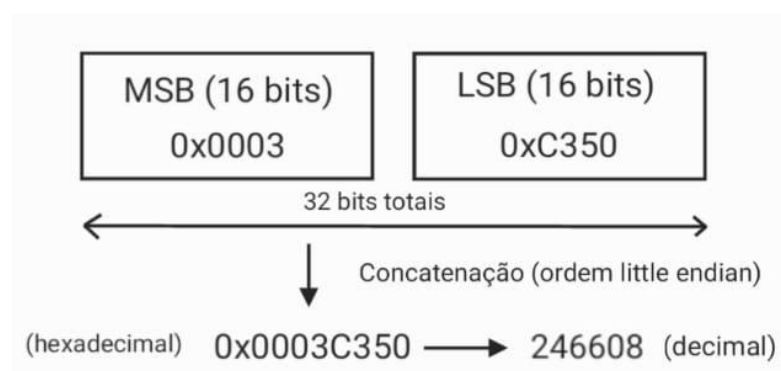


Figura 21 – Conversão de registos Modbus

O valor decimal representado na Figura 21 corresponde à energia ativa instântanea, devendo, no entanto, ser ajustado com o fator de escala definido pelo fabricante — por exemplo através da divisão por 10, i. e.,

$$Energia(Wh) = \frac{246064}{10} = 24606,4 Wh ; \quad (17)$$

Convertendo para kilowatt-hora, temos:

$$Energia(kWh) = \frac{24606,4}{1000} = 24,606kWh ; \quad (18)$$

Estes cálculos descritos em (17) e em (18) são aplicados individualmente a cada variável recolhida, assegurando que as grandezas obtidas correspondem aos valores físicos reais medidos no sistema elétrico. Importa notar que o fator de escala pode variar consoante o tipo de variável e o modelo do equipamento, sendo, por isso, essencial consultar o manual técnico de registos *Modbus* de cada dispositivo sempre que existam dúvidas sobre a forma correta de interpretar os valores brutos.

Com base no valor da potência ativa medida e na tarifa horária previamente determinada, o custo instantâneo é calculado diretamente no *script* de aquisição, em euros por hora (€/h), no momento da leitura de cada medidor. Esta unidade foi escolhida por proporcionar uma leitura mais direta e intuitiva nos *dashboards* da plataforma, evitando valores demasiado reduzidos. Para permitir o cálculo do custo acumulado mensal, a plataforma EmonCMS aplica internamente uma conversão do valor de €/h para €/min, ajustando-o ao intervalo de amostragem de um minuto. Essa conversão é realizada através do módulo *Input Process*, que permite definir cadeias automáticas de operações sobre os dados recebidos — como multiplicações, divisões, somas acumulativas — e guardar os resultados num *feed* distinto. Neste caso, a conversão é efetuada por uma multiplicação pelo fator 0.01(6) (equivalente a 1/60), e o valor resultante é acumulado ao longo do tempo.

3.2.3 AQUISIÇÃO DE DADOS VIA HTTP

Além dos dispositivos que comunicam via *Modbus*, o sistema também integra equipamentos que disponibilizam os seus dados através de interfaces HTTP. Nestes casos, a aquisição é

efetuada por meio de requisições diretas ao endereço IP do equipamento, dispensando a utilização de *gateways* ou protocolos industriais adicionais.

Em comparação com a leitura por *Modbus*, a aquisição via HTTP apresenta a vantagem dos dados já se encontrarem disponíveis num formato diretamente interpretável, geralmente em JSON. Este formato elimina a necessidade de converter registos binários, aplicar fatores de escala ou efetuar transformações *little-endian*. As grandezas elétricas são extraídas com as respetivas unidades físicas — como amperes, volts ou watts — o que simplifica significativamente o processo de integração.

Importa salientar que o tipo de interface utilizado por cada equipamento não resulta de uma escolha arbitrária, mas sim das características técnicas dos próprios medidores, definidas pelos seus fabricantes e modelos. Assim, a utilização do protocolo HTTP decorre naturalmente da especificação dos dispositivos, nomeadamente de inversores fotovoltaicos ou sistemas de monitorização como o SolarLog.

O *script* principal de aquisição de dados está preparado para lidar com este tipo de comunicação, realizando requisições periódicas à API dos equipamentos e interpretando diretamente os valores recebidos, que incluem tipicamente variáveis elétricas essenciais como correntes, tensões, potência ativa e potência reativa. Estes dados são integrados de forma harmoniosa com o restante sistema e tratados localmente utilizando a mesma estrutura aplicada aos restantes medidores, promovendo uniformidade na análise posterior.

Atualmente, esta forma de aquisição via HTTP está implementada exclusivamente nos medidores SolarLog associados aos sistemas de produção fotovoltaica, embora possa ser facilmente estendida a qualquer equipamento compatível com comunicação RESTful. Entre os dispositivos atualmente operacionais, este método está a ser utilizado apenas na monitorização da instalação fotovoltaica do edifício em U do Instituto Superior de Engenharia, no campus da Penha, onde é recolhida e armazenada unicamente a potência ativa, por ser a grandeza mais representativa da produção solar em tempo real.

O inversor monitorizado disponibiliza também outras variáveis, como as tensões e correntes por fase, acessíveis via API e potencialmente integráveis no sistema no futuro, caso sejam requeridas para uma análise mais detalhada ou para suportar novas funcionalidades.

3.2.4 MIGRAÇÃO DA BASE DE DADOS POSTGRESQL

Durante o processo de transição para a nova infraestrutura, procedeu-se à migração da base de dados PostgreSQL utilizada desde a fase inicial do sistema, anteriormente alojada no servidor

físico do Lab 160. Esta base de dados foi criada com o intuito de servir como repositório complementar ao EmonCMS, prevendo-se a sua integração com ferramentas de visualização como o Grafana.

Apesar de apresentar várias limitações, nomeadamente falhas recorrentes de registo e uma estrutura que não se mostrava totalmente adequada aos requisitos do sistema, optou-se nesta fase por preservar o conteúdo existente, garantindo a sua continuidade no novo ambiente, enquanto se estudava uma solução mais robusta e adaptada.

Para realizar a migração, foi efetuado um *dump* completo da base de dados, utilizando o comando `pg_dump`, nativamente incluído no PostgreSQL. Este comando permite gerar um ficheiro de *backup* contendo toda a estrutura (tabelas, esquemas, tipos de dados) e todo o conteúdo da base. A exportação foi feita diretamente no terminal do servidor de origem.

O ficheiro resultante, com a extensão `.dump` foi depois transferido para a nova máquina virtual utilizando o comando `SCP`, que permite a cópia segura de ficheiros entre máquinas ligadas em rede.

Na nova máquina, foi criada uma base de dados vazia com o mesmo nome da original, através do comando SQL `createdb`, associada a um utilizador previamente configurado para esse propósito. A partir daí, foi utilizada a ferramenta `pg_restore` para importar o conteúdo do ficheiro de *backup*, reconstituindo fielmente a base de dados na nova infraestrutura.

Esta operação permitiu preservar todos os dados históricos recolhidos até então, mesmo tratando-se de uma base provisória, assegurando a continuidade do trabalho de desenvolvimento e testes, enquanto se planeava a reestruturação total deste componente do sistema.

3.2.5 Migração da Base de Dados SQLite

Para além da base de dados PostgreSQL, o sistema dispunha de uma base de dados SQLite presente num ficheiro com extensão `.db`, localizado num diretório do sistema. Ambas funcionavam em paralelo, com o objetivo de registar simultaneamente as medições dos mesmos edifícios e variáveis. O *script* principal de recolha, responsável pela leitura dos medidores, encontrava-se diretamente ligado a esta base de dados. Para cada medição, o *script* procedia ao envio dos dados para o EmonCMS, ao registo na base PostgreSQL e, adicionalmente, à execução de uma função específica dedicada ao armazenamento local na base SQLite.

Apesar de o SQLite ser geralmente reconhecido pela sua leveza, portabilidade e facilidade de integração em sistemas embebidos, a implementação existente apresentava várias limitações (ver Secção 2.6.4). Tal como acontecia com a base PostgreSQL, não estavam a ser armazenados dados de forma uniforme para todos os edifícios nem para todas as variáveis monitorizadas. Verificavam-se lacunas significativas nos registos, bem como uma estrutura de tabelas com nomes pouco diferenciadores, desorganizados e tecnicamente desatualizados, o que dificultava a sua utilização posterior para análise ou integração com outras ferramentas. O processo de migração teve início com a identificação do ficheiro principal da base de dados SQLite, que foi posteriormente transferido para a nova máquina virtual utilizando o comando SCP. Após a transferência, foi verificada a integridade da base de dados com recurso à ferramenta `sqlite3`, executada a partir do terminal Linux. Para aceder ao conteúdo da base foi necessário recorrer ao ambiente da ferramenta através do comando `sqlite3 nome_base.db`, utilizando de seguida os comandos `.tables` e `.schema` para inspecionar a existência das tabelas e analisar a respetiva estrutura, confirmando assim o sucesso da migração. Contudo, a análise detalhada da organização e do conteúdo evidenciou limitações significativas na manutenção paralela desta base, sobretudo no que respeita à escalabilidade e integração com outras ferramentas. Em contrapartida, a nova estrutura em PostgreSQL revelou-se mais robusta, flexível e adequada às necessidades do sistema. Assim, a base de dados SQLite foi descartada na versão final, ficando a plataforma dependente exclusivamente da base PostgreSQL para o armazenamento estruturado das medições energéticas.

3.2.6 INTEGRAÇÃO ADICIONAL DE EDIFÍCIOS EXTERNOS

Para responder às necessidades de monitorização do edifício Tec Campus, estabeleceu-se uma comunicação direta com os responsáveis locais. Este processo incluiu o contacto com o Engenheiro Paulo Pacheco, responsável pelos dados energéticos do edifício, que forneceu as informações técnicas essenciais — nomeadamente os endereços IP das *gateways* Modbus e os IDs dos medidores instalados. À data, o edifício não se encontrava ainda integrado no sistema de monitorização da universidade, sendo esta a sua primeira ligação à plataforma.

Com base nesses dados, procedeu-se à configuração da rede interna da universidade, solicitando aos Serviços Informáticos o encaminhamento das VLANs onde se encontravam os medidores até à rede da máquina virtual que aloja o sistema de aquisição de dados, assegurando assim uma comunicação TCP/IP estável com os dispositivos instalados.

Com o Tec Campus já incorporado e o sistema estabilizado, foi possível iniciar o planeamento de uma nova fase de expansão, abrangendo vários edifícios externos da universidade. Entre os locais já identificados, alguns já possuem medidores instalados, mas aguardam a disponibilização dos dados de acesso por parte dos Serviços Informáticos — nomeadamente os endereços IP das *gateways* — para que a sua integração possa ser efetuada. Noutros casos, a expansão está ainda condicionada à instalação física de novos equipamentos de medição, de forma a garantir uma cobertura completa e representativa.

A integração destas infraestruturas dependerá, em todos os casos, da validação dos parâmetros técnicos de comunicação (endereço IP e identificador Modbus) e do mapeamento das variáveis energéticas disponibilizadas por cada dispositivo.

Entre os locais já identificados para futura integração encontram-se:

- Edifício 7 (geração fotovoltaica);
- Edifício 2 (geração fotovoltaica);
- Pavilhões de Alvenaria (geração fotovoltaica);
- Cantina de Gambelas (geração fotovoltaica);
- Ferragial (consumo e geração);
- Lote E (consumo e geração);
- Residência Albacor (consumo e geração).

O plano de expansão visa tirar partido da robustez e modularidade da arquitetura intermédia desenvolvida nesta fase do projeto, assegurando que o crescimento do sistema decorre de forma estruturada, escalável e compatível com os requisitos técnicos então implementados.

3.3 ARQUITETURA ATUAL DO PROJETO

A arquitetura atual do sistema assenta numa estrutura centralizada, consolidada numa máquina virtual dedicada, onde reside a versão mais estável e otimizada do código principal de aquisição, tratamento e organização de dados energéticos. Esta reorganização surge da necessidade de garantir maior fiabilidade e desempenho do sistema, sobretudo perante o aumento do número de medidores integrados e a complexidade da lógica de tratamento de dados. Entretanto, com o lançamento da versão 24.04 LTS do sistema operativo Ubuntu, e na sequência de uma sugestão automática do próprio sistema, foi realizada a atualização da versão 22.04 LTS anteriormente instalada, assegurando a continuidade de suporte, correções de segurança e compatibilidade com os requisitos técnicos da plataforma.

Para garantir a continuidade do serviço e aumentar a resiliência face a eventuais falhas, foi implementado um sistema redundante, baseado na criação de um espelho integral da instalação principal. Esse espelho foi configurado no PC do laboratório 160, que anteriormente desempenhava o papel de servidor principal, e passou a operar com a mesma versão atualizada do sistema operativo Ubuntu. Este processo envolveu a reinstalação do EmonCMS com todos os seus requisitos, bem como a transferência reversa da máquina virtual para o ambiente físico. A sincronização entre as duas instâncias foi realizada uma única vez no momento da criação do espelho, com o objetivo de garantir a equivalência estrutural e funcional entre ambas. Para tal, foram copiados manualmente todos os *feeds* do EmonCMS (ficheiros *.meta* e *.dat* do diretório PHPFina), bem como os ficheiros de configuração, scripts de aquisição, base de dados PostgreSQL e conteúdos associados. Esta sincronização pontual assegurou que ambas as instâncias partilham o mesmo estado inicial de dados e configuração.

Apesar de não existir uma sincronização contínua ou automática entre as instâncias, a plataforma EmonCMS dispõe de um módulo dedicado — *sync* — que permite, em caso de falha de uma das instâncias, recuperar os dados perdidos a partir da outra. Este módulo realiza a sincronização seletiva de *feeds* entre servidores distintos, permitindo restaurar medições ausentes e garantir a continuidade histórica dos dados, desde que uma das instâncias se mantenha operacional.

Importa referir que, ao contrário da instância principal alojada nos Serviços Informáticos do ISE, onde se encontrava inicialmente alojada a plataforma EmonCMS — a qual permanece acessível externamente para permitir o envio de dados provenientes de projetos localizados fora da rede interna — o espelho criado no laboratório 160 não apresenta qualquer exposição à internet pública. Este encontra-se totalmente isolado dentro da rede interna da universidade, servindo exclusivamente como sistema redundante e de testes locais. Assim, não se coloca neste caso a problemática associada à abertura externa da plataforma, uma vez que o seu acesso está restrito ao domínio interno, garantindo maior segurança e controlo sobre os dados armazenados.

A Figura 22 apresenta de forma esquemática esta nova arquitetura, evidenciando a coexistência das duas instâncias do sistema — a principal na máquina virtual dos Serviços Informáticos e a redundante no PC do laboratório 160 —, bem como os fluxos de comunicação estabelecidos com os medidores via Modbus/TCP/IP e com o sistema SolarLog através de API REST. A figura ilustra ainda a estrutura de ligação entre *gateways*, medidores e os servidores de recolha, evidenciando a replicação das funcionalidades essenciais entre as duas instâncias.

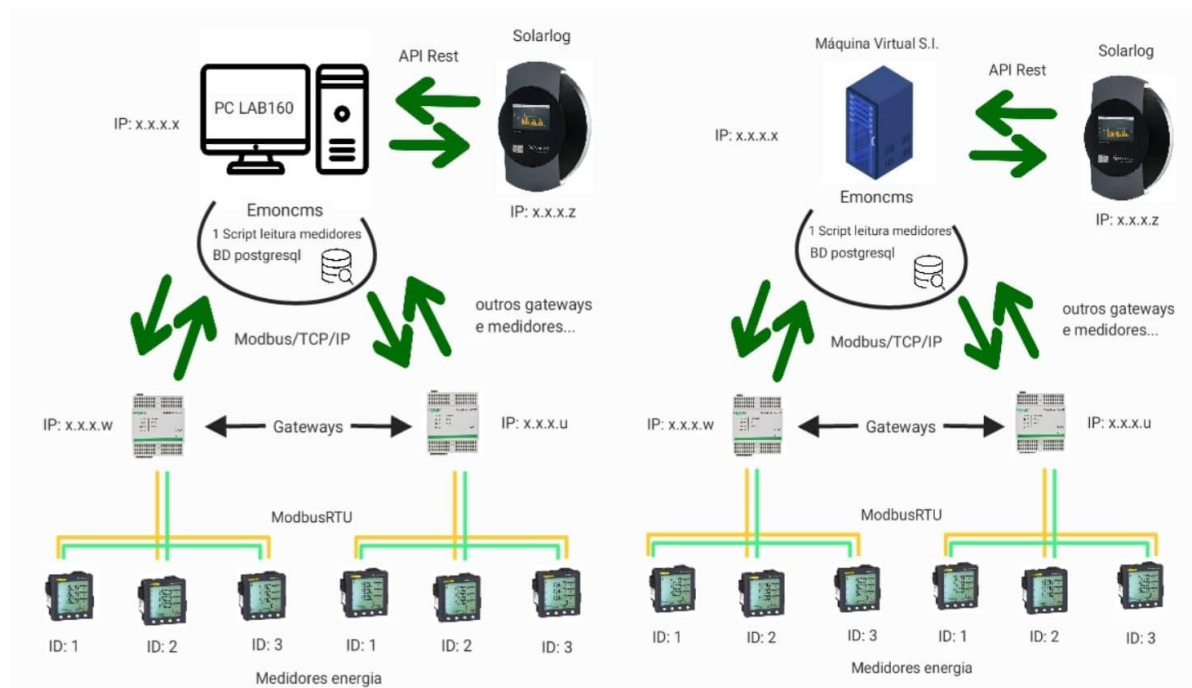


Figura 22 – Arquitetura na fase atual do sistema

3.3.1 REESTRUTURAÇÃO E OTIMIZAÇÃO DO CÓDIGO DE AQUISIÇÃO DE DADOS

Foi realizada uma reestruturação completa do código de aquisição de dados, que resultou numa versão mais robusta e compatível com as exigências atuais do sistema. O código anterior, que era uma junção de 3 códigos originalmente desenvolvidos entre 2019 e 2022, evoluiu o que levou à acumulação de estruturas heterogêneas, com abordagens distintas e por vezes redundantes na leitura e tratamento dos dados. Inicialmente, existiam três *scripts* distintos para diferentes zonas da infraestrutura. Estes foram unificados numa única estrutura durante a fase intermédia da arquitetura, onde foi também implementado o novo método de cálculo de custo energético. A presente versão representa a consolidação dessa transição, culminando numa lógica de recolha unificada e profundamente otimizada, mantendo a compatibilidade com os dispositivos existentes, mas com melhorias substanciais na fiabilidade da leitura, validação, cálculo e envio de variáveis.

O sistema continua a organizar-se de forma modular, com funções de leitura dedicadas a cada tipo de medidor Modbus (PM710, PM5100, TRV00210, entre outros). Contudo, essas funções foram alvo de uma revisão profunda. Uma das alterações estruturais mais relevantes consistiu na uniformização dos métodos de conversão dos registos binários. A versão anterior recorria a abordagens mistas e por vezes inconsistentes na transformação dos registos duplos Modbus de 16 *bits* em valores de 32 *bits* em formato de ponto flutuante, o que introduzia ambiguidade

e aumentava a probabilidade de erros difíceis de detetar, especialmente em dispositivos com ordens de *bytes* distintas.

Na versão atual, todas as conversões passaram a ser realizadas com recurso à biblioteca `struct`, utilizando chamadas explícitas que incluem a especificação da ordem dos bytes — por exemplo, `<HH` para ordem *little-endian* ou `>HH` para *big-endian*, consoante a arquitetura definida pelo fabricante. Esta padronização assegura uma interpretação inequívoca dos dados binários lidos, aumentando significativamente a fiabilidade da monitorização e a consistência entre os diferentes modelos de medidor. Além disso, a utilização sistemática desta biblioteca melhora a legibilidade e facilita a manutenção futura do código, ao eliminar construções inconsistentes anteriormente utilizadas e ao substituir o uso disperso de múltiplos métodos por uma abordagem única e coerente para a conversão de dados Modbus.

Para além da padronização dos processos de conversão, foi introduzido um conjunto de mecanismos transversais destinados a reforçar a fiabilidade da recolha. Uma das principais vulnerabilidades das versões anteriores residia na ausência de medidas preventivas contra leituras incompletas ou mal formatadas. Em situações de falha de comunicação com um dispositivo ou quando os registos retornavam valores vazios, *None* ou listas com tamanho inferior ao esperado, o sistema bloqueava temporariamente a execução da leitura desse dispositivo. Embora o ciclo de recolha continuasse funcional para os restantes medidores, estes bloqueios criavam atrasos e reduziam a previsibilidade da execução.

Para mitigar este problema foram implementadas duas camadas complementares de proteção. A primeira consiste numa verificação prévia de conectividade, realizada através de um pedido de *ping* ao endereço IP de cada medidor antes da tentativa de leitura dos registos Modbus. Se o dispositivo não responder dentro de três segundos, é automaticamente considerado indisponível e o ciclo de aquisição prossegue para o medidor seguinte, prevenindo tempos de espera excessivos em equipamentos desligados ou em avaria.

O mecanismo anteriormente existente, baseado apenas em instruções `try/except` durante a tentativa de obtenção das variáveis, foi assim complementado por uma camada adicional de robustez, permitindo reduzir atrasos e melhorar a previsibilidade da execução global do *script*. A segunda camada de proteção corresponde a uma função de verificação genérica chamada imediatamente após o recebimento de cada leitura. Esta função deteta automaticamente dados ausentes ou inválidos, substituindo-os por listas preenchidas com zeros. Esta abordagem impede que os dados inválidos sejam submetidos à conversão binária — processo que anteriormente podia provocar bloqueios — e assegura a continuidade do fluxo de execução.

Ainda que a substituição por zero não corrija a origem da falha de comunicação, permite atenuar os efeitos negativos de perdas esporádicas de pacotes Modbus.

