



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2014/09.02.20.20-TDI

CALIBRAÇÃO MULTIOBJETIVO DE MODELOS HIDROLÓGICO E DE SUPERFÍCIE ATMOSFÉRICO

Amarísio da Silva Araújo

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Haroldo Fraga de Campos Velho, aprovada em 28 de agosto de 2014.

URL do documento original:

<<http://urlib.net/8JMKD3MGP5W34M/3GUJQG5>>

INPE
São José dos Campos
2014

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE (RE/DIR-204):**Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Membros:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Maria Tereza Smith de Brito - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Maria Tereza Smith de Brito - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2014/09.02.20.20-TDI

CALIBRAÇÃO MULTIOBJETIVO DE MODELOS HIDROLÓGICO E DE SUPERFÍCIE ATMOSFÉRICO

Amarísio da Silva Araújo

Tese de Doutorado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Haroldo Fraga de Campos Velho, aprovada em 28 de agosto de 2014.

URL do documento original:

<<http://urlib.net/8JMKD3MGP5W34M/3GUJQG5>>

INPE
São José dos Campos
2014

Dados Internacionais de Catalogação na Publicação (CIP)

Araújo, Amarísio da Silva.

Ar15c Calibração multiobjetivo de modelos hidrológico e de superfície atmosférico / Amarísio da Silva Araújo. – São José dos Campos : INPE, 2014.

xxvi + 92 p. ; (sid.inpe.br/mtc-m21b/2014/09.02.20.20-TDI)

Tese (Doutorado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2014.

Orientador : Dr. Haroldo Fraga de Campos Velho .

1. Modelos de fluidos geofísicos. 2. Problemas inversos. 3. Calibração de parâmetros. 4. Otimização multiobjetivo. 5. Algoritmos genéticos. I.Título.

CDU 681.5:556



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de **Doutor(a)** em
Computação Aplicada

Dr. Ezzat Selim Chalhoub



Presidente / INPE / SJCampos - SP

Dr. Haroldo Fraga de Campos Velho



Orientador(a) / INPE / São José dos Campos - SP

Dr. Stephan Stephany



Membro da Banca / INPE / SJCampos - SP

Dr. Marcos Heil Costa



Convidado(a) / UFV / Viçosa - MG

Dr. Wagner Figueiredo Sacco



Convidado(a) / UFOPA / Santarém - PA

Este trabalho foi aprovado por:

() maioria simples

() unanimidade

Aluno (a): **Amarísio da Silva Araújo**

São José dos Campos, 28 de Agosto de 2014

“ O que é escrito, ordenado, factual nunca é suficiente para abarcar toda a verdade: a vida sempre transborda de qualquer cálice.”

BORIS PASTERNAK

A meus pais, Geraldo e Aita, e aos meus irmãos.

AGRADECIMENTOS

Agradeço, primeiramente, e principalmente, a meus pais, irmãos e familiares pelo amor incondicional e apoio constante na minha caminhada diária.

Agradeço às secretárias Margarida e Suely, da Pró-Reitoria de Pesquisa e Pós-Graduação da Universidade Federal de Viçosa (UFV), e aos meus colegas do Departamento de Matemática da UFV pelo suporte durante o meu período de afastamento para a realização do meu doutorado.

Agradeço ao meu orientador, Professor Haroldo, por seu competente, paciente e zeloso trabalho de orientação e importante contribuição na minha formação através das disciplinas por ele ministradas e dos constantes e profícuos diálogos científicos, culturais e políticos.

Agradeço a todos os professores e pesquisadores da CAP/INPE com os quais muito aprendi através das disciplinas do curso, dos seminários, das trocas de informações e da convivência fraterna.

Agradeço aos professores e pesquisadores Juan Martin Bravo e Walter Collischonn do Instituto de Pesquisas Hidráulicas da UFRGS que me receberam naquele instituto para uma aprendizagem sobre o modelo hidrológico IPH-II e me deram todo o suporte necessário para os testes de calibração do mesmo, disponibilizando dados hidrológicos e o código do programa computacional do modelo.

