
UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E
COMPUTAÇÃO CIENTÍFICA

**MODELOS DE PREDIÇÃO UTILIZANDO LÓGICA FUZZY:
UMA ABORDAGEM INSPIRADA NA INFERÊNCIA BAYESIANA¹**

FELIPO BACANI

Banca Examinadora:

Laécio Carvalho de Barros (orientador)

Jurandir Zullo Junior

Fernando Antonio Campos Gomide

Peter Sussner

Laércio Luis Vendite

CAMPINAS
2012

¹Trabalho realizado com o auxílio financeiro da CAPES.

**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E COMPUTAÇÃO
CIENTÍFICA**

FELIPO BACANI

**MODELOS DE PREDIÇÃO UTILIZANDO LÓGICA FUZZY:
UMA ABORDAGEM INSPIRADA NA INFERÊNCIA
BAYESIANA**

**DISSERTAÇÃO DE MESTRADO
APRESENTADA AO INSTITUTO DE MATEMÁTICA,
ESTATÍSTICA E COMPUTAÇÃO CIENTÍFICA
DA UNICAMP PARA OBTENÇÃO DO TÍTULO DE
MESTRE EM MATEMÁTICA APLICADA**

ORIENTADOR: LAÉCIO CARVALHO DE BARROS

**ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELO ALUNO FELIPO BACANI, E ORIENTADA PELO PROF. DR. LAÉCIO CARVALHO
DE BARROS**



Assinatura do Orientador

CAMPINAS, 2012

FICHA CATALOGRÁFICA ELABORADA POR
ANA REGINA MACHADO - CRB8/5467
BIBLIOTECA DO INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E
COMPUTAÇÃO CIENTÍFICA - UNICAMP

B12m Bacani, Felipo, 1985-
Modelos de predição utilizando lógica fuzzy : uma abordagem
inspirada na inferência bayesiana / Felipo Bacani. – Campinas, SP :
[s.n.], 2012.

Orientador: Laécio Carvalho de Barros.
Dissertação (mestrado) – Universidade Estadual de Campinas,
Instituto de Matemática, Estatística e Computação Científica.

1. Conjuntos fuzzy. 2. Teoria bayesiana de decisão
estatística. 3. Biomatemática. I. Barros, Laécio Carvalho de,
1954-. II. Universidade Estadual de Campinas. Instituto de
Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em inglês: Prediction models using fuzzy logic : an approach inspired in
the bayesian inference

Palavras-chave em inglês:

Fuzzy sets

Bayesian theory of statistical decision

Biomathematics

Área de concentração: Matemática Aplicada

Titulação: Mestre em Matemática Aplicada

Banca examinadora:

Laécio Carvalho de Barros [Orientador]

Jurandir Zullo Junior

Fernando Antônio Campos Gomide

Data de defesa: 29-05-2012

Programa de Pós-Graduação: Matemática Aplicada

Dissertação de Mestrado defendida em 29 de maio de 2012 e aprovada

Pela Banca Examinadora composta pelos Profs. Drs.



Prof.(a). Dr(a). LAÉCIO CARVALHO DE BARROS



Prof.(a). Dr(a). JURANDIR ZULLO JUNIOR



Prof.(a). Dr(a). FERNANDO ANTONIO CAMPOS GOMIDE

Dedico este trabalho a meu pai, cuja mãe
amiga estará sempre comigo através de seus
ensinamentos.

*“Knowing where the trap is - that’s the first
step in evading it. This is like single combat,
son, only on a larger scale - a feint within a
feint within a feint... seemingly without end.”*
(Frank Herbert - Dune)

Agradecimentos

Em primeiro lugar, gostaria de agradecer à minha família: meus pais Ruth e Mauro, minhas irmãs Rebeca e Nathália. Alessandra e família pelo apoio. Ângelo pela amizade.

A meus amigos de curso: Cajuru, Gaúcho, Modelo, Veronez, Sabão, Jaul (e Parlamento), Barrozo, Marcelinho, Sputnik, João baiano, Campello e Francismar. Aos amigos da faculdade, em especial Alê Martins, Julião, Werner, Fabi Anjos, Renan, Gustavo e Bia, Murilo, pessoal da Casa do Mato. Aos amigos recentes da pós-graduação: Varti Bart, Adson Edson, Lino Cabrobó, Marcus Pijaminha, Michael do pagode.

Ao pessoal da moradia, Julião [2], Matheuzinho, Vinícius, Cris, Du Virgílio, Ezequiel, Bigato, DVD, Ari e todos os que me receberam (e aturaram) na moradia da Unicamp durante minha turbulenta graduação. À moradia da Unicamp, sua história e memória, e a todos que lutam para preservar a assistência estudantil. À república CG pela recepção e irmandade: Jão Sorto, Viking, Tomé, Satan, Tião. A meus amigos de Jundiaí, Anas, Pedrão e Fer, que fizeram parte de um período muito importante de minha vida. Aos irmãos do som, Modelo [2], Gaúcho [2], Julião [3], Werner [2], Alê Martins [2], Fabi Anjos [2], Marcelinho [2], Bigato [2]. Ao coral Zíper na Boca da Unicamp pelo aprendizado, vivência e amizade.

Ao meu professor orientador Laécio, pela oportunidade de realizar um trabalho envolvendo idéias novas, mesmo consciente dos riscos envolvidos. Ao pessoal da secretaria da pós (Tânia, Lívia e Ednaldo) pela ajuda e paciência. À Dona Zefa e ao café no bandejão (necessariamente nessa ordem). Ao pessoal da limpeza e funcionários do IMECC. Ao Marcão e sua cantina (não esquecemos!). Aos professores da banca Fernando Gomide e Jurandir, pelas orientações. Ao professor Jurandir e à sua aluna Priscila, por terem gentilmente cedido os dados da aplicação deste trabalho. A Wander-son, pela amizade e por ter feito o contato que possibilitou a parceria com o professor Jurandir. Aos professores do IMECC, em especial algumas pessoas que me transpareceram humanidade através da seriedade de seu trabalho: professores San Martin, Ruffino, Joni, Márcia, Amélia, Margarida, Aurélio, Pulino, Yuri, Cheti, Adriano, André Assis (IFGW).

A Daniela, pelo companheirismo e suporte neste difícil fim de ciclo. Ao bandejão e a todas as pessoas que através dele (não é mentira!). A todos que fizeram parte deste processo de formação. A todos que sentem que deveriam estar citados aqui e não estão. A estes minhas desculpas.

Resumo

O presente trabalho tem por objetivo aplicar a teoria de conjuntos fuzzy a modelos de predição (inferência) de dados. O modelo utilizado baseia-se fortemente nas relações fuzzy em espaços contínuos (caso não matricial) e na regra de inferência modus ponens, utilizando t-normas (que neste contexto são similares à operação de cópula em estatística). É do modus ponens que surge o caráter “condicional” de alguns dos termos envolvidos, e a partir daí é que a analogia com a inferência bayesiana é feita. Entretanto, são apenas analogias conceituais: o presente trabalho não lida com nenhuma distribuição de probabilidades. Na verdade, conjuntos fuzzy são tratados como distribuições de possibilidades. A metodologia proposta é utilizada com o objetivo de tornar mais precisa a previsão de um especialista, levando em conta um registro histórico sobre o problema. Ou seja, melhorar a previsão do especialista levando em conta o que ocorreu com as previsões anteriores. Para testar a metodologia, utilizou-se dados meteorológicos de temperatura e umidade provenientes de lavouras de café. Os dados foram gentilmente cedidos pelo CEPAGRI/Unicamp. Os testes foram avaliados através de dois indicadores estatísticos, ‘ \mathcal{D} ’ de Willmott e MAPE (Mean Absolute Percentage Error), mostrando que a metodologia foi capaz de melhorar a previsão do especialista na maioria das situações estudadas.

Palavras-chave

Conjuntos fuzzy, Inferência bayesiana, Modelos de predição, Distribuições de possibilidade, Biomatemática

Abstract

This work aims to apply Fuzzy set theory in forecasting models. The modeling methodology is largely based on continuous fuzzy relations and in the modus ponens, using t-norms (that in this context are similar to the copula operations in statistics). It is from the modus ponens that arises the “conditional” interpretation of some of the terms involved, and it is from there that an analogy with the Bayesian inference is made. However, it is only a conceptual analogy: this work do not involve probability distributions. Actually, fuzzy sets are treated as possibility distributions. The methodology is used to improve the accuracy of expert forecasting considering a historic data. Namely, to improve expert prediction based on past performance. To evaluate the test the methodology, temperature and humidity data from coffee crop was used. The data was gently provided by CEPAGRI/Unicamp. Results were validated using two different statistic indicators, MAPE (mean absolute percentage error) and Willmott ‘ \mathcal{D} ’, showing that the methodology was able to improve the expert prediction in most cases.

Keywords

Fuzzy set theory, Bayesian inference, Prediction models, Possibility distributions, Biomatemathics

Sumário

1	Introdução	1
1.1	Sobre o problema abordado	1
1.2	Sobre a abordagem	2
1.3	Importância do estudo	2
1.4	Organização do texto	3
2	Fundamentos teóricos	5
2.1	Números fuzzy	5
2.2	Conectivos básicos da lógica fuzzy	7
2.2.1	T-norma	9
2.2.2	T-conorma	9
2.2.3	Implicação Fuzzy	10
2.3	Relações Fuzzy	11
2.3.1	Modus Ponens	13
2.4	Teoria de probabilidades/Inferência bayesiana	16
2.4.1	Variável aleatória	16
2.4.2	Distribuições de probabilidade	17
2.4.3	Método bayesiano de inferência	23
2.5	Distribuições de possibilidades	29
2.6	Relevância histórica da abordagem	29
2.6.1	Zadeh	30
2.6.2	Nguyen	32
2.6.3	Hisdal	34
2.6.4	Bouchon	36
2.6.5	Lapointe	38
3	Metodologia	40
3.1	Retomando o Modus Ponens / Analogia com a Inferência bayesiana	41
3.2	Obtendo a distribuição posteriori	44
3.2.1	Forma 1: Utilizando a Regra de Composição	44
3.2.2	Forma 2: Forçando a consistência da distribuição conjunta	45
3.2.3	Escolhendo entre as duas soluções	46
3.2.4	Normalidade da distribuição posteriori	50
3.2.5	Resumo das equações da metodologia	51
3.3	Exemplo	52

3.3.1	Obtendo as funções necessárias para aplicar o método	52
3.3.2	Distribuição Posteriori	56
3.4	Metodologia para obtenção de funções a partir do registro histórico	64
3.4.1	Metodologia para obtenção da distribuição priori	65
3.4.2	Metodologia para obtenção da função de verossimilhança	69
4	Aplicação	75
4.1	Sobre os dados utilizados e sua organização	75
4.1.1	Origem dos dados fornecidos	76
4.1.2	Organização dos dados utilizados	76
4.2	Metodologia de testes	77
4.2.1	Separação dos dados em grupo de treinamento e grupo de testes	77
4.2.2	As quatro formas de realizar os cálculos do método	78
4.2.3	Sobre o número de testes realizados	79
4.2.4	Calibração necessária para a realização de testes	80
4.2.5	Sobre a realização dos testes	81
4.3	Medidas estatísticas para análise dos testes	82
4.3.1	\mathcal{D} de Willmott	83
4.3.2	MAPE	84
4.4	Exemplo de testes	85
4.4.1	Conjunto de dados de exemplo	85
4.4.2	Obtenção das funções para dados da aplicação	86
4.4.3	Realização do teste de exemplo	89
4.4.4	Análise estatística do teste de exemplo	94
4.5	Análise estatística	96
4.5.1	Sobre os critérios escolhidos	96
4.5.2	Conclusões da análise estatística	97
4.6	Avanços metodológicos	99
4.7	Problemas encontrados e perspectivas para estudos posteriores	99
4.7.1	Sobre os dados de Umidade – Verão (Apêndice A.3)	99
4.8	Conclusões	100
A	Dados utilizados	102
A.1	Temperatura[K] – Verão	103
A.1.1	PS	103
A.1.2	SAF1	104
A.1.3	SAF2	105
A.1.4	SAF3	106
A.2	Temperatura[K] – Inverno	107
A.2.1	PS	107
A.2.2	SAF1	108
A.2.3	SAF2	109
A.2.4	SAF3	110
A.3	Umidade[%] – Verão	111
A.3.1	PS	111
A.3.2	SAF1	112

A.3.3	SAF2	113
A.3.4	SAF3	114
A.4	Umidade[%] – Inverno	115
A.4.1	PS	115
A.4.2	SAF1	116
A.4.3	SAF2	117
A.4.4	SAF3	118
B	Algoritmo computacional	119
B.1	Sobre o algoritmo computacional desenvolvido	119
B.1.1	Refazendo cálculos da dissertação usando o software	120
B.1.2	Rotina computacional para cálculo de mínimo entre funções	123
B.2	Funções principais	125
B.3	Funções auxiliares	128
B.4	Dados do programa	131

Capítulo 1

Introdução

1.1 Sobre o problema abordado

O Brasil é um dos maiores produtores e exportadores de café no mundo. Estudos recentes demonstram que o aumento das temperaturas globais poderá afetar negativamente a produção do grão. Uma das formas de adaptação da cultura do café diante das mudanças climáticas é a utilização da técnica de arborização gerando um sombreamento capaz de controlar a temperatura no microclima local. Tal técnica, chamada de Sistema Agro-Florestal (ou SAF), é utilizada para propor três formas alternativas de cultivo utilizando arborização, onde cada uma das formas consiste em utilizar tipos distintos de arborização¹. Dessa forma, tem-se quatro microclimas de cultivo de café:

- Três que utilizam arborização, chamados respectivamente de SAF1, SAF2 e SAF3;
- O usual, que consiste apenas do cafeeiro. Ela é chamada de PS (pleno sol).

Foram construídas as quatro configurações de cultivo, e em cada uma delas foi instalada uma estação meteorológica, que media, a cada hora, as variáveis temperatura e umidade relativa. Foram também feitas simulações dos valores de temperatura e

¹As três configurações de plantio utilizam árvores ou leguminosas distintas; um dos sistemas se refere ao plantio de café arborizado com as leguminosas “Feijão Guandú” e “Leucena”, outro sistema ao plantio de café com as leguminosas “Leucena” e “Gliricidia”, e o terceiro refere-se ao café consorciado com árvores de Macadâmia.

umidade, utilizando o *software* ENVI-met², e estas simulações serão comparadas às medidas geradas por cada uma das estações meteorológicas.

1.2 Sobre a abordagem

Propõe-se utilizar conjuntos fuzzy com o objetivo de melhorar as previsões de temperatura e umidade do problema descrito na seção 1.1. Isso é feito através de uma abordagem que incorpora informação histórica à previsão fornecida pelo especialista. Quer dizer, tenta-se melhorar a previsão do especialista levando em conta *o que ocorreu às previsões anteriores*.

Em suma, a abordagem resume-se à obtenção da solução de um problema relacional fuzzy contínuo, com esta solução sendo obtida através de funções ajustadas a partir do registro histórico. A analogia com a inferência bayesiana é feita quando se observa esta obtenção de solução como um processo bayesiano de inferência, identificando termos análogos às distribuições priori, posteriori e verossimilhança no problema fuzzy descrito.

1.3 Importância do estudo

É cada vez mais frequente a utilização de recursos computacionais para simular sistemas físicos, com o objetivo de fazer previsões que possam ser verificadas posteriormente em campo. É frequente também a idéia de acoplar dados provenientes de uma situação específica (no caso um microclima de plantio) a uma abordagem que lida apenas com os dados (e não com a dinâmica ou modelagem que os originaram, no caso abordada pelo software ENVI-met) com o objetivo de assim aumentar o conhecimento sobre a situação específica estudada (no caso originando previsões mais precisas que as fornecidas pelo software ENVI-met).

Esta citada junção de duas abordagens de naturezas distintas pode parecer sem propósito, uma vez que um especialista dificilmente faz uma previsão sem utilizar al-

²Para mais informações sobre o software ENVI-met, <http://www.envi-met.com/>.

gum tipo informação sobre o que ocorreu anteriormente no problema. Dessa forma, a princípio pode parecer redundante aliar uma abordagem que agrega informação histórica ao conhecimento de um especialista.

Um argumento que defende a junção dessas duas abordagens é que a informação histórica pode ser um banco de dados muito grande, o que dificulta a obtenção deste tipo de informação sem uma técnica específica para tal.

É importante que as previsões em mãos sejam as mais precisas possíveis, sabendo que ultrapassada uma determinada faixa-limite de temperatura o cafeeiro não floresce e a produção é comprometida, é essencial saber se alguma das configurações de plantio estudadas mantém a temperatura dentro desta faixa segura, logo qual configuração seria a mais adequada frente às possíveis novas condições climáticas (de temperaturas mais elevadas). No caso particular de aplicação, previsões mais confiáveis permitem incentivar o cultivo de café utilizando sistemas agro-florestais, o que pode se provar uma vantagem frente às situações climáticas futuras. Além disso, pode incentivar trabalhos acadêmicos semelhantes em outras culturas estratégicas (cana, por exemplo).

1.4 Organização do texto

No capítulo 1, fala-se sobre o problema específico abordado, sobre a importância do estudo e a recorrência da metodologia utilizada. No capítulo 2 estabelece-se os conceitos necessários para a compreensão da metodologia do presente trabalho, além de um panorama histórico sobre os questionamentos que originaram a abordagem apresentada. No capítulo 3, aplica-se os conceitos do capítulo 2 para motivar a metodologia (que é onde surge a relação com a metodologia bayesiana) utilizada e desenvolvê-la. Além disso, com o objetivo de ilustrar a metodologia proposta, os passos cálculos do método associados são feitos para um exemplo de previsão. No capítulo 4 versa especificamente sobre uma aplicação utilizando dados reais, como serão realizados e avaliados os testes do método, e é onde se realiza as análises estatísticas necessárias que permitem estabelecer conclusões sobre o comportamento do método em uma situação real. Por fim,

são feitos alguns comentários sobre os problemas enfrentados, os avanços metodológicos que foram realizados e algumas questões que podem ser investigadas futuramente.³

³Neste trabalho utiliza-se o formato americano para representar números, isto é, ‘.’ ao invés de ‘,’ como separador de decimais.

Capítulo 2

Fundamentos teóricos

O modelo matemático de predição proposto pelo presente trabalho utiliza sistemas que envolvem a teoria de conjuntos fuzzy. Os conceitos básicos necessários para o desenvolvimento deste trabalho são descritos a seguir. Denota-se a sentença “ A é subconjunto fuzzy do conjunto universo U ” por “ $A \in \mathcal{F}(U)$ ”.

2.1 Números fuzzy

Um número fuzzy é uma extensão de um número usual, no sentido que pode se referir não apenas a um único valor, mas a um conjunto conexo de valores possíveis, onde cada valor possível tem um “peso” entre 0 e 1. Tal “peso” é chamado de grau de pertinência, e a função que determina o “peso” de cada elemento é chamada função de pertinência do número fuzzy. Para que seja um número fuzzy, deve existir algum elemento cujo valor de pertinência é 1 (normalidade do conjunto). Assim como a teoria de conjuntos fuzzy é uma extensão da teoria de conjuntos clássica (onde os únicos “pesos” possíveis são 0 e 1), números fuzzy são uma extensão de números reais. Neste trabalho, eles números fuzzy serão utilizados para modelar as funções de pertinência dos conjuntos necessários para a aplicação do método proposto.

Definição 2.1 (Número fuzzy). *Subconjunto fuzzy dos conjuntos reais, convexo, com*

suporte limitado e (pelo menos) contínuo por partes.

Exemplo (Número fuzzy triangular). *É um número fuzzy cuja função de pertinência é dada pela equação (2.1.1), onde a, u e b são constantes positivas.*

$$\text{triang}(a, u, b)(y) = \begin{cases} \frac{(y-a)}{u-a} & \text{se } a \leq y < u \\ \frac{(y-b)}{u-b} & \text{se } u \leq y < b \\ 0 & \text{caso contrário} \end{cases} \quad (2.1.1)$$

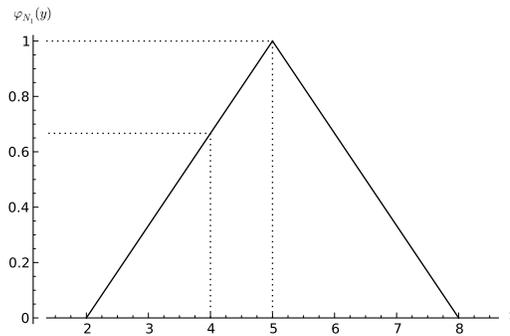
Suponha que o número fuzzy N_1 represente “Aproximadamente 5”. Sua função de pertinência, denotada por φ_{N_1} , é denotada por $\text{triang}(2, 5, 8)$. A função de pertinência correspondente é dada na equação 2.1.2, e a figura 2.1a representa graficamente este número fuzzy. Por exemplo, $\varphi_{N_1}(4) = \frac{2}{3}$ (2.1.2), ou seja, o valor 4 é compartilhável com o número fuzzy N_1 com grau $\frac{2}{3}$.

$$\varphi_{N_1}(y) = \text{triang}(2, 5, 8)(y) = \begin{cases} \frac{(y-2)}{3} & \text{se } 2 \leq y < 5 \\ \frac{(8-y)}{3} & \text{se } 5 \leq y < 8 \\ 0 & \text{caso contrário} \end{cases} \quad (2.1.2)$$

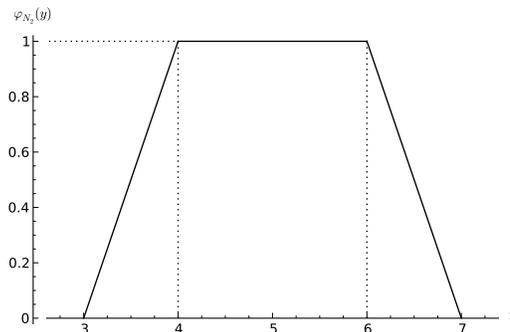
Exemplo (Número fuzzy trapezoidal). *Suponha o número fuzzy cuja função de pertinência seja dada pela equação (2.1.3), onde a, u, v e b são constantes positivas.*

$$\text{trapez}(a, u, v, b)(y) = \begin{cases} \frac{(y-a)}{u-a} & \text{se } a \leq y < u \\ 1 & \text{se } u \leq y \leq v \\ \frac{(y-b)}{v-b} & \text{se } v \leq y < b \\ 0 & \text{caso contrário} \end{cases} \quad (2.1.3)$$

Suponha agora que o número fuzzy N_2 modele “Entre 4 e 6”, e que sua função de pertinência seja $\varphi_{N_2}(y) = \text{trapez}(3, 4, 6, 7)(y)$. A figura 2.1b representa graficamente a função de pertinência do conjunto fuzzy N_2 .



(a) Número fuzzy triangular N_1 “Aproximadamente 5”, $\varphi_{N_1}(y) = \text{triang}(2, 5, 8)(y)$



(b) Número fuzzy trapezoidal N_2 “Entre 4 e 6”, $\varphi_{N_2}(y) = \text{trapez}(3, 4, 6, 7)(y)$

Figura 2.1: Exemplos de números fuzzy: triangular e trapezoidal.

2.2 Conectivos básicos da lógica fuzzy

Os primeiros passos em lógica matemática são realizados com o estudo dos conectivos “e”, “ou”, “não” e “implicação”. Tais conectivos são tipicamente usados na modelagem matemática através de sentenças do tipo:

“Se a está em A e b está em B ,
então c está em C ou d não está em D .”

Os valores lógicos para cada conectivo são estudados por meio de tabelas verdades. Assim, o valor lógico de uma sentença, formada a partir de duas ou mais proposições, é obtido por meio de composições das tabelas dos conectivos presentes nesta sentença.

Por exemplo, supondo que A e B sejam conjuntos, a proposição

“ a está em A e b está em B ”

é verdadeira apenas se for verdade que a está em A e também que b está em B . O valor lógico desta sentença é uma consequência da tabela verdade para o conectivo **e**.

Na lógica clássica, sentenças verdadeiras têm valor **1**, enquanto sentenças falsas têm valor **0**. Pensando na extensão para o caso fuzzy, usa-se aqui a notação \wedge (mínimo) para a conjunção **e**, \vee (máximo) para **ou**, \neg para **negação** e \implies para **implicação**.

Sejam p e q duas proposições. As tabelas verdade para os conectivos apresentados são dadas a seguir nas tabelas 2.1 e 2.2.

Tabela 2.1: Tabela verdade referente aos operadores lógicos “e”, “ou” e “implicação”

p	q	$p \wedge q$	$p \vee q$	$p \implies q$
1	1	1	1	1
1	0	0	1	0
0	1	0	1	1
0	0	0	0	1

Tabela 2.2: Tabela verdade referente ao operador lógico “negação”

p	$\neg p$
1	0
0	1

Cada um dos conectivos lógicos acima pode ser visto como operador matemático, cujos valores coincidem com as das respectivas tabelas verdade, e é esse fato que justifica as notações \wedge para o conectivo **e** e \vee para o **ou**.

Observamos que, para avaliar logicamente expressões como a que inicia esta seção, “Se a está em A e b está em B , **então** c está em C **ou** d **não** está em D ”, admitimos que a mesma poderia somente assumir os valores 0 e 1. Esta suposição é coerente com o fato dos conjuntos relacionados serem clássicos. Agora, se admitirmos que os conjuntos envolvidos possam ser fuzzy, como avaliar logicamente tal expressão?

Para realizar a avaliação lógica dos conectivos no sentido fuzzy, deve-se estendê-los. Tais extensões são obtidas por meio das t-normas e t-conormas, que são operadores que generalizam conectivos lógicos clássicos. Estes serão vistos em maiores detalhes adiante.

2.2.1 T-norma

Uma t-norma é um operador que generaliza o conectivo “e” da lógica clássica (binária). Recebe os graus de pertinência de duas proposições “ p ” e “ q ”, e retorna o grau de pertinência de “ p e q ”. No caso clássico, p e q podem apenas ter graus 0 ou 1. Contudo, no caso fuzzy, podem valer qualquer valor no intervalo $[0, 1]$, e o mesmo ocorre com o grau de pertinência de “ p e q ”.

Formalmente, o operador $T : [0, 1] \times [0, 1] \rightarrow [0, 1]$, é uma *t-norma* se satisfizer as seguintes condições[6, p. 86]:

- *Elemento neutro*: $T(1, x) = x$;
- *Comutatividade*: $T(x, y) = T(y, x)$;
- *Associatividade*: $T(x, T(y, z)) = T(T(x, y), z)$;
- *Monotonicidade*: se $x \leq u$ e $y \leq v$, então $T(x, y) \leq T(u, v)$.

Uma outra notação possível é denotar $T(x, y)$ por $x \Delta y$. Os exemplos abaixo reproduzem a segunda coluna da Tabela 2.1.

Exemplo (Mínimo). $T_1(x, y) = x \Delta_1 y = \min\{x, y\} = x \wedge y$.

Exemplo (Produto). $T_2(x, y) = x \Delta_2 y = xy$.

2.2.2 T-conorma

Analogamente à t-norma, t-conormas tem generaliza o conectivo lógico “ou”. Recebe também os graus de pertinência de duas proposições lógicas “ p ” e “ q ”, e retorna o grau de pertinência de “ p ou q ”. Como nas t-normas, p , q e “ p ou q ” podem assumir valores no intervalo $[0, 1]$.

Formalmente, o operador $S : [0, 1] \times [0, 1] \rightarrow [0, 1]$, é uma *t-conorma* se satisfizer as seguintes condições:

- *Elemento neutro*: $S(0, x) = x$;

- *Comutatividade:* $S(x, y) = S(y, x)$;
- *Associatividade:* $S(x, S(y, z)) = S(S(x, y), z)$;
- *Monotonicidade:* se $x \leq u$ e $y \leq v$, então $S(x, y) \leq S(u, v)$.

Uma outra notação possível é denotar $S(x, y)$ por $x \nabla y$. Os exemplos abaixo reproduzem a terceira coluna da Tabela 2.1.

Exemplo. $S_1(x, y) = x \nabla_1 y = \max\{x, y\} = x \vee y$.

Exemplo. $S_2(x, y) = x \nabla_2 y = x + y - xy$.

2.2.3 Implicação Fuzzy

Um operador $\Rightarrow: [0, 1] \times [0, 1] \rightarrow [0, 1]$ é uma implicação fuzzy se satisfaz as seguintes condições:

1. Reproduz a terceira coluna da Tabela 2.1, a tabela verdade da implicação clássica;
2. É decrescente na primeira variável, ou seja, para cada $x \in [0, 1]$, tem-se:

$$(a \Rightarrow x) \leq (b \Rightarrow x) \text{ se } a \geq b;$$

3. É crescente na segunda variável, ou seja, para cada $x \in [0, 1]$, tem-se:

$$(x \Rightarrow a) \geq (x \Rightarrow b) \text{ se } a \geq b;$$

Exemplo (Implicação de Gödel).

$$(x \Longrightarrow y) = g(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ y & \text{se } x > y \end{cases} .$$

Exemplo (Implicação de Goguen).

$$(x \implies y) = g_n(x, y) = \begin{cases} 1 & \text{se } x \leq y \\ \frac{y}{x} & \text{se } x > y \end{cases} .$$

2.3 Relações Fuzzy

O conceito de relação em matemática é formalizado a partir da teoria de conjuntos. Intuitivamente, pode-se dizer que a relação será *fuzzy* quando opta-se pela teoria dos conjuntos fuzzy, e será clássica quando for usada a teoria de conjuntos para conceituar a relação em estudo. Uma relação indica se há ou não alguma associação entre dois objetos, enquanto uma relação fuzzy, além de indicar se há ou não tal associação, indica também o grau dessa relação.

Definição 2.2 (Relação). *Uma relação \mathcal{R} sobre $U_1 \times U_2 \times \cdots \times U_n$ é qualquer subconjunto do produto cartesiano $U_1 \times U_2 \times \cdots \times U_n$.*

Como a relação \mathcal{R} é um subconjunto do produto cartesiano, então ela pode ser representada por sua *função característica* $\chi_{\mathcal{R}}$:

$$\chi_{\mathcal{R}} : U_1 \times U_2 \times \cdots \times U_n \longrightarrow \{0, 1\},$$

com

$$\chi_{\mathcal{R}}(x_1, x_2, \dots, x_n) = \begin{cases} 1 & \text{se } (x_1, x_2, \dots, x_n) \in \mathcal{R} \\ 0 & \text{se } (x_1, x_2, \dots, x_n) \notin \mathcal{R} \end{cases}$$

O conceito matemático de relação fuzzy é formalizado a partir do produto cartesiano usual entre conjuntos, estendendo a função característica de uma relação clássica para uma função de pertinência.

Definição 2.3 (Relação fuzzy). *Uma relação fuzzy \mathcal{R} sobre $U_1 \times U_2 \times \cdots \times U_n$ é qualquer subconjunto fuzzy de $U_1 \times U_2 \times \cdots \times U_n$. Assim, uma relação fuzzy \mathcal{R} é definida por uma função de pertinência $\varphi_{\mathcal{R}} : U_1 \times U_2 \times \cdots \times U_n \longrightarrow [0, 1]$.*

Se a função de pertinência da relação fuzzy \mathcal{R} for indicada por $\varphi_{\mathcal{R}}$, então o número

$$\varphi_{\mathcal{R}}(x_1, x_2, \dots, x_n) \in [0, 1]$$

indica o grau com que os elementos x_i , que compõem (x_1, x_2, \dots, x_n) , estão relacionados segundo a relação \mathcal{R} .

Definição 2.4 (Regra da composição). *Sejam U e V dois conjuntos, $\mathcal{F}(U)$ e $\mathcal{F}(V)$ as classes de subconjuntos fuzzy de U e V respectivamente, e \mathcal{R} uma relação fuzzy sobre $U \times V$. A relação \mathcal{R} define um funcional de $\mathcal{F}(U)$ em $\mathcal{F}(V)$ que, a cada elemento $A \in \mathcal{F}(U)$, faz corresponder o elemento $B \in \mathcal{F}(V)$ cuja função de pertinência é dada por*

$$\varphi_B(y) = \sup_{x \in U} \{ \min(\varphi_{\mathcal{R}}(x, y), \varphi_A(x)) \}. \quad (2.3.1)$$

Esta expressão é conhecida como regra da composição.

Uma aplicação da regra da composição, definida acima, é que ela permite que, conhecendo a relação \mathcal{R} e o conjunto A , seja possível descrever cada elemento y do conjunto B . Outra possibilidade de aplicação é, conhecendo os conjuntos A e B , obter a relação \mathcal{R} . Este problema é de particular relevância para o presente trabalho, pois a equação (2.3.1) modela matematicamente a regra de inferência *modus ponens*, o que se verá a seguir na seção 2.3.1.

A importância do *modus ponens* para o presente trabalho encontra-se na possibilidade de resolver um problema de relações fuzzy, só que interpretando os termos envolvidos como termos do *modus ponens*, o que mostra a natureza *condicional* de alguns desses termos. A partir daí é que surge a relação com a teoria de inferência bayesiana da estatística, pois o *modus ponens* permite concluir “saídas” de um sistema, e estas saídas, vistas como distribuições de possibilidade, são análogas à distribuições posteriores na metodologia bayesiana. Isso é exposto com os devidos detalhes na seção 3.1.

2.3.1 Modus Ponens

Modus ponens é uma regra de inferência lógica. Esta regra de inferência envolve duas sentenças. A primeira delas é a proposição ($p \implies q$) (uma implicação, conforme a seção 2.2.3), e a segunda sentença é p . A partir dessas duas sentenças pode ser deduzido q , o conseqüente da implicação, como mostra o esquema a seguir.

$p \implies q$:	$\overbrace{\text{“}x \text{ é } A\text{”}}^{\text{antecedente}}$	$\overbrace{\text{então } y \text{ é } B\text{”}}^{\text{consequente}}$	$\text{“Se hoje é segunda-feira, João vai trabalhar”}$
	$\text{“}x \text{ é } A\text{”}$		$\text{“Hoje é segunda-feira”}$
	$\text{“}y \text{ é } B\text{”}$		$\text{“João vai trabalhar”}$

Para que a sentença q seja verdadeira através do modus ponens, é necessário que ambas as sentenças, a proposição ($p \implies q$) e o antecedente p , sejam verdadeiros. Ou seja, que “ $(p \implies q) \wedge p$ ” seja verdadeiro (onde “ \wedge ” é o operador lógico “e”).

Tabela 2.3: Tabela verdade referente ao modus ponens

(“ x é A ”) p	(“ y é B ”) q	(Implicação) $p \implies q$	(Modus Ponens) $(p \implies q) \wedge p$
1	1	1	1
1	0	0	0
0	1	1	0
0	0	1	0

Para o caso clássico, suponha A e B subconjuntos de (respectivamente) U e V . As expressões “ x é A ” e “ y é B ” são respectivamente as proposições lógicas p e q (antecedente e conseqüente), que podem ser representadas matematicamente pelas funções indicadoras $\chi_A(x)$ e $\chi_B(y)$, com $\chi_A : U \rightarrow \{0, 1\}$ e $\chi_B : V \rightarrow \{0, 1\}$. Denota-se a proposição condicional “ $\chi_A(x) \implies \chi_B(y)$ ” por $\chi_{\mathcal{R}}$ pelo fato desta proposição ser uma *relação* \mathcal{R} sobre $U \times V$ (“relação” no sentido de relação entre conjuntos, como definido na seção 2.3.1), com função indicadora $\chi_{\mathcal{R}} : U \times V \rightarrow \{0, 1\}$ definida por:

$$\chi_{\mathcal{R}}(x, y) = \begin{cases} 1 & \text{se } (x \notin A \text{ e } y \text{ qualquer}) \text{ ou } (x \in A \text{ e } y \in B) \\ 0 & \text{se } (x \notin A \text{ e } y \in B), \end{cases} \quad (2.3.2)$$

que, como esperado, reproduz a terceira coluna da tabela 2.3. Dessa forma:

$$\sup_{x \in U} \{\chi_{\mathcal{R}}(x, y) \wedge \chi_A(x)\} = \begin{cases} 1 & \text{se } y \in B \\ 0 & \text{se } y \notin B \end{cases} = \chi_B(y),$$

onde “ $\chi_{\mathcal{R}}(x, y) \wedge \chi_A(x)$ ” representa a função indicadora da sentença “ $(p \implies q) \wedge p$ ”. Assim, para o caso clássico, a fórmula (2.3.3) abaixo representa a função indicadora do conseqüente $\chi_B(y)$ (q ou “ y é B ”):

$$\chi_B(y) = \sup_{x \in U} \{\chi_{\mathcal{R}}(x, y) \wedge \chi_A(x)\}, \quad (2.3.3)$$

que é exatamente a equação (2.3.1) da regra da composição, exposta na seção 2.3.

2.3.1.1 Caso fuzzy

Suponha $A \in \mathcal{F}(U)$ (A subconjunto fuzzy de U) e $B \in \mathcal{F}(V)$ (B subconjunto fuzzy de V), cujas funções de pertinência são definidas como $\varphi_A : U \rightarrow [0, 1]$ e $\varphi_B : V \rightarrow [0, 1]$. A implicação $p \implies q$, representada pela relação \mathcal{R} , é uma *relação fuzzy* sobre $U \times V$, com $\varphi_{\mathcal{R}} : U \times V \rightarrow [0, 1]$. Com estas considerações em mente, pode-se trocar em (2.3.3) as funções indicadoras por funções de pertinência e o operador \wedge por alguma t-norma Δ , alternativamente denotada por T , de forma a obter

$$\varphi_B(y) = \sup_{x \in U} \{\varphi_{\mathcal{R}}^T(x, y) \Delta \varphi_A(x)\} = \sup_{x \in U} \{T(\varphi_{\mathcal{R}}^T(x, y), \varphi_A(x))\}, \quad (2.3.4)$$

que é a função de pertinência do conseqüente “ y é B ” com relação ao modus ponens fuzzy.

Na equação (2.3.4) acima, denota-se $\varphi_{\mathcal{R}}^T$ ao invés de $\varphi_{\mathcal{R}}$ pelo fato de a relação fuzzy \mathcal{R} ser obtida em função da t-norma escolhida¹. No caso fuzzy, existem algumas complicações na obtenção de $\varphi_{\mathcal{R}}$. Um dos motivos para que isso ocorra é a possibilidade da escolha de t-normas no caso fuzzy (na equação 2.3.3 lida-se sempre com o mínimo).

¹No caso clássico, $\chi_{\mathcal{R}}$ (Eq. 2.3.2) é facilmente obtida a partir da Tabela 2.3

Supondo que os conjuntos A e B sejam dados, é possível encontrar a solução $\varphi_{\mathcal{R}}^T$ da equação (2.3.4). Mas pode ser que exista mais de uma solução que satisfaça a equação; Como critério de escolha, busca-se comumente a solução maximal[6, p. 174], que é a “maior” solução. Definindo este conceito formalmente:

Definição 2.5 (Solução maximal). *Seja $\overline{\varphi}_{\mathcal{R}}^T$ uma relação solução de um problema relacional fuzzy. Se ocorrer que*

$$\forall x \in U, y \in V, \quad \overline{\varphi}_{\mathcal{R}}^T(x, y) \geq \varphi_{\mathcal{R}}^T(x, y),$$

para toda solução $\varphi_{\mathcal{R}}^T$, então $\overline{\varphi}_{\mathcal{R}}^T$ é a solução maximal do problema em questão.

Um dos principais resultados teóricos deste trabalho é o Teorema 1, enunciado abaixo (ver [6, p.176]). É ele que permite obter a solução $\varphi_{\mathcal{R}}^T(x, y)$ da equação 2.3.4 de forma analítica.

Teorema 1 (Existência da solução). *Suponha a equação relacional fuzzy (2.3.4):*

$$\varphi_B(y) = \sup_{x \in U} \{T(\varphi_{\mathcal{R}}^T(x, y), \varphi_A(x))\}$$

Dada uma t-norma T e os conjuntos fuzzy $A \in \mathcal{F}(U)$ e $B \in \mathcal{F}(V)$, a relação \mathcal{R} em $(U \times V)$, que é solução maximal² do problema acima, tem função de pertinência dada pela expressão:

$$\varphi_{\mathcal{R}}^T(x, y) = \sup_z \{z \in [0, 1] : T(z, \varphi_A(x)) \leq \varphi_B(y)\} = \psi_T(\varphi_A(x), \varphi_B(y)),$$

onde $\psi_T(a, b) = \sup_z \{z \in [0, 1] : T(z, a) \leq b\}$

Demonstração. A prova é encontrada em [8].

²Se num problema de relações fuzzy houver uma única solução, esta solução é a maximal. Portanto se não houver solução maximal possível, não existe solução para o problema em questão.

Neste trabalho foram utilizadas duas t-normas: a do mínimo (\wedge) e a do produto (\cdot). Para estes dois casos particulares de t-norma, o Teorema 1 toma a seguinte forma:

Corolário 1 (Solução para t-norma do produto). *Se $T=produto$,*

$$\psi_{(\cdot)}(\varphi_A(x), \varphi_B(y)) = \begin{cases} 1 & \text{se } \varphi_A(x) < \varphi_B(y) \\ \frac{\varphi_B(y)}{\varphi_A(x)} & \text{se } \varphi_A(x) \geq \varphi_B(y) \end{cases}$$

Corolário 2 (Solução para t-norma do mínimo). *Se $T=mínimo$,*

$$\psi_{(\wedge)}(\varphi_A(x), \varphi_B(y)) = \begin{cases} 1 & \text{se } \varphi_A(x) \leq \varphi_B(y) \\ \varphi_B(y) & \text{se } \varphi_A(x) > \varphi_B(y) \end{cases}$$

2.4 Teoria de probabilidades/Inferência bayesiana

Definição 2.6 (Espaço de probabilidades). *Um espaço de probabilidades é uma tripla $(\Omega, \mathcal{A}, \mathbf{P})$, que consiste de:*

- *O espaço amostral Ω : conjunto arbitrário não vazio*
- *A σ -álgebra $\mathcal{A} \subseteq 2^\Omega$: conjunto formado por subconjuntos de Ω , denominados eventos*
- *A medida de probabilidades $\mathbf{P} : \mathcal{A} \rightarrow [0, 1]$*

2.4.1 Variável aleatória

É frequente o caso onde um experimento é realizado onde preocupa-se principalmente em alguma função da saída ao invés da saída de fato. Por exemplo, ao jogar um par de dados é frequente se interessar pela soma dos dois dados, e não realmente interessado no valor separado de cada um deles. Isto é, interessa-se em saber que a soma é 7 e não que se a saída foi (1,6), (2,5), (3,4), (4,3), (5,2) ou (6,1). Outro exemplo: Ao

lançar moedas, pode interessar apenas o número do total de coroas que ocorreram, e não a sequência cara/coroa que foi resultada. Estas quantidades, ou mais formalmente, estas funções reais definidas em um espaço amostral, são conhecidas como *variáveis aleatórias*.

Pelo fato do valor de uma variável aleatória ser determinado pela saída do experimento, pode-se associar probabilidades aos valores possíveis da variável aleatória [15, p. 126]. Formalmente:

Definição 2.7 (Variável aleatória real). *Seja $(\Omega, \mathcal{A}, \mathbf{P})$ um espaço de probabilidades. A função $X : \Omega \rightarrow \mathbb{R}$ é uma variável aleatória se:*

$$\{\omega : X(\omega) \leq r\} \in \mathcal{A} \quad \forall r \in \mathbb{R}.$$

2.4.2 Distribuições de probabilidade

2.4.2.1 Caso discreto

Suponha que $X : \Omega \rightarrow A \subseteq \mathbf{R}$ é uma variável aleatória discreta definida no espaço amostral S . Então a função de massa de probabilidade $f_x : A \rightarrow [0, 1]$ é definida como:

$$f_X(x) = \mathbf{P}(X = x) = \mathbf{P}(\{s \in S : X(s) = x\}).$$

A probabilidade total para todos $x \in A$ deve ser igual a um:

$$\sum_{x \in A} f_X(x) = 1$$

Exemplo (Distribuição binomial). *Se uma variável aleatória K tem uma distribuição binomial com parâmetros n e p , escreve-se $K \sim B(n, p)$. A probabilidade de, em n tentativas, obter exatamente k sucessos em um evento cuja probabilidade é p , é dada*

pela função de densidade:

$$f_B(k; n, p) = Pr(K = k) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k},$$

para $k = 1, 2, \dots, n$, onde

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

2.4.2.2 Caso contínuo

Uma variável aleatória X tem densidade f , se:

$$\mathbf{P}[a \leq X \leq b] = \int_a^b f(x) dx.$$

Assim, se F_X é a distribuição cumulativa de X , então:

$$F_X(x) = \int_{-\infty}^x f(u) du,$$

e, se f é contínua no ponto x ,

$$f_X(x) = \frac{d}{dx} F(x);$$

Intuitivamente, pode-se pensar em $f_X(x) \cdot dx$ como a probabilidade de X “atingir” o intervalo $[x, x + dx]$.