Adicionalmente, esta substituição por zero assegura que os pacotes JSON enviados para o EmonCMS são válidos, visto que esta plataforma rejeita variáveis com valores *null* ou *NaN*. Assim, o valor zero é utilizado como marcador visual de possível falha de leitura, permitindo a sua visualização na interface gráfica sem interromper a recolha. Contudo, é importante que os utilizadores do sistema interpretem estes zeros com precaução, distinguindo entre valores nulos provenientes de erros de comunicação e zeros legítimos de operação. Por exemplo, circuitos de iluminação exterior frequentemente apresentam zero durante o dia por estarem desligados, e sistemas fotovoltaicos naturalmente registam produção nula durante o período noturno. A presença de zeros em momentos ou contextos inesperados torna-se assim um indicador útil para deteção de anomalias, sem comprometer a integridade estrutural do sistema. A reestruturação estendeu-se também à organização interna do código. Enquanto as funções anteriores, devido à acumulação progressiva de funcionalidades, chegavam a ultrapassar as 200 a 300 linhas por função, na versão atual essas foram divididas, modularizadas e reorganizadas de forma lógica e mais concisa, com blocos funcionais que raramente ultrapassam as 100 linhas. Como reflexo global deste processo, o tamanho total do *script* principal foi reduzido de cerca de 1350 linhas para aproximadamente 750 linhas, sem comprometer as funcionalidades essenciais. Esta simplificação facilita a manutenção, permite atualizações localizadas e reduz significativamente a complexidade de análise do código.

Por fim, um dos reflexos mais notórios desta reestruturação, aliado à reformulação do cálculo de custo energético já discutida na fase intermédia (Secção 3.2.1), foi a drástica redução do tempo total de execução do *script*. Na versão anterior, o tempo médio de execução oscilava entre 45 a 55 segundos, próximo do limite de 60 segundos definido pelo intervalo de amostragem. Este limite é crítico, uma vez que a sobreposição de execuções pode causar interferência entre instâncias concorrentes do *script*. Com a nova versão, o tempo de execução reduziu-se para um intervalo entre 10 e 20 segundos. Este resultado foi validado com recurso à biblioteca *time*, medindo o tempo de início e fim da execução do *script*, assegurando que todo o processo decorre dentro dos limites temporais definidos.

No seu conjunto, estas alterações refletem uma maturação do código e da arquitetura lógica subjacente, transpondo o sistema de um modelo funcional, mas frágil, para uma solução mais resiliente, coerente e otimizada, capaz de lidar com falhas pontuais de forma controlada, e com melhor desempenho geral em ambientes de monitorização energética em larga escala.

3.3.2 CONSOLIDAÇÃO E ESTRUTURAÇÃO DA NOVA BASE DE DADOS ENERGÉTICA

Na fase atual do sistema, foi tomada a decisão estratégica de abandonar as duas bases de dados históricas herdadas da arquitetura inicial, que apresentavam inconsistências estruturais, lacunas temporais e uma organização pouco uniforme das variáveis. Embora funcionais, essas bases haviam evoluído de forma não estruturada, dificultando a manutenção, a análise sistemática e a integração com outras camadas do sistema. Em alternativa, optou-se pela criação de uma nova base de dados, desenvolvida com uma estrutura unificada e otimizada, alinhada com a realidade atual da monitorização energética e com os objetivos de exploração futura. Esta base foi implementada em PostgreSQL — tecnologia já anteriormente utilizada no sistema — selecionada pelas suas reconhecidas vantagens em termos de desempenho, escalabilidade e capacidade de lidar com acessos concorrentes em ambientes com elevado volume de dados. A nova base foi estruturada para conter apenas as variáveis essenciais ao acompanhamento energético e à análise técnica suportada pelo sistema, com especial enfoque nas que já estavam integradas nos *feeds* da plataforma EmonCMS. Foram assim mantidas colunas correspondentes às grandezas fundamentais: correntes por fase (A, B e C), tensões de fase (A, B e C), potência ativa, potência reativa, custo instantâneo e custo acumulado. As bases de dados anteriores incluíam outras variáveis — como mínimos e máximos, fator de potência, frequência e grandezas entre fases — que, apesar de tecnicamente relevantes, acrescentavam complexidade estrutural sem retorno prático direto. Estas variáveis não estavam a ser utilizadas nos *dashboards* nem na lógica de tratamento, e a sua análise sistemática não era realizada. Além disso, apresentavam lacunas, falhas pontuais e inconsistências temporais que comprometeram a sua fiabilidade. Importa ainda referir que estas variáveis adicionais não eram armazenadas no EmonCMS, o que impediu a obtenção de versões atualizadas durante o processo de migração. Esta limitação reforçou a decisão de centrar a nova base de dados nas variáveis com maior percentagem de registos válidos, provenientes do sistema de aquisição ativo.

Com este propósito, cada *feed* guardado no EmonCMS foi convertido individualmente para o formato CSV, utilizando um *script* desenvolvido em Python com recurso à biblioteca `emon-tools`, que dispõe de funções específicas para interpretar os ficheiros binários do EmonCMS. Foi assim possível extrair os dados armazenados em ficheiros `.meta` e `.data` do diretório `PHPFina` — que se encontravam mais atualizados e com menos lacunas do que os registos presentes nas bases anteriores — resultando num ficheiro por cada variável medida em cada edifício, pronto a ser inserido na nova base de dados.

Para proceder ao preenchimento da base, foi criado um segundo *script* em Python que automatiza todo o processo de criação e inserção de dados. A estrutura adotada para a nova base organiza os dados em tabelas distintas por edifício. Os nomes atribuídos às tabelas e colunas seguem a nomenclatura padronizada previamente definida nas fases anteriores do projeto, assegurando consistência com o restante sistema e facilitando a integração com os processos de análise e visualização já existentes.

Cada tabela inclui um campo *timestamp* como chave primária e colunas para cada variável energética.

A nova base representa atualmente a fonte mais fiável e completa de dados energéticos do sistema, integrando medições de potência, tensão, corrente, custo e geração elétrica ao longo de vários anos de operação. Muitos dos *feeds* históricos apresentam volumes significativos de informação, com algumas variáveis a ultrapassar os três milhões de registos individuais, o que justifica a adoção de um motor de base de dados relacional robusto e otimizado como o PostgreSQL.

A organização dos dados em tabelas por edifício permite que cada medição realizada pelo sistema seja imediatamente guardada na respetiva tabela. Esta abordagem assegura ainda que o armazenamento em tempo real decorre sem comprometer a integridade nem a acessibilidade dos registos históricos, permitindo múltiplos acessos concorrentes sem perda de desempenho. Apesar da nova base de dados já se encontrar funcional e preenchida com um volume elevado de dados históricos — extraídos, convertidos e estruturados com o apoio da biblioteca `emon-tools` — foi ainda necessário um processo contínuo de observação, verificação e tratamento das séries temporais armazenadas. Este processo contínuo foi essencial para assegurar a integridade, consistência e utilidade futura da base, especialmente no contexto de aplicações de aprendizagem automática. A sua utilização como *dataset* para modelos de *machine learning* depende da qualidade dos dados recolhidos e da sua granularidade temporal.

Para assegurar a escrita contínua dos dados recolhidos na nova base, foi desenvolvida uma função dedicada, responsável por inserir automaticamente as medições nas tabelas correspondentes. Esta função, integrada no *script* principal de aquisição, garante que, a cada novo ciclo de leitura, os dados sejam gravados na base PostgreSQL com o *timestamp* correto, ajustado ao fuso horário do sistema. A lógica implementada adapta-se dinamicamente à estrutura de cada tabela, permitindo o armazenamento de variáveis energéticas convencionais, bem como de configurações específicas, como as tabelas agregadas de geração fotovoltaica. Esta integração direta entre aquisição e armazenamento assegura a consistência do fluxo de dados e simplifica o processo de atualização contínua da base.

3.3.3 CUSTO ACUMULADO MENSAL E MECANISMO DE REINICIALIZAÇÃO

Nesta fase, o sistema realiza o cálculo automático do custo energético acumulado por edifício, atualizado a cada medição com base nos valores de custo instantâneo, obtidos segundo um tarifário tetra-horário ajustado à estação do ano e ao período do dia. Estes valores são somados continuamente, originando uma função crescente ao longo de cada mês, armazenada num *feed* específico na plataforma de monitorização, sendo esta representação essencial para a análise direta do impacto financeiro dos consumos energéticos.

Para garantir que o custo acumulado reflete exclusivamente o ciclo mensal em vigor, foi implementado um mecanismo automático de reinicialização, responsável por repor os *feeds* de custo a zero no início de cada mês. Este processo é assegurado por um script Python, agendado via `crontab` para execução diária às 00:00. Em cada execução, o script consulta um ficheiro `.txt` contendo a data do último *reset*. Caso o mês registado seja diferente do mês atual, procede-se à reinicialização de todos os *feeds* de custo acumulado.

Esta estratégia assegura a robustez do sistema perante falhas temporárias, como interrupções no fornecimento elétrico ou manutenções inesperadas. Se o servidor estiver inativo no momento previsto para o *reset* — por exemplo, no dia 1 —, o script continuará a ser executado nos dias seguintes. Assim que o sistema retomar a operação, a ausência de reinicialização no mês corrente será detetada e o processo será automaticamente executado, preservando a consistência da função acumulada. A nova data é então registada no ficheiro `.txt` de controlo, impedindo duplicações e garantindo a fiabilidade do processo.

3.3.4 DESENVOLVIMENTO E PLANEAMENTO DE *DASHBOARDS* INTERATIVOS

Com a nova base de dados estruturada e funcional, foram criadas as condições necessárias para o desenvolvimento de *dashboards* interativos mais avançados, com maior capacidade de análise e personalização do que os painéis anteriormente disponíveis. Embora os *dashboards* do EmonCMS cumpram um papel relevante na visualização inicial dos consumos energéticos, a sua flexibilidade e capacidade gráfica são limitadas, sobretudo no que respeita à apresentação de métricas derivadas, cruzamento de variáveis e configuração de visualizações por utilizador. Neste contexto, foi instalado localmente o Grafana no servidor físico do laboratório 160, aproveitando a sua capacidade nativa para se conectar à base de dados PostgreSQL. A ligação direta entre o Grafana e a base de dados PostgreSQL viabiliza a construção de visualizações interativas avançadas, com capacidade para aplicar filtros temporais, efetuar agregações estatísticas, definir alarmes e realizar comparações ao longo do tempo. Este tipo de interface

potência uma análise mais aprofundada e orientada tecnicamente dos dados energéticos, ajustada às necessidades de diferentes contextos de utilização. Paralelamente, o facto de a instalação estar localizada no servidor redundante permite explorar funcionalidades experimentais e desenvolver trabalho em ambiente controlado, sem risco para o sistema produtivo principal.

A estrutura modular da nova base de dados, organizada por edifício, facilita a construção de painéis direcionados a cada local monitorizado. Assim, foi iniciado o planeamento para a instalação de dispositivos Raspberry Pi com ecrãs nas entradas principais dos edifícios. Estes dispositivos estarão ligados à rede interna da Universidade do Algarve e configurados para aceder diretamente à instância local do Grafana, permitindo apresentar *dashboards* específicos com os consumos, custos e produções energéticas em tempo real para cada edifício. Para tal, cada Raspberry Pi será preparado com um sistema operativo baseado em Linux — como o Raspberry Pi OS — e configurado para arrancar automaticamente com um browser em modo quiosque (*kiosk mode*), abrindo a página web correspondente ao *dashboard* do edifício onde estiver instalado. Esta solução garante uma visualização contínua, sem necessidade de intervenção do utilizador.

Esta componente visual, ilustrada na Figura 23, tem como objetivo não apenas aumentar a transparência da informação energética, mas também reforçar a sensibilização da comunidade académica para comportamentos sustentáveis. A presença constante de informação atualizada sobre o desempenho energético de cada espaço incentiva à consciencialização coletiva e à adoção de práticas mais eficientes.

Paralelamente, a capacidade de criar *dashboards* personalizados por tipo de utilizador ou contexto abre caminho à introdução de mecanismos de gamificação. Estes mecanismos podem incluir, por exemplo, comparações de desempenho energético entre edifícios, desafios mensais para redução de consumos, ou o destaque de boas práticas adotadas por determinada unidade. Ao transformar o consumo energético num indicador visível e dinâmico, cria-se uma oportunidade para envolver os utilizadores numa lógica participativa e motivacional, promovendo uma cultura institucional de eficiência.

A integração entre a base de dados, a plataforma de visualização e os dispositivos de exibição distribuídos representa um avanço significativo na arquitetura do sistema. Ao disponibilizar dados atualizados, acessíveis e adaptados aos diferentes contextos de utilização, promove-se uma abordagem mais consciente, colaborativa e tecnicamente sustentada à gestão energética nos espaços da universidade.

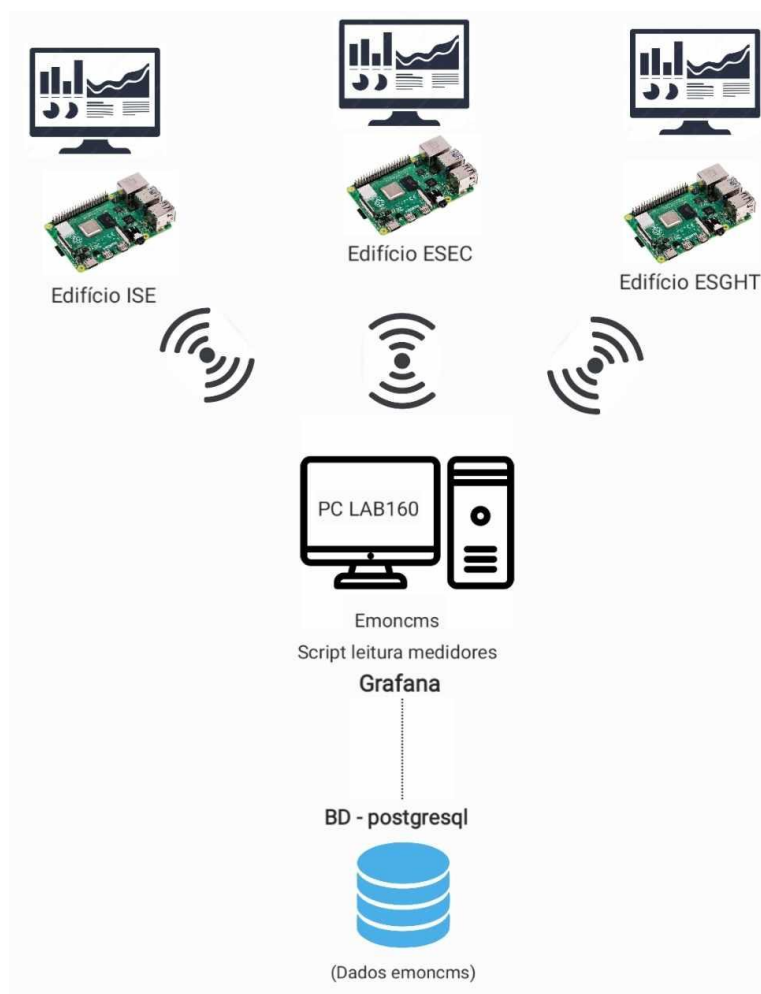


Figura 23 – Interligação entre Grafana e PostgreSQL com dashboards interativos

3.3.5 PARTICIONAMENTO DE MEDIÇÕES NO INTERIOR DOS EDIFÍCIOS MONITORIZADOS

Com a generalização da monitorização energética nos principais edifícios da Universidade do Algarve, tornou-se pertinente iniciar o planeamento de uma nova fase centrada na partição interna das medições, com o objetivo de obter uma caracterização mais detalhada do consumo dentro de cada edifício. Esta abordagem visa reforçar a capacidade de identificação de consumos excessivos ou comportamentos anómalos, segmentando a informação por áreas funcionais ou setores específicos, como departamentos, serviços administrativos ou infraestruturas técnicas.

O Edifício Principal do ISE, situado no campus da Penha, foi selecionado como caso piloto para este tipo de abordagem, não só por ter sido o ponto de partida do projeto de monitorização, mas também pela sua centralidade na estrutura do campus e pela acessibilidade técnica dos quadros elétricos. A intenção subjacente à sua escolha foi a de avaliar o impacto e a viabilidade

técnica do particionamento interno, antes de proceder à eventual replicação noutros edifícios da instituição.

Foi realizado um levantamento técnico detalhado dos quadros elétricos existentes e dos respetivos setores funcionais que poderiam ser individualmente monitorizados. Com base nesse estudo, foram definidos como pontos prioritários para futura instrumentação os Quadros Parciais dos Pisos 1 e 2, os Quadros de AVAC correspondentes, os Quadros da Secretaria, do Bar da AAUAlg, dos Serviços Académicos e o quadro do *Datacenter*, este último já com medidor previamente instalado. A escolha destes locais teve por base a sua relevância funcional e o potencial impacto no consumo global do edifício.

Para a execução da partição foi proposta a utilização de dispositivos Shelly Pro 3EM, adequados à medição trifásica, compatíveis com a arquitetura do sistema e com capacidade de comunicação via Ethernet ou Wi-Fi. Este modelo permite a utilização de transformadores de corrente externos com capacidade de medição de correntes até 120 A, o que é mais do que suficiente para os circuitos identificados nos edifícios da universidade. A sua instalação permitiria acompanhar separadamente os consumos por zona, oferecendo dados relevantes para a tomada de decisões ao nível da gestão energética interna.

Apesar de todo o planeamento e levantamento técnico terem sido concluídos, a implementação efetiva do particionamento encontra-se pendente de validação e decisão por parte da universidade, nomeadamente no que respeita ao orçamento e à priorização de recursos. Este processo poderá, no entanto, avançar numa fase posterior, estando já criada a base técnica necessária para a sua concretização futura.

3.3.6 PREVISÃO METEOROLÓGICA ATRAVÉS DE API

Neste sistema dispõe-se de dados meteorológicos em tempo real, recolhidos por um *script* dedicado a partir da estação meteorológica instalada no edifício principal do ISE, complementados por um histórico contínuo armazenado em PostgreSQL. Contudo, para suportar algoritmos de previsão de consumo e produção, tornou-se necessário acrescentar ao sistema um processo sistemático de obtenção de previsões meteorológicas de curto prazo. O objetivo é fornecer variáveis exógenas fiáveis para modelos de previsão de potência ativa, assegurando coerência com o fuso horário local.

Face a estes requisitos, foram analisadas algumas alternativas e optou-se por uma solução baseada num serviço online de previsões, acessível via API. Este tipo de abordagem permite parametrizar a localização geográfica, o fuso horário, a cadência temporal e as unidades físicas

diretamente no pedido, recebendo uma resposta estruturada por vetor de instantes e séries paralelas por variável.

A solução selecionada foi a Open-Meteo, uma API pública e gratuita para uso não comercial, sem necessidade de chave de autenticação, que disponibiliza previsões geradas por modelos de previsão numérica amplamente usados e reconhecidos. Indicando no pedido HTTP as coordenadas do local de interesse, a API seleciona os modelos adequados à localização e devolve as variáveis solicitadas com parametrização explícita de fuso horário e de unidades, respeitando a resolução espacial do modelo subjacente. As séries podem ser obtidas em várias cadências temporais, o que facilita o alinhamento com diferentes casos de uso.

Quanto ao horizonte temporal, a API permite que sejam efetuados pedidos de diferentes janelas de previsão de curto a longo prazo. Nesta implementação adotou-se um horizonte de 24 horas, pois permite rever diariamente as previsões do dia seguinte aproveitando as atualizações intradiárias dos modelos e evita o aumento desnecessário da incerteza associado a horizontes mais longos.

O pedido a esta API foi parametrizado para as coordenadas da cidade de Faro, com fuso horário local e taxa de amostragem de quinze minutos. As variáveis meteorológicas escolhidas para obtenção foram a temperatura do ar, a humidade relativa, a pressão atmosférica, a velocidade e a direção do vento, a precipitação, a radiação solar direta e a nebulosidade. Para manter consistência dimensional e evitar conversões posteriores, as unidades foram definidas na própria chamada, assegurando compatibilidade direta com as escalas utilizadas no resto do sistema. A resposta a este pedido é devolvida em formato JSON com um vetor de instantes e vetores paralelos por variável, o que simplifica a ingestão dos dados obtidos.

Para efetuar estes pedidos um novo *script* em Python foi desenvolvido, este desempenha a tarefa de consultar a API Open-Meteo, validar e normalizar os registos, encaminhá-los para a plataforma EmonCMS e, por fim, armazená-los na tabela *meteo_forecast* na base de dados.

Para garantir persistência de operação o *script* é acionado uma vez por dia via *crontab*. O resultado é uma série de previsões contínua e coerente, atualizada diariamente, pronta para confronto com as observações da estação meteorológica e para uso pelos modelos de aprendizagem automática, sem desalinhamentos temporais. A Figura 24 apresenta a interação entre o sistema de monitorização e a API Open-Meteo, ilustrando o envio de pedidos HTTP e a receção das respostas utilizadas no processo de obtenção de dados de previsão meteorológica.