Agradeço ao professor e pesquisador Marcos Heil, do Departamento de Engenharia Agrícola da UFV, e à sua orientanda de Mestrado Carla Camargo pela orientação e treinamento no software de calibração do modelo de superfície atmosférico IBIS, com todo o suporte necessário para os testes de calibração do IBIS e a disponibilização dos códigos e dados.

Agradeço à CAPES pelo auxílio financeiro através do Programa Pró-Doutoral.

Agradeço a todos os funcionários do INPE, em especial aos das secretarias do LAC e da CAP e aos da Biblioteca do INPE.

Deixo, aqui, um agradecimento especial ao meu grande amigo e colega de curso Vitor Conrado pelo seu precioso auxílio na discussão de ideias e na implementação do algoritmo genético AGE e o constante suporte computacional nos testes de calibração do IPH-II e do IBIS.

Por fim, agradeço a todos os meus colegas e amigos de pós-graduação do INPE, em especial aos amigos: Rodolfo Lotte, Anna Karina, Luiz Gustavo, Luciane Sato, Fernando Oliveira, Marcos Paulo, Fábio Dall Cortivo, Eduardo Luz, Luis França, Marlon da Silva, Rudinei Martins e Sóstenes Gomes.

RESUMO

Nos últimos anos, a ciência tem se beneficiado muito de modelos computacionais como ferramentas importantes e cada vez mais confiáveis em muitos tipos de investigações. Tais modelos estabelecem uma representação de sistemas físicos através de equações matemáticas que são resolvidas de forma aproximada por meio de técnicas numéricas codificadas em programas computacionais. Fenômenos de dinâmica de fluidos e nas áreas de meteorologia, de transferência de calor em estruturas de engenharia, por exemplo, são modelados com equações matemáticas, muitas vezes sob a forma de sistemas complexos de equações diferenciais. Para se garantir a aderência de um modelo à realidade do fenômeno modelado, torna-se necessário o ajuste de vários parâmetros do mesmo, procedimento este conhecido como *calibração de parâmetros*. No caso de modelos de fluidos geofísicos, por exemplo, em que há um número finito de parâmetros a serem ajustados, a calibração corresponde a um tipo de problema inverso - *problema inverso discreto* ou *estimação de parâmetros* - no qual, partindo-se dos efeitos (dados observados), tenta-se chegar às causas (parâmetros). Para se resolver este tipo de problema inverso, uma estratégia é a otimização de uma ou mais funções (objetivos) para avaliar o erro entre as variáveis de entrada e de saída do modelo. Este trabalho trata da calibração multiobjetivo dos parâmetros de dois modelos de fluidos geofísicos: o modelo hidrológico de pequenas bacias IPH-II e o modelo de superfície atmosférico IBIS. Para a calibração do IPH-II, foi implementado um algoritmo genético com epidemia. A calibração do IBIS é feita com um software de calibração multiobjetivo do IBIS, denominado Optis, que usa como técnica de otimização multiobjetivo o algoritmo NSGA-II, sendo, aqui, propostas algumas mudanças nas definições das funções objetivo, a implementação de um operador de epidemia no NSGA-II e o uso de uma estratégia de calibração baseada na escala temporal das variáveis.

MULTI-OBJECTIVE CALIBRATION OF HYDROLOGICAL AND SURFACE ATMOSPHERIC MODELS

ABSTRACT

In recent years, Science has greatly benefited from computational models as important and increasingly reliable tools for many types of investigations. These models establish a representation of physical systems through mathematical equations that are solved in an approximate way by means of numerical techniques encoded in computer programs. Phenomena of geophysical fluid dynamics, meteorological phenomena and heat transfer in engineering structures phenomena, for example, are modeled with mathematical equations, often in the form of complex systems of differential equations. To ensure the adherence of a model to the real modeled phenomenon, it is necessary to adjust several parameters of the model, a procedure known as *calibration of parameters*. In the case of geophysical fluid models, for example, where there is a finite number of parameters to be adjusted, the calibration corresponds to an inverse problem type - *discrete inverse problem* or *parameters estimation* - in which, starting from the effects (observed data), we try to determine the causes (parameters). To address this type of inverse problem, a possible strategy is to optimize one or more function (objective function) that evaluate the error between the input variables and the output variables of the model. This paper deals with the multi-objective calibration of the parameters of two models of geophysical fluids: the hydrological model of small watersheds IPH-II and the atmospheric surface model IBIS. For the parameters calibration of the IPH-II model, a genetic algorithm was implemented with an epidemic operator. The IBIS parameters calibration is done with a software for multi-objective calibration of the IBIS, called Optis, which uses as a technique for multi-objective optimization the NSGA-II algorithm. Here we are proposing some changes in the definitions of the objective functions, the implementation of an epidemic operator in NSGA-II and the use of a calibration strategy based on time scale of the observed data.