Analogamente ao caso discreto, tem-se que

$$\int_{-\infty}^{\infty} f_X(u) du = 1,$$

Exemplo (Distribuição normal). *A distribuição normal (ou Gaussiana) é uma densidade de probabilidades contínua que têm a função de densidade dada por:*

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \mathcal{N}(\mu, \sigma^2),$$

onde o parâmetro μ é a média (valor de pico) e σ^2 a variância. σ é conhecido como o

desvio padrão.

2.4.2.3 Distribuição conjunta de probabilidades

Dadas duas variáveis aleatórias X e Y que são definidas no mesmo espaço de probabilidades, a distribuição conjunta de X e Y define a probabilidade de eventos definidos em termos de ambas as variáveis aleatórias. A distribuição cumulativa para um par de variáveis aleatórias é definida em termos da distribuição conjunta de probabilidades:

$$F_{X,Y}(x, y) = \mathbf{P}(X \leq x, Y \leq y)$$

A distribuição de X pode ser obtida a partir da distribuição conjunta de X e Y da seguinte forma:

$$\begin{aligned} F_X(a) &= \mathbf{P}\{X \leq a\} = \mathbf{P}\{X \leq a, Y \leq \infty\} \\ &= \mathbf{P}\left(\lim_{b \rightarrow \infty} \{X \leq a, Y \leq b\}\right) = \lim_{b \rightarrow \infty} F(a, b) \\ F_X(a) &\equiv F(a, \infty) \end{aligned}$$

Caso discreto Quando X e Y são variáveis aleatórias discretas, pode-se definir a *função conjunta de massa de probabilidades* de X e Y como:

$$f_{X,Y}(x, y) = \mathbf{P}\{X = x, Y = y\}$$

A função de massa de probabilidades de X pode ser obtida a partir de $f_{X,Y}$, da seguinte forma:

$$f_X = \mathbf{P}\{X = x\} = \sum_y f_{X,Y}(x, y)$$

Similarmente,

$$f_Y = \sum_x f_{X,Y}(x, y)$$

Caso contínuo Quando X e Y são variáveis aleatórias contínuas, a função de densidade de X pode ser definida em função da acumulada $F_X(\cdot)$ obtida acima da seguinte forma:

$$f_X(x) = \frac{\partial}{\partial x} F_{X,Y}(x, \infty).$$

A distribuição f_X acima é chamada de distribuição *marginal* de X .

Analogamente à função de densidade f_X acima, define-se a função de densidade conjunta de X e Y :

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial y \partial x} F_{X,Y}(x, y)$$

Dessa forma, pode-se definir as distribuições marginais de X e Y a partir da densidade conjunta:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy \quad \text{e} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx.$$

2.4.2.4 Variáveis aleatórias independentes

As variáveis aleatórias X e Y são ditas *independentes* se, para quaisquer conjuntos de números reais A e B ,

$$\mathbf{P}(X \in A, Y \in B) = \mathbf{P}(X \in A) \cdot \mathbf{P}(Y \in B)$$

Em outras palavras, X e Y são independentes se, para todo A e B , os eventos $E_A = \{X \in A\}$ e $E_B = \{X \in B\}$ são independentes.

2.4.2.5 Distribuições condicionais

Caso discreto Para dois eventos E e F quaisquer, a probabilidade condicional de E dado F é definida, supondo que $\mathbf{P}(F) > 0$, por

$$\mathbf{P}(E|F) = \frac{\mathbf{P}(E \cap F)}{\mathbf{P}(F)}$$

Dessa forma, se X e Y forem variáveis aleatórias discretas, é natural definir a função de massa de probabilidades condicional de X dado Y , por:

$$f_{X|Y}(x|y) = \mathbf{P}\{X = x|Y = y\} = \frac{\mathbf{P}\{X = x, Y = y\}}{\mathbf{P}\{Y = y\}} = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Se X e Y forem variáveis aleatórias independentes, tem-se:

$$f_{X|Y}(x|y) = \mathbf{P}\{X = x|Y = y\} = \frac{\mathbf{P}\{X = x, Y = y\}}{\mathbf{P}\{Y = y\}} = \frac{\mathbf{P}\{X = x\} \cdot \mathbf{P}\{Y = y\}}{\mathbf{P}\{Y = y\}} = \mathbf{P}\{X = x\}$$

Caso contínuo Se X e Y têm uma função de densidade conjunta $f_{X,Y}(x, y)$, então a função de densidade condicional de X , dado que $Y = y$, é definido para todos os valores y tais que $f_Y(y) > 0$, por:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Para motivar esta definição, multiplicando o lado esquerdo da igualdade acima por dx e a o lado direito por $(dx dy)/dy$ de forma a obter

$$\begin{aligned} f_{X|Y}(x|y) dx &= \frac{f_{X,Y}(x, y) dx dy}{f_Y(y) dy} \\ &\approx \frac{\mathbf{P}\{x \leq X \leq x + dx, y \leq Y \leq y + dy\}}{\mathbf{P}\{y \leq Y \leq y + dy\}} \\ f_{X|Y}(x|y) dx &\approx \mathbf{P}\{x \leq X \leq x + dx | y \leq Y \leq y + dy\} \end{aligned}$$

Em outras palavras, para valores pequenos de dx e dy , “ $f_{X|Y}(x|y) dx$ ” representa a probabilidade condicional de que X esteja entre x e $x + dx$, dado que Y está entre

$y + dy$.

Analogamente ao caso discreto, tem-se que se X e Y forem variáveis aleatórias independentes:

$$f_{X|Y}(x|y) = \mathbf{P}\{X = x|Y = y\} = \frac{\mathbf{P}\{X = x, Y = y\}}{\mathbf{P}\{Y = y\}} = \frac{\mathbf{P}\{X = x\} \cdot \mathbf{P}\{Y = y\}}{\mathbf{P}\{Y = y\}}$$
$$f_{X|Y}(x|y) = \frac{f_X(x) \cdot f_Y(y)}{f_Y(y)} = f_X(x)$$

2.4.2.6 Teorema de Bayes

Resumindo as expressões obtidas até aqui para o caso contínuo:

Distribuição condicional A distribuição condicional de X dado Y é dada por:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \quad (2.4.1)$$

Isolando $f_{X,Y}(x,y)$ da equação 2.4.1 acima, obtém-se

$$f_{X,Y}(x,y) = f_{X|Y}(x|y) \cdot f_Y(y), \quad (2.4.2)$$

e analogamente para a condicional de Y dado X :

$$f_{X,Y}(x,y) = f_{Y|X}(y|x) \cdot f_X(x), \quad (2.4.3)$$

Distribuições marginais A distribuição marginal de X é dada por:

$$f_X(x) = \int_y f_{X,Y}(x,y) dy, \quad (2.4.4a)$$

e substituindo a Eq. 2.4.2 na equação 2.4.4a acima, tem-se:

$$f_X(x) = \int_y f_{X|Y}(x|y) \cdot f_Y(y) dy \quad (2.4.4b)$$

De forma análoga, pode-se definir a distribuição marginal de Y :

$$f_Y(y) = \int_x f_{X,Y}(x, y) dx, \quad (2.4.5a)$$

e substituindo a Eq. 2.4.3 na equação 2.4.5a acima, tem-se:

$$f_Y(y) = \int_x f_{Y|X}(x|y) \cdot f_X(x) dx \quad (2.4.5b)$$

Repete-se agora a equação da distribuição condicional de X dado Y (Eq. 2.4.6), para todo y tal que $f_Y(y) > 0$:

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$$

Dessa forma, substituindo as Eqs. 2.4.2 e 2.4.5b na equação 2.4.6 acima, obtém-se:

$$(2.4.6)$$

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(x|y) \cdot f_X(x)}{\int_y f_{Y|X}(x|y) \cdot f_X(x) dx} \quad (2.4.7)$$

Chama-se a equação (2.4.7) acima de *Teorema de Bayes*.

2.4.3 Método bayesiano de inferência

A realização de previsão utilizando a metodologia bayesiana é frequentemente utilizada em problemas em que há informação histórica escassa ou inútil no momento que a previsão inicial é necessária. Um exemplo deste processo é a previsão de vendas de

produtos sazonais que, por conta da obsolência dos modelos de produto, têm uma vida curta. Uma previsão inicial de vendas do produto é feita ao início da temporada, e à medida que a temporada passa e os pedidos vão sendo feitos, a previsão inicial é modificada de alguma forma.

Seja y uma variável aleatória com função de densidade f , caracterizada por um parâmetro desconhecido x . Escreve-se essa densidade como $f(y|x)$ para indicar que a distribuição de probabilidades depende do valor de x . O parâmetro x é uma variável aleatória com função de densidade $h_0(x)$, que é chamada a densidade *priori* de x . Esta distribuição de probabilidades mede a informação subjetiva (ou “grau de confiança”) sobre o valor de x .

Em uma situação de previsão, suponha que a estimativa inicial de x é dada como uma distribuição de probabilidades $h_0(x)$, e que a informação subsequente sobre o fenômeno é obtida na forma de uma variável aleatória y , cuja distribuição de probabilidades $f(y|x)$ depende de x . A nova estimativa de x será na forma de uma distribuição revisada, $h_1(x|y)$, chamada distribuição *posteriori*. Utilizando o teorema de Bayes (Eq. 2.4.7), tem-se

$$h_1(x|y) = \frac{f(y|x) h_0(x)}{\int_x f(y|x) h_0(x) dx} = \frac{f(y|x) h_0(x)}{g(y)}, \quad (2.4.8)$$

onde:

$h_0(x)$: distribuição *priori* de x (distribuição marginal de x)

$f(y|x)$: *verossimilhança* de y , dado x (distribuição condicional de y dado x)

$g(y)$: distribuição de y , pesada por todos os valores de x (distribuição marginal de y)

$h_1(x|y)$: distribuição *posteriori* de x , dada a informação y (distribuição condicional de x dado y)

Deve-se notar que se x for uma variável discreta, a integral na Eq. (2.4.8) deve ser substituída por um somatório[14, p. 241]. A equação (2.4.8) fornece meios de combinar o dado observado y com a informação *priori* $h_0(x)$, para obter uma descrição revisada da incerteza sobre x .

A distribuição *posteriori* pode ser vista como um “acordo formal” entre a *verossimilhança* $f(y|x)$, que resume a evidência nos dados por si só, e a distribuição *priori* $h_0(x)$,

que resume as evidências externas que sugerem frequências maiores.

2.4.3.1 Exemplo de cálculo

Nesta seção, realiza-se os cálculos para casos particulares de distribuição priori e função de verossimilhança. Ambas são distribuições normais:

$$h_0(x) = \mathcal{N}(\bar{x}', v'_x) \equiv (2\pi v'_x)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{(x - \bar{x}')^2}{v'_x} \right]$$

$$f(y|x) = \mathcal{N}(x, \sigma_y^2) \equiv (2\pi \sigma_y^2)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{(y - x)^2}{\sigma_y^2} \right]$$

Calculando a distribuição conjunta $f_{(x,y)}$ (de acordo com Eq. 2.4.2):

$$\begin{aligned} f_{(x,y)}(x, y) &= h_0(x) \cdot f(y|x) \\ &= \left((2\pi v'_x)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{(x - \bar{x}')^2}{v'_x} \right] \right) \cdot \left((2\pi \sigma_y^2)^{-\frac{1}{2}} \exp \left[-\frac{1}{2} \frac{(y - x)^2}{\sigma_y^2} \right] \right) \\ f_{(x,y)}(x, y) &= 2\pi (v'_x \sigma_y^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\frac{(x - \bar{x}')^2}{v'_x} + \frac{(y - x)^2}{\sigma_y^2} \right] \right\} \end{aligned} \quad (2.4.9)$$

Substituindo (2.4.9) em (2.4.5b), tem-se:

$$\begin{aligned} g(y) &= \int_{-\infty}^{\infty} \mathbf{P}_{(x,y)}(x, y) dx \\ g(y) &= \int_{-\infty}^{\infty} 2\pi (v'_x \sigma_y^2)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\frac{(x - \bar{x}')^2}{v'_x} + \frac{(y - x)^2}{\sigma_y^2} \right] \right\} dx \\ g(y) &= (2\pi \sigma^2 v'_x)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \left[\frac{(y - \bar{x}')^2}{v'_x + \sigma_y^2} \right] \right\} \end{aligned} \quad (2.4.10)$$

De acordo com (2.4.6), dividindo (2.4.9) por (2.4.10):

$$h_1(x|y) = \left(2\pi \frac{v'_x \sigma_y^2}{v'_x + \sigma_y^2}\right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{[x - (yv'_x + \bar{x}'\sigma_y^2)/(v'_x + \sigma_y^2)]^2}{v'_x \sigma_y^2 / (v'_x + \sigma_y^2)} \right\}$$

$$h_1(x|y) = \mathcal{N} \left(\frac{yv'_x + \bar{x}'\sigma_y^2}{v'_x + \sigma_y^2}, \frac{v'_x \sigma_y^2}{v'_x + \sigma_y^2} \right)$$

Para $\bar{x}' = 0$, $v'_x = 1.5$, $\sigma_y = 1$, $y = 0.5$, obtém-se que:

$$h_0(x) = \mathcal{N}(\bar{x}', v'_x) = \mathcal{N}(0, 1.5) \quad (2.4.11)$$

$$f(y|x) = f(0.5|x) = \mathcal{N}(x, \sigma_y^2) = \mathcal{N}(x, 1) \quad (2.4.12)$$

$$h_1(x|y) = h_1(x|0.5) = \mathcal{N} \left(\frac{yv'_x + \bar{x}'\sigma_y^2}{v'_x + \sigma_y^2}, \frac{v'_x \sigma_y^2}{v'_x + \sigma_y^2} \right) = \mathcal{N} \left(\frac{3}{13}, \frac{6}{13} \right) \quad (2.4.13)$$

As funções 2.4.11, 2.4.12 e 2.4.13 obtidas acima para o caso particular de valores são representadas graficamente na figura 2.2 a seguir:

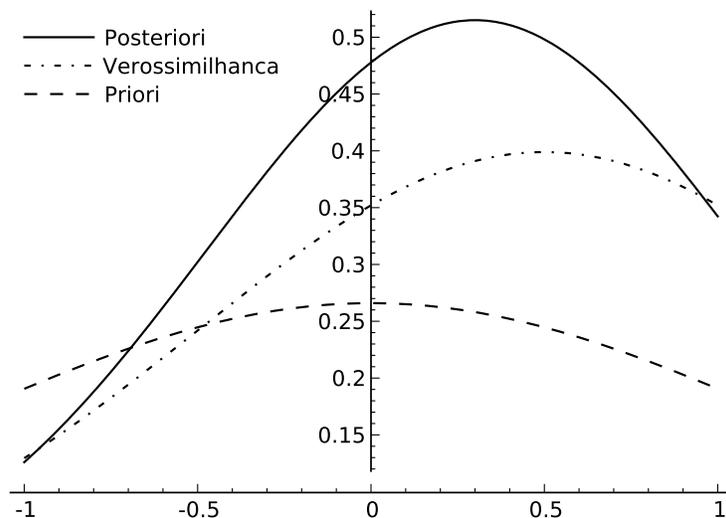


Figura 2.2: Gráfico da distribuição posteriori $h_1(x|y)$ do caso bayesiano (Eq. 2.4.13), juntamente com a função de verossimilhança (Eq. 2.4.12) e a distribuição priori (Eq. 2.4.11).

A interpretação de que a posteriori pode ser vista como um “acordo” entre as informações da distribuição priori e da verossimilhança é muito útil neste momento. Uma análise gráfica utilizando a figura 2.2 acima nota-se que, por exemplo, o ponto

de máximo da distribuição posteriori está entre os pontos de máximo da priori e da verossimilhança, o que intuitivamente valida a interpretação sobre a posteriori.

Interpretações como as acima são extremamente úteis na presente abordagem, pois mesmo que o presente trabalho não lide com distribuições de probabilidade, a analogia da abordagem com a metodologia bayesiana faz com que as interpretações sobre as distribuições priori, posteriori e a função de verossimilhança sejam mantidas. Dessa forma, análises da posteriori como a feita acima podem ser estendidas quando se fala de distribuições de possibilidades (caso fuzzy).

2.4.3.2 Sobre a função de verossimilhança $f(y|x)$

Nesta seção discute-se um pouco acerca do conceito de “verossimilhança” em probabilidades. Em uma linguagem não-técnica, “verossimilhança” é usualmente um sinônimo para “probabilidade”, mas em estatística ocorre uma distinção clara entre os dois conceitos. Pode-se perguntar (1) “Se eu lançar uma moeda honesta 100 vezes, qual a *probabilidade* da moeda cair com o lado da coroa para cima todas as vezes?”, ou (2) “Dado que eu lancei uma moeda 100 vezes e a moeda caiu com o lado da coroa para cima todas as vezes, qual é a *verossimilhança* de que a moeda é honesta?”, mas seria impróprio trocar “verossimilhança” por “probabilidade” nas duas sentenças.

A seção a seguir tem por objetivo ilustrar o *método da máxima verossimilhança*. Mesmo que este método não seja utilizado neste trabalho (nem nenhum método fuzzy análogo a ele), ter uma idéia intuitiva dos conceitos que envolvem a função de verossimilhança tem uma grande importância na compreensão deste trabalho.

Método da máxima verossimilhança: A essência deste método consiste em adotar para o parâmetro procurado o valor que maximize a *função de verossimilhança* correspondente ao resultado obtido na amostra[19]. Esclarece-se esse ponto a seguir.

Retirada uma amostra de uma população, a configuração dessa amostra irá, é claro, depender das características da população e, particularmente, do valor do parâmetro desconhecido x que se deseja estimar. Consideremos agora a probabilidade, ou densidade

de probabilidade, conforme o caso, de que uma particular amostra seja obtida. Essa probabilidade ou densidade de probabilidade irá depender, evidentemente, da amostra considerada e do valor do parâmetro x da população. *Fixada a amostra*, essa probabilidade ou densidade de probabilidade será em função de x , dita *função de verossimilhança correspondente a essa particular amostra*. Essa função admite, em geral, um único ponto de máximo, o qual fornecerá a estimativa de máxima verossimilhança do parâmetro x .

Exemplo. Suponha que uma caixa contenha dez bolas, das quais S são pretas e $10 - S$ são brancas. Uma amostra de quatro bolas com reposição é retirada dessa caixa, verificando-se que ela contém três bolas brancas e uma preta. Deve-se estimar o parâmetro S pelo método da máxima verossimilhança. Para tanto, deve-se determinar a função de verossimilhança correspondente ao resultado amostral obtido, a qual será dada pelas probabilidades de, em uma amostra de $n = 4$, sair exatamente uma bola preta ($k = 1$), dadas em função do parâmetro desconhecido S . Essas probabilidades podem ser obtidas pela aplicação de distribuição binomial, definida na seção 2.4.2.1. Designando $f(S|k = 1)$ a função de verossimilhança, tem-se:

$$f(S|k = 1) = f_B \left(k = 1; n = 4, p = \frac{S}{10} \right) = 4 \cdot \frac{S}{10} \left(\frac{10 - S}{10} \right)^3 = \frac{1}{2500} \cdot S(10 - S)^3$$

Na tabela, seguem os valores de $f(S|k = 1)$ calculados para todos os possíveis valores de S , verificando-se imediatamente que o valor de máxima verossimilhança é $S = 3$, o qual será, pois, a estimativa [17, p. 63].

S	$f(S k = 1)$	S	$f(S k = 1)$
0	0	6	384/2500
1	729/2500	7	189/2500
2	1024/2500	8	64/2500
3	1029/2500	9	9/2500
4	864/2500	10	0
5	625/2500		

2.5 Distribuições de possibilidades

Uma das questões mais controversas na modelagem de incertezas é a relação entre a teoria de probabilidades e a teoria de conjuntos fuzzy. Relacionada à teoria de conjuntos fuzzy é a teoria de possibilidades, presente na literatura desde os anos 70 (Zadeh [2], Dubois e Prade [11]). A teoria de possibilidades pode ser mais diretamente comparada à de probabilidades que conjuntos fuzzy pelo fato de possibilidades também proporem funções associadas a conjuntos que quantificam a incerteza de eventos[9]. Uma discussão mais aprofundada na questão sobre modelagem de incertezas é feita em Klir[12, 13], onde se analisa com mais detalhes a relação entre as teorias de probabilidades, possibilidades e conjuntos fuzzy.

Pode-se enxergar uma distribuição de possibilidades como uma representação da função de pertinência φ_X de um conjunto fuzzy X [2]. A distribuição de possibilidades de uma variável X , que toma valores em um conjunto universo U , é uma função de U em $[0, 1]$, denotada por $\pi_X(\cdot)$. Apesar disso, é importante fazer uma distinção entre a teoria de possibilidades e a teoria de conjuntos fuzzy: o objetivo da teoria de possibilidades é modelar incertezas induzidas por estados incompletos (deficientes) de informação sobre uma entidade existente, enquanto na teoria fuzzy o objetivo é modelar propriedades graduais[1, p. 121].

Em outras palavras, a teoria de possibilidades é a representação de incerteza a partir de alguns conjuntos fuzzy fixos e predeterminados[7, p. 238]. Isso permite que se possa definir o nível de sobreposição entre dois conjuntos fuzzy X (que toma valores em U) e Y (que toma valores em V), dando origem à distribuição conjunta de possibilidades destes conjuntos, denotada por $\pi_{X,Y}$ [2].

2.6 Relevância histórica da abordagem

Os questionamentos que tornaram o presente trabalho possível *não surgiram* tendo por objetivo propor modelos de predição (como o utilizado neste trabalho). Na verdade,

tais questionamentos surgiram com o objetivo de analisar uma antiga questão teórica dos conjuntos fuzzy: a de dependência e interatividade entre conjuntos fuzzy. Intuitivamente, tais questionamentos resumem-se a responder a seguinte pergunta: Como obter informação de um conjunto fuzzy a partir da informação sobre outro conjunto fuzzy? Ou seja, como obter o grau de pertinência de um conjunto fuzzy a partir do grau de pertinência de um outro conjunto fuzzy a ele relacionado? As seções a seguir expõem o trabalho de alguns autores sobre o assunto, incluindo trabalhos do próprio criador da teoria de conjuntos fuzzy, L.A. Zadeh.

2.6.1 Zadeh

Os primeiros questionamentos acerca da questão de interatividade/independência entre conjuntos fuzzy foram feitos por Zadeh[2] no final dos anos 70. Tais questionamentos surgiram da idéia de trabalhar com uma única função de pertinência que representasse a informação mais de um conjunto fuzzy simultaneamente. Suponha $Z = (X, Y) \in \mathcal{F}(U \times V)$ ³ e, representando as funções de pertinência como distribuições de possibilidade⁴, a função de pertinência do conjunto $Z = (X, Y)$, $\pi_Z \equiv \pi_{X,Y}$, é a função que representa a informação *conjunta* dos conjuntos fuzzy X e Y . Agora, suponha que a distribuição de possibilidades $\pi_{X,Y}$ seja dada. Como seria possível obter as informações de um dos dois conjuntos fuzzy, X por exemplo, a partir da distribuição $\pi_{X,Y}$? Zadeh propõe utilizar o conceito de projeção, ou seja, projetar a distribuição $\pi_{X,Y}$ em U . Isso é feito por ele da seguinte forma:

$$\pi_X(x) = \text{Proj}_U [\pi_{X,Y}] = \sup_{v \in V} \{\pi_{X,Y}(x, v)\} \quad (2.6.1)$$

$$\pi_Y(y) = \text{Proj}_V [\pi_{X,Y}] = \sup_{u \in U} \{\pi_{X,Y}(u, y)\} \quad (2.6.2)$$

³Na verdade Zadeh trabalhou com um número finito qualquer de conjuntos fuzzy, $(X_1, \dots, X_n) \in \mathcal{F}(U_1 \times \dots \times U_n)$. Utiliza-se aqui apenas dois conjuntos fuzzy, mas não há nenhuma perda de generalidade nos conceitos propostos pelo autor.

⁴Isto é, para $X \in \mathcal{F}(U)$, $\varphi_X(x) \equiv \pi_X(x)$.

Zadeh interpreta $\pi_X(x)$ e $\pi_Y(y)$ como as *distribuições marginais de possibilidade* dos conjuntos X e Y , respectivamente. Um caso particular que chamou a atenção do autor foi quando a distribuição $\pi_{X,Y}(x, y)$ era dada por

$$\pi_{X,Y}(x, y) = \pi_X(x) \wedge \pi_Y(y), \quad (2.6.3)$$

onde \wedge é o operador de mínimo.

Zadeh argumenta que quando a equação (2.6.3) vale para todo $x \in X, y \in Y$, ou seja, quando é possível “reconstruir” a distribuição do conjunto $X \times Y$ combinando as distribuições marginais dos conjuntos X e Y envolvidos, os conjuntos X e Y são *não-interativos*. Dessa forma:

Definição 2.8 (Conjuntos fuzzy não-interativos). *Sejam $X \in \mathcal{F}(U)$ e $Y \in \mathcal{F}(V)$, cujas distribuições de possibilidade são $\pi_X(x)$ e $\pi_Y(y)$. X e Y são não-interativos se ocorrer que, $\forall x \in U, y \in V$:*

$$\pi_{X,Y}(x, y) = \pi_X(x) \wedge \pi_Y(y),$$

onde $\pi_{X,Y}(x, y)$ é a distribuição conjunta de possibilidades de X, Y .

Intuitivamente, o conceito de não-interatividade pode ser visto através de um exemplo: suponha X e Y conjuntos fuzzy não-interativos, e que $\pi_X(x^*) = \alpha$ e $\pi_Y(y^*) = \beta < \alpha$. Dessa forma, pela equação (2.6.3), $\pi_{X,Y}(x^*, y^*) = \alpha \wedge \beta = \beta$. Agora, se o valor de $\pi_X(x^*)$ é aumentado para $\alpha + \delta$, $\delta > 0$, não é possível diminuir o valor de $\pi_Y(y^*)$ de forma que o valor de $\pi_{X,Y}(x^*, y^*)$ permaneça o mesmo. Assim, o fato de aumentar $\pi_X(x^*)$ *não pode ser compensado* por uma diminuição da possibilidade de $\pi_Y(y^*)$. Então, essencialmente, não-interatividade pode ser interpretada como a falta de capacidade das componentes da distribuição de possibilidades compensarem internamente as variações umas das outras, de forma que a distribuição que envolva todas as componentes mantenha suas características iniciais.

2.6.2 Nguyen

O trabalho de Nguyen[3] dá continuidade às investigações de Zadeh, e o seu principal resultado é que, para que o conceito de não-interatividade seja consistente, a distribuição condicional de possibilidades deve ser devidamente normalizada. Avanços são feitos na direção de obter explicitamente a distribuição condicional a partir da distribuição conjunta e das marginais.

A idéia da normalização surge do fato de, no caso probabilístico, a distribuição condicional ser um múltiplo da conjunta, com a constante multiplicadora $\alpha(\mathbf{P}_Y(y))$, sendo $\alpha(\mathbf{P}_Y(y)) = 1/\mathbf{P}_Y(y)$, como mostra a expressão (2.6.4) a seguir:

$$\mathbf{P}_{X|Y}(x|y) = \frac{\mathbf{P}_{X,Y}(x, y)}{\mathbf{P}_Y(y)} = \mathbf{P}_{X,Y}(x, y) \cdot \alpha(\mathbf{P}_Y(y)). \quad (2.6.4)$$

De forma análoga, se $\pi_X : U \rightarrow [0, 1]$, $\pi_Y : V \rightarrow [0, 1]$ e $\pi_{X,Y} : U \times V \rightarrow [0, 1]$ forem respectivamente as distribuições de possibilidades correspondentes à marginal de X , marginal de Y e à conjunta de X e Y , propõe-se uma expressão análoga ao caso probabilístico (Eq. 2.6.4) para a distribuição condicional $\pi_{X|Y}$ em função de $\alpha(\pi_X(u), \pi_Y(v))$:⁵

$$\pi_{X|Y}(u|v) = \pi_{X,Y}(u, v) \cdot \alpha(\pi_X(u), \pi_Y(v)), \quad (2.6.5)$$

onde $\pi_X(u) = \sup_{v \in V} \{\pi_{X,Y}(u, v)\}$ e $\pi_Y(v) = \sup_{u \in U} \{\pi_{X,Y}(u, v)\}$.

Agora deve-se escolher o valor de α de forma que sejam satisfeitos dois critérios, (i) que a distribuição condicional seja sempre normal, e (ii) que o conceito de não-interatividade de Zadeh (Def. 2.8) seja respeitado⁶. Ou seja, $\forall u \in U, v \in V$:

$$(i) \quad \pi_{X|Y}(u|v) = \pi_{X,Y}(u, v) \cdot \alpha(\pi_X(u), \pi_Y(v)) \in [0, 1]$$

$$(ii) \quad \pi_{X,Y}(u, v) = \pi_X(u) \wedge \pi_Y(v) \implies [\pi_X(u) \wedge \pi_Y(v)] \cdot \alpha(\pi_X(u), \pi_Y(v)) = \pi_X(u)$$

⁵Nota-se que se escolhe α em geral como função das duas distribuições marginais de possibilidades, e no caso probabilístico é apenas em função da marginal de Y .

⁶Ou seja, X e Y não-interativos $\implies \pi_{X|Y}(u|v) = \pi_X(u)$. Equivalentemente, $\forall u \in U, v \in V, \pi_{X,Y}(u, v) = \pi_X(u) \wedge \pi_Y(v) \implies \pi_{X,Y}(u, v) \cdot \alpha(\pi_X(u), \pi_Y(v)) = \pi_X(u)$

Deve-se portanto obter a constante $\alpha(\pi_X(u), \pi_Y(v))$ de forma a satisfazer as duas propriedades acima. Para satisfazer a condição (ii), deve valer que:

$$\forall a, b \in [0, 1], \quad (a \wedge b) \cdot \alpha(a, b) = a,$$

cuja solução para $\alpha(a, b)$ é:

$$\alpha(a, b) = \begin{cases} 1 & \text{se } a \leq b \\ \frac{a}{b} & \text{se } a > b \end{cases},$$

logo para $\alpha(\pi_X(u), \pi_Y(v))$ tem-se que

$$\alpha(\pi_X(u), \pi_Y(v)) = \begin{cases} 1 & \text{se } \pi_X(u) \leq \pi_Y(v) \\ \frac{\pi_X(u)}{\pi_Y(v)} & \text{se } \pi_X(u) > \pi_Y(v) \end{cases}. \quad (2.6.6)$$

Dessa forma, a distribuição condicional de possibilidades (Eq.(2.6.5), com a constante α definida conforme Eq. (2.6.6) toma a forma

$$\pi_{X|Y}(u|v) = \begin{cases} \pi_{X,Y}(u, v) & \text{se } \pi_X(u) \leq \pi_Y(v) \\ \pi_{X,Y}(u, v) \cdot \frac{\pi_X(u)}{\pi_Y(v)} & \text{se } \pi_X(u) > \pi_Y(v) \end{cases}. \quad (2.6.7)$$

Mostra-se que a condicional definida pela Eq. (2.6.7) acima é normal, satisfazendo a condição (i). Dessa forma, foi possível manter as definições feitas por Zadeh e dar um passo à frente na questão. Nguyen ilustra ainda a expressão obtida em uma aplicação simples de otimização de um sistema discreto que lida ao mesmo tempo com probabilidades e possibilidades.

2.6.3 Hisdal

Outros autores levaram a noção proposta por Zadeh adiante, e um caso importante é o estudado por Ellen Hisdal em [4]. Sua abordagem se caracteriza pela forte analogia com a teoria de probabilidades, e por proceder no sentido contrário ao que foi feito por Zadeh e Nguyen, no sentido que calcula a distribuição conjunta a partir da condicional e da marginal (enquanto tanto Zadeh quanto Nguyen obtém a distribuição condicional de possibilidades a partir das distribuições marginais e conjuntas). Primeiramente, o trabalho propõe uma expressão (em certo sentido) mais geral que a de Zadeh para a distribuição de possibilidades de conjuntos fuzzy, como se verá adiante. Ainda lidando com $X \in \mathcal{F}(U)$, $Y \in \mathcal{F}(V)$, tem-se:

$$\pi_{X,Y}(x, y) = \pi_Y(y) \wedge \pi_{X|Y}(x|y) \quad (2.6.8)$$

$$\pi_{X,Y}(x, y) = \pi_X(x) \wedge \pi_{Y|X}(y|x) \quad (2.6.9)$$

onde cada uma das duas equações acima é chamada de *distribuição conjunta de possibilidades* dos conjuntos fuzzy X e Y , enquanto $\pi_{X|Y}$ e $\pi_{Y|X}$ são chamadas de *distribuições condicionais de possibilidades*⁷. As distribuições marginais são definidas exatamente da mesma forma que Zadeh define em (2.6.1) e (2.6.2), apenas deve-se usar uma das distribuições conjuntas (2.6.9) ou (2.6.8).

Trabalha-se também com o conceito de não-interatividade da mesma forma que Zadeh o fez (Definição 2.8), e define-se um conceito similar a este: o de *independência* entre dois conjuntos fuzzy.

Definição 2.9. *O conjunto fuzzy X é independente do conjunto fuzzy Y se, $\forall x \in U, y \in V$,*

$$\pi_{X|Y}(x|y) = \pi_X(x).$$

⁷As definições de distribuição conjunta de possibilidades proposta por Hisdal é inspirada na teoria de probabilidades, e isso é ilustrado ao perceber que para obter a equação (2.6.9), $\pi_{X,Y}(x, y) = \pi_X(x) \wedge \pi_{Y|X}(y|x)$, é feita a substituição do operador de produto (\cdot) pelo de mínimo (\wedge) na equação probabilística $\mathbf{P}_{(X,Y)}(x, y) = \mathbf{P}_Y(y) \cdot \mathbf{P}_{X|Y}(x|y)$.

As expressões (2.6.8) e (2.6.9) foram ditas “mais gerais” que a (2.6.3) proposta por Zadeh pelo seguinte raciocínio que parte da equação (2.6.8):

$$\pi_{X,Y}(x, y) = \pi_Y(y) \wedge \pi_{X|Y}(x|y).$$

Se X é independente de Y , tem-se $\pi_{X|Y}(x|y) = \pi_X(x)$, obtendo:

$$\pi_{X,Y}(x, y) = \pi_Y(y) \wedge \pi_X(x), \quad (2.6.10)$$

que é exatamente a expressão (2.6.3) proposta por Zadeh. Assim, supondo Y independente de X , a expressão proposta por Hisdal recai no caso não-interativo de Zadeh (Definição 2.8). Dessa forma mostrou-se um fato importante: *independência implica não-interatividade*. O contrário não necessariamente ocorre, como é mostrado a seguir. Com o objetivo de isolar a distribuição condicional $\pi_{X|Y}(x|y)$ em (2.6.8), obtém-se:

$$\pi_{X|Y}(x|y) = \begin{cases} \alpha \in [\pi_{X,Y}(x, y), 1] & \text{se } \pi_Y(y) = \pi_{X,Y}(x, y), \\ \pi_{X,Y}(x, y) & \text{se } \pi_Y(y) > \pi_{X,Y}(x, y), \end{cases} \quad (2.6.11)$$

e se X e Y forem não-interativos (Definição 2.8), $\pi_{X,Y}(x, y) = \pi_Y(y) \wedge \pi_X(x)$, e para $\pi_Y(y) \geq \pi_X(x)$ tem-se que $\pi_{X,Y}(x, y) = \pi_X(x)$. Assim, a Eq.2.6.11 acima torna-se:

$$\pi_{X|Y}(x|y) = \begin{cases} \alpha \in [\pi_X(x), 1] & \text{se } \pi_Y(y) = \pi_X(x) \\ \pi_X(x) & \text{se } \pi_Y(y) > \pi_X(x) \end{cases} \quad (2.6.12)$$

Na equação 2.6.12, só ocorre que $\pi_{X|Y}(x|y) = \pi_X(x)$ para pares x, y tais que $\pi_Y(y) > \pi_X(x)$. Mas se existir algum par x^*, y^* tal que $\pi_Y(y^*) = \pi_X(x^*)$, em geral $\pi_{X|Y}(x^*|y^*) \neq \pi_X(x^*)$ (pois $\pi_{X|Y}(x^*|y^*)$ pode assumir qualquer valor entre $\pi_X(x^*)$ e 1), não satisfazendo portanto a Definição 2.9. Então conclui-se que *não-interatividade não implica*

*independência*⁸.

Muito embora se leve a questão adiante, mantendo o conceito de não-interatividade entre conjuntos fuzzy definido por Zadeh e ter proposto uma expressão mais geral para a distribuição conjunta de possibilidade (que tem como caso particular o não interativo), este trabalho apresenta algumas questões a serem investigadas:

1. Pela Eq. (2.6.11), se $\pi_X(x) \leq \pi_{X,Y}(x, y)$, a condicional $\pi_{Y|X}(y|x)$ pode assumir qualquer valor no intervalo $[\pi_{X,Y}(x, y), 1]$, isto é, a distribuição condicional não é sempre única, o que é um problema da abordagem da autora. Uma forma de contornar tal problema é escolher caso a caso um valor específico do intervalo que represente o valor a ser calculado, o que pode ser uma tarefa trabalhosa.
2. Como foi mostrado, os conceitos de não-interatividade (Definição 2.8) e independência (Definição 2.9) não são equivalentes. Como independência implica não interatividade, no caso possibilístico independência é um conceito mais forte que não-interatividade.
3. A definição de independência entre conjuntos fuzzy proposta não é simétrica, no sentido que *não ocorre* que $(X \text{ independente de } Y) \iff (Y \text{ independente de } X)$. Esta é uma propriedade importante, satisfeita na teoria de probabilidades.

2.6.4 Bouchon

O trabalho de Bernadette Bouchon é de 1985, portanto entre os de Zadeh/Nguyen/Hisdal (seções 2.6.1/2.6.2/2.6.3) e o de Lapointe (seção 2.6.5). Situa-se entre esses dois extremos também do ponto de vista conceitual, na medida que parte da abordagem de Zadeh ao representar funções de pertinência como distribuições de possibilidades, e aproxima-se de Lapointe por ter sido aparentemente o primeiro trabalho a visualizar uma implicação fuzzy como uma distribuição *condicional* de possibilidades, que é uma característica essencial da abordagem de Lapointe.

⁸Este resultado é importante pelo fato de que na teoria de probabilidades, se X e Y forem variáveis aleatórias, X e Y independentes $\iff X$ e Y não-interativos.

Define-se $f_1 : U_1 \rightarrow [0, 1]$ e $f_2 : U_2 \rightarrow [0, 1]$ como respectivamente as distribuições de possibilidades dos conjuntos fuzzy X_1 (que toma valores em U_1) e X_2 (que toma valores em U_2). A partir de $f_1(x)$ e $f_2(y)$ propõe-se um processo de inferência, com o objetivo de deduzir uma caracterização de X_2 a partir de uma observação de X_1 . A quantidade $q(y|x)$, $x \in U_1$, $y \in U_2$, é considerada uma implicação fuzzy de X_1 a X_2 . Então o modus ponens combina $q(y|x)$ com $f_1(x)$ para descrever uma caracterização de X_2 .

Define-se $\{q(y|x), x \in U_1, y \in U_2\}$ como uma implicação fuzzy de X_1 para X_2 , e será trabalhado com implicações de dois tipos[6, p. 89]: As do tipo R, que satisfazem

$$q_R(y|x) = \hat{F}(f_1(x), f_2(y)),$$

onde $\hat{F}(u, v) = \sup\{a \in [0, 1], T(u, a) \leq v\}$ para uma T-norma contínua T , e as implicações do tipo S, que satisfazem

$$q_S(y|x) = \max(N(f_1(x)), f_2(y)),$$

onde N é uma negação⁹.

Bouchon obtém explicitamente expressões para $q_R(y|x) = q(y|x)$, chegando às expressões abaixo¹⁰.

$$T(u, v) = \min(u, v), \quad q(y|x) = \begin{cases} 1 & \text{se } f_1(x) \leq f_2(y) \\ f_2(y) & \text{se } f_1(x) > f_2(y) \end{cases}$$

$$T(u, v) = uv, \quad q(y|x) = \min\left(\frac{f_2(y)}{f_1(y)}, 1\right)$$

As expressões acima são as mesmas que foram obtidas por Lapointe, como se verá na metodologia deste texto. De fato, Lapointe utiliza muitos dos resultados de Bouchon no seu trabalho, utilizando não só sua definição de distribuições condicionais de

⁹No artigo de Bouchon representa-se as R e S implicações (neste texto denotadas por q_R e q_S) respectivamente por q_F e q_N .

¹⁰Na verdade Bouchon obtém expressões também para as S-implicações, mas pelo fato de Lapointe trabalhar apenas com R-implicações, neste texto comenta-se apenas os resultados de Bouchon envolvendo implicações do tipo R.

possibilidades, mas também utilizando as mesmas definições de distribuições conjuntas e marginais de possibilidades que Bouchon (embora as argumentações envolvidas sejam distintas).

A principal escolha que diferencia o trabalho de Bouchon do de Stéphane é o fato de Bouchon, assim como Zadeh, Nguyen e Hisdal, utilizar as noções de não-interatividade e independência para o caso possibilístico, o que não é feito por Lapointe.

2.6.5 Lapointe

Esta é a principal referência do presente trabalho. O artigo [1] de Stéphane Lapointe leva adiante a discussão sobre o assunto, só que dessa vez utilizando a inferência bayesiana como fonte de analogias conceituais e as relações fuzzy como base teórica. Partindo do ponto de vista de Bouchon[5] (seção 2.6.4), o artigo contextualiza o problema em termos das relações fuzzy (expostas na seção 2.3), o que permite utilizar suas ferramentas para resolver este problema. Dessa forma, a parte teórica baseia-se agora em operações entre conjuntos fuzzy e não em analogias probabilísticas, pois os autores anteriores obtiveram suas expressões para distribuições condicionais trocando a operação de produto pelo mínimo nas equações da teoria de probabilidades. As interpretações sobre como surge o caráter “condicional” de algumas das distribuições envolvidas (levando à definição sobre distribuições condicionais de possibilidades) vêm também da teoria fuzzy, especificamente do Modus Ponens (seção 2.3.1).

A abordagem assemelha-se mais ao trabalho de Hisdal que de Zadeh/Nguyen, no sentido que obtém a conjunta a partir da condicional e marginal (e não o contrário, como feito por Zadeh e Nguyen). Um avanço com relação à abordagem de Hisdal é que, ao contextualizar o problema utilizando relações fuzzy, é possível obter a distribuição condicional para outras t-normas além do mínimo. Além disso, uma vez que se procura sempre pela solução maximal (a “maior” delas, conforme Definição 2.3), isso permite também que o problema de falta de unicidade da solução da abordagem de Hisdal (item 1 do final da seção 2.6.3) seja resolvido, pois para o caso de t-norma do mínimo,

a distribuição condicional torna-se

$$\pi_{X|Y}(x|y) = \begin{cases} 1 & \text{se } \pi_Y(y) = \pi_{X,Y}(x, y) \\ \pi_{X,Y}(x, y) & \text{se } \pi_Y(y) > \pi_{X,Y}(x, y) \end{cases},$$

que corresponde à equação 2.6.11 obtida por Ellen Hisdal, só que tomando o maior valor possível do intervalo $[\pi_{X,Y}(x, y), 1]$ (isto é, $\pi_{X|Y}(x|y) = 1$) quando $\pi_Y(y) \leq \pi_{X,Y}(x, y)$.

Uma característica muito restritiva das abordagens anteriores era definir formalmente a independência/interatividade entre os dois conjuntos fuzzy envolvidos, de forma análoga à teoria de probabilidades. Este problema é contornado através da abordagem bayesiana, onde a definição de independência no sentido formal tem menor relevância. Dessa forma, este trabalho não depende das definições de dependência/interatividade entre conjuntos fuzzy propostas por Zadeh e Hisdal, o que faz com que os problemas descritos nos itens 2 e 3 do final da seção 2.6.3 sejam contornados.

