

**Faculdade de Engenharia da Universidade do Porto**



**FEUP**

**Sistemas de deteção de eventos anómalos no  
padrão do consumo**

Daniel Fernandes Fonte

VERSÃO DE TRABALHO

Dissertação realizada no âmbito do  
Mestrado Integrado em Engenharia Eletrotécnica e de Computadores  
Major Energia

Orientador: Prof. Dr. Cláudio Domingos Martins Monteiro

Junho 2012



# Resumo

Atualmente surgiram novos mercados relacionados com a eficiência energética, pois a nível industrial o custo do produto final é influenciado pelo custo da energia.

Deste modo, as decisões nas empresas são tomadas tendo em conta os gastos em energia, uma vez que este recurso tem sofrido um aumento de preços, e torna-se fundamental para o desenvolvimento económico e social das sociedades.

Desta forma, nos últimos anos surgiu um crescente interesse na monitorização do consumo energético, dando origem a uma nova gama de produtos de inteligência de monitorização. Estes sistemas permitem detetar alterações no comportamento do consumo ou desvios relativamente a padrões típicos, e são utilizados como ferramenta de alerta para gestores de energia, para deteção de consumos ativos e quantificação de poupanças. Neste contexto é possível assegurar uma maior racionalidade na utilização da energia, reduzindo-se custos e, conseqüentemente, trazendo vantagens ao nível de competitividade das empresas no mercado atual.

Neste projeto desenvolveu-se uma ferramenta de gestão de energia que analisa o consumo energético da instalação. Esta ferramenta permite em tempo real detetar anomalias de consumo, bem como a comparação do consumo energético com os padrões históricos do consumo.

Este tipo de ferramenta permitirá uma análise e melhoramento fundamentado dos consumos energéticos da empresa, permitindo desta forma reduzir despesas e melhorar a sua competitividade, como referido anteriormente.

**Palavras-chave:** Eficiência energética; Padronização do consumo; Monitorização do consumo; Deteção de eventos anómalos no padrão do consumo das indústrias;



# Abstract

*Nowadays arise markets related to energy efficiency. On an industrial level the final costs of the product are influenced by the costs of energy.*

*Therefore, the companies' decisions are made taking into account the amount of energy wasted, since the value of this resource has been increasing, and becomes critical to both economic and social development of societies.*

*As a result, over the last years, a growing interest emerged from the monitoring of energy consumption, giving rise to a new range of intelligent monitoring products. These systems are able to detect changes in the consumption behavior or deviations of standard consumption patterns, and are used as an alert tool to energy managers, for the detection of active consumptions and savings quantifications.*

*Accordingly, it is possible to ensure a more rational way to use the energy, by reducing costs and, consequently, bringing advantages in the actual competitive company's market.*

*In this project, it has been developed an energy management tool that analyzes the energy consumption of the industry. This tool allows it's user to detect, in real time, consumption anomalies, as well as comparing between the energetic consumption and historic consumption patterns. This type of tool will also allow the analysis and grounded improvement of the energetic consumptions of the company, in order to reduce expenditures and improve it's competitiveness, as told previously.*



# Agradecimentos

Serve o presente momento para agradecer às pessoas e instituições que de algum modo contribuíram para a boa elaboração desta dissertação.

Ao meu orientador, Professor Doutor **Cláudio Monteiro**, meu muito obrigado por todo o apoio e orientação, assim como todo o conhecimento que me transmitiu não só na orientação desta dissertação mas também ao longo do curso. Contra todas as adversidades impostas e pelas inúmeras tarefas que é responsável, encontrou sempre disponibilidade para me orientar e auxiliar na execução desta dissertação. Por ser um exemplo de docente e engenheiro que me marcou durante todo o meu percurso académico.

Agradeço à empresa **SmartWatt** pela disponibilização de todos os dados, sem os quais não seria possível a realização desta dissertação.

Aos meus **pais** por todos os sacrifícios que fizeram, por tudo o que me ensinaram, pela força que ainda hoje me transmitem, algo que indubitavelmente me ajudou a atingir os meus objetivos. É graças a eles que neste momento estou pronto para realizar um dos meus maiores sonhos, e é com toda esta força e motivação que me ajudam a alcançar todos os meus objetivos na vida.

Não podendo esquecer todos os meus **alunos**, que me ajudaram neste percurso com todo o seu apoio e compreensão. Foi graças a toda esta compreensão que consegui ultrapassar alguns dos momentos mais complicados nesta faculdade.

Um obrigado muito especial á **Rita** por toda a ajuda e compreensão disponibilizada. Por todo o apoio dado nos momentos mais complicados, “...onde a luz ao fundo do túnel era tão pequenina...”. Por ser a pessoa que é, e ter o fantástico “poder” de na hora e momento certo ter capacidade de transmitir a força necessário para continuar em frente.

A todos os colegas de curso, em especial ao **Bruno Pires, Luís Oliveira, Justino, Hélder Costa, Pedro Carvalho, Nuno Moreira, Fábio Freitas, Magalhães, Andry, Filipe Bouça**, por

terem sido verdadeiros companheiros neste percurso e por me terem proporcionado momentos tão especiais.

# Índice

Resumo .....	iii
Abstract.....	v
Agradecimentos .....	vii
Índice.....	ix
Lista de figuras .....	xi
Lista de tabelas .....	xiii
Abreviaturas e Símbolos .....	xv
<b>Capítulo 1 .....</b>	<b>17</b>
Introdução.....	17
1.1 - Enquadramento .....	17
1.2 - Motivação .....	19
1.3 - Objetivo .....	19
1.4 - Informação Utilizada .....	20
1.5 - Estrutura .....	20
<b>Capítulo 2 .....</b>	<b>21</b>
Estado da arte .....	21
2.1 - Eficiência energética.....	21
2.2 - Gestão de energia na indústria.....	22
2.3 - Indicadores de Desempenho .....	22
2.4 - Monitorização e controlo .....	24
2.5 - Padronização da produção .....	24
2.6 - Necessidades da indústria .....	25
2.6.1 - Fontes de energia para as indústrias .....	25
2.6.2 - Processos de produção eficientes.....	25
2.6.3 - Utilização da energia dissipada .....	26
2.6.4 - Gestão do consumo energético.....	26
2.7 - Necessidades das indústrias na área de gestão e controlo do consumo energético .....	27
2.8 - Softwares de gestão de consumos .....	28
2.8.1 - <i>Energy Brain®</i> - software de contagem de energia .....	28
2.8.2 - <i>Power Studio - Energy monitoring and control software - Circutor SA</i> .....	31
<b>Capítulo 3 .....</b>	<b>35</b>

Metodologia.....	35
3.1 - Introdução .....	35
3.1.1 - Abordagem .....	36
3.1.2 - Objetivos .....	37
3.2 - Modelo de detecção de anomalias .....	37
3.2.1 - Recolha do histórico.....	37
3.2.2 - Análise visual e interpretativa da série do consumo .....	38
3.2.3 - Caracterização qualitativa de anomalias.....	42
3.2.4 - Pré-processamento da série do consumo.....	43
3.2.5 - Metodologia de reconhecimento de padrões .....	43
3.2.6 - Criação da base dados do modelo .....	47
3.2.7 - Cálculo das pertenças.....	49
3.2.8 - Indicadores de anomalias .....	51
3.2.8.1 - <i>Indicador 1</i> .....	51
3.2.8.2 - <i>Indicador 2</i> .....	54
<b>Capítulo 4 .....</b>	<b>61</b>
Simulação para as séries estudadas .....	61
4.1 - Introdução .....	61
4.1.1 - Abordagem .....	62
4.2 - Dados .....	63
4.3 - Determinação das Pertenças .....	63
4.4 - Detecção de Anomalias .....	64
4.5 - Mensagens de anomalias .....	65
4.6 - Conclusões .....	68
<b>Capítulo 5 .....</b>	<b>69</b>
Validação e Conclusões .....	69
5.1 - Introdução .....	69
5.2 - Indicador de Validação .....	70
5.3 - Validação.....	71
5.4 - Conclusões e trabalhos futuros.....	71
<b>Referências .....</b>	<b>75</b>

## Lista de figuras

Figura 1.1 - Partilha dos proveitos entre o cliente e a ESCO [3] .....	18
Figura 2.1 - Contribuição da eficiência energética para os três principais aspetos da produção sustentável [5].....	22
Figura 2.2- <i>Energy Brain</i> ® [21] .....	28
Figura 2.3 - Sistema <i>energy brain</i> ® (exemplo 1) [21] .....	29
Figura 2.4 - Sistema <i>energy brain</i> ® (exemplo 2) [21] .....	30
Figura 2.5 - <i>Circuitor - Power Studio</i> [24].....	32
Figura 2.6 - Sistema <i>Power Studio - Circuitor</i> [25] .....	33
Figura 3.1 - Fluxograma da abordagem seguida no desenvolvimento do modelo .....	36
Figura 3.2 - Série do consumo (Savinor) .....	37
Figura 3.3 - Consumo ao Domingo .....	38
Figura 3.4 - Consumo às Segundas .....	39
Figura 3.5 - Consumo às Terças-feiras .....	39
Figura 3.6 - Consumo às Quartas-feiras.....	40
Figura 3.7 - Consumo às Quintas-feiras .....	40
Figura 3.8 - Consumo às Sexta-feiras.....	41
Figura 3.9 - Consumo aos Sábados .....	41
Figura 3.10 - Caracterização qualitativa de anomalias .....	42
Figura 3.11 - Fases do processo de reconhecimento de padrões.....	45
Figura 3.12 - Representação gráfica dos cinco <i>clusters</i> .....	47
Figura 3.13 - Diagrama do processo de atribuição dos índices aos <i>clusters</i> 1 e 3.....	49
Figura 3.14 - Fluxograma com o algoritmo da determinação do <i>cluster</i> de maior pertença ...	50

Figura 3.15 - Indicadores utilizados .....	51
Figura 3.16 - Fluxograma do algoritmo de detecção de eventos anómalos .....	52
Figura 3.17 - Fluxograma do algoritmo do indicador 1 .....	52
Figura 3.18 - Processo de descodificação do <i>cluster</i> de maior pertença .....	53
Figura 3.19 - Fluxograma da comparação entre os dias da semana .....	53
Figura 3.20 - Fluxograma do processo de avaliação do indicador 1 .....	54
Figura 3.21 - Fluxograma do algoritmo do indicador 2 .....	55
Figura 3.22 - Curva da distribuição normal .....	58
Figura 3.23 - Distribuição acumulada .....	58
Figura 3.24 - Distribuição normal .....	59
Figura 3.25 - Modelo para a mudança de escala do indicador 2 .....	59
Figura 3.26 - Fluxograma com os limites do algoritmo de avaliação do indicador 2 .....	60
Figura 4.1 - Fluxograma da abordagem seguida na simulação do modelo .....	62
Figura 4.2 - Simulação do indicador 1 .....	64
Figura 4.3 - Simulação do Indicador 2 .....	65
Figura 4.4 - Mensagens referentes ao indicador 1 .....	66
Figura 4.5 - Mensagens referentes ao indicador 2 .....	67
Figura 4.6 - Simulação 06/05/2011 (sexta-feira) .....	68
Figura 5.1 - Indicador_de_Validação .....	70

# Lista de tabelas

Tabela 2.1 - Exemplos de indicadores de eficiência energética [5] .....	24
Tabela 2.2 - Visão geral dos temas de investigação na área de "Gestão e Controlo do Consumo de Energia" [5] .....	26
Tabela 3.1- Parte da matriz DATA .....	43
Tabela 3.2 - Algoritmos para o reconhecimento de padrões [26] .....	44
Tabela 3.3 - Centros dos clusters .....	48
Tabela 3.4 - Relação dos <i>clusters</i> com os dias da semana .....	48
Tabela 3.5 - Demonstração do indicador 1 .....	53
Tabela 3.6 - Médias e desvios padrões dos <i>clusters</i> .....	57
Tabela 3.7 - Demonstração do indicador 2 .....	60
Tabela 4.1 - Cluster de maior pertença .....	63
Tabela 5.1- Validação .....	71



# Abreviaturas e Símbolos

Lista de abreviaturas (ordenadas por ordem alfabética)

CEN	Comité Europeu de Normalização
DEEC	Departamento de Engenharia Electrotécnica e de Computadores
EPC	Energy Performance Contracts
ESCO	<i>Energy Service Company</i>
ESE	Empresa de Serviços Snergéticos
FCM	<i>fuzzy c-means</i>
FEUP	Faculdade de Engenharia da Universidade do Porto
KPIs	<i>Key Performance Indicators</i>
MATLAB	<i>MATrix LABoratory</i> , software interativo de alta performance voltado para o cálculo numérico
MIEEC	Mestrado em Engenharia Eletrotécnica e de Computadores
SAVINOR	Sociedade Avícola do Norte
TIC	Tecnologias de informação e comunicação

Lista de símbolos

$\Delta$	Delta, é a quarta letra do alfabeto grego e significa “variação”
$\sigma$	Desvio Padrão
$\alpha$	Ângulo



# Capítulo 1

## Introdução

Esta dissertação foi realizada no âmbito do Mestrado Integrado em Engenharia Eletrotécnica e de Computadores, na Faculdade de Engenharia da Universidade do Porto (FEUP).

Neste trabalho de dissertação foi realizado um modelo que deteta eventos anómalos no padrão do consumo de uma determinada instalação. Para a elaboração do projeto foram determinados diferentes momentos:

- Recolha do histórico da série de consumo na instalação real em monitorização. Análise visual e interpretativa das séries de consumos, identificando por observação eventos anómalos.
- Identificação de indicadores para a caracterização dos eventos anómalos.
- Criação dos modelos de parametrização e treino. Criação dos modelos operacionais de deteção de anomalias.
- Simulação para as séries estudadas. Integração do modelo funcional na plataforma de serviços de monitorização. Aplicação do modelo funcional, operando em tempo real.

Este capítulo apresenta uma visão da temática e define quais os objetivos e motivação inerentes à sua realização. No final do capítulo apresenta-se uma breve explicação da organização estrutural da dissertação.

### 1.1 - Enquadramento

A melhoria da eficiência energética é uma forma básica, mas significativa, de abordar a poupança energética e as preocupações ambientais.

A nível industrial existem diversas medidas para melhorar o desempenho energético, com diferentes finalidades e aplicações. A crescente preocupação com o consumo energético trouxe a eficiência energética para os assuntos de primeira ordem, uma vez que este contribui diretamente para a redução do consumo de energia. Devido à complexidade dos

processos industriais projetar índices de eficiência robustos e flexíveis é extremamente difícil.

Assim a chave para a gestão dos consumos nas indústrias passa pela criação de modelos que detetem em tempo real pequenos desvios do consumo, de modo a diminuir o tempo de atuação na eliminação dos mesmos [1].

Esta gestão é possível através da implementação de tecnologias que permitem a monitorização da instalação e recolha de dados para ser efetuada a sua análise. Desta análise são obtidos perfis de consumo para a instalação, sendo utilizados para monitorizar os níveis de consumo em relação ao previsto, e fornecer dados relativos ao imediato. Com o acesso a estes perfis, é possível detetar anomalias e promover um consumo mais eficiente através da redução de desperdícios [1]. Fundamentalmente, a introdução de inteligência e tecnologias de informação na gestão de consumos, é um fator de peso o crescimento da eficiência energética [2].

A introdução destas tecnologias leva ao aparecimento de um mercado de eficiência energética em que é necessário medir, analisar e ajustar os consumos de forma a verificar poupanças energéticas. As empresas que operam neste mercado são as ESCO, *Energy Service Companys*, que estabelecem contratos de performance, EPC - *Energy Performance Contracts*, com os seus clientes, com a premissa de realizarem uma gestão das suas instalações tendo em vista o aumento da eficiência energética [3].

## Distribuição dos proveitos

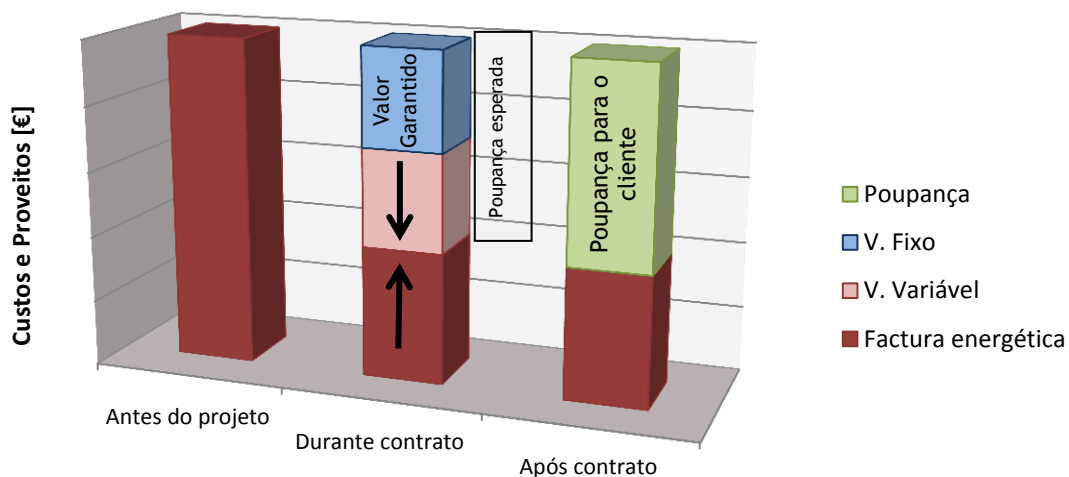


Figura 1.1 - Partilha dos proveitos entre o cliente e a ESCO [3]

## 1.2 - Motivação

Analisando as ferramentas existentes a nível industrial com o objetivo de melhorar o desempenho energético, nota-se a existência de diversas lacunas. Algumas das ferramentas apenas monitorizam o consumo, não tendo qualquer alerta direcionado ao consumidor, enquanto outras não operam em tempo real. Esta dissertação tem como motivação contribuir com uma ferramenta que opere em tempo real, analisando os desvios do consumo energético identificando anomalias, preenchendo deste modo as lacunas nos programas existentes.

O modelo deve ter em conta a variabilidade existente no universo industrial, apresentando uma metodologia robusta e de fácil implementação, de forma a operar sobre os variados perfis do consumo. Assim a ferramenta desenvolvida deverá ser capaz de construir modelos automaticamente e de ser aplicado a situações futuras. A proposta de participação deste tipo de modelos em sistemas de gestão de energia é mais um ponto inovador, que permite que as indústrias aumentem os seus proveitos e que a energia tenha uma maior racionalização. Produzir ferramentas através desta nova metodologia permite aproveitar ao máximo os recursos energéticos e ainda facultar um conjunto de indicadores decisivos para a gestão ativa dos consumos.

Em conclusão, esta metodologia permite reduzir o consumo alertando o consumidor para consumos anómalos e permite avaliar o desempenho energético da instalação.

## 1.3 - Objetivo

Pretende-se nesta dissertação desenvolver um sistema de processamento de medições de consumo energético para a deteção de eventos anómalos.

Desta forma, foi pensada uma ferramenta com os seguintes objetivos:

- **Operar em tempo real:**

Esta ferramenta fará uma análise em tempo real da informação de monitorização de consumos, sendo de fácil parametrização.

- **Reconhecer padrões de consumo:**

O algoritmo desenvolvido deverá analisar séries de consumo histórico de modo a identificar diferentes perfis do consumo energético.

- **Detetar eventos anómalos no consumo:**

O modelo terá associado indicadores que permitem identificar desvios do consumo. Os algoritmos desenvolvidos para cada indicador deverão ser robustos e de fácil parametrização.

- **O modelo deverá ser testado em casos reais**

Deste modo o modelo criado será aplicado a instalações reais já em monitorização de modo a ser estudada a fiabilidade do modelo desenvolvido.

- **Possibilidade de integração em sistemas de gestão de energia:**

Numa fase final esta ferramenta deverá ser otimizada de modo a ser possível a sua integração num sistema de gestão de energia.

## 1.4 - Informação Utilizada

Para a elaboração desta dissertação foi necessário informação relativa ao consumo nos diversos sectores da empresa em estudo e posteriormente deslocações ao local de modo a identificar alguns dos processos envolventes da empresa. Toda a informação utilizada foi disponibilizada pela *Smartwatt*.

Os dados fornecidos são relativos a:

- Séries de consumos de todos os sectores da empresa SAVINOR;

O fornecimento dos dados foi fundamental para a elaboração desta dissertação.

## 1.5 - Estrutura

Esta dissertação apresenta cinco capítulos, sendo o presente capítulo o introdutório. Neste capítulo encontra-se a motivação e os objetivos que levaram à elaboração da dissertação.

O capítulo dois apresenta o estado da arte referente ao consumo energético nas indústrias, onde é feita uma abordagem às questões da eficiência energética. Do mesmo modo serão analisados modelos já desenvolvidos nesta área da gestão ativa dos consumos.

No terceiro capítulo será descrita a metodologia de reconhecimento de padrões desenvolvida, sendo apresentadas todas as características e funcionalidades do modelo criado.

Relativamente ao quarto capítulo são apresentados todos os resultados obtidos, provenientes da simulação do modelo com séries reais.

No quinto capítulo será apresentada a validação do modelo, bem como todas as conclusões e trabalhos futuros.

# Capítulo 2

## Estado da arte

No presente capítulo pretende-se abordar assuntos relacionados com o tema desta dissertação. Assim, começa-se por tratar alguns dos conceitos atualmente abordados que se encontram fortemente ligados à gestão ativa de consumos. Também será feito um breve resumo dos métodos utilizados no desenvolvimento do modelo.

### 2.1 - Eficiência energética

Os preços crescentes da energia, as alterações climáticas e a energia não fornecida são temas de elevada importância na atualidade. Embora os recursos renováveis, tais como tecnologias fotovoltaicas, eólicas ou hídricas possam ser soluções a longo prazo, o uso mais eficiente da energia pode ser a solução a curto prazo mais rentável [4].

Sobre o ponto de vista das empresas existem três fatores que proporcionam melhorias na eficiência energética, sendo eles fatores económicos, sociais e ambientais [5].

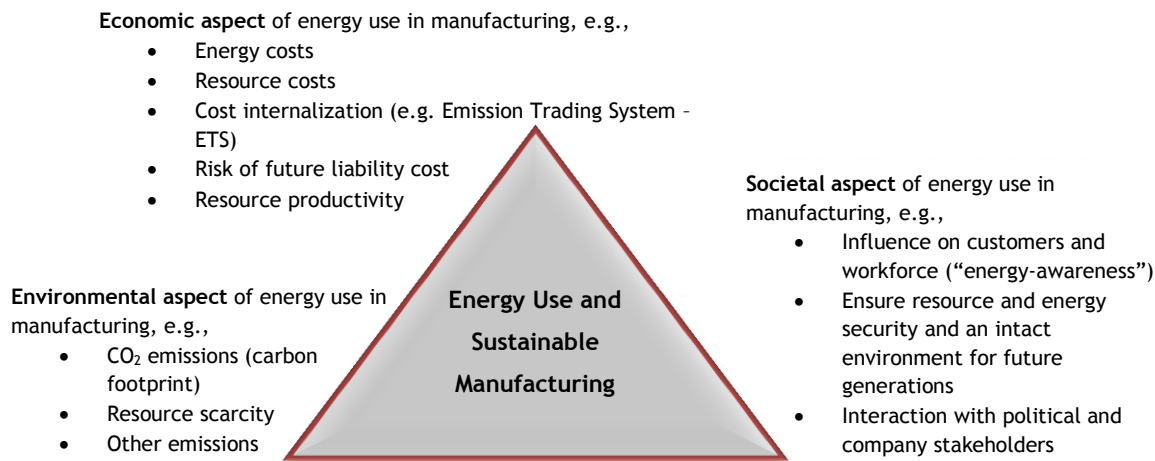
Dentro dos fatores económicos destaca-se o crescimento exponencial do preço da energia. Este deve-se, entre outros, à inflação do preço do petróleo, do gás natural e de outros combustíveis fósseis como o carvão, tornando-se cada vez mais importante a redução do consumo de energia a nível industrial. Neste âmbito, a energia pode representar mais de 60% dos custos de operação, assumindo assim um forte fator de competitividade [6].