Keywords: geophysical fluid models. inverse problems. calibration of parameters. multi-objective optimization. genetic algorithms.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 (a) Relação entre Evapotranspiração potencial e umidade do solo; (b) Reservatório de perdas por interceptação e depressões do solo.	8
2.2 Esquema modular do IBIS 2.6.	12
3.1 Dois exemplos de trajetórias no espaço de parâmetros para (a) $k = 2$ e (b) $k = 3$	25
3.2 Mapeamento do espaço de decisão no espaço objetivo.	27
3.3 Dominância em uma minimização de duas funções.	29
3.4 Minimização de duas funções exemplos g e h	30
3.5 Espaço objetivo para as funções exemplo g e h	31
3.6 Ordenamento de Pareto no Espaço Objetivo.	36
3.7 Passos de reflexão e contração em um <i>Complexo</i> no MOCOM.	40
3.8 Cálculo das distâncias de aglomeração nas soluções intermediárias.	41
3.9 Esquema do Algoritmo NSGA-II.	42
4.1 Espaço objetivo resultante da calibração multiobjetivo do IPH-II, com $ns = 200$	48
4.2 Espaço objetivo resultante da calibração multiobjetivo do IPH-II, com $ns = 50$	50
4.3 Conjunto de Pareto resultante de uma calibração multiobjetivo do IPH-II, com o AGE.	51
4.4 Vazões observadas e vazões calculadas com o conjunto de Pareto.	52
5.1 Resultados de PARo em um dia típico - calibração hierárquica.	59
5.2 Resultados de fAPAR com dados mensais durante um ano - calibração hierárquica.	60
5.3 Resultados de Rn em um dia típico - calibração hierárquica.	60
5.4 Resultados de u^* em um dia típico - calibração hierárquica.	61
5.5 Resultados de NEE em um dia típico - calibração hierárquica.	61
5.6 Resultados de HE em um dia típico - calibração hierárquica.	62
5.7 Resultados de LE em um dia típico - calibração hierárquica.	62
5.8 Resultados de LAI com dados mensais durante um ano - calibração hierárquica.	63
5.9 Resultados de PARo em um dia típico - calibração em escala temporal.	68
5.10 Resultados de Rn em um dia típico - calibração em escala temporal.	68
5.11 Resultados de u^* em um dia típico - calibração em escala temporal.	69

5.12	Resultados de NEE em um dia típico - calibração em escala temporal. . .	69
5.13	Resultados de HE em um dia típico - calibração em escala temporal. . .	70
5.14	Resultados de LE em um dia típico - calibração em escala temporal. . . .	70
5.15	Frentes de Pareto resultantes de calibrações do IBIS obtidas usando-se o NSGA-II com e sem epidemia - $(F_1, F_2) = (\Phi_1, \Phi_2)$	71
5.16	Resultados de PARo em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	73
5.17	Resultados de Rn em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	74
5.18	Resultados de u^* em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	74
5.19	Resultados de NEE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	75
5.20	Resultados de HE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	75
5.21	Resultados de LE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.	76
5.22	(a) Resultados de PARo em um dia típico - escala temporal (NSGA- II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de PARo em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II). . .	77
5.23	(a) Resultados de Rn em um dia típico - escala temporal (NSGA-II epi- dêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epi- dêmico); (c) Resultados de Rn em um dia típico - escala temporal (NSGA- II); (d) Erro absoluto entre observado e calculado (NSGA-II).	77
5.24	(a) Resultados de u^* em um dia típico - escala temporal (NSGA-II epi- dêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epi- dêmico); (c) Resultados de u^* em um dia típico - escala temporal (NSGA- II); (d) Erro absoluto entre observado e calculado (NSGA-II).	78
5.25	(a) Resultados de NEE em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epi- dêmico); (c) Resultados de NEE em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II). . .	78
5.26	(a) Resultados de HE em um dia típico - escala temporal (NSGA-II epi- dêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epi- dêmico); (c) Resultados de HE em um dia típico - escala temporal (NSGA- II); (d) Erro absoluto entre observado e calculado (NSGA-II).	79