A analogia com a abordagem bayesiana surge não no sentido teórico (na medida que o presente trabalho não utiliza distribuições de probabilidades), mas para levar adiante as interpretações sobre as distribuições condicionais de possibilidade, e estabelecer a relação entre o processos de inferência fuzzy proposto com o bayesiano. Isto é, utilizando uma analogia do processo de inferência proposto com o bayesiano permite identificar uma analogia entre algumas distribuições de possibilidade do processo de inferência da presente abordagem como distribuição priori, posteriori e verossimilhança (conforme exposto na seção 2.4.3).

O processo de inferência análogo ao bayesiano é utilizado com o objetivo de obter informações de uma variável X com base em uma previsão Y de um especialista. Ou seja, “acoplar” à previsão do especialista Y (representada através da função de verossimilhança) o que já se conhece anteriormente sobre a variável X (priori), de forma a conhecer melhor o comportamento atual da variável X (posteriori). Dessa forma, a posteriori deve (ou deveria) fornecer mais informações sobre o comportamento atual de X do que a priori e a verossimilhança separadamente (mais detalhes na seção 2.4.3).

Capítulo 3

Metodologia

O presente trabalho utiliza o método proposto em [1], que tem por objetivo melhorar a previsão fornecida por um especialista utilizando um registro histórico, de forma a também levar em conta *o que ocorreu com as previsões anteriores*.

Dessa forma, tem-se por objetivo propor um modelo matemático que “capte” o conhecimento do especialista e, se possível, melhore as previsões feitas por ele. Isso será feito utilizando um registro histórico que contém informações sobre as previsões anteriores do especialista e as medidas correspondentes. Designando:

- X : Variável que se pretende prever (temperatura, por exemplo)
- Y : Previsão da variável X , fornecida por um especialista,

pretende-se melhorar a previsão Y sobre a variável X , levando em conta o registro histórico.

Usualmente em tais problemas utiliza-se um modelo equivalente proveniente da teoria de inferência bayesiana. Algumas das vantagens metodológicas do método proposto, se comparado ao modelo equivalente bayesiano:

- Uma solução analítica pode ser obtida no caso fuzzy sempre, o que é garantido pelo Teorema 1. Isso se deve principalmente de o modelo fuzzy passar em nenhum momento por resolução de integrais, o que ocorre no caso bayesiano e leva

frequentemente à utilização de aproximações numéricas. Por exemplo, no caso bayesiano é muito frequente a utilização de algoritmos computacionais aproximativos como o de Monte Carlo;

- Ao lidar com poucos dados, o método bayesiano requer modelos demasiadamente específicos cujos parâmetros precisam ser precisamente estimados. Por outro lado, na abordagem fuzzy, pouca informação muitas vezes é suficiente para obter modelos satisfatórios;

3.1 Retomando o Modus Ponens / Analogia com a Inferência bayesiana

Retoma-se o modus ponens, exposto na seção 2.3.1, que é uma *regra de inferência lógica* que tem por objetivo obter informações sobre um conseqüente (saída) “ q ” com base no conhecimento sobre o antecedente “ p ” e em uma regra condicional do tipo “ $p \implies q$ ” (implicação). Representando p , q e a implicação “ $p \implies q$ ” respectivamente através de funções de pertinência $\varphi_A(x)$, $\varphi_B(y)$ e $\varphi_{\mathcal{R}}^T(x, y)$ (onde $A \in \mathcal{F}(U)$, $B \in \mathcal{F}(V)$ e $\mathcal{R} \in \mathcal{F}(U \times V)$), a fórmula que representa a função de pertinência do conseqüente “ y é B ” a partir do modus ponens é, conforme (2.3.4):

$$\varphi_B(y) = \sup_{x \in U} \{ \varphi_{\mathcal{R}}^T(x, y) \Delta \varphi_A(x) \}, \quad (3.1.1)$$

onde $\varphi_A(x)$ corresponde ao antecedente “ x é A ” e $\varphi_{\mathcal{R}}^T(x, y)$ à implicação “ $(x \text{ é } A) \implies (y \text{ é } B)$ ”.

No método proposto, pretende-se obter informações sobre a variável X com base em Y , uma previsão de X . Então, se esta situação é descrita a partir do modus ponens, Y é o antecedente “ p ” e X o conseqüente “ q ”. Portanto, neste caso, a proposição “ $p \implies q$ ” representa a implicação “ $Y \implies X$ ”.

Representando X e Y como as distribuições de possibilidade $\pi_X \equiv \varphi_B$ e $\pi_Y \equiv \varphi_A$, a implicação “ $p \implies q$ ” torna-se “ $\pi_Y \implies \pi_X$ ”. Como toda implicação é uma proposição

lógica *condicional* do tipo “Se p então q ”, é possível denotar a função de pertinência da relação fuzzy $\varphi_{\mathcal{R}}^T$ como a *distribuição condicional de possibilidades* $\pi_{X|Y}^T$, pois representa o conhecimento sobre a variável X *dada* (condicionada à) informação sobre a variável Y . Reescrevendo a fórmula (3.1.1) em função das distribuições de possibilidades, tem-se:

$$\pi_X^T(x) = \sup_y \{ \pi_{X|Y}^T(x|y) \Delta \pi_Y(y) \} = \sup_y \{ T(\pi_{X|Y}^T(x|y), \pi_Y(y)) \}. \quad (3.1.2)$$

Observando a equação da distribuição conjunta de probabilidades (2.4.2),

$$\mathbf{P}_{X,Y}(x, y) = \mathbf{P}_{X|Y}(x|y) \cdot \mathbf{P}_Y(y),$$

e comparando seu lado direito com o termo “ $\pi_{X|Y}(x|y) \Delta \pi_Y(y)$ ” da equação (3.1.2), define-se

$$\pi_{X,Y}^T(x, y) = \pi_{X|Y}^T(x|y) \Delta \pi_Y(y) = T(\pi_{X|Y}^T(x|y), \pi_Y(y)) \quad (3.1.3)$$

como a *distribuição conjunta de possibilidades* de X, Y . De forma análoga, pode-se também definir

$$\pi_{X,Y}^T(x, y) = \pi_{Y|X}^T(y|x) \Delta \pi_X(x) = T(\pi_{Y|X}^T(y|x), \pi_X(x)). \quad (3.1.4)$$

É importante notar que as distribuições probabilísticas foram trocadas pelas análogas possibilísticas (fuzzy) nas equações (3.1.3) e (3.1.4), e o operador de produto foi trocado por uma t-norma T arbitrária. Agora, observando a equação da distribuição marginal de possibilidades da variável aleatória X , conforme (2.4.4a):

$$\mathbf{P}_X(x) = \int_y \mathbf{P}(x|y) \cdot \mathbf{P}_Y(y) dy,$$

e substituindo o operador de integral em y por supremo em y e comparando com a

equação (3.1.2), pode-se designar

$$\pi_X^T(x) = \sup_y \{ \pi_{X|Y}^T(x|y) \Delta \pi_Y(y) \} = \sup_{y \in U} \{ T(\pi_{X|Y}^T(x|y), \pi_Y(y)) \} \quad (3.1.5a)$$

$$\pi_X^T(x) = \sup_y \{ \pi_{X,Y}^T(x, y) \} \quad (3.1.5b)$$

como a *distribuição marginal de possibilidades* de X . De forma análoga, define-se a distribuição marginal de Y :

$$\pi_Y^T(y) = \sup_x \{ \pi_{Y|X}^T(y|x) \Delta \pi_X(x) \} = \sup_x \{ T(\pi_{Y|X}^T(y|x), \pi_X(x)) \} \quad (3.1.6a)$$

$$\pi_Y^T(y) = \sup_x \{ \pi_{X,Y}^T(x, y) \}, \quad (3.1.6b)$$

onde $\pi_{Y|X}^T(y|x)$ é a distribuição condicional de possibilidades de Y dado X , segundo a t -norma $T = \Delta$.

A abordagem proposta é baseada num padrão *bayesiano* de inferência, discutido na seção 2.4.3. Contextualizando o problema de previsão aqui abordado neste padrão, tem-se que:

- Variável que se quer obter informação: X
- Variável que fornece informação subsequente: Y .

Assim, as distribuições de possibilidades definidas nesta seção (marginais, condicionais de X dado Y) serão, em analogia à nomenclatura bayesiana, referenciadas como:

$$h_0(x) \sim \pi_X(x) \quad \equiv \pi(x) \quad : \text{distribuição } \textit{priori} \text{ de } X$$

$$f(y|x) \sim \pi_{Y|X}(y|x) \equiv \pi(y|x) \quad : \textit{verossimilhança} \text{ de } Y$$

$$g(y) \sim \pi_Y^T(y) \quad \equiv \pi^T(y) \quad : \text{distribuição marginal de possibilidades de } Y$$

$$h_1(x|y) \sim \pi_{X,Y}^T(x|y) \equiv \pi^T(x|y) : \text{distribuição } \textit{posteriori} \text{ de } X, \text{ dada a informação } Y=y$$

As interpretações da função de verossimilhança e das distribuições *priori* e *posteriori* têm neste trabalho o mesmo significado que no caso bayesiano. Dessa forma, os dois padrões de inferência são análogos, como mostra a Figura 3.1 a seguir.

$$\left. \begin{array}{l} \text{(priori)} \ h_0(x) \\ \text{(verossimilhança)} \ f(y|x) \end{array} \right\} h_1(x|y) \text{ (posteriori)}$$

$$\left. \begin{array}{l} \text{(priori)} \ \pi(x) \\ \text{(verossimilhança)} \ \pi(y|x) \end{array} \right\} \pi^T(x|y) \text{ (posteriori)}$$

Figura 3.1: Analogia entre os processos de inferência bayesiano (à esquerda) e o proposto (à direita).

Nota-se também que na presente abordagem a priori $\pi(x)$ e a função de verossimilhança $\pi(y|x)$ *não dependem da t -norma T* , pois suas expressões são obtidas a partir do registro histórico, de acordo com o procedimento das seções 3.4.1 e 3.4.2.

3.2 Obtendo a distribuição posteriori

Como será visto adiante, existem duas formas possíveis de se obter a distribuição posteriori $\pi^T(x|y)$. Com base em alguns argumentos, apenas uma das duas formas será escolhida.

3.2.1 Forma 1: Utilizando a Regra de Composição

Lembrando que a equação (3.1.2),

$$\mathbf{Condição 1:} \quad \pi(x) = \sup_y \{T(\pi^T(x|y), \pi^T(y))\}, \quad (3.2.1)$$

é apenas uma reinterpretação da equação (3.1.1) do modus ponens, pode-se, utilizando o Teorema 1, determinar também $\pi^T(x|y)$ em função de $\pi(x)$ e $\pi^T(y)$. Ou seja, o Teorema 1 permite obter as soluções maximais dos dois problemas:

$$\varphi_{\mathcal{R}}^T(x, y) = \psi_T(\varphi_A(x), \varphi_B(y)) \iff \pi^T(x|y) = \psi_T(\pi^T(y), \pi(x)),$$

onde $\psi_T(a, b) = \sup_{z \in [0,1]} \{z : T(a, z) \leq b\}$. Assim, é possível obter a primeira fórmula para a distribuição posteriori, denotada por $\pi_{\mathbf{1}}^T(x|y)$:

$$\pi_{\mathbf{1}}^T(x|y) = \psi_T(\pi^T(y), \pi(x)). \quad (3.2.2)$$

Como a distribuição $\pi^T(y)$ pode ser obtida a partir da equação $\pi^T(y) = \sup_x \{T(\pi(y|x), \pi(x))\}$ (Eq. 3.1.6a), a fórmula (3.2.2) acima fornece a distribuição posteriori $\pi(x|y)$ a partir da distribuição priori $\pi(x)$ e da função de verossimilhança $\pi(y|x)$, portanto a obtenção da posteriori é consistente com o processo de inferência segundo a Figura 3.1.

3.2.2 Forma 2: Forçando a consistência da distribuição conjunta

Procura-se por soluções em que haja consistência entre as duas formas de definir a distribuição conjunta, ou seja, que as equações Eqs. (3.1.3) e (3.1.4) sejam iguais. Ou seja, repetindo as duas equações da conjunta,

$$\pi^T(x, y) = T(\pi^T(x|y), \pi^T(y)) \quad (3.2.3)$$

$$\pi^T(x, y) = T(\pi(y|x), \pi(x)), \quad (3.2.4)$$

exige-se que ocorra:

$$\text{Condição 2: } \forall x, y, \quad T(\pi^T(x|y), \pi^T(y)) = T(\pi(y|x), \pi(x)) \quad (3.2.5)$$

As distribuições posteriori $\pi^T(x|y)$ que satisfazem a Condição 2 (Eq. 3.2.5) acima¹ são os elementos “z” do conjunto

$$\{z \in [0, 1] : T(z, \pi^T(y)) = T(\pi(y|x), \pi(x))\}. \quad (3.2.6)$$

¹Deve-se notar que, neste trabalho, $\pi(x|y)$ é obtida pois $\pi(y|x)$ e $\pi(x)$ são dados. Exatamente o mesmo procedimento da seção 3.2.2 pode ser realizado para obter $\pi(y|x)$ se $\pi(x|y)$ e $\pi(y)$ forem dados.

Como se procura apenas pela “maior” posteriori (ou solução maximal, no contexto de relações fuzzy), toma-se o supremo entre as soluções do conjunto definido na Eq. 3.2.6 acima (ou seja, toma-se supremo em z), de forma a obter \bar{z} :

$$\bar{z} = \sup_z \{z \in [0, 1] : T(z, \pi^T(y)) = T(\pi(y|x), \pi(x))\},$$

e, pela monotonicidade da t-norma T [1, p. 128],

$$\bar{z} = \sup_z \{z \in [0, 1] : T(z, \pi^T(y)) \leq T(\pi(y|x), \pi(x))\}. \quad (3.2.7)$$

Representando o conjunto da equação 3.2.7 acima através da notação do Teorema 1:

$$\bar{z} = \psi_T(\pi^T(y), T(\pi(y|x), \pi(x))). \quad (3.2.8)$$

A equação 3.2.8 permite obter a “maior solução” \bar{z} utilizando o mesmo Teorema 1 que foi utilizado na “Forma 1” (seção 3.2.1), pois ambas estão definidas utilizando o conjunto ψ_T . Assim, a Eq. 3.2.8 é a segunda forma de se obter a distribuição posteriori:

$$\pi_2^T(x|y) = \psi_T(\pi^T(y), T(\pi(y|x), \pi(x))) \quad (3.2.9)$$

Como $\pi^T(y) = \sup_x \{T(\pi(y|x), \pi(x))\}$ (Eq. 3.1.6a), a equação 3.2.9 acima fornece meios de obter a posteriori a partir da priori $\pi(x)$ e da função de verossimilhança $\pi(y|x)$, portanto a obtenção da posteriori é consistente com o processo de inferência esquematizado na Figura 3.1.

3.2.3 Escolhendo entre as duas soluções

Duas formas de obter a distribuição posteriori $\pi(x|y)$ foram obtidas, respectivamente $\pi_1^T(x|y)$ (3.2.2) e $\pi_2^T(x|y)$ (3.2.9), ambas em função do conjunto ψ_T (Teorema 1).

A escolha entre as soluções $\pi_1^T(x|y)$ e $\pi_2^T(x|y)$ leva dois fatos em consideração:

1. Mostra-se (conforme a observação abaixo) que, em geral, nem toda posteriori obtida através da “Forma 1”, ou seja, que satisfaça a Condição 1 (Eq. 3.2.1) satisfaz a Condição 2 (Eq. 3.2.5);
2. Por outro lado, mostra-se (Teorema 2) que *toda* posteriori obtida pela “Forma 2”, ou seja, que satisfaça a Condição 2 (Eq. 3.2.5) satisfaz também a Condição 1 (Eq. 3.2.1).

Observação: Uma posteriori $\pi_1^T(x|y)$ dada pela Eq. 3.2.2, solução maximal que satisfaz a Condição 1 (Eq. 3.2.1), em geral não satisfaz a Condição 2 (Eq. 3.2.5). Mostra-se este resultado através de um contra-exemplo:

Suponha a t-norma do produto (\cdot) , e a posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ dada pela Eq. 3.2.2, que por construção satisfaz a Condição 1 3.2.1. Seja $\pi^{(\cdot)}(x|y)$ tal que:

1. $\forall x, y, \pi(x) \geq \pi^{(\cdot)}(y) > 0$
2. Existe \bar{x} tal que, $\forall y, \pi(\bar{x}) > \pi^{(\cdot)}(\bar{x}, y)$

Aplicando o Corolário 1 à equação 3.2.2 (e usando o item 1 acima), obtém-se uma distribuição posteriori contínua:

$$\hat{\pi}_1^{(\cdot)}(x|y) = \frac{\pi(x)}{\pi^{(\cdot)}(y)}.$$

Resta saber se a distribuição posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ satisfaz a Condição 2 (Eq. 3.2.5).

Aplicando $\hat{\pi}_1^{(\cdot)}(x|y)$ no ponto \bar{x} no lado esquerdo da Condição 2,

$$\begin{aligned} T\left(\hat{\pi}_1^{(\cdot)}(\bar{x}|y), \pi^{(\cdot)}(y)\right) &= \hat{\pi}_1^{(\cdot)}(\bar{x}|y) \cdot \pi^{(\cdot)}(y) \\ &= \left(\frac{\pi(\bar{x})}{\pi^{(\cdot)}(y)}\right) \cdot \pi^{(\cdot)}(y) \\ T\left(\hat{\pi}_1^{(\cdot)}(\bar{x}|y), \pi^{(\cdot)}(y)\right) &= \pi(\bar{x}) > \pi^{(\cdot)}(\bar{x}, y), \end{aligned}$$

com $\pi(\bar{x}) > \pi^{(\cdot)}(\bar{x}, y)$ conforme o item 2 definido acima.

Então $T\left(\hat{\pi}_1^{(\cdot)}(x|y), \pi^T(y)\right) \neq \pi^{(\cdot)}(x, y)$ para algum x , ou seja, a Condição 2 (Eq. 3.2.5) não foi satisfeita. Dessa forma foi obtida uma posteriori $\hat{\pi}_1^{(\cdot)}(x|y)$ que satisfaz a Condição 1 (Eq. 3.2.1) mas que não a Condição 2 (Eq. 3.2.5).

Teorema 2. *Uma posteriori $\pi_2^T(x|y)$ dada pela equação (3.2.9), solução maximal que satisfaz a Condição 2 (Eq. 3.2.5), sempre satisfaz a Condição 1 (Eq. 3.2.1).*

Demonstração. Suponha $\pi_2^T(x|y)$ uma distribuição posteriori dada pela Eq. 3.2.9. Por construção ela satisfaz a Condição 2 (Eq. 3.2.5):

$$T\left(\pi_2^T(x|y), \pi^T(y)\right) = T\left(\pi(y|x), \pi(x)\right)$$

Utilizando a fórmula $\pi^T(x, y) = T\left(\pi(y|x), \pi(x)\right)$ (Eq. 3.1.4),

$$T\left(\pi_2^T(x|y), \pi^T(y)\right) = \pi^T(x, y)$$

Tomando supremo em y e substituindo $\pi(x) = \sup_y \{\pi^T(x, y)\}$ (Eq. 3.1.6b),

$$\begin{aligned} \sup_y \{T\left(\pi_2^T(x|y), \pi^T(y)\right)\} &= \sup_y \{\pi^T(x, y)\} \\ \sup_y \{T\left(\pi_2^T(x|y), \pi^T(y)\right)\} &= \pi(x), \end{aligned}$$

de onde se conclui $\pi_2^T(x|y)$ satisfaz a Condição 1 (Eq. 3.2.1). □

Escolha pela “Forma 2” de obter a posteriori:

Com base na observação acima e no Teorema 2, escolhe-se obter sempre a distribuição posteriori através da “Forma 2” (Eq. 3.2.9). Denota-se então a partir de agora $\pi_2^T(x|y) = \pi^T(x|y)$, ou seja:

$$\pi^T(x|y) = \psi_T\left(\pi^T(y), T\left(\pi(y|x), \pi(x)\right)\right), \quad (3.2.10)$$

onde $\psi_T(a, b) = \sup_{z \in [0,1]} \{z : T(a, z) \leq b\}$

A seguir são obtidas as distribuições posteriori de possibilidade acima para os casos particulares de t-norma do produto e mínimo, respectivamente $\pi^{(\cdot)}(x|y)$ (Corolário 1) e $\pi^{(\wedge)}(x|y)$ (Corolário 2).

Posteriori $\pi^{(\cdot)}(x|y)$, t-norma=produto:

Aplicando o Corolário 1 para $\varphi_A(x) = \pi^{(\cdot)}(y)$ e $\varphi_B(y) = \pi(y|x) \cdot \pi(x) = \pi^{(\cdot)}(x, y)$, obtém-se:

$$\pi^{(\cdot)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\cdot)}(y) < \pi^{(\cdot)}(x, y) \\ \frac{\pi(y|x) \cdot \pi(x)}{\pi^{(\cdot)}(y)} & \text{se } \pi^{(\cdot)}(y) \geq \pi^{(\cdot)}(x, y) \end{cases}$$

Como $\forall x, y$ tem-se que pela Eq. (3.1.6b), $\pi^{(\cdot)}(y) = \sup_x \{\pi^{(\cdot)}(x, y)\} = \sup_x \{\pi^{(\cdot)}(x, y)\} \geq \pi^{(\cdot)}(x, y)$. Dessa forma, a parte onde a posteriori vale sempre “1” é descartada. Sendo assim:

$$\pi^{(\cdot)}(x|y) = \frac{\pi(y|x) \cdot \pi(x)}{\pi^{(\cdot)}(y)} \quad (3.2.11)$$

Posteriori $\pi^{(\wedge)}(x|y)$, t-norma=mínimo:

Aplicando o Corolário 2 para $\varphi_A(x) = \pi^{(\wedge)}(y)$ e $\varphi_B(y) = \pi(y|x) \wedge \pi(x) = \pi^{(\wedge)}(x, y)$, obtém-se:

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) \leq \pi^{(\wedge)}(x, y) \\ \pi(y|x) \wedge \pi(x) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases}$$

Da Eq. 3.1.6b, conclui-se que $\pi^{(\wedge)}(y) \geq \pi^{(\wedge)}(x, y)$, pois $\pi^{(\wedge)}(y) = \sup_x \{\pi^{(\wedge)}(x, y)\} \geq \pi^{(\wedge)}(x, y)$. Dessa forma, a equação acima é simplificada:

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) = \pi^{(\wedge)}(x, y) \\ \pi(y|x) \wedge \pi(x) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases} \quad (3.2.12)$$

Comparação do processo de inferência da presente abordagem e do bayesiano:

Conclui-se que, para o caso particular de t-norma do produto, a distribuição posteriori $\pi^{(\cdot)}(x|y)$ solução assume a forma (Eq. 3.2.11):

$$\pi^{(\cdot)}(x|y) = \frac{\pi(y|x) \pi(x)}{\sup_x \{\pi(y|x) \pi(x)\}},$$

onde $\pi(x)$ é a distribuição priori e $\pi(y|x)$ a verossimilhança. Nota-se a clara semelhança entre esta expressão e a posteriori $h_1(x|y)$ do caso bayesiano (Eq. 2.4.8):

$$h_1(x|y) = \frac{f(y|x) h_0(x)}{\int_x f(y|x) h_0(x) dx},$$

onde $h_0(x)$ é a priori e $f(y|x)$ a verossimilhança. Há também a correspondência entre a distribuição priori ($\pi(x) \sim h_0(x)$) e verossimilhança ($\pi(y|x) \sim f(y|x)$) nos dois casos, o que permite notar mais claramente a analogia das duas metodologias, conforme o esquema da Figura 3.1.

3.2.4 Normalidade da distribuição posteriori

Definição 3.1 (Normalidade). *Suponha uma distribuição de possibilidades $\pi(u)$ que toma valores em um espaço U . $\pi(u)$ é dita normal se existir $u_0 \in U$ tal que:*

$$\pi(u_0) = 1$$

A seguir demonstra-se que a distribuição posteriori é sempre normal, uma propriedade muito desejável, pois dessa forma existirá sempre (pelo menos) um valor x^* da posteriori que tem pertinência máxima. Prova-se a seguir que a posteriori é normal, independentemente da priori ou a verossimilhança serem normais.

Teorema 3 (Normalidade da posteriori). *A solução da distribuição posteriori $\pi^T(x|y)$, dada pela Eq. (3.2.10), é normal para toda T-norma T .*

Demonstração. $\forall y \in V, \exists x_0 \in U$, tal que $\pi^T(y) = \pi^T(x_0, y)$, uma vez que $\pi^T(y) = \sup_x \{\pi^T(x, y)\}$ (Eq. 3.1.6b). Dessa forma, da equação (3.2.10):

$$\pi^T(x_0|y) = \sup_{z \in [0,1]} \{T(\pi^T(y), z) \leq \pi^T(x_0, y)\} = \psi_T(\pi^T(y), \pi^T(x_0, y))$$

$$\pi^T(x_0|y) = \sup_{z \in [0,1]} \{T(\pi^T(y), z) \leq \pi^T(y)\}$$

$$\pi^T(x_0|y) = \sup_{z \in [0,1]} \{\pi^T(y) \leq z \leq 1\}$$

$$\pi^T(x_0|y) = 1. \quad \square$$

3.2.5 Resumo das equações da metodologia

Com o objetivo de facilitar a consulta, as equações utilizadas na metodologia são reproduzidas a seguir.

$$\pi^T(x, y) = T(\pi(y|x), \pi(x)) \quad (3.2.13)$$

$$\pi^T(y) = \sup_x \{\pi^T(x, y)\} \quad (3.2.14a)$$

$$\pi^T(y) = \sup_x \{T(\pi(y|x), \pi(x))\} \quad (3.2.14b)$$

$$\pi^{(\wedge)}(x|y) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(y) = \pi^{(\wedge)}(x, y) \\ \pi^{(\wedge)}(x, y) & \text{se } \pi^{(\wedge)}(y) > \pi^{(\wedge)}(x, y) \end{cases} \quad (3.2.15a)$$

$$\pi^{(\cdot)}(x|y) = \frac{\pi^{(\cdot)}(x, y)}{\pi^{(\cdot)}(y)} \quad (3.2.15b)$$

3.3 Exemplo

Nesta seção, as equações da metodologia (seção 3.2.5) serão aplicadas a um conjunto de dados de exemplo, com o objetivo de ilustrar a metodologia proposta. Obtém-se, para um conjunto de dados de exemplo, as funções necessárias para a aplicação do método: a distribuição priori $\pi(x)$ (seção 3.3.1.1) e a função de verossimilhança $\pi(y|x)$ (seção 3.3.1.2). A partir destas, a distribuição posteriori é obtida, para os dois casos de t-norma estudados (mínimo e produto), na seção 3.3.2.

3.3.1 Obtendo as funções necessárias para aplicar o método

Optando por representar (modelar) as distribuições de possibilidades $\pi(y|x)$ e $\pi(x)$ respectivamente por números fuzzy triangulares e trapezoidais, obter tais conjuntos fuzzy resume-se a escolher os parâmetros (a, u, b) (caso triangular, $\pi(y|x)$, Eq. 2.1.1) e (a, u, v, b) (caso trapezoidal, $\pi(x)$, Eq. 2.1.3). Tal escolha de parâmetros baseia-se em informações retiradas do registro histórico. O método de obtenção de tais parâmetros para os dois conjuntos, para um exemplo de registro histórico, é explicado a seguir.

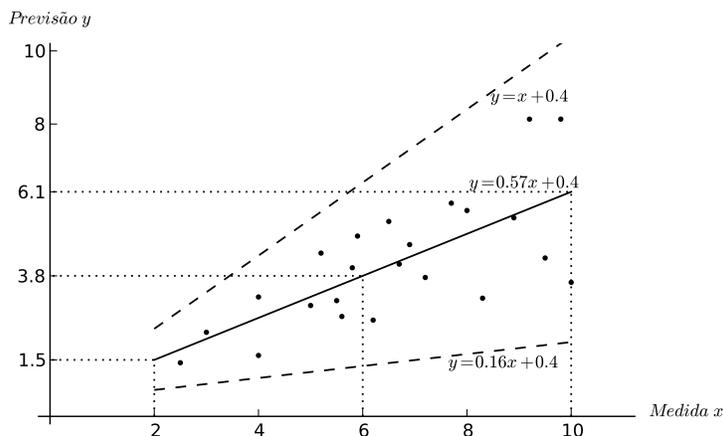


Figura 3.2: Registro histórico: gráfico $X \times Y$ (medida \times previsão)

3.3.1.1 Distribuição priori

Uma vez que esta distribuição de possibilidades representa o comportamento da variável X do problema sem levar em conta os valores de previsão Y , a informação sobre

quais valores são relevantes para a variável X (e com que pesos) pode ser fornecida, por exemplo, por um histograma. Como foi escolhido um número fuzzy trapezoidal (Eq. 2.1.3) para modelar a distribuição $\pi(x)$, é preciso então escolher os parâmetros (a, u, v, b) .

Suponha que um especialista informe que os valores da variável X geralmente se encontram no intervalo $[5,8]$. Além disso, em alguns casos mais raros o valor é menor que 5 mas nunca abaixo de 2, e maior que 8 mas nunca acima de 10. A afirmação do especialista pode ser validada utilizando o registro histórico da Figura 3.2. Dessa forma, tem-se que $u = 5$, $v = 8$, $a = 2$ e $b = 10$. O resultado da escolha de parâmetros é mostrado na equação 3.3.1 abaixo e na figura 3.3.

$$\pi(x) = \begin{cases} \frac{(x-2)}{3} & \text{se } 2 \leq x < 5 \\ 1 & \text{se } 5 \leq x \leq 8 \\ \frac{(10-x)}{2} & \text{se } 8 \leq x < 10 \\ 0 & \text{caso contrário} \end{cases} \quad (3.3.1)$$

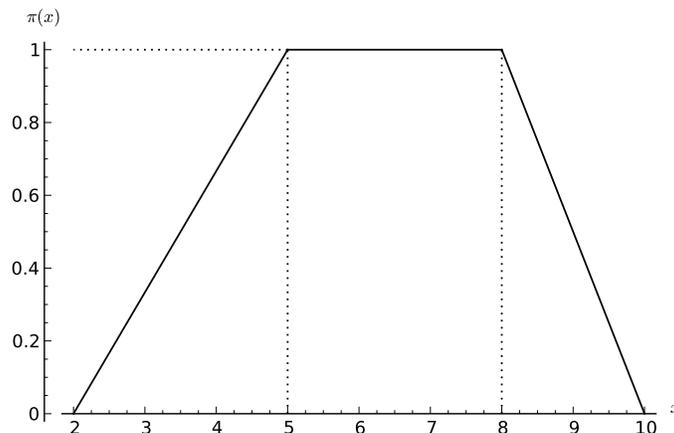


Figura 3.3: Gráfico da distribuição priori $\pi(x)$, Eq. 3.3.1.

A informação do especialista torna simples a obtenção da distribuição priori $\pi(\cdot)$, mas nem sempre esta informação é disponível. A seção 3.4.1 é dedicada à tarefa de obter a distribuição priori com base *apenas* no registro histórico como o da figura 3.2

utilizando histogramas.

3.3.1.2 Função de verossimilhança

Levando em consideração que a verossimilhança representa a *relação* entre as medidas X e as previsões Y do problema, os parâmetros são obtidos a partir de análises de um gráfico $X \times Y$, de forma que cada par ordenado (x^*, y^*) corresponde à medida x^* e sua previsão y^* .

Uma análise do gráfico da figura 3.2 utilizando: **(1)** o ajuste linear de mínimos quadrados dos dados e **(2)** duas retas-limite (uma inferior e outra superior) que englobam os pontos experimentais (em tracejado) é suficiente para obter a de verossimilhança $\pi(y|x)$. As citadas “retas-limite” pontilhadas são obtidas aumentando ou diminuindo o coeficiente linear do ajuste linear até que se consiga englobar todos os pontos experimentais do gráfico. Uma forma de obter tais retas para um conjunto de dados $X \times Y$ arbitrário é proposta na seção 3.4.2.

Como a função de verossimilhança $\pi(y|x)$ será modelada utilizando um número fuzzy triangular (Eq. 2.1.1), os parâmetros (u, a, b) serão obtidos em analogia com os do número fuzzy triangular N_1 (Figura 2.1a), que modela “em torno de 5”. Para tal, deve-se observar que: Na Eq. (2.1.2)/Fig 2.1a, o valor $u = 5$ é o mais representativo deste conjunto (pois tem pertinência igual a 1), e comparativamente, os “valores mais representativos” do gráfico da figura 3.2 agrupam-se em torno do ajuste linear $y = 0.57x + 0.4$ (correspondente à linha cheia, sem pontilhados), é razoável concluir que o parâmetro u seja dado exatamente pelo ajuste linear do gráfico. Dessa forma:

$$u = 0.57x + 0.4. \tag{3.3.2}$$

Quanto ao parâmetro a , nota-se também que na figura 2.1a:

1. Entre a e u , os valores de pertinência variam entre o menor e o maior valor de pertinência possível de forma *crescente*;

2. Segue então do item (1) acima que (para os valores entre a e u) quanto mais próximo de a , menor a relevância.

3. É possível observar comportamento análogo ao do item (2) no gráfico da figura 3.2, pois os pontos abaixo do ajuste linear vão se tornando cada vez menos relevantes à medida que se afastam do ajuste e se aproximam da linha pontilhada $y = 0.16x + 0.4$.

Os fatores acima permitem concluir que a corresponde à reta-limite (pontilhada) inferior. Ou seja,

$$a = 0.16x + 0.4. \quad (3.3.3)$$

Quanto ao parâmetro b , procede-se analogamente ao caso do parâmetro a , notando que nesse caso os valores entre u e b tornam-se menos relevantes à medida que se aproximam de b . Dessa forma, de forma análoga ao caso do parâmetro a conclui-se que

$$b = x + 0.4 \quad (3.3.4)$$

onde $y = x + 0.4$ é exatamente a reta-limite superior do gráfico da figura 3.2.

Sendo a verossimilhança um número fuzzy triangular, ela tem portanto a função de pertinência abaixo (Eq. 2.1.1)

$$\pi(y|x) = \begin{cases} \frac{(y-a)}{u-a} & \text{se } a \leq y < u \\ \frac{(y-b)}{u-b} & \text{se } u \leq y < b \\ 0 & \text{caso contrário,} \end{cases} \quad (3.3.5)$$

e substituindo os parâmetros “ $u = 0.57x + 0.4$ ” (Eq. 3.4.13), “ $a = 0.16x + 0.4$ ” (Eq. 3.4.11) e “ $b = x + 0.4$ ” (Eq. 3.4.12) na equação (3.3.5) acima, e deixando os intervalos

em função da variável x , obtém-se:

$$\pi(y|x) = \begin{cases} \frac{2.35x - 2.35y + 0.93}{x} & \text{se } y - 0.4 \leq x \leq 1.74y - 0.69 \\ \frac{-0.4x + 2.43y - 0.96}{x} & \text{se } 1.74y - 0.69 < x \leq 6.11y - 2.41 \\ 0 & \text{caso contrário} \end{cases} \quad (3.3.6)$$

A figura 3.4 ilustra o comportamento de $\pi(y|x)$ (Eq. 3.3.6) para alguns casos de previsão y .

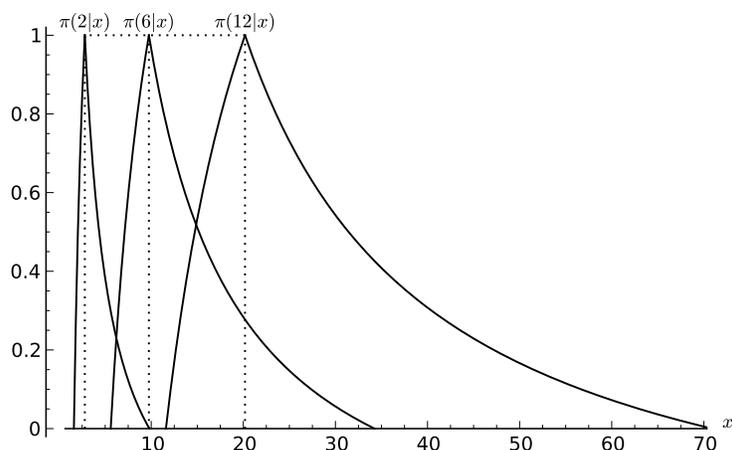


Figura 3.4: Gráficos da função de verossimilhança $\pi(y|x)$ (Eq. 3.3.6) para diversos valores de previsão Y .

Nota-se que o gráfico da figura 3.2 já continha as retas-limite (superior e inferior) necessárias para a obtenção da função de verossimilhança. Na seção 3.4.2 é proposta uma forma de, a partir *apenas dos pontos experimentais* de um gráfico $X \times Y$, obter $u(x)$, $a(x)$ e $b(x)$.

3.3.2 Distribuição Posteriori

Suponha que, utilizando as funções obtidas na seção 3.3.1 (ou seja, a função de verossimilhança e a distribuição priori dadas respectivamente pelas equações 3.3.6 e 3.3.1), deseja-se obter a distribuição posteriori de X para uma previsão particular $Y = 2$.

Como a distribuição priori $\pi(x)$ diz respeito somente ao comportamento de X sem

levar em conta Y , ela permanece a mesma da equação (3.3.1). Quanto à função de verossimilhança $\pi(y|x)$, basta substituir $y = 2$ em (3.3.6):

$$\pi(2|x) = \begin{cases} \frac{2.35x-3.77}{x} & \text{se } 1.60 \leq x \leq 2.79 \\ \frac{-0.4x+3.9}{x} & \text{se } 2.79 < x \leq 9.80 \\ 0 & \text{caso contrário} \end{cases} \quad (3.3.7)$$

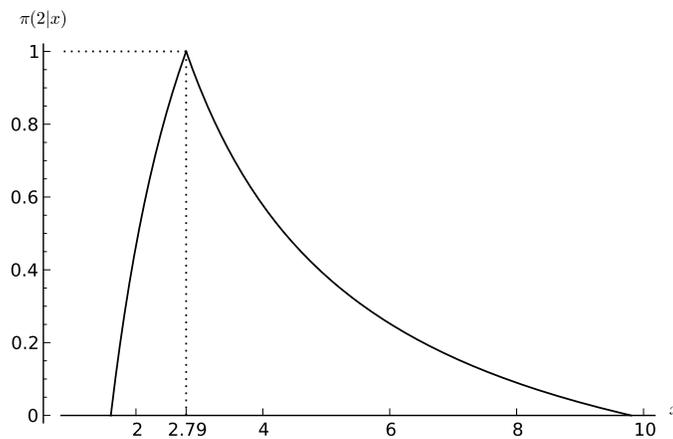


Figura 3.5: Gráfico da função de verossimilhança $\pi(2|x)$, para a previsão $Y = 2$ (Eq. 3.3.7). O valor $\bar{x}^* = 2.79$, tal que $\pi(2|\bar{x}^*) = 1$, é destacado.

A função de verossimilhança, Eq. 3.3.7/Fig. 3.5, pode ser vista como: dada a previsão $Y = 2$, os valores que se espera para a variável X estão em torno de $\bar{x}^* = 2.79$, que de acordo com a Eq. 3.3.7 é o valor com maior grau de pertinência. A distribuição dos valores em torno de $\bar{x}^* = 2.79$ não é simétrica, pois à esquerda tem-se um intervalo de valores com pertinência maiores que zero (entre 1.60 e 2.79) menor do que à direita (entre 2.79 e 9.80). Nota-se também que o crescimento (entre 1.60 e 2.79) e o decréscimo (entre 2.79 e 9.80) ocorrem de forma não-linear.

Conforme a seção 2.4.3, a distribuição posteriori pode ser vista como um “acordo” entre a informação da função de verossimilhança e a da distribuição priori. Para a previsão $Y = 2$, a verossimilhança de Y (Eq. 3.3.7) informa que o valor de X com maior possibilidade é 2.79, com uma maior tolerância para valores maiores do que para

menores que $\bar{x}^* = 2.79$. Mas qual a relevância destes valores com relação à variável X , ou seja: qual a relevância destes valores em relação à priori? Por exemplo, de acordo com a verossimilhança, o valor mais relevante é $\bar{x}^* = 2.79$. Só que este valor não é tão relevante (no sentido de possibilidade) para a variável X , pois $\pi(x = 2.79) = 0.3$ (Eq. 3.3.1, Fig. 3.3). Então, pode-se dizer que dificilmente o valor $\bar{x}^* = 2.79$ seria o mais relevante da distribuição posteriori $\pi(x|2)$, pois o citado “acordo” que origina a posteriori deve levar em consideração a relativa baixa relevância deste valor para a distribuição priori. Quando a posteriori for calculada será possível avaliar se de fato tal hipótese faz sentido.

Utilizando a verossimilhança obtida na seção 3.3.1.2 e a distribuição priori obtida na seção 3.3.1.1, os cálculos da distribuição posteriori são realizados a seguir, para a t-norma do produto na seção 3.3.2.1 e para t-norma do mínimo na seção 3.3.2.2.

3.3.2.1 Posteriores em $Y = 2$, t-norma=produto (\cdot)

Definições da distribuição priori e da função de verossimilhança

De acordo com Eq. (3.3.1) e (3.3.7):

$$\pi(x) = \begin{cases} \frac{(x-2)}{3} & \text{se } 2 \leq x < 5 \\ 1 & \text{se } 5 \leq x \leq 8 \\ \frac{(10-x)}{2} & \text{se } 8 < x \leq 10 \\ 0 & \text{caso contrário} \end{cases} \quad \pi(2|x) = \begin{cases} \frac{2.35x-3.77}{x} & \text{se } 1.60 \leq x \leq 2.79 \\ \frac{-0.4x+3.9}{x} & \text{se } 2.79 < x \leq 9.80 \\ 0 & \text{caso contrário} \end{cases}.$$

i. Distribuição conjunta em $Y = 2$

Partindo da fórmula (3.2.13):

$$\begin{aligned} \pi^{(\cdot)}(x, 2) &= \pi(x) \cdot \pi(2|x), \\ \pi^{(\cdot)}(x, 2) &= \begin{cases} \frac{(x-2)}{3} \cdot \left(\frac{2.35x-3.77}{x}\right) & \text{se } 2 \leq x \leq 2.79 \\ \frac{(x-2)}{3} \cdot \left(\frac{-0.4x+3.9}{x}\right) & \text{se } 2.79 < x \leq 5 \\ 1 \cdot \left(\frac{-0.4x+3.9}{x}\right) & \text{se } 5 < x \leq 8 \\ \frac{(10-x)}{2} \cdot \left(\frac{-0.4x+3.9}{x}\right) & \text{se } 8 < x \leq 9.80 \\ 0 & \text{caso contrário} \end{cases} \\ \pi^{(\cdot)}(x, 2) &= \begin{cases} \frac{(x-2) \cdot (2.35x-3.77)}{3x} & \text{se } 2 \leq x \leq 2.79 \\ \frac{(x-2) \cdot (-0.4x+3.9)}{3x} & \text{se } 2.79 < x \leq 5 \\ \frac{-0.4x+3.9}{x} & \text{se } 5 < x \leq 8 \\ \frac{(10-x) \cdot (-0.4x+3.9)}{2x} & \text{se } 8 < x \leq 9.80 \\ 0 & \text{caso contrário} \end{cases} \end{aligned} \quad (3.3.8)$$

ii. Marginal em $Y = 2$

Substituindo (3.3.8) na fórmula (3.2.14a):

$$\begin{aligned}
\pi^{(\cdot)}(y = 2) &= \sup_x \{ \pi^{(\cdot)}(x, 2) \}, \\
\pi^{(\cdot)}(y = 2) &= \max \left(\sup_{2 \leq x \leq 2.79} \left\{ \frac{(x-2) \cdot (2.35x - 3.77)}{3x} \right\}, \sup_{2.79 \leq x \leq 5} \left\{ \frac{(x-2) \cdot (-0.4x + 3.9)}{3x} \right\}, \right. \\
&\quad \left. \sup_{5 \leq x \leq 8} \left\{ \frac{-0.4x + 3.9}{x} \right\}, \sup_{8 \leq x \leq 9.80} \left\{ \frac{(10-x) \cdot (-0.4x + 3.9)}{2x} \right\} \right) \\
\pi^{(\cdot)}(y = 2) &= \max(0.26, 0.39, 0.38, 0.09) \\
\pi^{(\cdot)}(y = 2) &= 0.39 \tag{3.3.9}
\end{aligned}$$

iii. Distribuição posteriori em $Y = 2$

De acordo com (3.2.15b), $\pi^{(\cdot)}(x|2) = \frac{\pi^{(\cdot)}(x,2)}{\pi^{(\cdot)}(2)}$, onde $\pi^{(\cdot)}(x, 2)$ e $\pi^{(\cdot)}(2)$ são respectivamente (3.3.8) (3.3.9):

$$\pi^{(\cdot)}(x|2) = \begin{cases} \frac{0.85(x-2)(2.35x-3.77)}{x} & \text{se } 2 \leq x \leq 2.79 \\ \frac{0.85(-0.34x+3.90)(x-2)}{x} & \text{se } 2.79 < x \leq 5 \\ \frac{2.56(-0.4x+3.90)}{x} & \text{se } 5 < x \leq 8 \\ -\frac{1.28(-0.4x+3.90)(x-10)}{x} & \text{se } 8 < x \leq 9.80 \\ 0 & \text{caso contrário} \end{cases} \tag{3.3.10}$$

O gráfico da posteriori de X para a previsão $Y = 2$, $\pi^{(\cdot)}(x|2)$ (3.3.10), é mostrado na figura 3.6.