Relativamente aos fatores ambientais, foram criados, por exemplo, o protocolo de Kyoto em 1997 [7] e o acordo de Copenhaga em 2009 [8] que visam regulamentar as emissões de CO<sub>2</sub>. Uma vez que o sector industrial é responsável por cerca de 1/3 das emissões globais de CO<sub>2</sub>, foram criadas diferentes iniciativas políticas e regras orientadas pelo mercado como taxas, subsídios, limites de emissões aceitáveis e certificados “verdes”[9]. Empresas que melhorem a sua eficiência energética e, conseqüentemente, a sua emissão de CO<sub>2</sub> podem melhorar a sua posição para encarar novos desafios e os custos resultantes das regulamentações de CO<sub>2</sub> atuais e no futuro[5].

A nível social, torna-se cada vez mais evidentes as alterações dos comportamentos de compra dos consumidores também em função do impacto ambiental da produção, favorecendo a compra de “produtos verdes”. Este tipo de produtos provêm de indústrias que

promovem a eficiência energética nas fases de produção, tornando-se uma vantagem de reputação e competitividade [5].

Os aspetos mencionados acima tornam a eficiência energética num forte pilar para uma estrutura de produção sustentável (fig 2.1)



**Figura 2.1 - Contribuição da eficiência energética para os três principais aspetos da produção sustentável [5]**

## 2.2 - Gestão de energia na indústria

Embora o sector de produção esteja já a fazer melhorias contínuas a nível da eficiência energética, do ponto de vista económico o seu potencial não está ainda totalmente explorado [10]. As principais barreiras para a implementação de medidas de melhoramento da eficiência energética nas indústrias são as decisões baseadas nos períodos de retorno em vez do cálculo da taxa de interesse, o capital limitado e a baixa prioridade dada à eficiência energética pela gestão e a falta de informação [5]. Alguns estudos identificam ainda um baixo nível de gestão de energia como uma barreira para a eficiência energética [11]. A gestão energética pode ser um benefício para as indústrias por razões económicas, ambientais e sociais. A implementação de uma gestão de energia pode ser uma forma para a redução do consumo de energia e emissões de CO<sub>2</sub> associadas. Essencialmente esta envolve a monitorização, medição, registo, análise, exame crítico, controlo e redireccionamento da energia e fluxos de material através de sistemas para que menos energia seja gasta para alcançar os melhores objetivos [5].

## 2.3 - Indicadores de Desempenho

A medição do consumo de energia de um processo permite avaliar os potenciais de otimização e suporta a visualização de benefícios para as medidas de melhoramento. O

desenvolvimento e aplicação de indicadores de eficiência energética dependem do objetivo para os quais são aplicados. Usualmente os indicadores são ratios que descrevem a relação entre uma atividade e a energia requerida. No sector industrial, as atividades com os processos de produção de um produto podem ser descritas em termos económicos ou físicos resultando em indicadores económicos ou físicos [12]. Os indicadores económicos são uteis num nível agregado, como para comparar diferentes sectores; contudo, para obter uma visão de processos de produção particulares, os indicadores físicos são mais adequados [13].

Devido à variedade de processos industriais e à sua complexidade, existe uma grande variedade de indicadores estruturais e explanatórios, isto é, indicadores designados para um sector de produção específico. Cuidados especiais têm que ser tidos em conta quando comparam indicadores de eficiência energética internacionalmente e também quando comparam diferentes industrias, sectores ou companhias. Os resultados da comparação podem variar fortemente dependendo das diferentes medidas de consumo de energia, e medidas de atividade de produtos heterogêneos[5].

Não existe nenhum indicador de eficiência energética único que possa ser aplicado em qualquer situação, estes devem ser definidos de acordo com as decisões que têm de tomar. A tabela 2.1 apresenta uma visão geral de diferentes indicadores, das suas aplicações e fórmulas/unidades.

Referência	Indicador	Tipo de indicador	Aplicação	Fórmula / Unidade
(Phylipsen et al., 1997) [13]	Energy Intensity	Economic	Aggregated level	Energy consumption/economic term
	Specific energy consumption	Physical	Disaggregated level	GJ per t
(Irrek e Thomas, 2006) [4]	Energy intensity	Macro-economic	Aggregated level	Energy consumption/monetary variables (GDP)
	Degree of efficiency	Engineering view	Aggregated level	Net energy/used primary energy
	Final energy efficiency improvement	Physical	National level	Energy savings per year
(Patterson, 1996) [14]	Thermodynamic energy efficiency	Thermodynamic	Measurements derived from the science of thermodynamics	Actual energy usage related to an 'ideal' process
	Physical-thermodynamic energy efficiency	Hybrid	Measure the service or delivery of the process	Actual energy usages/tonnes or passenger miles
	Economic-thermodynamic energy efficiency	Hybrid	Measure the service or delivery of the process	Energy usage in conventional thermodynamic units/output in terms of market price
	Economic energy efficiency	Economic	Measure in terms of market value	Energy input in monetary terms/output in monetary terms
(Farla et al.,	Energy efficiency	Economic	Activity of a sector	Energy consumption/value added

Referência	Indicador	Tipo de indicador	Aplicação	Fórmula / Unidade
<a href="#">1997</a> ) [15]	measurement			or value of shipments
	Specific energy consumption	Physical	Process level, cross country comparison	Energy use/physical unit of production
( <a href="#">Boyd et al., 2008</a> ) [16]	Energy performance indicator	Statistical	On plant level	Percentile ranking of the energy efficiency

**Tabela 2.1 - Exemplos de indicadores de eficiência energética [5]**

## 2.4 - Monitorização e controlo

O acompanhamento e análise do consumo de energia das máquinas e processos de fabrico são os primeiros passos para aumentar a eficiência energética. Uma monitorização insuficiente pode resultar em empresas que não têm noção do seu potencial a nível de investimentos em energia rentável. Logo a monitorização da eficiência energética, e auditorias são importantes para a gestão de energia, uma vez que permitem que os gestores identifiquem oportunidades de melhorar o consumo energético e do mesmo modo acompanhar os efeitos das suas decisões [17].

Contudo os programas das auditorias não levam automaticamente a altas taxas de aprovação, isto é as taxas de adoção de medidas de eficiência energética foram encontradas em cerca de 40-50% incluindo medidas na área de aquecimentos, ventilação, iluminação, motores, compressores de ar, vapor ou de recuperação de calor [18] [19]. As principais barreiras para investimentos em eficiência energética passam pelos limitados recursos financeiros e falta de informação.

A monitorização da energia para as máquinas é uma porção importante na monitorização global da indústria e pode ajudar a entender, caracterizar e reduzir a energia utilizada no processo de fabrico. Contudo, novas tecnologias de sensores inteligentes e pequenos dispositivos que integram os processos vão possibilitar medições aos gestores de energia quase em tempo real. Assim, estas novas tecnologias permitem fornecer informações precisas de monitorização proporcionando uma base sustentável para as decisões na gestão da energia na produção [5].

## 2.5 - Padronização da produção

As normas de gestão existentes visam aumentar o desempenho ambiental e a eficiência energética das empresas, melhorando os processos de organização existentes. Há diversos anos que existem em diferentes países padrões nacionais para a gestão da energia, por exemplo na Irlanda, Coreia, EUA, Espanha, Tailândia, Dinamarca, Suécia. Serve como exemplo, a norma europeia EN 16001 para sistemas de gestão de energia lançada em 2009

pelo Comité Europeu de Normalização (CEN) [5]. A norma fornece um quadro para a economia da energia e atividades orientadas para ações práticas no que diz respeito à gestão da energia [20].

Deste modo a padronização aumenta a transparência dos processos das empresas, assim torna-se a avaliação por parte externa mais facilitada. Em geral as normas de gestão de energia podem suportar significativamente a gestão energética dos processos de fabrico. No entanto este tipo de normas pode contribuir para a criação de mecanismos de medição, controlo e de melhoria contínua da eficiência energética.

## **2.6 - Necessidades da indústria**

O tema eficiência energética está cada vez mais presente no mundo industrial, deste modo existe uma crescente necessidade em estudos que visam criar respostas para as mais diversas questões existentes na indústria.

### **2.6.1 - Fontes de energia para as indústrias**

Esta área de pesquisa inclui tópicos de investigação em tecnologias e estruturas de produção que permitam a Auto-dependência na geração de eletricidade e de diferentes conceitos para a captação de energia. O uso dos potenciais da energia, em menor escala, para alimentar dispositivos como sensores e controladores pode resultar em consideráveis economias de energia. Por exemplo, o controlo pode criar melhorias nos processos de fabrico [5].

### **2.6.2 - Processos de produção eficientes**

Reduzir o consumo de energia utilizada nos processos de fabrico e aumentar a produção significa o aumento da produtividade e eficiência. Novas abordagens tecnológicas em equipamentos de produção podem reduzir o consumo de energia:

- Pesquisa sobre os tópicos relacionados na redução de emissões de CO<sub>2</sub>;
- Formas alternativas da redução do tamanho da partícula;
- Otimizar a produção de veículos no futuro (devido ao sector automóvel ser um dos mais importantes da Europa, e possuir uma elevada participação no volume de negócios global e um grande numero de postos de trabalhos criados) [5];

### 2.6.3 - Utilização da energia dissipada

A dissipação da energia sob a forma de calor ou subprodutos é geralmente entendida como saída de resíduos. Muitas vezes é possível a reutilização economicamente viável dessa energia dissipada em outros processos de produção. Empresas e indústrias precisam de ter a capacidade de colaborar numa base intersectorial com o intuito de usar fluxos de energia e resíduos de uma forma simbiótica. Isto é, as empresas e indústria devem investir numa relação mutualmente vantajosa conhecida no sector por “simbiose industrial”.

Como exemplo a utilização do calor residual que é disponibilizado em diversos sectores pode apresentar elevadas vantagens na poupança energética [5].

### 2.6.4 - Gestão do consumo energético

Sistemas de medição e controlo são partes integrantes dos sistemas de fabricação. Novos conceitos de gestão de energia formam uma base para as decisões de melhoria da eficiência energética. Para desenvolver novos conceitos de gestão de energia, a atenção deve ser dada aos sensores e dispositivos de controlo, os KPIs, e às interfaces técnico-humano. A eficiência energética também deve ser representada nas Tecnologias de informação e comunicação (TIC) dos sistemas de produção [5].

Devido a novas opções para uma maior colaboração, novas economias de energia podem ser realizadas nas cadeias de abastecimento (tabela 2.2).

	Nr.	Título do tópico de pesquisa	Descrição
Área de pesquisa: gestão e controle do consumo de energia	RT1	Energy-aware manufacturing processes - measurement and control	An effective energy control system has to be developed, using the information of sensors and in-process measurement and a suitable energy efficiency performance measuring system. This control system focuses on concepts that facilitate the evaluation, control, and improvement of energy efficiency in manufacturing processes.
	RT2	Integrating energy efficiency in production information systems	A novel framework that manages and optimizes energy efficiency with respect to production planning and control needs to be developed and implemented in enterprise control and information systems, such as Enterprise Resource Planning (ERP), Manufacturing Execution Systems (MES), and Distributed Control Systems (DCS).
	RT3	Product tags for holistic value chain improvement	Product related information about the inputs and outputs of manufacturing processes make the value chain transparent for its stakeholders. The transparency allows process improvements to be coordinated to increase the overall value chain performance (in terms of e.g., energy, efficiency, costs, delivery time).

**Tabela 2.2 - Visão geral dos temas de investigação na área de "Gestão e Controlo do Consumo de Energia" [5]**

## 2.7 - Necessidades das indústrias na área de gestão e controlo do consumo energético

Medir e controlar a eficiência energética é o primeiro passo para avaliar e implementar medidas de melhoria. Ao estabelecer sistemas de informação na produção, os gestores de energia terão acesso às informações relativas aos impactos no desempenho energético resultantes dos novos planeamentos da produção e possíveis decisões de negócio.

Nas indústrias existe uma necessidade de KPIs de eficiência energética para acompanhar as mudanças e melhorias nos processos energéticos globais da empresa. As principais necessidades de medição da eficiência energética incluem [5]:

### - KPIs:

- Indicadores de eficiência energética para identificarem ineficiências na utilização da energia (ex. perfis do consumo de energia);
- Facilitar o controlo de alterações e melhorias na eficiência energética;
- Medição da eficiência energética diretamente em valores monetários de modo a comunicar diretamente, onde o dinheiro pode ser salvo;

Em relação a monitorização e controlo, as indústrias têm a necessidade da integração de dados em tempo real e da análise dos processos. Métodos de avaliação, visualização e novos conceitos de gestão para promover a eficiência energética são os pontos necessários a serem desenvolvidos para a gestão futura da energia na produção das indústrias [5].

Algumas das necessidades na área de controlo e melhoria da eficiência energética incluem:

### - Monitorização:

- Maior desenvolvimentos de sistemas que utilizam dados em tempo real para a monitorização operacional e para facilitar as decisões estratégicas de eficiência energética;
- Processos mais baseados nas capacidades das máquinas, de modo a alcançar um melhor rendimento;
- Sistemas de automação devem incluir *software* para monitorizar processos e o consumo de energia;

### - Estruturas de avaliação:

- Avaliação de novos conceitos para aumentar a eficiência energética, considerando os custos envolvidos, o que permite transparência e a identificação de componentes de elevada intensidade energética;
- Novos conceitos de gestão de modo a adequar estratégias de melhoria da eficiência energética;

- Desenvolver uma ferramenta de cálculo de custos que tenha em conta os recursos energéticos gastos no processo de fabrico bem como a matéria-prima utilizada;

## 2.8 - Softwares de gestão de consumos

A energia é um recurso fundamental no desenvolvimento económico e social que tem sofrido um aumento sustentado do seu custo. Associado às crescentes preocupações ambientais, esse aumento leva, cada vez mais, à adoção de medidas de eficiência energética por parte dos promotores e proprietários dos edifícios, indústrias ou outras instalações consumidoras de energia. O aumento da eficiência energética e a identificação das oportunidades de redução de custos com a energia é efetuado através da gestão e monitorização dos consumos de energia.

O controlo do consumo de energia revela-se uma ferramenta essencial para assegurar uma maior racionalidade no seu uso, tendo grande impacto na estrutura de custos das empresas e, conseqüentemente, na sua competitividade.

A título de exemplo serão abordados os seguintes *softwares*:

- *Energy Brain*®;
- *Power Studio*®;

### 2.8.1 - *Energy Brain*® - software de contagem de energia

O *energy brain*® é um sistema para aquisição e simulação de dados de consumos.

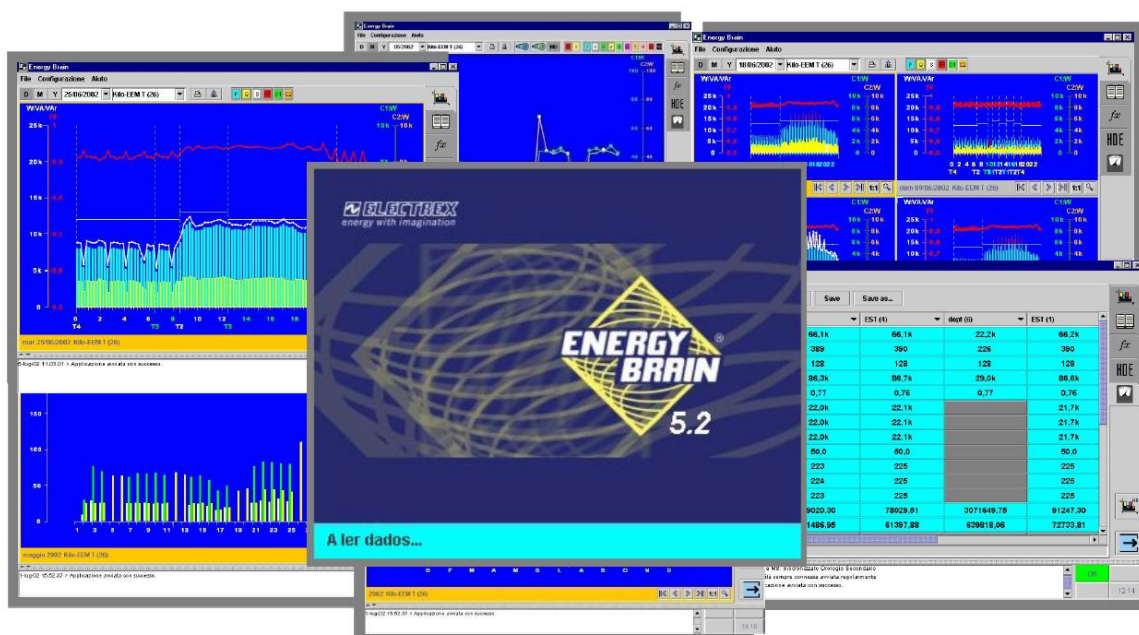


Figura 2.2- *Energy Brain*® [21]

O sistema *Energy Brain*® tem como principais objetivos [22]:

- **Análise da fatura da energia** - otimização do contrato de aquisição de energia;
- **Localização de consumos anómalos ou evitáveis** - eliminação ou minimização de encargos associados a situações anómalas que derivam de avaria, desconhecimento ou á utilização de recursos;
- **Aquisição, análise e comparação de dados** - verificação da eficiência de um setor ou equipamento por comparação com outros de funcionamento semelhante;

O *Energy Brain*® para além dos dados de energia eléctrica também tem a possibilidade de registar outros consumos, como água, gás, vapor, entre outros. Basta instalar contadores deste tipo de fluidos com saídas de impulsos [21].

O sistema *Energy Brain*® é constituído por [23]:

- Um conjunto de contadores de energia que se colocam nos pontos de consumo relevante;
- Um concentrador que assegura as comunicações com os contadores;
- Um *software* residente num PC, onde está instalada a base de dados;

Para integração com o *software* existe uma variada gama de equipamentos, desde o mais simples contador de energia com saída de impulsos ao sofisticado analisador de energia com registo de eventos [22]. Na figura 2.3 e 2.4 estão ilustrados dois exemplos da interação de diferentes equipamentos com o sistema *energy brain*®.

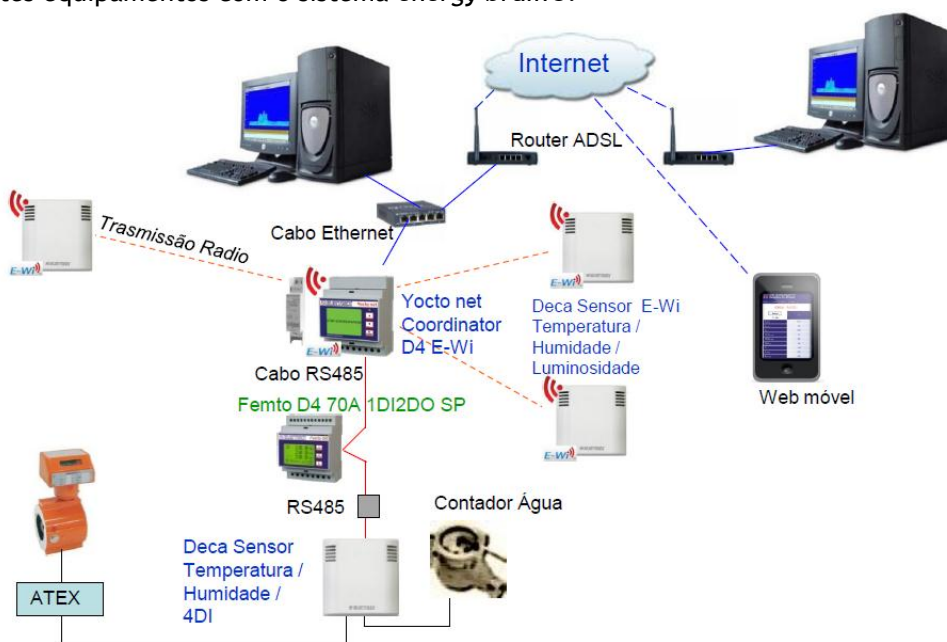


Figura 2.3 - Sistema *energy brain*® (exemplo 1) [21]

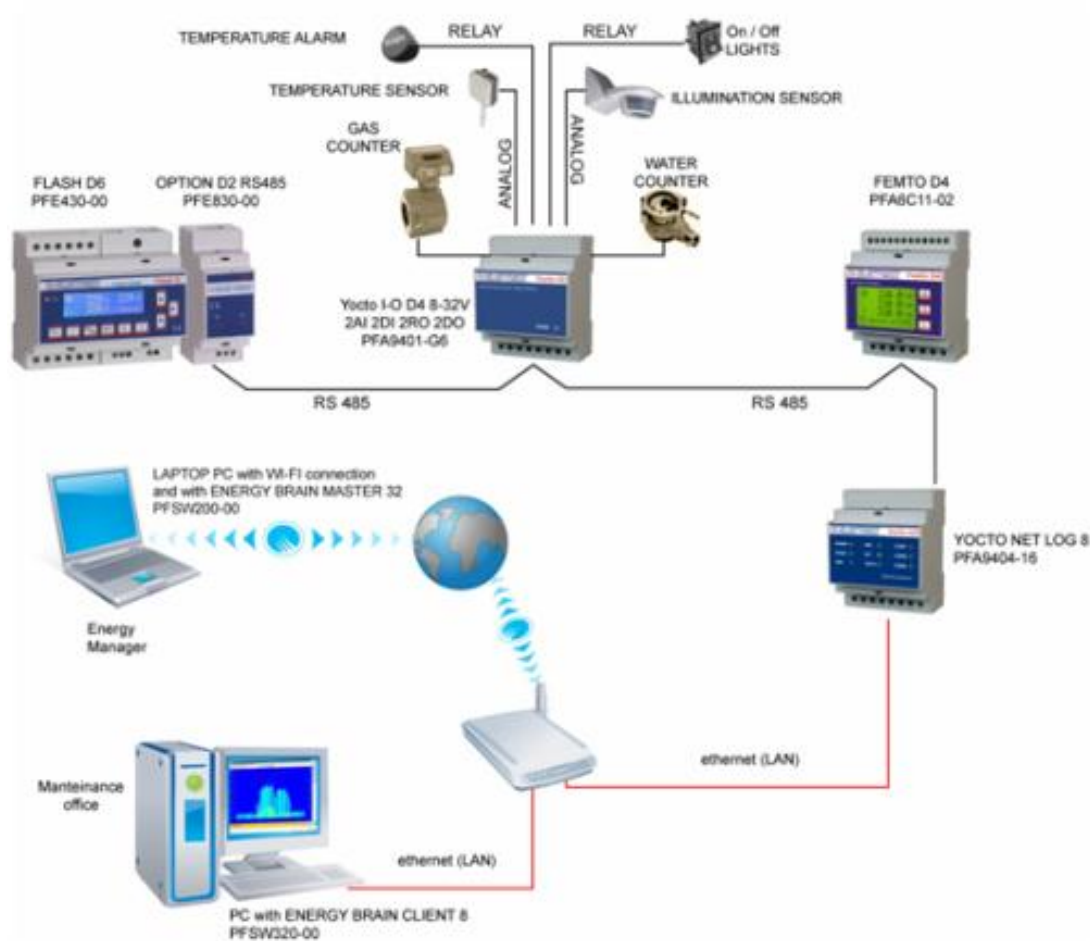


Figura 2.4 - Sistema *energy brain*® (exemplo 2) [21]

Algumas das **potencialidades** associadas a este sistema são [22]:

- Controlo remoto através da recolha de dados calendarizados;
- Monitorização da: Energia Ativa, Energia Reativa, Potência Aparente e Fator de Potência;
- Relatórios diários dos consumos divididos pelas tarifas horárias;
- Relatórios mensais dos consumos divididos pelas tarifas horárias;
- Relatórios de “picos” de potência pelas tarifas horárias;
- Gráficos de perfil de cargas com intervalos de 15min;
- Gráficos de consumos com intervalos de 1h;
- Gráficos editáveis para adição de informações adicionais como luminosidade, temperatura, etc.;
- Criação de fórmulas para análise de dados (custos com energia, divisão por centros de custo, médias, etc.);
- Editor para calendário e adequação de tarifas;

- Gravação de eventos e qualidade de energia (cavas, sobretensões e abaixamentos de tensão) para os equipamentos da família X3M;
- Configuração de equipamentos e rede;
- Criação e envio para um endereço de e-mail automaticamente de um relatório Excel com os consumos mensais e perfis de dados (gráficos com detalhes diários + sumário mensal) e com as leituras detalhadas com indicação de tempo para cada intervalo;
- Criação de um relatório com os valores de contagem da energia (kWh e kVarh) de todos os contadores de uma determinada localização, formato HTML e CSV;

O sistema pode também obter informações sobre processos de trabalho e eficiência das máquinas. O consumo de energia ativa é, de facto, um indicador do trabalho real, logo, a curva de consumo corresponderá também à curva relativa à eficiência do departamento. A energia reativa corresponderá à potência necessária para o funcionamento das máquinas [21].