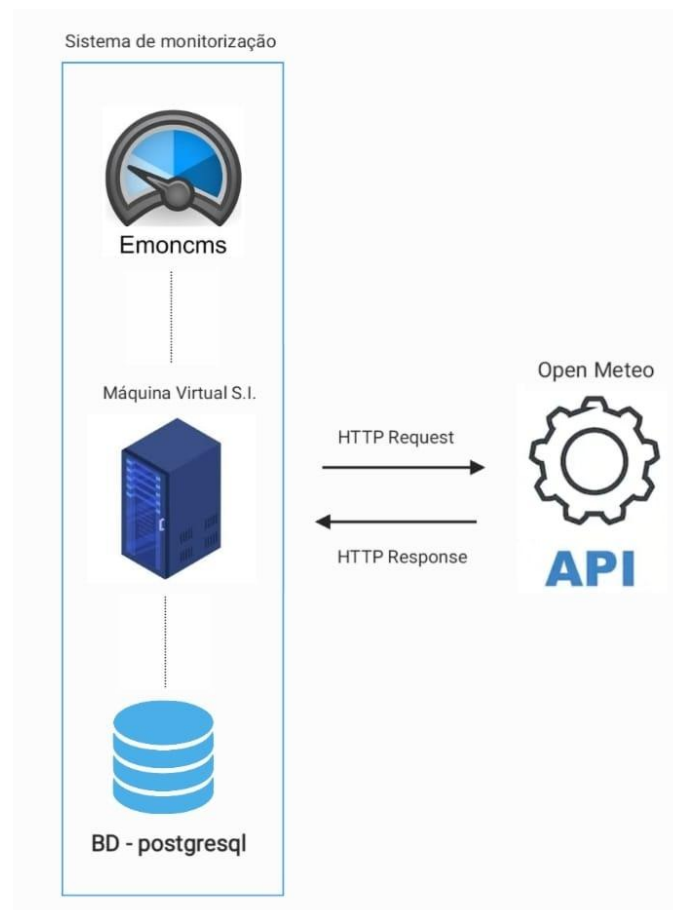


Figura 24 – Integração entre sistema de monitorização e API meteorológica

3.4 MACHINE LEARNING PARA PREVISÃO ENERGÉTICA

A aplicação de técnicas de aprendizagem automática em sistemas de monitorização energética permite antecipar padrões de consumo e identificar desvios operacionais com impacto direto na eficiência da gestão energética. No âmbito deste trabalho, explorou-se a modelação de séries temporais aplicada à previsão da potência ativa em edifícios monitorizados, bem como à identificação de comportamentos de consumo anómalos, caracterizados por desvios significativos face aos perfis históricos esperados. Como ponto de partida, adotou-se o modelo Prophet, apresentado na Secção 2.9.2, como modelo de referência para a previsão, decisão que condicionou a forma como o pré-processamento foi definido em função dos requisitos específicos do modelo. Sendo o modelo Prophet compatível com variáveis regressoras externas, a sua utilização exige não só a disponibilização dessas variáveis ao longo de todo o conjunto de dados de treino, mas também ao longo de todo o horizonte de previsão. Este requisito impõe uma articulação rigorosa entre a preparação dos dados, a calendarização das sazonalidades, a obtenção de previsões para as regressoras e a construção de um conjunto de dados de entrada consistente. Esta opção requer um planeamento cuidadoso das fontes de dados

e das etapas de limpeza, imputação e alinhamento temporal, garantindo coerência entre a granularidade das medições, a janela histórica considerada e o horizonte futuro de previsão. Para a detecção de anomalias abordou-se um processo em dois níveis. No primeiro nível é calculado um *score* de decisão a partir dos resíduos entre a potência observada e a prevista e no segundo, aplica-se o modelo *Isolation Forest*, apresentado na Seção 2.9.3, para refinar a triagem, atuando como filtro adicional sobre as detecções propostas pelo *score*.

3.4.1 ANÁLISE PRELIMINAR DOS DADOS

Antes da aplicação de qualquer modelo de previsão, torna-se fundamental avaliar a qualidade dos dados disponíveis, dado que a fiabilidade das previsões depende diretamente da consistência das séries temporais utilizadas como entrada. Com este propósito, foi elaborada a tabela apresentada no Apêndice A, a qual sintetiza a análise preliminar da integridade dos registos por edifício. Esta tabela apresenta o número total de registos existentes em cada tabela da base de dados, o número de amostras em falta, o número de amostras inválidas e a percentagem de registos considerados válidos no intervalo temporal existente. A análise evidenciou que, embora a maioria dos medidores apresente taxas de validade superiores a 90%, existem casos com percentagens significativamente inferiores, como alguns edifícios nos quais se verificaram falhas estruturais de recolha e registos inválidos em larga escala. Esta avaliação inicial revelou-se determinante não só para orientar o processo de pré-processamento, mas também para a escolha de um edifício piloto sobre o qual aplicar as primeiras previsões de consumo, garantindo que o modelo fosse validado em condições de dados suficientemente consistentes.

3.4.2 CONSTRUÇÃO DOS CONJUNTOS DE DADOS

Após a análise preliminar, foi selecionado o edifício `gel_fct_cintal_edificio_1`, por apresentar uma elevada percentagem de registos válidos e um perfil de consumo considerado representativo. A partir desta seleção construíram-se quatro conjuntos de dados explicitamente orientados para a aplicação do modelo Prophet, partindo do princípio de que o pré-processamento deve ser parametrizado pelas limitações e necessidades do algoritmo adotado. Três desses conjuntos correspondem ao conjunto energético histórico, ao conjunto meteorológico histórico e ao conjunto de previsões meteorológicas, todos extraídos diretamente da base de dados PostgreSQL. O conjunto energético histórico, com amostragem

de um minuto, é constituído pelos registos de consumo do edifício, enquanto o conjunto meteorológico histórico agrega as variáveis ambientais medidas pela estação meteorológica instalada na universidade, também com amostragem de um minuto. O conjunto de previsões meteorológicas tem origem na tabela da base de dados que diariamente armazena os dados obtidos através da API externa Open-Meteo, apresenta amostragem de quinze minutos e foi harmonizado para conter exatamente as mesmas variáveis exógenas do conjunto meteorológico histórico, assegurando compatibilidade e disponibilidade ao longo do horizonte de previsão. O quarto conjunto, relativo ao calendário académico, foi definido manualmente em código, estruturado com intervalos de quinze minutos e enriquecido com indicadores que assinalam períodos letivos, épocas de exames, pausas e eventos institucionais relevantes. Para a sua construção, recorreu-se aos calendários letivos oficiais da Universidade do Algarve para os anos académicos de 2018/2019 até 2025/2026, garantindo a inclusão de informação de calendário relativa ao passado e ao futuro.

3.4.3 PRÉ-PROCESSAMENTO E LIMPEZA

Na preparação dos quatro conjuntos de dados optou-se por excluir variáveis energéticas que, embora presentes nas tabelas originais, não eram adequadas à aplicação do modelo. Tendo sido selecionado o modelo Prophet, impôs-se a exigência de que todas as regressoras externas estivessem disponíveis ao longo do horizonte de previsão. Nesse contexto, grandezas sem valores futuros conhecidos foram consideradas inadequadas, pois a sua utilização implicaria estimá-las previamente, acumulando erro e potenciando fuga de informação. Assim, no *dataset* energético manteve-se apenas a potência ativa como variável alvo e eliminaram-se correntes, tensões, potência reativa e custos energéticos.

Para sustentar a seleção das variáveis exógenas meteorológicas efetuou-se um estudo de correlação no conjunto de treino, calculando a correlação de Spearman entre a potência ativa e cada variável externa ao longo de múltiplos desfasamentos temporais e considerando todas as observações disponíveis no período de análise [92]. A análise em diferentes desfasamentos teve como objetivo captar atrasos físicos e operacionais entre a evolução das condições meteorológicas e a resposta de consumo do edifício. Uma subida ou descida da temperatura exterior não se reflete de forma imediata na temperatura interior devido à inércia térmica da envolvente e a ativação dos sistemas de aquecimento ou ar condicionado pode ocorrer apenas algum tempo depois. Nestes casos a potência ativa pode apresentar maior correlação com valores passados das variáveis meteorológicas do que com o valor observado no próprio

instante.

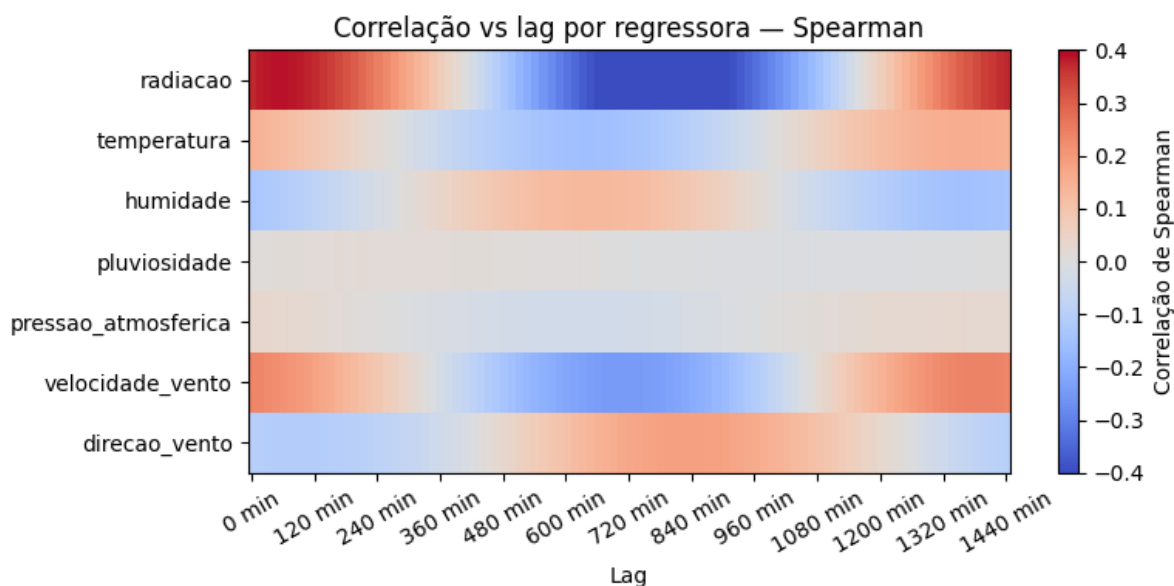


Figura 25 – Mapa de calor da correlação de Spearman em função do lag

A Figura 25 mostra um mapa de calor com uma linha por variável e o desfasamento no eixo horizontal. A cor indica a força e o sinal da associação. Tons quentes significam correlação positiva. Tons frios significam correlação negativa. Zonas próximas do neutro indicam pouca relação. A leitura faz-se procurando em cada linha os desfasamentos onde a intensidade é mais elevada em valor absoluto. Esses pontos sugerem atrasos onde a variável é mais informativa para a potência ativa. Padrões em faixa que se repetem ao longo do dia revelam a presença de ciclos diários partilhados pelas séries. Nesta fase o mapa foi utilizado com um propósito sobretudo exploratório e interpretativo, servindo para evidenciar a importância de estudar as regressoras externas ao longo de diferentes desfasamentos temporais e para apoiar a escolha das variáveis a incluir no modelo. Apesar desta análise com desfasamentos, na modelação implementada neste trabalho as variáveis exógenas foram incluídas sem qualquer desfasamento temporal, sendo utilizadas no mesmo instante que a potência ativa, ficando o estudo sistemático de *lags* específicos como direção natural para trabalhos futuros.

Optou-se por Spearman por captar associações monotónicas mesmo quando não lineares e por ser mais robusto a valores extremos, uma vez que assenta na ordenação dos dados e não na sua amplitude. No cálculo destas correlações a direção do vento, originalmente expressa em graus, foi convertida em duas componentes contínuas que representam o seno e o cosseno do ângulo, o que permite traduzir a natureza circular desta variável e avaliar de forma coerente a sua relação com a potência ativa e com as restantes séries meteorológicas. No mapa observa-

se um padrão diário claro e uma parte relevante dessa associação decorre da estrutura temporal comum da potência ativa e das variáveis meteorológicas ao longo do dia e da semana. Optou-se por incluir na modelação como regressoras exógenas a radiação, a temperatura, a humidade e a pluviosidade por motivos específicos. A radiação capta desvios ao padrão diário associados à nebulosidade e pode influenciar de forma imediata a carga através da iluminação. A temperatura e a humidade refletem efeitos térmicos que podem influenciar os sistemas de climatização. A pluviosidade, embora rara em Faro e com correlação global próxima de zero, pode ser útil de forma pontual como indicador de ocorrência para captar impactos localizados em períodos específicos. Em contraste, a velocidade e a direção do vento e a pressão atmosférica foram excluídas por se revelarem redundantes face às restantes variáveis e por não evidenciarem, no contexto analisado, uma relação operacional direta com a variação da carga elétrica no edifício, apesar de terem sido consideradas na análise exploratória representada no mapa de calor.

O modelo Prophet padroniza as exógenas e aplica regularização aos seus coeficientes, reduzindo automaticamente o impacto de variáveis com baixo sinal informativo. Assim, mesmo quando pouco relevantes, estas variáveis não devem diminuir a qualidade da previsão. O *dataset* de previsões meteorológicas foi harmonizado para conter exatamente estas quatro variáveis exógenas, assegurando compatibilidade com o histórico. Por fim, o quarto conjunto, correspondente ao calendário académico, foi definido através da criação de intervalos representativos de períodos letivos, épocas de exames e pausas. A partir desses intervalos construiu-se uma tabela de feriados e eventos, com as colunas `ds` e `holiday`, que é fornecida diretamente ao modelo como parâmetro *holidays*. Esta abordagem integra nativamente os efeitos de calendário no processo de previsão e preserva a coerência temporal ao longo do horizonte considerado.

Esta função automatiza três etapas essenciais: em primeiro lugar, normaliza os registos temporais para uma representação temporal consistente, assegurando uma referência uniforme ao longo dos diferentes conjuntos de dados, em segundo lugar, ordena as entradas e elimina duplicados, preservando apenas o último valor em caso de conflito e, por fim, valida as colunas existentes, retendo apenas as variáveis especificadas. O resultado obtido foram *datasets* cronologicamente coerentes e com as variáveis de interesse devidamente identificadas.

Numa segunda etapa de tratamento, foi aplicada uma função dedicada à limpeza de valores inválidos e lacunas. No caso do *dataset* energético, removeram-se todos os registos da variável alvo com valores nulos ou iguais a zero. Esta decisão visou evitar a introdução de estimativas artificiais na variável alvo, garantindo que apenas dados reais serviam de base ao modelo. O

modelo Prophet não realiza imputação automática da variável alvo, pelo que observações sem nenhum valor são, na prática, excluídas do processo de ajuste. A criação de valores artificiais para preencher essas lacunas poderia distorcer a estimação dos parâmetros. Além disso, a análise exploratória indicou que, para o edifício em estudo, valores exatamente iguais a zero correspondiam maioritariamente a falhas de medição ou a interrupções pontuais na aquisição de dados, e não a situações físicas de ausência de consumo, o que reforça a opção por descartá-los em vez de os imputar. No *dataset* meteorológico histórico, a função recorreu a um processo de interpolação temporal controlada para preencher falhas de curta duração, complementado por técnicas de *forward fill* e *backward fill*. No primeiro caso, valores em falta são substituídos pelo último valor válido anterior, enquanto no segundo são preenchidos com o valor válido seguinte, assegurando a continuidade das séries em períodos de pequenas falhas pontuais. Por fim, eliminaram-se registos residuais que permaneciam incompletos após este processo, de modo a garantir que apenas séries totalmente válidas prosseguiam para as fases seguintes da análise.

Posteriormente, os conjuntos de dados foram transformados numa estrutura diretamente compatível com o modelo de previsão, composta pela coluna temporal *ds*, a variável alvo *y* e pelas variáveis meteorológicas. Os valores de todas as colunas foram arredondados para duas casas decimais e as séries reamostradas para intervalos de quinze minutos, utilizando a média como critério de agregação. Esta decisão teve duas motivações principais. Em primeiro lugar, atenuar a variabilidade dentro de cada intervalo e suavizar o perfil de carga face às leituras minuto a minuto, preservando simultaneamente a contribuição relativa de picos e vales dentro de cada bloco de quinze minutos. Em segundo lugar, assegurar a compatibilidade com as variáveis meteorológicas previstas, cuja disponibilidade futura se limita precisamente a registos de quinze em quinze minutos. Optou-se pela média, em vez da mediana, por esta preservar informação sobre o nível médio de potência em cada intervalo e manter a proporcionalidade com a energia consumida nesse período. Desta forma, energia e meteorologia passaram a partilhar a mesma cadência temporal, condição indispensável para que as séries pudessem ser combinadas num único quadro de análise e exploradas pelo modelo em questão. A Tabela 9 ilustra três registos do *dataset* preparado para o modelo Prophet, após limpeza, arredondamento e reamostragem.

Tabela 9 – Exemplo de três registos do dataset preparado para o modelo Prophet – *ds* (timestamp), *y* (potência ativa), radiação (W/m_2), temperatura ($^{\circ}C$), humidade (%), pluviosidade (binária 0/1)

<i>ds</i>	<i>y</i>	radiação	temperatura	humidade	pluviosidade
2025-09-23 15:00:00	40.55	740	27.54	31	0
2025-09-23 15:15:00	42.35	735	27.43	31.5	0
2025-09-23 15:30:00	41.25	730	27.2	33	0

3.4.4 ESTRATÉGIA DE TREINO E OTIMIZAÇÃO

Com os conjuntos de dados consolidados e preparados para integração no modelo, passou-se à fase de treino e calibração do algoritmo. Para este fim, foi adotada uma estratégia em duas etapas que combina otimização Bayesiana com reavaliação em validação temporal. Numa primeira etapa, recorreu-se à biblioteca Mango[93] para explorar um espaço de hiperparâmetros que cobrisse os principais fatores de flexibilidade do modelo. A definição deste espaço de procura possibilitou avaliar de forma estruturada o equilíbrio entre a suavidade da tendência, a complexidade das sazonalidades e o grau de regularização aplicado às regressoras. Foram incluídos parâmetros ligados à tendência, às sazonalidades, às regressoras exógenas e ao modo de combinação de cada componente. No que respeita à tendência, considerou-se a intensidade da regularização aplicada às quebras (`changepoint_prior_scale`), a fração do histórico em que estas podem ocorrer (`changepoint_range`) e o número máximo de pontos de mudança candidatos (`n_changepoints`). Estes hiperparâmetros são determinantes para controlar o grau de suavidade ou de adaptabilidade da curva de tendência. Ao nível das sazonalidades, testaram-se as ordens de Fourier diária, semanal e anual (`daily_fourier`, `weekly_fourier`, `yearly_fourier`), que determinam a granularidade da representação de padrões periódicos, bem como as respetivas escalas de regularização (`daily_prior_scale`, `weekly_prior_scale`, `yearly_prior_scale`), responsáveis por limitar a amplitude permitida de cada componente sazonal. A escala associada a feriados (`holidays_prior_scale`) foi igualmente incluída, de modo a ajustar o impacto de datas especiais na série temporal.

Foram também integrados hiperparâmetros relacionados com regressoras externas, em particular a escala de regularização (`regressor_prior_scale`) e o modo de combinação (`regressor_mode`), permitindo calibrar o peso e a forma de contribuição destas variáveis

adicionais no modelo. Foi definido o modo aditivo ou multiplicativo das sazonalidades de maior frequência (`daily_mode`, `weekly_mode`), refletindo diferentes formas de interação com o nível da série, e um modo aditivo para a sazonalidade anual (`yearly_mode`).

O espaço de procura dos hiperparâmetros baseou-se nos valores de referência sugeridos na documentação oficial do algoritmo. A Tabela 10 apresenta os intervalos concretos considerados para cada hiperparâmetro.

Tabela 10 – Espaço de procura dos hiperparâmetros e respectivas distribuições de amostragem

Hiperparâmetros	Distribuição	Espaço de procura	Tipo
<code>changepoint_prior_scale</code>	log-uniforme	[0.01, 0.20]	contínuo
<code>changepoint_range</code>	uniforme	[0.70, 0.95]	contínuo
<code>n_changepoints</code>	inteiros	{8,...,35}	discreto
<code>holidays_prior_scale</code>	log-uniforme	[0.05, 5.0]	contínuo
<code>daily_fourier</code>	inteiros	{20,...,45}	discreto
<code>weekly_fourier</code>	inteiros	{7,...,20}	discreto
<code>yearly_fourier</code>	inteiros	{5,...,15}	discreto
<code>daily_prior_scale</code>	log-uniforme	[0.10, 5.0]	contínuo
<code>weekly_prior_scale</code>	log-uniforme	[0.10, 5.0]	contínuo
<code>yearly_prior_scale</code>	log-uniforme	[0.10, 5.0]	contínuo
<code>daily_mode</code>	categórica	{ <i>additive, multiplicative</i> }	categórico
<code>weekly_mode</code>	categórica	{ <i>additive, multiplicative</i> }	categórico
<code>yearly_mode</code>	categórica	{ <i>additive, multiplicative</i> }	categórico
<code>regressor_prior_scale</code>	log-uniforme	[0.5, 3.0]	contínuo
<code>regressor_mode</code>	categórica	{ <i>additive, multiplicative</i> }	categórico

A opção pela otimização Bayesiana (BO), em detrimento de um *gridsearch* tradicional, deveu-se à elevada dimensionalidade do espaço de procura e ao custo computacional associado, sendo a abordagem Bayesiana mais eficiente na identificação de regiões promissoras com menos avaliações.

Na fase inicial foram avaliadas 16 combinações de hiperparâmetros, recorrendo a validação temporal parcial com base em cortes fixos do histórico (*cutoffs*). Os *cutoffs* candidatos foram gerados de 60 em 60 dias ao longo de toda a série, totalizando 33. As últimas 24h ficaram

reservadas e não entraram no treino para servir de *holdout* (conjunto de teste reservado) final de avaliação. O modelo foi treinado, nessa seleção de cortes, com todo o histórico disponível até esses pontos e avaliado nas 24 horas seguintes. A Figura 27 representa a validação temporal usada com cortes a cada 60 dias, treino até ao *cutoff* e avaliação nas 24 horas seguintes ao longo de toda a série temporal. As áreas a vermelho correspondem aos dados de treino e as áreas a cinzento correspondem às janelas de validação. A área a verde assinala o conjunto de teste reservado correspondente ao último dia presente no conjunto de dados.