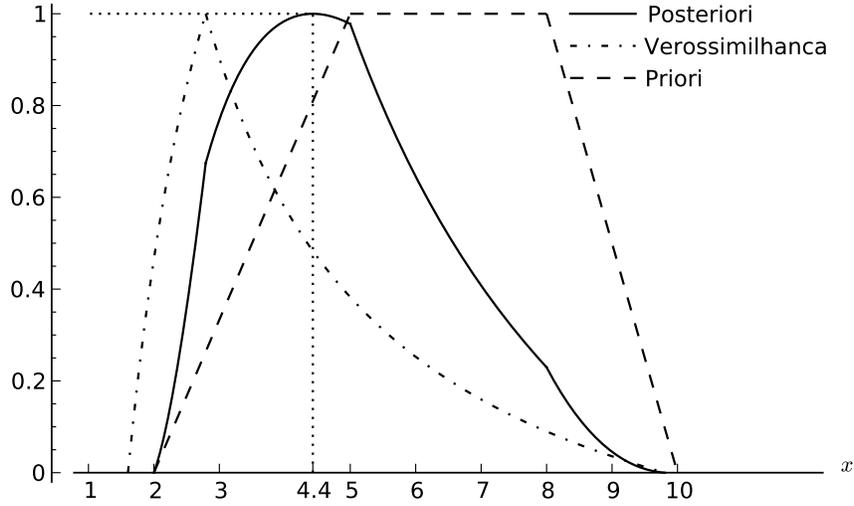


Figura 3.6: Distribuição posteriori para a previsão $Y = 2$, t-norma=produto (Eq. 3.3.10). O valor $\bar{x}^{(\cdot)} = 4.4$, tal que $\pi^{(\cdot)}(\bar{x}^{(\cdot)}|2) = 1$, é destacado.

3.3.2.2 Posteriori em $Y = 2$, t-norma=mínimo (\wedge)

Definições da distribuição priori e da função de verossimilhança

As funções utilizadas são as mesmas dos cálculos para a t-norma do produto, ou seja, Eq. 3.3.1 para a priori e 3.3.7 para a verossimilhança.

i. Distribuição conjunta em $Y = 2$

A distribuição conjunta (3.2.13) para a t-norma do mínimo, $\pi^{(\wedge)}(x, 2) = \min \{\pi(x), \pi(2|x)\}$, exige o cálculo de mínimo entre as duas funções. Esta é uma tarefa que aqui será feita graficamente (Fig. 3.7), mas pode também ser feita de forma analítica².

A figura 3.7 destaca o mínimo entre as duas funções, a priori $\pi(x)$ entre 2 e 3.855, e a verossimilhança $\pi(2|x)$ entre 3.85 e 9.80, ou seja:

$$\pi^{(\wedge)}(x, 2) = \begin{cases} \frac{x-2}{3} & \text{se } 2 \leq x \leq 3.85 \\ \frac{-0.40x+3.90}{x} & \text{se } 3.85 < x \leq 9.8 \\ 0 & \text{caso contrário} \end{cases} \quad (3.3.11)$$

²Na rotina computacional implementada neste trabalho, de fato se utiliza um método analítico para obter a função que corresponde ao mínimo entre duas funções. O algoritmo desenvolvido que realiza tais cálculos é descrito na seção B.1.2.

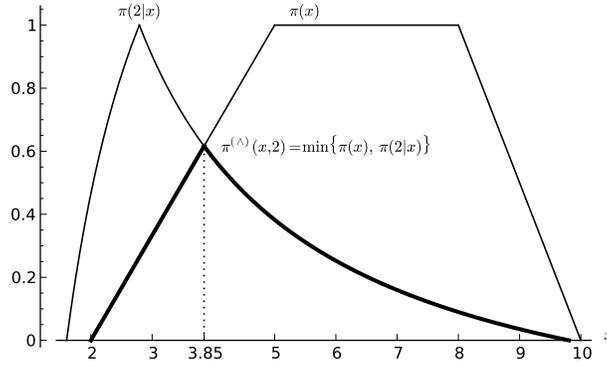


Figura 3.7: Em destaque $\pi^{(\wedge)}(x, 2)$, a distribuição conjunta em $Y = 2$ para a t-norma do mínimo.

ii. Marginal em $Y = 2$

Partindo da fórmula (3.2.14a),

$$\pi^{(\wedge)}(y = 2) = \sup_x \{ \pi^{(\wedge)}(x, 2) \},$$

e substituindo acima a equação (3.3.11):

$$\begin{aligned} \pi^{(\wedge)}(y = 2) &= \max \left(\sup_{2 \leq x \leq 3.85} \left\{ \frac{x - 2}{3} \right\}, \sup_{3.85 \leq x \leq 9.8} \left\{ \frac{-0.40x + 3.90}{x} \right\} \right) \\ \pi^{(\wedge)}(y = 2) &= \max(0.62, 0.62) = 0.62 \end{aligned} \quad (3.3.12)$$

iii. Distribuição posteriori em $Y = 2$

Utilizando as Eqs. (3.3.11) e (3.3.12), nota-se que $\pi(3.85, 2) = 0.62 = \pi(y = 2)$. Então, pela fórmula (3.2.15a), a distribuição posteriori para o caso da t-norma de mínimo é dada por:

$$\pi^{(\wedge)}(x|2) = \begin{cases} 1 & \text{se } \pi^{(\wedge)}(2) = \pi^{(\wedge)}(x, 2) \\ \pi^{(\wedge)}(x, 2) & \text{se } \pi^{(\wedge)}(2) > \pi^{(\wedge)}(x, 2) \\ 0 & \text{caso contrário} \end{cases}$$

E como pela Figura 3.7, para $x^* = 3.85$ tem-se que $\pi^{(\wedge)}(2) = \pi^{(\wedge)}(x^*, 2) = 0.62$, tem-se:

$$\pi^{(\wedge)}(x|2) = \begin{cases} 1 & \text{se } x = 3.85 \\ \frac{x-2}{3} & \text{se } 2 \leq x < 3.85 \\ \frac{-0.40x+3.90}{x} & \text{se } 3.85 < x \leq 9.8 \\ 0 & \text{caso contrário} \end{cases} \quad (3.3.13)$$

O gráfico da posteriori de X para a previsão $Y = 2$, $\pi^{(\wedge)}(x|2)$ (Eq. 3.3.13), é mostrado na figura 3.8.

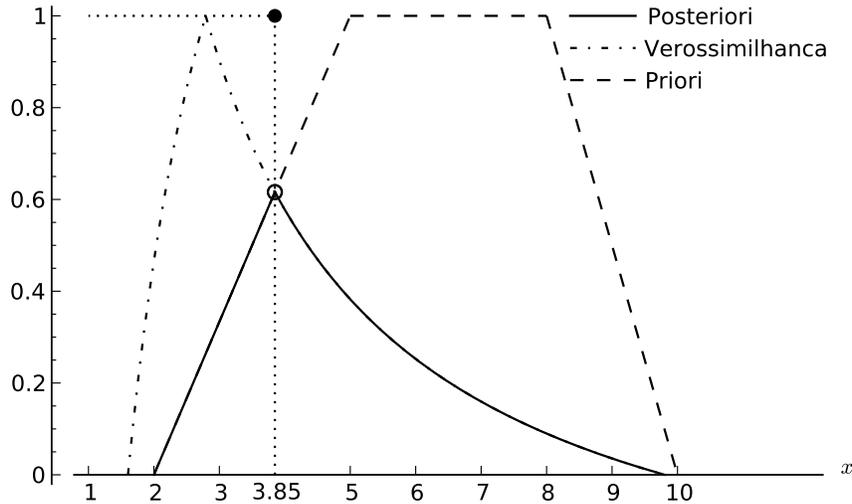


Figura 3.8: Distribuição posteriori para a previsão $Y = 2$, t-norma=mínimo (Eq. 3.3.13). O valor $\bar{x}^{(\wedge)} = 3.85$, tal que $\pi^{(\wedge)}(\bar{x}^{(\wedge)}|2) = 1$, é destacado.

3.3.2.3 Comentários sobre as distribuições posteriori obtidas

Os gráficos das distribuições posteriori para as t-normas do produto (Eq. 3.3.10) e do mínimo (Eq. 3.3.13), respectivamente calculadas nas seções 3.3.2.1 e 3.3.2.2, utilizando distribuição priori (3.3.1) e função de verossimilhança (3.3.7) obtidas a partir de um registro histórico na seção 3.3.1, são mostrados abaixo nas figuras 3.6 e 3.8. Pelo fato das distribuições que originaram a posteriori estarem no mesmo gráfico, é possível compreender melhor a noção de que a posteriori seria um “acordo” (conforme exposto

em 2.4.3.1) entre a informação da distribuição priori e a da verossimilhança.

No final da seção 3.3.2, após algumas considerações simples, parecia correto afirmar que o valor $\bar{x}^* = 2.89$ (valor mais relevante da verossimilhança, Eq. 3.3.7) não seria o mais relevante da distribuição posteriori de X . Tal intuição é confirmada quando se percebe que o valor mais relevante (cuja pertinência vale “1”) de ambas as distribuições posteriori (figuras 3.6 e 3.8, acima) *não é* $\bar{x}^* = 2.89$.

Baseando-se nos valores mais relevantes de ambos os casos da posteriori, “ $\bar{x}^{(\cdot)} = 4.4$ ” para a t-norma do produto e “ $\bar{x}^{(\wedge)} = 3.85$ ” para a do mínimo (ambos destacados em seus respectivos gráficos), conclui-se então que, de acordo com a metodologia:

“Para uma previsão $Y = 2$, o registro histórico sugere que se espera que a variável X assuma um valor *maior que o da previsão.*”

3.4 Metodologia para obtenção de funções a partir do registro histórico

Para que fosse possível concentrar-se mais na metodologia proposta, no exemplo da seção 3.3 procedeu-se da seguinte forma para obter as duas funções necessárias para a aplicação do método:

- Utilizou-se a informação do especialista para modelar a distribuição priori $\pi(x)$ na seção 3.3.1.1;
- O gráfico utilizado para obter a função de verossimilhança $\pi(y|x)$ na seção 3.3.1.2 (Figura 3.2) *já continha as retas-limite superior e inferior* (em pontilhado).

Esta seção tem por objetivo propor metodologias para, a partir de um conjunto genérico de pontos $X \times Y$ e seu ajuste linear de mínimos quadrados correspondente, obter a distribuição priori $\pi(x)$ e a função de verossimilhança $\pi(y|x)$. Para ilustrar a metodologia, as duas funções (distribuição priori e função de verossimilhança) serão obtidas para os dados do exemplo da seção 3.3, representados na Figura 3.9. Dessa

forma, a priori e a verossimilhança obtidas através da metodologia proposta nesta seção devem coincidir exatamente com as obtidas nas seções 3.3.1.1 e 3.3.1.2, respectivamente equações (3.3.1) e (3.3.6).

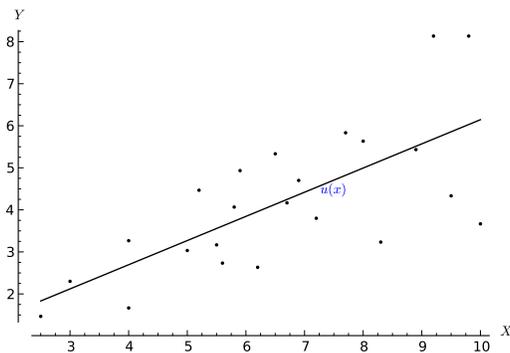


Figura 3.9: Gráfico $X \times Y$, com os pontos e o ajuste linear correspondentes aos dados do exemplo da seção 3.3.

3.4.1 Metodologia para obtenção da distribuição priori

No exemplo da seção 3.3.1.1, utilizou-se um número trapezoidal Fuzzy para representar $\pi(x)$, a função marginal de X . Com o objetivo de ter uma variação maior nos resultados oferecidos, escolheu-se, além da função trapezoidal, uma outra forma de representar a distribuição marginal de X : uma parábola com concavidade para baixo. A parábola foi escolhida como opção pelo fato de ser mais suave que uma função trapezoidal, assim presumivelmente mais sensível a variações bruscas. Os cálculos do método serão realizados para as duas distribuições priori separadamente. Dessa forma, tem-se as seguintes opções de funções para modelar a distribuições priori:

1. $\pi^{(1)}(x)$: Um número fuzzy trapezoidal, definido na seção 2.1;
2. $\pi^{(2)}(x)$: Uma parábola normalizada, definida a seguir nesta seção.

Primeiro caso: Trapézio

Uma ferramenta estatística apropriada para a obtenção da distribuição priori é o *histograma*³. Para gerar o histograma correspondente, utiliza-se apenas as medidas experimentais X , ou seja, apenas a primeira coordenada de cada ponto da figura 3.9. A quantidade e o tamanho das barras do histograma foram ajustadas de forma ótima pelo programa utilizado.

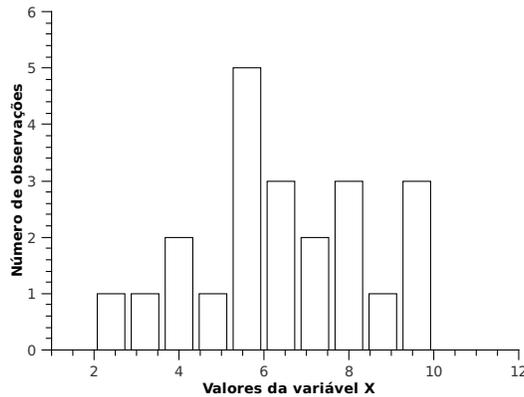


Figura 3.10: Histograma da variável X correspondente aos dados desta seção.

Informações importantes também podem ser obtidas a partir de um *histograma cumulativo*, que mostra de que forma o total de valores estão distribuídos. Um histograma cumulativo para os dados de exemplo desta seção é mostrado a seguir na figura 4.2b:

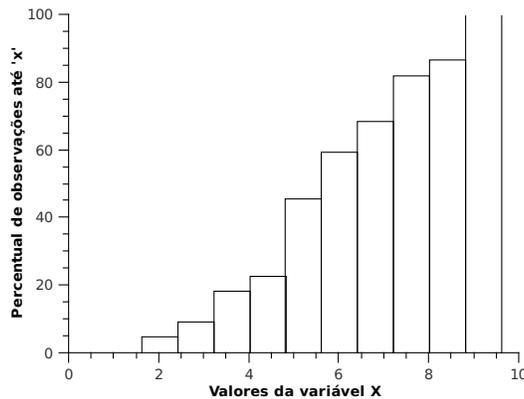


Figura 3.11: Histograma cumulativo da variável X correspondente aos dados desta seção.

³Para gerar os histogramas necessários utilizou-se o programa de software livre *Qtplot* (<http://soft.proindependent.com/qtplot.html>), análogo ao programa proprietário Origin®

Um histograma cumulativo pode servir de grande ajuda, pois informa mais facilmente o quanto cada uma das barras contribui na quantidade total de pontos. A primeira opção de distribuição priori é um número fuzzy trapezoidal. Para obtê-lo, é preciso escolher os parâmetros a , u , v e b de acordo com a equação (2.1.3), repetida abaixo. Observando a figura 3.11, nota-se que a maior quantidade de valores (o maior salto do histograma cumulativo) ocorre entre os valores $X = 5$ e $X = 8$ (40% dos valores está neste intervalo, pois a porcentagem varia de 40 a 80%). Assim, escolhe-se que $u = 5$ e $v = 8$. De forma a englobar os outros valores, escolhe-se $a = 2$ e $b = 10$. Então, o primeiro caso de distribuição priori, $\pi^{(1)}(x)$, é descrito pelo número fuzzy trapezoidal de parâmetros $(2, 5, 8, 10)$. Substituindo o valor na equação do número fuzzy trapezoidal, obtém-se

$$\text{trapez}(a, u, v, b)(x) = \begin{cases} \frac{(x-a)}{u-a} & \text{se } a \leq x < u \\ 1 & \text{se } u \leq x \leq v \\ \frac{(b-x)}{b-v} & \text{se } v \leq x < b \\ 0 & \text{caso contrário} \end{cases} \quad (3.4.1)$$

$$\pi^{(1)}(x) = \text{trapez}(2, 5, 8, 10)(x) = \begin{cases} \frac{(x-2)}{3} & \text{se } 2 \leq x < 5 \\ 1 & \text{se } 5 \leq x \leq 8 \\ \frac{(10-x)}{2} & \text{se } 8 \leq x < 10 \\ 0 & \text{caso contrário} \end{cases}, \quad (3.4.2)$$

que é exatamente a equação (3.3.1), obtida com o auxílio de um especialista na seção 3.3. O gráfico da priori trapezoidal segue na figura 3.12.

Segundo caso: Parábola

Também pode-se descrever a distribuição marginal de X por uma parábola com concavidade para baixo, onde suas duas raízes são os mesmos valores extremos do caso trapezoidal (Eq. 3.4.2): $a = 2$ e $b = 10$. A parábola é normalizada de forma a ter “1”

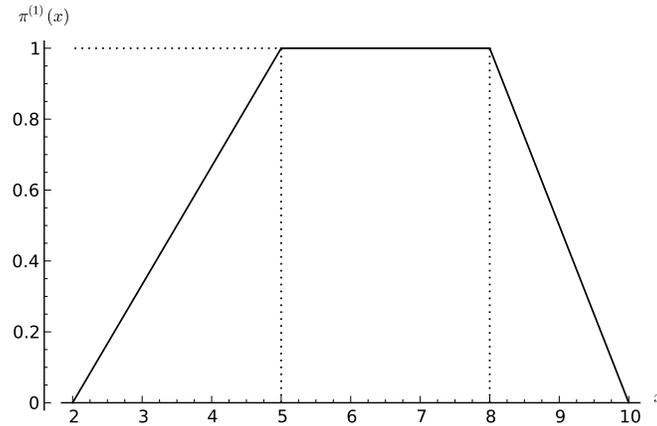


Figura 3.12: Primeira escolha de distribuição priori: $\pi_1(x) = \text{trapez}(2,5,8,10)$ (Eq. 3.4.2)

como valor máximo assumido. A equação da parábola é dada a seguir:

$$\text{parabolaNorm}(a, b)(x) = \frac{(x - a)(x - b)}{\left(\frac{a+b}{2}\right)^2} \quad (3.4.3)$$

$$\pi^{(2)}(x) = \text{parabolaNorm}(2, 10)(x) = \frac{(x - 2)(x - 10)}{36}. \quad (3.4.4)$$

O gráfico da priori trapezoidal segue na Figura 3.12.

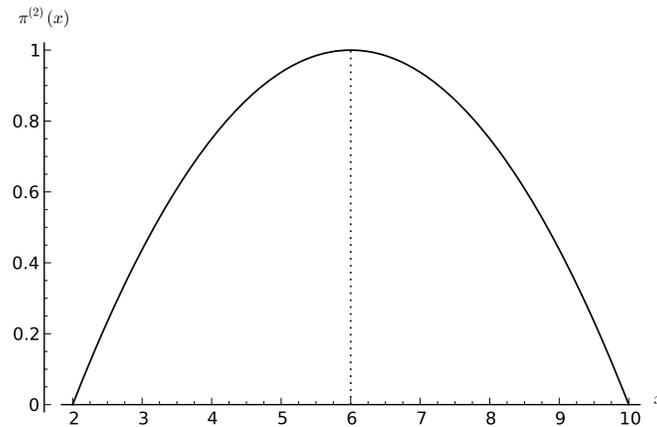


Figura 3.13: Segunda escolha de distribuição priori: $\pi_2(x) = \text{parabolaNorm}(2,10)$ (Eq. 3.4.4)

Uma observação relevante: apesar da parábola ser mais suave (portanto presumivelmente mais sensível a mudanças), tem-se que ela é simétrica, com seu vértice exatamente no meio das duas raízes (isto é, entre a temperatura mínima e a máxima). Uma função

trapezoidal não necessariamente é simétrica, como por exemplo a da figura 3.13. O que pode ser uma vantagem: é possível uma situação onde exista uma variação pequena abaixo da temperatura média, mas uma variação grande acima da média. Tal situação, a princípio, é difícil de se descrever usando uma parábola, mas poderia ser feito sem muitos problemas usando uma função trapezoidal. A maior liberdade de escolha que a função trapezoidal oferece é uma vantagem muito grande, mas por outro lado a suavidade da parábola pode oferecer maior robustez ao método. Isso faz com que as duas escolhas sejam de certa forma complementares na abordagem, o que justifica a escolha de trabalhar com as duas.

3.4.2 Metodologia para obtenção da função de verossimilhança

A partir do gráfico da Figura 3.9, que contém os pontos experimentais e seu ajuste linear de mínimos quadrados correspondente, serão obtidas as retas-limite. As duas retas devem “englobar” todos os pontos experimentais do gráfico da Figura 3.9, e além disso devem dar alguma “folga” aos dados. Essa “folga” é o que permite que o sistema se adapte bem à situações novas. A questão mais importante no ajuste das retas limite é justamente a de definir *o quanto* de folga deve ser dado aos dados. Mais adiante esta questão será vista em detalhes.

Como foi dito, as retas-limite devem englobar todos os pontos do gráfico. A seguir descreve-se um procedimento para englobar todos os pontos acima do ajuste linear:

- Para cada ponto p_i que está *acima* do ajuste linear, aumenta-se o coeficiente angular do ajuste de forma a interceptar o ponto p_i , obtendo o coeficiente angular a_i^4 ;
- O ponto cujo coeficiente angular for o maior entre todos os pontos acima do ajuste linear é chamado \hat{p}_M , e o coeficiente é chamado \hat{a}_M .

⁴Ou seja, se a função do ajuste linear for dada por $u(x) = A \cdot x + B$, e se $p_i = (x_i, y_i)$, o coeficiente angular a_i é dado por: $a_i = \frac{y_i - B}{x_i}$

Ao proceder desta forma, a função “ $y_{MAX} = \hat{a}_M \cdot x + B$ ”, onde B é o coeficiente linear do ajuste linear “ $u(x) = A \cdot x + B$ ”, engloba superiormente todos os pontos do gráfico. Se os coeficientes a_i forem obtidos exatamente da mesma forma do procedimento acima para os pontos *abaixo* do ajuste linear, e o ponto cujo coeficiente angular for o menor entre todos os pontos abaixo do ajuste for denominado \hat{p}_m (de coeficiente angular \hat{a}_m), a função “ $y_{min} = \hat{a}_m \cdot x + B$ ” engloba inferiormente todos os pontos do gráfico. A figura 3.14 a seguir ilustra as funções y_{MAX} e y_{min} , bem como os pontos \hat{p}_M e \hat{p}_m :

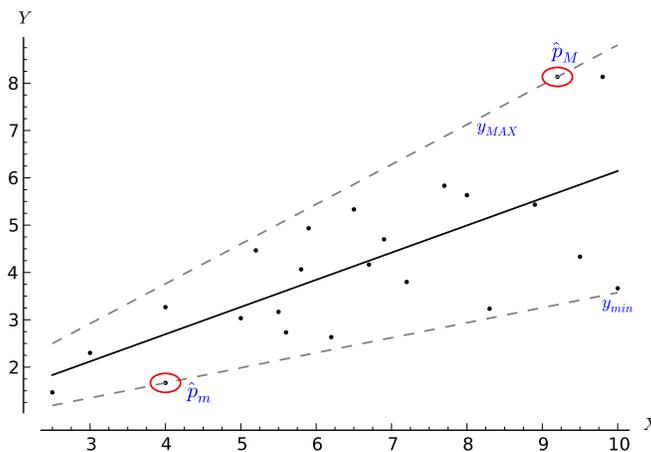


Figura 3.14: Gráfico $X \times Y$, correspondente aos dados desta seção. Contém além disso as retas y_{MAX} e y_{min} , explicadas nesta seção, bem como os pontos \hat{p}_M e \hat{p}_m .

Volta-se então à questão da “folga” dada aos dados pelas retas-limite. Deve-se notar que na Figura 3.14, as retas y_{MAX} e y_{min} interceptam exatamente os pontos mais distantes do ajuste linear (isto é, os pontos \hat{p}_M e \hat{p}_m), logo as retas y_{MAX} e y_{min} não oferecem nenhuma “folga” aos dados. Uma das formas de definir a citada folga é adicionar algum valor ao coeficiente angular de y_{MAX} , e (analogamente) diminuir algum valor ao coeficiente angular de y_{min} .

Para determinar esta “quantidade” que irá definir as folgas superiores e inferiores, é razoável supor que esta quantidade deve ser um múltiplo de $|\hat{a}_M - A|$ para y_{MAX} e $|\hat{a}_m - A|$ para y_{min} . Ou seja, a “folga” será proporcional ao quanto se alterou o coeficiente angular do ajuste linear para que se pudesse ajustar y_{MAX} e y_{min} . A constante que definirá esta proporcionalidade será chamada C , de forma que o coeficiente angular

de y_{MAX} receberá um aumento de “ $C \cdot |\hat{a}_M - A|$ ”, enquanto o de y_{min} receberá uma diminuição de “ $C \cdot |\hat{a}_m - A|$ ”. Dessa forma, as retas-limite $b(x)$ (superior) e $a(x)$ (inferior) são definidas por:

$$a(x) = (\hat{a}_m - C \cdot |\hat{a}_m - A|)x + B \quad (3.4.5)$$

$$b(x) = (\hat{a}_M + C \cdot |\hat{a}_M - A|)x + B \quad (3.4.6)$$

Na Figura 3.15 abaixo, seguem as retas-limite $b(x)$ e $a(x)$ para dois casos de constante C : $C = 0$ e $C = 1.5$.

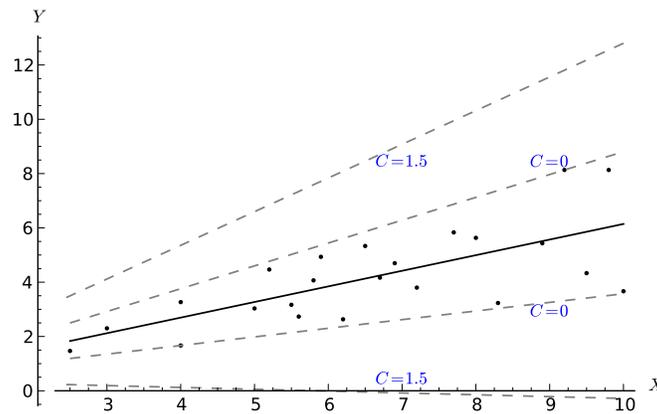


Figura 3.15: Gráfico $X \times Y$ com as respectivas retas-limite $b(x)$ e $a(x)$, definidas de acordo com as fórmulas 3.4.5 e 3.4.6, para dois casos de constante C , $C = 0$ e $C = 1.5$.

Observando as figuras em 3.15 acima, deve-se notar que a constante C deve ser obrigatoriamente positiva ($C \geq 0$), pois senão as retas-limite não englobarão todos os pontos experimentais. No caso $C = 0$, $b(x) = y_{MAX}(x)$ e $a(x) = y_{min}(x)$ (basta substituir $C = 0$ em 3.4.5 e 3.4.6).

Resta então escolher qual será a constante C de acordo com algum critério razoável. Experimentos numéricos sugeriram que:

- Nenhuma folga (Figura 3.15 para $C = 0$) não é recomendada, no sentido que torna o modelo pouco tolerante à novas situações. Sendo assim, folgas muito pequenas ($C \approx 0$) mostram-se incapazes de incorporar a generalidade da situação estudada,

tornando o modelo assim inadequado para melhorar previsões de um especialista de forma satisfatória;

- Valores de C que se aproximam de 2 (Figura 3.15 para $C = 1.5$) oferecem, em geral, uma folga exagerada. Isso faz com que o modelo seja excessivamente tolerante, tornando-o assim incapaz de diferenciar previsões que retratam o comportamento usual da situação estudada de previsões que representam anormalidades ou situações-limite. Dessa forma, não pode-se esperar que as previsões de um especialista sejam melhoradas por um modelo com excesso de folga.

De acordo com os dois pontos acima, conclui-se que $0 < C < 2$, com C mais próximo de 1 do que de 2. Uma hipótese razoável é supor que exista alguma relação entre a qualidade do ajuste linear e a quantidade de folga oferecida, no sentido que quanto melhor for o ajuste linear, menor a folga. Uma forma de avaliar a qualidade de um ajuste linear é utilizando o coeficiente $R = r^2$, onde r é o coeficiente de correlação linear de Pearson[17, p.182]. $R = r^2$ é chamado *coeficiente de determinação* (e $1 - R$ chamado de *coeficiente de indeterminação*), pois seus valores são indicativos de quanto o ajuste linear fica bem determinado em função da correlação entre os pontos experimentais. Assim, por exemplo, no caso ideal em que $R=1$, não haveria variação residual, e todos os pontos estariam alinhados[17, p.200]. Dessa forma, é razoável supor que a constante C envolva o termo $(1 - R)$. Assim, quanto melhor o ajuste linear, mais C se aproxima de 0, e quanto pior for o ajuste, mais C se aproxima de 1. Além disso, melhores resultados se apresentaram se for oferecida sempre um mínimo de folga, independentemente do coeficiente R . Este valor de mínimo, chamado de m , permite também que o valor de C sempre seja estritamente maior que 0 (um dos dois pontos levantados acima). Assim, é seguro definir que a constante positiva C seja definida por:

$$C = (1 - R) + m, \tag{3.4.7}$$

onde R é o coeficiente de determinação do ajuste linear do gráfico e $m > 0$.

Tendo em mente o segundo ponto levantado acima, conclui-se que m deve ser um valor pequeno e menor que 1 (pois $0 \leq 1 - R \leq 1$). Dentre os testes realizados, $m = 0.10$ apresentou, de forma geral, os melhores resultados. Substituindo $m = 0.10$ em 3.4.7, utilizou-se neste trabalho, portanto, o valor de C a seguir:

$$C = (1 - R) + 0.10. \quad (3.4.8)$$

Substituindo C da equação 3.4.8 acima nas equações das retas-limite 3.4.5 e 3.4.6, obtém-se as equações das duas retas limite:

$$a(x) = (\hat{a}_m - (1.1 - R)|\hat{a}_m - A|)x + B \quad (3.4.9)$$

$$b(x) = (\hat{a}_M + (1.1 - R)|\hat{a}_M - A|)x + B \quad (3.4.10)$$

No caso particular dos dados deste exemplo, os valores numéricos correspondentes são:

$$\begin{array}{ll} \text{Ajuste linear:} & u(x) = 0.57 \cdot x + 0.40 \\ & A = 0.57 \\ & B = 0.40 \\ \text{Constantes:} & \hat{a}_M(x) = 0.84 \\ & \hat{a}_m(x) = 0.32 \\ & R = 0.50 \\ & C = 0.60 \end{array}$$

Substituindo os valores acima nas equações 3.4.10 e 3.4.9, obtém-se:

$$a(x) = 0.16 \cdot x + 0.40 \quad (3.4.11)$$

$$b(x) = x + 0.40 \quad (3.4.12)$$

Onde o ajuste linear $u(x)$ é dado por:

$$u(x) = 0.57 \cdot x + 0.40 \quad (3.4.13)$$

O gráfico para os dados de exemplo com as retas-limite correspondentes segue na Figura 3.16.

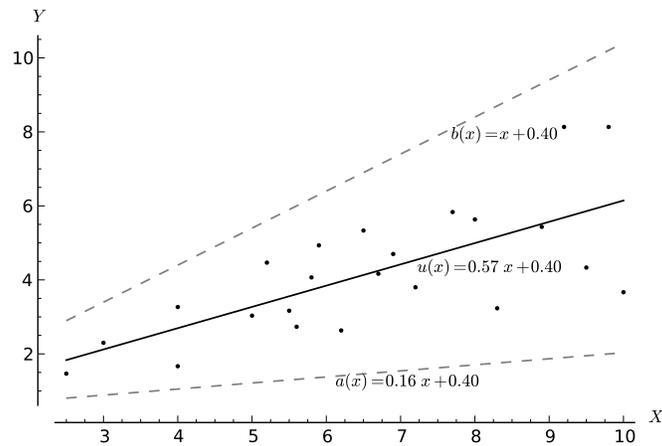


Figura 3.16: Gráfico $X \times Y$, com o ajuste linear correspondente aos dados desta seção, com as retas-limite $a(x)$ e $b(x)$ definidas de acordo com a metodologia proposta nesta seção (Eq. 3.4.12 e 3.4.11).

Deve-se notar que as retas-limite da Figura 3.16 são exatamente as mesmas da figura 3.2 do exemplo. Dessa forma, substituindo $u = u(x)$ (Eq. 3.4.13), $a(x)$ (Eq. 3.4.11) e $b(x)$ (Eq. 3.4.12) na equação (3.3.5), obtém-se exatamente a equação (3.3.6).

A metodologia descrita neste tópico, que recebe um conjunto de pontos (x, y) e retorna o ajuste linear $u(x)$ e as retas-limite $a(x)$ e $b(x)$, foi implementada computacionalmente. Os comandos que obtém as retas-limite referentes a este exemplo podem ser encontrados na seção B.1.1.1.

Capítulo 4

Aplicação

4.1 Sobre os dados utilizados e sua organização

Os dados referem-se à temperatura e umidade relativa referentes à lavouras de café localizadas no sul de Minas Gerais. Com o objetivo de melhor adaptar a cultura do grão ao aumento da temperatura global, pode-se consorciá-lo com outras espécies vegetais com o objetivo de criar um microclima mais propício ao cultivo do café. Essa técnica é chamada de Sistema Agro-Florestal (SAF).

Dessa forma, além da técnica usual de plantio que consiste apenas do cafeeiro, chamada pleno sol (“PS”), a técnica SAF foi utilizada para criar três microclimas adicionais para o cultivo de café: “SAF1”, “SAF2” e “SAF3”. Em cada microclima foram feitas, a cada hora, medidas de temperatura e umidade. Cada uma dessas medidas têm uma previsão correspondente, pois cada microclima tem suas características físicas descritas no software de modelagem de microclimas ENVI-met, permitindo assim que se possa simular os valores de temperatura e umidade (ver seção 1.1).

Aplicou-se a metodologia descrita no capítulo 3 com o objetivo de melhorar as previsões de temperatura e umidade fornecidas pelo software ENVI-met. Isto é, as previsões do ENVI-met de um determinado microclima são as chamadas “previsões do especialista” que se pretende revisar levando em conta um registro histórico sobre o

comportamento das previsões e medidas neste dado microclima.

4.1.1 Origem dos dados fornecidos

Os dados referem-se ao trabalho[20], proveniente da Faculdade de Engenharia Agrícola (UNICAMP). A instalação das estações meteorológicas foi financiada pela EMBRAPA CAFÉ e FUNARBE (Fundação Arthur Bernardes). O local onde foram realizadas as medidas é a fazenda da EPAMIG (Empresa de Pesquisa Agropecuária de Minas Gerais), localizada no Município de São Sebastião do Paraíso, sul de Minas Gerais.

4.1.2 Organização dos dados utilizados

Os dados fornecidos pelo especialista consistem de um conjunto de pares (x^*, y^*) , onde x^* é uma medida experimental da variável temperatura ou umidade, e y^* é a previsão correspondente à medida x^* . Um “Conjunto de dados” corresponde às medidas e previsões de uma variável (temperatura ou umidade), em um dos quatro microclimas (“PS”, “SAF1”, “SAF2” ou “SAF3”), numa determinada época do ano (verão ou inverno). Portanto, dois exemplos de conjuntos de dados seriam “Temperatura – Inverno – PS” (tópico A.2.1 do apêndice) e “Umidade – Verão – SAF1” (tópico A.3.2 do apêndice). Ou seja, cada conjunto de dados corresponde à uma combinação destes três fatores, representados na Tabela 4.1 abaixo.

Tabela 4.1: Combinação dos fatores que dividem os dados fornecidos pelo especialista em conjuntos de dados menores.

Fator	Escolhas possíveis	Número de opções
Variável	Temperatura Umidade	2
Época do ano	Verão Inverno	2
Configuração de plantio (microclima)	PS SAF1 SAF2 SAF3	4
Total:		16 combinações

Cada um dos 16 conjuntos de dados definidos acima consiste de 3 dias de medidas

experimentais e de previsões. Para cada um destes dias são feitas, a cada hora, medições da variável em questão (temperatura ou umidade). Cada medida têm uma previsão correspondente, gerada pelo *software* ENVI-met (ver seção 1.1). Dessa forma, cada dia de um conjunto de dados corresponde a 24 pares da forma “medida/previsão”, e cada conjunto de dados corresponde a 72 pares.

4.2 Metodologia de testes

Nesta seção é descrita a forma que, utilizando os dados fornecidos pelo especialista, testou-se a metodologia proposta. Desta forma, um teste é bem sucedido quando as previsões obtidas pela metodologia proposta aproximam-se mais dos dados experimentais que as previsões do especialista. Existem medidas estatísticas especificamente criadas para avaliar o quanto um conjunto de previsões aproxima-se dos dados experimentais, e duas dessas medidas foram utilizadas neste trabalho (mais detalhes na seção 4.3).

4.2.1 Separação dos dados em grupo de treinamento e grupo de testes

Uma forma de realizar testes em um determinado conjunto de dados é separá-lo em dois subconjuntos (grupos) disjuntos: grupo de treinamento e de testes. Dessa forma, o modelo utilizado é “treinado” utilizando os dados do grupo de treinamento, e seus resultados são avaliados no grupo de testes. Portanto:

- **Grupo de treinamento:** Dados que serão utilizados para obter os parâmetros necessários à aplicação do método (calibração do método).
- **Grupo de testes:** Dados utilizados com o objetivo de avaliar como a metodologia se comporta frente a novas informações, pois tais dados não foram usados no treinamento (calibração) do modelo.

Esta separação dos dados é interessante na medida que, como os grupos de treinamento e testes são disjuntos (não têm nenhum dado em comum), pode-se observar como

o método proposto lida com situações para as quais não foi preparado (treinado). Dessa forma, se o modelo for capaz de reproduzir bem o comportamento do grupo de testes, ele foi capaz de “captar” o comportamento global do fenômeno em questão usando as informações de treinamento.

Como cada conjunto de dados corresponde a três dias de medidas e previsões (ver seção 4.1.2), para cada conjunto de dados foram feitos três testes, onde em um deles o grupo de testes foi um dos três dias, conforme a Tabela 4.2.

Tabela 4.2: Os três testes a serem realizados em cada conjunto de dados.

Grupo de testes	Grupo de treinamento
1º dia	2º e 3º dias
2º dia	1º e 3º dias
3º dia	1º e 2º dias

Logo, a divisão dos dados para cada conjunto de dados é de 66/33%, ou seja, dois dias de treino para um de testes. Esta proporção é reconhecida pela literatura[7, p. 268] como uma prática segura.

4.2.2 As quatro formas de realizar os cálculos do método

No exemplo da seção 3.3, o fato de se trabalhar sempre com duas t-normas (mínimo e produto) fez com que, para uma previsão do especialista, o método proposto retornou *duas distribuições posteriori distintas*. Isto é, a priori e a verossimilhança utilizadas para obter as distribuições posteriori eram exatamente as mesmas, a única diferença foi que uma delas foi obtida através da t-norma do mínimo (Eq. 3.3.13, dada pela Eq. 3.2.15a) e a outra da do produto (Eq. 3.3.10, dada pela Eq. 3.2.15b).

Escolheu-se trabalhar simultaneamente com duas distribuições posteriori no exemplo citado acima para que se pudesse ter uma variabilidade maior de resultados. E, uma vez que a distribuição posteriori é obtida em função de uma t-norma arbitrária T (conforme Eq. 3.2.10), uma forma natural de gerar resultados variados foi escolhendo trabalhar com mais de uma t-norma. Com o objetivo de obter uma variabilidade ainda maior que no exemplo da seção 3.3, ao lidar com os dados reais deste capítulo propõe-se que,

além de se trabalhar com as duas t-normas do mínimo e do produto, que se trabalhasse também com duas opções de distribuição priori:

1. Uma função trapezoidal (o mesmo tipo de função que foi usada no exemplo da seção 3.3), conforme Eq. (3.4.1);
2. Uma parábola com concavidade para baixo normalizada (de forma que seu valor máximo fosse 1), conforme Eq. (3.4.3);

Dessa forma, para cada previsão do especialista obteve-se 4 distribuições posteriori distintas, cada uma correspondente a uma combinação de t-norma e de priori, como ilustra a Tabela 4.3.

Tabela 4.3: As quatro combinações priori/t-norma, que, para uma mesma previsão, geram distribuições posteriori distintas.

Configuração	Priori	t-norma
1	Trapézio	Produto (\cdot)
2	Parábola	Produto (\cdot)
3	Trapézio	Mínimo (\wedge)
4	Parábola	Mínimo (\wedge)

4.2.3 Sobre o número de testes realizados

Conforme definido na seção 4.1.2, cada conjunto de dados corresponde a uma combinação de três fatores: variável (temperatura ou umidade), configuração de plantio (PS, SAF1, SAF2 ou SAF3) e estação do ano (Verão ou Inverno), totalizando 16 conjuntos de dados (ver Tabela 4.1). Cada conjunto continha (nessa ordem) as medidas experimentais, as previsões do especialista e as quatro previsões retornadas pelo método proposto, correspondentes às quatro formas de obter a distribuição posteriori, conforme a Tabela 4.3.

Dessa forma, como para cada um dos 16 conjuntos de dados foram feitos três testes distintos (ver Tabela 4.2), sendo portanto realizados

$$16 \cdot 3 = 48 \text{ testes.}$$

Durante a realização dos testes, notou-se que *não foi possível realizar os testes para os dados de “Umidade – Verão”*, para nenhuma das quatro configurações de plantio. Isto é, não foram realizados testes para as seções A.3.1, A.3.2, A.3.3. A.3.4 do apêndice. A seção 4.7.1 é dedicada à explicação dos fatores que fizeram com que isso ocorresse, sugerindo algumas medidas que resolveriam este problema. Como cada um dos 4 conjuntos de dados corresponde a três testes, diminui-se 12 testes dos 48 iniciais, de forma que na presente abordagem foram realizados $T = 36$ testes.

4.2.4 Calibração necessária para a realização de testes

Como exposto na seção 4.2.1, para cada conjunto de dados (conforme seção 4.1.2), o modelo proposto foi calibrado utilizando seus dados do grupo de treinamento. O esquema de aplicação do método, conforme a Figura 3.1, indica que, neste trabalho, “calibração” significa definir a distribuição priori e a verossimilhança.

Então, dado um conjunto de dados e qual teste se pretende realizar, define-se o grupo de treinamento do teste correspondente (conforme a Tabela 4.2). A partir daí foi obtida a verossimilhança $\pi(y|x)$ e as distribuições priori $\pi^{(1)}(x)$ e $\pi^{(2)}(x)$ (lembrando que trabalha-se com duas distribuições priori: trapézio e parábola).

Obtendo a função de verossimilhança O grupo de treinamento, composto por 48 pares “medida/previsão” (dados correspondentes a dois dias), foi utilizado para obter a função de verossimilhança conforme a metodologia da seção 3.4.2. Ou seja, os dados do grupo de treinamento foram utilizados para gerar um gráfico $(X \times Y)$, e a partir deste gráfico obteve-se a função de verossimilhança.