A relação entre as duas potências poderá determinar a eficiência das máquinas, isto é, se a componente reativa for superior à componente ativa, os motores estão a trabalhar mas não estão a produzir. O *Energy Brain*® torna a análise dos dados de energia fácil para qualquer utilizador [21].

O *software Energy Brain*® apresenta a possibilidade em editar **outras opções**, tais como [23]:

- Páginas WEB - *standard* ou personalizadas;
- *Energy Brain Configurator* - Configuração e dados em tempo real de todos os analisadores da ELECTREX;
- Contadores *Energy Brain* - Relatórios periódicos dos contadores/analisadores;
- Opção *Energy Brain Personal Report* - Relatórios automáticos a partir de cada canal;
- Opção *Energy Brain Log Report* - Transmissão automática de uma mensagem de *e-mail* com *status* dos canais e anomalias;

## **2.8.2 -Power Studio - Energy monitoring and control software - Circutor SA**

A *Circutor SA* oferece produtos e soluções que vão desde a medição elétrica e controlo, protecção e controlo de qualidade e medição para correção do fator de potência e filtragem harmônica [24].



Figura 2.5 - Circutor - Power Studio [24]

*Power Studio* é um *software* poderoso e simples que pode ser usado para o estudo da energia a grande nível. Permite analisar os consumos energéticos e proporciona um controlo completo em diferentes sectores dos processos industriais [24].

Algumas das **potencialidades** associadas a este sistema são [25]:

- Operar em *Web server* (multiusuários);
- Visualização em tempo real de todos os parâmetros elétricos;
- Visualização de gráficos de energia;
- Exibi os dados em diferentes tabelas;
- Exibe o histórico (dia, semana, mês) com elevada resolução;
- Possibilidade de exportar dados (excel ou XML)
- Possibilidade de exportar parâmetros em tempo real através do servidor XML
- Máxima conetividade interna (intranet) e externa (internet)

Na figura 2.7 está representado uma das possíveis **configurações** de operação do sistema *Power Studio*.

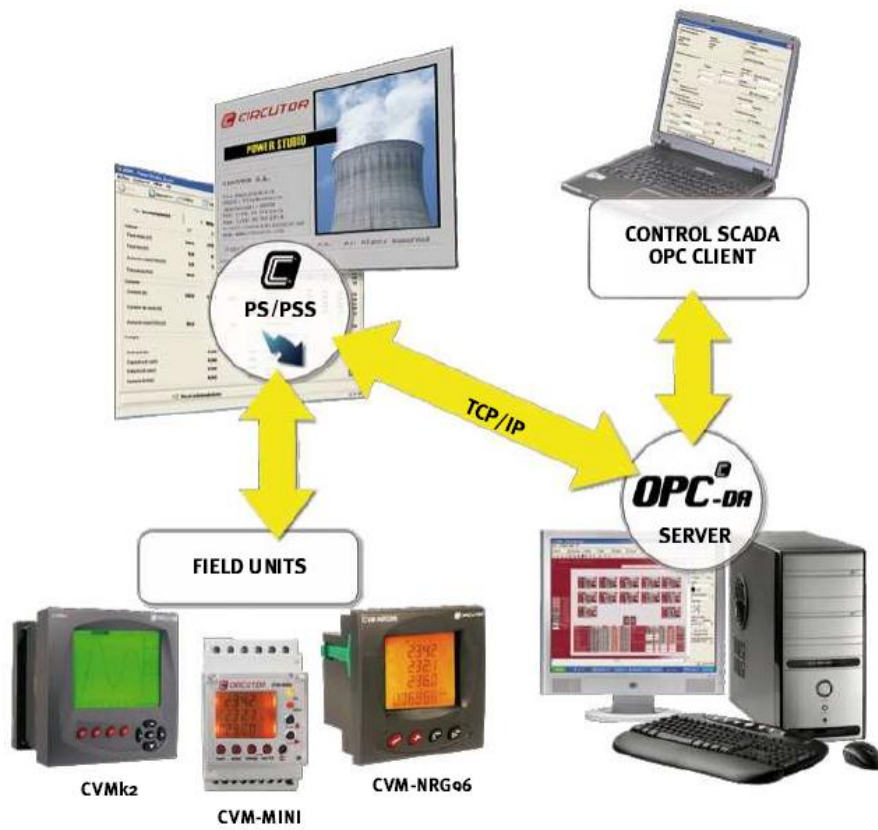


Figura 2.6 - Sistema Power Studio - Circutor [25]



# Capítulo 3

## Metodologia

### 3.1 - Introdução

Neste capítulo é exposta a metodologia desenvolvida para a deteção de eventos anómalos no padrão do consumo. Numa primeira fase, foi realizada uma recolha do histórico da série de consumo da instalação real e uma posterior análise visual e interpretativa das séries, identificando por observação eventos anómalos. Posteriormente foram identificados os indicadores para a caracterização dos eventos anómalos, e por último, criados os modelos de parametrização e treino, que irão detetar automaticamente as anomalias. Em fase de conclusão, toda esta metodologia foi testada com as séries reais de modo a ser possível validar o modelo criado.

O modelo foi desenvolvido com o intuito de ser facilmente implementado em plataformas de gestão de energia reais integrando-se num serviço de monitorização de consumos.

A aplicação deste modelo depende apenas dos consumos do passado, que podem ser adquiridos com equipamentos de monitorização presentes na instalação.

### 3.1.1 - Abordagem

A abordagem utilizada para o desenvolvimento do modelo de deteção de eventos anómalos foi representada de forma esquemática no seguinte diagrama:

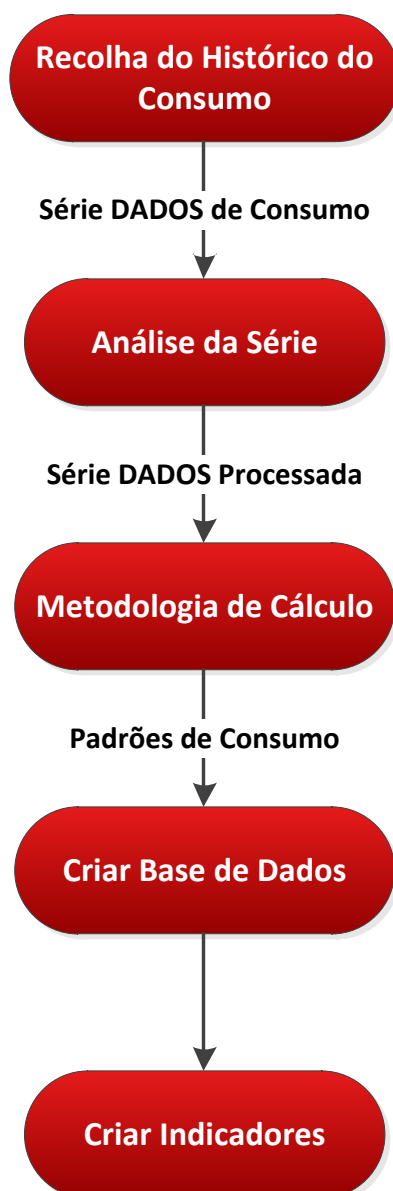


Figura 3.1 - Fluxograma da abordagem seguida no desenvolvimento do modelo

### 3.1.2 - Objetivos

O sistema de inteligência foi desenvolvido tendo em vista a sua integração em serviços de monitorização de consumos. Esta necessidade surge no âmbito da obtenção de uma maior eficiência energética, isto é de uma atuação mais precisa, rápida e eficaz nas diferentes anomalias que podem surgir nas instalações, tornando o consumidor um elemento mais ativo nas operações da rede. Tendo em conta as premissas referidas anteriormente foram estipulados objetivos a atingir, sendo eles:

- O modelo deverá ser fácil de implementar nas mais diversas séries de consumo;
- As “*baselines*” construídas no modelo deverão representar com o máximo de pormenor o consumo da instalação;
- O modelo deve ser aplicado em casos reais de modo a ser possível testar o seu desempenho;
- O modelo deverá ser robusto;

Todos estes objetivos foram criados com o intuito de garantir o máximo de fiabilidade e robustez do modelo desenvolvido, uma vez que posteriormente o sistema possa ser integrado em serviços de monitorização de consumos em tempo real, devendo ser fiável e apresentar um elevado nível de credibilidade para as diferentes situações a que será submetido.

## 3.2 - Modelo de deteção de anomalias

### 3.2.1 - Recolha do histórico

Para desenvolver este modelo foram utilizadas séries de consumos da empresa Savinor referentes ao intervalo de tempo entre Maio de 2011 e Janeiro de 2012. Os dados foram medidos hora a hora, tendo sido representados no seguinte gráfico.

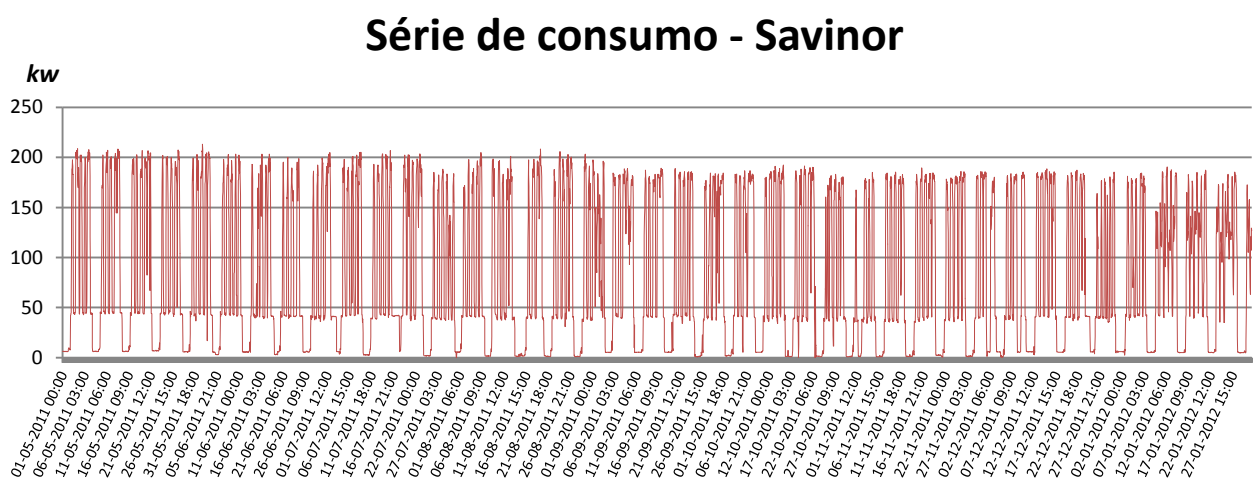


Figura 3.2 - Série do consumo (Savinor)

Analisando a série representada foi possível verificar a existência de um padrão de consumo, detectando-se consumos tendencialmente mais baixos aos fins de semana, assim como alguns feriados e pontes. O consumo máximo da empresa até ao mês de Setembro era na ordem dos 200 kw, notando-se que após esta data houve uma pequena redução do consumo máximo para valores próximos de 180 kw. Ao questionar a empresa acerca desta redução concluiu-se que esta diminuição se deveu a uma reestruturação da empresa no âmbito da poupança energética, que inclui alteração de luminárias e substituição que equipamentos com baixo rendimento.

Para melhor entender o consumo da empresa foi necessário repartir toda a série de modo a poder criar um padrão do consumo mais fiel ao original. Deste modo no ponto 3.2.2 será feita uma análise detalhada da série.

### 3.2.2 - Análise visual e interpretativa da série do consumo

Neste ponto a série do consumo foi dividida e analisada individualmente para cada dia da semana, de modo a perceber ao pormenor qual o modo de operar da empresa.

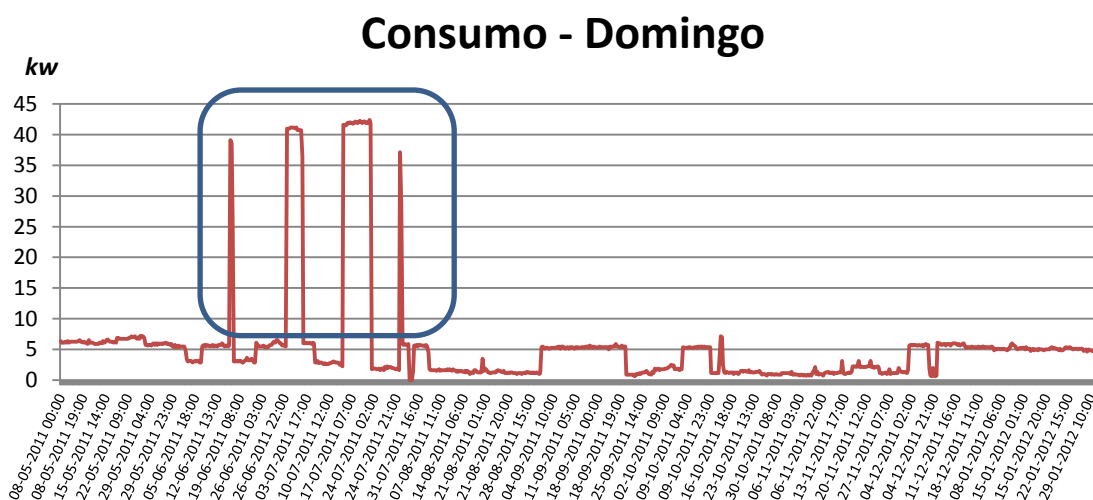


Figura 3.3 - Consumo ao Domingo

Analisando os domingos isoladamente podemos concluir que este representa um dia onde não há produção, isto é o consumo apresenta valores reduzidos.

No entanto, foi possível identificar a existência de valores de consumo fora do normal, como por exemplo, os detectados no gráfico apresentado acima, identificados com retângulo azul. Neste caso verificam-se consumos superiores a 40 kw, que por comparação com os restantes dias, é considerado anormal e se deve ao facto de neste dia o consumo se ter comportado como um dia onde se prepara a os equipamentos para dar início ao processo de fabrico.

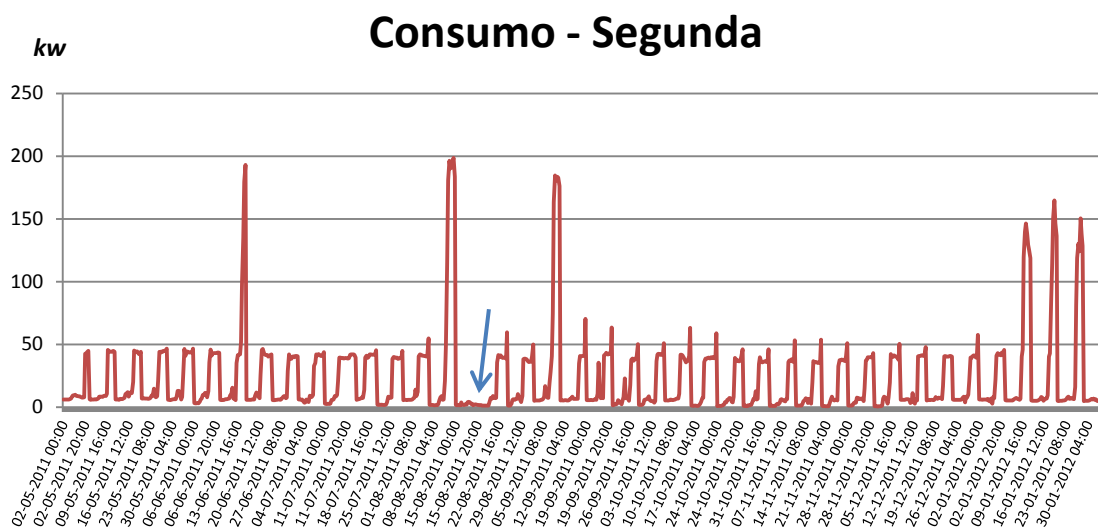


Figura 3.4 - Consumo às Segundas

Por norma às segundas-feiras o perfil do consumo varia entre um máximo de 50 kw e um mínimo de 0,6 kw, estando em média entre as 0h e as 16h no valor mínimo e entre as 17 horas e 23 horas no valor máximo, não invalidando os picos existentes. Os operadores da indústria descreveram que a segunda-feira é o dia onde todos os equipamentos são ligados e colocados a postos para no dia seguinte a indústria dar início ao processo de fabrico. Fazendo a análise do perfil de consumo pode-se identificar que em alguns dias a empresa apresenta consumos elevados, isto deve-se ao excesso de matéria-prima. Deste modo a empresa opta por dar início ao processo de fabrico mais cedo. Outro exemplo interessante é o feriado de 15 de Agosto (assinalado com a seta azul) onde não existe consumo.

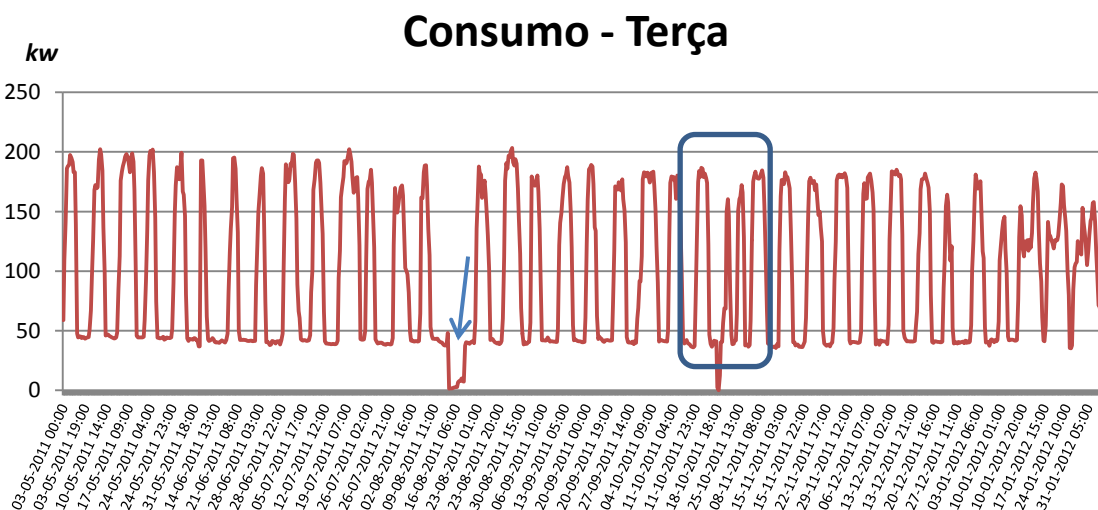


Figura 3.5 - Consumo às Terças-feiras

No que diz respeito ao consumo energético às terças-feiras conclui-se que este dia representa um dia de produção, isto é um dia onde o consumo apresenta os valores máximos de aproximadamente 200kw. Mais um vez o perfil apresenta consumos irregulares, como por

exemplo o ponto assinalado com a seta. Este representa o dia seguinte ao feriado de 15 de agosto, em conversa com a empresa concluiu-se que foi um dia utilizado para manutenção de máquinas tal como aconteceu na zona selecionada com o retângulo azul.