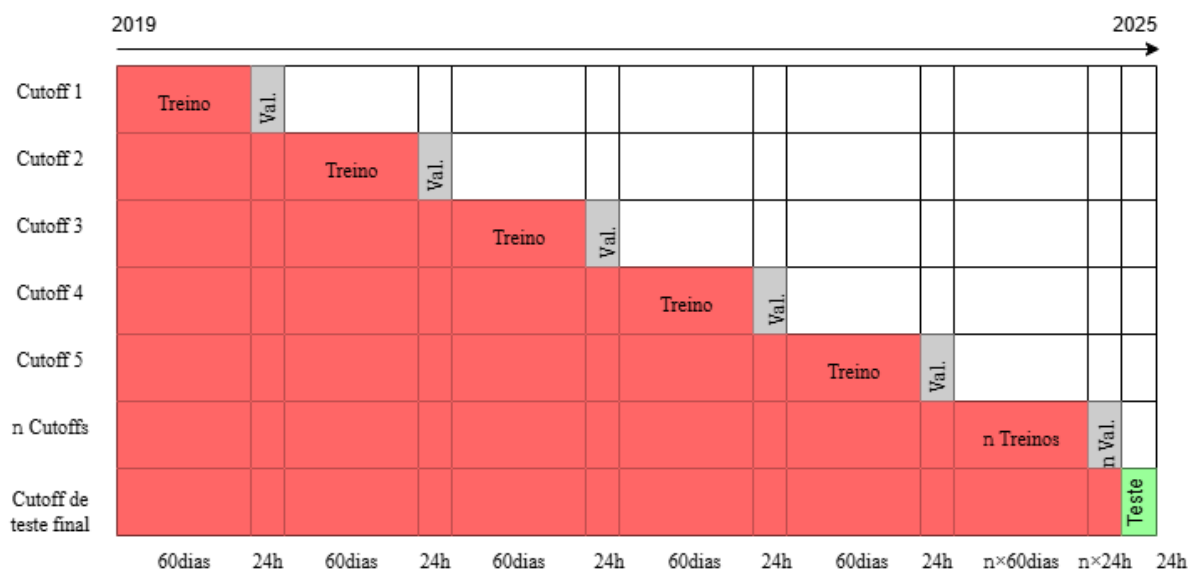


Figura 26 – Esquema de validação temporal do treino na primeira fase da otimização do Prophet

Na segunda fase da otimização foram avaliadas mais 200 combinações adicionais de hiperparâmetros propostas pela abordagem Bayesiana, recorrendo a um subconjunto de quatro *cutoffs* selecionados a partir dos gerados na fase inicial. Estes *cutoffs* foram escolhidos de forma ordenada e distribuída ao longo da série, garantindo espaçamento temporal e representatividade de diferentes períodos, evitando concentrações em trimestres adjacentes. Para reduzir o custo computacional, o histórico foi limitado a aproximadamente 365 dias anteriores ao corte mais antigo da seleção, assegurando a presença de um ciclo anual completo sem necessidade de carregar toda a série temporal. Dentro desse recorte temporal, cada *cutoff* utiliza todo o passado disponível até à sua data ($365 + n$). O *cutoff* mais antigo observa cerca de um ano de dados, enquanto os *cutoffs* mais recentes acumulam intervalos progressivamente maiores e todos são avaliadas no mesmo intervalo de 24h de validação. Esta estratégia permite manter diversidade temporal nas validações com menor consumo de memória e tempo de execução. A Figura 27 representa esta segunda fase.

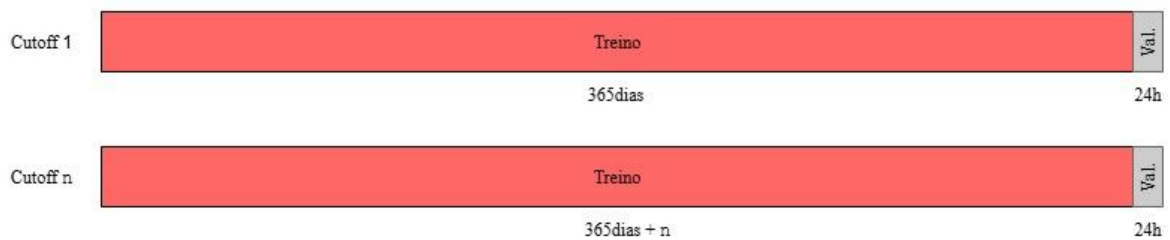


Figura 27 – Esquema de validação temporal do treino na segunda fase da otimização do Prophet

Os resultados de cada validação foram sintetizados numa pontuação que combina métricas determinísticas como MAE, RMSE e WAPE com métricas probabilísticas como cobertura dos intervalos de previsão e *sharpness*. Estes resultados orientaram a seleção de novas combinações promissoras no espaço de procura.

Concluída a otimização, passou-se à fase seguinte onde apenas a melhor configuração de parâmetros foi reavaliada de forma abrangente. O procedimento de treino e teste, neste caso, foi aplicado aos 33 pontos de corte válidos gerados. Em cada corte, o modelo foi treinado com todo o histórico disponível até essa data e depois avaliado nas 24 horas seguintes. Este passo forneceu a principal estimativa independente e abrangente do desempenho, uma vez que os hiperparâmetros permaneceram fixos e os dados desta etapa foram usados exclusivamente para avaliar a configuração selecionada. A cobertura de todos os pontos de corte válidos ao longo da série assegura representatividade de diferentes períodos e padrões de carga. As métricas são agregadas com peso igual entre cortes, o que estabiliza a estimativa e aproxima o resultado do desempenho médio esperado nas previsões.

Como passo complementar, o modelo foi treinado com todos os dados exceto o período de teste reservado, correspondente às últimas 24 horas do conjunto de dados, gerando previsões para esse intervalo. Calcularam-se as mesmas métricas da validação cruzada neste conjunto independente — MAE, RMSE, WAPE, cobertura, *sharpness* e perda *pinball* por quantis. Esta verificação adicional mitiga o otimismo da estimativa, confirma a generalização em dados não usados na otimização e ilustra o comportamento do modelo num dia concreto do horizonte de operação, constituindo apenas uma das várias janelas de teste consideradas ao longo da validação temporal. Todos estes resultados foram continuamente apresentados no terminal e registados em ficheiros de *log*, contendo tanto as métricas de desempenho como os tempos de execução, o que permitiu acompanhar a evolução do processo em tempo real.

3.4.5 RESULTADOS DO TREINO E OTIMIZAÇÃO

Para ordenar as configurações obtidas, o BO forneceu um score composto, definido pela combinação das métricas de avaliação. Considerou-se a seguinte função objetivo escalar:

$$S = \alpha MAE + \beta RMSE + \gamma |Coverage - 0.90| + \delta Sharpness \quad (19)$$

em que MAE e RMSE quantificam o erro pontual das previsões, enquanto *coverage* e *sharpness* caracterizam, respetivamente, a proporção de observações contidas nos intervalos de previsão e a sua largura média. Neste trabalho os pesos foram fixados em $\alpha = 1$, $\beta = 0,1$, $\gamma = 0,2$ e $\delta = 0,05$, escolhidos de forma a atribuir primazia ao MAE, reduzir a influência relativa do RMSE e colocar a penalização da cobertura e da largura dos intervalos na mesma ordem de grandeza das restantes parcelas. Os valores destes pesos foram definidos de forma empírica, após iterações preliminares sobre o comportamento das métricas neste conjunto de dados. Em avaliação de modelos é comum reduzir vários critérios a uma única métrica por meio de pesos pré-definidos que servem de função objetivo. A construção adotada segue o princípio de compromisso entre calibração e concentração dos intervalos preditivos, favorecendo configurações em que os intervalos são tão estreitos quanto possível desde que a *coverage* se mantenha próxima do nível alvo de 90%. A penalização da diferença entre a *coverage* observada e o nível nominal promove calibração. A penalização da *sharpness* desencoraja intervalos excessivamente largos e preserva o conteúdo informativo. Para completar a caracterização do desempenho, foram ainda calculadas as métricas WAPE e perda *pinball* quantílica, que não entram diretamente na definição do score, mas podem ser analisadas a posteriori como medidas adicionais de qualidade previsional. A fiabilidade do ranking foi confirmada por análise de sensibilidade aos pesos, preservando a ordenação das melhores configurações. O Apêndice B apresenta os resultados obtidos durante a BO, numa tabela com as dez melhores combinações testadas, ordenadas por score crescente, com as respetivas métricas de desempenho. A melhor combinação apresenta as métricas apresentadas na Tabela 11.

Tabela 11 – Métricas dos melhores hiperparâmetros nos 4 cutoffs

MÉTRICA	MAE	RMSE	WAPE%	COVERAGE	SHARPNESS	PINBALL
RESULTADO	5.63209	8.07956	7.94948	0.94444	35.09299	1.02182

3.4.6 RESULTADOS DA PREVISÃO

Finalizada a otimização, procedeu-se à reavaliação temporal ao longo dos 33 *cutoffs* gerados com a melhor combinação de hiperparâmetros, de forma a garantir que esta combinação mantinha os resultados ao longo do *dataset* completo. As métricas da Tabela 12 traduzem o desempenho agregado dos 33 cortes. Todas as previsões e observações de validação foram reunidas num único conjunto e, sobre esse conjunto, calcularam-se as métricas de avaliação. Esta abordagem evita agregações intermédias que podem distorcer a leitura e espelha o desempenho global da configuração na validação temporal.

Tabela 12 – Métricas agregadas dos melhores hiperparâmetros nos 33 *cutoffs*

MÉTRICA	MAE	RMSE	WAPE%	COVERAGE	SHARPNESS	PINBALL
RESULTADO	8.20165	11.06142	12.26716	0.85628	31.74599	1.29536

Em comparação com os quatro *cutoffs* usados na seleção, o agregado dos 33 *cutoffs* evidencia a degradação expectável ao ampliar o horizonte temporal e a diversidade de condições. Os erros aumentam de forma clara, com MAE, RMSE e WAPE mais elevados, enquanto a cobertura baixa para valores abaixo do nível nominal. Em contrapartida, os intervalos tornam-se mais estreitos, já que a *sharpness* desce, o que ajuda a explicar a perda de cobertura. A perda *pinball* também sobe no agregado, coerente com o aumento do erro e com intervalos menos conservadores. Ainda assim, o comportamento permanece estável, sustentando a generalização da configuração escolhida no conjunto completo.

Por fim, procedeu-se à avaliação nos dados que foram postos de parte para teste em *holdout*. A Figura 28 mostra a sobreposição das séries real e prevista no período de teste, com a banda de incerteza do modelo sombreada. Observa-se concordância das tendências intradiárias, sem desvios sistemáticos.

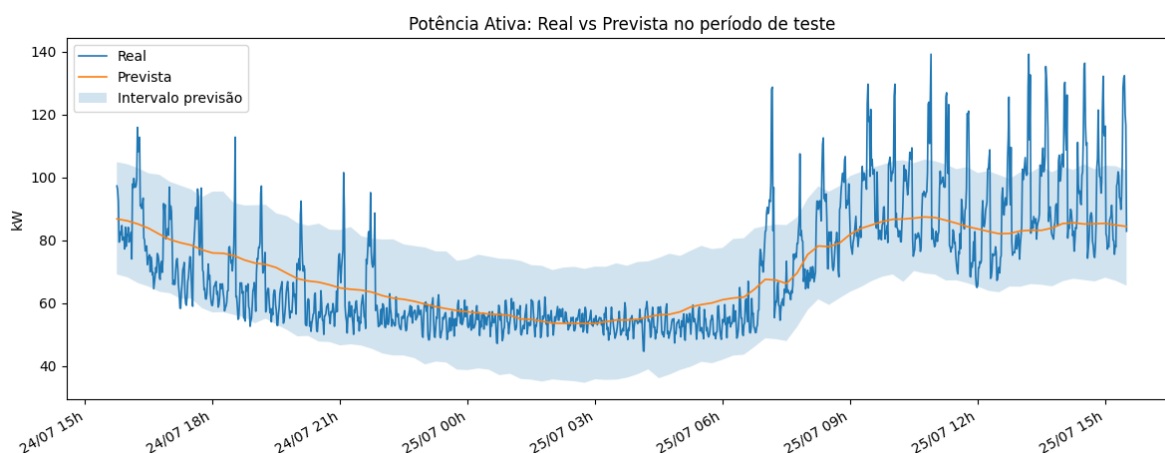


Figura 28 – Comparação da potência ativa real com a previsão em 24 horas do conjunto em holdout

A comparação entre o *holdout* e a avaliação em múltiplos *cutoffs* evidencia uma diferença clara de escala e de exigência. A avaliação em *cutoffs* abrange múltiplos períodos de teste distribuídos no tempo e capta melhor a variabilidade do processo, enquanto o *holdout* aqui considerado cobre apenas um dia. Os resultados do *holdout*, demonstrados na Tabela 13, apontam para boa generalização em dados recentes não treinados, mas a confiança nessa leitura é limitada pelo tamanho do teste. É recomendável ampliar a avaliação com mais dias e com mais horizontes de dados não treinados em diferentes alturas do ano para consolidar a evidência. Neste *holdout* observa-se MAE e RMSE na mesma ordem, o que indica ausência de picos relevantes e estabilidade do desvio. O erro relativo é moderado face ao nível médio de potência e à variabilidade usual do processo, sem sinais de deriva ao longo do horizonte em estudo. A cobertura permanece próxima do nível alvo de 90% e os intervalos mostram boa concentração, o que aponta para calibração adequada sem alargamento excessivo. O *score* de *pinball* valida a coerência entre os quantis previstos e os valores observados.

Tabela 13 – Métricas dos melhores hiperparâmetros no conjunto de teste (24h)

MÉTRICA	MAE	RMSE	WAPE%	COVERAGE	SHARPNESS	PINBALL
RESULTADO	6.16645	8.28539	8.70125	0.9375	36.17268	0.99279

A Figura 29 mostra as curvas da MAE e da RMSE ao longo das 24 horas de teste e permite analisar como o erro evolui durante o dia. Foi possível concluir que as duas métricas não seguem uma tendência ao longo das 24 horas, o que indica ausência de degradação sistemática da precisão. As oscilações observadas concentram-se em períodos de maior variabilidade da carga e não revelam persistência temporal. A proximidade entre as métricas sugere influência limitada dos extremos da distribuição dos resíduos, sustentando a estabilidade do desempenho previsional no período analisado.

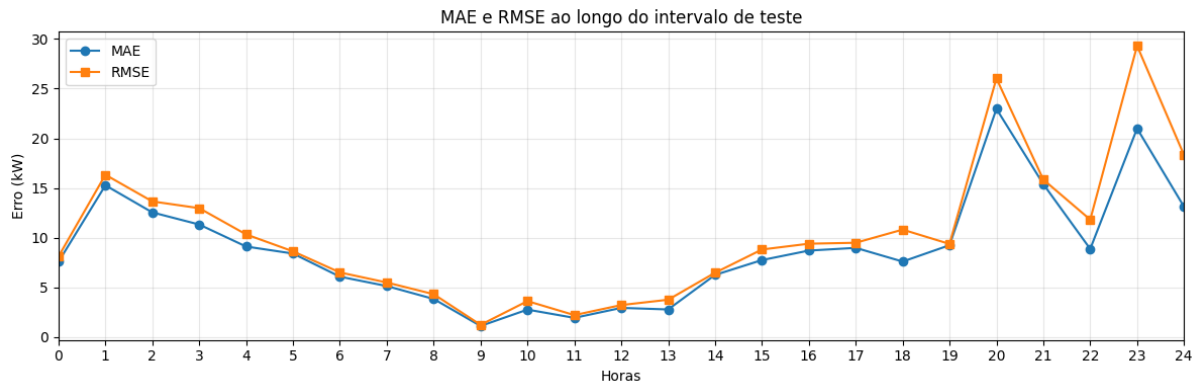


Figura 29 – Evolução das métricas MAE e RMSE durante o período de teste (24h)

Para ultrapassar a limitação associada à avaliação num único dia de teste, a mesma configuração ótima foi aplicada a um *backtest* contínuo ao longo de todo o mês de outubro de 2025. O modelo gerou, para cada um dos trinta e um dias, trajetórias de previsão com horizonte de vinte e quatro horas e resolução de quinze minutos em regime totalmente fora de amostra, utilizando como treino apenas o histórico disponível até ao início de cada dia. As métricas médias dos respetivos horizontes diários, resumidas na Tabela 14, mantêm a ordem de grandeza observada na avaliação em múltiplos *cutoffs* e no teste em *holdout* e evidenciam uma ligeira melhoria nas métricas determinísticas, com erro relativo estável e cobertura mais próxima do nível alvo.

Tabela 14 – Métricas médias do modelo no mês de outubro de 2025

MÉTRICA	MAE	RMSE	WAPE%	COVERAGE	SHARPNESS	PINBALL
RESULTADO	5.456	6.703	9.083	0.990	36.108	0.921

Em conjunto, estes resultados mostram que o modelo preserva a capacidade preditiva quando exposto a um horizonte contínuo e mais exigente de dados não treinados, reforçando a robustez da configuração selecionada para utilização regular.

A Figura 30 ilustra a evolução da potência ativa real e prevista ao longo do mês de outubro de 2025, incluindo os intervalos de previsão associados. Observa-se que a forma global da curva é bem reproduzida, com boa correspondência entre os picos diários de consumo e as trajetórias centrais previstas, e que os intervalos englobam a maior parte das oscilações observadas, em particular nos períodos de maior variabilidade. Em alguns dias verificam-se discrepâncias pontuais nos máximos instantâneos, com leve tendência para subestimação dos valores de pico, o que é consistente com a suavização inerente ao modelo. No entanto, esse efeito não compromete de forma significativa a qualidade global da previsão e a evolução conjunta das séries real, prevista e dos intervalos ao longo de todo o mês é compatível com o desempenho sintetizado na Tabela 14.

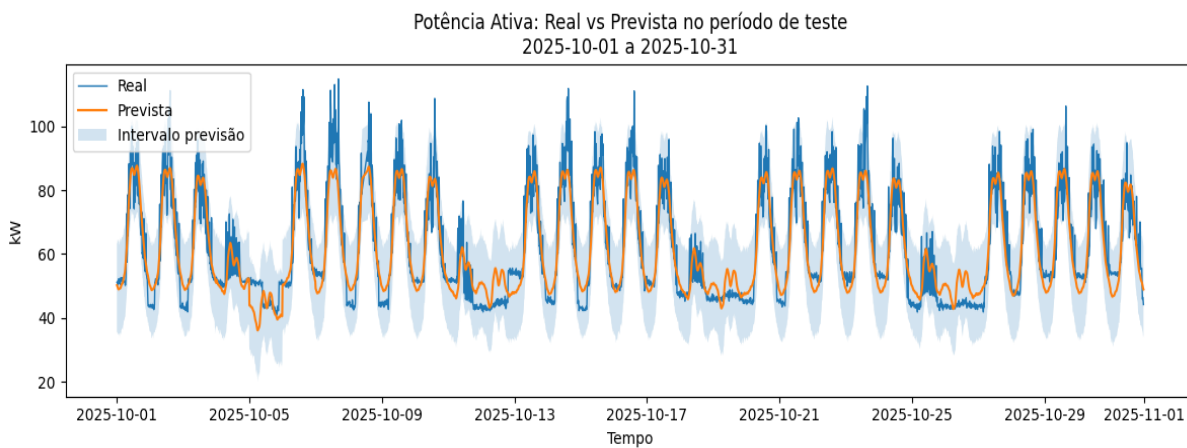


Figura 30 – Potência ativa real e prevista no mês de outubro de 2025, com intervalos de previsão sombreados

3.4.7 IMPLEMENTAÇÃO OPERACIONAL DO MODELO

Concluída a otimização e avaliação, foi desenvolvido um script em Python que carrega a configuração vencedora, volta a treinar o modelo com todo o histórico disponível e gera previsões para as 24 horas seguintes ao momento de execução. O script foi agendado em crontab com periodicidade diária. No final da sua execução os resultados são gravados na tabela `gel_fct_cintal_edificio_1_previsao_consumo` da base de dados e enviados para o EmonCMS. O treino é efetuado em cada execução, incorporando os dados mais recentes e garantindo que as previsões refletem informação atual.

A Figura 31 comprova a integração da previsão no EmonCMS ao mostrar a potência ativa observada e a série prevista com a respetiva banda de incerteza. O alinhamento visual confirma que o *pipeline* de previsão está incorporado no ambiente de monitorização e que os resultados

podem ser consultados e validados diretamente na plataforma. A banda entre $yhat_lower$ e $yhat_upper$ é o intervalo de previsão para a incerteza do modelo, assinalando a região onde o valor observado deverá ocorrer. Quando a série observada permanece maioritariamente dentro

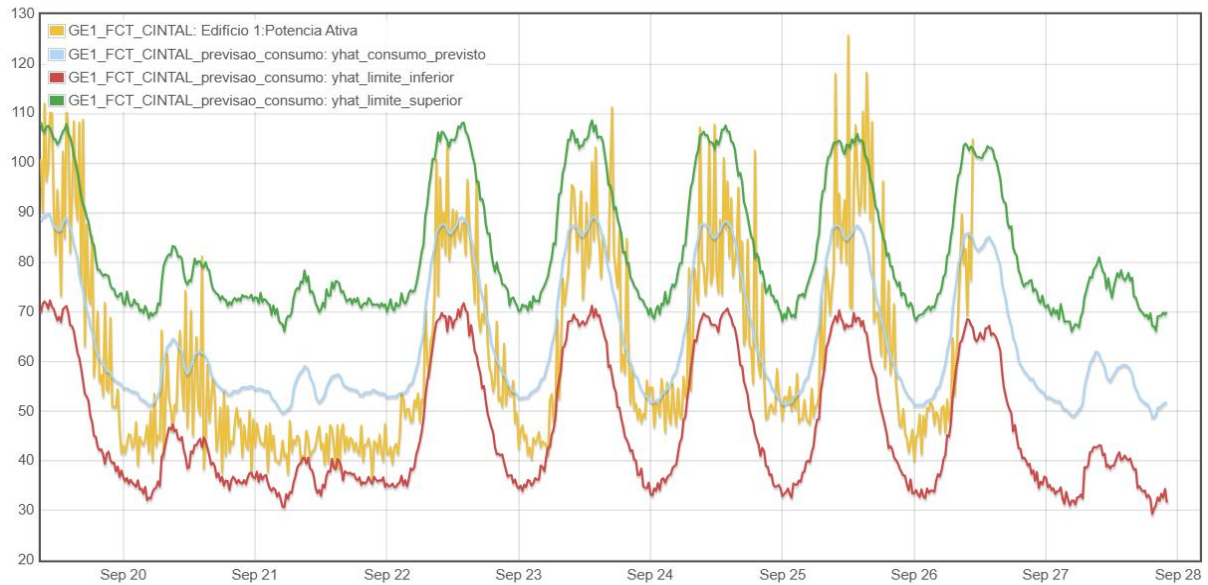


Figura 31 – Séries temporais de potência ativa observada e previsões e limites de consumo do Edifício 1 no EmonCMS

desta banda infere-se que a calibração probabilística do modelo é adequada.