5.27 (a) Resultados de LE em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de LE em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II). 79

LISTA DE TABELAS

	<u>Pág.</u>
2.1 Descrição das variáveis do IBIS consideradas neste trabalho.	18
2.2 Descrição dos parâmetros do IBIS considerados neste trabalho.	19
4.1 Limites máximo e mínimo dos parâmetros do modelo IPH-II para a calibração automática.	47
4.2 Faixas de variação dos parâmetros do IPH-II obtidas por uma série de calibrações com o AGE e com o MOCOM, usando população de tamanho $ns = 50$	51
5.1 Funções objetivo, variáveis e parâmetros calibrados em cada nível na calibração hierárquica.	56
5.2 Resultados da calibração hierárquica.	57
5.3 Grupos de calibração em escala temporal.	65
5.4 Resultados da calibração em escala temporal.	66
5.5 Resultados das calibrações em escala temporal com o NSGA-II e com o NSGA-II epidêmico.	72
A.1 Parâmetros com índice de influência não-nulo na escala mensal.	90
A.2 Parâmetros com índice de influência não-nulo na escala diária.	90
A.3 Parâmetros com índice de influência não-nulo na escala horária.	91
A.4 Grupos de calibração em escala temporal.	92

LISTA DE ABREVIATURAS E SIGLAS

AG	–	Algoritmo Genético
AGE	–	Algoritmo Genético Epidêmico
IBIS	–	Integrated Biosphere Simulator
MOCOM	–	Multi-objective Complex Evolution Algorithm
NSGA-II	–	Elitist Non-dominated Sorting Genetic Algorithm II

LISTA DE SÍMBOLOS

A_g	–	Fotossíntese bruta do dossel
$aleaf$	–	Fração de alocação de carbono nas folhas
$alogl_coef$	–	Coefficiente para o cálculo da rugosidade do dossel inferior
$alogu_coef$	–	Coefficiente para o cálculo da rugosidade do dossel superior
$aroot$	–	Fração de alocação de carbono nas raízes finas
$avmuir_coef$	–	Coefficiente para o cálculo da emissividade do dossel
$awood$	–	Fração de alocação de carbono na madeira
$beta2$	–	Parâmetro relacionado à distribuição de raízes finas
$Bmax$	–	Viés máximo da soma cumulativa
$chifuz$	–	Fator de orientação da folha no dossel superior
chs	–	Capacidade térmica dos galhos do dossel superior
chu	–	Capacidade térmica das folhas do dossel superior
chl	–	Capacidade térmica das folhas e galhos do dossel inferior
$clitll_coef$	–	Carbono inicial no reservatório de liteira de folha - lignina
$clitlm_coef$	–	Carbono inicial no reservatório de liteira de folha - metabólico
$clitls_coef$	–	Carbono inicial no reservatório de liteira de folha - estrutural
$clitrl_coef$	–	Carbono inicial no reservatório de liteira de raízes finas - lignina
$clitrm_coef$	–	Carbono inicial no reservatório de liteira de raízes finas - metabólico
$clitrs_coef$	–	Carbono inicial no reservatório de liteira de raízes finas - estrutural
$clitwl_coef$	–	Carbono inicial no reservatório de liteira de madeira - lignina
$clitwm_coef$	–	Carbono inicial no reservatório de liteira de madeira - metabólico
$clitws_coef$	–	Carbono inicial no reservatório de liteira de madeira - estrutural
$coefmub$	–	Coefficiente relacionado à condutância estomática
$csoipas_coef$	–	Carbono inicial no solo - húmus passivo
$csoislon_coef$	–	Carbono inicial no solo - húmus não-protetido lento
$csoislop_coef$	–	Carbono inicial no solo - húmus protegido lento
d_i	–	Efeito elementar do parâmetro i
$dispu_coef$	–	Altura de deslocamento do plano zero para o dossel superior
$fAPAR$	–	Fração absorvida da radiação fotossinteticamente ativa
$funca_coef$	–	Coefficiente da função de temperatura da biomassa dos galhos
$funcb_coef$	–	Coefficiente da função de temperatura da biomassa das raízes
H	–	Parâmetro que caracteriza o decaimento da curva exponencial de infiltração
HE	–	Fluxo de calor sensível
I	–	Índice de sensibilidade do parâmetro
I_b	–	Capacidade mínima de infiltração no solo
I_o	–	Capacidade inicial de infiltração no solo
$kfactor$	–	Fator de multiplicação das constantes de declínio dos reservatórios
Ks	–	Parâmetro que representa o tempo de retardo do escoamento superficial
$Ksub$	–	Parâmetro que representa o tempo médio de esvaziamento do reservatório sub
LAI	–	Índice de área foliar