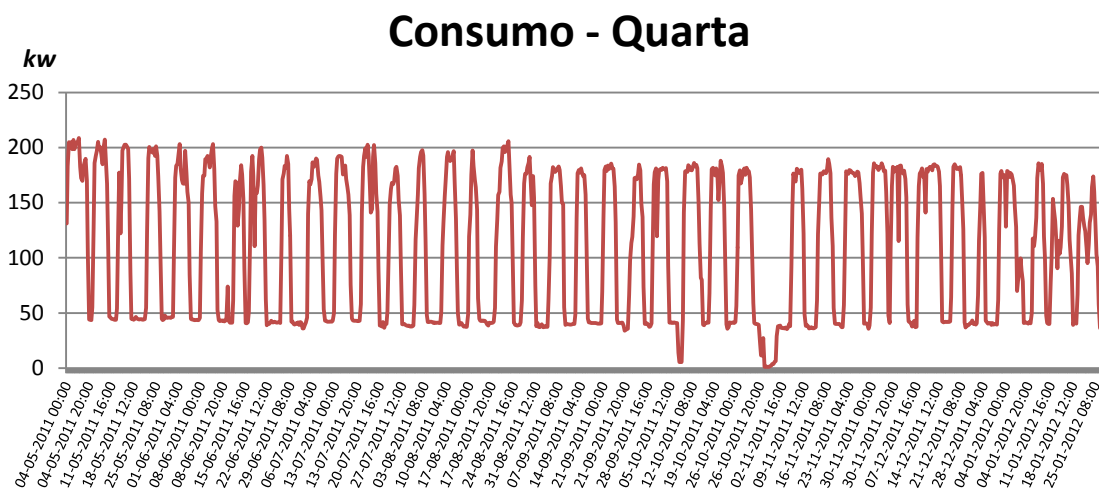


Figura 3.6 - Consumo às Quartas-feiras

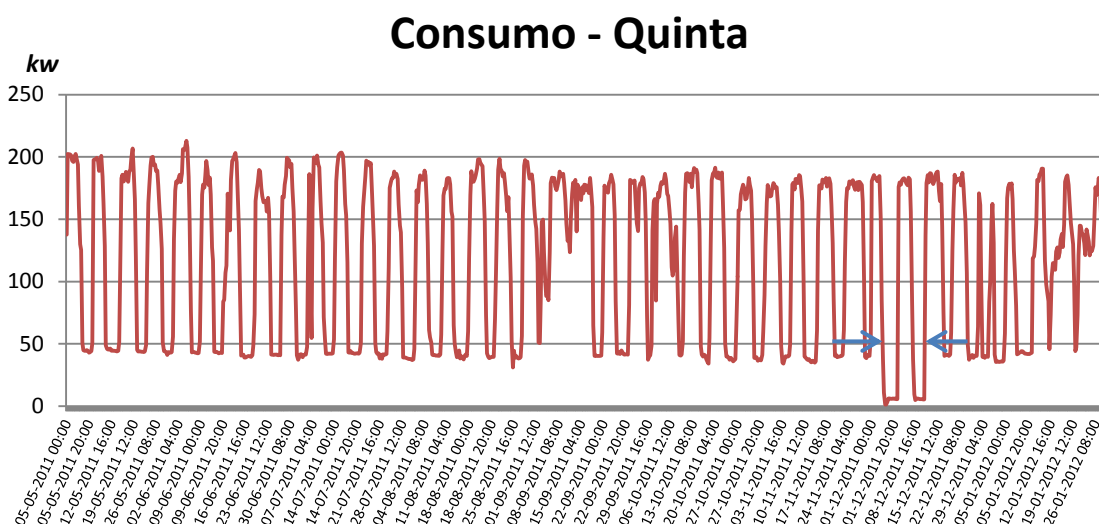


Figura 3.7 - Consumo às Quintas-feiras

Relativamente ao comportamento do consumo nas quartas e quintas-feiras conclui-se que correspondem tal como as terças-feiras a dias de produção. Através destes perfis podemos concluir o tipo de “base line” que vai corresponder aos dias de produção, sendo nestes dias que se vai concentrar o maior consumo energético. No gráfico de quinta-feira pode-se identificar dois pontos fora do consumo normal (identificados com as setas azuis), estes correspondem aos feriados de 1 e 8 de Dezembro. Deste modo podemos também concluir em conformidade com o que foi dito pelos responsáveis pela empresa que nos feriados mais importantes a industria para a produção.

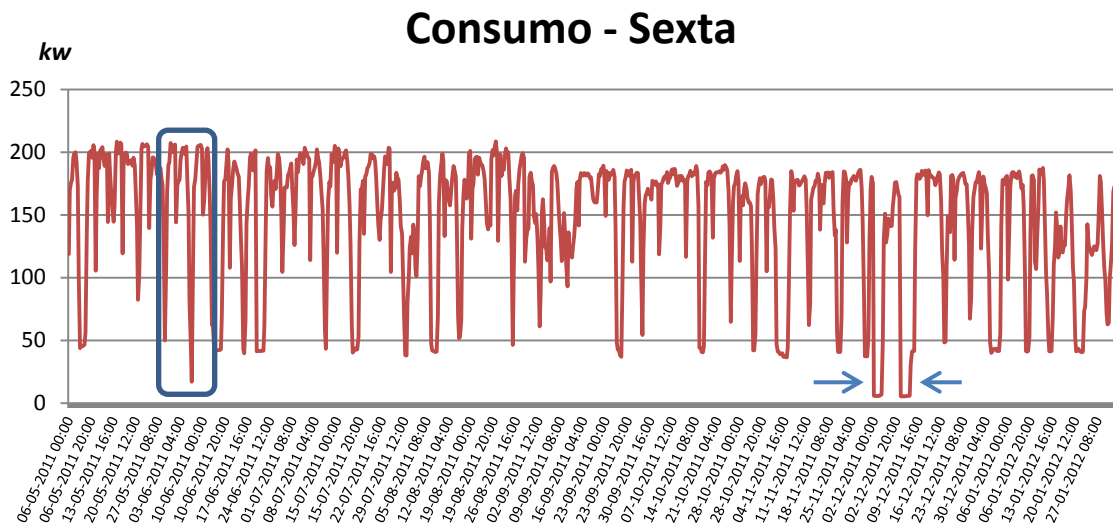


Figura 3.8 - Consumo às Sexta-feiras

O perfil de consumo que representa a sexta-feira apresenta um padrão bastante particular, uma vez que embora tenha um máximo dentro da mesma ordem dos dias de produção, não apresenta um padrão tão definido como os restantes dias. Contudo através deste perfil podemos tirar algumas conclusões no que diz respeito à área dentro do rectângulo azul, relativo ao feriado de 10 de Junho, onde podemos analisar que o consumo não varia por ser feriado. Já nos dois pontos assinalados pelas duas setas que representam os dois dias seguintes aos feriados de 1 e 8 de Dezembro conclui-se que a empresa opta por fazer ponte.

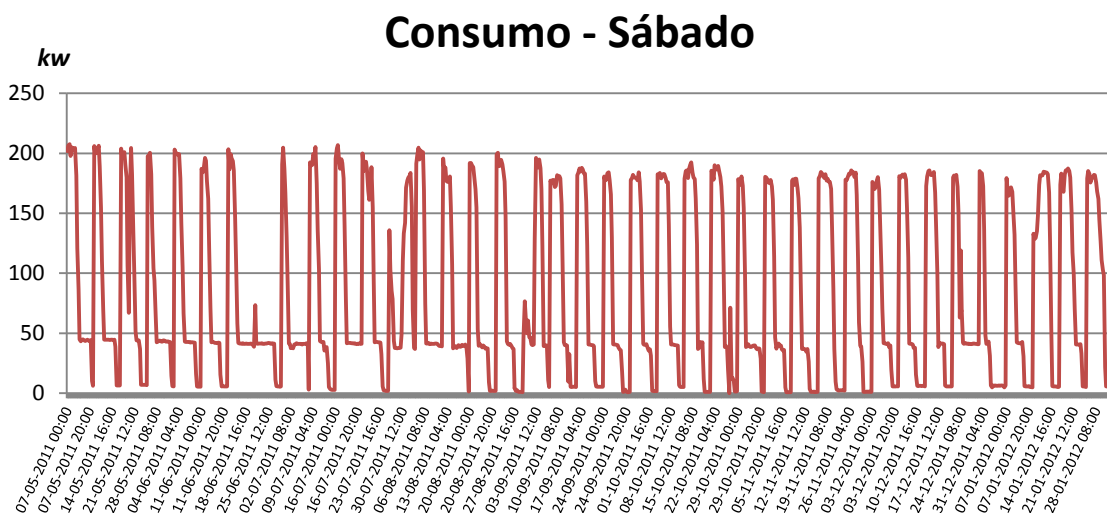


Figura 3.9 - Consumo aos Sábados

No caso do sábado, pode-se ver que este representa um dia de fim de produção. Neste caso a partir das 7h o consumo de energia tem uma diminuição intensa. Relativamente aos pontos onde se nota uma variação do perfil do consumo, foi abordado mais uma vez o assunto com os operadores da empresa e conclui-se que representam dias onde possivelmente ocorreu alguma anomalia ou manobras de manutenção.

### 3.2.3 - Caracterização qualitativa de anomalias

Durante a análise interpretativa do consumo surgiram diversas dúvidas no que diz respeito a pequenos pontos onde o consumo tinha um comportamento anormal.

Logo para realizar a caracterização qualitativa das anomalias foi necessário recorrer a contactos com os gestores da indústria, uma vez que através desta interação foi possível concluir o que levou a esses comportamentos anormais. Em alguns dos casos esses comportamentos podem ter sido premeditados, isto é, podem corresponder a manutenções agendadas, a feriados ou pontes ou até mesmo a dias onde a matéria-prima não era suficiente para dar início ao ciclo de produção.

Perante esta situação foi desenvolvido um *layout* onde o consumo estava dividido por dias da semana. Esta folha continha um gráfico de linhas com o consumo energético e um gráfico de dispersão onde era possível analisar o consumo em função da hora. Em conjunto com os gráficos era também calculado o valor máximo e mínimo da série do consumo de modo a identificar rapidamente valores anormais.

Esta folha de cálculo teve como objetivo tornar a interação com a empresa mais eficaz, tornando deste modo o número de reuniões reduzidas. Neste caso concreto foi necessário apenas uma reunião para conseguir identificar todos os pontos anómalos e o modo em que a empresa opera. De seguida está apresentado o *layout* utilizado:

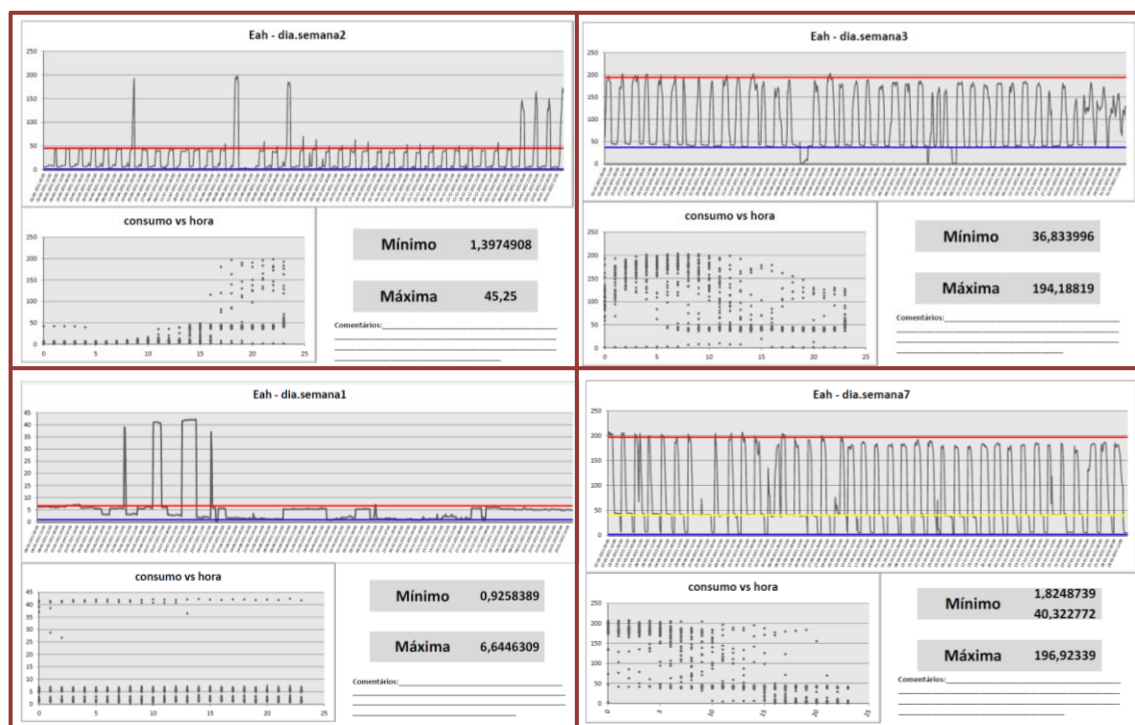


Figura 3.10 - Caracterização qualitativa de anomalias

### 3.2.4 - Pré-processamento da série do consumo

Para implementar uma metodologia de reconhecimento de padrões é necessário eliminar os pontos anômalos presentes na série do consumo original, de modo ao desenvolvimento do modelo não se basear em pontos anômalos. Desta forma todo o processo anteriormente descrito é de extrema importância para a fiabilidade no modelo. Após determinar a série DADOS processada, onde todos os pontos anômalos foram eliminados foi realizada uma análise visual e interpretativa de modo a identificar o possível número de *clusters* em que a série se poderia dividir.

Com o pré-processamento da série do consumo concretizado, o ponto seguinte consiste na realização da metodologia que vai realizar o reconhecimento de padrões (*clustering*). Na tabela 3.1 está representada parte da matriz que engloba a série DADOS processada que será utilizada posteriormente na realização da metodologia. A série DADOS processada em conjunto com a data (hora, mês, dia da semana) são os *inputs* do método.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
01-05-2011 00:00	6,3	6,3	6,3	6,1	6,3	6,1	6,3	6,3	6,3	6,3	6,0	6,0	6,1	6,1	6,1	6,3	6,3	6,1	6,4	6,0	6,1	6,1	6,3	5,9
02-05-2011 00:00	6,1	6,1	6,0	6,1	6,0	6,1	6,3	6,8	8,3	9,4	9,4	9,9	8,9	8,8	8,5	8,6	8,0	7,6	7,5	7,9	42,5	42,6	44,0	44,8
03-05-2011 00:00	58,9	109,0	141,4	186,3	188,1	189,1	197,5	194,9	191,5	182,9	183,0	138,9	47,5	44,3	45,1	45,0	43,6	44,6	44,6	43,1	44,1	44,4	44,6	50,9
04-05-2011 00:00	131,0	183,0	205,0	199,9	205,0	198,6	206,9	198,4	203,8	203,9	205,6	208,6	182,4	172,3	169,8	182,3	187,0	189,9	170,1	102,6	44,0	43,9	43,5	55,1
05-05-2011 00:00	137,8	202,4	202,3	202,1	201,1	197,0	196,0	200,0	202,4	198,5	193,9	160,9	129,9	125,4	50,6	45,0	44,4	44,4	45,0	44,4	42,9	43,4	44,1	50,8
06-05-2011 00:00	118,9	169,8	175,9	178,0	195,8	198,8	199,9	192,4	173,0	100,9	43,9	45,1	45,3	45,5	46,6	56,1	133,4	176,5	199,4	199,8	201,3	195,5	205,5	201,1
07-05-2011 00:00	201,9	206,6	207,6	197,9	204,9	203,0	204,3	204,4	181,4	120,6	86,6	45,0	43,4	44,6	45,0	44,0	43,5	44,1	44,5	43,9	43,4	44,3	10,1	6,3
08-05-2011 00:00	6,4	6,1	6,1	6,1	6,3	6,1	6,4	6,1	6,3	6,3	6,3	6,3	6,3	6,3	6,3	6,4	6,5	6,3	6,3	6,1	6,3	6,1	6,1	5,9
09-05-2011 00:00	6,3	5,9	6,0	6,0	6,0	6,3	6,1	6,3	6,8	8,4	8,4	8,4	8,0	8,8	9,0	8,8	10,5	45,6	44,0	44,4	44,0	44,4	44,8	44,1
10-05-2011 00:00	66,8	94,3	126,5	168,3	172,3	169,5	170,3	193,8	202,1	197,9	184,1	136,8	99,1	45,9	46,4	46,6	45,6	45,1	44,5	44,0	43,4	43,5	44,3	48,3
11-05-2011 00:00	126,4	186,0	192,6	196,9	205,3	198,6	200,3	194,1	184,8	199,8	207,1	190,4	167,3	119,9	46,8	46,0	44,6	44,8	44,0	43,6	43,6	54,0	130,4	177,3
12-05-2011 00:00	197,5	198,1	198,3	199,3	196,4	188,9	195,5	200,8	184,3	162,6	128,5	48,8	46,1	45,4	45,4	46,3	44,5	44,8	44,5	44,5	44,1	44,0	44,5	51,4
13-05-2011 00:00	105,4	165,8	198,8	187,0	201,3	202,1	204,0	195,4	189,0	198,1	196,3	144,4	198,8	182,3	165,1	154,8	144,5	158,3	199,8	208,3	198,9	204,1	207,5	206,5
14-05-2011 00:00	206,1	200,0	200,1	202,9	206,3	189,1	154,9	110,1	69,4	44,8	44,8	44,8	44,8	44,5	44,3	44,5	44,6	44,4	44,6	38,8	6,5	6,5	6,5	6,4
15-05-2011 00:00	6,5	6,1	6,1	6,1	6,0	5,9	5,9	5,9	5,9	6,1	6,1	6,0	6,4	6,1	6,3	6,5	6,6	6,4	6,3	6,3	6,3	6,1	6,3	6,1
16-05-2011 00:00	6,0	6,1	6,0	5,9	6,5	6,6	6,3	6,8	7,4	10,4	10,9	11,8	8,3	10,8	11,9	11,1	20,0	45,1	43,9	44,4	44,4	42,5	42,0	44,1

Tabela 3.1- Parte da matriz DATA

A matriz DATA processada possui dimensão M por N, onde o M é o número de dias e N as 24h.

### 3.2.5 - Metodologia de reconhecimento de padrões

Nesta dissertação, foi implementada uma metodologia de reconhecimento de padrões (*clustering*), o qual é definido como um processo pelo qual se procura classificar objetos em categorias.

Um problema de análise de *cluster* pode ser tratado segundo diferentes abordagens, entre elas, a abordagem convencional (*crisp*), na qual cada objeto deve ser classificado única e exclusivamente numa determinada categoria, e a abordagem *fuzzy*, mais flexível, na qual um objeto pode ser classificado em vários grupos, com diferentes graus de associação a cada um deles (pertença). Por exemplo, um certo ponto que se encontre perto do centro de um *cluster* terá um elevado grau de pertença a esse *cluster*, já no caso de um outro ponto que se encontra distante do centro de um determinado *cluster* terá um baixo grau de pertença. A tabela 3.2 representa alguns dos algoritmos existentes para o reconhecimento de padrões:

Category		Partitional					
Name	Type of data	Complexity <sup>a</sup>	Geometry	Outliers, noise	Input parameters	Results	Clustering criterion
K-Mean	Numerical	$O(n)$	Non-convex shapes	No	Number of clusters	Center of clusters	$\min_{v_1, v_2, \dots, v_k}(E_k)$ $E_k = \sum_{i=1}^k \sum_{j=1}^n d^2(x_j, v_i)$
K-mode	Categorical	$O(n)$	Non-convex shapes	No	Number of clusters	Modes of clusters	$\min_{Q_1, Q_2, \dots, Q_k}(E_k)$ $E = \sum_{i=1}^k \sum_{j=1}^n d(X_i, Q_i)$ $D(X_i, Q_i)$ = distance between categorical objects $X_i$ , and modes $Q_i$
PAM	Numerical	$O(k(n-k)^2)$	Non-convex shapes	No	Number of clusters	Medoids of clusters	$\min(TC_{ih})$ $TC_{ih} = \sum_j C_{jih}$
CLARA	Numerical	$O(k(40+k)^2 + k(n-k))$	Non-convex shapes	No	Number of clusters	Medoids of clusters	$\min(TC_{ih})$ $TC_{ih} = \sum_j C_{jih}$ ( $C_{jih}$ = the cost of replacing center $i$ with $h$ as far as $O_j$ is concerned)
CLARANS	Numerical	$O(kn^2)$	Non-convex shapes	No	Number of clusters, maximum number of neighbors examined	Medoids of clusters	$\min(TC_{ih})$ $TC_{ih} = \sum_j C_{jih}$
FCM	Numerical	$O(n)$	Non-convex shapes	No	Number of clusters	Center of cluster, beliefs	$\min_{U, v_1, v_2, \dots, v_k}(J_m(U, V))$
Fuzzy C-Means							$J_m(U, V) = \sum_{i=1}^k \sum_{j=1}^n U_{ik}^m d^2(x_j, v_i)$

<sup>a</sup>  $n$  is the number of points in the dataset and  $k$  the number of clusters defined.

**Tabela 3.2 - Algoritmos para o reconhecimento de padrões [26]**

A metodologia criada segue o caso específico de *Fuzzy Clustering*, onde diferentes algoritmos podem ser utilizados. Este modelo foi criado através do algoritmo *fuzzy c-means*, desenvolvido por J. C. Bezdek. No diagrama seguinte estão esquematizadas as fases do processo para o reconhecimento de padrões:

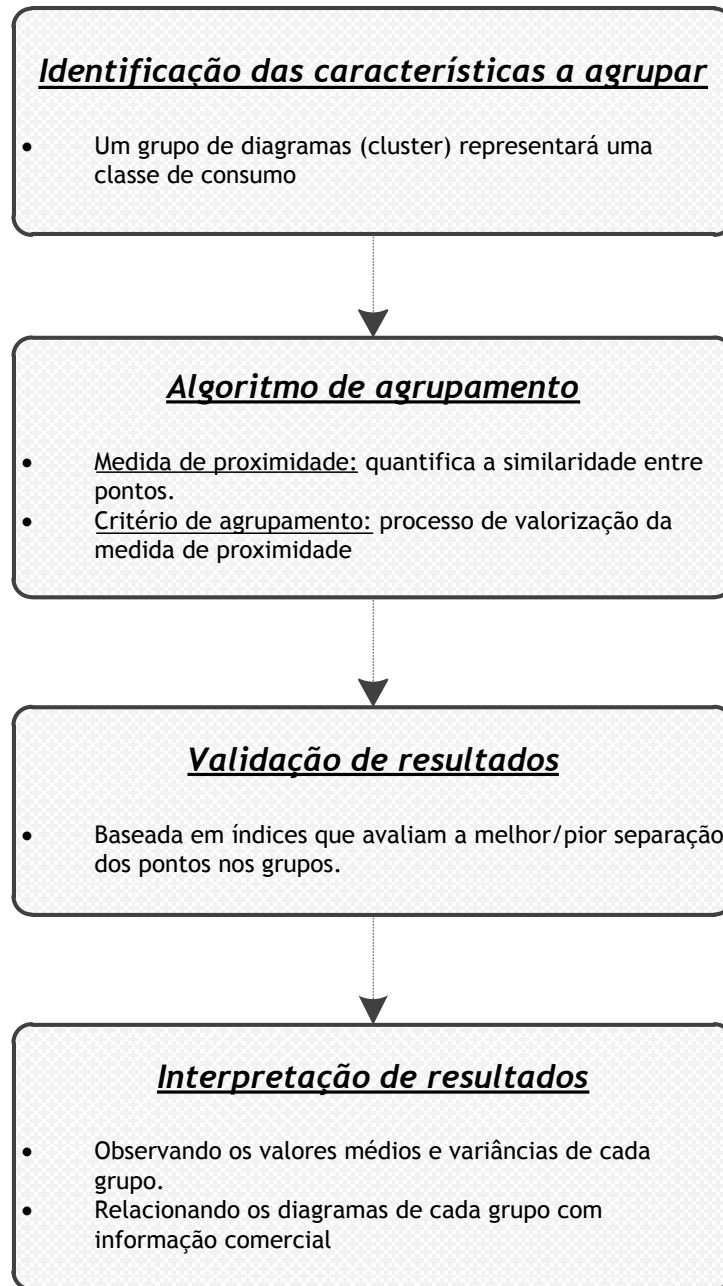


Figura 3.11 - Fases do processo de reconhecimento de padrões

Nesta dissertação, para o desenvolvimento da metodologia foi adotado o algoritmo de clusterização *fuzzy* → **fuzzy c-means**. Após detalhadas as fases do processo o algoritmo desenvolvido segue o seguinte raciocínio:

Supondo a série de DADOS processada,  $X = \{x_1, \dots, x_n\}$  em que cada valor  $X_k \in \mathfrak{R}$ ,  $k = 1, \dots, n$ , e deseja-se organizá-los em  $c$  clusters,  $C = \{C_1, \dots, C_c\}$ .

O algoritmo *fuzzy c-means* (FCM) é um algoritmo não hierárquico de *clusterização* cujo objetivo é fornecer uma partição *fuzzy* de um conjunto de valores em  $c$  *clusters*.