3.5 MACHINE LEARNING PARA DETECÇÃO DE ANOMALIAS

A detecção de anomalias foi implementada como arquitetura híbrida em duas camadas. A primeira camada baseia-se em resíduos entre potência observada e prevista, normalizados de forma robusta para estabilizar a escala do erro e avaliados com regras de persistência temporal. A segunda camada aplica o *Isolation Forest* sobre um conjunto de características derivadas dos resíduos e da dinâmica local do sinal. A decisão final resulta da interseção entre a violação residual persistente, apurada na análise dos resíduos, e a sinalização do *Isolation Forest*, reduzindo falsos positivos e preservando a interpretabilidade no domínio físico da potência.

3.5.1 RESÍDUOS E DECISÃO POR PERSISTÊNCIA

Nesta camada a série de resíduos é obtida como a diferença entre a potência observada e a potência prevista. Para estabilizar a escala do erro, aplica-se normalização robusta, em que cada resíduo é recentrado na mediana e escalado pelo desvio absoluto mediano (MAD), ambos calculados numa janela deslizando que considera os n últimos registos, sendo o comprimento da janela ajustado empiricamente a partir de dados históricos. O MAD é uma medida robusta

de dispersão amplamente utilizada em métodos de detecção de *outliers* e anomalias em séries temporais, incluindo aplicações em consumo energético [94]. Esta padronização produz um *score* que quantifica a intensidade do desvio face ao comportamento típico recente e apresenta baixa sensibilidade a valores extremos. Um ponto torna-se elegível para alarme quando o *score* excede um limiar calibrado no histórico recente e a violação persiste por uma janela mínima, impondo persistência temporal e mitigando picos isolados ou ruído de curta duração. O limiar é definido por calibração periódica com base em percentis dos *scores* de resíduo e pode ser ajustado por edifício ou por ponto de medição. Após a detecção pontual, pontos consecutivos são agregados por continuidade temporal para formar eventos com início e fim.

Seja $r(t) = y(t) - \hat{y}(t)$ o resíduo. Seja W_t a janela deslizante retrospectiva usada em t . Define-se:

$$m(t) = \text{mediana}\{r(\tau) : \tau \in W_t\}, \quad MAD = \text{mediana}\{|r(\tau) - m(t)| : \tau \in W_t\}.$$

Onde o *score* é calculado como:

$$s(t) = \frac{r(t) - m(t)}{c \cdot MAD(t)}, \quad (20)$$

onde c é uma constante de escala, fixada em 1,4826, que permite interpretar o *score* como um desvio normalizado em unidades aproximadamente equivalentes ao desvio padrão. A decisão por persistência exige que $|s(t)| > \theta$ durante pelo menos M amostras consecutivas, com θ calibrado no histórico recente.

A aplicação prática do procedimento encontra-se ilustrada na Figura 32, que resume o comportamento da camada de resíduos ao longo de um período de várias semanas, desde o início de setembro até ao final de outubro de 2025, para o edifício em estudo. No gráfico superior representa-se a potência ativa medida e a potência prevista, enquanto as regiões sombreadas a cinzento assinalam os intervalos em que o *score* residual excede o limiar durante pelo menos duas amostras consecutivas. A parametrização utilizada recorreu a uma janela deslizante de quarenta e oito amostras, correspondente a doze horas com passo de quinze minutos, e a um limiar θ obtido como percentil 95 do histórico recente dos valores absolutos do *score* dentro da janela, impondo ainda uma persistência mínima M igual a 2. O painel inferior mostra a evolução temporal do *score*, o limiar simétrico e as excedências pontuais assinaladas a vermelho, permitindo observar a relação direta entre a ultrapassagem persistente do limiar e os intervalos sombreados no painel superior. Neste período obtiveram-se 4916 amostras com observação e previsão em simultâneo à resolução de quinze minutos. Com essa

configuração foram detetados quarenta e dois eventos persistentes ao longo do período analisado. Estes eventos distribuem-se maioritariamente em torno dos períodos de maior carga, o que confirma a tendência do método para realçar desvios sustentados face à trajetória prevista e atenuar oscilações isoladas de curta duração.

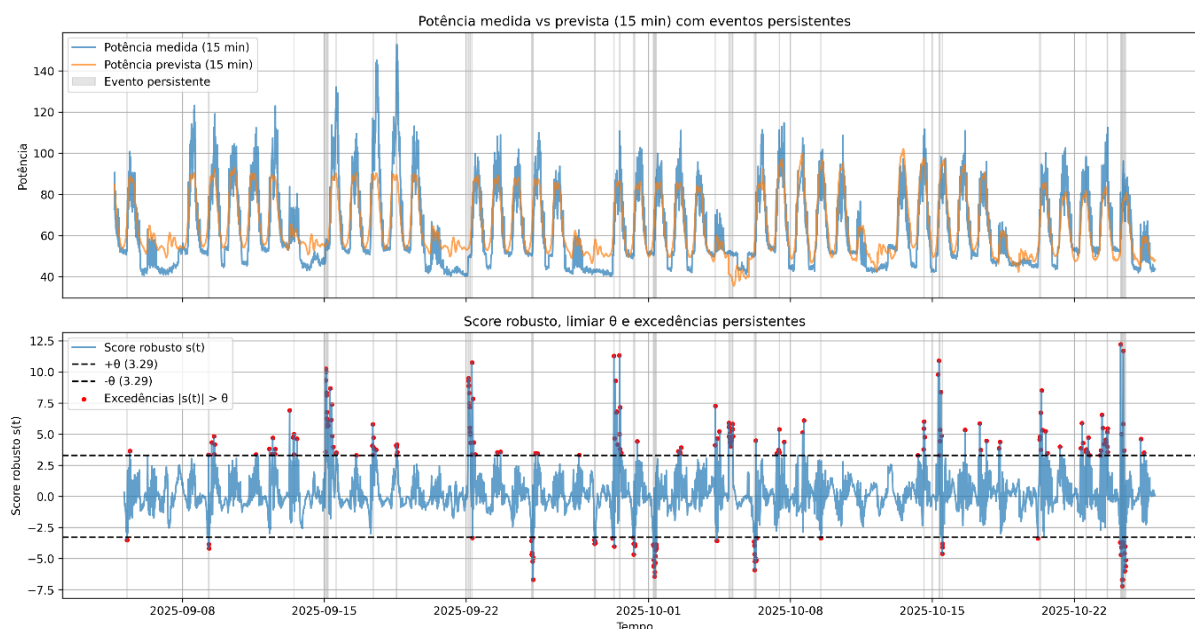


Figura 32 – Detecção de eventos anómalos através de resíduos com persistência

A análise evidencia que a deteção é sensível aos parâmetros definidos, em particular ao tamanho da janela de suavização, ao quantil usado para calibrar o limiar e ao número mínimo de excedências consecutivas exigidas. Uma janela mais curta aumenta a reatividade do *score*, conduzindo a maior número de excedências, mas também a maior suscetibilidade a oscilações de curta duração. Pelo contrário, uma janela mais longa estabiliza o comportamento e favorece a identificação apenas de desvios mais sustentados, mas ao custo de menor sensibilidade a alterações rápidas. O limiar, ao ser ajustado em percentis mais elevados do histórico, reduz o número de eventos detetados, concentrando-se em desvios de grande magnitude, enquanto valores mais baixos tornam o sistema mais permissivo. A persistência mínima atua como filtro temporal, valores mais elevados restringem a deteção a desvios prolongados, reduzindo a probabilidade de falsos alarmes, mas podendo omitir episódios curtos. Valores mais baixos aumentam a sensibilidade a flutuações rápidas, com maior suscetibilidade a assinalar variações transitórias que não correspondem a anomalias estruturais. Neste enquadramento, os parâmetros foram fixados de forma empírica, com um compromisso intermédio entre

sensibilidade e robustez, servindo sobretudo para preparar a etapa seguinte em que o *Isolation Forest* terá a decisão final sobre a classificação das anomalias.

3.5.2 CLASSIFICAÇÃO FINAL COM *ISOLATION FOREST*

No âmbito deste trabalho, o *Isolation Forest* foi aplicado como camada decisória sobre a análise residual. Em primeiro lugar, as amostras são avaliadas pela regra de resíduos persistentes e, em seguida, pelo algoritmo de isolamento. Só é detetada anomalia quando ambos os critérios convergem na classificação de *outlier*, descartando casos em que apenas um dos métodos sinaliza. Para o caso de estudo, o *Isolation Forest* foi treinado com o último ano disponível de dados de potência ativa, de modo a refletir o regime atual de utilização do edifício e evitar que padrões antigos, potencialmente desatualizados, condicionassem a aprendizagem. Em cada amostra, o vetor de entrada é constituído pelo valor instantâneo de potência ativa, pela hora do dia e pelo dia da semana. Desta forma, o modelo aprende diretamente a fronteira de comportamento normal a partir da combinação entre o nível de carga e o contexto de calendário, sem assumir uma forma paramétrica pré-definida para a relação entre estas variáveis, o que o torna mais adaptável a diferentes regimes de operação. A configuração do *Isolation Forest* envolve parâmetros com impacto direto na sensibilidade e na estabilidade do modelo. Entre os parâmetros mais relevantes destacam-se o `n_estimators`, que influencia a estabilidade dos *scores* e a consistência entre execuções, uma vez que um maior número de árvores aumenta a estabilidade estatística do resultado, a fração `contamination`, que regula a sensibilidade do método ao definir a proporção de amostras que se espera que sejam classificadas como anómalas, afetando diretamente o limiar de decisão. Além desses, existe o parâmetro `max_samples` que controla o tamanho da amostra usada em cada árvore, influenciando a variância do modelo e o custo computacional, o parâmetro `max_features` que define o número de variáveis consideradas em cada divisão, aumentando a diversidade das árvores e a capacidade de explorar padrões em subespaços, e por fim o `random_state` que assegura a reprodutibilidade das execuções.

Nesta aplicação, os parâmetros foram configurados para garantir consistência nos resultados e avaliar a adequação do algoritmo ao contexto. O número de árvores (`n_estimators`) foi fixado em 200, assegurando estabilidade estatística dos *scores*. O parâmetro `max_samples` foi mantido automático, permitindo que cada árvore fosse treinada com subconjuntos de dados suficientemente representativos, sem necessidade de ajuste. A seleção de variáveis em cada divisão, controlada pelo `max_features`, foi mantida no valor padrão, assegurando diversidade entre as árvores. O `random_state` foi configurado para garantir que as escolhas

aleatórias do algoritmo se mantêm consistentes entre execuções. A fração de contaminação (contamination) foi fixada em 1%, alinhada com a baixa incidência expectável de anomalias, mas suficiente para permitir a sua deteção. A calibração final destes valores deve ser realizada com base em validação retrospectiva com dados reais de cada edifício, procurando um compromisso adequado entre cobertura de eventos relevantes e minimização de falsos positivos. Na prática, o sistema apenas emite alarme quando a função de avaliação do *Isolation Forest* desce abaixo do valor de corte definido e, em simultâneo, ocorre violação persistente do limiar residual.

A Figura 33 ilustra o funcionamento do modelo ao longo do mesmo intervalo considerado na camada de resíduos, representando a série de potência ativa com os instantes classificados como *outliers* assinalados a vermelho. Verifica-se que o algoritmo concentra as deteções em episódios em que o consumo se afasta de forma marcada do patamar típico, sobretudo em torno de períodos de carga elevada e de algumas transições rápidas, refletindo a sensibilidade do modelo a desvios no nível e na dinâmica local da série. No período de deteção considerado, correspondente a 4916 instantes, o modelo assinalou 53 amostras como anómalas, cerca de 1,08 % do total, valor alinhado com a fração de contaminação configurada.

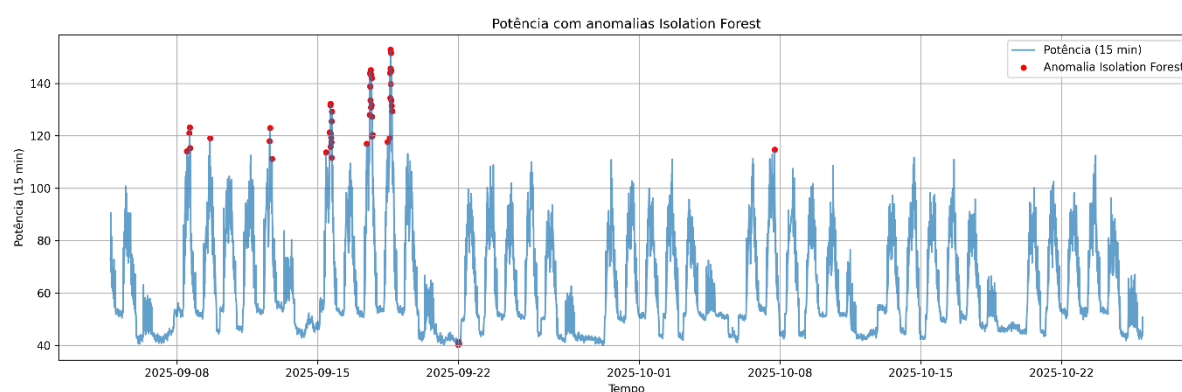


Figura 33 – Potência ativa com deteções do algoritmo Isolation Forest, entre 4 de setembro e 26 de outubro de 2025

A decisão final resulta da interseção entre as janelas de resíduos persistentes e as amostras classificadas como *outliers* pelo *Isolation Forest*. A Figura 34 apresenta a série de potência ativa ao longo de todo o período analisado, com os alarmes finais assinalados com marcadores a preto. Observa-se que, apesar do elevado número de pontos avaliados, apenas dois intervalos curtos são sinalizados, visíveis na figura como dois grupos contíguos de marcadores sucessivos, o que ilustra o carácter seletivo da combinação entre os dois métodos.

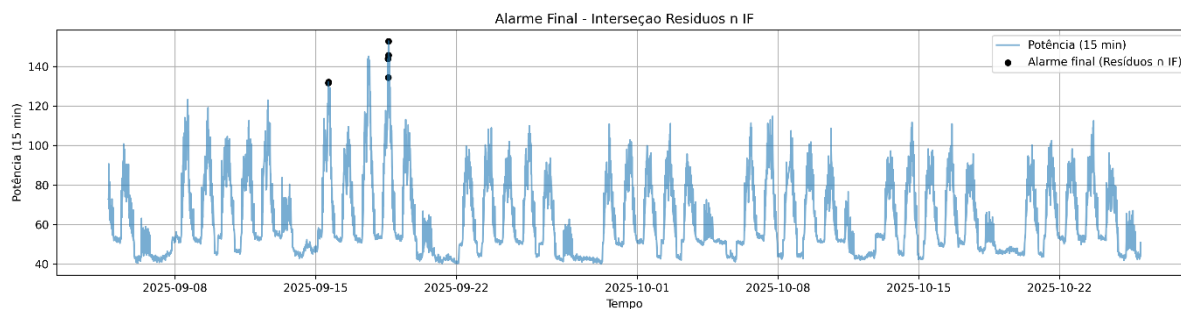


Figura 34 – Potência ativa no edifício 1, com os alarmes finais obtidos pela interseção entre resíduos persistentes e *Isolation Forest*, entre 4 de setembro e 26 de outubro de 2025

A Figura 35 mostra um zoom ao intervalo de 15 a 22 de setembro da Figura 34, permitindo inspecionar em detalhe os episódios detetados. Verifica-se que os alarmes se concentram em dois intervalos curtos, coincidentes com picos de consumo significativamente superiores ao padrão típico. Este comportamento confirma o carácter conservador da arquitetura em duas camadas, que filtra a maioria das sinalizações isoladas de cada método e concentra a atenção do operador em desvios simultaneamente persistentes no tempo e atípicos no espaço de características.

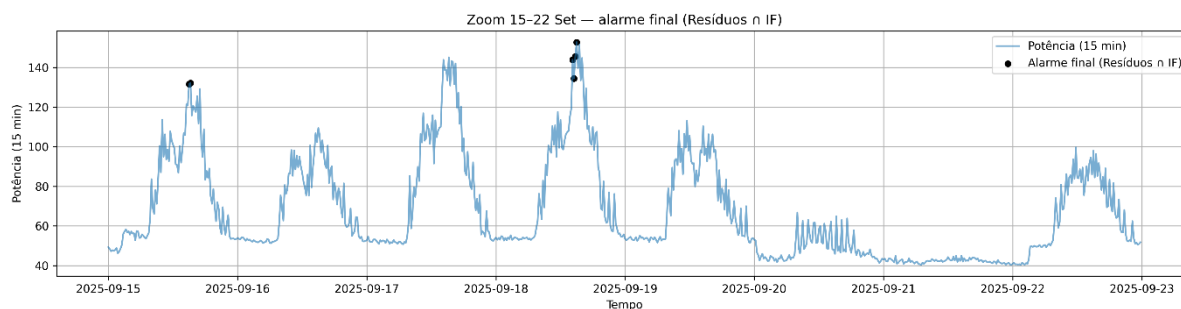


Figura 35 – Detalhe do período de 15 a 22 de setembro de 2025, com pormenor dos alarmes finais resultantes da interseção entre resíduos persistentes e *Isolation Forest*

O script responsável por estas detecções é executado a cada quinze minutos, agendado no *crontab*, avaliando o consumo recente segundo os critérios definidos e registrando as ocorrências para consulta no EmonCMS e na base de dados.

3.5.3 DETECÇÃO DE ANOMALIAS DE REDE E INFRAESTRUTURA

Os métodos aplicados à série de consumo, discutidos anteriormente, são adequados para detetar desvios no comportamento dos consumos dos edifícios, mas não cobrem falhas técnicas no percurso de recolha, transporte e persistência dos dados. Em cenários de perda de conectividade dos dispositivos que integram o sistema, ou de erros de escrita na base de dados, a série pode apresentar lacunas ou padrões artificiais que não configuram por si um *outlier* estatístico. Noutros casos, podem nem existir observações suficientes para calcular resíduos. Sem uma camada específica de monitorização da integridade da infraestrutura, estas situações conduzem a diagnósticos ambíguos, aumentam os falsos positivos da detecção baseada na série e atrasam a correção das causas.

Para colmatar essa lacuna, integrou-se uma rotina específica de detecção de anomalias de rede e de infraestrutura que, de forma periódica, monitoriza a conectividade ao EmonCMS e a todos os medidores, bem como o estado de atualização dos registos na base de dados. O procedimento verifica a disponibilidade do servidor e dos dispositivos com o comando *ping*, identifica tabelas vazias ou desatualizadas por mais de três intervalos de execução e assinala situações de estagnação caracterizadas por sequências prolongadas de valores repetidos ou pela ausência de novos registos. Cada ocorrência é registada com identificador único e guardada em ficheiro JSON, incluindo endereço IP, causa provável e tabela afetada, o que permite gerar relatórios periódicos de diagnóstico.

Esta separação por eixos, consumo anómalo por um lado e integridade operacional por outro, melhora a robustez do sistema em três dimensões. Em primeiro lugar, reduz falsos positivos na camada de série ao filtrar previamente problemas de dados e de comunicação. Em segundo, evita falsos negativos em que a ausência de dados mascara incidentes reais. Por fim, acelera o processo de triagem, ao fornecer indicadores diretos sobre a origem provável do incidente, orientando a intervenção para a rede, para o dispositivo ou para o armazenamento, e preservando a integridade global do *pipeline* de monitorização e previsão.

A verificação de rede e infraestrutura também é executada de quinze em quinze minutos, via *crontab*.

4

RESULTADOS EXPERIMENTAIS

A arquitetura atual do sistema revelou-se mais estável do que a versão anterior, fruto da consolidação da plataforma numa máquina virtual dedicada, onde ocorre centralmente a recolha, o processamento e o armazenamento dos dados energéticos. Em paralelo, foi mantido um sistema funcional no servidor físico do laboratório 160, configurado como instância redundante, capaz de operar autonomamente em caso de falha do servidor principal. Esta abordagem contribuiu para aumentar a resiliência do sistema e garantir a continuidade da recolha de dados em contextos de falha, validando o princípio de redundância previsto no planeamento.

Em termos de aquisição de dados, o novo código Python unificado demonstrou fiabilidade ao longo dos meses de teste subsequentes à sua implementação. As leituras dos diferentes dispositivos começaram a ser corretamente interpretadas, convertidas e registadas sem perdas significativas. Embora este script constitua o núcleo do processo de recolha, o sistema integra ainda outros módulos especializados, desenvolvidos sob a forma de scripts autónomos, responsáveis por tarefas como a leitura da estação meteorológica, a realização de cópias de segurança, a reinicialização mensal do custo acumulado por edifício, a previsão de consumo e a deteção de anomalias. Esta arquitetura modular garante a distribuição equilibrada das funções e reforça a resiliência do sistema, facilitando a evolução e manutenção.