Obtendo as distribuições priori Como uma distribuição priori representa o comportamento da variável X (seja ela temperatura ou umidade) de uma forma geral, no sentido que representa quais são os valores esperados e possíveis para X , escolhe-se utilizar as informações tanto do grupo de testes quanto do grupo de treinamento. Dessa forma, utiliza-se a primeira coordenada de todos os pares do conjunto de dados atual

(72 pares, dados correspondentes aos três dias) para obter as duas distribuições priori (trapézio e parábola, conforme seção 4.2.2) de acordo com a metodologia proposta na seção 3.4.1. Assim, as duas distribuições priori são as mesmas para um mesmo conjunto de dados, não importando qual é o grupo de treinamento.

4.2.5 Sobre a realização dos testes

Escolhido um conjunto de dados e qual dos três dias seria o grupo de testes, foi definido também (Tabela 4.2) o grupo de treinamento correspondente (ver seção 4.2.1). Obtendo a distribuição priori e a verossimilhança para este grupo de treinamento (conforme seção 4.2.4), pode-se portanto aplicar o método às previsões do grupo de testes.

De acordo com a seção 4.2.2, cada previsão corresponde a quatro distribuições posteriores distintas (ver Tabela 4.3). Dessa forma, o grupo de testes, que consiste de 24 pares “medida/previsão” (um dia de dados), tem quatro distribuições posteriores para cada uma das 24 previsões do dia de testes. Portanto, se o método proposto foi aplicado à previsão y_{i^*} do grupo de testes ($i^* = 1, \dots, 24$), obteve-se quatro distribuições posteriores correspondentes a essa previsão.

Avaliou-se se o método foi capaz de melhorar a previsão do especialista y_{i^*} comparando suas quatro distribuições posteriores à medida experimental x_{i^*} (medida correspondente à previsão y_{i^*}). Esta comparação foi possível utilizando *somente o valor de máximo de cada posteriori*, como os valores do exemplo da seção 3.3, “ $\bar{x}^{(\cdot)} = 4.4$ ” e “ $\bar{x}^{(\wedge)} = 3.85$ ” (respectivamente destacados nas figuras 3.6 e 3.8). De forma similar definiu-se então que $\bar{x}_{i^*}^{(1)}$ é o ponto de máximo da “configuração 1” (trapézio/produto, conforme Tabela 4.3) da posteriori correspondente à previsão y^* , e assim por diante definiu-se de forma análoga os pontos de máximo das outras configurações, $\bar{x}_{i^*}^{(2)}$, $\bar{x}_{i^*}^{(3)}$ e $\bar{x}_{i^*}^{(4)}$.

Todos os testes realizados estão contidos no apêndice A. Cada um dos três dias corresponde a 24 linhas do apêndice, com cada linha contendo 6 colunas. A Figura 4.1 mostra como as informações citadas nesta seção se dispõem no apêndice (destacando a

linha da previsão y_{i^*} , citada acima):

	Exp	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
linha i^* :	x_{i^*}	y_{i^*}	$\bar{x}_{i^*}^{(1)}$	$\bar{x}_{i^*}^{(2)}$	$\bar{x}_{i^*}^{(3)}$	$\bar{x}_{i^*}^{(4)}$
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

Figura 4.1: Padrão de colunas de um conjunto de dados do apêndice. (x_{i^*}, y_{i^*}) é o par “medida/previsão”, e $\bar{x}_{i^*}^{(1)}$, $\bar{x}_{i^*}^{(2)}$, $\bar{x}_{i^*}^{(3)}$ e $\bar{x}_{i^*}^{(4)}$ são as previsões revisadas pelo método obtidas respectivamente pela configuração 1, 2, 3 e 4 (segundo Tabela 4.3).

O padrão de colunas da Figura 4.1 acima foi mantido para todas as 72 linhas de cada conjunto de dados, agrupando estas linhas em três grupos de 24 (três dias).

4.3 Medidas estatísticas para análise dos testes

Suponha que os cálculos do método proposto tenham sido realizados para algum teste de um conjunto de dados, de acordo com a seção 4.2, para uma das 4 configurações de obtenção da posteriori (ver seção 4.2.2). Dessa forma, dispõe-se de três colunas de 24 linhas cada: coluna **(1)** corresponde às medidas experimentais, **(2)** às simulações dadas pelo especialista (onde (1) e (2) são respectivamente a primeira e segunda colunas da Figura 4.1), e **(3)** às simulações do especialista modificadas pelo método proposto (por exemplo, se for escolhida a “Configuração 2”, é a quarta coluna da Figura 4.1).

Uma forma de saber se o método comportou-se bem no teste realizado é verificar se as previsões do método (3) aproximaram mais dos dados experimentais (1) que as previsões do especialista (2). Entre as possíveis métricas usadas para comparar colunas dessa forma, uma característica importante é que a métrica escolhida lide bem com a ocorrência de *outliers* (pontos muito distantes da média), muito comuns em grandezas oriundas de problemas ambientais (como temperatura e umidade). Existem medidas estatísticas especificamente criadas com o objetivo de comparar medidas experimentais com previsões. Duas delas foram utilizadas neste trabalho, descritas a seguir.

4.3.1 \mathcal{D} de Willmott

Esta é uma métrica proposta especificamente com o objetivo de avaliar melhor a precisão de um dado modelo (ou seja, a precisão das simulações retornadas por esse modelo) aos dados experimentais¹ (ver [16]). Tal coeficiente baseia-se essencialmente no *erro médio relativo* $E \in [0, 1]$, definido por:

$$E(\mathbf{x}, \mathbf{z}) = \frac{\sum_j |z_j - x_j|^2}{\sum_j (|z_j - \bar{x}| + |x_j - \bar{x}|)^2}, \quad (4.3.1)$$

onde $\mathbf{x} = (x_1, x_2, \dots, x_n)$ é um vetor de medidas experimentais, $\mathbf{z} = (z_1, z_2, \dots, z_n)$ é um vetor de simulações e \bar{x} é a média do vetor \mathbf{x} . Nota-se que o vetor de medidas \mathbf{x} deve ser sempre a primeira coordenada, pois E não é simétrica nas variáveis (em geral $E(\mathbf{x}, \mathbf{z}) \neq E(\mathbf{z}, \mathbf{x})$).

Tem-se que $E = 1$ é o maior valor de erro que se pode ter nessa medida. Ou seja, para $E = 1$, a aproximação da simulação aos dados experimentais (de acordo com essa medida estatística) é a pior possível. Dessa forma, para definir um índice que represente o quanto um vetor \mathbf{z} de simulações aproxima-se de um vetor de medidas \mathbf{x} , define-se o coeficiente \mathcal{D} de Willmott a partir do erro médio relativo (4.3.1):

$$\mathcal{D}(\mathbf{x}, \mathbf{z}) = 1 - E(\mathbf{x}, \mathbf{z}) = 1 - \frac{\sum_j |z_j - x_j|^2}{\sum_j (|z_j - \bar{x}| + |x_j - \bar{x}|)^2}, \quad (4.3.2)$$

e dessa forma, se $\mathcal{D}(\mathbf{x}, \mathbf{z}) = 1$ tem-se que $E(\mathbf{x}, \mathbf{z}) = 0$, logo a previsão é a melhor possível.

¹Mais precisamente, o coeficiente \mathcal{D} de Willmott foi proposto com o objetivo de avaliar simulações dadas por problemas da área de geofísica.

4.3.2 MAPE

O erro percentual absoluto médio (MAPE, Mean Absolute Percentage Error)² expressa precisão como uma porcentagem, definida pela fórmula 4.3.3:

$$M = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| = \frac{100\%}{n} \sum_{t=1}^n \left| 1 - \frac{F_t}{A_t} \right|, \quad (4.3.3)$$

onde A_t é o valor verdadeiro (medida) e F_t é o valor previsto.

A diferença entre A_t e F_t é dividida pelo valor verdadeiro A_t novamente. O valor absoluto neste cálculo é somado para todo ponto previsto ou ajustado e dividido novamente pelo número de pontos n . Multiplicando o resultado por 100 torna MAPE um erro percentual. Quanto mais próximo de zero é o valor do MAPE, menor o erro, logo melhor é a previsão³.

Muito embora o conceito do MAPE seja simples e convincente, há dois problemas relevantes na sua aplicação prática[18]:

- Se algum A_t vale zero, haverá uma divisão por zero;
- Quando há um ajuste perfeito, MAPE é zero. Mas não há limite superior para a medida.

Quanto ao primeiro ponto problemático, este problema não ocorre para o problema abordado aqui, pois os valores reais nunca atingirão o valor zero, seja a temperatura (em Kelvin, sempre estritamente maior que zero) ou a umidade, que também assume sempre valores estritamente positivos.

O segundo ponto (a falta de limite superior) não é um problema para a abordagem do presente trabalho, pois o que realmente importa é a comparação do MAPE gerado a partir dos dados experimentais e a simulação do especialista com o MAPE gerado

²Também conhecido como desvio percentual médio absoluto (MAPD, Mean Absolute Percentage Deviation)

³Ao contrário do \mathcal{D} de Willmott, onde quanto mais próximo de zero o valor da previsão, *pior* é a previsão.

usando os (mesmos) dados experimentais e as previsões retornadas pelo método proposto.

4.4 Exemplo de testes

Uma vez exposta a metodologia de testes e as medidas estatísticas utilizadas para avaliá-los, com o objetivo de exemplificar tal metodologia foram feitos todos os cálculos envolvidos na realização de um teste específico. Ou seja, foram feitos os seguintes passos:

1. *Seção 4.4.1:* Escolher o conjunto de dados a ser utilizado, e em seguida o grupo (dia) de testes. A partir do grupo de testes, definir o grupo de treinamento;
2. *Seção 4.4.2:* Obter as duas distribuições priori e a função de verossimilhança para o grupo de treinamento definido;
3. *Seção 4.4.3:* Realizar os testes utilizando as distribuições priori e a verossimilhança obtidas no item 2 acima;
4. *Seção 4.4.4:* Interpretar os resultados dos testes, e analisá-los com relação às duas medidas estatísticas escolhidas, \mathcal{D} de Willmott e MAPE.

Os itens 2 e 3 acima foram realizados utilizando um software que foi desenvolvido como parte integrante do presente trabalho. Sendo assim, todas as equações das seções 4.4.2 e 4.4.3 foram obtidas utilizando o software citado, e os respectivos comandos foram organizados na seção B.1.1.2.

4.4.1 Conjunto de dados de exemplo

O conjunto de dados utilizado foi o “Temperatura – Verão – SAF2” (variável Temperatura, época do ano Verão e microclima SAF2, conforme seção 4.1.2), correspondente aos dados A.1.3 do apêndice.

Grupo de testes e de treinamento Os testes foram realizados quando o terceiro dia foi o grupo de testes. Dessa forma (pela Tabela 4.2), o grupo de treinamento constituiu-se dos dois primeiros dias.

4.4.2 Obtenção das funções para dados da aplicação

Conforme a seção 4.2.4, foram obtidas as duas distribuições priori (trapezoidal e parábola) e a função de verossimilhança correspondentes aos dados definidos na seção 4.4.1 conforme a metodologia descrita na seção 3.4.

4.4.2.1 Distribuição Priori

De acordo com a metodologia proposta na seção 3.4.1, foram obtidas as duas opções de distribuição priori para os dados da seção 4.4.1. De acordo com a seção 4.2.4, para obter as distribuições priori utilizou-se as medidas experimentais dos três dias do conjunto de dados em questão. Dessa forma, utilizou-se a primeira coluna (x_i , conforme Figura 4.1) de todas as 72 linhas (referentes aos três dias) do apêndice A.1.3 para gerar os histogramas necessários, que seguem na Figura 4.2:

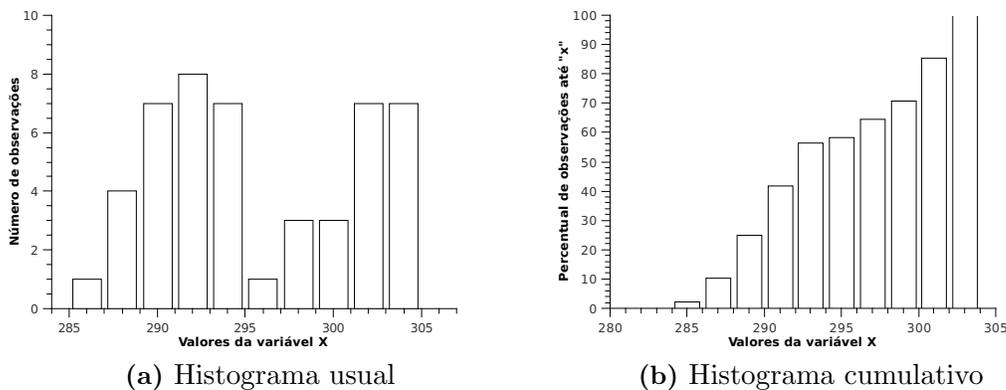


Figura 4.2: Histogramas da variável X correspondentes aos dados definidos na seção 4.4.1

Observando o histograma cumulativo da Figura 4.2b, nota-se que entre $X = 290$ e $X = 296$ houve um aumento percentual de quase 40%, o que torna razoável a escolha de que $u = 290$ e $v = 296$. Ajustando a e b para os valores extremos do histograma, obteve-se que $a = 284$ e $b = 307$ (nota-se que não são exatamente os valores extremos, mas foi

dada uma folga). Dessa forma, a priori $\pi^{(1)}(x)$ é descrita pelo número Fuzzy trapezoidal de parâmetros (284, 290, 296, 307). Substituindo o valor na equação do número fuzzy trapezoidal, obteve-se:

$$\text{trapez}(a, u, v, b)(x) = \begin{cases} \frac{(x-a)}{u-a} & \text{se } a \leq x < u \\ 1 & \text{se } u \leq x \leq v \\ \frac{(b-x)}{b-v} & \text{se } v \leq x < b \\ 0 & \text{caso contrário} \end{cases}$$

$$\pi^{(1)}(x) = \text{trapez}(284, 290, 296, 307)(x) = \begin{cases} \frac{(x-284)}{6} & \text{se } 284 \leq x < 290 \\ 1 & \text{se } 290 \leq x \leq 296 \\ \frac{(307-x)}{11} & \text{se } 296 \leq x < 307 \\ 0 & \text{caso contrário} \end{cases} \quad (4.4.1)$$

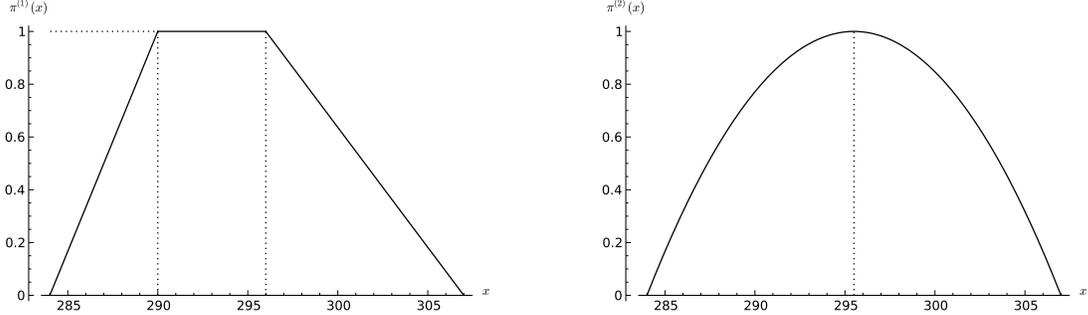
Para o caso de priori da parábola, $\pi^{(2)}(x)$, definindo os valores $a = 284$ e $b = 307$ da função trapezoidal (Eq. 4.4.1) definida acima como os extremos de uma parábola, define-se:

$$\text{parabolaNorm}(a, b)(x) = \frac{(x-a)(x-b)}{\left(\frac{a+b}{2}\right)^2}$$

$$\pi^{(2)}(x) = \text{parabolaNorm}(284, 307)(x) = \frac{(x-284)(x-307)}{\left(\frac{284+307}{2}\right)^2} \quad (4.4.2)$$

4.4.2.2 Função de Verossimilhança

Segundo o procedimento da seção 3.4.2, para obter a verossimilhança é necessário um gráfico $X \times Y$ com o ajuste linear de mínimos quadrados correspondente $u(x)$. Este gráfico, conforme a seção 4.2.4, deve utilizar os dados do grupo de treinamento correspondente ao teste realizado. Conforme a seção 4.4.1, o grupo de treinamento consiste dos dois primeiros dias, logo os pares (x, y) correspondem às duas primeiras



(a) Primeiro caso de distribuição priori, $\pi^{(1)}(x) = \text{trapez}(284, 290, 296, 307)(x)$ (b) Segundo caso de distribuição priori, $\pi^{(2)}(x) = \text{parabolaNorm}(284, 307)(x)$

Figura 4.3: As duas distribuições priori correspondentes ao teste definido na seção 4.4.1

colunas (x_i e y_i , conforme Figura 4.1) dos 48 primeiros pares do apêndice A.1.3. O ajuste linear de mínimos quadrados correspondente aos pontos $u(x)$ é calculado da forma usual [17, p. 191], obtendo como resultado o gráfico da Figura 4.4.

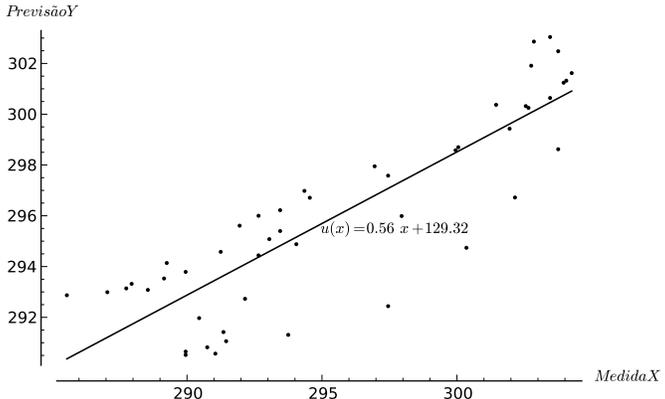


Figura 4.4: Gráfico $X \times Y$, com os pontos e o ajuste linear que correspondem aos dados que tiveram seus testes explicitados.

De acordo com a metodologia proposta na seção 3.4.2, foram obtidas computacionalmente (conforme comandos da seção B.1.1.2) as funções $a(x)$ e $b(x)$ (retas-limite) para a Figura 4.4. Dessa forma, foi obtido que:

$$a(x) = 0.54 \cdot x + 129.32 \tag{4.4.3}$$

$$b(x) = 0.58 \cdot x + 129.32, \tag{4.4.4}$$

onde o ajuste linear é:

$$u(x) = 0.56 \cdot x + 129.32. \quad (4.4.5)$$

Procedendo agora de forma análoga à seção 3.3.1.2, lembrando que se utilizou um número fuzzy triangular (2.1.1) para modelar a função de verossimilhança:

$$\pi(y|x) = \begin{cases} \frac{(y-a)}{u-a} & \text{se } a \leq y < u \\ \frac{(y-b)}{u-b} & \text{se } u \leq y \leq b \\ 0 & \text{caso contrário,} \end{cases} \quad (4.4.6)$$

e substitui-se os valores de (u, a, b) (respectivamente Eqs. 4.4.5, 4.4.3 e 4.4.4) na equação 4.4.6 acima e isolou-se a variável x nos intervalos, obteve-se a função verossimilhança correspondente ao teste da seção 4.4.1:

$$\pi(y|x) = \begin{cases} \frac{47.33 \cdot x - 82.16 \cdot y + 10624.99}{x} & \text{se } 1.74 \cdot y - 224.47 \leq x < 1.77 \cdot y - 229.30 \\ \frac{-25.87 \cdot x + 47.65 \cdot y - 6162.40}{x} & \text{se } 1.77 \cdot y - 229.30 \leq x < 1.84 \cdot y - 238.17 \cdot \\ 0 & \text{caso contrário} \end{cases} \quad (4.4.7)$$

Os códigos usados para obter a verossimilhança (4.4.7) constam na seção B.1.1.2. Segue na Figura 4.5 o gráfico da verossimilhança para alguns casos de previsão:

4.4.3 Realização do teste de exemplo

Uma vez que foram obtidas as funções necessárias para a aplicação do método (seção 4.4.2) para o teste referente aos da seção 4.4.1, a seguir são mostrados os resultados do teste correspondente utilizando o padrão de colunas do apêndice (Figura 4.1). O código que gera este teste consta na seção B.1.1.2.

Os valores da tabela 4.4 correspondem exatamente ao terceiro dia do apêndice A.1.3. Para que se possa ter idéia do comportamento geral dos resultados dos testes realizados,

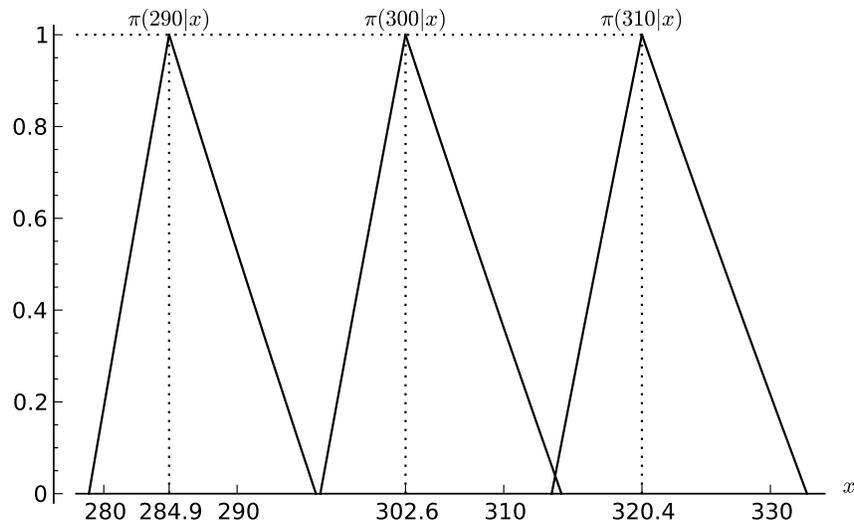


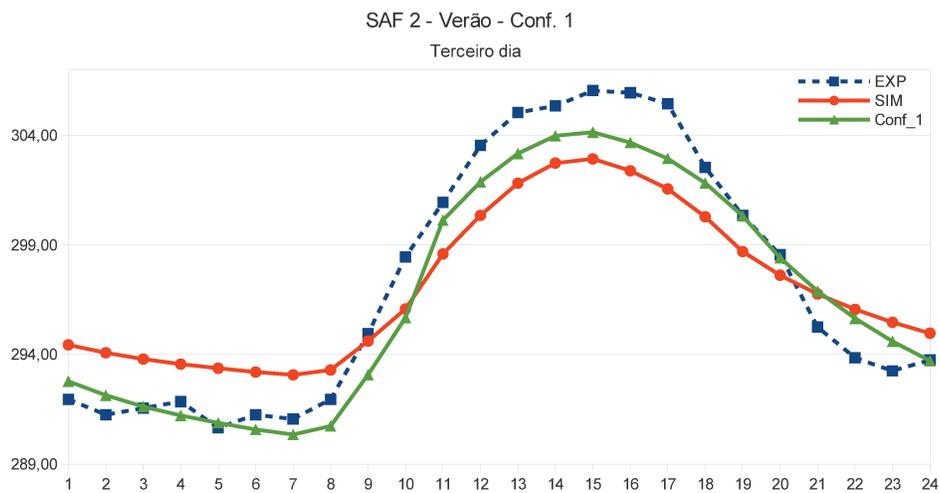
Figura 4.5: Eq. 4.4.7: Função de verossimilhança $\pi(y|x)$ para os dados do grupo de treinamento do teste da seção 4.4.1, para diversos casos de previsão Y .

Tabela 4.4: Resultados do teste correspondente ao grupo de testes da seção 4.4.1.

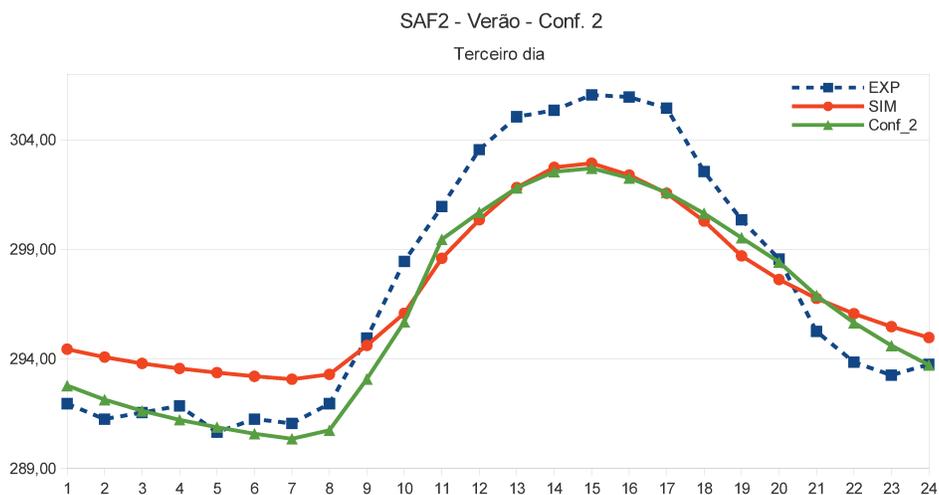
	Exp	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4
1	291,95	294,44	292,77	292,77	292,77	292,77
	291,25	294,08	292,14	292,14	292,14	292,14
	291,55	293,79	291,62	291,62	291,62	291,62
	291,85	293,56	291,21	291,21	291,21	291,21
	290,65	293,37	290,88	290,88	290,88	290,88
6	291,25	293,20	290,58	290,58	290,58	290,58
	291,05	293,07	290,35	290,35	290,35	290,35
	291,95	293,29	290,74	290,74	290,74	290,74
	294,95	294,61	293,08	293,08	293,08	293,08
	298,45	296,08	295,68	295,68	295,45	295,68
12	300,95	298,59	300,13	299,45	298,34	300,13
	303,55	300,35	301,88	300,68	300,36	303,25
	305,05	301,82	303,18	301,80	302,04	299,40
	305,35	302,74	303,98	302,54	303,08	301,00
	306,05	302,93	304,15	302,70	303,30	301,33
18	305,95	302,39	303,68	302,26	302,69	300,39
	305,45	301,56	302,95	301,60	301,74	298,95
	302,55	300,29	301,83	300,64	300,29	303,15
	300,35	298,70	300,33	299,53	298,47	300,33
	298,55	297,62	298,41	298,41	297,22	298,41
24	295,25	296,76	296,89	296,89	296,23	296,89
	293,85	296,06	295,65	295,65	295,42	295,65
	293,25	295,47	294,60	294,60	294,60	294,60
	293,75	294,97	293,71	293,71	293,71	293,71

gerou-se figuras que permitem visualizar o comportamento de (1) as medidas experimentais, (2) das previsões simuladas pelo especialista e (3) das previsões revisadas pelo método (segundo a terminologia da seção 4.3). Dessa forma, nos quatro gráficos das Figura 4.6 e 4.7, (1) e (2) correspondem respectivamente às colunas “Exp” e “Sim” da

Tabela 4.4, e em cada um dos gráficos tem-se que (3) é a coluna da configuração de posteriori correspondente na Tabela 4.4.

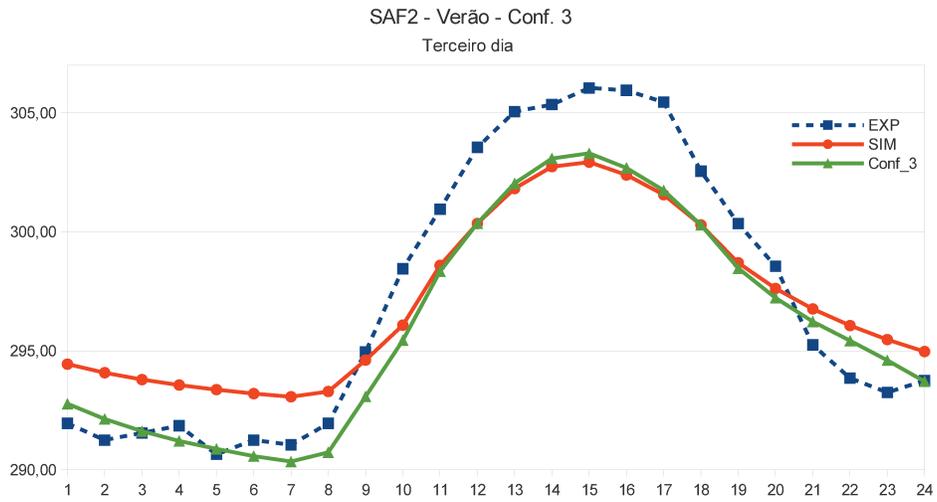


(a) Configuração 1: Trapézio (4.4.1)/Produto

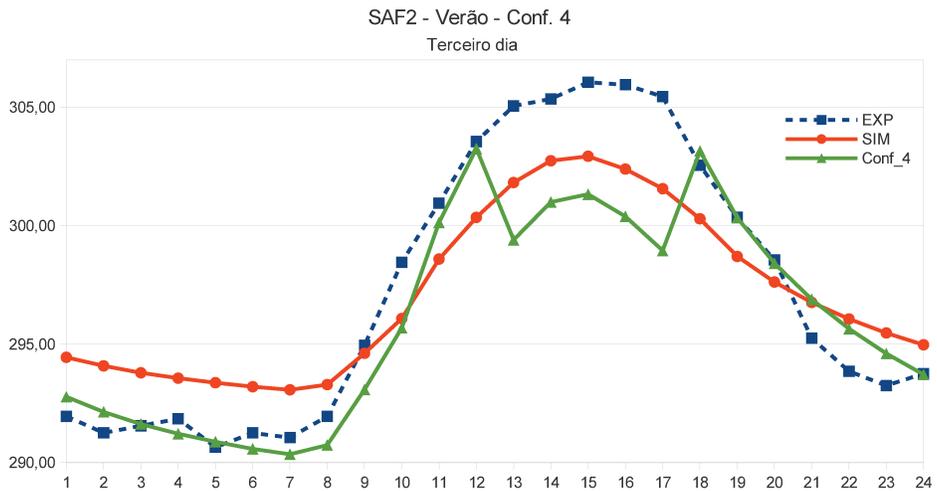


(b) Configuração 2: Parábola (4.4.2)/Produto

Figura 4.6: Gráficos com os dados experimentais de temperatura (“Exp”), as previsões dadas pelo especialista (“Sim”) e as previsões revisadas pelo método proposto (correspondendo às Configurações 1 e 2, segundo a Tabela 4.3). Os dados utilizados são os da Tabela 4.4 (dados da seção 4.4.1). Os eixos são: tempo (em horas) por temperatura (em Kelvin).



(a) Configuração 3: Trapézio (4.4.1)/Mínimo



(b) Configuração 4: Parábola (4.4.2)/Mínimo

Figura 4.7: Gráficos com os dados experimentais de temperatura (“Exp”), as previsões dadas pelo especialista (“Sim”) e as previsões revisadas pelo método proposto (correspondendo às Configurações 3 e 4, segundo Tabela 4.3). Os dados utilizados são os da Tabela 4.4 (dados da seção 4.4.1). Os eixos são: tempo (em horas) por temperatura (em Kelvin).

Observando com mais detalhes o gráfico da “Configuração 1” (Figura 4.6a), nota-se que nas primeiras 8 horas as previsões revisadas pelo método proposto (“Conf1”) acompanharam melhor os dados experimentais (“Exp”) que as previsões do especialista (“Sim”). Em seguida há um grande aumento, seguido de uma grande queda de temperatura. Nesta situação, as previsões do método também comportaram-se melhor que as simuladas pelo especialista.

Mas, apesar de gráficos como os das figuras 4.6 e 4.7 poderem ilustrar o comportamento da metodologia proposta e darem uma idéia de que se foi possível melhorar as previsões do especialista, esta não é uma maneira segura (nem prática) de avaliação. No tópico a seguir é feita a análise estatística da configuração que foi analisada graficamente acima (Configuração 1), para verificar se realmente foi possível melhorar as previsões do especialista neste caso.

4.4.4 Análise estatística do teste de exemplo

Conforme a seção 4.3, foram utilizadas duas medidas estatísticas para avaliar os testes realizados: \mathcal{D} de Willmott e MAPE (respectivamente Eqs. 4.3.2 e 4.3.3). Ambos foram calculados para a “Configuração 1”. Assim, foi possível determinar se essa configuração foi capaz de melhorar as previsões do especialista.

Para a avaliação do coeficiente \mathcal{D} de Willmott, seguiu-se o procedimento abaixo:

- Realizou-se dois cálculos:
 - \mathcal{D}_{Sim} : comparou as medidas experimentais (“Exp”) com as previsões do especialista (“Sim”);
 - \mathcal{D}_{C1} : comparou as medidas experimentais (“Exp”) com as previsões do método (“Conf. 1”)
- Considerou-se que o método proposto melhorou as previsões do especialista quando

a porcentagem $\mathcal{D}_\%$ foi positiva, onde

$$\mathcal{D}_\% = \frac{\mathcal{D}_{C1} - \mathcal{D}_{Sim}}{\mathcal{D}_{Sim}}.$$

Quanto ao coeficiente MAPE, seguiu-se um procedimento análogo:

- Realizou-se também dois cálculos:
 - \mathcal{M}_{Sim} : comparou as medidas experimentais (“Exp”) com as previsões do especialista (“Sim”);
 - \mathcal{M}_{C1} : comparou as medidas experimentais (“Exp”) com as previsões do método (“Conf. 1”)
- Considerou-se que o método proposto melhorou as previsões do especialista quando a porcentagem $\mathcal{M}_\%$ foi positiva, onde

$$\mathcal{M}_\% = \frac{\mathcal{M}_{Sim} - \mathcal{M}_{C1}}{\mathcal{M}_{Sim}}.^4$$

Deve-se tomar o cuidado que, tanto no caso do \mathcal{D} de Willmott como no MAPE, a aplicação das fórmulas correspondentes deve levar em conta a ordem dos vetores: o primeiro argumento é sempre o vetor de medidas “Exp”, o segundo é o vetor de previsões (“Sim” ou “Conf.1”). A Tabela 4.5 mostra os valores dos índices para os dados discutidos nesta seção.

Tabela 4.5: Índices estatísticos referentes à “Configuração 1” da Tabela 4.6, comparados com as previsões do especialista.

	Sim	Conf. 1	Melhora %
\mathcal{D}	0.92	0.98	5.7%
MAPE	0.75%	0.39%	48.30%

⁴Nota-se que o sinal da porcentagem do MAPE é invertido pelo fato de o \mathcal{D} de Willmott ser maior quanto melhor for a aproximação, enquanto o MAPE é uma medida de erro, logo um MAPE melhor é o menor possível.

Dessa forma, pela tabela tem-se que $\mathcal{D}_{\text{Sim}} = 0.92$, $\mathcal{D}_{C_1} = 0.98$, $\mathcal{M}_{\text{Sim}} = 0.75\%$ e $\mathcal{M}_{C_1} = 0.39\%$. Assim, teve-se que

$$\mathcal{D}_{\%} = \frac{\mathcal{D}_{C_1} - \mathcal{D}_{\text{Sim}}}{\mathcal{D}_{\text{Sim}}} = 5.7\% > 0$$

$$\mathcal{M}_{\%} = \frac{\mathcal{M}_{\text{Sim}} - \mathcal{M}_{C_1}}{\mathcal{M}_{\text{Sim}}} = 48.30\% > 0,$$

de onde conclui-se que a “Configuração 1” *foi capaz de melhorar as previsões do especialista*, como foi sugerido pela análise gráfica do final da seção 4.4.3. Os coeficientes calculados foram os correspondentes ao “dia 3” do apêndice A.1.3.

4.5 Análise estatística

4.5.1 Sobre os critérios escolhidos

Na seção 4.4.4, foi realizada a análise estatística de uma configuração específica em um teste específico. Se fossem realizados os testes para as outras três configurações para este caso, poderia-se pensar em formas de definir se este teste, de forma geral (isto é, sem pensar em alguma das quatro configurações de posteriori em particular), foi “bem sucedido”. Uma forma de avaliar se um teste foi bem sucedido é se *alguma* das quatro configurações foi capaz de melhorar as previsões do especialista (se para alguma das configurações ocorrer $\mathcal{D}_{\%} > 0$ e $\mathcal{M}_{\%} > 0$, conforme seção 4.4.4). Uma outra forma, mais restritiva (porém muito mais informativa), é definir que um teste é bem sucedido quando *todas* as quatro configurações foram capazes de melhorar as previsões do especialista (para todas as configurações ocorrer $\mathcal{D}_{\%} > 0$ e $\mathcal{M}_{\%} > 0$). Estes dois critérios, respectivamente chamados de critérios C_1 e C_2 , foram usados para classificar não só o teste de exemplo da seção 4.4, mas todos os 36 testes (conforme seção 4.2.3) realizados. Estes critérios são definidos formalmente a seguir:

Critério C_1 : **Pelo menos uma** das quatro formas de revisar previsões utilizando o método proposto (“Conf.1”, “Conf.2”, “Conf.3” ou “Conf.4”, conforme Tabela 4.3) foi

capaz de apresentar melhores resultados que as previsões do especialista nas duas medidas utilizadas (\mathcal{D} de Willmott e MAPE, $\mathcal{D}\% > 0$ e $\mathcal{M}\% > 0$).

Critério C_2 : **Todas** as quatro formas de realizar os cálculos do método proposto (“Conf.1”, “Conf.2”, “Conf.3” ou “Conf.4”, conforme Tabela 4.3) foram capazes de apresentar melhores resultados que as previsões do especialista nas duas medidas utilizadas (\mathcal{D} de Willmott e MAPE, $\mathcal{D}\% > 0$ e $\mathcal{M}\% > 0$).

4.5.2 Conclusões da análise estatística

A tabela 4.6 tem o objetivo de detalhar o comportamento dos testes com relação aos dois critérios definidos acima, denotados respectivamente por C_1 e C_2 . Por exemplo, para os testes correspondentes à variável temperatura, estação inverno e configuração de plantio PS, os três dias de testes foram aprovados no critério C_1 acima, enquanto apenas dois dias de testes foram aprovados no critério C_2 . O total “ C_1 ” de cada linha da tabela é obtido somando todos os valores “ C_1 ” de uma mesma linha, analogamente para o total de “ C_2 ”. Os valores $T^{(1)}$ e $T^{(2)}$ foram obtidos somando todos os totais correspondentes de cada linha.

Tabela 4.6: Número de testes (dias) bem sucedidos nos critérios C_1 e C_2 . $T^{(1)}$ é o número de testes bem sucedidos com relação ao critério C_1 , enquanto $T^{(2)}$ é o número de testes bem sucedidos no critério C_2

Temperatura										
	PS		SAF1		SAF2		SAF3		Total	
	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
Verão	2	1	2	2	2	1	3	2	9	6
Inverno	3	2	3	3	3	3	3	3	12	11
Testes bem sucedidos para a variável temperatura:									21	17
Umidade										
	PS		SAF1		SAF2		SAF3		Total	
	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2	C_1	C_2
Verão	Não foram realizados testes neste caso								–	
Inverno	2	0	3	2	3	3	3	3	11	8
Testes bem sucedidos para a variável umidade:									11	8
Total de testes bem sucedidos:									32	25
									$T^{(1)}$	$T^{(2)}$

Com base na Tabela 4.6, usando as informações $T^{(1)} = 32$, $T^{(2)} = 25$ e $T = 36$ (onde T é o total de testes realizados, ver seção 4.2.3), conclui-se que:

- $\frac{T^{(1)}}{T} = \frac{32}{36} \approx \underline{89\%}$ dos testes realizados foram bem sucedidos com relação ao critério C_1 . Isso significa que, para $T^{(1)} = 32$ testes, pelo menos uma das quatro configurações que permitem realizar os cálculos do método (Configurações 1, 2, 3 e 4, conforme seção 4.2.2) tem uma melhora percentual positiva tanto do \mathcal{D} de Willmott quanto do MAPE.
- $\frac{T^{(2)}}{T} = \frac{25}{36} \approx \underline{69\%}$ dos testes realizados foram bem sucedidos com relação ao critério C_1 . Isso quer dizer que, para $T^{(2)} = 25$ testes, todas as quatro formas de realizar os cálculos do método proposto têm uma melhora percentual positiva tanto do \mathcal{D} de Willmott quanto do MAPE.

4.6 Avanços metodológicos

Além de aplicações da metodologia do presente trabalho, proposta no artigo [1], foram obtidos alguns avanços com relação ao que foi proposto por Lapointe. Estes avanços são:

1. Desenvolvimento de um software que realiza os cálculos do método proposto que, pela utilização da matemática simbólica, foi capaz de trabalhar apenas com expressões analíticas sem utilizar aproximações;
2. Criação de uma forma de obter a função de verossimilhança fuzzy a partir de um registro histórico (ou seja, obtém as retas-limite para um gráfico $X \times Y$). Isso foi feito de forma totalmente independente do método, fazendo com que outras formas de obter as retas-limite possam ser propostas e acopladas ao código já existente do Apêndice B;
3. O software desenvolvido e a metodologia para a obtenção da verossimilhança permitiram que pudessem ser feitos testes em larga escala;
4. Avaliação dos testes utilizando duas medidas estatísticas distintas, obtendo resultados consistentes.

4.7 Problemas encontrados e perspectivas para estudos posteriores

4.7.1 Sobre os dados de Umidade – Verão (Apêndice A.3)

O método proposto não foi capaz de revisar as previsões do especialista para o conjunto de dados correspondente à variável Umidade/estação Verão (Apêndice A.3, cujas medidas foram realizadas no dia 16/09) pelo fato de que, para que seja possível realizar os cálculos do método (obter a distribuição posteriori), é preciso que o suporte da distribuição priori $\pi(x)$ e da função de verossimilhança $\pi(y|x)$ se interceptem em

algum intervalo. Se isso não ocorre, o produto ou o mínimo entre as duas funções será sempre zero. A Figura 4.8 a seguir ilustra a situação descrita:

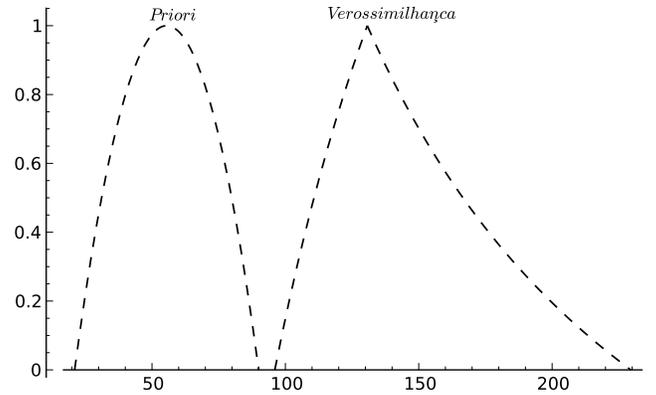


Figura 4.8: Gráfico contendo a distribuição priori de X e a função de verossimilhança $f(y|x)$, cujos suportes não se interceptam.

Possíveis soluções Se a distribuição priori ou a função de verossimilhança tivessem suporte infinito, o problema encontrado não ocorreria. Uma vez que já existe uma metodologia para obtenção da verossimilhança a partir de um conjunto de dados, o que não ocorre para a distribuição priori (que neste trabalho é obtida através de um histograma), inicialmente parece mais adequado utilizar uma função com suporte infinito para descrever a priori, como uma gaussiana por exemplo.

4.8 Conclusões

- Existem muitas potencialidades de aplicação do método proposto, como por exemplo dados meteorológicos, climáticos, que envolvem diagnósticos médicos, além de qualquer área que se tenha por objetivo melhorar uma previsão dada por um especialista levando em conta informações de um registro histórico.
- Foi desenvolvido um *software* que realiza todos os passos do método proposto, que permite que o método seja melhor compreendido e disseminado. Isso é particularmente útil quando se trata de um método relativamente novo e sem out-

ras implementações computacionais conhecidas. A documentação disponível no código anexo e o fato de ter sido usado um programa de código aberto (portanto gratuito) tem por objetivo facilitar tais tarefas.

- O método proposto, de acordo com a metodologia de testes e com os critérios estipulados, apresentou comportamento estatístico positivo. Uma vez que não se definiu qual das quatro formas é a mais adequada para realizar os cálculos da abordagem (de acordo com a Tabela 4.3), entre os dois critérios possíveis de avaliação (conforme a seção 4.5.2), **o critério C_2 é o mais confiável**. O fato de se ter utilizado duas medidas estatísticas distintas no mesmo conjunto de testes aumenta a confiabilidade destes resultados.