Logo o FCM define e minimiza uma função objetivo, a qual mede a adaptação entre os dados e os *clusters*. A função objetivo é definida como (Bezdek, 1981):

$$J_m(\mathbf{U}, \mathbf{G}) = \sum_{k=1}^n \sum_{i=1}^c (\mathbf{u}_{ik})^m \phi(\mathbf{X}_k, \mathbf{g}_i) = \sum_{k=1}^n \sum_{i=1}^c (\mathbf{u}_{ik})^m \sum_{l=1}^p (\mathbf{x}_{kl} - \mathbf{g}_{il})^2 \quad (1)$$

$$\phi(\mathbf{x}_k, \mathbf{g}_i) = \sum_{l=1}^p (\mathbf{x}_{kl} - \mathbf{g}_{il})^2 \quad (2)$$

Onde,

$\mathbf{U} \in \mathbf{M}_{fc}$  = Matriz de pertença  $\{u_{ik}\}$  ou apenas pertença, do valor k ao *cluster*  $C_i$ ;

$\mathbf{G}$  = Vector dos centros dos *clusters*  $(g_1, \dots, g_k), g_i = (g_{i1}, \dots, g_{ip})$  ;

$m \in ]1, +\infty[$  = É o peso atribuído à pertença. Quanto maior o m menor será a contribuição dos pontos mais afastados dos centros ;

$\phi(\mathbf{x}_k, \mathbf{g}_i)$  = É o quadrado da distância euclidiana entre o ponto e o centro do *cluster*  $c$ .

Os dados de entrada são n pontos  $\mathbf{X}_k$  (diagrama). Cada ponto é caracterizado por um conjunto de dimensões  $h$ ,  $\mathbf{X}_k = \{x_1, \dots, x_h\}$  (discretização do diagrama em horas).

Relativamente aos resultados estes são os centros de cada *cluster*  $\mathbf{G}$  e a pertença de cada ponto a cada *cluster*, sendo estas também as variáveis da minimização.

Portanto escolhido o numero de *clusters*  $c$  e o peso atribuído à pertença  $m$ :

- Define-se uma matriz inicial de pertenças  $\mathbf{U}$ , que somem 1 para cada um dos pontos;
- De seguida encontra-se os centros  $\mathbf{G}$  usando a fórmula:

$$g_{il} = \frac{\sum_{k=1}^n (\mathbf{u}_{ik})^m \mathbf{x}_{kl}}{\sum_{k=1}^n (\mathbf{u}_{ik})^m} \text{ para } l = 1, \dots, p \quad (3)$$

- Recalculer as funções de pertença;  
(este ponto será abordado na secção 3.2.7 Cálculo das pertenças)
- O processo repete-se até que o somatório da variação da pertença entre iterações seja inferior ao erro.

$$\sum_{k=1}^n \sum_{i=1}^c (\mathbf{u}_{ik}^t - \mathbf{u}_{ik}^{t-1}) < \varepsilon \quad (4)$$

### 3.2.6 - Criação da base dados do modelo

Após a realização da metodologia anteriormente descrita os resultados obtidos são guardados numa base de dados, de modo a que seja possível nas etapas seguintes aceder aos resultados obtidos no algoritmo.

A primeira tabela a integrar a base de dados é:

- Tabela da matriz CENTER;

A matriz CENTER detém a informação relativa ao centro de cada *cluster*. Deste modo esta é uma das tabelas importantes para o funcionamento do modelo, sendo a partir desta que diversos calculos e análises vão ser desenvolvidos. Na figura e tabela seguintes estão representados os cinco *clusters* obtidos:

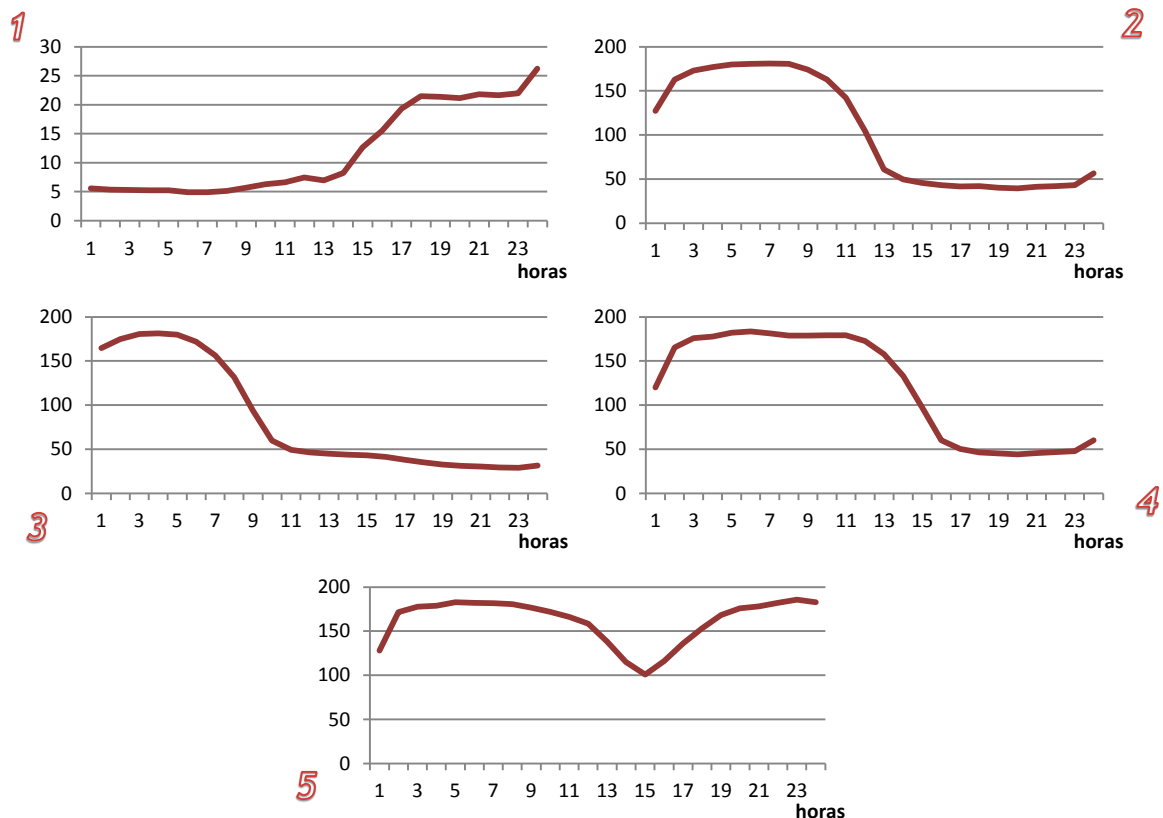


Figura 3.12 - Representação gráfica dos cinco *clusters*

		<i>Clusters</i>				
		1	2	3	4	5
1		5,535306	127,1454	164,6129	120,1931	127,9072
2		5,324179	162,9306	174,7455	165,2948	171,542
3		5,280007	173,1274	180,6169	175,8786	177,8507
4		5,205735	176,8719	181,4418	177,5846	178,9402
5		5,210935	180,0984	179,9198	182,0213	182,7163

6	4,870072	180,8115	171,7593	183,4719	182,1792
7	4,873407	180,9947	156,2311	181,4195	181,7935
8	5,106669	180,5036	131,7656	178,8223	180,4889
9	5,670241	174,1845	93,40306	178,9982	176,6151
10	6,28907	162,9721	59,89951	179,3584	171,7714
11	6,588473	142,2454	49,3602	179,2875	166,1431
12	7,457016	104,463	46,33462	172,7224	158,5136
13	6,957462	61,02544	44,92687	157,873	138,3755
14	8,184699	49,8149	43,74467	133,0441	115,0881
15	12,61107	45,53839	42,99014	97,46762	100,6658
16	15,51404	43,19036	41,33483	60,05583	116,128
17	19,3408	41,67876	38,1242	50,38454	136,0178
18	21,46653	42,17814	34,99446	46,45077	152,8957
19	21,34078	40,38011	32,61361	45,15026	168,2005
20	21,14883	39,54248	31,25683	44,34782	175,8066
21	21,77877	41,24132	30,25364	45,69894	178,0673
22	21,64472	41,9586	29,16092	46,60112	181,9191
23	21,98136	42,98215	28,87443	48,00915	185,7737
24	26,21326	56,38522	31,50994	60,28972	182,7941

Tabela 3.3 - Centros dos clusters

Obtidos os *clusters*, o passo seguinte foi analisar os gráficos de modo a relacionar os diferentes dias da semana com cada *cluster* obtido. Isto é os perfis dos diferentes dias vão ter maior pertença a determinado *cluster*. Este procedimento é importante pois o indicador de anomalia está diretamente ligado à relação previamente fixada entre os *clusters* e os dias da semana. Após esta análise a relação fixada foi:

1 - A	• Segunda - Feira
1 - B	• Domingo
2	• Terça - Feira • Quarta - Feira
3	• Terça - Feira • Quarta - Feira
3 - Sabado	• Sábado
4	• Quinta - Feira
5	• Sexta - Feira

Tabela 3.4 - Relação dos *clusters* com os dias da semana

Pode-se reparar que alguns dos *clusters* foram divididos em dois grupos, foi o caso do cluster 1 e 3. Isto foi necessário devido ao perfil do consumo de um dia como segunda-feira ou domingo estarem a ser englobados no mesmo *cluster*, deste modo desagregou-se o cluster 1. Como exemplo se o perfil do consumo estiver com maior pertença ao *cluster* 1, neste caso a partir das 15h se o consumo tender a aumentar vai ser atribuído o *cluster* 1-A, caso o consumo se mantenha será atribuído o *cluster* 1-B.

No que diz respeito ao *cluster* 3 este foi processado da mesma forma, isto é se o perfil do consumo de um determinado dia for identificado com o *cluster* 3 então será feita uma desagregação idêntica à descrita anteriormente. Na figura 3.13 está representado o diagrama que traduz o processo anteriormente descrito:

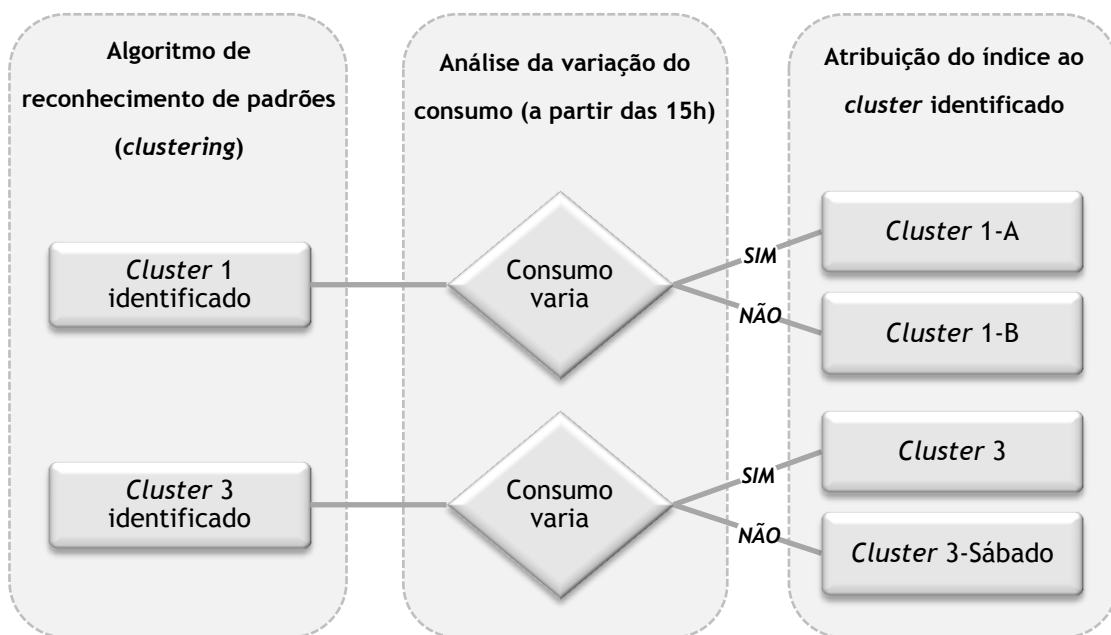


Figura 3.13 - Diagrama do processo de atribuição dos índices aos *clusters* 1 e 3

Em suma a base de dados vai conter a matriz CENTER que sai do algoritmo de reconhecimento de padrões (*clustering*) com as coordenadas dos centros dos 5 *clusters* e a tabela 3.3 com a correspondência entre os *clusters* e os dias da semana.

### 3.2.7 - Cálculo das pertenças

Após a obtenção da matriz dos centros dos *clusters* realiza-se o cálculo das funções de pertença  $U_{pc}$ :

$$U_{pc} = \frac{(1/d_{pc})^{2/(m-1)}}{\sum_{c=1}^k (1/d_{pc})^{2/(m-1)}} \quad (5)$$

Em que

$$d_{pc} = \sqrt{x_i - \text{Centro}_{\text{cluster}1}} \quad (6)$$

$$\sum_{c=1}^k (U_{pc}) = 1 \quad (7)$$

Os valores de pertença variam entre 0 e 1, onde 0 corresponde a pontos que não pertencem a determinado cluster, enquanto que valores de pertença 1 significam que o ponto tem pertença total ao *cluster*.

Com o cálculo das pertenças realizado o passo seguinte foi concretizar o modelo que vai determinar o *cluster* de maior pertença. Isto é cada ponto da série DADOS vai deter um valor de pertença alusivo a cada *cluster* e o somatório de todas essas pertenças será 1. Logo para cada ponto é necessário identificar o *cluster* de maior pertença. Na **figura 3.14** está representado o fluxograma do algoritmo desenvolvido para determinar o *cluster* de maior pertença.

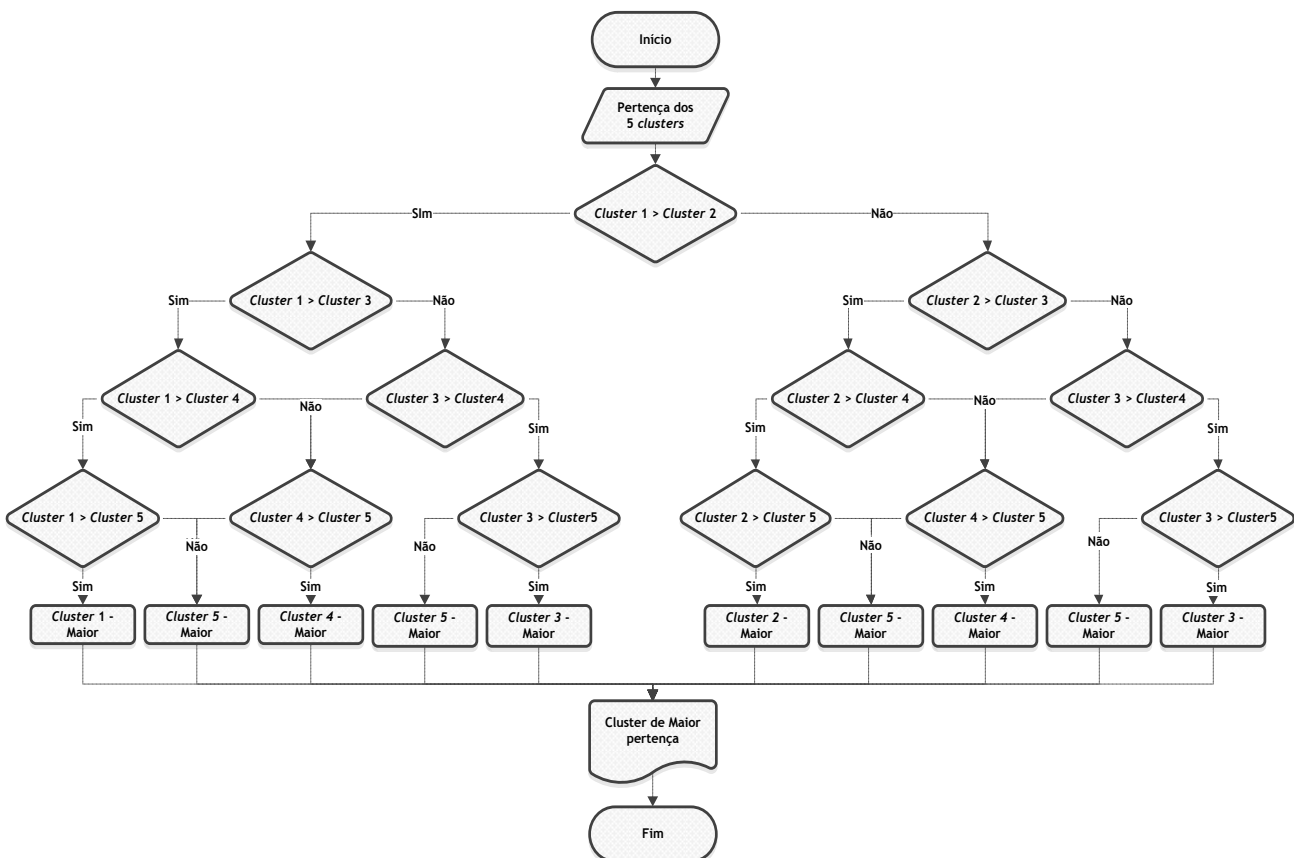


Figura 3.14 - Fluxograma com o algoritmo da determinação do *cluster* de maior pertença

Tendo alcançado todos estes pontos já estão reunidas todas as informações necessárias para elaborar os indicadores que vão identificar as anomalias.

### 3.2.8 - Indicadores de anomalias

Para a deteção das anomalias foram desenvolvidos dois indicadores, estes vão trabalhar em conjunto de modo a complementarem a informação relativa à anomalia. O objetivo é que, para além da deteção da anomalia, o indicador consiga identificar qual o tipo e se dentro da anomalia esperada o consumo apresenta uma discrepância elevada.

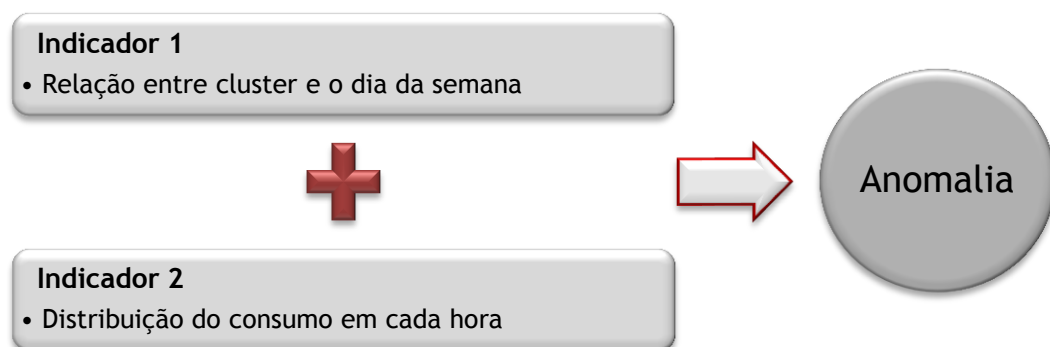


Figura 3.15 - Indicadores utilizados

#### 3.2.8.1 - Indicador 1

Este indicador tem como função alertar possíveis eventos anómalos no padrão do consumo. Mensagens serão impressas em conjunto com o indicador de modo a completar a informação editada pelo indicador.

O princípio de funcionamento consiste numa comparação entre o dia da semana em que o modelo está a operar e o respetivo *cluster* identificado. Para tal o indicador vai recorrer à tabela 3.4 gravada na base de dados que tem a informação que relaciona os *clusters* com os diferentes dias da semana. Deste modo o indicador vai constantemente avaliar se *cluster* de maior pertença identificado caracteriza o dia da semana em análise ou se equivale a um diferente dia. Os fluxogramas a baixo representados pretendem demonstrar o funcionamento do indicador 1 integrado na metodologia criada:

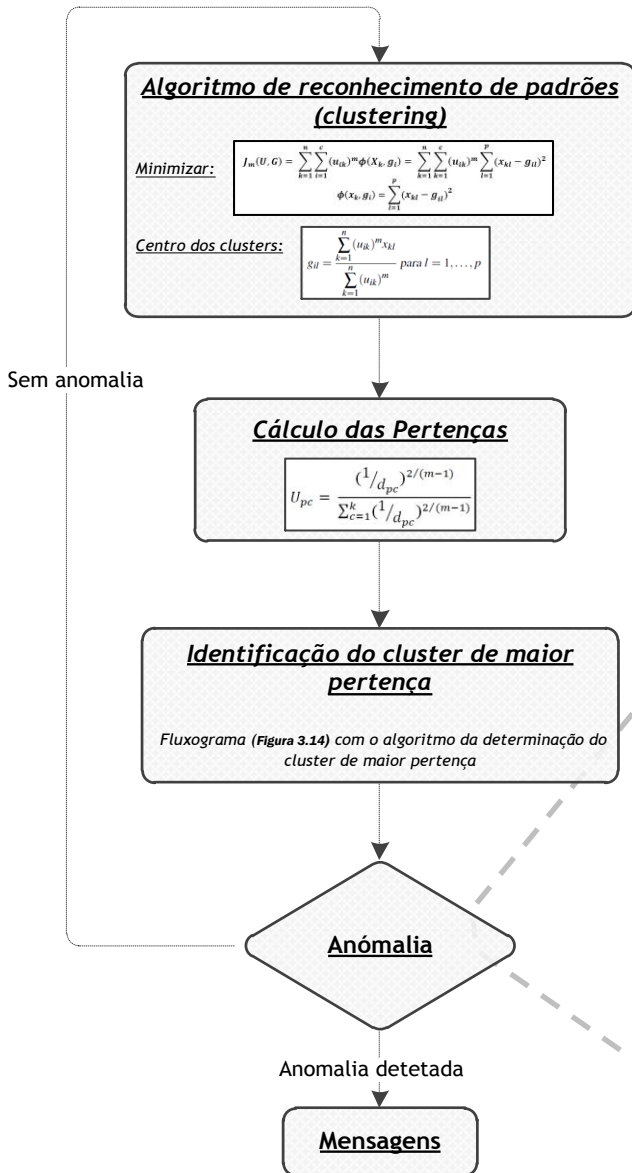


Figura 3.16 - Fluxograma do algoritmo de detecção de eventos anómalos

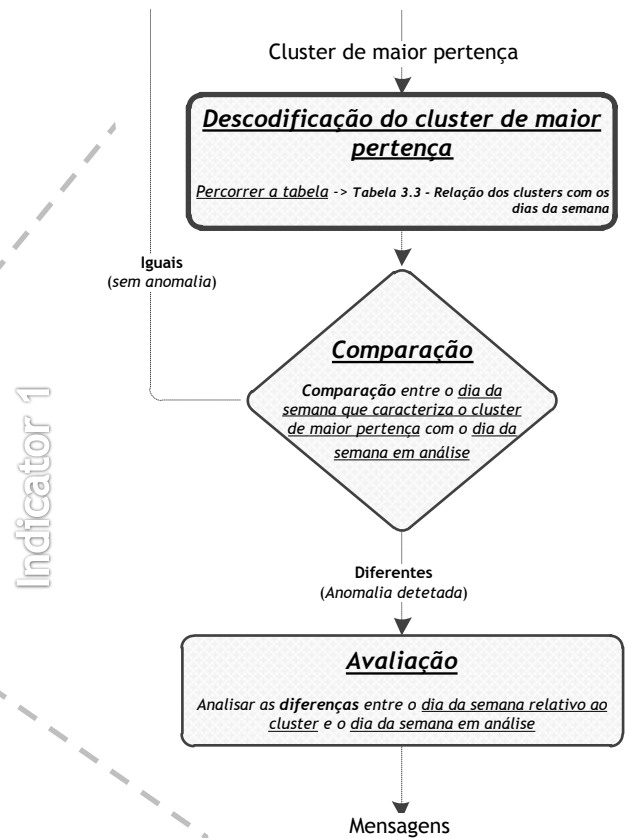


Figura 3.17 - Fluxograma do algoritmo do indicador 1

Através do fluxograma da figura 3.17 identifica-se as várias fases pelo qual o algoritmo é desenvolvido.