As situações de falha de comunicação começaram a ser identificadas e registadas em *logs*, permitindo monitorizar os problemas de conectividade.

A criação da nova base de dados PostgreSQL constituiu um dos avanços mais significativos do sistema. A estrutura implementada, com tabelas padronizadas por cada edifício e nomenclatura consistente para as variáveis energéticas, permitiu uniformizar a organização da informação e eliminar as redundâncias e incoerências presentes nas bases de dados anteriores. A base de dados foi concebida como repositório central e como suporte a análises futuras, tendo sido o seu funcionamento validado através de testes de integridade e consistência sobre as séries históricas importadas. Integra milhões de registos por edifício, permite a interligação a plataformas de visualização e sustenta a integração dos modelos de previsão. A plataforma

EmonCMS, instalada localmente, permanece como o núcleo de visualização e disponibilização de dados energéticos, expondo APIs REST sobre HTTP e permitindo integração com plataformas institucionais como a UAlgNet. Apesar das alterações estruturais implementadas ao longo do projeto, os *dashboards* pré-existentes foram continuamente integrados na nova arquitetura e foi garantido o seu pleno funcionamento.

Após se ter garantido o funcionamento efetivo da recolha e do armazenamento de dados a evolução natural do projeto culminou na integração da camada de *machine learning* com visualização interativa em Grafana. Esta plataforma complementa o EmonCMS na exploração analítica ao permitir consultas SQL nativas, composição de painéis com variáveis e filtragem dinâmica por edifício e período, sobreposição de séries com intervalos de previsão e marcação de anomalias, transformações no próprio painel e anotações de eventos. O Grafana facilita ainda a correlação entre múltiplas fontes e a reprodução de *dashboards* que podem ser inseridos em páginas *web* institucionais. O EmonCMS mantém-se como núcleo de aquisição e disponibilização operacional dos dados, enquanto o Grafana disponibiliza um ambiente mais flexível e orientado à decisão. Com vista à validação final deste conceito, foi desenvolvido um *dashboard* dedicado do edifício piloto usado na previsão, o Edifício 1, no Campus das Gambelas. O painel está organizado em quatro secções. A primeira apresenta variáveis instantâneas em indicadores do tipo *gauge*, permitindo acompanhar em tempo real as tensões e correntes trifásicas, a potência ativa e reativa e o custo horário. A segunda secção é dedicada aos consumos energéticos, disponibilizando séries temporais das mesmas variáveis elétricas, complementadas pelas curvas de potência ativa e reativa, essenciais para a caracterização do perfil de carga. A terceira secção apresenta variáveis de energia relacionadas com custos, incluindo a evolução do custo instantâneo e do custo acumulado mensal. Por fim, a quarta secção integra os resultados do modelo de previsão com deteção de anomalias, representando em simultâneo a potência ativa real e prevista, acompanhada dos intervalos de confiança e com marcação visual dos pontos anómalos detetados.

As Figuras 36 a 39 apresentam por secções o *dashboard* desenvolvido no Grafana, integrando a monitorização em tempo real, a análise histórica e a previsão de consumos com deteção de anomalias.

Este painel representa o culminar do processo de desenvolvimento, transformando o sistema de monitorização num verdadeiro sistema de apoio à decisão.

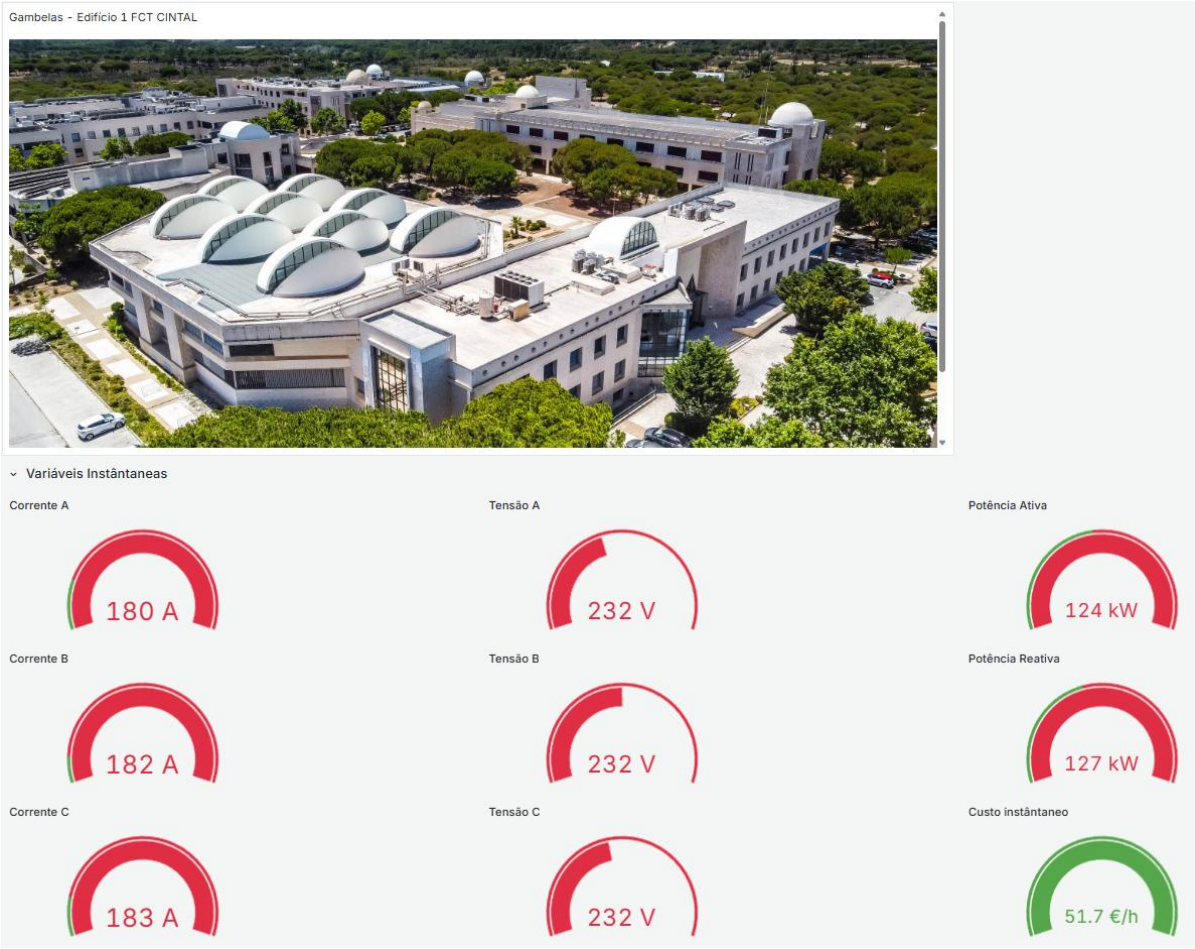


Figura 36 – Visualizações do tipo gauge para monitorizar variáveis instantâneas

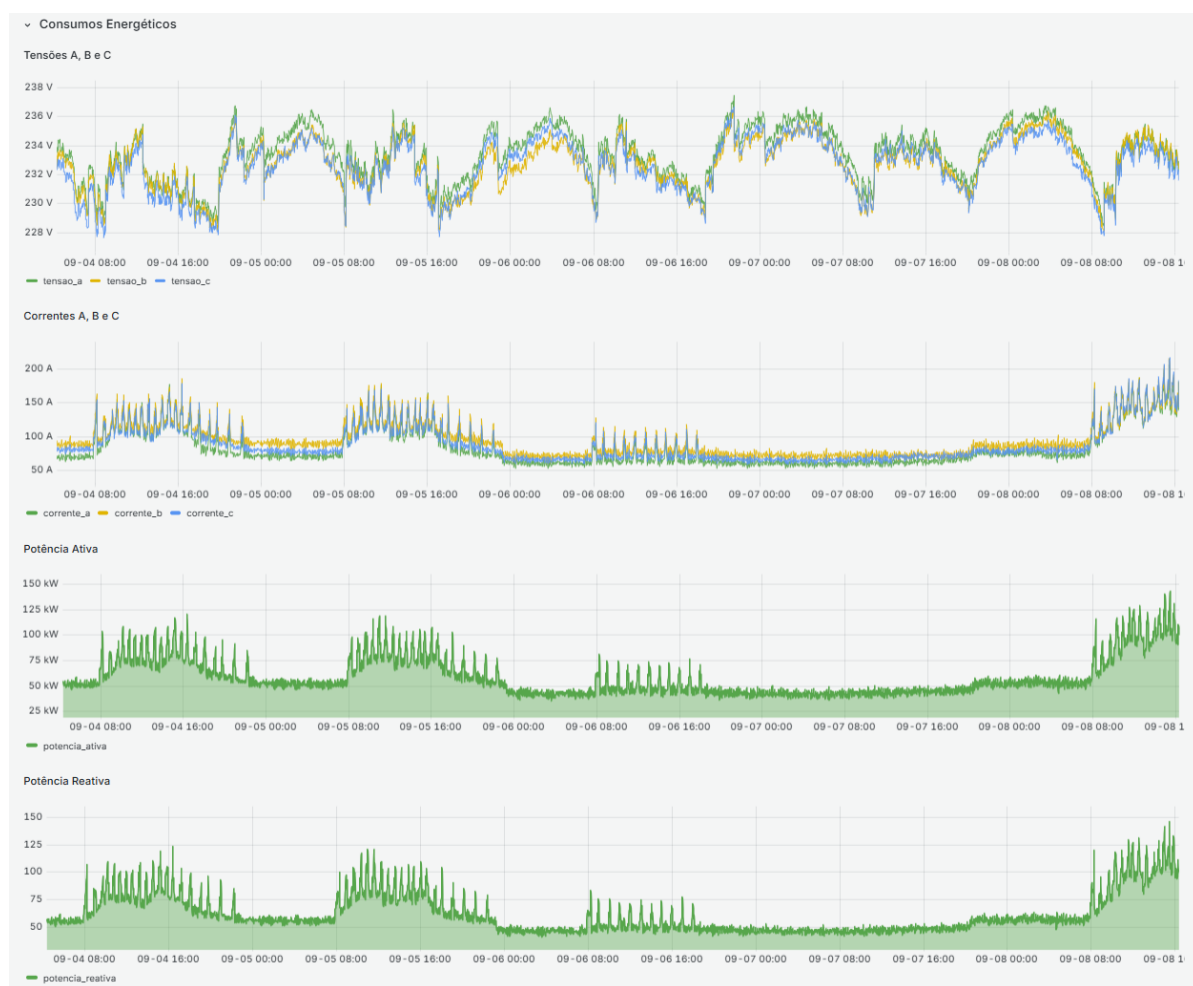


Figura 37 – Visualização em série temporal para variáveis relativas a consumo energético

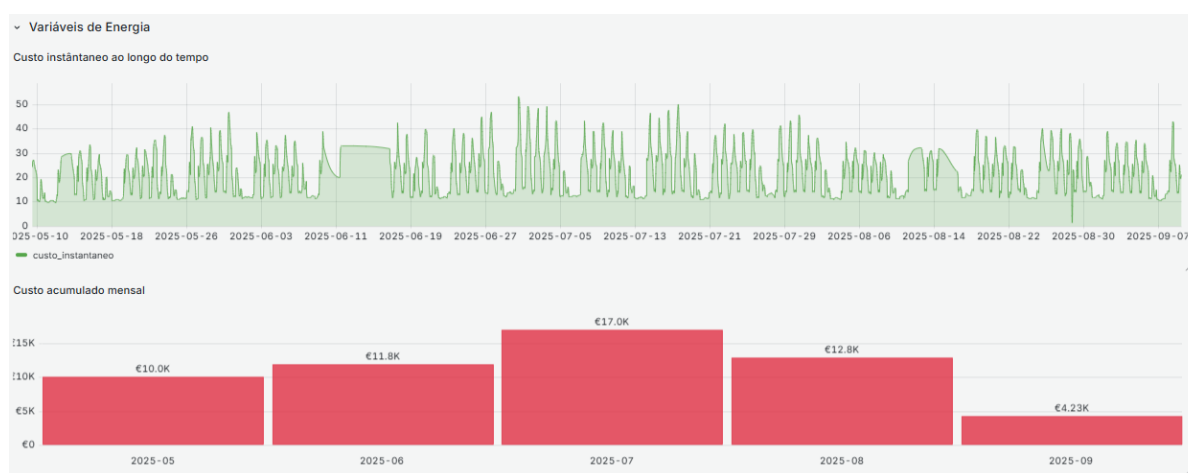


Figura 38 – Visualização do custo instantâneo em série temporal e do acumulado em gráfico de barras

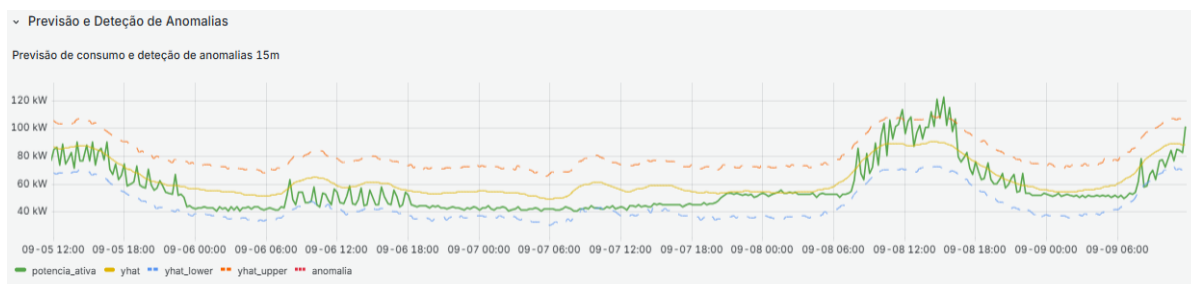


Figura 39 – Visualização da curva de consumo real e de previsão com limites de erro definidos

A integração do modelo de previsão no ambiente operacional confirmou estabilidade e coerência entre valores observados e estimados. A seleção de hiperparâmetros por validação cruzada revelou métricas que mantiveram a ordem de grandeza com cobertura próxima do alvo e intervalos compatíveis com a calibração conforme a Tabela 11. A otimização bayesiana convergiu para uma região estável que equilibrou as métricas e produziu ganhos face às configurações iniciais. O resumo das melhores combinações e respectivas métricas encontra-se no Apêndice B.

Quando a mesma configuração foi reavaliada nos 33 *cutoffs* distribuídos ao longo da série temporal, observou-se a degradação expectável das métricas e uma ligeira redução da cobertura, acompanhada por intervalos de previsão mais estreitos. Tanto a redução da cobertura como o aumento da perda *pinball* são compatíveis com a maior diversidade de padrões em treino e validação expondo o modelo a cenários mais heterogêneos. A Tabela 12 sintetiza estes resultados e mostra que as métricas se mantêm em faixas operacionais para o edifício piloto, sustentando a capacidade de generalização do modelo em cenários mais diversos e a estabilidade do seu desempenho global.

No teste em *holdout* reservado às últimas 24 horas, a comparação visual entre a série real e a prevista confirma alinhamento de tendência e ausência de desvios sistemáticos como se observa na Figura 28. As métricas mantêm a ordem de grandeza verificada nos passos anteriores de validação com cobertura próxima do alvo e intervalos de previsão de largura controlada coerentes com a calibração probabilística. Estes resultados constam da Tabela 13 e reforçam a boa generalização em dados não utilizados durante os treinos da otimização. A Figura 29 evidencia estabilidade horária das curvas de erro, sem degradação ao longo de um dia, o que sugere ausência de deriva temporal no horizonte analisado. Para ultrapassar a limitação associada à avaliação num único dia, a mesma configuração ótima foi aplicada a um *backtest* contínuo ao longo de todo o mês de outubro de 2025. As métricas médias obtidas, apresentadas na Tabela 14 e ilustradas na Figura 30, mantêm a ordem de grandeza das

avaliações anteriores e revelam ligeira melhoria nas métricas determinísticas, com cobertura próxima de 100% e intervalos de previsão bem concentrados, o que confirma a robustez do modelo em regime operacional prolongado.

Na componente de detecção de anomalias, a aplicação da estratégia por resíduos evidenciou a sua capacidade para destacar desvios sustentados. Com a parametrização adotada ao longo do período de detecção, compreendido entre 4 de setembro e 26 de outubro de 2025, foram identificados quarenta e dois eventos persistentes na camada de resíduos, concentrados sobretudo em torno dos períodos de maior carga, tal como ilustrado na Figura 32.

No mesmo intervalo de detecção, o *Isolation Forest*, treinado com o último ano de dados de potência ativa, classificou cinquenta e três amostras como anómalas, cerca de 1,08 % dos pontos avaliados, valor alinhado com a fração de contaminação configurada. Ao exigir a interseção entre a violação residual persistente e a detecção de *outlier* pelo algoritmo, o conjunto de candidatos foi reduzido a apenas dois episódios de alarme, visíveis na Figura 34 sob a forma de dois grupos contíguos de marcadores sucessivos. Este comportamento confirma o perfil conservador da arquitetura em duas camadas, que filtra a maioria das sinalizações isoladas de cada método e concentra a atenção do operador em desvios simultaneamente persistentes no tempo e atípicos no espaço de características.

Em síntese, o modelo manteve um desempenho consistente nas diferentes fases de validação e teste, com cobertura próxima do alvo e erros controlados face à variabilidade da carga. A estratégia por resíduos com persistência isolou desvios sustentados e mostrou ser ajustável ao perfil de carga do edifício em estudo, equilibrando sensibilidade e redução de falsos alarmes. A interseção com o *Isolation Forest* atuou como verificação independente que reforça a decisão e mantém a interpretabilidade, com execução periódica e registo contínuo em base de dados e EmonCMS.

A Figura 40 sintetiza a arquitetura final e a circulação de variáveis entre os principais scripts e serviços do sistema. A representação é propositadamente simplificada e centrada nos fluxos essenciais entre os módulos. Explicita as trocas de dados entre *scripts*, base de dados e plataformas de visualização evidenciando como aquisição, armazenamento, previsão e detecção de anomalias se articulam num circuito único.

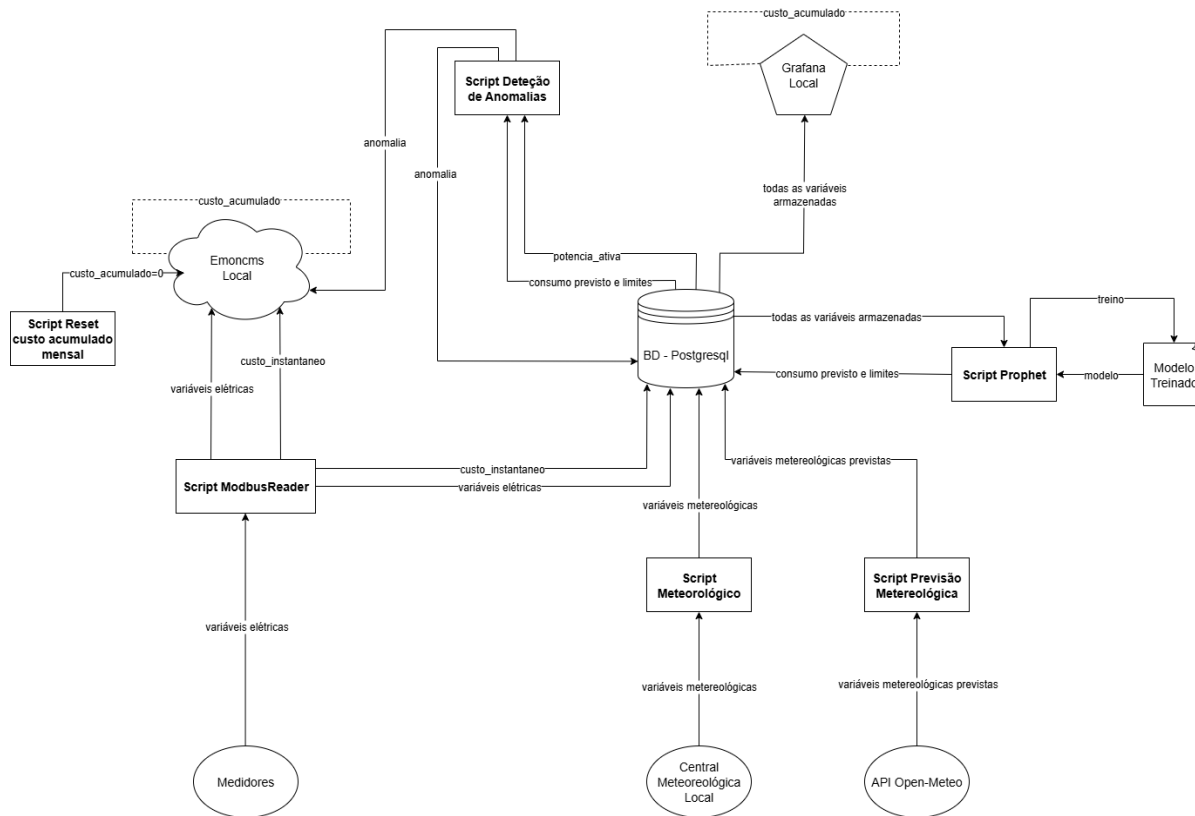


Figura 40 – Arquitetura final do sistema com fluxos de dados

4.1 DISCUSSÃO

O resultado do desenvolvimento deste projeto comprova a evolução de um sistema básico de monitorização para um sistema modular de apoio à decisão. A criação de uma base de dados estruturada, a implementação de mecanismos de redundância e a disponibilização de *dashboards* dinâmicos consolidaram a infraestrutura como uma solução estável e consistente no contexto institucional.