Apêndice A

Dados utilizados

A.1 Temperatura[K] – Verão

A.1.1 PS

Exp	Sim	Conf1	Conf2	Conf3	Conf4
292,75	292,74	289,80	289,80	290,15	289,80
291,85	291,99	288,55	288,55	289,26	288,55
292,45	291,44	287,63	287,63	288,61	287,63
292,05	291,09	287,05	287,05	288,19	287,05
291,25	290,84	286,74	286,63	287,90	286,63
290,65	290,67	286,59	286,43	287,69	286,35
290,25	290,54	286,49	286,34	287,54	286,13
292,05	290,59	286,53	286,37	287,60	286,22
295,05	291,33	287,45	287,45	288,48	287,45
297,25	292,46	289,33	289,33	289,82	289,33
299,85	294,81	293,24	293,24	293,24	293,24
301,55	296,86	296,66	296,66	295,89	296,66
303,35	298,70	299,72	299,72	298,06	299,72
303,05	300,24	301,37	302,28	299,87	302,28
304,05	301,19	302,16	303,52	300,98	303,87
304,05	301,29	302,24	303,59	301,10	304,03
303,35	300,64	301,71	302,95	300,34	302,95
301,85	299,43	300,70	300,93	298,92	300,93
296,85	297,94	298,45	298,45	297,16	298,45
294,05	296,98	296,86	296,86	296,03	296,86
293,55	296,23	295,61	295,61	295,14	295,61
293,25	295,61	294,57	294,57	294,41	294,57
293,75	295,09	293,71	293,71	293,71	293,71
291,65	294,65	292,98	292,98	292,98	292,98
290,75	294,25	294,22	294,22	294,15	294,22
291,25	293,90	293,68	293,68	293,68	293,68
290,15	293,59	293,21	293,21	293,21	293,21
289,05	293,34	292,83	292,83	292,83	292,83
289,35	293,13	292,51	292,51	292,51	292,51
288,85	292,97	292,26	292,26	292,26	292,26
285,95	292,84	292,06	292,06	292,06	292,06
290,15	293,05	292,38	292,38	292,38	292,38
293,95	294,45	294,52	294,52	294,36	294,52
297,55	296,12	297,07	297,07	296,11	297,07
299,45	298,83	300,55	301,21	298,93	301,21
301,35	300,53	301,84	303,27	300,69	303,81
302,45	302,01	302,96	304,18	302,22	306,07
302,95	302,93	303,66	304,78	303,17	302,32
303,15	303,07	303,76	304,88	303,31	302,53
303,95	302,48	303,32	304,49	302,71	301,65
304,35	301,62	302,67	303,93	301,82	305,47
302,85	300,30	301,67	303,14	300,45	303,46
300,35	298,68	300,44	300,98	298,77	300,98
298,15	297,58	299,30	299,30	297,63	299,30
295,75	296,72	297,99	297,99	296,73	297,99
294,55	296,02	296,92	296,92	296,00	296,92
294,85	295,43	296,02	296,02	295,39	296,02
295,05	294,94	295,27	295,27	294,87	295,27
292,55	294,51	293,64	293,64	293,64	293,64
292,95	294,15	293,06	293,06	293,06	293,06
292,55	293,84	292,56	292,56	292,56	292,56
292,55	293,60	292,17	292,17	292,17	292,17
292,05	293,39	291,84	291,84	291,84	291,84
291,95	293,23	291,58	291,58	291,58	291,58
291,35	293,09	291,36	291,36	291,36	291,36
292,75	293,30	291,69	291,69	291,69	291,69
295,65	294,62	293,81	293,81	293,81	293,81
298,25	296,17	296,30	296,30	295,46	296,30
300,95	298,77	299,52	300,48	298,09	300,48
303,35	300,42	300,84	302,52	299,75	303,13
305,05	301,82	301,95	303,36	301,16	305,37
305,35	302,69	302,64	303,91	302,03	306,77
306,15	302,84	302,76	304,01	302,18	300,54
306,25	302,29	302,33	303,65	301,63	306,13
305,45	301,47	301,67	303,14	300,80	304,81
302,85	300,20	300,66	302,40	299,53	302,77
300,85	298,65	299,43	300,28	297,97	300,28
299,95	297,59	298,58	298,58	296,90	298,58
296,35	296,76	297,25	297,25	296,06	297,25
294,95	296,07	296,14	296,14	295,36	296,14
294,55	295,51	295,24	295,24	294,79	295,24
294,45	295,02	294,45	294,45	294,29	294,45

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.894	0.860	0.870	0.849	0.870	1
	0.918	0.946	0.948	0.940	0.940	2
	0.937	0.958	0.984	0.925	0.981	3
Melhora % “D”		-3.83%	-2.74%	-5.05%	-2.69%	1
		2.96%	3.16%	2.31%	2.38%	2
		2.27%	5.04%	-1.29%	4.68%	3
MAPE	0.76%	1.15%	1.09%	1.12%	1.09%	1
	0.68%	0.61%	0.66%	0.61%	0.70%	2
	0.64%	0.49%	0.33%	0.62%	0.29%	3
Melhora % MAPE		-52.62%	-44.47%	-47.48%	-43.88%	1
		9.87%	2.02%	10.47%	-3.88%	2
		23.04%	48.69%	2.62%	55.18%	3
Melhora % MAPE		-52.62%	-44.47%	-47.48%	-43.88%	1
		9.87%	2.02%	10.47%	-3.88%	2
		23.04%	48.69%	2.62%	55.18%	3
Data de medidas:		16/09/2009				

A.1.2 SAF1

Exp	Sim	Conf1	Conf2	Conf3	Conf4
292.25	292.72	288.90	288.90	289.84	289.58
291.45	291.97	287.66	287.58	289.31	288.21
291.45	291.42	287.18	286.92	288.92	287.20
291.65	291.06	286.86	286.66	288.66	286.55
290.55	290.81	286.64	286.47	288.49	286.09
290.15	290.64	286.49	286.34	288.36	285.78
289.95	290.51	286.38	286.24	288.27	285.54
291.75	290.56	286.42	286.28	288.31	285.63
295.05	291.31	287.08	286.84	288.84	287.00
297.35	292.44	288.41	288.41	289.64	289.07
299.85	294.76	292.47	292.47	293.30	293.30
301.15	296.75	295.96	295.96	296.57	296.94
303.75	298.67	299.33	299.33	298.70	300.44
304.25	300.30	302.19	302.19	300.50	303.42
305.45	301.30	303.12	303.50	301.60	305.24
305.45	301.38	303.19	303.56	301.69	305.39
304.45	300.71	302.61	302.90	300.95	304.17
302.95	299.48	300.75	300.75	299.60	301.92
297.25	297.99	298.14	298.14	297.95	299.20
292.25	297.01	296.42	296.42	296.86	297.41
291.65	296.25	295.09	295.09	296.01	296.02
291.35	295.63	294.00	294.00	294.89	294.89
291.35	295.10	293.07	293.07	293.92	293.92
290.75	294.62	292.23	292.23	293.05	293.05
290.85	294.17	293.47	293.47	292.23	292.23
289.95	293.80	292.85	292.85	291.55	291.55
288.85	293.51	292.36	292.36	291.02	291.02
287.85	293.29	291.99	291.99	290.62	290.62
287.55	293.12	291.70	291.70	290.31	290.31
286.85	292.97	291.45	291.45	290.03	290.03
285.05	292.85	291.24	291.24	289.93	289.82
289.55	293.06	291.60	291.60	290.20	290.20
293.95	294.44	293.93	293.93	292.72	292.72
297.75	296.01	296.58	296.58	295.59	295.59
299.55	298.63	300.32	301.01	298.66	300.37
301.25	300.44	301.85	302.54	300.65	303.67
302.55	301.97	303.13	303.50	302.33	306.47
302.75	302.92	303.92	304.15	303.37	307.00
303.15	303.10	304.07	304.28	303.57	307.00
303.75	302.53	303.59	303.88	302.95	307.00
303.95	301.66	302.87	303.30	301.99	305.90
302.55	300.35	301.77	302.49	300.56	303.51
299.25	298.72	300.40	301.16	298.76	300.53
296.05	297.61	299.29	299.29	297.53	298.51
295.05	296.73	297.80	297.80	296.55	296.90
294.55	296.01	296.58	296.58	295.59	295.59
293.85	295.42	295.59	295.59	294.51	294.51
293.85	294.90	294.71	294.71	293.56	293.56
292.05	294.45	292.74	292.74	292.74	292.74
292.45	294.08	292.06	292.06	292.06	292.06
292.45	293.78	291.51	291.51	291.51	291.51
291.95	293.55	291.09	291.09	291.09	291.09
291.65	293.35	290.73	290.73	290.73	290.73
291.45	293.19	290.44	290.44	290.44	290.44
290.85	293.06	290.20	290.32	290.20	290.20
292.45	293.28	290.60	290.60	290.60	290.60
295.75	294.61	293.03	293.03	293.03	293.03
298.25	296.09	295.73	295.73	295.73	295.73
300.65	298.63	300.25	300.37	298.66	300.37
303.35	300.38	301.84	302.53	300.59	303.56
305.15	301.86	303.18	303.54	302.21	306.27
305.45	302.78	304.01	304.23	303.22	307.00
306.45	302.96	304.17	304.36	303.42	307.00
306.55	302.42	303.68	303.96	302.83	307.00
305.75	301.59	302.93	303.35	301.92	305.77
302.75	300.31	301.77	302.49	300.51	303.44
299.35	298.72	300.33	300.53	298.76	300.53
297.55	297.64	298.56	298.56	297.56	298.56
295.65	296.78	296.99	296.99	296.60	296.99
294.75	296.07	295.70	295.70	295.70	295.70
294.25	295.49	294.64	294.64	294.64	294.64
293.95	294.98	293.70	293.70	293.70	293.70

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D”	0.861	0.864	0.864	0.858	0.878	1
Willmott	0.888	0.942	0.942	0.970	0.960	2
	0.937	0.978	0.982	0.962	0.990	3
Melhora %		0.29%	0.29%	-0.35%	1.96%	1
“D”		5.99%	6.04%	9.22%	8.01%	2
		4.43%	4.83%	2.71%	5.70%	3
MAPE	0.96%	1.29%	1.30%	1.20%	1.24%	1
	0.84%	0.70%	0.73%	0.51%	0.73%	2
	0.66%	0.44%	0.40%	0.53%	0.32%	3
Melhora %		-34.36%	-34.91%	-24.78%	-28.66%	1
MAPE		16.22%	13.25%	39.31%	13.11%	2
		33.01%	39.08%	19.14%	52.02%	3
Melhora %		-34.36%	-34.91%	-24.78%	-28.66%	1
MAPE		16.22%	13.25%	39.31%	13.11%	2
		33.01%	39.08%	19.14%	52.02%	3
Data de medidas:			16/09/2009			

A.1.3 SAF2

Exp	Sim	Conf1	Conf2	Conf3	Conf4
292,15	292,73	288,58	288,58	289,08	288,58
290,45	291,97	287,21	287,21	288,19	287,21
291,35	291,42	286,27	286,21	287,55	286,21
291,45	291,06	285,95	285,56	287,12	285,56
290,75	290,82	285,73	285,13	286,84	285,13
289,95	290,65	285,58	284,83	286,64	284,83
289,95	290,52	285,46	284,59	286,49	284,59
291,05	290,57	285,50	284,68	286,55	284,68
293,75	291,31	286,17	286,02	287,42	286,02
297,45	292,44	288,05	288,05	288,74	288,05
300,35	294,74	292,20	292,20	292,20	292,20
302,15	296,72	295,77	295,77	295,56	295,77
303,75	298,62	299,19	299,19	298,03	299,19
302,65	300,25	302,13	300,97	300,14	302,13
303,95	301,24	303,12	301,75	301,42	299,28
304,05	301,32	303,19	301,81	301,52	299,42
303,45	300,64	302,58	301,27	300,64	298,22
301,95	299,43	300,65	300,36	299,08	300,65
296,95	297,95	297,99	297,99	297,16	297,99
294,35	296,98	296,24	296,24	295,90	296,24
293,45	296,22	294,87	294,87	294,87	294,87
291,95	295,61	293,77	293,77	293,77	293,77
293,05	295,08	292,81	292,81	292,81	292,81
291,25	294,58	291,91	291,91	291,91	291,91
289,25	294,14	293,39	293,39	293,39	293,39
289,95	293,79	292,82	292,82	292,82	292,82
289,15	293,53	292,40	292,40	292,40	292,40
287,95	293,32	292,05	292,05	292,05	292,05
287,75	293,14	291,76	291,76	291,76	291,76
287,05	292,99	291,52	291,52	291,52	291,52
285,55	292,87	291,32	291,32	291,32	291,32
288,55	293,08	291,66	291,66	291,66	291,66
292,65	294,44	293,89	293,89	293,89	293,89
297,95	295,99	296,42	296,42	296,00	296,42
299,95	298,58	300,65	300,21	298,99	300,65
301,45	300,37	302,76	301,44	301,04	298,58
302,75	301,91	304,02	302,57	302,80	301,06
302,85	302,86	304,79	303,30	303,88	302,59
303,45	303,04	304,93	303,44	304,09	302,88
303,75	302,48	304,48	303,01	303,45	301,98
304,25	301,62	303,78	302,35	302,47	300,59
302,55	300,32	302,72	301,40	300,98	298,50
300,05	298,70	300,85	300,29	299,13	300,85
297,45	297,58	299,02	299,02	297,84	299,02
294,55	296,71	297,60	297,60	296,84	297,60
292,65	296,00	296,43	296,43	296,02	296,43
293,45	295,40	295,45	295,45	295,32	295,45
294,05	294,88	294,60	294,60	294,60	294,60
291,95	294,44	292,77	292,77	292,77	292,77
291,25	294,08	292,14	292,14	292,14	292,14
291,55	293,79	291,62	291,62	291,62	291,62
291,85	293,56	291,21	291,21	291,21	291,21
290,65	293,37	290,88	290,88	290,88	290,88
291,25	293,20	290,58	290,58	290,58	290,58
291,05	293,07	290,35	290,35	290,35	290,35
291,95	293,29	290,74	290,74	290,74	290,74
294,95	294,61	293,08	293,08	293,08	293,08
298,45	296,08	295,68	295,68	295,45	295,68
300,95	298,59	300,13	299,45	298,34	300,13
303,55	300,35	301,88	300,68	300,36	303,25
305,05	301,82	303,18	301,80	302,04	299,40
305,35	302,74	303,98	302,54	303,08	301,00
306,05	302,93	304,15	302,70	303,30	301,33
305,95	302,39	303,68	302,26	302,69	300,39
305,45	301,56	302,95	301,60	301,74	298,95
302,55	300,29	301,83	300,64	300,29	303,15
300,35	298,70	300,33	299,53	298,47	300,33
298,55	297,62	298,41	298,41	297,22	298,41
295,25	296,76	296,89	296,89	296,23	296,89
293,85	296,06	295,65	295,65	295,42	295,65
293,25	295,47	294,60	294,60	294,60	294,60
293,75	294,97	293,71	293,71	293,71	293,71

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D”	0,891	0,861	0,843	0,854	0,818	1
Willmott	0,882	0,942	0,937	0,938	0,914	2
	0,929	0,982	0,960	0,961	0,926	3
Melhora %		-3,42%	-5,47%	-4,20%	-8,25%	1
“D”		6,83%	6,26%	6,42%	3,61%	2
		5,70%	3,33%	3,39%	-0,36%	3
MAPE	0,79%	1,19%	1,32%	1,17%	1,41%	1
	0,93%	0,78%	0,72%	0,73%	0,88%	2
	0,75%	0,39%	0,54%	0,56%	0,60%	3
Melhora %		-50,54%	-66,68%	-47,94%	-77,87%	1
MAPE		16,57%	22,40%	21,51%	5,46%	2
		48,30%	28,50%	25,83%	20,51%	3
Melhora %		-50,54%	-66,68%	-47,94%	-77,87%	1
MAPE		16,57%	22,40%	21,51%	5,46%	2
		48,30%	28,50%	25,83%	20,51%	3
Data de medidas:			16/09/2009			

A.1.4 SAF3

Exp	Sim	Conf1	Conf2	Conf3	Conf4
289.95	292.52	286.88	286.56	286.08	284.70
289.85	291.82	285.99	285.80	285.37	289.00
289.35	291.30	285.32	285.21	284.84	287.65
288.95	290.95	284.87	284.80	284.48	286.74
290.15	290.72	284.57	284.52	284.24	286.15
288.85	290.55	284.35	284.31	284.07	285.70
288.55	290.42	284.18	284.15	283.94	285.37
290.05	290.45	284.22	284.19	283.97	285.44
293.85	291.14	285.11	285.02	284.67	287.24
297.05	292.25	286.54	286.27	285.81	284.02
302.05	294.46	289.64	289.64	289.64	289.64
303.65	296.34	294.42	294.42	293.90	294.42
305.55	298.12	298.95	298.95	297.44	298.95
303.35	299.65	302.84	302.84	300.47	302.84
304.85	300.57	305.03	305.18	302.28	305.18
305.05	300.64	305.12	305.36	302.42	305.36
304.25	300.00	303.73	303.73	301.16	303.73
301.05	298.86	300.83	300.83	298.91	300.83
297.45	297.50	297.37	297.37	296.21	297.37
295.35	296.65	295.21	295.21	294.52	295.21
291.85	295.98	293.50	293.50	293.18	293.50
290.85	295.44	292.13	292.13	292.10	292.13
289.45	294.95	290.88	290.88	290.88	290.88
288.45	294.54	289.84	289.84	289.84	289.84
288.05	294.19	292.41	292.41	292.30	292.41
286.45	293.89	291.74	291.74	291.74	291.74
285.75	293.64	291.18	291.18	291.18	291.18
284.35	293.43	290.71	290.71	290.71	290.71
284.05	293.25	290.30	290.30	290.30	290.30
284.25	293.09	289.95	290.15	289.95	289.95
283.35	292.96	289.65	290.07	289.65	289.65
285.65	293.10	289.97	290.15	289.97	289.97
290.95	294.27	292.59	292.59	292.43	292.59
297.55	295.71	295.81	295.81	294.82	295.81
301.25	298.16	301.29	301.29	298.85	301.29
304.25	299.81	304.24	304.97	301.56	304.98
304.05	301.21	305.80	306.17	303.84	308.11
305.35	302.10	306.79	307.00	305.28	303.61
305.15	302.32	307.03	307.21	305.64	304.09
306.45	301.83	306.49	306.74	304.85	303.02
305.95	301.02	305.59	306.00	303.53	307.69
303.05	299.80	304.23	304.96	301.54	304.96
298.85	298.31	301.63	301.63	299.10	301.63
295.25	297.33	299.43	299.43	297.49	299.43
292.55	296.55	297.69	297.69	296.21	297.69
291.65	295.90	296.23	296.23	295.13	296.23
291.25	295.36	295.02	295.02	294.24	295.02
290.55	294.89	293.97	293.97	293.46	293.97
290.15	294.50	291.67	291.67	291.67	291.67
289.55	294.18	290.87	291.15	290.87	290.87
290.55	293.90	290.16	291.00	290.16	290.16
290.35	293.67	289.59	290.88	289.59	289.59
289.85	293.47	289.08	290.76	289.08	289.08
289.85	293.30	288.66	290.66	288.66	288.66
290.35	293.16	288.30	290.58	288.30	288.30
291.45	293.31	288.68	290.67	288.68	288.68
295.05	294.45	291.54	291.54	291.54	291.54
298.55	295.83	295.01	295.01	294.08	295.01
302.55	298.23	301.04	301.04	298.22	301.04
306.05	299.85	303.42	304.40	301.00	305.11
307.15	301.20	305.11	305.62	303.30	308.50
307.45	302.06	306.17	306.48	304.76	310.00
307.25	302.26	306.42	306.69	305.10	310.00
308.05	301.78	305.83	306.19	304.28	309.96
306.45	300.99	304.85	305.42	302.94	307.98
302.25	299.78	303.34	304.34	300.88	304.94
299.05	298.33	301.29	301.29	298.39	301.29
295.55	297.37	298.88	298.88	296.74	298.88
293.35	296.60	296.95	296.95	295.41	296.95
291.55	295.96	295.34	295.34	294.31	295.34
290.45	295.42	293.98	293.98	293.37	293.98
290.25	294.97	292.85	292.85	292.59	292.85

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D”	0.818	0.860	0.859	0.823	0.860	1
Willmott	0.769	0.931	0.929	0.927	0.923	2
	0.831	0.969	0.973	0.953	0.973	3
Melhora %		5.11%	4.92%	0.54%	5.13%	1
“D”		20.97%	20.77%	20.50%	19.92%	2
		16.63%	17.04%	14.68%	17.07%	3
MAPE	1.17%	1.27%	1.30%	1.55%	1.14%	1
	1.71%	1.12%	1.16%	1.08%	1.23%	2
	1.28%	0.70%	0.62%	0.80%	0.73%	3
Melhora %		-9.03%	-10.92%	-32.61%	2.88%	1
MAPE		34.33%	31.91%	36.87%	27.93%	2
		45.06%	51.56%	37.57%	42.90%	3
Melhora %		-9.03%	-10.92%	-32.61%	2.88%	1
MAPE		34.33%	31.91%	36.87%	27.93%	2
		45.06%	51.56%	37.57%	42.90%	3
Data de medidas:			16/09/2009			

A.2 Temperatura[K] – Inverno

A.2.1 PS

Exp	Sim	Conf1	Conf2	Conf3	Conf4
285.95	285.64	287.29	287.29	287.29	287.29
286.35	285.40	286.99	286.99	286.99	286.99
286.45	285.18	286.72	286.72	286.72	286.72
285.55	284.98	286.47	286.47	286.47	286.47
286.05	284.81	286.26	286.26	286.26	286.26
286.55	284.65	286.06	286.06	286.06	286.06
286.95	284.51	285.89	285.89	285.89	285.89
286.85	284.39	285.74	285.74	285.74	285.74
288.85	284.94	286.42	286.42	286.42	286.42
291.25	286.43	288.27	288.27	288.27	288.27
293.35	288.11	290.35	290.35	290.28	290.35
295.55	290.61	293.44	293.44	292.74	293.44
297.55	292.89	296.26	296.26	294.98	296.26
298.35	294.57	298.34	298.34	296.63	298.34
299.55	295.22	299.01	299.15	297.26	299.15
299.45	294.76	298.58	298.58	296.81	298.58
299.55	293.59	297.13	297.13	295.67	297.13
299.65	291.78	294.89	294.89	293.89	294.89
295.85	290.26	293.01	293.01	292.40	293.01
291.75	289.24	291.75	291.75	291.39	291.75
291.85	288.43	290.74	290.74	290.59	290.74
290.35	287.75	289.90	289.90	289.90	289.90
290.75	287.15	289.16	289.16	289.16	289.16
289.55	286.61	288.49	288.49	288.49	288.49
288.35	286.13	287.69	287.69	287.69	287.69
288.25	285.70	287.09	287.09	287.09	287.09
288.35	285.33	286.56	286.56	286.56	286.56
287.35	285.02	286.13	286.13	286.13	286.13
287.35	284.75	285.75	285.75	285.75	285.75
287.35	284.53	285.44	285.45	285.44	285.44
287.25	284.34	285.17	285.35	285.17	285.17
286.95	284.19	285.00	285.26	284.99	284.96
289.05	284.96	286.04	286.04	286.04	286.04
291.75	287.08	289.03	289.03	289.03	289.03
294.05	289.21	292.02	292.02	291.54	292.02
296.25	291.87	295.77	295.77	294.38	295.77
297.35	294.51	298.82	299.48	297.19	299.48
297.85	295.97	299.84	300.53	298.73	301.53
298.25	296.24	300.03	300.69	299.02	298.07
299.25	295.54	299.54	300.27	298.28	300.93
299.95	294.18	298.59	299.02	296.84	299.02
297.35	292.29	296.36	296.36	294.83	296.36
295.05	290.72	294.15	294.15	293.15	294.15
292.15	289.67	292.67	292.67	292.03	292.67
291.25	288.84	291.50	291.50	291.14	291.50
289.95	288.14	290.52	290.52	290.39	290.52
288.65	287.54	289.67	289.67	289.67	289.67
288.45	287.02	288.94	288.94	288.94	288.94
288.05	286.55	289.16	289.16	289.16	289.16
288.25	286.14	288.63	288.63	288.63	288.63
286.85	285.79	288.17	288.17	288.17	288.17
286.95	285.50	287.80	287.80	287.80	287.80
286.45	285.24	287.46	287.46	287.46	287.46
285.15	285.03	287.18	287.18	287.18	287.18
285.75	284.85	286.95	286.95	286.95	286.95
285.35	284.70	286.75	286.75	286.75	286.75
288.65	285.36	287.61	287.61	287.61	287.61
291.25	287.19	290.00	290.00	290.00	290.00
293.85	289.02	292.38	292.38	291.81	292.38
295.75	291.46	295.56	295.56	294.21	295.56
296.55	293.95	298.45	298.80	296.65	298.80
297.25	295.30	299.33	300.10	297.97	300.56
297.55	295.56	299.50	300.24	298.23	300.90
297.15	294.94	299.09	299.91	297.62	300.09
296.85	293.72	298.30	298.50	296.43	298.50
295.75	292.02	296.29	296.29	294.76	296.29
292.45	290.60	294.44	294.44	293.36	294.44
289.95	289.64	293.19	293.19	292.42	293.19
289.35	288.86	292.17	292.17	291.65	292.17
287.75	288.21	291.32	291.32	291.01	291.32
287.25	287.65	290.60	290.60	290.45	290.60
286.85	287.15	289.94	289.94	289.94	289.94

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.828	0.964	0.964	0.935	0.964	1
	0.864	0.974	0.972	0.964	0.970	2
	0.918	0.951	0.946	0.957	0.943	3
Melhora % “D”		16.44%	16.46%	12.95%	16.46%	1
		12.77%	12.54%	11.59%	12.31%	2
		3.56%	3.07%	4.27%	2.76%	3
MAPE	1.16%	0.48%	0.47%	0.63%	0.47%	1
	1.05%	0.46%	0.49%	0.50%	0.49%	2
	0.64%	0.59%	0.63%	0.49%	0.65%	3
Melhora % MAPE		59.04%	59.21%	45.98%	59.21%	1
		56.25%	53.80%	52.63%	53.92%	2
		7.53%	1.24%	23.86%	-1.62%	3
Melhora % MAPE		59.04%	59.21%	45.98%	59.21%	1
		56.25%	53.80%	52.63%	53.92%	2
		7.53%	1.24%	23.86%	-1.62%	3
Data de medidas:			20/06/2009			

A.2.2 SAF1

Exp	Sim	Conf1	Conf2	Conf3	Conf4
285.75	285.62	287.65	287.65	287.65	287.65
286.05	285.38	287.20	287.20	287.20	287.20
286.55	285.15	286.76	286.76	286.76	286.76
285.85	284.95	286.38	286.38	286.38	286.38
285.95	284.78	286.06	286.06	286.06	286.06
286.35	284.62	285.76	285.92	285.76	285.76
286.55	284.48	285.49	285.81	285.49	285.49
286.45	284.36	285.26	285.72	285.26	285.26
288.45	284.78	286.06	286.06	286.06	286.06
291.15	285.77	287.93	287.93	287.93	287.93
293.15	286.88	290.04	290.04	290.03	290.04
295.45	288.50	293.10	293.10	292.32	293.10
297.05	289.83	295.62	295.62	294.20	295.62
298.35	290.80	297.46	297.46	295.57	297.46
299.45	291.24	298.29	298.29	296.19	298.29
300.85	291.02	297.87	297.87	295.88	297.87
299.55	290.37	296.64	296.64	294.96	296.64
298.95	289.22	294.47	294.47	293.34	294.47
292.55	288.25	292.63	292.63	291.97	292.63
290.85	287.61	291.42	291.42	291.06	291.42
290.15	287.08	290.41	290.41	290.31	290.41
289.05	286.62	289.54	289.54	289.54	289.54
288.55	286.21	288.77	288.77	288.77	288.77
288.65	285.85	288.09	288.09	288.09	288.09
288.05	285.52	287.06	287.06	287.06	287.06
287.95	285.21	286.38	286.38	286.38	286.38
287.75	284.92	285.74	286.10	285.74	285.74
287.45	284.65	285.14	285.86	285.14	285.14
287.35	284.42	285.00	285.66	284.88	284.63
287.25	284.22	285.00	285.47	284.74	284.19
286.75	284.05	285.00	285.31	284.61	283.82
286.35	283.92	285.00	285.19	284.52	287.71
288.65	284.52	285.00	285.75	284.95	284.85
291.45	285.91	287.92	287.92	287.92	287.92
293.55	287.33	291.05	291.05	290.77	291.05
295.75	289.22	295.22	295.22	293.80	295.22
296.85	290.74	298.32	298.58	296.23	298.58
297.85	291.67	299.35	299.74	297.71	300.63
298.65	291.94	299.64	299.97	298.14	301.23
299.85	291.61	299.28	299.69	297.61	300.50
299.45	290.84	298.43	298.80	296.39	298.80
296.35	289.62	296.11	296.11	294.44	296.11
292.95	288.63	293.92	293.92	292.86	293.92
290.85	287.96	292.44	292.44	291.78	292.44
289.05	287.41	291.23	291.23	290.90	291.23
288.55	286.93	290.17	290.17	290.13	290.17
287.55	286.50	289.22	289.22	289.22	289.22
288.35	286.12	288.38	288.38	288.38	288.38
287.75	285.77	288.51	288.51	288.51	288.51
287.95	285.45	287.84	287.84	287.84	287.84
286.75	285.14	287.19	287.19	287.19	287.19
286.45	284.88	286.65	286.65	286.65	286.65
286.15	284.66	286.19	286.19	286.19	286.19
285.25	284.49	285.84	285.86	285.84	285.84
285.55	284.36	285.57	285.75	285.57	285.57
285.45	284.25	285.34	285.66	285.34	285.34
288.35	284.81	286.51	286.51	286.51	286.51
290.75	286.03	289.05	289.05	289.05	289.05
293.65	287.27	291.63	291.63	291.31	291.63
295.45	288.99	295.21	295.21	294.17	295.21
296.55	290.47	298.29	298.29	296.63	298.29
296.55	291.39	299.96	300.21	298.15	300.21
296.85	291.66	300.24	300.45	298.60	300.77
296.55	291.36	299.93	300.15	298.10	300.15
295.95	290.64	298.65	298.65	296.91	298.65
294.05	289.50	296.27	296.27	295.02	296.27
290.25	288.58	294.36	294.36	293.49	294.36
288.35	287.96	293.07	293.07	292.46	293.07
287.55	287.43	291.96	291.96	291.57	291.96
286.75	286.98	291.03	291.03	290.82	291.03
286.95	286.57	290.17	290.17	290.14	290.17
286.25	286.21	289.42	289.42	289.42	289.42

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D”	0.667	0.961	0.961	0.919	0.961	1
Willmott	0.699	0.964	0.970	0.950	0.956	2
	0.762	0.923	0.921	0.940	0.921	3
Melhora %		44.11%	44.17%	37.81%	44.11%	1
“D”		37.95%	38.74%	35.97%	36.83%	2
		21.17%	20.92%	23.35%	20.88%	3
MAPE	1.4%	0.48%	0.47%	0.63%	0.48%	1
	1.4%	0.55%	0.51%	0.61%	0.63%	2
	0.96%	0.70%	0.71%	0.55%	0.71%	3
Melhora %		66.80%	67.74%	56.16%	66.80%	1
MAPE		61.21%	64.53%	57.52%	55.89%	2
		27.39%	25.93%	43.03%	25.91%	3
Melhora %		66.80%	67.74%	56.16%	66.80%	1
MAPE		61.21%	64.53%	57.52%	55.89%	2
		27.39%	25.93%	43.03%	25.91%	3
Data de medidas:			20/06/2009			

A.2.3 SAF2

Exp	Sim	Conf1	Conf2	Conf3	Conf4
283.75	285.62	287.08	286.86	288.06	286.86
284.95	285.38	286.83	286.36	287.75	286.36
284.85	285.16	286.60	286.19	287.46	285.91
284.35	284.96	286.39	286.03	287.20	285.49
284.95	284.78	286.20	285.88	286.97	285.11
285.75	284.63	286.04	285.76	286.78	284.80
286.25	284.49	285.90	285.64	286.59	284.51
285.95	284.36	285.76	285.53	286.43	284.24
287.75	284.77	286.19	285.88	286.96	285.09
290.45	285.76	287.22	287.16	288.24	287.16
293.35	286.86	289.45	289.45	289.66	289.45
295.75	288.47	292.80	292.80	292.80	292.80
297.75	289.80	295.57	295.57	295.32	295.57
298.55	290.76	297.13	297.57	296.46	297.57
298.95	291.19	297.58	298.20	296.97	298.46
299.35	290.98	297.36	298.02	296.72	298.02
299.05	290.33	296.67	296.67	295.95	296.67
297.95	289.19	294.30	294.30	294.30	294.30
293.05	288.23	292.30	292.30	292.30	292.30
290.55	287.59	290.97	290.97	290.97	290.97
289.95	287.07	289.88	289.88	289.93	289.88
288.75	286.61	288.93	288.93	289.34	288.93
288.15	286.19	288.05	288.05	288.79	288.05
288.15	285.81	287.28	287.26	288.30	287.26
286.95	285.47	286.96	286.46	287.69	286.11
287.05	285.16	286.58	286.18	287.25	285.36
286.55	284.87	286.23	285.91	286.83	284.66
286.35	284.60	285.90	285.65	286.44	284.01
286.45	284.38	285.64	285.43	286.13	283.48
286.35	284.21	285.43	285.25	285.88	287.88
286.25	284.07	285.26	285.11	285.68	287.54
285.95	283.96	285.13	284.99	285.52	287.27
287.85	284.54	285.83	285.59	286.36	283.87
290.75	285.87	287.44	287.08	288.27	287.08
293.65	287.20	290.29	290.29	290.29	290.29
295.65	289.00	294.63	294.63	294.63	294.63
297.35	290.53	297.33	298.03	296.84	298.32
297.55	291.51	298.51	298.88	298.14	300.69
297.85	291.82	298.88	299.18	298.55	301.43
298.65	291.49	298.49	298.87	298.11	300.64
298.95	290.74	297.59	298.21	297.12	298.83
295.95	289.54	295.93	295.93	295.52	295.93
291.95	288.56	293.57	293.57	293.57	293.57
290.65	287.90	291.98	291.98	291.98	291.98
288.75	287.35	290.65	290.65	290.65	290.65
288.15	286.88	289.51	289.51	289.71	289.51
286.95	286.46	288.50	288.50	289.11	288.50
287.05	286.08	287.70	287.58	288.57	287.58
286.45	285.70	287.67	287.67	288.48	287.67
286.95	285.34	286.87	286.87	287.95	286.87
285.55	285.03	286.44	286.18	287.50	286.18
285.85	284.79	286.17	285.86	287.15	285.65
285.05	284.62	285.98	285.71	286.91	285.27
283.95	284.48	285.83	285.58	286.70	284.96
284.85	284.37	285.70	285.48	286.54	284.72
284.45	284.28	285.60	285.40	286.41	284.52
286.85	284.81	286.19	285.88	287.18	285.69
289.95	286.01	288.35	288.35	288.93	288.35
293.45	287.20	290.99	290.99	290.99	290.99
296.05	288.88	294.72	294.72	294.72	294.72
297.05	290.35	297.61	297.98	296.78	297.98
297.05	291.29	298.65	298.99	298.01	300.06
297.25	291.59	298.98	299.26	298.41	300.73
296.65	291.29	298.65	298.99	298.01	300.06
296.05	290.58	297.86	298.40	297.08	298.49
294.15	289.46	296.01	296.01	295.60	296.01
289.35	288.54	293.97	293.97	293.97	293.97
287.65	287.91	292.57	292.57	292.57	292.57
286.55	287.40	291.44	291.44	291.44	291.44
285.55	286.94	290.42	290.42	290.42	290.42
286.05	286.54	289.53	289.53	289.69	289.53
285.35	286.17	288.71	288.71	289.16	288.71

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.684	0.957	0.942	0.937	0.962	1
	0.730	0.961	0.976	0.977	0.956	2
Melhora % “D”		0.74%	0.93%	0.917	0.930	3
		39.80%	37.71%	36.94%	40.57%	1
MAPE		31.72%	33.73%	33.93%	31.06%	2
		19.99%	20.08%	18.11%	19.85%	3
	1.35%	0.54%	0.50%	0.64%	0.54%	1
MAPE	1.24%	0.38%	0.43%	0.37%	0.63%	2
	0.90%	0.71%	0.72%	0.76%	0.74%	3
		59.85%	62.87%	53.00%	60.41%	1
Melhora % MAPE		69.43%	65.48%	69.81%	49.16%	2
		21.15%	20.17%	15.06%	17.75%	3
		59.85%	62.87%	53.00%	60.41%	1
Melhora % MAPE		69.43%	65.48%	69.81%	49.16%	2
		21.15%	20.17%	15.06%	17.75%	3
		59.85%	62.87%	53.00%	60.41%	1
Data de medidas:		20/06/2009				

A.2.4 SAF3

Exp	Sim	Conf1	Conf2	Conf3	Conf4
282.85	285.56	285.85	286.10	285.85	285.85
282.55	285.32	285.18	285.91	285.18	285.18
281.95	285.10	284.56	285.72	284.56	284.56
282.65	284.90	284.00	285.54	284.00	284.00
283.55	284.73	283.52	285.38	283.52	283.52
284.05	284.57	283.07	285.23	283.07	283.07
284.65	284.44	283.00	285.11	282.93	282.70
284.75	284.31	283.00	284.98	282.84	282.34
287.95	284.69	283.41	285.35	283.41	283.41
293.15	285.66	286.14	286.19	286.14	286.14
295.15	286.70	289.06	289.06	288.80	289.06
296.45	288.18	293.23	293.23	291.90	293.23
298.55	289.43	296.75	296.75	294.51	296.75
299.45	290.33	299.28	299.28	296.39	299.28
299.55	290.74	300.43	299.81	297.24	300.43
299.45	290.54	299.87	299.62	296.82	299.87
299.25	289.89	298.04	298.04	295.47	298.04
295.75	288.77	294.89	294.89	293.14	294.89
292.05	287.85	292.30	292.30	291.21	292.30
289.85	287.26	290.64	290.64	289.97	290.64
287.55	286.78	289.29	289.29	288.96	289.29
287.75	286.37	288.13	288.13	288.10	288.13
286.55	286.00	287.09	287.09	287.09	287.09
286.55	285.62	286.02	286.15	286.02	286.02
286.75	285.23	284.14	285.80	284.14	284.14
286.45	284.92	283.15	285.49	283.15	283.15
285.95	284.70	283.00	285.26	282.87	282.45
285.75	284.53	283.00	285.07	282.74	281.91
285.85	284.39	283.00	284.92	282.64	281.46
285.95	284.28	283.00	284.79	282.56	281.11
285.45	284.18	283.00	284.68	282.48	280.80
284.95	284.08	283.00	284.56	282.41	280.48
288.35	284.57	283.00	285.12	282.77	282.04
293.35	285.73	285.73	286.26	285.73	285.73
295.35	286.91	289.48	289.48	289.12	289.48
296.95	288.53	294.63	294.63	293.00	294.63
298.15	289.94	299.12	299.12	296.35	299.12
298.75	290.86	302.05	300.48	298.53	302.05
298.15	291.18	302.81	300.86	299.29	303.06
298.35	290.89	302.14	300.52	298.60	302.14
298.75	290.17	299.85	299.71	296.90	299.85
295.85	289.04	296.26	296.26	294.21	296.26
291.55	288.13	293.36	293.36	292.04	293.36
289.65	287.54	291.48	291.48	290.63	291.48
287.85	287.05	289.93	289.93	289.45	289.93
285.95	286.63	288.59	288.59	288.45	288.59
285.35	286.25	287.38	287.38	287.38	287.38
285.85	285.82	286.01	286.34	286.01	286.01
284.55	285.45	286.27	286.27	286.27	286.27
283.65	285.19	285.52	285.58	285.52	285.52
284.65	285.00	284.97	285.40	284.97	284.97
284.55	284.85	284.53	285.26	284.53	284.53
283.55	284.72	284.15	285.14	284.15	284.15
282.65	284.60	283.81	285.02	283.81	283.81
283.45	284.49	283.49	284.91	283.49	283.49
283.75	284.38	283.17	284.79	283.17	283.17
287.25	284.82	284.44	285.23	284.44	284.44
292.65	285.90	287.58	287.58	287.58	287.58
295.85	286.99	290.74	290.74	290.15	290.74
296.85	288.53	295.20	295.20	293.65	295.20
297.75	289.90	299.18	299.18	296.75	299.18
297.35	290.78	301.73	300.81	298.73	301.73
297.25	291.08	302.60	301.15	299.41	302.60
296.75	290.80	301.79	300.83	298.78	301.79
295.95	290.11	299.78	299.78	297.22	299.78
293.35	289.03	296.65	296.65	294.79	296.65
289.05	288.16	294.13	294.13	292.81	294.13
286.95	287.59	292.48	292.48	291.52	292.48
284.95	287.12	291.11	291.11	290.45	291.11
283.75	286.72	289.95	289.95	289.54	289.95
283.75	286.30	288.74	288.74	288.58	288.74
284.25	285.89	287.55	287.55	287.55	287.55

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.622	0.962	0.960	0.935	0.962	1
	0.651	0.944	0.965	0.934	0.928	2
Melhora % “D”		54.69%	54.30%	50.24%	54.56%	1
		45.05%	48.34%	43.52%	42.66%	2
MAPE		31.54%	31.63%	33.28%	31.54%	3
	1.46%	0.64%	0.64%	0.84%	0.65%	1
	1.31%	0.97%	0.64%	0.89%	1.13%	2
Melhora % MAPE		1.08%	1.10%	0.87%	1.08%	3
		56.36%	56.28%	42.56%	55.40%	1
		26.14%	51.23%	32.03%	13.93%	2
Melhora % MAPE		6.75%	5.08%	25.52%	6.75%	3
		56.36%	56.28%	42.56%	55.40%	1
		26.14%	51.23%	32.03%	13.93%	2
Melhora % MAPE		6.75%	5.08%	25.52%	6.75%	3
	Data de medidas:		20/06/2009			

A.3 Umidade[%] – Verão

A.3.1 PS

Exp	Sim	Conf1	Conf2	Conf3	Conf4
77	69.01	73.10	74.84	72.28	74.58
81	72.43	75.09	76.35	74.39	76.60
78	74.52	76.29	77.30	75.66	77.76
80	75.46	76.83	77.74	76.23	78.27
81	75.72	76.98	77.86	76.38	78.41
84	75.62	76.92	77.81	76.32	78.36
86	75.35	76.76	77.68	76.16	78.21
83	74.64	76.36	77.36	75.73	77.83
74	72.20	74.96	76.25	74.25	76.46
67	68.40	72.74	74.58	71.90	74.21
62	59.90	67.62	71.13	66.49	68.40
53	51.10	61.15	61.15	60.61	61.15
37	43.29	51.14	51.14	51.14	51.14
36	37.82	44.13	44.13	44.13	44.13
32	34.74	40.18	40.18	40.18	40.18
26	33.80	40.00	38.97	39.51	38.97
27	34.74	40.18	40.18	40.18	40.18
31	36.86	42.90	42.90	42.90	42.90
43	39.76	46.61	46.61	46.61	46.61
50	41.64	49.02	49.02	49.02	49.02
53	43.12	50.92	50.92	50.92	50.92
51	44.37	52.52	52.52	52.52	52.52
51	45.46	53.92	53.92	53.92	53.92
60	46.44	55.18	55.18	55.18	55.18
61	47.35	56.34	56.34	56.34	56.34
57	48.20	57.43	57.43	57.43	57.43
62	48.95	58.39	58.39	58.39	58.39
68	49.57	59.19	59.19	59.19	59.19
68	50.06	59.82	59.82	59.82	59.82
71	50.45	60.32	60.32	60.17	60.32
82	50.75	60.70	60.70	60.37	60.70
72	50.02	59.76	59.76	59.76	59.76
60	45.80	54.36	54.36	54.36	54.36
54	41.29	48.57	48.57	48.57	48.57
51	34.97	40.47	40.47	40.47	40.47
48	31.59	38.92	36.70	38.13	36.14
41	28.93	37.12	35.49	36.44	32.73
42	27.40	36.07	34.74	35.45	30.77
44	27.15	35.90	34.61	35.28	30.45
36	28.05	36.52	35.06	35.87	31.60
34	29.48	37.50	35.75	36.79	33.44
39	31.82	39.07	36.80	38.28	36.44
43	35.00	40.51	40.51	40.51	40.51
46	37.36	43.54	43.54	43.54	43.54
59	39.34	46.08	46.08	46.08	46.08
61	41.04	48.25	48.25	48.25	48.25
59	42.52	50.15	50.15	50.15	50.15
62	43.82	51.82	51.82	51.82	51.82
70	44.97	53.29	53.29	53.29	53.29
71	45.98	54.59	54.59	54.59	54.59
74	46.85	55.70	55.70	55.70	55.70
75	47.58	56.64	56.64	56.64	56.64
78	48.18	57.41	57.41	57.41	57.41
78	48.67	58.03	58.03	58.03	58.03
81	49.08	58.56	58.56	58.56	58.56
78	48.46	57.76	57.76	57.76	57.76
70	44.68	52.92	52.92	52.92	52.92
64	40.66	47.77	47.77	47.77	47.77
57	34.77	40.22	40.22	40.22	40.22
46	31.54	38.88	36.68	38.10	36.08
34	29.08	37.23	35.56	36.54	32.92
33	27.64	36.24	34.86	35.60	31.08
29	27.39	36.07	34.73	35.44	30.76
29	28.28	36.68	35.18	36.02	31.90
29	29.67	37.63	35.84	36.92	33.68
38	31.95	39.16	36.85	38.36	36.60
39	35.02	40.54	40.54	40.54	40.54
40	37.30	43.46	43.46	43.46	43.46
54	39.23	45.93	45.93	45.93	45.93
56	40.88	48.05	48.05	48.05	48.05
58	42.33	49.91	49.91	49.91	49.91
57	43.60	51.54	51.54	51.54	51.54