➤ Descodificação do cluster de maior pertença:

Nesta etapa o algoritmo vai percorrer a primeira coluna da tabela 3.4 de modo a encontrar o cluster identificado, para depois imprimir o dado correspondente da segunda coluna (dia da semana).

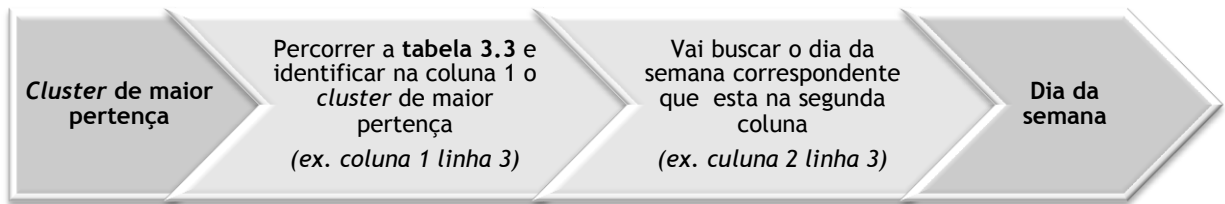


Figura 3.18 - Processo de descodificação do *cluster* de maior pertença

➤ Comparação:

Após o algoritmo ter descodificado e obtido o dia da semana que traduz o *cluster* de maior pertença, o passo seguinte é realizar a comparação entre o dia da semana adquirido no processo anterior e o dia da semana em análise.

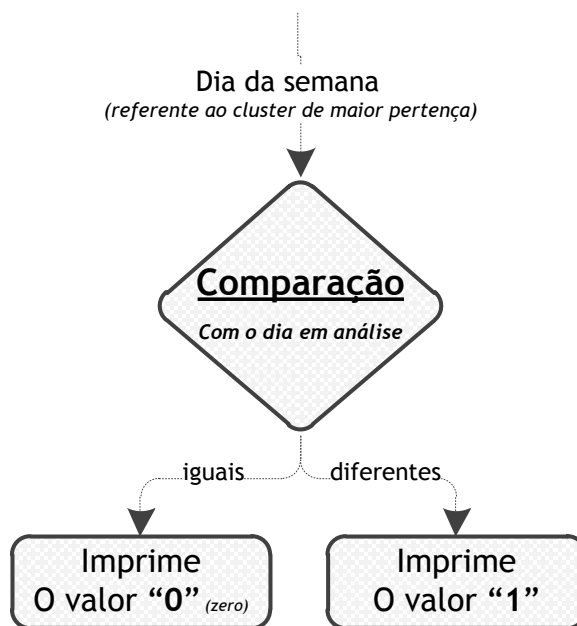


Figura 3.19 - Fluxograma da comparação entre os dias da semana

No fluxograma representado na **figura 3.19** identifica-se que o modelo após concretizar a comparação vai devolver o valor **0** se existir igualdade entre os dias da semana, e o valor **1** nas situações em que os dias da semana não são compatíveis.

Através da tabela seguinte (**tabela 3.5**) pode-se identificar o processo anteriormente descrito de forma a compreender o modelo desenvolvido.

<i>Hora</i>	<i>Dia da Semana</i>	<i>Data</i>	<i>Dados</i>	<i>Pertença</i>	<i>Cluster</i>	<i>Indicador 1</i>
0	1	01-05-2011 00:00	6,25	0,99987	1 - B	0
1	1	01-05-2011 01:00	6,25	0,99987	1 - B	0
2	1	01-05-2011 02:00	6,25	0,99987	1 - B	0
3	1	01-05-2011 03:00	6,125	0,99988	1 - B	0
4	1	01-05-2011 04:00	6,25	0,99987	1 - B	0

Tabela 3.5 - Demonstração do indicador 1

Na tabela identifica-se que o *cluster* obtido é o **1-B** para cada hora, deste modo o indicador vai percorrer a tabela 3.4 de forma a obter os dias da semana a que o perfil do *cluster 1-B* representa. Com esta informação vai efetivar a comparação com o dia da semana em que esta a ocorrer a análise, neste caso dia da semana 1.

Analisando este caso em particular o *cluster 1-B* caracteriza um perfil relativo ao dia de Domingo, e o dia da semana em análise é 1 (Domingo) deste modo o indicador será 0 pois existe conformidade entre o *cluster* e o dia da semana.

### ➤ Avaliação:

Neste ponto o modelo deve interpretar a comparação feita no ponto anterior, de modo a imprimir a mensagem de anomalia mais adequada.

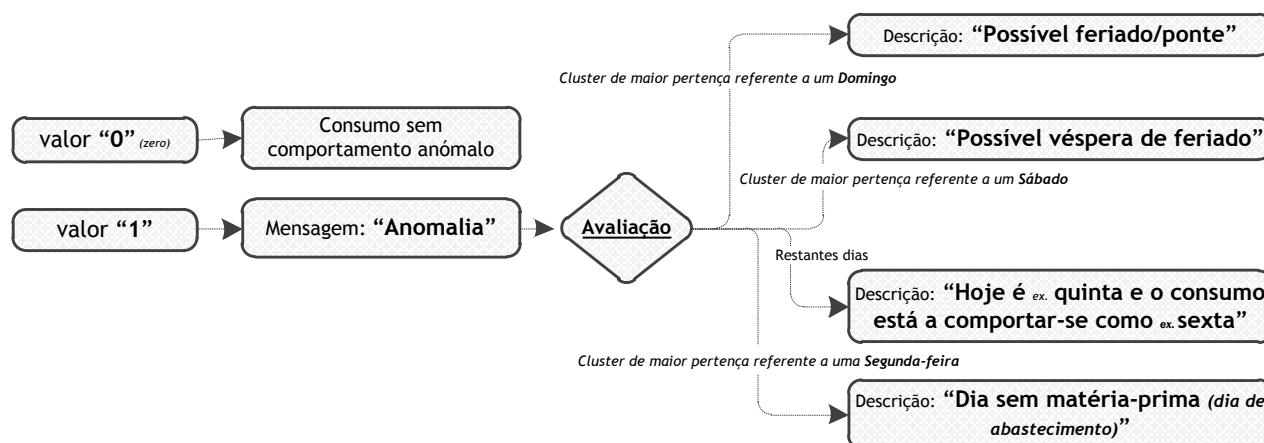


Figura 3.20 - Fluxograma do processo de avaliação do indicador 1

Analisando o fluxograma anterior observa-se que o modelo desenvolvido para avaliar as anomalias está diretamente ligado ao *cluster* de maior pertença. É a partir desta informação que o modelo vai determinar a descrição mais adequada para cada anomalia detetada.

As mensagens descritivas guardadas na base de dados são impressas de acordo com as decisões tomadas pelo modelo de avaliação. Como por exemplo, no caso do modelo de avaliação tomar a decisão do *cluster* de maior pertença ser referente a segunda-feira, aí o modelo irá imprimir a mensagem de: "**Anomalia** - *Dia sem matéria-prima (dia de abastecimento)*".

### 3.2.8.2 - Indicador 2

O indicador 2 tem como principal objetivo compreender o comportamento do consumo em relação ao perfil identificado pelo *cluster*, isto é para cada hora é analisada a distribuição

do consumo. Esta análise foi estudada de diversas formas e como conclusão o modelo desenvolvido para o indicador 2 usa como ferramenta a **distribuição normal**.

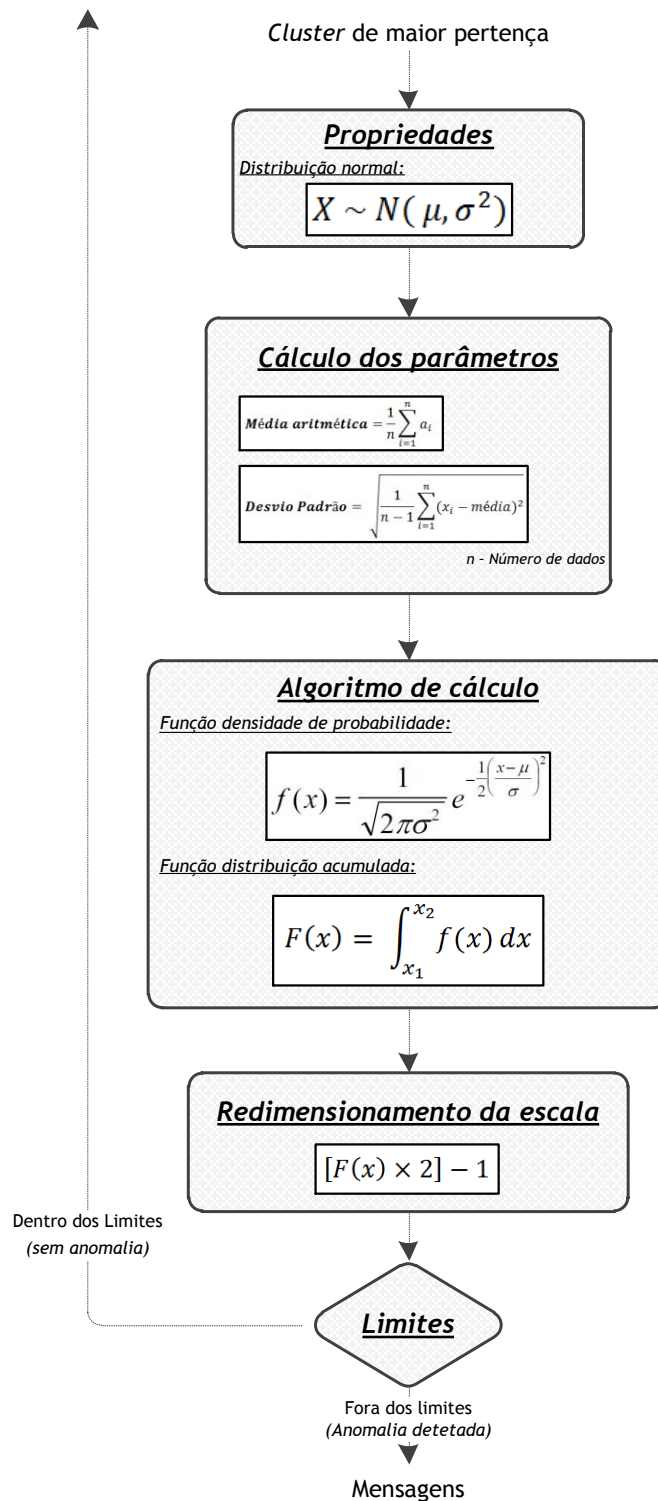


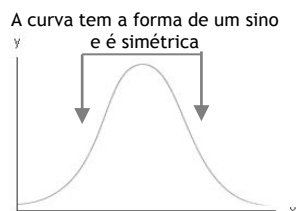
Figura 3.21 - Fluxograma do algoritmo do indicador 2

No fluxograma da figura 3.21 estão presentes as várias fases pelo qual o algoritmo é desenvolvido.

➤ Propriedades:

Uma variável aleatória contínua segue uma distribuição normal se a sua distribuição é:

- Simétrica;
- Apresenta um gráfico com a forma de um sino;



Se a variável aleatória  $X$  segue uma distribuição normal então define-se  $X$  como:

$$X \sim N(\mu, \sigma^2) \quad (8)$$

De um modo geral, a maior parte dos fenómenos probabilísticos de natureza contínua, e mesmo alguns de natureza discreta, tendem a seguir uma lei de distribuição designada por função de **distribuição normal**, ou de Gauss.

Esta lei de distribuição estabelece que os valores mais frequentes (isto é, os valores a que correspondem as maiores probabilidades) se encontram em torno da média da variável aleatória. Quanto mais afastados os valores estão da média (este afastamento é quantificado em termos de variância ou segundo momento em relação à média), quer acima quer abaixo desta, menos frequentes são. Esta interpretação da lei de distribuição normal é coerente com o que se passa com a maior parte dos fenómenos que ocorrem na natureza.

➤ Calculo dos parâmetros:

Para determinar a média, a variância e o desvio padrão foi utilizado o histórico de dados utilizado para desenvolver o modelo, deste modo foram filtrados os valores referentes a cada *cluster* e assim calculados os parâmetros para cada hora.

$$\text{Média aritmética} = \frac{1}{n} \sum_{i=1}^n a_i \quad (9)$$

$$\text{Desvio Padrão} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \text{média})^2} \quad (10) \quad n - \text{Número de dados}$$

$$\text{Variância} = s^2 = \frac{\sum_{i=1}^n (x_i - \text{média})^2}{(n-1)} \quad (11)$$

Após o cálculo de todas as médias e desvios padrões dos *clusters*, estes valores são organizados numa tabela que será integrada na base de dados. Com isto o cálculo do indicador apenas vai necessitar de recorrer à **tabela 3.6** para extrair a média e desvio padrão referente à hora e *cluster* em análise.

De seguida está representada a tabela criada e guardada na base de dados:

	1 - A		1 - B		2		3		3 - Sábado		4		5	
	média	desvioPAD	média	desvioPAD	média	desvioPAD	média	desvioPAD	média	desvioPAD	média	desvioPAD	média	desvioPAD
0	3,615384615	2,309470469	3,649193548	2,214869107	119,0603448	27,69798094	145,2583333	42,31006578	189,0833333	10,46909426	117,3225806	19,38593171	128,7569444	14,41834719
1	3,629807692	2,284316301	3,560483871	2,224595227	161,3189655	22,74130491	167,95	26,92038527	186,5416667	9,757742652	163,3346774	21,44968403	173,2291667	11,24961192
2	3,586538462	2,302924562	3,544354839	2,205709539	172,7672414	16,58123011	178,0833333	19,05640515	185,975	9,659516033	176,8890645	12,09523561	180,3541667	9,479975994
3	3,629807692	2,276367932	3,487903226	2,197383586	177,7284483	11,16440998	179,3166667	14,20975415	185,6333333	8,647546445	179,5604839	9,041044773	181,4652778	10,16226852
4	3,822115385	2,265825754	3,512096774	2,223069115	181,1422414	10,89308452	178,0833333	21,51233381	183,0166667	16,01832563	183,5282258	8,83186169	185,4097222	10,20901108
5	4,139423077	2,309850353	3,592741935	2,173698783	182,6724138	9,893259215	171,2333333	31,12768443	174,2833333	21,8983277	185,5564516	8,434356742	185,5902778	10,66215869
6	4,418269231	2,113037597	3,508064516	2,227885648	184,5387931	10,23251467	151,5083333	46,20107129	160,5833333	30,15372322	183,7943548	8,74788953	187,0416667	10,71728949
7	5,024038462	1,678920498	3,572580645	2,151495167	185,1465517	8,943402706	129,4166667	39,87451485	127,1	45,5705448	182,2056452	11,54391743	186,7708333	10,50982121
8	6,25	1,862122445	3,681451613	2,270700771	178,6551724	12,52402402	85,90833333	40,54787918	89,63333333	40,52816665	183,4677419	10,16780908	184,9305556	9,966231793
9	7,514423077	2,414882741	3,725806452	2,255183694	167,637931	15,15247449	50,75833333	19,83449751	53,21666667	25,86294002	184,7016129	11,03364539	182,625	14,11917501
10	7,894230769	1,984846439	3,548387097	2,168591473	145,1422414	24,61166527	41,55	1,764480903	41,01666667	2,24943776	183,516129	13,3914518	175,8333333	20,00491668
11	8,278846154	2,945612774	3,544354839	2,17479606	101,8362069	32,23916802	41,5	1,583959776	40,95	2,169080615	177,3548387	16,63928261	164,9861111	23,35916375
12	7,158653846	3,042317721	3,584677419	2,185746225	52,36206897	19,42069338	41,25833333	1,50924729	40,85833333	2,097972377	162,9153226	22,68790489	142,6458333	51,22472924
13	9,538461538	7,446372374	3,608870968	2,176039487	44,06465517	9,152637981	41,36666667	1,347506516	40,18333333	2,568641016	137,0524194	32,16266881	116,5	55,49844327
14	20,61538462	15,56409823	3,596774194	2,195570749	42,50431034	4,671876673	41,525	1,18510096	39,68333333	2,702869947	96,61693548	44,01417782	102,6180556	55,49457565
15	27,57692308	15,53588897	3,794354839	2,135403062	41,07758621	2,449128351	41,175	0,925578969	37,05	7,303955631	56,93145161	28,54466235	125,0625	40,49805392
16	35,27403846	11,02196779	3,629032258	2,180998403	40,76293103	2,474276001	40,675	0,990805949	29,425	15,87681373	46,04032258	13,83847998	148,4236111	36,55717065
17	39,85576923	2,909401207	3,439516129	2,200821572	40,61637931	2,434849987	40,61666667	1,220606797	21,88333333	17,14457085	40,91532258	2,321185737	165,0138889	25,01297049
18	40,1875	2,589703168	3,39516129	2,234692553	40,38793103	2,610150351	40,98333333	0,720965787	16,5	18,37627547	41,08064516	2,044185959	181,7361111	14,37157177
19	40,02403846	2,712475629	3,266129032	2,232277876	40,02155172	2,78189077	40,36666667	1,568002217	13,95	16,4270438	40,89112903	1,982407505	187,7638889	17,50981284
20	39,94711538	2,669544221	3,362903226	2,196198145	42,09051724	9,276680666	40,275	1,679179221	10,49166667	14,34394679	40,81854839	2,043240561	186,5138889	17,43211038
21	39,81730769	2,678906206	3,27016129	2,232593964	42,46982759	10,19265619	40,15833333	1,295137426	6,416666667	9,274541126	40,53629032	2,335904881	190,3819444	11,15082695
22	40,51442308	2,413329371	3,366935484	2,245582581	44,31465517	15,59220203	40,675	1,227293771	5,366666667	6,887078135	41,16129032	2,266879635	194,7013889	10,98906031
23	49,02403846	7,37800949	3,177419355	2,197719982	60,98706897	15,79437996	48,56666667	10,30948188	3,416666667	2,280383479	54,93951613	6,939085477	192,2291667	11,95389812

**Tabela 3.6 - Médias e desvios padrões dos *clusters***

Com isto já estão reunidas todas as condições para criar o algoritmo de cálculo do segundo indicador.

➤ Algoritmo de cálculo:

Esquemáticamente, esta distribuição tem uma curva em forma de sino, **simétrica em torno da média**, tendo o eixo das abcissas como assíntota horizontal (12).

Pela definição de assíntota horizontal:

$$\lim_{x \rightarrow \pm\infty} f(x) = 0 \quad (12)$$

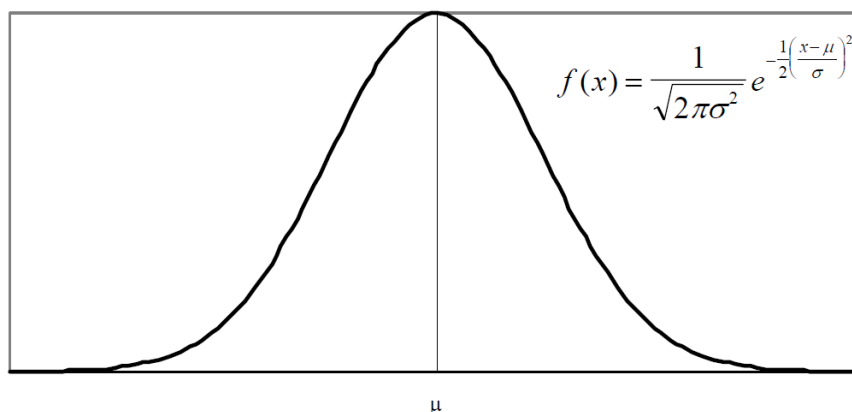


Figura 3.22 - Curva da distribuição normal

A curva da distribuição normal tem a seguinte função de densidade:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (13)$$

que depende de dois parâmetros: média  $\mu$ , e desvio padrão  $\sigma$ , que caracterizam cada uma das curvas em particular. Como anteriormente foi referido, a curva é simétrica em torno da média.

Neste algoritmo desenvolvido utiliza-se a **função distribuição acumulada**.

$$F(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x) dx \quad (14)$$

$$F(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \quad (15)$$

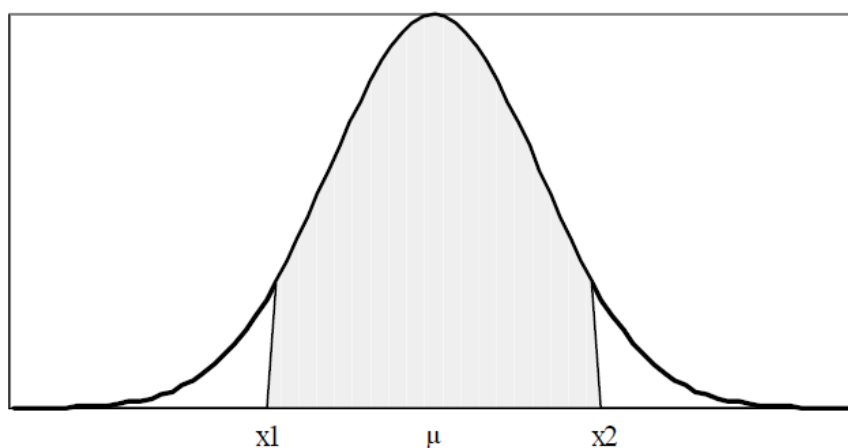


Figura 3.23 - Distribuição acumulada

A probabilidade de ocorrência de um dado valor  $X$  num determinado intervalo  $[x_1, x_2]$  equivale à área delimitada pela curva e pelo eixo das abcissas.

A utilização da distribuição normal é devido a esta ser inteiramente descrita pelos seus parâmetros de média e desvio padrão, ou seja, conhecendo-se estes consegue-se determinar qualquer probabilidade.

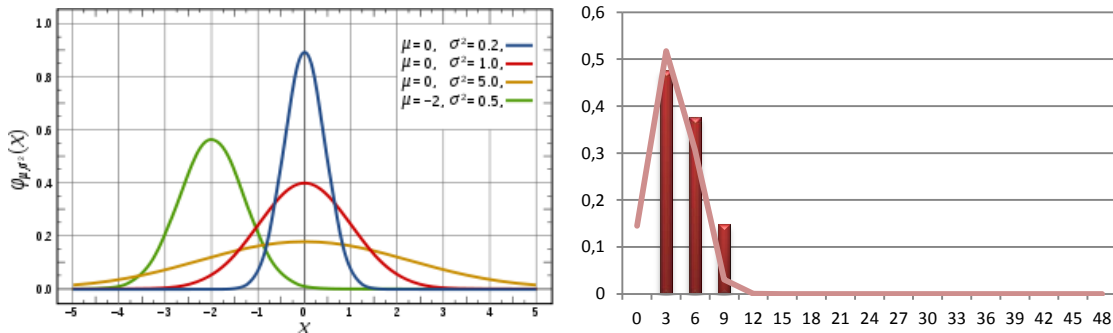


Figura 3.24 - Distribuição normal

Nos gráficos em cima representados identifica-se do lado esquerdo os gráficos da distribuição normal a variarem com a média e desvio padrão. Do lado esquerdo a distribuição normal da primeira hora do *cluster* 1.