A integração de modelos preditivos baseados em aprendizagem automática e a introdução de mecanismos de deteção de anomalias constituíram um avanço relevante, mas a deteção atualmente realizada deve ser complementada com a possível confirmação das anomalias e da sua etiquetagem. Embora seja possível identificar desvios e padrões atípicos com estes métodos não supervisionados, ainda não existe validação que permita concluir se correspondem a anomalias operacionais ou a variações normais de consumo. Para isso vai ser necessária colaboração entre quem faz a análise de dados e os responsáveis pelas cargas para confirmar eventos no terreno, registar causa e impacto e etiquetar ocorrências, criando a base para afinar limiares e no futuro treinar classificadores supervisionados.

Do ponto de vista económico o sistema cria valor de forma clara. Ajuda a reduzir consumos ao permitir decisões rápidas perante desvios e anomalias o que evita desperdícios e penalizações. Baixa os custos operacionais porque concentra a informação num único ponto e automatiza análises que antes eram manuais. Torna o consumo mais consciente ao disponibilizar dados em tempo real à comunidade académica e apoia a gestão da procura ao facilitar a deslocação de cargas para períodos com tarifação mais reduzida.

O retorno neste caso pode ser medido através da contagem do consumo de energia mitigado por cada ação desencadeada pelo sistema, pelo valor das penalizações evitadas por potência reativa e por picos excessivos, pelas horas técnicas poupadas em recolha validação e reporte dos dados e pela redução do tempo médio até à resolução de incidentes.

5

CONCLUSÕES E TRABALHOS FUTUROS

A implementação da nova arquitetura de monitorização energética consolidou um sistema robusto, capaz de responder às necessidades institucionais tanto em termos de recolha de dados como de integridade e consistência das séries temporais. A unificação do código de aquisição, a integração de diferentes dispositivos de medição e a criação de uma base de dados estruturada constituíram avanços decisivos, permitindo ultrapassar limitações anteriormente associadas a perdas de registos, heterogeneidade de variáveis e lacunas de armazenamento.

O sistema encontra-se atualmente numa fase de maturidade que alia estabilidade operacional a uma camada analítica. A automatização dos processos de aquisição e cálculo de custos, a incorporação de dados meteorológicos e a utilização de modelos de previsão colocaram o projeto para além da simples monitorização, abrindo caminho para uma abordagem orientada à análise. A deteção de anomalias acrescentou uma dimensão crítica ao diagnóstico, ainda que dependente de validação adicional para garantir clareza na explicação dos resultados.

Em síntese, o trabalho desenvolvido demonstra a viabilidade de um sistema de monitorização energética modular, adaptado ao contexto académico e com relevância prática comprovada. A arquitetura implementada constitui uma base sólida para a evolução futura, permitindo transformar a monitorização em gestão ativa, sustentada em dados de elevada resolução temporal e qualidade assegurada pela precisão dos medidores em uso.

5.1 TRABALHOS FUTUROS

A continuidade do projeto deverá incidir no aperfeiçoamento da componente analítica e na expansão gradual da infraestrutura de monitorização a novos pontos de medição. Com o culminar desta futura expansão, por exemplo através da inclusão de novos edifícios ou partição dos já monitorizados, o número de tabelas na base de dados pode crescer exponencialmente, o que torna aconselhável a evolução do modelo atual com uma tabela por edifício para um modelo com o esquema relacional normalizado como o ilustrado na Figura 41.

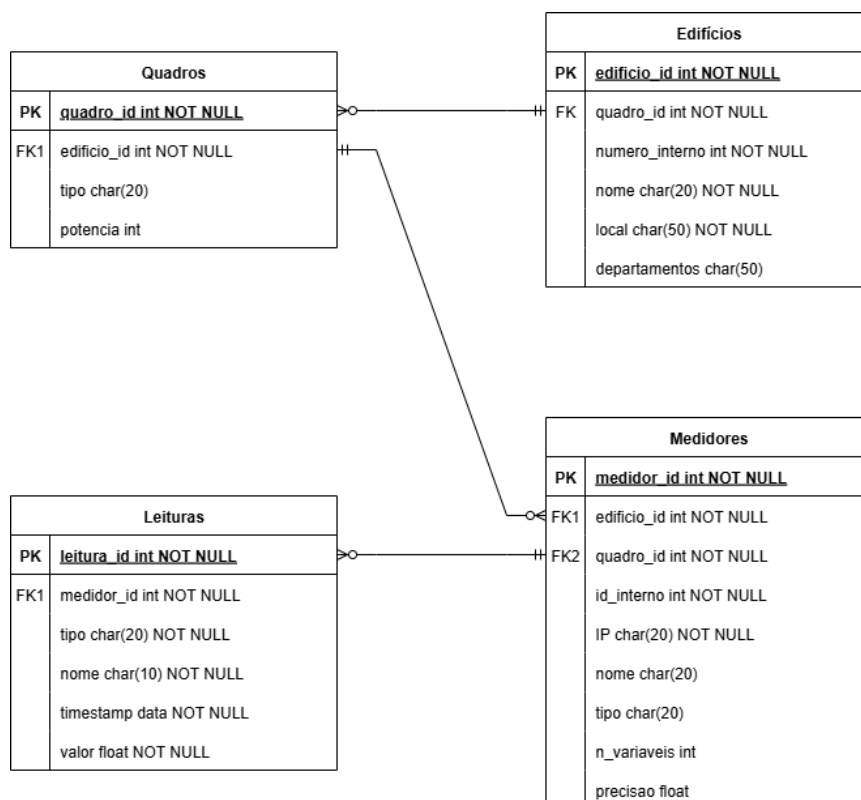


Figura 41 – Proposta de esquema relacional da base de dados no futuro

Este modelo normaliza as entidades de contexto Edifícios, Quadros e Medidores e desacopla as séries temporais para a tabela Leituras, assegurando integridade referencial, eliminação de redundância e escalabilidade da ingestão contínua. A organização num conjunto reduzido de tabelas relacionadas facilita ainda consultas transversais com otimização por índices e partições temporais.

No domínio da previsão, é de importância elevada a tentativa de refinar ainda mais o modelo já implementado e experimentar outros para comparação de resultados. Deve ainda aprofundar-se a engenharia de variáveis, não só incorporando novas variáveis exógenas com maior correlação com a variável alvo, como também explorando de forma sistemática desfasamentos temporais dessas variáveis relativamente à potência ativa, de modo a captar

atrasos físicos e operacionais na resposta de consumo. A precisão poderá aumentar de forma significativa com a inclusão de informação sobre a atividade interna dos edifícios, como o número de utilizadores ligados à rede ou outros indicadores de ocupação. Relativamente à deteção de anomalias, torna-se necessário o estudo das deteções e a sua associação com problemas reais ocorrentes no edifício.

A expansão da camada de aprendizagem automática a novos edifícios e a disponibilização de *dashboards* dedicados no Grafana para todos os pontos de monitorização são prioridades imediatas, estendendo a cobertura e a capacidade analítica demonstradas no edifício piloto a todos os edifícios da instituição.

A partição interna dos edifícios por zonas ou departamentos permitirá igualmente alcançar uma visão mais granular dos consumos, identificando áreas críticas e facilitando a definição de estratégias de eficiência energética.

Por fim, a disponibilização dos *dashboards* em dispositivos de exibição local, através da programação de dispositivos Raspberry Pi ligados a ecrãs instalados nas entradas dos edifícios, poderá aproximar a monitorização dos utilizadores e fomentar uma maior consciencialização energética em toda a comunidade académica.

REFERÊNCIAS

- [1] M. Domínguez, J. I. San Martín, A. Perallos e J. A. Gutiérrez, “Power monitoring system for university buildings,” **Energy and Buildings**, vol. 66, pp. 72–77, Nov. 2013, doi: 10.1016/j.enbuild.2013.07.013. [Online]. Disponível em: <https://doi.org/10.1016/j.enbuild.2013.07.013>. [Acedido em: 19-Set-2025].
- [2] GAIA Project, “Green Awareness in Action,” *Horizon 2020*, 2019. [Online]. Disponível em: <https://gaia-project.eu>. [Acedido em: 19-Set-2025].
- [3] G. Mylonas, D. Amaxilatis, S. Tsampas, L. Pocero e J. Gunneriusson, “A Methodology for Saving Energy in Educational Buildings Using an IoT Infrastructure,” *arXiv preprint arXiv:1907.07760*, 2019. [Online]. Disponível em: <https://arxiv.org/abs/1907.07760>. [Acedido em: 19-Set-2025].
- [4] E. C. Quispe, J. A. P. Guerrero, J. C. C. Caicedo e S. Quispe, “Energy management systems in higher education institutions’ buildings,” **Energies**, vol. 18, no. 7, p. 1810, 2025, doi: 10.3390/en18071810. [Online]. Disponível em: <https://doi.org/10.3390/en18071810>. [Acedido em: 19-Set-2025].
- [5] Cranfield University, “Carbon and energy management — ISO 50001 energy management system certification,” 2019. [Online]. Disponível em: <https://www.cranfield.ac.uk>. [Acedido em: 19-Set-2025].
- [6] M. Domínguez, J. J. Fuertes, S. Alonso, M. A. Prada, A. Morán e P. Barrientos, “Power monitoring system for university buildings: Architecture and advanced analysis tools,” *Energy and Buildings*, vol. 59, pp. 152–160, Apr. 2013. [Online]. Disponível em: <https://doi.org/10.1016/j.enbuild.2012.12.032>. [Acedido em: 19-Set-2025].
- [7] J. Ma, S. Jørgensen, A. W. Colombo e R. Harrison, “Energy management systems for smart buildings: A review and case studies,” *Energy and AI*, vol. 1, Art. no. 100005, 2020. [Online]. Disponível em: <https://doi.org/10.1016/j.egyai.2020.100005>. [Acedido em: 19-Set-2025].
- [8] M. F. M. Elias, M. A. M. Radzi, N. F. Mailah e M. Z. A. Ab Kadir, “Low-Cost, Open-Source, Emoncms-Based SCADA System for a PV Plant with Local Data Logging,” *Sensors*, vol. 22, no. 19, p. 7461, 2022. [Online]. Disponível em: <https://doi.org/10.3390/s22197461>. [Acedido em: 19-Set-2025].
- [9] J. Figueiredo e J. Sá da Costa, “A SCADA system for energy management in intelligent buildings,” **Energy and Buildings**, vol. 49, pp. 85–98, Jun. 2012, doi: 10.1016/j.enbuild.2012.01.041. [Online]. Disponível em: <https://doi.org/10.1016/j.enbuild.2012.01.041>. [Acedido em: 19-Set-2025].
- [10] K. Heussen et al., “ERIGrid Holistic Test Description for Validating Cyber-Physical Energy Systems,” *Energies*, vol. 12, no. 14, p. 2722, 2019. [Online]. Disponível em: <https://doi.org/10.3390/en12142722>. [Acedido em: 19-Set-2025].
- [11] Power Factors, “EMS and SCADA Provide the Backbone for One of the Largest Solar Plus Storage Projects,” *Power Factors*, 2022. [Online]. Disponível em: <https://www.powerfactors.com/customer-stories/power-factors-ems-and-scada-provide-the-backbone-for-one-of-the-largest-solar-plus-storage-projects>. [Acedido em: 19-Set-2025].
- [12] GreenPowerMonitor, “GreenPowerMonitor Introduces New Energy Management System for Enhanced Renewable Power Plant Control,” *GreenPowerMonitor*, 2023. [Online]. Disponível em: <https://www.greenpowermonitor.com/news/greenpowermonitor-introduces-new-energy-management-system-for-enhanced-renewable-power-plant-control>. [Acedido em: 19-Set-2025].

-
- [13] M. N. Dat, K. D. Trung, P. V. Minh, C. D. Van, Q. T. Tran e T. N. Ngoc, “Assessment of Energy Efficiency Using an Energy Monitoring System: A Case Study of a Major Energy-Consuming Enterprise in Vietnam,” *Energies*, vol. 16, no. 13, p. 5214, 2023. [Online]. Disponível em: <https://doi.org/10.3390/en16135214>. [Acedido em: 19-Set-2025].
- [14] CEN – European Committee for Standardization, *EN 17267:2019 – Energy measurement and monitoring plan – Design and implementation – Principles for energy data collection*, Brussels, Belgium: CEN, 2019. [Online]. Disponível em: <https://webstore.ansi.org/standards/ds/dsen172672019>. [Acedido em: 19-Set-2025].
- [15] M. B. V. Granadeiro, “Sistema de monitorização energética e comportamental SMEC: caracterização do comportamento adaptativo humano e influência no conforto e consumo de energia de um gabinete,” Dissertação de Mestrado, Faculdade de Ciências, Universidade de Lisboa, 2017. [Online]. Disponível em: <http://hdl.handle.net/10451/31516>. [Acedido em: 19-Set-2025].
- [16] Modbus Organization, “Modbus Application Protocol Specification V1.1b3,” 2006. [Online]. Disponível em: https://modbus.org/docs/Modbus_Application_Protocol_V1_1b3.pdf. [Acedido em: 19-Set-2025].
- [17] Open4Tech, “Application of the Modbus Protocol with Microcontrollers,” 2020. [Online]. Disponível em: <https://open4tech.com/application-modbus-protocol-microcontrollers/>. [Acedido em: 19-Set-2025].
- [18] Texas Instruments, “RS-485 basics: Introduction,” 2015. [Online]. Disponível em: <https://www.ti.com/lit/pdf/ssztc4>. [Acedido em: 19-Set-2025].
- [19] Simply Modbus, “Modbus TCP/IP Explained,” 2023. [Online]. Disponível em: <https://www.simplymodbus.ca/TCP.htm>. [Acedido em: 19-Set-2025].
- [20] F. Ferrández-Pastor, H. Mora e C. Fritzen, “RESTful IoT Integration Frameworks for Smart Energy Monitoring and Management Systems,” *Sensors*, vol. 22, no. 3, p. 1048, 2022. [Online]. Disponível em: <https://www.mdpi.com/1424-8220/22/3/1048>. [Acedido em: 19-Set-2025].
- [21] Make Computer Science Great Again, “Understanding REST API – A Comprehensive Guide,” *Medium*, 2021. [Online]. Disponível em: <https://medium.com/@MakeComputerScienceGreatAgain/understanding-rest-api-a-comprehensive-guide-52fc10f6c9ed>. [Acedido em: 19-Set-2025].
- [22] H. Mann, “REST API Basics – 4 Things you Need to Know,” *mannhowie.com*, Mar. 24, 2023. [Online]. Disponível em: <https://mannhowie.com/rest-api>. [Acedido em: 19-Set-2025].
- [23] A. Manowska, A. Wycisk, A. Nowrot e J. Pielot, “The use of the MQTT protocol in measurement, monitoring and control systems as part of the implementation of energy management systems,” **Electronics**, vol. 12, no. 1, p. 17, Jan. 2023, doi: 10.3390/electronics12010017. [Online]. Disponível em: <https://doi.org/10.3390/electronics12010017>. [Acedido em: 19-Set-2025].
- [24] Nordic Semiconductor, “MQTT Protocol – Lesson 4, Cellular IoT Fundamentals,” *Nordic Developer Academy*, 2023. [Online]. Disponível em: <https://academy.nordicsemi.com/courses/cellular-iot-fundamentals/lessons/lesson-4-cellular-fundamentals/topic/lesson-4-mqtt-protocol/>. [Acedido em: 19-Set-2025].
- [25] International Electrotechnical Commission, *IEC 62053-21:2020 – Electricity metering equipment (a.c.) – Particular requirements – Part 21: Static meters for active energy (classes 0.5, 1 and 2)*. Geneva, Switzerland: IEC, 2020. [Online]. Disponível em: <https://webstore.iec.ch/en/publication/28660>. [Acedido em: 19-Set-2025].
- [26] International Electrotechnical Commission, *IEC 62053-22:2020 – Electricity metering equipment (a.c.) – Particular requirements – Part 22: Static meters for active energy (classes 0.1S, 0.2S and 0.5S)*. Geneva, Switzerland: IEC, 2020. [Online]. Disponível em: <https://webstore.iec.ch/en/publication/29987>. [Acedido em: 19-Set-2025].

-
- [27] Schneider Electric, *PowerLogic Power and Energy Meters Catalog*, Document No. 3000CT1701, Jun. 2019. [Online]. Disponível em: <https://www.se.com/us/en/download/document/3000CT1701/>. [Acedido em: 19-Set-2025].
- [28] Schneider Electric, “PowerLogic PM710 Datasheet,” 2023. [Online]. Disponível em: <https://www.se.com/id/en/product/PM710MG/power-meter-pm710-basic-readings-thd-+-min-max-+-rs485/>. [Acedido em: 19-Set-2025].
- [29] Schneider Electric, “PM5100 / PM5110 – Product Information,” 2023. [Online]. Disponível em: <https://www.se.com/uk/en/product/METSEPM5100/pm5100-meter-without-communication-up-to-15th-h-1do-33-alarms/>. [Acedido em: 19-Set-2025].
- [30] Schneider Electric, “PowerLogic PM3250 - Technical Datasheet,” 2023. [Online]. Disponível em: <https://www.se.com/pt/pt/product/METSEPM3250/pm3250-contador-de-energia-rs485/>. [Acedido em: 19-Set-2025].
- [31] R. El-Khozondar, H. El-Khozondar, S. Alajerami, A. Al-Agha e M. Qandil, “A smart energy monitoring system using ESP32 microcontroller,” **e-Prime – Advances in Electrical Engineering, Electronics and Energy**, vol. 4, 100288, 2024, doi: 10.1016/j.prime.2024.100288. [Online]. Disponível em: <https://doi.org/10.1016/j.prime.2024.100288>. [Acedido em: 19-Set-2025].
- [32] Shelly, “Shelly EM - Wi-Fi Operated Energy Meter and Contactor Control,” *Shelly*, 2024. [Online]. Disponível em: <https://kb.shelly.cloud/knowledge-base/shelly-em>. [Acedido em: 19-Set-2025].
- [33] Shelly, “Shelly 3EM - Professional 3-Phase Energy Meter,” *Shelly*, 2024. [Online]. Disponível em: <https://kb.shelly.cloud/knowledge-base/shelly-3em>. [Acedido em: 19-Set-2025].
- [34] Shelly, “Shelly Pro 3EM - Professional 3-Phase Energy Meter,” *Shelly Knowledge Base*, 2023. [Online]. Disponível em: <https://kb.shelly.cloud/knowledge-base/shelly-pro-3em>. [Acedido em: 19-Set-2025].
- [35] Huawei, “SUN2000 Hybrid Inverter – Product Documentation,” *Huawei Solar*, 2024. [Online]. Disponível em: <https://solar.huawei.com/pt/professionals/all-products>. [Acedido em: 19-Set-2025].
- [36] Solar-Log GmbH, “Solar-Log™ WEB Enerest™ 4 – PV Monitoring Solutions,” 2024. [Online]. Disponível em: <https://www.solar-log.com/en/products-solutions/monitoring-solar-log-web-eneresttm>. [Acedido em: 19-Set-2025].
- [37] Tecno-Logistica, “Solar-Log 2000,” *Tecno-Logistica*, 2024. [Online]. Disponível em: <https://tecno-logistica.it/negozio/it/accessori/480-solar-log-2000.html>. [Acedido em: 19-Set-2025].
- [38] International Electrotechnical Commission, *IEC 61000-4-30:2015 – Testing and measurement techniques – Power quality measurement methods*. Geneva, Switzerland: IEC, 2015. [Online]. Disponível em: <https://webstore.iec.ch/publication/22215>. [Acedido em: 19-Set-2025].
- [39] C. I. Ciontea e F. Iov, “A Study of Load Imbalance Influence on Power Quality Assessment for Distribution Networks,” **Electricity**, vol. 2, no. 1, pp. 77–90, 2021, doi: 10.3390/electricity2010005. [Online]. Disponível em: <https://doi.org/10.3390/electricity2010005>. [Acedido em: 19-Set-2025].
- [40] IEEE, *IEEE Standard 1459-2010: Definitions for the Measurement of Electric Power Quantities Under Sinusoidal, Non-sinusoidal, Balanced, or Unbalanced Conditions*. New York, USA: IEEE, 2010. [Online]. Disponível em: <https://doi.org/10.1109/IEEESTD.2010.5439063>. [Acedido em: 19-Set-2025].
- [41] IEEE, *IEEE Recommended Practice for Monitoring Electric Power Quality – IEEE Std 1159-2019*. New York, USA: IEEE, 2019. [Online]. Disponível em: <https://doi.org/10.1109/IEEESTD.2019.8758232>. [Acedido em: 19-Set-2025].