A.3.2 SAF1

Exp	Sim	Conf1	Conf2	Conf3	Conf4
78	69.12	74.06	76.35	73.50	75.95
80	72.58	76.19	78.06	75.72	78.14
80	74.72	77.48	79.14	77.08	79.42
79	75.67	78.05	79.63	77.68	79.98
83	75.94	78.22	79.77	77.85	80.14
83	75.85	78.16	79.72	77.79	80.08
85	75.58	78.00	79.59	77.62	79.93
82	74.86	77.57	79.22	77.17	79.50
72	72.41	76.08	77.97	75.61	78.03
67	68.57	73.72	76.08	73.14	75.59
60	60.06	68.33	72.16	67.49	69.44
54	51.33	62.53	62.72	61.43	62.72
37	43.39	51.90	51.90	51.90	51.90
34	37.71	44.15	44.15	44.15	44.15
31	34.54	39.83	39.83	39.83	39.83
26	33.62	38.58	38.58	38.58	38.58
26	34.59	39.90	39.90	39.90	39.90
29	36.74	42.83	42.83	42.83	42.83
44	39.66	46.81	46.81	46.81	46.81
57	41.57	49.42	49.42	49.42	49.42
60	43.08	51.47	51.47	51.47	51.47
57	44.33	53.18	53.18	53.18	53.18
58	45.44	54.69	54.69	54.69	54.69
62	46.53	56.18	56.18	56.18	56.18
59	47.62	57.66	57.66	57.66	57.66
60	48.55	58.93	58.93	58.93	58.93
65	49.25	59.88	59.88	59.88	59.88
72	49.77	60.59	60.59	60.31	60.59
73	50.18	61.15	61.15	60.61	61.15
78	50.51	61.60	61.60	60.84	61.60
85	50.79	61.98	61.98	61.04	61.98
74	50.03	60.95	60.95	60.50	60.95
61	45.87	55.28	55.28	55.28	55.28
53	41.61	49.47	49.47	49.47	49.47
51	35.43	41.05	41.05	41.05	41.05
47	31.78	36.07	36.07	36.07	36.07
39	29.01	32.30	32.30	32.30	32.30
43	27.42	30.13	30.13	30.13	30.13
44	27.12	30.00	29.72	29.88	29.72
37	27.99	30.91	30.91	30.91	30.91
32	29.41	32.84	32.84	32.84	32.84
37	31.74	36.02	36.02	36.02	36.02
47	34.93	40.37	40.37	40.37	40.37
51	37.32	43.62	43.62	43.62	43.62
58	39.33	46.36	46.36	46.36	46.36
59	41.06	48.72	48.72	48.72	48.72
62	42.57	50.78	50.78	50.78	50.78
64	43.93	52.63	52.63	52.63	52.63
70	45.16	54.31	54.31	54.31	54.31
70	46.20	55.73	55.73	55.73	55.73
73	47.04	56.87	56.87	56.87	56.87
75	47.73	57.81	57.81	57.81	57.81
77	48.30	58.59	58.59	58.59	58.59
78	48.79	59.26	59.26	59.26	59.26
81	49.20	59.81	59.81	59.81	59.81
78	48.53	58.90	58.90	58.90	58.90
70	44.70	53.68	53.68	53.68	53.68
61	40.85	48.43	48.43	48.43	48.43
56	35.08	40.57	40.57	40.57	40.57
43	31.62	35.85	35.85	35.85	35.85
33	29.01	32.30	32.30	32.30	32.30
34	27.50	30.24	30.24	30.24	30.24
29	27.21	30.00	29.84	29.93	29.84
30	28.07	31.02	31.02	31.02	31.02
28	29.46	32.91	32.91	32.91	32.91
36	31.74	36.02	36.02	36.02	36.02
45	34.86	40.27	40.27	40.27	40.27
44	37.21	43.47	43.47	43.47	43.47
54	39.19	46.17	46.17	46.17	46.17
54	40.89	48.49	48.49	48.49	48.49
58	42.38	50.52	50.52	50.52	50.52
56	43.71	52.33	52.33	52.33	52.33

A.3.3 SAF2

Exp	Sim	Conf1	Conf2	Conf3	Conf4
78	69.08	73.53	76.77	74.51	76.60
85	72.53	75.63	78.37	76.45	78.74
82	74.65	76.89	79.39	77.63	79.97
81	75.61	77.46	79.85	78.15	80.52
82	75.87	77.62	79.98	78.30	80.67
85	75.78	77.56	79.93	78.25	80.62
86	75.52	77.41	79.81	78.10	80.47
85	74.79	76.98	79.45	77.70	80.05
77	72.36	75.52	78.29	76.36	78.63
64	68.54	73.20	76.52	74.20	76.25
58	60.08	67.85	72.85	69.24	70.14
50	51.40	62.50	62.50	62.50	62.50
35	43.50	51.54	51.54	51.54	51.54
38	37.82	43.66	43.66	43.66	43.66
33	34.65	39.26	39.26	39.26	39.26
28	33.74	38.00	38.00	38.00	38.00
28	34.72	39.36	39.36	39.36	39.36
31	36.85	42.31	42.31	42.31	42.31
41	39.75	46.34	46.34	46.34	46.34
47	41.64	48.96	48.96	48.96	48.96
51	43.14	51.04	51.04	51.04	51.04
54	44.38	52.76	52.76	52.76	52.76
51	45.50	54.32	54.32	54.32	54.32
62	46.66	55.93	55.93	55.93	55.93
66	47.74	57.42	57.42	57.42	57.42
59	48.58	58.59	58.59	58.59	58.59
63	49.21	59.46	59.46	59.46	59.46
70	49.70	60.14	60.14	60.14	60.14
72	50.11	60.71	60.71	60.71	60.71
76	50.45	61.19	61.19	61.19	61.19
82	50.74	61.59	61.59	61.59	61.59
77	49.98	60.53	60.53	60.53	60.53
63	45.89	54.86	54.86	54.86	54.86
51	41.67	49.00	49.00	49.00	49.00
49	35.54	40.49	40.49	40.49	40.49
46	31.90	35.44	35.44	35.44	35.44
39	29.13	31.60	33.80	31.60	31.60
42	27.53	30.00	32.90	29.82	29.38
42	27.21	30.00	32.71	29.69	28.93
38	28.07	30.13	33.21	30.13	30.13
32	29.48	32.08	33.99	32.08	32.08
37	31.79	35.29	35.29	35.29	35.29
42	34.98	39.72	39.72	39.72	39.72
47	37.37	43.03	43.03	43.03	43.03
61	39.38	45.82	45.82	45.82	45.82
69	41.10	48.21	48.21	48.21	48.21
62	42.61	50.31	50.31	50.31	50.31
61	43.97	52.19	52.19	52.19	52.19
70	45.18	53.87	53.87	53.87	53.87
76	46.19	55.27	55.27	55.27	55.27
78	47.01	56.41	56.41	56.41	56.41
76	47.69	57.36	57.36	57.36	57.36
82	48.26	58.15	58.15	58.15	58.15
80	48.76	58.84	58.84	58.84	58.84
80	49.17	59.41	59.41	59.41	59.41
80	48.50	58.48	58.48	58.48	58.48
71	44.72	53.23	53.23	53.23	53.23
59	40.89	47.92	47.92	47.92	47.92
54	35.15	39.95	39.95	39.95	39.95
44	31.69	35.15	35.15	35.15	35.15
37	29.08	31.53	33.78	31.53	31.53
35	27.56	30.00	32.92	29.83	29.42
30	27.26	30.00	32.74	29.71	29.00
29	28.12	30.20	33.24	30.20	30.20
28	29.51	32.13	34.01	32.13	32.13
36	31.78	35.28	35.28	35.28	35.28
39	34.91	39.62	39.62	39.62	39.62
42	37.25	42.87	42.87	42.87	42.87
56	39.23	45.61	45.61	45.61	45.61
59	40.94	47.99	47.99	47.99	47.99
61	42.43	50.06	50.06	50.06	50.06
58	43.76	51.90	51.90	51.90	51.90

A.3.4 SAF3

Exp	Sim	Conf1	Conf2	Conf3	Conf4
86	70.60	75.84	78.01	82.30	81.14
87	73.76	78.24	79.79	83.73	82.94
88	75.70	79.68	80.92	84.59	83.98
89	76.55	80.31	81.42	84.97	84.43
85	76.77	80.48	81.55	85.06	84.54
88	76.65	80.39	81.48	85.01	84.48
91	76.37	80.18	81.31	84.89	84.33
91	75.64	79.64	80.88	84.56	83.95
78	73.25	77.85	79.49	83.50	82.65
67	69.51	75.00	77.41	81.81	80.49
55	61.39	75.00	73.26	78.00	75.01
48	53.03	70.69	69.61	70.69	67.45
33	45.24	57.58	55.62	57.76	55.62
36	39.48	51.49	44.47	51.63	44.47
33	36.23	47.71	40.64	47.83	38.18
28	35.27	46.53	40.21	46.65	36.32
27	36.24	47.72	40.65	47.85	38.20
31	38.34	50.19	42.26	50.33	42.26
39	41.07	53.24	47.55	53.40	47.55
45	42.75	55.03	50.80	55.20	50.80
58	44.08	56.41	53.37	56.58	53.37
59	45.17	57.51	55.48	57.69	55.48
63	46.22	58.56	57.51	58.74	57.51
71	47.13	59.45	59.27	59.64	59.27
73	47.90	60.77	60.77	60.77	60.77
75	48.56	62.04	62.04	62.04	62.04
79	49.13	63.15	63.15	63.15	62.52
85	49.61	64.07	64.07	64.07	63.21
88	50.02	64.87	64.87	64.87	63.78
88	50.37	65.55	65.55	65.55	64.25
91	50.67	66.13	66.13	66.13	64.64
90	50.13	65.08	65.08	65.08	63.93
77	46.59	58.92	58.23	59.11	58.23
54	42.58	54.85	50.47	55.02	50.47
47	36.59	48.14	40.80	48.27	38.87
41	33.10	43.76	39.06	43.86	32.12
39	30.44	40.10	37.19	40.18	26.97
39	28.84	37.73	35.76	37.80	25.00
40	28.44	37.11	35.35	37.18	25.00
34	29.22	38.30	36.12	38.38	25.00
31	30.61	40.34	37.33	40.43	27.30
37	32.87	43.45	38.92	43.56	31.67
46	35.85	47.25	40.48	47.37	37.44
59	38.00	49.80	41.60	49.94	41.60
69	39.82	51.87	45.13	52.02	45.13
73	41.39	53.59	48.16	53.74	48.16
75	42.78	55.06	50.86	55.23	50.86
78	44.01	56.34	53.24	56.51	53.24
80	45.06	57.40	55.27	57.58	55.27
82	45.97	58.31	57.03	58.49	57.03
81	46.75	59.08	58.54	59.26	58.54
80	47.41	59.82	59.82	59.91	59.82
84	47.99	60.94	60.94	60.94	60.94
85	48.49	61.91	61.91	61.91	61.91
83	48.93	62.76	62.76	62.76	62.76
82	48.49	61.91	61.91	61.91	61.91
72	45.22	57.56	55.58	57.74	55.58
61	41.57	53.78	48.51	53.94	48.51
51	36.00	47.43	40.55	47.55	37.73
40	32.70	43.23	38.81	43.33	31.35
32	30.21	39.77	37.00	39.85	26.53
34	28.73	37.56	35.65	37.63	25.00
30	28.39	37.03	35.30	37.10	25.00
29	29.19	38.26	36.09	38.33	25.00
30	30.58	40.30	37.31	40.38	27.24
41	32.80	43.36	38.88	43.46	31.54
45	35.74	47.11	40.43	47.23	37.23
59	37.85	49.63	41.31	49.76	41.31
64	39.65	51.68	44.80	51.83	44.80
70	41.21	53.39	47.82	53.55	47.82
75	42.58	54.85	50.47	55.02	50.47
73	43.80	56.12	52.83	56.29	52.83

A.4 Umidade[%] – Inverno

A.4.1 PS

81	83.40	84.87	80.93	84.87	79.00
80	84.43	86.04	81.38	86.04	79.67
81	85.39	87.14	81.81	87.06	80.29
86	86.24	88.11	82.19	87.96	80.83
86	86.97	88.95	82.51	88.74	81.28
84	87.61	89.68	82.80	89.41	81.67
80	88.17	90.32	83.05	90.01	82.01
80	88.68	90.90	83.28	90.54	82.32
75	85.84	87.65	82.01	87.54	80.58
70	78.91	79.74	79.00	79.74	75.88
65	71.57	71.35	71.35	71.35	71.35
58	61.43	59.77	59.77	59.85	59.77
53	53.01	50.15	50.15	53.28	50.15
51	47.16	43.47	43.47	48.38	43.47
44	44.79	41.56	40.76	46.31	40.76
41	45.83	42.16	41.95	47.23	41.95
41	49.00	45.57	45.57	49.96	45.57
47	54.57	51.93	51.93	54.54	51.93
59	59.85	57.96	57.96	58.65	57.96
73	63.63	62.28	62.28	62.28	62.28
75	66.82	65.93	65.93	65.93	65.93
76	69.67	69.18	69.18	69.18	69.18
79	72.32	72.21	72.21	72.21	70.69
83	74.79	75.03	75.03	75.03	72.73
87	77.09	75.47	75.47	75.47	73.42
87	79.19	77.87	77.87	77.87	75.15
87	81.06	80.00	80.00	80.00	76.62
85	82.68	81.84	80.88	81.84	77.85
85	84.08	83.44	81.52	83.44	78.88
83	85.28	84.80	82.07	84.80	79.74
83	86.31	85.98	82.54	85.98	80.47
83	87.16	86.95	82.94	86.89	81.05
78	82.82	82.00	80.95	82.00	77.95
72	72.03	69.71	69.71	69.71	69.71
64	62.80	59.19	59.19	59.44	59.19
58	53.04	48.07	48.07	51.34	48.07
53	45.03	39.62	38.94	44.10	38.94
46	41.16	37.31	37.00	40.37	43.50
42	40.48	36.89	36.63	39.69	42.53
41	42.25	37.98	37.58	41.44	35.77
40	45.89	40.12	39.92	44.91	39.92
54	51.59	46.41	46.41	50.08	46.41
60	56.90	52.46	52.46	54.63	52.46
64	60.81	56.92	56.92	57.84	56.92
73	64.12	60.69	60.69	60.69	60.69
70	67.03	64.01	64.01	64.01	64.01
70	69.66	67.01	67.01	67.01	67.01
73	72.09	69.77	69.77	69.77	69.77
75	74.31	74.72	74.72	74.72	72.49
73	76.30	76.80	76.80	76.80	73.93
79	78.07	78.66	78.53	78.66	75.15
78	79.60	80.27	79.12	80.27	76.18
81	80.93	81.66	79.64	81.66	77.04
83	82.09	82.88	80.09	82.88	77.78
81	83.09	83.93	80.49	83.93	78.40
82	83.92	84.80	80.82	84.80	78.91
72	80.31	81.01	79.40	81.01	76.64
66	71.28	71.54	71.54	71.54	71.54
57	63.35	63.22	63.22	63.22	63.22
49	54.28	53.70	53.70	55.43	53.70
42	46.51	45.54	45.54	49.17	45.54
36	42.80	41.65	41.65	46.03	41.65
37	42.12	40.94	40.94	45.44	40.94
35	43.76	42.66	42.66	46.86	42.66
38	47.15	46.22	46.22	49.70	46.22
42	52.41	51.74	51.74	53.96	51.74
54	57.27	56.84	56.84	57.73	56.84
57	60.89	60.63	60.63	60.63	60.63
61	64.00	63.90	63.90	63.90	63.90
63	66.72	66.75	66.75	66.75	66.75
66	69.18	69.33	69.33	69.33	69.33
68	71.43	71.69	71.69	71.69	71.69

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.956	0.953	0.967	0.946	0.964	1
	0.977	0.957	0.956	0.965	0.945	2
	0.972	0.976	0.976	0.955	0.973	3
Melhora % “D”		-0.3%	1.1%	-1.1%	0.8%	1
		-2.0%	-2.2%	-1.2%	-3.3%	2
		0.4%	0.4%	-1.7%	0.1%	3
MAPE	8.02%	8.69%	6.78%	9.15%	6.83%	1
	5.80%	9.02%	8.89%	7.61%	8.83%	2
	8.63%	8.02%	7.85%	10.72%	8.32%	3
Melhora % MAPE		-8.34%	15.41%	-14.03%	14.82%	1
		-55.70%	-53.33%	-31.25%	-52.37%	2
		7.03%	9.11%	-24.17%	3.66%	3
Melhora % MAPE		-8.34%	15.41%	-14.03%	14.82%	1
		-55.70%	-53.33%	-31.25%	-52.37%	2
		7.03%	9.11%	-24.17%	3.66%	3
Data de medidas:		20/06/2009				

A.4.2 SAF1

Exp	Sim	Conf1	Conf2	Conf3	Conf4
79	83.53	75.87	77.71	72.82	74.50
79	84.58	76.57	78.30	73.72	75.43
78	85.55	77.21	78.84	74.55	76.26
83	86.42	77.77	79.33	75.28	77.00
84	87.15	78.25	79.74	75.90	77.61
83	87.79	78.66	80.11	76.43	78.13
80	88.35	79.02	80.42	76.90	78.59
79	88.86	79.34	80.71	77.32	79.00
74	86.65	77.92	79.46	75.48	77.19
69	81.89	74.77	76.81	71.40	73.01
63	76.81	70.24	70.24	66.88	68.04
59	69.73	60.33	60.33	60.23	60.33
54	64.16	52.53	52.53	52.53	52.53
51	60.07	46.81	46.81	46.81	46.81
45	58.04	43.97	43.97	43.97	43.97
41	58.49	44.60	44.60	44.60	44.60
44	60.60	47.55	47.55	47.55	47.55
46	64.87	53.53	53.53	53.53	53.53
62	68.70	58.89	58.89	58.89	58.89
74	71.35	62.60	62.60	61.79	62.60
75	73.61	65.76	65.76	63.93	65.76
75	75.62	68.57	68.57	65.79	68.57
77	77.46	71.15	71.15	67.47	68.71
83	79.18	72.91	73.55	69.01	70.43
84	80.79	74.51	74.51	70.38	71.66
85	82.32	75.86	76.75	71.91	73.29
84	83.79	76.92	78.60	73.37	74.80
83	85.17	77.91	79.45	74.72	76.18
82	86.43	78.79	80.22	75.95	77.40
81	87.51	79.54	80.89	76.99	78.43
82	88.39	80.15	81.44	77.83	79.25
82	89.09	80.63	81.87	78.49	79.89
77	85.58	78.19	79.70	75.12	76.58
72	78.09	70.54	70.54	67.63	68.67
65	71.16	60.37	60.37	60.28	60.37
58	62.93	48.29	48.29	48.29	48.29
52	57.11	39.74	39.74	39.74	39.74
46	53.83	34.93	34.93	34.93	34.93
40	52.88	33.53	33.53	33.53	33.53
41	53.98	35.15	35.15	35.15	35.15
43	56.63	39.04	39.04	39.04	39.04
51	61.15	45.67	45.67	45.67	45.67
59	65.13	51.52	51.52	51.52	51.52
65	67.97	55.69	55.69	55.69	55.69
72	70.42	59.28	59.28	59.28	59.28
71	72.62	62.51	62.51	61.86	62.51
71	74.64	65.48	65.48	64.02	65.48
71	76.52	68.24	68.24	66.00	68.24
72	78.27	71.27	73.27	68.14	69.70
71	79.94	72.36	74.89	69.46	71.17
77	81.54	73.40	75.70	70.71	72.53
77	82.98	74.32	76.44	71.82	73.72
78	84.18	75.08	77.06	72.74	74.67
81	85.13	75.67	77.55	73.46	75.41
78	85.88	76.14	77.94	74.02	75.98
79	86.48	76.51	78.25	74.47	76.43
71	83.36	74.56	76.63	72.11	74.02
66	76.88	70.34	71.38	67.02	71.38
57	70.95	63.31	63.31	62.10	63.31
47	63.48	53.15	53.15	53.15	53.15
41	57.80	45.42	45.42	45.42	45.42
40	54.54	40.99	40.99	40.99	40.99
38	53.60	39.71	39.71	39.71	39.71
35	54.65	41.14	41.14	41.14	41.14
37	57.17	44.57	44.57	44.57	44.57
42	61.44	50.38	50.38	50.38	50.38
53	65.15	55.42	55.42	55.42	55.42
57	67.83	59.07	59.07	59.07	59.07
59	70.16	62.24	62.24	61.43	62.24
63	72.25	65.08	65.08	63.21	65.08
64	74.18	67.71	67.71	64.82	67.71
65	75.97	69.73	70.14	66.29	70.14

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D”	0.849	0.956	0.959	0.932	0.950	1
Willmott	0.935	0.950	0.954	0.929	0.940	2
	0.832	0.984	0.985	0.982	0.984	3
Melhora %		12.7%	13.0%	9.7%	11.9%	1
“D”		1.6%	2.0%	-0.6%	0.6%	2
		18.3%	18.4%	18.1%	18.3%	3
MAPE	14.3%	6.79%	6.60%	7.91%	7.04%	1
	10.5%	9.99%	9.60%	11.69%	10.73%	2
	22.69%	7.02%	6.78%	6.92%	7.23%	3
Melhora %		52.64%	53.97%	44.87%	50.90%	1
MAPE		4.84%	8.59%	-11.39%	-2.19%	2
		69.06%	70.10%	69.50%	68.14%	3
Melhora %		52.64%	53.97%	44.87%	50.90%	1
MAPE		4.84%	8.59%	-11.39%	-2.19%	2
		69.06%	70.10%	69.50%	68.14%	3
Data de medidas:						20/06/2009

A.4.3 SAF2

Exp	Sim	Conf1	Conf2	Conf3	Conf4
89	83.50	83.12	84.40	82.32	80.93
84	84.55	83.95	85.43	83.19	82.03
86	85.52	84.71	86.01	83.99	83.02
89	86.38	85.37	86.54	84.70	83.87
89	87.12	85.94	86.99	85.30	84.59
86	87.75	86.42	87.38	85.81	85.19
82	88.31	86.84	87.73	86.26	85.72
82	88.82	87.23	88.04	86.66	86.19
77	86.68	85.60	86.72	84.94	84.16
71	81.94	81.88	81.89	81.00	79.22
63	76.88	73.76	73.76	73.76	73.76
57	69.85	62.46	62.46	62.46	62.46
51	64.29	53.52	53.52	55.67	53.52
49	60.23	47.00	47.00	51.07	47.00
42	58.22	45.14	43.77	48.68	43.77
42	58.66	45.50	44.47	49.21	44.47
40	60.74	47.82	47.82	51.66	47.82
46	64.98	54.63	54.63	56.43	54.63
59	68.79	60.75	60.75	60.75	60.75
75	71.41	64.97	64.97	64.97	64.97
76	73.65	68.57	68.57	68.57	68.57
78	75.69	71.84	71.84	71.84	71.84
79	77.59	74.90	74.90	74.90	74.90
84	79.37	77.76	77.76	77.76	77.76
89	81.03	78.73	78.73	78.73	78.73
89	82.60	81.41	81.41	80.77	78.99
89	84.11	83.28	83.98	82.16	80.78
89	85.48	84.43	85.80	83.40	82.33
87	86.65	85.40	86.56	84.45	83.60
86	87.59	86.17	87.18	85.28	84.60
85	88.33	86.77	87.67	85.93	85.36
84	88.89	87.23	88.04	86.42	85.93
81	85.49	84.44	85.81	83.41	82.34
75	78.29	74.07	74.07	74.07	74.07
63	71.77	62.96	62.96	62.96	62.96
56	63.87	49.50	49.50	52.80	49.50
50	57.88	42.62	41.02	45.12	39.30
45	54.38	39.26	38.09	40.24	41.65
42	53.31	38.17	37.08	38.67	39.37
42	54.39	39.27	38.10	40.25	41.67
40	57.02	41.82	40.35	43.95	47.27
52	61.49	45.84	45.45	49.84	45.45
67	65.43	52.16	52.16	54.68	52.16
66	68.25	56.96	56.96	57.98	56.96
74	70.67	61.09	61.09	61.09	61.09
75	72.84	64.78	64.78	64.78	64.78
75	74.84	68.19	68.19	68.19	68.19
77	76.74	71.43	71.43	71.43	71.43
79	78.63	76.61	76.61	76.61	76.61
76	80.53	79.65	79.65	79.65	77.19
82	82.18	80.07	82.29	81.06	78.93
80	83.46	81.06	83.30	81.99	80.21
84	84.43	81.80	83.83	82.69	81.15
87	85.20	82.39	84.26	83.24	81.87
83	85.81	82.85	84.60	83.67	82.43
84	86.34	83.25	84.89	84.05	82.91
78	83.34	80.97	83.23	81.90	80.09
70	77.03	74.05	74.05	74.05	74.05
56	71.29	64.86	64.86	64.86	64.86
46	63.96	53.13	53.13	55.45	53.13
38	58.24	45.42	43.98	48.94	43.98
36	54.88	42.55	40.96	44.84	38.60
37	53.86	41.64	40.19	43.55	46.86
36	54.89	42.56	40.97	44.85	38.62
38	57.38	44.70	42.68	47.92	42.60
43	61.62	49.39	49.39	52.85	49.39
57	65.33	55.33	55.33	56.93	55.33
61	68.01	59.61	59.61	59.75	59.61
65	70.33	63.33	63.33	63.33	63.33
69	72.41	66.66	66.66	66.66	66.66
70	74.33	69.73	69.73	69.73	69.73
71	76.16	72.66	72.66	72.66	72.66

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
“D” Willmott	0.862	0.963	0.964	0.951	0.964	1
	0.940	0.962	0.961	0.965	0.953	2
Melhora % “D”	0.859	0.944	0.998	0.995	0.993	3
		11.7%	11.8%	10.3%	11.8%	1
MAPE	14.92%	2.4%	2.3%	2.7%	1.4%	2
		9.8%	16.2%	15.8%	15.5%	3
		7.97%	7.77%	9.45%	7.83%	1
Melhora % MAPE	10.13%	8.31%	8.71%	7.37%	8.84%	2
		18.90%	7.34%	6.48%	9.20%	3
		46.56%	47.91%	36.66%	47.56%	1
Melhora % MAPE	17.89%	14.00%	14.00%	27.21%	12.67%	2
		61.15%	65.73%	51.32%	65.30%	3
		46.56%	47.91%	36.66%	47.56%	1
Melhora % MAPE	17.89%	14.00%	14.00%	27.21%	12.67%	2
		61.15%	65.73%	51.32%	65.30%	3
		Data de medidas: 20/06/2009				

A.4.4 SAF3

Exp	Sim	Conf1	Conf2	Conf3	Conf4
90	83.90	80.00	79.99	82.40	80.08
90	84.94	80.00	80.51	82.97	80.94
93	85.90	80.00	81.00	83.50	81.71
92	86.75	80.00	81.43	83.96	82.36
94	87.49	80.00	81.81	84.36	82.92
92	88.12	80.00	82.14	84.70	83.38
89	88.68	80.00	82.44	85.00	83.78
88	89.18	80.21	82.70	85.27	84.14
80	87.17	80.00	81.65	84.19	82.68
63	82.49	80.00	79.31	81.61	78.85
58	77.62	76.15	76.15	76.15	76.15
55	71.06	63.86	63.86	63.86	63.86
50	65.79	53.98	53.98	56.29	53.98
48	61.96	48.19	46.81	51.60	46.81
44	60.01	46.27	43.85	49.08	43.16
43	60.46	46.72	44.17	49.67	44.00
44	62.63	48.83	48.06	52.44	48.06
52	66.90	56.06	56.06	57.60	56.06
61	70.63	63.05	63.05	63.05	63.05
77	73.10	67.68	67.68	67.68	67.68
85	75.16	71.54	71.54	71.54	71.54
82	77.00	74.99	74.99	74.99	74.99
85	78.70	78.17	77.54	78.17	75.10
90	80.53	80.00	78.38	80.49	77.00
89	82.48	80.00	79.52	81.58	78.85
89	84.06	80.00	80.39	82.58	80.39
90	85.21	80.00	81.04	83.30	81.45
90	86.06	80.00	81.52	83.82	82.19
88	86.74	80.00	81.92	84.24	82.77
86	87.32	80.00	82.26	84.60	83.26
87	87.83	80.00	82.56	84.90	83.67
88	88.30	80.40	82.83	85.19	84.05
82	85.50	80.00	81.20	83.48	81.70
66	79.21	78.28	77.80	78.28	75.23
60	73.29	66.05	66.05	66.05	66.05
55	65.94	50.90	50.86	54.46	50.86
49	60.24	44.62	42.62	46.54	39.09
46	56.78	40.34	39.06	41.17	43.98
44	55.61	38.78	37.65	39.22	40.65
43	56.61	40.12	38.86	40.89	43.50
39	59.20	43.38	41.64	44.98	36.94
54	63.55	48.37	45.93	51.26	45.93
65	67.32	53.71	53.71	56.23	53.71
72	69.93	59.11	59.11	59.48	59.11
77	72.16	63.71	63.71	63.71	63.71
82	74.12	67.76	67.76	67.76	67.76
82	75.95	71.54	71.54	71.54	71.54
81	78.10	75.98	75.98	75.98	75.98
85	80.02	80.00	78.51	80.07	76.42
87	81.39	80.00	79.18	80.91	77.82
86	82.38	80.00	79.67	81.50	78.78
84	83.19	80.00	80.08	81.98	79.53
86	83.90	80.00	80.44	82.40	80.16
91	84.56	80.00	80.79	82.79	80.74
87	85.18	80.00	81.11	83.15	81.27
86	85.77	80.00	81.42	83.50	81.76
79	83.31	80.00	80.14	82.06	79.64
55	77.65	75.76	75.76	75.76	75.76
52	72.31	65.70	65.70	65.70	65.70
47	65.47	52.81	52.81	55.73	52.81
41	60.01	46.79	44.22	49.08	42.53
37	56.76	43.29	41.57	44.75	36.40
38	55.69	42.08	40.57	43.25	47.85
39	56.69	43.22	41.51	44.65	36.27
39	59.20	45.95	43.62	48.03	41.00
45	63.37	50.16	48.86	53.25	48.86
58	66.97	55.64	55.64	57.44	55.64
65	69.46	60.33	60.33	60.33	60.33
74	71.61	64.38	64.38	64.38	64.38
78	73.50	67.94	67.94	67.94	67.94
77	75.52	71.75	71.75	71.75	71.75
74	77.60	75.67	75.67	75.67	75.67

Análise estatística:						
	Sim	Conf. 1	Conf. 2	Conf. 3	Conf. 4	Dia
"D"	0.835	0.903	0.918	0.919	0.921	1
Willmott	0.908	0.935	0.942	0.954	0.947	2
	0.833	0.945	0.952	0.945	0.949	3
Melhora %		8.1%	9.9%	10.1%	10.3%	1
"D"		3.0%	3.8%	5.1%	4.3%	2
		13.5%	14.3%	13.4%	13.9%	3
MAPE	16.94%	11.27%	10.44%	11.61%	10.36%	1
	12.51%	11.00%	10.88%	8.71%	9.63%	2
	19.87%	11.07%	9.85%	11.71%	9.97%	3
Melhora %		33.47%	38.37%	31.46%	38.83%	1
MAPE		12.08%	13.06%	30.42%	23.05%	2
		44.28%	50.42%	41.04%	49.79%	3
Melhora %		33.47%	38.37%	31.46%	38.83%	1
MAPE		12.08%	13.06%	30.42%	23.05%	2
		44.28%	50.42%	41.04%	49.79%	3
Melhora %		33.47%	38.37%	31.46%	38.83%	1
MAPE		12.08%	13.06%	30.42%	23.05%	2
		44.28%	50.42%	41.04%	49.79%	3
Data de medidas:						20/06/2009

Apêndice B

Algoritmo computacional

B.1 Sobre o algoritmo computacional desenvolvido

Software utilizado

As computações simbólicas feitas pelo algoritmo computacional foram feitas utilizando o software *SAGE* (<http://www.sagemath.org>), que é um software matemático *open-source* que fornece uma interface que permite, através da linguagem de programação *Phyton*, a utilização de várias bibliotecas e ferramentas de software livre já existentes. Seu objetivo é criar uma alternativa *open-source* ao Magma, Maple, Mathematica e Matlab. A sintaxe da linguagem utilizada (Python) é muito similar à outras linguagens como Matlab ou C. Tecnicamente, a rotina desenvolvida realiza operações algébricas, demandando a utilização de recursos da matemática simbólica (realizada por softwares comerciais como *Mathematica* e *Maple*).

Justificativa da rotina computacional

Observando o exemplo da seção 3.3, nota-se que o procedimento algébrico envolvido deve ser repetido para cada uma das previsões. Como será explicado na seção 4.1.2, a análise da eficácia da metodologia não é realizada para cada previsão individualmente, mas para um conjunto de previsões horárias correspondentes a um dia inteiro. Ou seja, o processo do exemplo 3.3 seria repetido 24 vezes para que se pudesse fazer a análise de eficácia mínima do trabalho, o que inviabiliza a realização de uma rotina de testes sem

que haja uma rotina computacional que realize o procedimento algébrico envolvido.

Comentários sobre tempo de execução

Cada conjunto de dados, correspondente a três dias de testes (conforme a seção 4.1.2), demorou-se em média 10 minutos para realizar seus cálculos. Isto é, aproximadamente 10 minutos de processamento computacional foi suficiente para, fornecidas as duas distribuições prioris e os 72 pares (três dias) de medidas experimentais e previsões, retornar as quatro últimas colunas da tabela à esquerda de algum dos itens do apêndice A.

Como os cálculos correspondentes a cada conjunto de dados leva em torno de 10 minutos, e foram realizados os cálculos em 12 conjuntos de dados (conforme a seção 4.2.3), todos os testes deste trabalho levaram em média 2 horas (120 minutos) para serem realizados.

Uma questão interessante que surge sobre tempo de execução é se ele aumenta se forem utilizados mais dados para testar a abordagem. A resposta é que o tempo essencialmente seria o mesmo, uma vez que apenas a obtenção da verossimilhança é alterada pela quantidade de dados utilizados para testes. Uma quantidade maior de dados de treinamento implica somente em um cálculo de coeficiente de mínimos quadrados utilizando mais pontos, e cálculo de mínimos quadrados é uma tarefa extremamente simples do ponto de vista computacional se comparado ao tempo de processamento necessário para obter a posteriori nas quatro combinações t-norma/priori. Dessa forma, não há nenhum tipo de proporcionalidade entre a quantidade de dados utilizados para o teste e o tempo de execução.

B.1.1 Refazendo cálculos da dissertação usando o software

Nesta seção, os cálculos do método realizados durante a dissertação são repetidos utilizando o *software* desenvolvido. Assim pode-se observar como as entradas do programa são fornecidas, e como são retornadas as saídas.

B.1.1.1 Exemplo da seção 3.3

Distribuição posteriori Obtém-se abaixo as duas distribuições posteriori correspondentes aos dois casos de t-norma (produto e mínimo).

```
# Dados do exemplo da seção 3.3:
D= [(2.5, 1.466666666666667), (3, 2.3), (4, 5/3), (4, 3.266666666666667), (5, 3.033333333333333),
    (5.2, 4.466666666666667), (5.5, 3.166666666666667), (5.6, 2.733333333333333), (5.8, 4.066666666666667),
    (5.9, 4.933333333333333), (6.2, 2.633333333333333), (6.5, 5.333333333333333), (6.7, 4.166666666666667),
    (6.9, 4.7), (7.2, 3.8), (7.7, 5.833333333333333), (8, 5.633333333333333), (8.3, 3.233333333333333),
    (8.9, 5.433333333333333), (9.2, 8.133333333333333), (9.5, 4.333333333333333), (9.8, 8.133333333333333),
    (10, 11/3)]

# definição da priori trapezoidal
priori = trapez(2,5,8,10) # (EQ. 3.3.1)
# obtém-se as funções u(x), a(x) e b(x):
u,a,b= cria_func(D)
print '>>> u(x)=',u
>> u(x)= 0.57491725090567836*x + 0.395254849835 # (EQ. 3.4.13)
print '>>> a(x)=',a
>> a(x)= 0.163668899019*x + 0.395254849835 # (EQ. 3.4.11)
print '>>> b(x)=',b
>> b(x)= 1.00074597331699*x + 0.395254849835 # (EQ. 3.4.12)
f = likelihood(u,a,b)
print '>>>', f # (EQ. 3.3.6, FIG. 3.4)
>> [[0.999254582744*y - 0.39496022005, 1.73938075162*y - 0.687498677788], (2.35011383838*x
- 2.34836202297*y + 0.928201478747)/x],
    [[1.73938075162*y - 0.687498677788, 6.10989629669*y - 2.41496614326], (-0.397980680696*x
+ 2.43162068714*y - 0.961109869551)/x]]
# previsão 'Y':
Y = 2
# substitui-se o valor de previsão 'Y=2' na função de verossimilhança
f.Y = likelihood.y(f,Y)
print '>>>', f.Y # (EQ. 3.3.7, FIG. 3.5)
>> Piecewise defined function with 2 parts,[
    [(1.60354894544, 2.79126282545), (2.35011383838*x - 3.76852256719)/x],
    [(2.79126282545, 9.80482645013), (-0.397980680696*x + 3.90213150473)/x]]
# obtém a posteriori para uma das duas cada t-Normas
post_produto= posteriori(priori,f.Y,'produto')
print '>>>', post_produto # (EQ. 3.3.10, FIG. 3.6)
>> Piecewise defined function with 6 parts,[
    [(1.60354894544, 2), 0], [(2, 2.79126282545), 0.852257669034*(x - 2)*(2.35011383838*x - 3.76852256719)/x],
    [(2.79126282545, 5), 0.852257669034*(-0.397980680696*x + 3.90213150473)*(x - 2)/x],
    [(5, 8), 2.5567730071*(-0.397980680696*x + 3.90213150473)/x],
    [(8, 9.80482645013), -1.27838650355*(-0.397980680696*x + 3.90213150473)*(x - 10)/x],
    [(9.80482645013, 10), 0]]
post_minimo = posteriori(priori,f.Y,'minimo')
print '>>>', post_minimo # (EQ. 3.3.13, FIG. 3.8)
>> Piecewise defined function with 4 parts,[
    [(1.6035489454392151, 2.0), 0],
    [(2, 3.848145346584749), 1/3*x - 2/3],
```

```

      [(3.848145346584749, 9.80482645013), (-0.397980680696*x + 3.90213150473)/x],
      [(9.80482645013, 10), 0]]
# escreve o valor de máximo das distribuições posteriori
print '>>> produto: ',sup.Piecewise(post_produto),'; mínimo:',sup.Piecewise(post_minimo)
>>> produto: 4.42827876135 ; mínimo: 3.84814545686
# exatamente os pontos de máximo das posteriori das Figuras 3.6 e 3.8

```

B.1.1.2 Exemplo de testes da aplicação (seção 4.4)

Função de verossimilhança De acordo com a metodologia proposta na seção 3.4.2, são obtidas as funções $u(x)$, $a(x)$ e $b(x)$ (ajuste linear e retas-limite) para a figura 4.4.

```

# utiliza-se os 48 primeiros pares do apêndice A.1.3, conforme a seção 4.4.1:
D= [[292.15,292.73],[290.45,291.97],[291.35,291.42],[291.45,291.06],[290.75,290.82],[289.95,290.65],
    [289.95,290.52],[291.05,290.57],[293.75,291.31],[297.45,292.44],[300.35,294.74],[302.15,296.72],
    [303.75,298.62],[302.65,300.25],[303.95,301.24],[304.05,301.32],[303.45,300.64],[301.95,299.43],
    [296.95,297.95],[294.35,296.98],[293.45,296.22],[291.95,295.61],[293.05,295.08],[291.25,294.58],
    [289.25,294.14],[289.95,293.79],[289.15,293.53],[287.95,293.32],[287.75,293.14],[287.05,292.99],
    [285.55,292.87],[288.55,293.08],[292.65,294.44],[297.95,295.99],[299.95,298.58],[301.45,300.37],
    [302.75,301.91],[302.85,302.86],[303.45,303.04],[303.75,302.48],[304.25,301.62],[302.55,300.32],
    [300.05,298.70],[297.45,297.58],[294.55,296.71],[292.65,296.00],[293.45,295.40],[294.05,294.88]]
# obtém-se as funções u(x), a(x) e b(x):
u,a,b= cria_func(D)
# escrevendo as funções:
print '>>> u(x)=',u
>>> u(x)= 0.56397026364630365*x + 129.323616181
print '>>> a(x)=',a
>>> a(x)= 0.54298433330412000*x + 129.323616181
print '>>> b(x)=',b
>>> b(x)= 0.57614191411161200*x + 129.323616181
f=likelihood(u,a,b)
print '>>> f=',Piecewise(f)
# abaixo escreve-se a função de verossimilhança:
>>> f= Piecewise defined function with 2 parts, [
    [[1.73568347573*y - 224.464863628, 1.77314313264*y - 229.309281921], (47.3347403258*x
- 82.1581266115*y + 10624.9860321)/x],
    [[1.77314313264*y - 229.309281921, 1.84167376233*y - 238.17191077], (-25.8737318027*x
+ 47.6509729945*y - 6162.39614221)/x]] # (EQ. 4.4.7)

```

Testes

```

# onde a configuração "Temperatura, SAF2 - Verão" denota-se no programa por "temp_S2_1"
# (se fosse a mesma configuração para inverno seria "temp_S2_2")
testaDias('temp_S2.1',nTeste=3) # "nTeste=3" pelo fato de o dia de teste ser o terceiro
292,77 292,77 292,77 292,77
292,14 292,14 292,14 292,14
291,62 291,62 291,62 291,62
291,21 291,21 291,21 291,21

```

290,88	290,88	290,88	290,88
290,58	290,58	290,58	290,58
290,35	290,35	290,35	290,35
290,74	290,74	290,74	290,74
293,08	293,08	293,08	293,08
295,68	295,68	295,45	295,68
300,13	299,45	298,34	300,13
301,88	300,68	300,36	303,25
303,18	301,80	302,04	299,40
303,98	302,54	303,08	301,00
304,15	302,70	303,30	301,33
303,68	302,26	302,69	300,39
302,95	301,60	301,74	298,95
301,83	300,64	300,29	303,15
300,33	299,53	298,47	300,33
298,41	298,41	297,22	298,41
296,89	296,89	296,23	296,89
295,65	295,65	295,42	295,65
294,60	294,60	294,60	294,60
293,71	293,71	293,71	293,71

Os valores retornados pela função `testaDias('temp_S2_1',nTeste=3)` correspondem exatamente aos 24 últimos registros (terceiro dia) das colunas “Conf1”, “Conf2”, “Conf3”, “Conf4” do apêndice A.1.3. Por exemplo, o comando `testaDias('temp_S2_1',nTeste=2)`, correspondem à situação onde o segundo dia é o grupo de testes (e o grupo de treinamento usado para calibração corresponde ao primeiro e terceiro dias, conforme Tabela 4.2). O comando `testaDias('temp_S2_1')`, sem especificação nenhuma do dia de testes, realiza os três testes. Ou seja, os comandos `testaDias('temp_S2_1',nTeste=1)`, `testaDias('temp_S2_1',nTeste=2)` e `testaDias('temp_S2_1',nTeste=3)` executados em sequência trazem exatamente o mesmo resultado que apenas o comando `testaDias('temp_S2_1')`, reproduzindo exatamente as quatro colunas “Conf1”, “Conf2”, “Conf3”, “Conf4” do apêndice A.1.3.