➤ Redimensionamento da escala:

O indicador numa primeira fase detinha uma escala entre 0 e 1, que é a saída natural da função acumulada da distribuição normal ( $F(x)$ ).

Para a leitura do indicador ser mais clara de analisar foi mudada a escala, passando a ter uma escala entre -1 e 1. Deste modo os valores próximos de zero significam que o consumo está próximo do valor médio, sendo o caso em que o indicador apresenta valores vizinhos de 1 ou -1 significar que o consumo está muito acima do normal ou muito a baixo respetivamente.

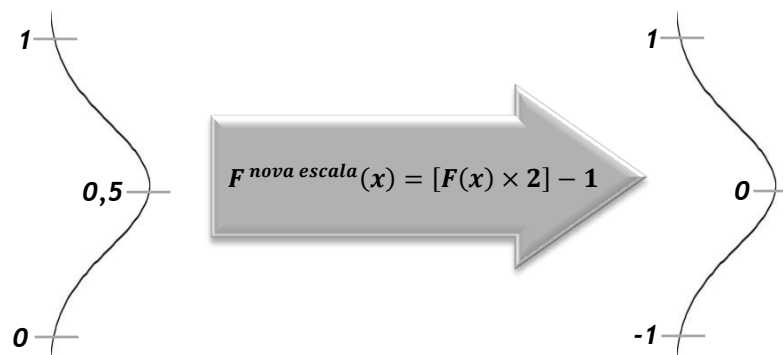


Figura 3.25 - Modelo para a mudança de escala do indicador 2

➤ Limites:

É neste ponto que são definidos os limites que caracterizam o indicador 2. São eles que vão limitar o algoritmo, isto é quando o modelo imprimir valores exteriores aos limites fixados significa que o consumo esta com um comportamento fora do normal.

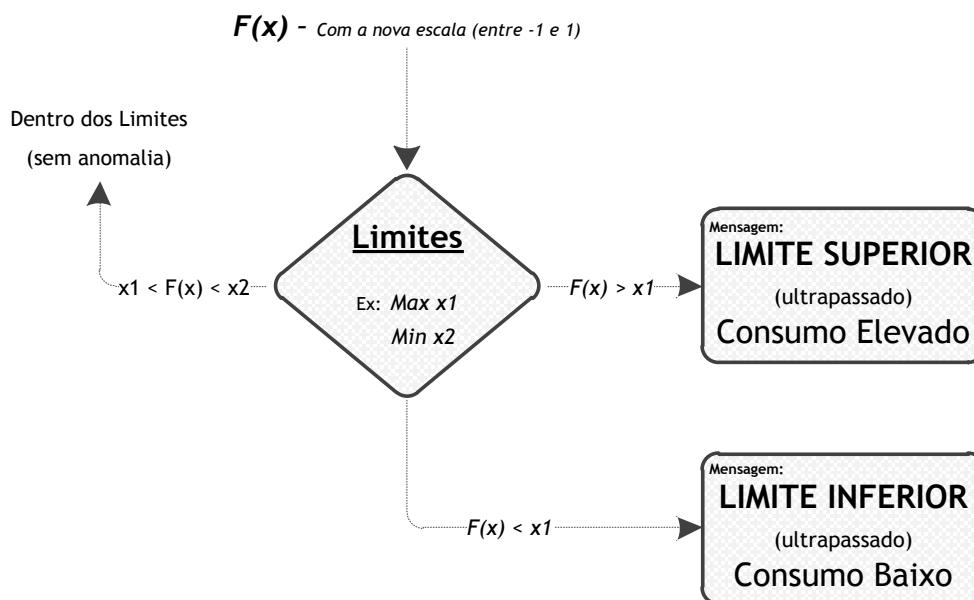


Figura 3.26 - Fluxograma com os limites do algoritmo de avaliação do indicador 2

Serve como exemplo a tabela 3.7 para demonstrar o algoritmo desenvolvido para o indicador 2. Na tabela estão presentes as cinco primeiras horas do dia 1 de Maio de 2011.

Hora	Dia da Semana	Data	Dados	Pertença	Cluster	Indicador 2	$\Delta$ indicador
0	1	01-05-2011 00:00	6,25	0,99987	1 - B	0,7597044	
1	1	01-05-2011 01:00	6,25	0,99987	1 - B	0,7733338	0,006814671
2	1	01-05-2011 02:00	6,25	0,99987	1 - B	0,7800478	0,003357008
3	1	01-05-2011 03:00	6,125	0,99988	1 - B	0,7699025	0,005072675
4	1	01-05-2011 04:00	6,25	0,99987	1 - B	0,7818966	0,005997092
5	1	01-05-2011 05:00	6,125	0,99986	1 - B	0,7559622	0,01296722

Tabela 3.7 - Demonstração do indicador 2

# Capítulo 4

## Simulação para as séries estudadas

### 4.1 - Introdução

Neste capítulo é demonstrado o processo de simulação do modelo desenvolvido e descrito no capítulo anterior.

No desenrolar deste capítulo serão abordados os diferentes passos na simulação da metodologia desenvolvida de modo a validar o modelo criado. A possível integração deste tipo de metodologia em sistemas de gestão de energia leva a que o modelo criado tenha de possuir uma elevada fiabilidade, assim a sua validação é de extrema importância.

A aplicação desta metodologia depende apenas dos consumos do passado, que podem ser adquiridos com equipamentos de monitorização presentes nas instalações.

### 4.1.1 - Abordagem

A abordagem seguida na simulação do modelo desenvolvido e posterior operação do gestor de energia foi representada de forma esquemática no seguinte diagrama:

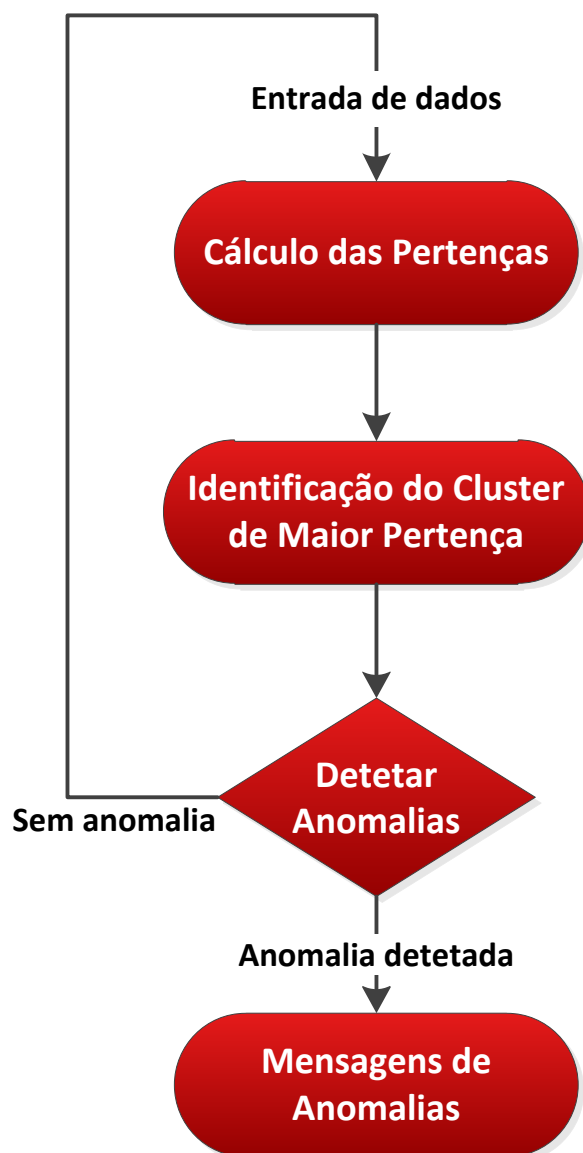


Figura 4.1 - Fluxograma da abordagem seguida na simulação do modelo

## 4.2 - Dados

Para a validação do modelo desenvolvido foi utilizada a série de consumo da empresa Savinor referente ao intervalo de tempo entre Maio de 2011 e Janeiro de 2012. A utilização da série de consumo da Savinor deve-se a esta ser a mais completa em comparação com outras séries disponíveis.

Relativamente á Savinor a área de negócio é:

- Criação, abate e preparação de carcaças de aves, obtenção de carnes de aves, obtenção de carnes diversas em peças ou pedaços (por corte e desmancha) e de miudezas comestíveis, refrigerados ou congelados, embalados ou não, e respetiva comercialização. Inclui também a transformação de subprodutos de peixes e de subprodutos resultantes do abate e desmancha de aves e de mamíferos, visando a produção de farinhas e a fusão de gorduras animais para fins industriais e respetiva comercialização.

Toda a informação utilizada foi disponibilizada pela empresa *Smartwatt*.

## 4.3 - Determinação das Pertenças

O primeiro cálculo que é realizado pelo algoritmo é o cálculo das pertenças. Deste modo, ao longo que se dá a entrada dos dados por cada hora o programa vai calcular automaticamente a pertença aos diferentes *clusters*. Para tal vai ser necessário calcular as distâncias entre o ponto de *input* e os diferentes clusters, isto requer a tabela guardada na base de dados com os centros dos clusters (tabela 3.3). Esta distância é determinada sobre a forma de:

$$d_{pc} = \sqrt{x_i - Centro_{cluster1}} \quad (6)$$

O dado que o algoritmo processa em seguida é a determinação do *cluster* de maior pertença, é a partir deste resultado que começam a ser possíveis identificar alguns desvios ou anomalias no consumo.

Na **tabela 4.1** está representada a simulação do modelo que permite calcular o *cluster* de maior pertença.

Data	Upc1	Upc2	Upc3	Upc4	Upc5	Max	C1	C2	C3	C4	C5	Cluster
01-05-2011 00:00	0,999871	0,000035	0,000020	0,000039	0,000035	0,999871	1					1
01-05-2011 01:00	0,9998713	3,49433E-05	2,55801E-05	3,57319E-05	3,24714E-05	0,9998713	1					1

Tabela 4.1 - Cluster de maior pertença

Na tabela estão representadas as duas primeiras horas do dia 1 de Maio de 2011, com as respetivas pertenças a cada *cluster* ( $U_{pc1}$ ,  $U_{pc2}$ ,  $U_{pc3}$ ,  $U_{pc4}$ ,  $U_{pc4}$ ). Em ambos os casos o *cluster* de maior pertença é o *cluster 1* como se pode verificar no resultado do modelo desenvolvido e ilustrado na tabela 4.1.

#### 4.4 - Detecção de Anomalias

Com o *cluster* de maior pertença identificado o primeiro indicador já pode entrar em funcionamento. Para tal, ele terá de recorrer à tabela 3.4 de modo a comparar o dia da semana em questão com o dia da semana correspondente ao perfil do *cluster* identificado. Com isto detém a informação necessária para identificar se para o dia e hora em análise o consumo apresenta algum tipo de desvio. Caso seja detetada uma anomalia o indicador imprime uma pequena descrição com o tipo de discrepância identificada. Estas descrições foram previamente definidas de acordo com a análise feita á empresa e serão exemplificadas no ponto seguinte (4.5).

hora	dia semana	Data	Dados (kw)	Pertença	Cluster			diferença pertença	Indicador 1	Anómalia
22	3	03-05-2011 22:00	44,625	0,58843	2	0,5920	2	-0,0036	0	
23	3	03-05-2011 23:00	50,875	0,59202	2	0,5920	2	0,0000	0	
0	4	04-05-2011 00:00	131	0,57644	5	0,3432	4	0,2332	1	Anómalia
1	4	04-05-2011 01:00	183	0,56017	5	0,3432	4	0,2169	1	Anómalia
2	4	04-05-2011 02:00	205	0,35726	5	0,3432	4	0,0140	1	Anómalia
3	4	04-05-2011 03:00	199,875	0,32896	5	0,3432	4	-0,0143	1	Anómalia
4	4	04-05-2011 04:00	205	0,31681	5	0,3432	4	-0,0264	1	Anómalia
5	4	04-05-2011 05:00	198,625	0,31754	5	0,3432	4	-0,0257	1	Anómalia
6	4	04-05-2011 06:00	206,875	0,32551	5	0,3432	4	-0,0177	1	Anómalia
7	4	04-05-2011 07:00	198,375	0,34122	5	0,3432	4	-0,0020	1	Anómalia
8	4	04-05-2011 08:00	203,75	0,34823	5	0,3432	4	0,0050	1	Anómalia
9	4	04-05-2011 09:00	203,875	0,34954	4	0,3432	4	0,0063	0	
10	4	04-05-2011 10:00	205,625	0,40353	4	0,3432	4	0,0603	0	
11	4	04-05-2011 11:00	208,625	0,46155	4	0,3432	4	0,1183	0	
12	4	04-05-2011 12:00	182,375	0,50860	4	0,3432	4	0,1654	0	
13	4	04-05-2011 13:00	172,25	0,52821	4	0,3432	4	0,1850	0	
14	4	04-05-2011 14:00	169,75	0,47940	4	0,3432	4	0,1362	0	
15	4	04-05-2011 15:00	182,25	0,44398	5	0,3432	4	0,1008	1	Anómalia
16	4	04-05-2011 16:00	187	0,50895	5	0,3432	4	0,1657	1	Anómalia
17	4	04-05-2011 17:00	189,875	0,56098	5	0,3432	4	0,2178	1	Anómalia
18	4	04-05-2011 18:00	170,125	0,59807	5	0,3432	4	0,2548	1	Anómalia
19	4	04-05-2011 19:00	102,625	0,56186	5	0,3432	4	0,2186	1	Anómalia
20	4	04-05-2011 20:00	44	0,45244	5	0,3432	4	0,1092	1	Anómalia
21	4	04-05-2011 21:00	43,875	0,37505	5	0,3432	4	0,0318	1	Anómalia
22	4	04-05-2011 22:00	43,5	0,32644	4	0,3432	4	-0,0168	0	
23	4	04-05-2011 23:00	55,125	0,34323	4	0,3432	4	0,0000	0	
0	5	05-05-2011 00:00	137,75	0,43185	5	0,6376	4	-0,2057	1	Anómalia
1	5	05-05-2011 01:00	202,375	0,33621	5	0,6376	4	-0,3013	1	Anómalia
2	5	05-05-2011 02:00	202,25	0,31230	5	0,6376	4	-0,3253	1	Anómalia
3	5	05-05-2011 03:00	202,125	0,29807	5	0,6376	4	-0,3395	1	Anómalia
4	5	05-05-2011 04:00	201,125	0,29778	5	0,6376	4	-0,3398	1	Anómalia
5	5	05-05-2011 05:00	197	0,30404	5	0,6376	4	-0,3335	1	Anómalia
6	5	05-05-2011 06:00	196	0,32252	5	0,6376	4	-0,3150	1	Anómalia
7	5	05-05-2011 07:00	200	0,34593	5	0,6376	4	-0,2916	1	Anómalia
8	5	05-05-2011 08:00	202,375	0,35762	5	0,6376	4	-0,2799	1	Anómalia
9	5	05-05-2011 09:00	198,5	0,36102	5	0,6376	4	-0,2765	1	Anómalia

Figura 4.2 - Simulação do indicador 1

Relativamente ao segundo indicador este também tem todas as informações necessárias guardadas na base de dados. O conhecimento que ele necessita oriundo da base de dados é a tabela 3.6 que possui os dados das médias e desvios padrões dos diferentes clusters. Logo através dessa tabela em conjunto com o valor de entrada e com o *cluster* de maior pertença identificado o modelo já possui as condições essenciais para iniciar o cálculo do indicador 2. Em conjunto com o indicador é editada uma mensagem que traduz o indicador por palavras,

isto é o algoritmo vai interpretar o indicador e imprimir uma pequena descrição do comportamento do consumo, caso este esteja em torno dos limites do indicador 2 previamente fixados. Esta mensagem descritiva será abordada no ponto seguinte.

hora	dia semana	Data	Dados (kw)	Pertença	Cluster		diferença pertença		Anómalia	Indicador 2	Varição indicador 2
16	2	09-05-2011 16:00	10,5	0,99802	1 - B	0,9576	1 - B	0,0404	Anómalia	0,998369402	0,009336028
17	2	09-05-2011 17:00	45,625	0,99070	1 - B	0,9576	1 - B	0,0331	Anómalia	1	0,000815299
18	2	09-05-2011 18:00	44	0,98445	1 - B	0,9576	1 - B	0,0268	Anómalia	1	0
19	2	09-05-2011 19:00	44,375	0,97807	1 - B	0,9576	1 - B	0,0205	Anómalia	1	0
20	2	09-05-2011 20:00	44	0,97240	1 - B	0,9576	1 - B	0,0148	Anómalia	1	0
21	2	09-05-2011 21:00	44,375	0,96661	1 - B	0,9576	1 - B	0,0090	Anómalia	1	0
22	2	09-05-2011 22:00	44,75	0,96095	1 - B	0,9576	1 - B	0,0033	Anómalia	1	0
23	2	09-05-2011 23:00	44,125	0,95761	1 - B	0,9576	1 - B	0,0000	Anómalia	1	0
0	3	10-05-2011 00:00	66,75	0,27723	4	0,4962	2	-0,2190	Anómalia	-0,990911997	0,995455999
1	3	10-05-2011 01:00	94,25	0,25453	4	0,4962	2	-0,2417	Anómalia	-0,99872157	0,003904786
2	3	10-05-2011 02:00	126,5	0,26625	4	0,4962	2	-0,2300	Anómalia	-0,999968937	0,000623683
3	3	10-05-2011 03:00	168,25	0,28085	4	0,4962	2	-0,2154	Anómalia	-0,789071045	0,105448946
4	3	10-05-2011 04:00	172,25	0,28566	4	0,4962	2	-0,2106	Anómalia	-0,798395291	0,004662123
5	3	10-05-2011 05:00	169,5	0,28645	4	0,4962	2	-0,2098	Anómalia	-0,943050219	0,072327464
6	3	10-05-2011 06:00	170,25	0,28781	4	0,4962	2	-0,2084	Anómalia	-0,878449846	0,032300186
7	3	10-05-2011 07:00	193,75	0,29529	4	0,4962	2	-0,2009	Anómalia	0,682707829	0,780578838
8	3	10-05-2011 08:00	202,125	0,31150	4	0,4962	2	-0,1847	Anómalia	0,933484523	0,125388347
9	3	10-05-2011 09:00	197,875	0,33196	4	0,4962	2	-0,1643	Anómalia	0,767494183	0,08299517
10	3	10-05-2011 10:00	184,125	0,35348	4	0,4962	2	-0,1427	Anómalia	0,036265031	0,365614576
11	3	10-05-2011 11:00	136,75	0,34263	4	0,4962	2	-0,1536	Anómalia	-0,985324924	0,510794978
12	3	10-05-2011 12:00	99,125	0,31062	4	0,4962	2	-0,1856	Anómalia	-0,995071105	0,00487309
13	3	10-05-2011 13:00	45,875	0,35893	2	0,4962	2	-0,1373		0,156794439	0,575932772
14	3	10-05-2011 14:00	46,375	0,38177	2	0,4962	2	-0,1144		0,592617431	0,217911496
15	3	10-05-2011 15:00	46,625	0,40019	2	0,4962	2	-0,0960		0,97649078	0,191936674
16	3	10-05-2011 16:00	45,625	0,42197	2	0,4962	2	-0,0742		0,95054985	0,012970465
17	3	10-05-2011 17:00	45,125	0,44244	2	0,4962	2	-0,0538		0,935931615	0,007309118
18	3	10-05-2011 18:00	44,5	0,45943	2	0,4962	2	-0,0368		0,884839223	0,025546196
19	3	10-05-2011 19:00	44	0,47172	2	0,4962	2	-0,0245		0,847318517	0,018760353
20	3	10-05-2011 20:00	43,375	0,48058	2	0,4962	2	-0,0156		0,110125978	0,36859627
21	3	10-05-2011 21:00	43,5	0,48732	2	0,4962	2	-0,0089		0,080505158	0,01481041
22	3	10-05-2011 22:00	44,25	0,49260	2	0,4962	2	-0,0036		-0,003308527	0,041906842
23	3	10-05-2011 23:00	48,25	0,49621	2	0,4962	2	0,0000		-0,580005305	0,288348389
0	4	11-05-2011 00:00	126,375	0,78813	2	0,4930	4	0,2952	Anómalia	0,20828652	0,394145913
1	4	11-05-2011 01:00	186	0,50328	5	0,4930	4	0,0103	Anómalia	0,743718669	0,267716074
2	4	11-05-2011 02:00	192,625	0,43503	5	0,4930	4	-0,0579	Anómalia	0,804470976	0,030376154
3	4	11-05-2011 03:00	196,875	0,38012	5	0,4930	4	-0,1129	Anómalia	0,870573262	0,033051143

Figura 4.3 - Simulação do Indicador 2

## 4.5 - Mensagens de anomalias

Como foi exposto anteriormente caso o programa detete algum desvio ou anomalia no consumo este imprime uma mensagem discriminando o tipo de erro reconhecido. Estas mensagens têm o principal objetivo de simplificar a interpretação dos indicadores, pois todas as informações expostas são impressas de acordo com o resultado obtido pelo indicador. Vão transformar as saídas numéricas que os indicadores transmitem em frases curtas e descritivas desse valor.

No caso do primeiro indicador, este ao detetar uma possível anomalia vai fazer com que o processo de criação das mensagens analise o porquê da existência dessa anomalia. Tal análise vai ser baseada nas diferentes comparações que são efetuadas anteriormente pelo indicador 1 para concluir a existência da anomalia. Uma das análises em que se vai fundar é a questão da comparação entre o dia e hora em análise e o perfil do *cluster* de maior pertença identificado. Porque se os dias não coincidirem significa que o dia em estudo apresenta um comportamento do consumo que se assemelha a um outro dia diferente, no caso particular de coincidir a um domingo pode-se estar presente de um feriado ou ponte. Pois ao domingo é o dia em que a empresa não apresenta consumo energético.

Foram também identificados dias onde o consumo foi nulo podendo nestes casos existir diversas conclusões. Estes dias podem ser dias de manutenção mas não deixando de parte a hipótese de uma possível anomalia no sistema de transmissão de dados.

Na figura seguinte identifica-se algumas das mensagens de erro referentes ao indicador 1 que foram apresentadas do decorrer da simulação do modelo.

Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	0,97631311
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	0,954942067
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	0,9428004
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	0,933770845
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	0,961043806
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	1
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	1
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	1
Anómalia	É <u>Segunda</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	1
Anómalia	É <u>Terça</u> e foi identificado um cluster referente ao dia	<u>Domingo</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Domingo</u>	→ Poderá ser feriado/Ponte	1
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,988694857
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,996672112
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,540642487
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,397268318
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,32777528
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,882823901
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,727573541
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		0,570454363
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,131488396
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,030744199
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,979254809
Anómalia	É <u>Terça</u> e neste momento foi detectada uma anomalia		→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Quinta</u>		-0,197687224
						0,016156399
						0,42516844
						0,890745955
						0,752531286
						0,900308737
						0,895475685
						0,735402133
						0,173592596
						0,148273651
						0,015879925
						-0,47797828
Anómalia	É <u>Quarta</u> e foi identificado um cluster referente ao dia	<u>Sexta</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Sexta</u>		0,123627659
Anómalia	É <u>Quarta</u> e foi identificado um cluster referente ao dia	<u>Sexta</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Sexta</u>		0,614905855
Anómalia	É <u>Quarta</u> e foi identificado um cluster referente ao dia	<u>Sexta</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Sexta</u>		0,990671592
Anómalia	É <u>Quarta</u> e foi identificado um cluster referente ao dia	<u>Sexta</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Sexta</u>		0,929948271
Anómalia	É <u>Quarta</u> e foi identificado um cluster referente ao dia	<u>Sexta</u>	→ Nesta hora o consumo esta a comportar-se como um(a)	<u>Sexta</u>		0,945005569

Figura 4.4 - Mensagens referentes ao indicador 1

No que diz respeito ao segundo indicador, este também imprime uma pequena mensagem que vai traduzir o valor calculado. Mesmo alterando a escala do indicador para uma escala entre -1 e +1 que facilita a interpretação, foi desenvolvido um algoritmo que acrescente ao indicar uma pequena mensagem descritiva.

A mensagem editada vai ser impressa de acordo com os limites fixados na parte do desenvolvimento do modelo, a mensagem vai ser apresentada em conformidade com o valor determinado pelo indicador 2.

Por exemplo, ao ser identificado o *cluster* de maior pertença o indicador vai analisar o consumo em comparação com o perfil do cluster identificado e imprime um valor próximo de 1. Neste caso será apresentada uma mensagem indicando que o consumo está no LIMITE SUPERIOR, isto é o consumo está acima do normal.

Para demonstrar as mensagens correspondentes ao segundo indicador está representada a seguinte figura que diz respeito ao momento de simulação do modelo.

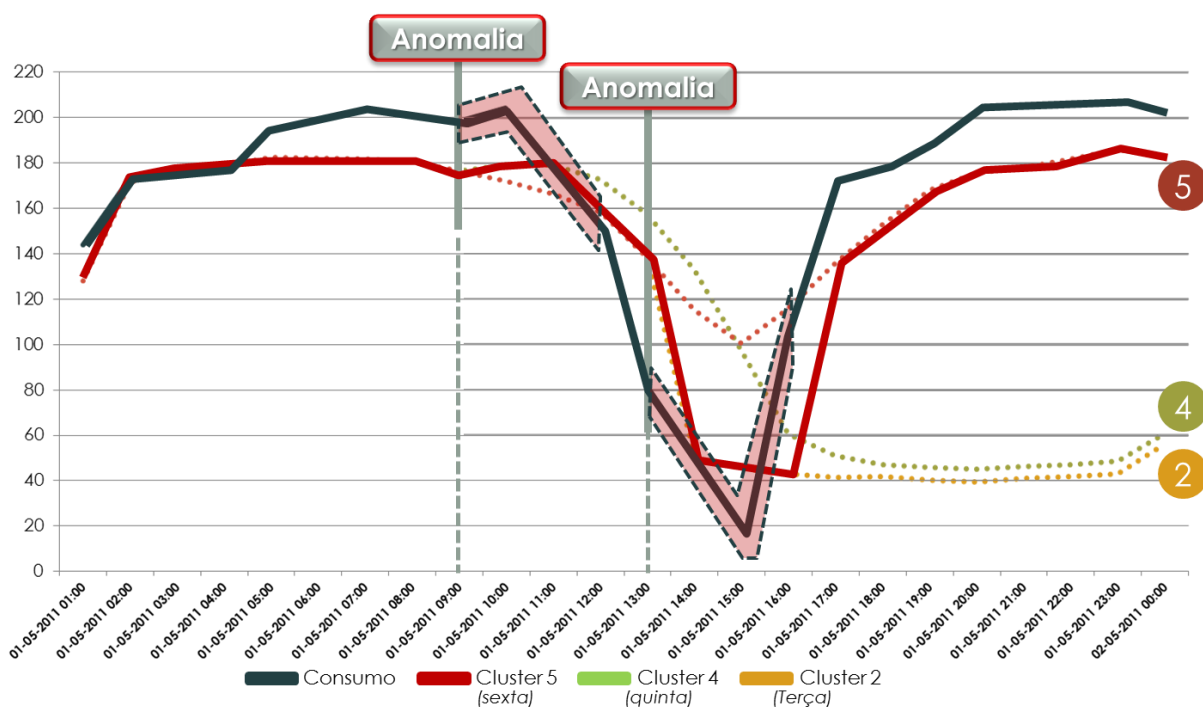
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,728861958	0,124722707	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,830575169	0,050856606	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,904982277	0,037203554	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,96433884	0,029678282	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,999690893	0,017676026	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,999999896	0,000154501	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	1	5,21294E-08	LIMITE SUPERIOR -> CONSUMO ELEVADO
		-0,974986311	0,987493156	LIMITE INFERIOR -> CONSUMO MÍNIMO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0,987493156	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Domingo</u>	→ Poderá ser feriado/Ponte	1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Quinta</u>		-0,610629474	0,805314737	
<u>Quinta</u>		-0,988968283	0,189169404	LIMITE INFERIOR -> CONSUMO MÍNIMO
<u>Quinta</u>		-0,0410645	0,473951891	
<u>Quinta</u>		0,964166087	0,502615294	LIMITE SUPERIOR -> CONSUMO ELEVADO
		0,644423107	0,15987149	
		0,769587379	0,062582136	
		0,799813538	0,015113079	
		0,76733316	0,016240189	
		-0,284310119	0,52582164	
		-0,099535919	0,0923871	
		0,114496463	0,107016191	
		0,999999756	0,442751646	LIMITE SUPERIOR -> CONSUMO ELEVADO
		1	1,21938E-07	LIMITE SUPERIOR -> CONSUMO ELEVADO
		1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
		1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
		1	0	LIMITE SUPERIOR -> CONSUMO ELEVADO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	-0,924113011	0,962056505	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	-0,999490805	0,037688897	LIMITE INFERIOR -> CONSUMO MÍNIMO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	-0,999996568	0,000252881	LIMITE INFERIOR -> CONSUMO MÍNIMO
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	-0,357010191	0,321493188	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,604423457	0,480716824	
<u>Sábado</u>	→ Poderá ser uma véspera de feriado	0,726243174	0,060909858	
		0,894967702	0,084362264	

Figura 4.5 - Mensagens referentes ao indicador 2

Todas estas mensagens após a simulação foram analisadas de forma a entender se estavam a interpretar bem os indicadores e do mesmo modo perceber se estes identificavam convenientemente as anomalias.

Mais uma vez este ponto de simulação das mensagens de erro é de extrema importância para a veracidade do modelo, pois é através delas que o algoritmo desenvolvido vai transmitir os resultados obtidos, isto é as diferentes interpretações relativas aos desvios do consumo. Todas as mensagens estão guardadas na base de dados, e de acordo com a análise dos indicadores serão impressas de modo a caracteriza-lo da melhor forma como foi demonstrado nas duas figuras anteriores (4.4 e 4.5).

Em suma, a figura 4.6 representa a título de exemplo um possível *layout* de saída de todo o algoritmo criado. Identifica-se as áreas onde foram detetadas anomalias em conjunto com os diferentes *clusters* identificados.



## 4.6 - Conclusões

Um dos aspetos importantes a concluir após a simulação do modelo é que este apresenta um elevado grau de flexibilidade, conseguindo assim ajustar-se às diferentes circunstâncias a que possa estar sujeito. A aplicação do modelo desenvolvido a casos reais permitiu também verificar a sua fiabilidade e a sua capacidade de integrar plataformas de gestão de energia. A análise feita ao modelo permitiu verificar que é um modelo viável e que pode ser aplicado com grande fiabilidade.

Relativamente aos resultados obtidos na simulação, provocam grande entusiasmo para a sua aplicação em sistemas de gestão de energia.

# Capítulo 5

## Validação e Conclusões

### 5.1 - Introdução

Neste capítulo é demonstrado o processo de validação do modelo desenvolvido. O programa contém uma função de validação do modelo que é executado ao fim das 24 horas de cada dia. Esta função de validação só é executada ao fim das 24 horas do dia devido aos dados de entrada que necessita.

Numa segunda parte é comparado o número de anomalias identificadas pelo modelo com o número de anomalias que efetivamente existiram. De modo a validar o modelo criado e suas competências.

Por fim, após ter obtido todos os dados de validação foram tiradas algumas conclusões e propostas de trabalhos futuros associadas ao processo de criação deste modelo.

## 5.2 - Indicador de Validação

Como foi descrito na introdução o programa tem um modelo de validação automático que é efetuado ao fim das 24 horas de cada dia. Esta função é um género de um indicador\_de\_validação que vai demonstrar o funcionamento do modelo ao fim das 24 horas de cada dia. Este indicador também está diretamente relacionado com o *cluster* de maior pertença como os indicadores anteriormente referidos (indicador 1 e indicador 2).

Como foi descrito na *secção 3.2.6 (Calculo das pertenças)* o cálculo da pertença depende do somatório das distâncias entre o consumo de entrada e o centro dos *clusters*.

$$U_{pc} = \frac{(1/d_{pc})^{2/(m-1)}}{\sum_{c=1}^k (1/d_{pc})^{2/(m-1)}} \quad (5)$$

$$d_{pc} = \sqrt{x_i - \text{Centro}_{\text{cluster}1}} \quad (6)$$

Só ao final das 24 horas do dia é que as distâncias para o cálculo estão todas determinadas, logo é nesse ponto que a pertença tem o máximo de dados reunidos. Dessa forma o cluster de maior pertença identificado no final do dia é o que efetivamente se aproximou mais do consumo analisado.

Com este indicador definido é possível no final de cada dia analisar as possíveis anomalias que não foram detetadas ao longo do dia. Este indicador pode ser inserido numa nova coluna (*cluster no final 24h*) com o cluster calculado às 24 horas, de modo a executar a comparação entre o cluster obtido ao longo do dia com o cluster calculado ao final das 24 horas. Para exemplificar este modelo de validação esta apresentada a seguinte figura:

hora	dia semana	Data	Dados	Pertença	Cluster	Cluster no final 24h	diferença pertença	Indicador Validação	Indicador 1	Anómalia	
0	3	03-05-2011 00:00	58,875	0,31069	1 - B	0,5920	2	-0,2813	1	1	Anómalia
1	3	03-05-2011 01:00	109	0,27243	4	0,5920	2	-0,3196	1	0	
2	3	03-05-2011 02:00	141,375	0,28676	4	0,5920	2	-0,3053	1	0	
3	3	03-05-2011 03:00	186,25	0,29703	4	0,5920	2	-0,2950	1	0	
4	3	03-05-2011 04:00	188,125	0,30094	4	0,5920	2	-0,2911	1	0	
5	3	03-05-2011 05:00	189,125	0,30389	4	0,5920	2	-0,2881	1	0	
6	3	03-05-2011 06:00	197,5	0,30715	4	0,5920	2	-0,2849	1	0	
7	3	03-05-2011 07:00	194,875	0,31272	4	0,5920	2	-0,2793	1	0	
8	3	03-05-2011 08:00	191,5	0,32578	4	0,5920	2	-0,2662	1	0	
9	3	03-05-2011 09:00	182,875	0,34067	4	0,5920	2	-0,2514	1	0	
10	3	03-05-2011 10:00	183	0,36562	4	0,5920	2	-0,2264	1	0	
11	3	03-05-2011 11:00	138,875	0,35372	4	0,5920	2	-0,2383	1	0	
12	3	03-05-2011 12:00	47,5	0,40761	2	0,5920	2	-0,1844	0	0	
13	3	03-05-2011 13:00	44,25	0,46051	2	0,5920	2	-0,1315	0	0	
14	3	03-05-2011 14:00	45,125	0,48131	2	0,5920	2	-0,1107	0	0	
15	3	03-05-2011 15:00	45	0,49908	2	0,5920	2	-0,0929	0	0	
16	3	03-05-2011 16:00	43,625	0,52024	2	0,5920	2	-0,0718	0	0	
17	3	03-05-2011 17:00	44,625	0,53959	2	0,5920	2	-0,0524	0	0	
18	3	03-05-2011 18:00	44,625	0,55574	2	0,5920	2	-0,0363	0	0	
19	3	03-05-2011 19:00	43,125	0,56792	2	0,5920	2	-0,0241	0	0	
20	3	03-05-2011 20:00	44,125	0,57651	2	0,5920	2	-0,0155	0	0	
21	3	03-05-2011 21:00	44,375	0,58312	2	0,5920	2	-0,0089	0	0	
22	3	03-05-2011 22:00	44,625	0,58843	2	0,5920	2	-0,0036	0	0	
23	3	03-05-2011 23:00	50,875	0,59202	2	0,5920	2	0,0000	0	0	

Figura 5.1 - Indicador\_de\_Validação

Na figura identifica-se que o modelo não conseguiu detetar anomalias ao longo do dia, embora através do indicador\_de\_validação possa-se observar que nas primeiras 11 horas o consumo teve um comportamento anormal.

### 5.3 - Validação

Após a simulação do modelo o passo seguinte e de extrema importância é determinar a autenticidade do modelo, para isso foi necessário comparar o número de anomalias detetadas com o número efetivo de anomalias.

Neste ponto a série de consumo utilizada foi estudada detalhadamente com o auxílio dos elementos da *smartwatt* que monitorizam o consumo da empresa, de modo a identificar todas as anomalias presentes na série e suas originas.

O processo de validação baseou-se na análise das anomalias detetadas ao longo da simulação, e posterior comparação com verdadeiras anomalias existentes. No decorrer da validação desta metodologia foram identificados momentos em que o modelo não conseguiu detetar anomalia, devido ao facto de o indicador\_de\_validação, tirar conclusões no final de cada dia.

Por forma a avaliar o modelo foi realizado o seguinte quadro onde-se pode reconhecer os diferentes casos identificados.

		<i>Anomalia real</i>	
		<b>Anomalia existente</b>	<b>Anomalia inexistente</b>
<i>Modelo deteta</i>	<b>Anomalia detetada</b>	2851 (43%)	528 (8%)
	<b>Anomalia não detetada</b>	65 (1%)	3180 (48%)

Tabela 5.1- Validação

Os resultados obtidos permitem validar o sistema porque 91% dos casos estudados foram corretamente categorizados (anomalias e não anomalias). O percentual de anomalias detetadas e inexistentes (8%) pode estar relacionado com o nível de sensibilidade utilizado nos indicadores do modelo. Desta forma considera-se que o modelo elaborado possui um elevado nível de fiabilidade.

### 5.4 - Conclusões e trabalhos futuros

A metodologia desenvolvida nesta dissertação cumpre os objetivos previamente fixados. O algoritmo criado tem a vantagem de operar em tempo real conseguindo tirar conclusões no imediato, ao contrário de outros sistemas já desenvolvidos que necessitam de toda a informação em torno da anomalia. Outro dos aspetos importantes do modelo é que este

apenas necessita da série de consumo histórico, pois é através desta que parametriza o modelo de forma aos indicadores terem a sensibilidade particularizada para a instalação em análise. Deste modo, o modelo tendo apenas a série de consumo histórico como entrada faz com que tenha um elevado grau de flexibilidade, conseguindo assim ajustar-se às diferentes circunstâncias a que possa estar sujeito.

O algoritmo de reconhecimento de padrões criado permite representar da melhor forma os diferentes padrões do consumo, de modo que os indicadores desenvolvidos consigam identificar as anomalias com grande fiabilidade como foi demonstrado na validação do modelo. Portanto analisando a metodologia desenvolvida verifica-se que o modelo é viável tendo grande potencialidade para integrar sistemas de gestão de energia, cumprindo assim um dos principais objetivos estipulados na elaboração desta dissertação.

Contudo no desenvolvimento do modelo identificou-se algumas melhorias que podiam ser realizadas de modo ao algoritmo funcionar sobre uma plataforma de software. Alguns desses trabalhos pensados para o futuro são:

- Automatizar o modelo de modo a diminuir os tempos de processo:

Ao testar o modelo identificou-se que nas fases iniciais a análise da série de consumo histórico, é bastante demorada. Com isto a metodologia foi apenas criada para uma só série de consumo. Na verdade era necessário otimizar o processo de modo ao modelo ser de rápida replicabilidade para os diferentes padrões de consumo.

- Desenvolver mecanismos automáticos para a análise qualitativa da série de consumo histórico:

A análise qualitativa da série é das fases mais demoradas no desenvolvimento do modelo, deste modo seria de grande importância automatizar este processo. Isto é desenvolver um mecanismo que realize automaticamente a análise de forma a eliminar os pontos anómalos existentes na série de consumo histórico inicial. Para tal este mecanismo deveria padronizar a série total do consumo histórico.

- Automatizar o cálculo dos *clusters*:

O cálculo dos *clusters* é de extrema complexidade, deste modo o desenvolvimento de uma ferramenta que realiza-se o cálculo automático dos *clusters* iria aumentar a agilidade da metodologia criada. Desta forma a ferramenta apenas necessitava da série de consumo histórico processada e automaticamente executava o algoritmo de reconhecimento de padrões, imprimindo a tabela referente aos centros dos *clusters*.

- Plataforma de *software*:

Todas as propostas referidas têm como objetivo automatizar a metodologia criada de modo a esta ser desenvolvida numa plataforma de *software*. Deste modo a metodologia pode ser inserida em sistemas de gestão de energia, onde o gestor apenas terá de abrir o *software*

desenvolvido e inserir os dados necessários para começar a executar o programa. Com isto o gestor de energia poderá analisar o consumo energético em tempo real e intervir de forma fundamentada sobre as diferentes anomalias que sejam detetadas pelos indicadores presentes no modelo.

Em suma, a associação desta metodologia desenvolvida aos novos projetos associados à eficiência energética, é essencial para o aproveitamento de todas as potencialidades da energia. É no entanto necessário a criação de modelos fiáveis e que sejam mais-valias para a otimização do sistema de gestão de energia.



# Referências

1. Vine, E., *An international survey of the energy service company (ESCO) industry*. Energy Policy, 2005. 33(5): p. 691-704.
2. EWEN. *Eficiencia energética na industria*. Available from: <http://www.ewen-energy.com/index.php?id=19>.
3. Azevedo, R. and C. Monteiro, *Desenvolvimento de um Sistema de Ajuda à Negociação de Contratos de Performance para Medidas de Eficiência Energética*. 2008.
4. Irrek, W. and S. Thomas, *Der EnergieSparFonds für Deutschland*, ed. e.d.H.B. Stiftung2006.
5. Bunse, K., et al., *Integrating energy efficiency performance in production management - gap analysis between industrial needs and scientific literature*. Journal of Cleaner Production, 2011. 19(6-7): p. 667-679.
6. IEA, *Feedstock Substitutes, Energy Efficient Technology and CO2 Reduction for Petrochemical Products*. 2007a.
7. UN. *Kyoto Protocol to the United Nations Framework Convention on Climate Change*. 1998; Available from: <http://unfccc.int/2860.php>.
8. UNFCCC. *Report of the Conference of the Parties on Its Fifteenth Session, Held in Copenhagen from 7 to 19 December 2009 at Its Fifteenth Session. Part Two: Action Taken by the Conference of the Parties*. 2010; Available from: <http://unfccc.int/2860.php>.
9. Jovane, F., et al., *The incoming global technological and industrial revolution towards competitive sustainable manufacturing*. CIRP Annals - Manufacturing Technology, 2008. 57(2): p. 641-659.
10. EC. *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Analysis of Options to Move beyond 20% Greenhouse Gas Emission Reductions and Assessing the Risk of Carbon Leakage*. 2012; Available from: <http://ec.europa.eu/>.
11. Thollander, P. and M. Ottosson, *Energy management practices in Swedish energy-intensive industries*. Journal of Cleaner Production, 2010. 18(12): p. 1125-1133.
12. Ang, B.W., *Monitoring changes in economy-wide energy efficiency: From energy-GDP ratio to composite efficiency index*. Energy Policy, 2006. 34(5): p. 574-582.
13. Phylipsen, G.J.M., K. Blok, and E. Worrell, *International comparisons of energy efficiency-Methodologies for the manufacturing industry*. Energy Policy, 1997. 25(7-9): p. 715-725.
14. Patterson, M.G., *What is energy efficiency?: Concepts, indicators and methodological issues*. Energy Policy, 1996. 24(5): p. 377-390.
15. Farla, J., K. Blok, and L. Schipper, *Energy efficiency developments in the pulp and paper industry: A cross-country comparison using physical production data*. Energy Policy, 1997. 25(7-9): p. 745-758.
16. Boyd, G., E. Dutrow, and W. Tunnessen, *The evolution of the ENERGY STAR® energy performance indicator for benchmarking industrial plant manufacturing energy use*. Journal of Cleaner Production, 2008. 16(6): p. 709-715.
17. Kannan, R. and W. Boie, *Energy management practices in SME--case study of a bakery in Germany*. Energy Conversion and Management, 2003. 44(6): p. 945-959.

18. Thollander, P., M. Danestig, and P. Rohdin, *Energy policies for increased industrial energy efficiency: Evaluation of a local energy programme for manufacturing SMEs*. Energy Policy, 2007. **35**(11): p. 5774-5783.
19. Anderson, S.T. and R.G. Newell, *Information programs for technology adoption: the case of energy-efficiency audits*. Resource and Energy Economics, 2004. **26**(1): p. 27-50.
20. Vesma, V., *Energy Management Principles and Practice - A Companion to BS EN 16001:2009*. British Standards Institution, 2009.
21. QEnergia. *Energy Brain - software de contagem de energia*. Available from: [www.qenergia.pt](http://www.qenergia.pt).
22. QEnergia, *Especificações de sistema de gestão de consumos Electrex - Solução EnergyBrain*, 2011.
23. QEnergia, *Apresentação - QEnergia, Sistemas para Qualidade e Gestão de Energia, Lda.*, 2011.
24. Circutor. *Power Studio - Energy monitoring and control software*. Available from: [www.circutor.com](http://www.circutor.com).
25. Circutor, *Energy management software*, 2011.
26. Halkidi, M., Y. Batistakis, and M. Vazirgiannis, *On Clustering Validation Techniques*. Journal of Intelligent Information Systems, 2001. **17**(2): p. 107-145.
27. ABESCO. *ESCO*. Available from: <http://www.abesco.com.br/datarobot/sistema/paginas/pagebody2.asp?id=22&msecu ndario=42>.
28. Monteiro, C., *Caracterização de consumos (Reconhecimento de padrões)*, 2011.
29. Canno, R.L., J.V. Dave, and J.C. Bezdek, *Efficient Implementation of the Fuzzy c-Means Clusteng Algor nthms*. 1986.
30. Vargas, R.R.d. and B.R.C. Bedregal, *Uma nova forma de calcular o centro dos clusters no algoritmo fuzzy c-means*. Publicação da SBMAC, 2010.
31. Pal, N.R. and J.C. Bezdek, *On cluster validity for the fuzzy c-means model*. IEE transactions on fuzzy systems, 1995. **vol. 3, No. 3**.