-
- [42] J. Ting, B. Wang, R. W. Kenyon e A. Hoke, “Evaluating Methods for Measuring Grid Frequency in Low-Inertia Power Systems,” *National Renewable Energy Laboratory (NREL)*, Report No. 81400, 2022. [Online]. Disponível em: <https://www.nrel.gov/docs/fy23osti/81400.pdf>. [Acedido em: 19-Set-2025].
- [43] IEEE, *IEEE Recommended Practice and Requirements for Harmonic Control in Electric Power Systems – IEEE Std 519-2014*. New York, USA: IEEE, 2014. [Online]. Disponível em: <https://doi.org/10.1109/IEEESTD.2014.6826459>. [Acedido em: 19-Set-2025].
- [44] OpenEnergyMonitor, *EmonCMS User Guide*, OpenEnergyMonitor Documentation, 2023. [Online]. Disponível em: <https://docs.openenergymonitor.org/emoncms/index.html>. [Acedido em: 19-Set-2025].
- [45] C. Alvarado, D. Carrera e A. L. Martínez, “Energy monitoring system using EmonCMS and Node-RED for industrial facilities,” **Energies**, vol. 14, no. 19, p. 6357, Sep. 2021, doi: 10.3390/en14196357. [Online]. Disponível em: <https://doi.org/10.3390/en14196357>. [Acedido em: 19-Set-2025].
- [46] F. Z. Kausar, A. H. Abdullah, M. F. Zolkapli e R. Jidin, “Development of smart energy monitoring system based on open-source platform,” **Journal of Telecommunication, Electronic and Computer Engineering (JTEC)**, vol. 10, no. 1–10, pp. 51–55, 2018. [Online]. Disponível em: <https://jtec.utem.edu.my/jtec/article/view/4420>. [Acedido em: 19-Set-2025].
- [47] Grafana Labs, “About Grafana,” *Grafana Documentation*, 2024. [Online]. Disponível em: <https://grafana.com/docs/grafana/latest/introduction/>. [Acedido em: 19-Set-2025].
- [48] Grafana Labs, “Grafana,” 2024. [Online]. Disponível em: <https://grafana.com/grafana/>. [Acedido em: 19-Set-2025].
- [49] Grafana Labs, “Grafana Dashboards – Examples,” 2024. [Online]. Disponível em: <https://grafana.com/grafana/dashboards>. [Acedido em: 19-Set-2025].
- [50] MathWorks, “ThingSpeak IoT Analytics Platform,” 2024. [Online]. Disponível em: <https://thingspeak.com>. [Acedido em: 19-Set-2025].
- [51] Arduino, “Arduino Cloud,” 2024. [Online]. Disponível em: <https://cloud.arduino.cc/>. [Acedido em: 19-Set-2025].
- [52] Home Assistant, “Monitorização de Energia com Home Assistant,” 2024. [Online]. Disponível em: <https://www.home-assistant.io/docs/energy/>. [Acedido em: 19-Set-2025].
- [53] PostgreSQL Global Development Group, “PostgreSQL 17.5 Documentation,” 2025. [Online]. Disponível em: <https://www.postgresql.org/docs/current/index.html>. [Acedido em: 19-Set-2025].
- [54] Paradigma Software, “Valentina Studio – Database Management,” 2025. [Online]. Disponível em: <https://valentina-db.com/en/database-management>. [Acedido em: 19-Set-2025].
- [55] InfluxData, “InfluxDB for Energy and Utilities,” *InfluxData*, 2024. [Online]. Disponível em: <https://www.influxdata.com/solutions/by-industries/energy-and-utilities/>. [Acedido em: 19-Set-2025].
- [56] M. Giacobbe, C. Chaouch, M. L. Scarpa e A. Puliafito, “An Implementation of InfluxDB for Monitoring and Analytics in Distributed IoT Environments,” in *Proc. Int. Conf. on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT)*, 2018. [Online]. Disponível em: https://www.researchgate.net/publication/334399754_An_Implementation_of_InfluxDB_for_Monitoring_and_Analytics_in_Distributed_IoT_Environments. [Acedido em: 19-Set-2025].
- [57] InfluxData, “Using Deduplication for Eventually Consistent Transactions,” *InfluxData Blog*, 2023. [Online]. Disponível em: <https://www.influxdata.com/blog/using-deduplication-eventually-consistent-transactions/>. [Acedido em: 19-Set-2025].

-
- [58] Timescale Inc., “Hypertables – Timescale Documentation,” *Timescale*, 2025. [Online]. Disponível em: <https://docs.timescale.com/use-timescale/latest/hypertables/>. [Acedido em: 19-Set-2025].
- [59] SQLite Consortium, “SQLite Documentation,” *SQLite*, 2024. [Online]. Disponível em: <https://www.sqlite.org/docs.html>. [Acedido em: 19-Set-2025].
- [60] D. A. V. Romero, E. Villalvazo Laureano, R. O. J. Betancourt e E. Navarro Álvarez, “An open source IoT edge-computing system for monitoring energy consumption in buildings,” *Results in Engineering*, vol. 21, art. 101875, 2024, doi: 10.1016/j.rineng.2024.101875. [Online]. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2590123024001282>. [Acedido em: 19-Set-2025].
- [61] Canonical Ltd., “Ubuntu release cycle,” *Ubuntu*, 2025. [Online]. Disponível em: <https://ubuntu.com/about/release-cycle>. [Acedido em: 19-Set-2025].
- [62] freedesktop.org, “systemd for Administrators,” *Systemd Documentation*, 2025. [Online]. Disponível em: <https://www.freedesktop.org/wiki/Software/systemd/>. [Acedido em: 19-Set-2025].
- [63] OpenSSH, “OpenSSH Manual Pages – ssh and scp,” *OpenSSH Project*, 2025. [Online]. Disponível em: <https://www.openssh.com/manual.html>. [Acedido em: 19-Set-2025].
- [64] Canonical Ltd., “CronHowto – Ubuntu Community Help Wiki,” *Ubuntu Documentation*, 2023. [Online]. Disponível em: <https://help.ubuntu.com/community/CronHowto>. [Acedido em: 19-Set-2025].
- [65] Canonical Ltd., “How to back up using shell scripts,” *Ubuntu Documentation*, 2025. [Online]. Disponível em: <https://documentation.ubuntu.com/server/how-to/backups/back-up-using-shell-scripts/>. [Acedido em: 19-Set-2025].
- [66] J. Fernández-Morales et al., “Statistical Dataset and Data Acquisition System for Monitoring the Voltage and Frequency of the Electrical Network in an Environment Based on Python and Grafana,” *Data*, vol. 7, no. 6, p. 77, 2022. [Online]. Disponível em: <https://www.mdpi.com/2306-5729/7/6/77>. [Acedido em: 19-Set-2025].
- [67] B. Heidrich et al., “pyWATTS: Python Workflow Automation Tool for Time Series,” *arXiv preprint arXiv:2106.10157*, 2021. [Online]. Disponível em: <https://arxiv.org/abs/2106.10157>. [Acedido em: 19-Set-2025].
- [68] L. Thurner et al., “pandapower: an Open Source Python Tool for Convenient Modeling, Analysis and Optimization of Electric Power Systems,” *arXiv preprint arXiv:1709.06743*, 2017. [Online]. Disponível em: <https://arxiv.org/abs/1709.06743>. [Acedido em: 19-Set-2025].
- [69] T. Lea, “emon-tools: Python tools for reading and writing EmonCMS feed data,” *OpenEnergyMonitor, GitHub Repository*, 2021. [Online]. Disponível em: <https://github.com/emoncms/emon-tools>. [Acedido em: 19-Set-2025].
- [70] ERSE – Entidade Reguladora dos Serviços Energéticos, *Tarifas e Preços para a Energia Elétrica e Outros Serviços em 2025*, Lisboa, Portugal, Dez. 2024. [Online]. Disponível em: <https://www.erse.pt/media/xczdo4as/tep-se-2025.pdf>. [Acedido em: 19-Set-2025].
- [71] EDP Comercial, “Como calcular o consumo de energia elétrica da sua empresa?,” *EDP Soluções*, 2023. [Online]. Disponível em: <https://solucoes.edp.com.br/blog/consumo-energia-eletrica-empresas/>. [Acedido em: 19-Set-2025].
- [72] H. M. Ali, A. E. Kabeel e R. Saidur, “Optimization of dynamic electricity tariffs and real-time energy cost calculation in smart grids,” *Scientific Reports*, vol. 15, no. 12234, pp. 1–14, 2025. [Online]. Disponível em: <https://doi.org/10.1038/s41598-025-12234>. [Acedido em: 19-Set-2025].
- [73] C. Chen, Z. Gao, X. Zhou, M. Wang e J. Yan, “Energy consumption prediction and energy-saving suggestions of public buildings based on machine learning,” **Energy and Buildings**, vol. 320, art. 114585, 2024, doi: 10.1016/j.enbuild.2024.114585. [Online]. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378778824007011>. [Acedido em: 19-Set-2025].

-
- [74] S. J. Taylor e B. Letham, “Forecasting at Scale,” *The American Statistician*, vol. 72, no. 1, pp. 37-45, 2018, doi: 10.1080/00031305.2017.1380080. [Online]. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/00031305.2017.1380080> [Acedido em: 19-Set-2025].
- [75] F. T. Liu, K. M. Ting e Z.-H. Zhou, “Isolation Forest,” in *Proceedings of the 2008 Eighth IEEE International Conference on Data Mining (ICDM)*, Pisa, Italy, 15-19 Dez. 2008, pp. 413-422, doi: 10.1109/ICDM.2008.17. [Online]. Disponível em: <https://doi.org/10.1109/ICDM.2008.17> [Acedido em: 19-Set-2025].
- [76] T. Mendes, P. J. S. Cardoso, J. Monteiro e J. Raposo, “Anomaly Detection of Consumption in Hotel Units: A Case Study Comparing Isolation Forest and Variational Autoencoder Algorithms,” *Applied Sciences*, vol. 13, no. 1, art. 314, 2023, doi: 10.3390/app13010314. [Online]. Disponível em: <https://doi.org/10.3390/app13010314> [Acedido em: 19-Set-2025].
- [77] M. M. Breunig, H.-P. Kriegel, R. T. Ng e J. S. Sander, “LOF: Identifying Density-Based Local Outliers,” in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, Dallas, TX, USA, 2000, pp. 93-104, doi: 10.1145/342009.335388. [Online]. Disponível em: <https://doi.org/10.1145/342009.335388> [Acedido em: 19-Set-2025].
- [78] G. E. P. Box, G. M. Jenkins, G. C. Reinsel e G. M. Ljung, *Time Series Analysis: Forecasting and Control*, 5th ed. Hoboken, NJ, USA: Wiley, 2015.
- [79] K. Szostek, J. Nowakowski, R. Jachimczyk e A. Gola, “Analysis of the Effectiveness of ARIMA, SARIMA, and SVR in the Forecasting of Wind Farm Generation in Poland,” *Energies*, vol. 17, no. 19, p. 4803, 2024, doi: 10.3390/en17194803. [Online]. Disponível em: <https://doi.org/10.3390/en17194803>. [Acedido em: 19-Set-2025].
- [80] S. J. Taylor e B. Letham, *Prophet: Forecasting at Scale*. Meta Open Source, 2017. [Online]. Disponível em: <https://facebook.github.io/prophet/>. [Acedido em: 19-Set-2025].
- [81] S. Sulandari, S. Yusuf, R. I. Kurnianto, M. H. Purnomo e H. Hermawan, “Implementation of Prophet in American Electricity Forecasting with and without Parameter Tuning,” *International Journal of Artificial Intelligence Research*, vol. 8, no. 1, pp. 60–72, 2024. [Online]. Disponível em: <https://www.researchgate.net/publication/387431551>. [Acedido em: 19-Set-2025].
- [82] A. Albahli, “LSTM vs. Prophet: Achieving Superior Accuracy in Short-Term Electricity Consumption Forecasting in Ontario, Canada,” *Energies*, vol. 18, no. 2, p. 278, 2025, doi: 10.3390/en18020278. [Online]. Disponível em: <https://doi.org/10.3390/en18020278>. [Acedido em: 19-Set-2025].
- [83] P. Labura, T. Antic e T. Capuder, “Time and Frequency Domain-based Anomaly Detection in Smart Meter Data for Distribution Network Studies,” *arXiv preprint arXiv:2504.18231*, 2025. [Online]. Disponível em: <https://arxiv.org/abs/2504.18231>. [Acedido em: 19-Set-2025].
- [84] N. Zangrando, P. Fraternali, M. Petri, N. O. P. Vago, S. L. H. González, et al., “Anomaly detection in quasi-periodic energy consumption data series: a comparison of algorithms,” *Energy Informatics*, vol. 5, Supplement 4, art. number 62, 2022. [Online]. Disponível em: <https://energyinformatics.springeropen.com/articles/10.1186/s42162-022-00230-7>. [Acedido em: 19-Set-2025].
- [85] R. J. Hyndman e A. B. Koehler, “Another look at measures of forecast accuracy,” *International Journal of Forecasting*, vol. 22, n.º 4, pp. 679–688, 2006. doi: 10.1016/j.ijforecast.2006.03.001. Disponível em: Elsevier/ScienceDirect. [Acedido em: 19-Set-2025].
- [86] R. J. Hyndman e G. Athanasopoulos, *Forecasting: Principles and Practice*, 3.ª ed. OTexts, 2021. doi: —. Disponível em: otexts.com/fpp3. [Acedido em: 19-Set-2025]
- [87] T. Gneiting e A. E. Raftery, “Strictly Proper Scoring Rules, Prediction, and Estimation,” *Journal of the American Statistical Association*, vol. 102, n.º 477, pp. 359–378, 2007. doi: 10.1198/016214506000001437. Disponível em: Taylor & Francis / cópia de autor: sites.stat.washington.edu/raftery/Research/PDF/Gneiting2007jasa.pdf. [Acedido em: 19-Set-2025].

-
- [88] T. Gneiting e A. E. Raftery, “Probabilistic forecasts, calibration and sharpness,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 69, n.º 2, pp. 243–268, 2007. doi: 10.1111/j.1467-9868.2007.00587.x. Disponível em: <https://sites.stat.washington.edu/raftery/Research/PDF/Gneiting2007jrssb.pdf>. [Acedido em: 19-Set-2025].
- [89] D. C. Montgomery, *Introduction to Statistical Quality Control*, 7.^a ed. Hoboken, NJ: Wiley, 2012. ISBN: 978-1118146811. Disponível em: <https://www.uaar.edu.pk/fs/books/12.pdf>. [Acedido em: 19-Set-2025].
- [90] B. Yurdakul e J. D. Naranjo, “Statistical properties of the population stability index,” *Journal of Risk Model Validation*, vol. 14, n.º 4, pp. 89–100, 2020. doi: 10.21314/JRMV.2020.227. Disponível em: <https://www.risk.net/journal-of-risk-model-validation/7725371/statistical-properties-of-the-population-stability-index>. [Acedido em: 19-Set-2025].
- [91] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*. London: Springer, 2001. ISBN: 978-1-85233-459-8. doi: 10.1007/978-1-4471-3675-0. Disponível em: <https://link.springer.com/book/10.1007/978-1-4471-3675-0>. [Acedido em: 19-Set-2025].
- [92] H. Chen, M. Zhu, X. Hu, J. Wang, Y. Sun, J. Yang, B. Li e X. Meng, “Multifeature Short-Term Power Load Forecasting Based on GCN-LSTM,” *International Transactions on Electrical Energy Systems*, vol. 2023, art. 8846554, 2023. doi: 10.1155/2023/8846554. Disponível em: <https://doi.org/10.1155/2023/8846554>. [Acedido em: 19-Set-2025].
- [93] S. S. Sandha, M. Aggarwal, I. Fedorov e M. Srivastava, “MANGO: A Python Library for Parallel Hyperparameter Tuning,” *arXiv preprint arXiv:2005.11394*, 2020. doi: 10.48550/arXiv.2005.11394. Disponível em: <https://arxiv.org/abs/2005.11394>. [Acedido em: 19-Set-2025].
- [94] J.-A. Solís-Villarreal, V. Soto-Mendoza, J. A. Navarro-Acosta e E. Ruiz-y-Ruiz, “Energy Consumption Outlier Detection with AI Models in Modern Cities: A Case Study from North-Eastern Mexico,” *Algorithms*, vol. 17, n.º 8, art. 322, 2024. doi: 10.3390/a17080322. Disponível em: <https://www.mdpi.com/1999-4893/17/8/322>. [Acedido em: 19-Set-2025].

APÊNDICES

- Apêndice A – Tabela de Análise da Integridade dos Registos
- Apêndice B – Tabela do Top10 de Combinações de Hiperparâmetros



TABELA DE ANÁLISE DA INTEGRIDADE DOS REGISTOS

Nº Edifício	Nome	Tipo de Medição	Nº Registos total	Nº Registos em falta	Nº Registos Inválidos	% de Registos Válidos
1	FCT CINTAL	Consumo	3237597	41329	271179	91.62%
2	FMCB CBMR	Consumo	3237568	41330	274284	91.53%
3	CGD Portaria	Consumo	3210665	68219	273152	91.49%
4	Complexo Pedagógico Gamb.	Consumo	3210664	68214	277471	91.36%
5	Biblioteca/Reitoria	Consumo	3210661	68215	275489	91.42%
6	Cantina Gambelas	Consumo	3210658	68215	275329	91.42%
7	CCMAR	Consumo	3209347	69534	607344	81.08%
8	FE FCT	Consumo	3209349	69536	606816	81.09%
8	FE FCT	Produção	2629642	69597	579736	81.94%
9	Horto e Leoa	Consumo	3237601	41330	272909	91.57%
11	Pavilhão Madeira Jotas	Consumo	3237613	41292	311360	90.38%
12	Zona 2 Pavilhão Alvenaria D H F	Consumo	3237611	41287	306409	90.54%
12	Zona 3 Pavilhão Alvenaria E	Consumo	3237607	41294	307537	90.50%
12	Zona 5 Biotério	Consumo	3237611	41299	314335	90.29%
12	Zona 6 Pavilhão Alvenaria G	Consumo	3237612	41300	314270	90.29%
12	Zona 7 Pavilhão L Bar Álvaro	Consumo	3237613	41303	314863	90.27%
12	Zona Geral + Zonas 2 a 7	Consumo	323765	41281	299908	90.74%
22	ESGHT	Consumo	3552907	38149	2801147	21.16%
22	ESGHT	Produção	1378906	40061	253049	84.49%
23	ESEC	Consumo	3552913	38151	2815432	20.76%
23	ESEC	Produção	1385680	40062	246275	84.91%
26	Cantina Penha	Consumo	3552913	38149	525535	85.21%
27	ISE – Edifício Principal	Consumo	3552925	38151	499198	85.95%
27	ISE – Edifício Principal	Produção	1354311	40062	277644	82.99%
27	ISE – Edifício Principal	Meteorologia	2543007	1431	733046	77.62%
28	ISE – Edifício em U	Consumo	535644	38151	3017287	15.08%
28	ISE – Edifício em U	Produção	774991	40062	856964	47.49%
28	ISE – Edifício em U	Bateria Condensadores	567811	38153	41314	92.72%
29	Complexo Pedagógico Penha	Consumo	3552919	38151	498679	85.96%
30	Tec Campus	Consumo	606056	39865	62555	90.64%
30	Tec Campus	Produção	668635	39844	55972	91.63%
32	Edifício Engenharia Civil	Consumo	3552898	38159	537406	84.87%
33	Serviços Técnicos / Oficinas	Consumo	3552901	38149	546308	84.62%
34	Residência Penha	Consumo	3552909	38150	536056	84.91%
Exterior	Iluminação exterior Penha	Consumo	3552917	38150	500156	85.92%
Exterior	Iluminação Exterior Gambelas PT1	Consumo	3237614	41330	271810	91.60%
Exterior	Iluminação Exterior Gambelas PT2	Consumo	3210684	68217	272756	91.50%
TOTAL de medidores: 37						

Quantificação dos medidores e análise preliminar da integridade dos registos por edifício realizada no dia 16 de agosto de 2025

B

TABELA DO TOP 10 DE COMBINAÇÕES DE HIPERPARÂMETROS

HIPERPARÂMETROS/MÉTRICAS	TOP1	TOP2	TOP3	TOP4	TOP5	TOP6	TOP7	TOP8	TOP 9	TOP 10
CHANGEPOINT_PRIOR_SCALE	0.03314	0.02067	0.01408	0.17182	0.05702	0.05846	0.01089	0.08756	0.03064	0.01363
CHANGEPOINT_RANGE	0.88956	0.90527	0.83333	0.88121	0.83518	0.94278	0.86440	0.82836	0.83133	0.94389
N_CHANGEPOINTS	15	10	32	15	10	16	32	10	11	12
DAILY_FOURIER	28	35	38	25	34	41	37	35	41	44
DAILY_PRIOR_SCALE	0.88323	2.78007	2.01180	3.34977	3.38781	1.58056	0.81290	2.73977	0.24341	1.44739
DAILY_MODE	additive	additive	additive	multiplicative	additive	additive	additive	additive	multiplicative	multiplicative
WEEKLY_FOURIER	14	11	10	17	11	19	11	11	10	20
WEEKLY_PRIOR_SCALE	0.73642	0.36524	0.74033	0.96718	0.56985	0.11566	0.74148	1.81587	0.43966	0.38092
WEEKLY_MODE	multiplicative	multiplicative	multiplicative	multiplicative	multiplicative	additive	multiplicative	multiplicative	additive	additive
YEARLY_FOURIER	13	13	13	15	14	13	14	13	8	12
YEARLY_PRIOR_SCALE	3.28514	0.74975	0.35247	0.27206	0.63496	0.19736	0.99351	1.16029	0.27557	0.35994
YEARLY_MODE	multiplicative	multiplicative	multiplicative	additive	multiplicative	additive	additive	multiplicative	multiplicative	additive
HOLIDAYS_PRIOR_SCALE	0.05672	1.32205	0.15704	0.14420	0.63611	0.06903	0.25390	0.73870	3.25403	2.01715
REGRESSOR_PRIOR_SCALE	2.56482	0.81023	0.62358	0.59900	1.97218	0.50644	0.94568	1.61897	0.54818	2.25969
REGRESSOR_MODE	additive	additive	additive	multiplicative	additive	multiplicative	additive	additive	multiplicative	additive
CV_SCORE	8.20358	8.25424	8.26866	8.28955	8.29086	8.34891	8.39289	8.40199	8.41633	8.42272
CV_MAE	5.63209	5.64728	5.69242	5.73153	5.69566	5.77225	5.81559	5.79208	5.76814	5.83030
CV_RMSE	8.07956	7.94638	8.48594	8.01684	7.91527	7.74867	8.48466	8.09537	8.28593	8.09462
CV_WAPE_%	7.94948	7.97093	8.03464	8.08984	8.03921	8.13732	8.20849	8.17530	8.14151	8.22925
CV_COVERAGE	0.94444	0.95139	0.92361	0.93750	0.95486	0.96528	0.92708	0.94792	0.94444	0.95486
CV_SHARPNESS	35.09299	36.04086	34.45852	34.97671	35.85409	35.77480	34.46843	35.81580	36.21430	35.43968
CV_PINBALL	1.02182	1.03495	1.06284	1.04573	1.03721	0.96528	1.04711	1.05299	1.05766	1.01580

Tabela com Top10 de combinações de Hiperparâmetros ordenadas por score