B.1.2 Rotina computacional para cálculo de mínimo entre funções

Na seção 3.3.2.2, a função, a distribuição conjunta $\pi^{(\wedge)}(x, 2) = \min\{\pi(x), \pi(x|2)\}$ foi calculada utilizando um gráfico (Figura 3.7). Computacionalmente, a abordagem gráfica é inviável; Uma forma analítica de obter a distribuição conjunta para a t-norma

do mínimo é criar uma partição para os valores de x utilizando o conjunto

$$P = S_1 \cup S_2 \cup S_3 = \{z : \pi(x = z) = \pi(2|z)\} \cup \{z : \pi(x = z) = 0\} \cup \{z : \pi(2|z) = 0\}.$$

Para cada par de elementos consecutivos z_i e z_{i+1} da partição P , deve-se checar, no intervalo (z_i, z_{i+1}) , qual das duas funções é $f_i(x)$, a menor entre as duas. A função $\pi^{(\wedge)}(x, 2) = \min\{\pi(x), \pi(x|2)\}$ é definida por “ $f_i(x)$ se $x \in (z_i, z_{i+1})$ ” para cada par da partição P .

Para o caso do exemplo (Fig. 3.7), $\pi(x)$ dada por (3.3.1) e $\pi(y|z)$ dada por (3.3.7) tem-se:

$$P = \{z : \pi(x = z) = \pi(2|x = z)\} \cup \{z : \pi(x = z) = 0\} \cup \{z : \pi(2|x = z) = 0\}$$

$$P = \{3.85\} \cup \{2, 10\} \cup \{1.6, 9.8\}$$

$$P = \{1.6, 2, 3.85, 9.8, 10\}.$$

Calculando o mínimo para cada um dos intervalos:

- **Intervalo** $[z_1, z_2) = [1.6, 2)$: $\forall x \in [1.6, 2)$, $\pi(x) = \pi(2|x) = 0 \implies f_1(x) = 0$
- **Intervalo** $[z_2, z_3) = [2, 3.85)$: $\forall x \in [2, 3.85)$, $\pi(x) \leq \pi(2|x) \implies f_2(x) = \pi(x)$
- **Intervalo** $[z_3, z_4) = [3.85, 9.8)$: $\forall x \in [3.85, 9.8)$, $\pi(2|x) \leq \pi(x) \implies f_3(x) = \pi(2|x)$
- **Intervalo** $[z_4, z_5] = [9.8, 10]$: $\forall x \in [9.8, 10]$, $\pi(x) = \pi(2|x) = 0 \implies f_4(x) = 0$

Dessa forma,

$$\pi^{(\wedge)}(x, 2) = \begin{cases} f_1 = 0 & \text{se } 1.6 \leq x < 2 \\ f_2 = \pi(x) & \text{se } 2 \leq x < 3.85 \\ f_3 = \pi(2|x) & \text{se } 3.85 \leq x < 9.8 \\ f_4 = 0 & \text{se } 9.8 \leq x \leq 10 \end{cases}$$

$$\pi^{(\wedge)}(x, 2) = \begin{cases} \frac{x-2}{3} & \text{se } 2 < x \leq 3.85 \\ \frac{-0.40x+3.90}{x} & \text{se } 3.85 < x < 9.8, \\ 0 & \text{caso contrário} \end{cases}$$

que é exatamente a equação (3.3.11).

Na rotina computacional desenvolvida neste trabalho, a tarefa descrita nesta seção é realizada pela função `min_Piecewise`.

B.2 Funções principais

```
# Importação do módulo estatístico
import scipy.stats as Stat
x,y = var('x,y')
*****
*****
def testaDias(param,nTeste=None):
# Esta é a função principal do programa: é a que reúne todas as outras funções e
# realiza os testes.
    priori = []; D=[]
    D,priori= dados(param,D,priori)
    n_partes= 3
    # Caso nenhum dia de teste seja especificado, faz-se as os três testes correspondentes
    if nTeste==None:
        P= range(n_partes)
    else:
        nTeste=nTeste-1
        P= [nTeste]
    for nTeste in P:
        dTeste,dTreino= separaDados(parteDadosEmDias(D),nTeste)
        u,a,b = cria_func(D=dTreino)
        f = likelihood(u,a,b)
        for Y in dTeste:
            post_0,post_1= [0,0],[0,0]
            # substitui-se o valor particular de 'y' na função de verossimilhança
            f_Y = likelihood_y(f,Y)
            # obtém-se as posteriores para cada uma das quatro combinações priori/tNorma
            post_0= [posteriori(priori[0],f_Y,'produto'), posteriori(priori[0],f_Y,'minimo')]
            post_1= [posteriori(priori[1],f_Y,'produto'), posteriori(priori[1],f_Y,'minimo')]
            # escreve-se os valores de máximo das posteriores
            print sup_Piecewise(post_0[0]), sup_Piecewise(post_1[0]), \
```

```

sup_Piecewise(post_0[1]), sup_Piecewise(post_1[1])

#####
def likelihood(u,a,b):
# função que isola os intervalos da função de verossimilhança em função de 'x', não de 'y'.
# isto é, ao invés de a<y<b, fica a'<x<b'
f = [[(y-b)/(u-b), u<y, y<b],[(y-a)/(u-a), a<y, y<u]]
g = []
for n in range(len(f)):
    int= [0,0]
    k = 0
    S = solve_ineq([f[n][1], f[n][2]],[x,y])
    # só interessam os intervalos, não igualdades
    if len(S[k])==1:
        k=k+1 # se for igualdade, passamos ao próximo
    # para cada intervalo da desigualdade, temos:
    for i in range(len(S[k])):
        z= list(S[k][i].iterator())
        # como queremos o apenas a expressão 'expr' em função de y de "x<expr",
        # se tivermos z=[x,expr], fazemos int[1]=expr
        if z[0]==x:
            int[1]=z[1]
        # e se tivermos z=[expr,x], fazemos int[0]=expr
        elif z[1]==x:
            int[0]=z[0]
        # guardamos o intervalo já ordenado
        g.append([ int, f[n][0].full_simplify() ])
return g

#####
def likelihood_y(f,Y):
# função que substitui o valor de 'Y' na função de verossimilhança
g=[]
for i in range(len(f)):
    I = f[i][0]
    g.append([(I[0].subs(y=Y),I[1].subs(y=Y)),(f[i][1].subs(y=Y).full_simplify())])
return Piecewise(g)

#####
def posteriori(f1,f2,T_norm):
# função que calcula a distribuição posteriori
# une-se os pontos do intervalo das duas funções para que se tenha
# as partes mínimas possíveis
I = union(h0.end_points(),f.end_points())
h0= h0.extend_by_zero_to( min(I), max(I) )
f = f.extend_by_zero_to( min(I), max(I) )
I = ord_interv(union(h0.end_points(),f.end_points()))
g = []; gi= []
# não é necessário trabalhar com a descontinuidade da posteriori para
# a t-norma do mínimo; o máximo da parte contínua coincide com o da
# função com a descontinuidade
if T_norm =='mínimo':
    for i in range(len(I)-1):
        gi=min_Piecewise(h0,f,(I[i],I[i+1]))
        if type(gi)==list:
            g.extend(gi)
    g= Piecewise(g)
elif T_norm=='produto':
    for i in range(len(I)-1):
        l= (I[i]+I[i+1])/2
        g.append([(I[i],I[i+1]), h0.which_function(l)*f.which_function(l)])
    g= Piecewise(g)
g= oper_PieceF(f=g ,K=1/g(sup_Piecewise(f=g)),oper='mult')

```

```
return fix_interv(g,0)
#####
```

B.3 Funções auxiliares

```
#####
def coefAng_maxMin(D,u):
# Função que informa quais são os coeficientes angulares das funções
# a(x) (coeficiente a_m) e b(x) (coeficiente a_M) se a constante 'C' fosse zero.
# Mais informações sobre a constante 'C' na função 'cria_func'
    a_M,a_m=0,0
    # Para cada par ordenado de 'D':
    for i in range(len(D)):
        Xi= D[i][0]; Yi= D[i][1]
        # Distância entre o ponto experimental (Xi,Yi) e o (Xi,u(Xi))
        Ai= (Yi - u(x=0))/Xi
        # Caso o ponto atual esteja ACIMA do ajuste linear 'u(x)':
        if Yi - u(x=Xi)>0:
            if Ai>a_M:
                a_M= Ai
                k_M= i
        # e caso o ponto esteja ABAIXO:
        else:
            if a_m==0:
                a_m= Ai
            elif Ai<a_m:
                a_m= Ai
                k_m= i
    # print k_M,k_m
    return a_M,a_m

#####
def cria_func(D):
# Função que recebe um vetor de coordenadas X/Y (medidas/simulações)
# (vetor 'D'), e retorna as funções 'u(x)', 'a(x)' e 'b(x)'.
    # Realiza o ajuste linear dos dados
    A, B, r, tmp= Stat.linregress(D)
    u= A*x + B
    # calcula a constante 'C', que define
    # o quanto de folga será oferecida aos dados
    C = (1-r^2)+.10
    a_M,a_m= coefAng_maxMin(D=D,u=u)
    # Define as funções 'a(x)' e 'b(x)'
    a = u(0) + (a_m - abs(a_m-A))*C*x
    b = u(0) + (a_M + abs(a_M-A))*C*x
    return u,a,b

#####
def parteDados(dados,n_partes):
    m=len(dados)/n_partes; d=[]
    for i in range(n_partes):
        d.append(dados[m*i:m*(i+1)])
    return d

#####
def parteDadosEmDias(dados):
    return parteDados(dados,n_partes=3)

#####
def separaDados(dados, nTeste):
# Esta função recebe os dados de três dias de medidas/simulações de uma
# mesma configuração de plantio, e retorna estes dados separados em dados
# de treino e dados de teste. 'nTeste'=0,1,2 é o número que define qual será o
# dia de treino. Por exemplo, nTeste=1 informa que o dia de testes será o segundo dia,
```

```

# logo o primeiro e terceiro dias serão utilizados para treinar o algoritmo (=dTreino)
dTreino= []
# Para o dia de testes, só será necessário o vetor de simulações,
# o de medidas experimentais não é utilizado
dTeste = [l[1] for l in dados[nTeste]]
P      = range(len(dados)); P.pop(nTeste)
# Constrói a matriz com os dados que serão usados para treino
# do algoritmo
for i in P:
    for elem in dados[i]:
        dTreino.append(elem)
return dTeste, dTreino

#####
def trapez(a,b,c,d):
    return Piecewise([ [(a,b),(x-a)/(b-a)], [(b,c),1], [(c,d),(d-x)/(d-c)] ])

#####
def triang(a,b,c):
    return Piecewise([ [(a,b),(x-a)/(b-a)], [(b,c),(c-x)/(c-b)] ])

#####
def parabola(a,b):
    return parabolaG(a,b,1)

#####
def min_Piecewise(f1,f2,I):
# Função que aplica o operador de mínimo simbólico entre duas funções f1
# e f2 num determinado intervalo I, e retorna as funções resultantes da
# operação na forma de lista usual
    h = []; S= []
    # 'S' será a partição utilizada pela função definida por partes no caso
    # não-trivial de termos intersecções entre as duas funções no intervalo 'I' dado.
    l = (I[0] + I[1])/2
    f1= f1.which_function(l)
    f2= f2.which_function(l)
    # Trata casos patológicos...
    if (f1==1 and f2==0) or (f1==0 and f2==1):
        return [[(I[0],I[1]), 0]]
    else:
        d= f2-f1
    # Procura-se pelas intersecções entre as funções
    # (quem é máximo no próximo trecho da partição vira mínimo, e vice-versa)
    S_= solve(d,x)
    if len(S_)>0:
        for i in range(len(S_)):
            S_[i]=list(S_[i].iterator())[1]
            # considera-se apenas as soluções dentro do domínio aberto
            for i in range(len(S_)):
                z = S_[i]
                if z>I[0] and z<I[1]:
                    S.append(z)

    # caso não tenha intersecção entre f1 e f2, é fácil
    if S==[] or len(S)==0:
        if d(x=l)>=0:
            h.append([(I[0],I[1]), f1])
        else:
            h.append([(I[0],I[1]), f2])
    # agora o caso mais complicado
    else:

```

```

# incluímos os extremos
S.insert(0,I[0])
S.append(I[1])
for i in range(len(S)-1):
    a= float(S[i]); b= float(S[i+1])
    l= (a+b)/2
    if d(x=l)>=0:
        h.append([(a,b), f1])
    else:
        h.append([(a,b), f2])
return h

#####
def fix_interv(f,k):
# caso tenha-se duas partes de uma função por partes com a mesma função,
# esta função une as duas partes em uma só
F= f.list()
if k==len(F)-1:
    return Piecewise(F)
else:
    f1= F[k][1]; i1= F[k][0]
    f2= F[k+1][1]; i2= F[k+1][0]
    # caso a função seja repetida...
    if f1==f2:
        #...apaga-se uma delas
        F.pop(k)
        #...e emenda-se os intervalos!
        F[k]=[(i1[0],i2[1]),f1]
        return fix_interv(Piecewise(F),k)
    else:
        return fix_interv(Piecewise(F),k+1)

#####
def sup_Piecewise(f):
# função que encontra o valor x* tal que a função em x*
# (em todas as suas partes) tenha valor máximo
M= None
L= f.list()
for i in range(len(L)):
    if L[i][1]!=0:
        m=L[i][1].find_maximum_on_interval(L[i][0][0], L[i][0][1])[1]
        if M==None:
            M=m
        else:
            if f(m)>f(M):
                M=m
return M

#####
def oper_PieceF(f, K, oper):
# função que realiza operações algébricas básicas: soma um determinado
# número ou multiplica por um escalar toda
L=f.list()
for i in range(len(L)):
    if oper=='mult':
        L[i][1]=L[i][1]*K
    if oper=='soma':
        L[i][1]=L[i][1]+K
return Piecewise(L)

#####

```

```

def ord_interv(I):
# função que recebe um intervalo e o ordena
# (função importante para trabalhar com funções por partes)
k= len(I)
for i in range(k-1):
    for j in range(i+1,k):
        if i>=k or j>=k:
            break
        if I[i]==I[j] and i!=j:
            I.pop(j)
            k=k-1
        elif I[i]>I[j]:
            t = I[i]
            I[i]= I[j]
            I[j]= t
return I

#####
def graf_lin(D,u,a,b):
# Função que gera um gráfico com os pontos experimentais do vetor 'D',
# juntamente com o ajuste linear u(x) e as retas-limite a(x) e b(x).
P,F=Graphics(),Graphics()
P= plot(point(D, rgbcolor='black', size=5))
d= P.get_axes_range(); x_m, x_M = d['xmin'], d['xmax']
F= line([(x_m,u(x=x_m)),(x_M,u(x=x_M))], rgbcolor='red')
if a!=None:
    F= line([(x_m,a(x=x_m)),(x_M,a(x=x_M))], rgbcolor='gray', linestyle='--') + F
if b!=None:
    F= line([(x_m,b(x=x_m)),(x_M,b(x=x_M))], rgbcolor='gray', linestyle='--') + F

return P+F
#####

```

B.4 Dados do programa

```

def dados(param,D,P_x=None):
    retornaPriori=1
    if P_x==None:
        retornaPriori=0
    if param[0:4]=='temp':
        if param[5:7]=='PS':
            if param=='temp_PS_1':
                D = [[292.75,292.74], [291.85,291.99], [292.45,291.44], [292.05,291.09], [291.25,290.84],
                    [290.65,290.67], [290.25,290.54], [292.05,290.59], [295.05,291.33], [297.25,292.46],
                    [299.85,294.81], [301.55,296.86], [303.35,298.70], [303.05,300.24], [304.05,301.19],
                    [304.05,301.29], [303.35,300.64], [301.85,299.43], [296.85,297.94], [294.05,296.98],
                    [293.55,296.23], [293.25,295.61], [293.75,295.09], [291.65,294.65], [290.75,294.25],
                    [291.25,293.90], [290.15,293.59], [289.05,293.34], [289.35,293.13], [288.85,292.97],
                    [285.95,292.84], [290.15,293.05], [293.95,294.45], [297.55,296.12], [299.45,298.83],
                    [301.35,300.53], [302.45,302.01], [302.95,302.93], [303.15,303.07], [303.95,302.48],
                    [304.35,301.62], [302.85,300.30], [300.35,298.68], [298.15,297.58], [295.75,296.72],
                    [294.55,296.02], [294.85,295.43], [295.05,294.94], [292.55,294.51], [292.95,294.15],
                    [292.55,293.84], [292.55,293.60], [292.05,293.39], [291.95,293.23], [291.35,293.09],
                    [292.75,293.30], [295.65,294.62], [298.25,296.17], [300.95,298.77], [303.35,300.42],
                    [305.05,301.82], [305.35,302.69], [306.15,302.84], [306.25,302.29], [305.45,301.47],
                    [302.85,300.20], [300.85,298.65], [299.95,297.59], [296.35,296.76], [294.95,296.07],
                    [294.55,295.51], [294.45,295.02]]
                P_x = [trapez(283,291,299,308), parabola(284,307)]
            elif param=='temp_PS_2':
                D = [[285.95,285.64], [286.35,285.40], [286.45,285.18], [285.55,284.98], [286.05,284.81],
                    [286.55,284.65], [286.95,284.51], [286.85,284.39], [288.85,284.94], [291.25,286.43],
                    [293.35,288.11], [295.55,290.61], [297.55,292.89], [298.35,294.57], [299.55,295.22],
                    [299.45,294.76], [299.55,293.59], [299.65,291.78], [295.85,290.26], [291.75,289.24],

```

```

[291.85,288.43],[290.35,287.75],[290.75,287.15],[289.55,286.61],[288.35,286.13],
[288.25,285.70],[288.35,285.33],[287.35,285.02],[287.35,284.75],[287.35,284.53],
[287.25,284.34],[286.95,284.19],[289.05,284.96],[291.75,287.08],[294.05,289.21],
[296.25,291.87],[297.35,294.51],[297.85,295.97],[298.25,296.24],[299.25,295.54],
[299.95,294.18],[297.35,292.29],[295.05,290.72],[292.15,289.67],[291.25,288.84],
[289.95,288.14],[288.65,287.54],[288.45,287.02],[288.05,286.55],[288.25,286.14],
[286.85,285.79],[286.95,285.50],[286.45,285.24],[285.15,285.03],[285.75,284.85],
[285.35,284.70],[288.65,285.36],[291.25,287.19],[293.85,289.02],[295.75,291.46],
[296.55,293.95],[297.25,295.30],[297.55,295.56],[297.15,294.94],[296.85,293.72],
[295.75,292.02],[292.45,290.60],[289.95,289.64],[289.35,288.86],[287.75,288.21],
[287.25,287.65],[286.85,287.15]]
P_x =[trapez(284,285,290,302), parabola(282,303)]
elif param[5:7]=='S1':
    if param =='temp_S1_1':
        D =[[292.25,292.72],[291.45,291.97],[291.45,291.42],[291.65,291.06],[290.55,290.81],
[290.15,290.64],[289.95,290.51],[291.75,290.56],[295.05,291.31],[297.35,292.44],
[299.85,294.76],[301.15,296.75],[303.75,298.67],[304.25,300.30],[305.45,301.30],
[305.45,301.38],[304.45,300.71],[302.95,299.48],[297.25,297.99],[292.25,297.01],
[291.65,296.25],[291.35,295.63],[291.35,295.10],[290.75,294.62],[290.85,294.17],
[289.95,293.80],[288.85,293.51],[287.85,293.29],[287.55,293.12],[286.85,292.97],
[285.05,292.85],[289.55,293.06],[293.95,294.44],[297.75,296.01],[299.55,298.63],
[301.25,300.44],[302.55,301.97],[302.75,302.92],[303.15,303.10],[303.75,302.53],
[303.95,301.66],[302.55,300.35],[299.25,298.72],[296.05,297.61],[295.05,296.73],
[294.55,296.01],[293.85,295.42],[293.85,294.90],[292.05,294.45],[292.45,294.08],
[292.45,293.78],[291.95,293.55],[291.65,293.35],[291.45,293.19],[290.85,293.06],
[292.45,293.28],[295.75,294.61],[298.25,296.09],[300.65,298.63],[303.35,300.38],
[305.15,301.86],[305.45,302.78],[306.45,302.96],[306.55,302.42],[305.75,301.59],
[302.75,300.31],[299.35,298.72],[297.55,297.64],[295.65,296.78],[294.75,296.07],
[294.25,295.49],[293.95,294.98]]
P_x =[trapez(284,290,296,307), parabola(284,307)]
elif param=='temp_S1_2':
    D =[[285.75,285.62],[286.05,285.38],[286.55,285.15],[285.85,284.95],[285.95,284.78],
[286.35,284.62],[286.55,284.48],[286.45,284.36],[288.45,284.78],[291.15,285.77],
[293.15,286.88],[295.45,288.50],[297.05,289.83],[298.35,290.80],[299.45,291.24],
[300.85,291.02],[299.55,290.37],[298.95,289.22],[292.55,288.25],[290.85,287.61],
[290.15,287.08],[289.05,286.62],[288.55,286.21],[288.65,285.85],[288.05,285.52],
[287.95,285.21],[287.75,284.92],[287.45,284.65],[287.35,284.42],[287.25,284.22],
[286.75,284.05],[286.35,283.92],[288.65,284.52],[291.45,285.91],[293.55,287.33],
[295.75,289.22],[296.85,290.74],[297.85,291.67],[298.65,291.94],[299.85,291.61],
[299.45,290.84],[296.35,289.62],[292.95,288.63],[290.85,287.96],[289.05,287.41],
[288.55,286.93],[287.55,286.50],[288.35,286.12],[287.75,285.77],[287.95,285.45],
[286.75,285.14],[286.45,284.88],[286.15,284.66],[285.25,284.49],[285.55,284.36],
[285.45,284.25],[288.35,284.81],[290.75,286.03],[293.65,287.27],[295.45,288.99],
[296.55,290.47],[296.55,291.39],[296.85,291.66],[296.55,291.36],[295.95,290.64],
[294.05,289.50],[290.25,288.58],[288.35,287.96],[287.55,287.43],[286.75,286.98],
[286.95,286.57],[286.25,286.21]]
P_x =[trapez(283,285,290,303), parabola(283,303)]
elif param[5:7]=='S2':
    if param =='temp_S2_1':
        D =[[292.15,292.73],[290.45,291.97],[291.35,291.42],[291.45,291.06],[290.75,290.82],
[289.95,290.65],[289.95,290.52],[291.05,290.57],[293.75,291.31],[297.45,292.44],
[300.35,294.74],[302.15,296.72],[303.75,298.62],[302.65,300.25],[303.95,301.24],
[304.05,301.32],[303.45,300.64],[301.95,299.43],[296.95,297.95],[294.35,296.98],
[293.45,296.22],[291.95,295.61],[293.05,295.08],[291.25,294.58],[289.25,294.14],
[289.95,293.79],[289.15,293.53],[287.95,293.32],[287.75,293.14],[287.05,292.99],
[285.55,292.87],[288.55,293.08],[292.65,294.44],[297.95,295.99],[299.95,298.58],
[301.45,300.37],[302.75,301.91],[302.85,302.86],[303.45,303.04],[303.75,302.48],
[304.25,301.62],[302.55,300.32],[300.05,298.70],[297.45,297.58],[294.55,296.71],
[292.65,296.00],[293.45,295.40],[294.05,294.88],[291.95,294.44],[291.25,294.08],
[291.55,293.79],[291.85,293.56],[290.65,293.37],[291.25,293.20],[291.05,293.07],
[291.95,293.29],[294.95,294.61],[298.45,296.08],[300.95,298.59],[303.55,300.35],
[305.05,301.82],[305.35,302.74],[306.05,302.93],[305.95,302.39],[305.45,301.56],
[302.55,300.29],[300.35,298.70],[298.55,297.62],[295.25,296.76],[293.85,296.06],
[293.25,295.47],[293.75,294.97]]
P_x =[trapez(282,290,295,307), parabola(278,304)]

```

```

elif param=='temp_S2_2':
    D = [[283.75,285.62],[284.95,285.38],[284.85,285.16],[284.35,284.96],[284.95,284.78],
          [285.75,284.63],[286.25,284.49],[285.95,284.36],[287.75,284.77],[290.45,285.76],
          [293.35,286.86],[295.75,288.47],[297.75,289.80],[298.55,290.76],[298.95,291.19],
          [299.35,290.98],[299.05,290.33],[297.95,289.19],[293.05,288.23],[290.55,287.59],
          [289.95,287.07],[288.75,286.61],[288.15,286.19],[288.15,285.81],[286.95,285.47],
          [287.05,285.16],[286.55,284.87],[286.35,284.60],[286.45,284.38],[286.35,284.21],
          [286.25,284.07],[285.95,283.96],[287.85,284.54],[290.75,285.87],[293.65,287.20],
          [295.65,289.00],[297.35,290.53],[297.55,291.51],[297.85,291.82],[298.65,291.49],
          [298.95,290.74],[295.95,289.54],[291.95,288.56],[290.65,287.90],[288.75,287.35],
          [288.15,286.88],[286.95,286.46],[287.05,286.08],[286.45,285.70],[286.95,285.34],
          [285.55,285.03],[285.85,284.79],[285.05,284.62],[283.95,284.48],[284.85,284.37],
          [284.45,284.28],[286.85,284.81],[289.95,286.01],[293.45,287.20],[296.05,288.88],
          [297.05,290.35],[297.05,291.29],[297.25,291.59],[296.65,291.29],[296.05,290.58],
          [294.15,289.46],[289.35,288.54],[287.65,287.91],[286.55,287.40],[285.55,286.94],
          [286.05,286.54],[285.35,286.17]]
    P_x=[trapez(283,290,295,302), parabola(283,302)]
elif param[5:7]=='S3':
    if param == 'temp_S3_1':
        D = [[289.95,292.52],[289.85,291.82],[289.35,291.30],[288.95,290.95],[290.15,290.72],
              [288.85,290.55],[288.55,290.42],[290.05,290.45],[293.85,291.14],[297.05,292.25],
              [302.05,294.46],[303.65,296.34],[305.55,298.12],[303.35,299.65],[304.85,300.57],
              [305.05,300.64],[304.25,300.00],[301.05,298.86],[297.45,297.50],[295.35,296.65],
              [291.85,295.98],[290.85,295.44],[289.45,294.95],[288.45,294.54],[288.05,294.19],
              [286.45,293.89],[285.75,293.64],[284.35,293.43],[284.05,293.25],[284.25,293.09],
              [283.35,292.96],[285.65,293.10],[290.95,294.27],[297.55,295.71],[301.25,298.16],
              [304.25,299.81],[304.05,301.21],[305.35,302.10],[305.15,302.32],[306.45,301.83],
              [305.95,301.02],[303.05,299.80],[298.85,298.31],[295.25,297.33],[292.55,296.55],
              [291.65,295.90],[291.25,295.36],[290.55,294.89],[290.15,294.50],[289.55,294.18],
              [290.55,293.90],[290.35,293.67],[289.85,293.47],[289.85,293.30],[290.35,293.16],
              [291.45,293.31],[295.05,294.45],[298.55,295.83],[302.55,298.23],[306.05,299.85],
              [307.15,301.20],[307.45,302.06],[307.25,302.26],[308.05,301.78],[306.45,300.99],
              [302.25,299.78],[299.05,298.33],[295.55,297.37],[293.35,296.60],[291.55,295.96],
              [290.45,295.42],[290.25,294.97]]
        P_x=[trapez(283,287,292,310), parabola(283,310)]
    elif param=='temp_S3_2':
        D = [[282.85,285.56],[282.55,285.32],[281.95,285.10],[282.65,284.90],[283.55,284.73],
              [284.05,284.57],[284.65,284.44],[284.75,284.31],[287.95,284.69],[293.15,285.66],
              [295.15,286.70],[296.45,288.18],[298.55,289.43],[299.45,290.33],[299.55,290.74],
              [299.45,290.54],[299.25,289.89],[295.75,288.77],[292.05,287.85],[289.85,287.26],
              [287.55,286.78],[287.75,286.37],[286.55,286.00],[286.55,285.62],[286.75,285.23],
              [286.45,284.92],[285.95,284.70],[285.75,284.53],[285.85,284.39],[285.95,284.28],
              [285.45,284.18],[284.95,284.08],[288.35,284.57],[293.35,285.73],[295.35,286.91],
              [296.95,288.53],[298.15,289.94],[298.75,290.86],[298.15,291.18],[298.35,290.89],
              [298.75,290.17],[295.85,289.04],[291.55,288.13],[289.65,287.54],[287.85,287.05],
              [285.95,286.63],[285.35,286.25],[285.85,285.82],[284.55,285.45],[283.65,285.19],
              [284.65,285.00],[284.55,284.85],[283.55,284.72],[282.65,284.60],[283.45,284.49],
              [283.75,284.38],[287.25,284.82],[292.65,285.90],[295.85,286.99],[296.85,288.53],
              [297.75,289.90],[297.35,290.78],[297.25,291.08],[296.75,290.80],[295.95,290.11],
              [293.35,289.03],[289.05,288.16],[286.95,287.59],[284.95,287.12],[283.75,286.72],
              [283.75,286.30],[284.25,285.89]]
        P_x=[trapez(280,283,288,310), parabola(280,305)]
elif param[0:4]=='humd':
    if param[5:7]=='PS':
        if param == 'humd_PS_1':
            D = [[77,69.01],[81,72.43],[78,74.52],[80,75.46],[81,75.72],[84,75.62],[86,75.35],
                  [83,74.64],[74,72.20],[67,68.40],[62,59.90],[53,51.10],[37,43.29],[36,37.82],
                  [32,34.74],[26,33.80],[27,34.74],[31,36.86],[43,39.76],[50,41.64],[53,43.12],
                  [51,44.37],[51,45.46],[60,46.44],[61,47.35],[57,48.20],[62,48.95],[68,49.57],
                  [68,50.06],[71,50.45],[82,50.75],[72,50.02],[60,45.80],[54,41.29],[51,34.97],
                  [48,31.59],[41,28.93],[42,27.40],[44,27.15],[36,28.05],[34,29.48],[39,31.82],
                  [43,35.00],[46,37.36],[59,39.34],[61,41.04],[59,42.52],[62,43.82],[70,44.97],
                  [71,45.98],[74,46.85],[75,47.58],[78,48.18],[78,48.67],[81,49.08],[78,48.46],
                  [70,44.68],[64,40.66],[57,34.77],[46,31.54],[34,29.08],[33,27.64],[29,27.39],
                  [29,28.28],[29,29.67],[38,31.95],[39,35.02],[40,37.30],[54,39.23],[56,40.88],

```

```

    [58,42.33], [57,43.60]]
    P_x=[trapez(20,35,70,95), parabola(21,90)]
elif param=='humd_PS_2':
    D = [[81,83.40], [80,84.43], [81,85.39], [86,86.24], [86,86.97], [84,87.61], [80,88.17],
          [80,88.68], [75,85.84], [70,78.91], [65,71.57], [58,61.43], [53,53.01], [51,47.16],
          [44,44.79], [41,45.83], [41,49.00], [47,54.57], [59,59.85], [73,63.63], [75,66.82],
          [76,69.67], [79,72.32], [83,74.79], [87,77.09], [87,79.19], [87,81.06], [85,82.68],
          [85,84.08], [83,85.28], [83,86.31], [83,87.16], [78,82.82], [72,72.03], [64,62.80],
          [58,53.04], [53,45.03], [46,41.16], [42,40.48], [41,42.25], [40,45.89], [54,51.59],
          [60,56.90], [64,60.81], [73,64.12], [70,67.03], [70,69.66], [73,72.09], [75,74.31],
          [73,76.30], [79,78.07], [78,79.60], [81,80.93], [83,82.09], [81,83.09], [82,83.92],
          [72,80.31], [66,71.28], [57,63.35], [49,54.28], [42,46.51], [36,42.80], [37,42.12],
          [35,43.76], [38,47.15], [42,52.41], [54,57.27], [57,60.89], [61,64.00], [63,66.72],
          [66,69.18], [68,71.43]]
    P_x=[trapez(32,60,86,90), parabola(32,90)]
elif param[5:7]=='S1':
    if param == 'humd_S1_1':
        D = [[78,69.12], [80,72.58], [80,74.72], [79,75.67], [83,75.94], [83,75.85], [85,75.58],
              [82,74.86], [72,72.41], [67,68.57], [60,60.06], [54,51.33], [37,43.39], [34,37.71],
              [31,34.54], [26,33.62], [26,34.59], [29,36.74], [44,39.66], [57,41.57], [60,43.08],
              [57,44.33], [58,45.44], [62,46.53], [59,47.62], [60,48.55], [65,49.25], [72,49.77],
              [73,50.18], [78,50.51], [85,50.79], [74,50.03], [61,45.87], [53,41.61], [51,35.43],
              [47,31.78], [39,29.01], [43,27.42], [44,27.12], [37,27.99], [32,29.41], [37,31.74],
              [47,34.93], [51,37.32], [58,39.33], [59,41.06], [62,42.57], [64,43.93], [70,45.16],
              [70,46.20], [73,47.04], [75,47.73], [77,48.30], [78,48.79], [81,49.20], [78,48.53],
              [70,44.70], [61,40.85], [56,35.08], [43,31.62], [33,29.01], [34,27.50], [29,27.21],
              [30,28.07], [28,29.46], [36,31.74], [45,34.86], [44,37.21], [54,39.19], [54,40.89],
              [58,42.38], [56,43.71]]
        P_x=[trapez(22,45,65,88), parabola(20,88)]
    elif param=='humd_S1_2':
        D = [[79,83.53], [79,84.58], [78,85.55], [83,86.42], [84,87.15], [83,87.79], [80,88.35],
              [79,88.86], [74,86.65], [69,81.89], [63,76.81], [59,69.73], [54,64.16], [51,60.07],
              [45,58.04], [41,58.49], [44,60.60], [46,64.87], [62,68.70], [74,71.35], [75,73.61],
              [75,75.62], [77,77.46], [83,79.18], [84,80.79], [85,82.32], [84,83.79], [83,85.17],
              [82,86.43], [81,87.51], [82,88.39], [82,89.09], [77,85.58], [72,78.09], [65,71.16],
              [58,62.93], [52,57.11], [46,53.83], [40,52.88], [41,53.98], [43,56.63], [51,61.15],
              [59,65.13], [65,67.97], [72,70.42], [71,72.62], [71,74.64], [71,76.52], [72,78.27],
              [71,79.94], [77,81.54], [77,82.98], [78,84.18], [81,85.13], [78,85.88], [79,86.48],
              [71,83.36], [66,76.88], [57,70.95], [47,63.48], [41,57.80], [40,54.54], [52,61.60],
              [35,54.65], [37,57.17], [42,61.44], [53,65.15], [57,67.83], [59,70.16], [63,72.25],
              [64,74.18], [65,75.97]]
        P_x=[trapez(22,30,60,86), parabola(20,88)]
elif param[5:7]=='S2':
    if param == 'humd_S2_1':
        D = [[78,69.08], [85,72.53], [82,74.65], [81,75.61], [82,75.87], [85,75.78], [86,75.52],
              [85,74.79], [77,72.36], [64,68.54], [58,60.08], [50,51.40], [35,43.50], [38,37.82],
              [33,34.65], [28,33.74], [28,34.72], [31,36.85], [41,39.75], [47,41.64], [51,43.14],
              [54,44.38], [51,45.50], [62,46.66], [66,47.74], [59,48.58], [63,49.21], [70,49.70],
              [72,50.11], [76,50.45], [82,50.74], [77,49.98], [63,45.89], [51,41.67], [49,35.54],
              [46,31.90], [39,29.13], [42,27.53], [42,27.21], [38,28.07], [32,29.48], [37,31.79],
              [42,34.98], [47,37.37], [61,39.38], [69,41.10], [62,42.61], [61,43.97], [70,45.18],
              [76,46.19], [78,47.01], [76,47.69], [82,48.26], [80,48.76], [80,49.17], [80,48.50],
              [71,44.72], [59,40.89], [54,35.15], [44,31.69], [37,29.08], [35,27.56], [30,27.26],
              [29,28.12], [28,29.51], [36,31.78], [39,34.91], [42,37.25], [56,39.23], [59,40.94],
              [61,42.43], [58,43.76]]
        P_x=[trapez(25,30,65,87), parabola(24,90)]
    elif param=='humd_S2_2':
        D = [[89,83.50], [84,84.55], [86,85.52], [89,86.38], [89,87.12], [86,87.75], [82,88.31],
              [82,88.82], [77,86.68], [71,81.94], [63,76.88], [57,69.85], [51,64.29], [49,60.23],
              [42,58.22], [42,58.66], [40,60.74], [46,64.98], [59,68.79], [75,71.41], [76,73.65],
              [78,75.69], [79,77.59], [84,79.37], [89,81.03], [89,82.60], [89,84.11], [89,85.48],
              [87,86.65], [86,87.59], [85,88.33], [84,88.89], [81,85.49], [75,78.29], [63,71.77],
              [56,63.87], [50,57.88], [45,54.38], [42,53.31], [42,54.39], [40,57.02], [52,61.49],
              [67,65.43], [66,68.25], [74,70.67], [75,72.84], [75,74.84], [77,76.74], [79,78.63],
              [76,80.53], [82,82.18], [80,83.46], [84,84.43], [87,85.20], [83,85.81], [84,86.34],

```

```

        [78,83.34], [70,77.03], [56,71.29], [46,63.96], [38,58.24], [36,54.88], [37,53.86],
        [36,54.89], [38,57.38], [43,61.62], [57,65.33], [61,68.01], [65,70.33], [69,72.41],
        [70,74.33], [71,76.16]]
    P_x = [trapez(37,60,80,97), parabola(35,97)]
elif param[5:7]=='S3':
    if param == 'humd_S3_1':
        D = [[86,70.60], [87,73.76], [88,75.70], [89,76.55], [85,76.77], [88,76.65], [91,76.37],
            [91,75.64], [78,73.25], [67,69.51], [55,61.39], [48,53.03], [33,45.24], [36,39.48],
            [33,36.23], [28,35.27], [27,36.24], [31,38.34], [39,41.07], [45,42.75], [58,44.08],
            [59,45.17], [63,46.22], [71,47.13], [73,47.90], [75,48.56], [79,49.13], [85,49.61],
            [88,50.02], [88,50.37], [91,50.67], [90,50.13], [77,46.59], [54,42.58], [47,36.59],
            [41,33.10], [39,30.44], [39,28.84], [40,28.44], [34,29.22], [31,30.61], [37,32.87],
            [46,35.85], [59,38.00], [69,39.82], [73,41.39], [75,42.78], [78,44.01], [80,45.06],
            [82,45.97], [81,46.75], [80,47.41], [84,47.99], [85,48.49], [83,48.93], [82,48.49],
            [72,45.22], [61,41.57], [51,36.00], [40,32.70], [32,30.21], [34,28.73], [30,28.39],
            [29,29.19], [30,30.58], [41,32.80], [45,35.74], [59,37.85], [64,39.65], [70,41.21],
            [75,42.58], [73,43.80]]
        P_x = [trapez(25,60,75,92), parabola(25,92)]
    elif param == 'humd_S3_2':
        D = [[90,83.90], [90,84.94], [93,85.90], [92,86.75], [94,87.49], [92,88.12], [89,88.68],
            [88,89.18], [80,87.17], [63,82.49], [58,77.62], [55,71.06], [50,65.79], [48,61.96],
            [44,60.01], [43,60.46], [44,62.63], [52,66.90], [61,70.63], [77,73.10], [85,75.16],
            [82,77.00], [85,78.70], [90,80.53], [89,82.48], [89,84.06], [90,85.21], [90,86.06],
            [88,86.74], [86,87.32], [87,87.83], [88,88.30], [82,85.50], [66,79.21], [60,73.29],
            [55,65.94], [49,60.24], [46,56.78], [44,55.61], [43,56.61], [39,59.20], [54,63.55],
            [65,67.32], [72,69.93], [77,72.16], [82,74.12], [82,75.95], [81,78.10], [85,80.02],
            [87,81.39], [86,82.38], [84,83.19], [86,83.90], [91,84.56], [87,85.18], [86,85.77],
            [79,83.31], [55,77.65], [52,72.31], [47,65.47], [41,60.01], [37,56.76], [38,55.69],
            [39,56.69], [39,59.20], [45,63.37], [58,66.97], [65,69.46], [74,71.61], [78,73.50],
            [77,75.52], [74,77.60]]
        P_x = [trapez(37,60,80,97), parabola(35,97)]
if retornaPriori:
    return (D,P_x)
else:
    return D

```

Referências Bibliográficas

- [1] S. Lapointe, B. Bobée, “Revision of possibility distributions: A Bayesian inference pattern”, *Fuzzy Sets and Systems*, (2000) 116(2):119–140.
- [2] L. A. Zadeh, “Fuzzy sets as a basis for a theory of possibility” *Fuzzy Sets and Systems*, (1978) 1(1):3–28
- [3] H. T. Nguyen, “On conditional possibility distributions” *Fuzzy Sets and Systems*, (1978) 1(4):299-307
- [4] E. Hisdal, “Conditional possibilities independence and noninteraction” *Fuzzy Sets and Systems*, (1978) 1(4):283-297
- [5] B. Bouchon, “Fuzzy inferences and conditional possibility distributions”, *Fuzzy Sets and Systems*, (1987) 23:33–41.
- [6] L. C. Barros; R. C. Bassanezi, *Tópicos de lógica fuzzy e biomatemática - 2ª edição*, Coleção IMECC – Textos Didáticos, Volume 5, 2010.
- [7] F. Gomide, W. Pedrycz, *Fuzzy Systems Engineering: Toward Human-Centric Computing*, John Wiley & Sons, 2007.
- [8] W. Pedrycz, “Fuzzy relational equations with triangular norms and their resolution”, *Busefal*, (1982) 11:24-32
- [9] D. Dubois, S. Moral, H. Prade, “A semantics for possibility theory based on likelihoods”, *J. Math. Anal. Appl.*, (1997) 205:359–380.
- [10] D. Dubois, H. Prade, “Bayesian conditioning in possibility theory”, *Fuzzy Sets and Systems*, (1997) 92:223–240.
- [11] D. Dubois, H. Prade, “Possibility Theory: An Approach to Computerized Processing of Uncertainty”, Plenum, New York, 1988.
- [12] G. J. Klir, *Uncertainty and Information: Foundations of Generalized Information Theory - 1ª edição*, Wiley-IEEE Press, 2005

- [13] G. J. Klir, M.J. Wierman, *Uncertainty-Based Information: Elements of Generalized Information Theory* - Studies in Fuzziness and Soft Computing - 2ª edição, Physica-Verlag HD, 1999.
- [14] D. C. Montgomery, *Forecasting and time series analysis* New York; St. Louis: McGraw-Hill, 1976.
- [15] S. Ross, *A First Course in Probability* - 5ª edição Prentice Hall Inc, 1998.
- [16] C. J. Willmott, “Statistics for the Evaluation and Comparison of Models” *Journal of Geophysical Research*, (1985) 90(I5):8995–9005
- [17] P. L. O. Costa, *Estatística*, Edgard Blücher Ltda., 1977.
- [18] J. S. Armstrong, F. Collopy, “Error measures for generalizing about forecasting methods: empirical comparisons” *International Journal of Forecasting*, (1992) 8:69–80
- [19] F. Gomide, “Estimação de Parâmetros: Máxima Verosimilhança e Bayes – Notas de aula” *Tópicos em Aprendizagem de Máquina e Classificação de Padrões*, http://www.dca.fee.unicamp.br/~gomide/courses/CT720/transp/CT720EstimacaoParametros_3.pdf
- [20] P. P. Coltri, “Simulação do microclima de cafeeiros a pleno sol e arborizados com o software ENVI-MET”
Tese de doutorado - Faculdade de Engenharia Agrícola, UNICAMP, 2012.
Processo: 565.387/2008-1 – Edital: MCT/CNPq n°27/2007