LE	– Fluxo de calor latente
L_{in}	– Radiação de onda longa incidente
L_{out}	– Radiação de onda longa emitida pela superfície
MAE	– Erro absoluto médio
NEE	– Troca líquida de CO_2 do ecossistema
NIR	– Infravermelho próximo
NPP	– Produção primária líquida
NSGA-II	– Non-dominated Sorted Genetic Algorithm (versão 2)
PAR_o	– Radiação fotossinteticamente ativa refletida
$rgrowth_coef$	– Coeficiente de respiração de crescimento
$rhoveg_NIR$	– Reflectância da folha no dossel superior - NIR
$rhoveg_vis$	– Reflectância da folha no dossel superior - visível
RMAX	– Capacidade máxima de armazenamento em um reservatório de perdas
RMSE	– Raiz do erro quadrático médio
Rn	– Saldo de radiação
$rroot_coef$	– Coeficiente de manutenção de respiração das raízes
$rwood_coef$	– Coeficiente de manutenção de respiração dos troncos
SAI	– Índice de área dos galhos
S_{in}	– Fluxo de radiação solar incidente
$specla$	– Área foliar específica
$stressf_coef$	– Coeficiente relacionado ao estresse hídrico do solo
$tauleaf$	– Constante de tempo de retorno da biomassa de folhas
$tauroot$	– Constante de tempo de retorno da biomassa de raízes finas
$tauveg_NIR$	– Tramitância da folha no dossel superior - NIR
$tauveg_vis$	– Tramitância da folha no dossel superior - visível
$tauwood$	– Constante de tempo de retorno da biomassa de madeira
$tempvm_coef$	– Parâmetro de estresse térmico do V_{max}
u^*	– Velocidade de fricção do vento
$vmax_pft$	– Capacidade de atividade máxima da enzima Rubisco
$wsoi_coef$	– umidade do solo inicial
α	– Parâmetro que representa o escoamento direto de áreas impermeáveis
μ	– Efeito médio do parâmetro na saída do modelo
μ^*	– Efeito absoluto médio do parâmetro na saída do modelo
σ	– Efeito não-linear ou de interações dos parâmetros na saída do modelo

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
2 MODELOS DE DINÂMICA DE FLUIDOS GEOFÍSICOS	7
2.1 O Modelo Hidrológico IPH-II	7
2.1.1 Evaporação e interceptação	7
2.1.2 Separação dos volumes	8
2.1.3 Escoamentos superficial e subterrâneo	9
2.2 O Modelo de Superfície para Atmosfera IBIS	10
3 CALIBRAÇÃO DE MODELOS	21
3.1 Problema Inverso Formulado como um Problema de Otimização	21
3.2 Análise de Sensibilidade: Método de Morris	22
3.3 Otimização Multiobjetivo	26
3.3.1 Formalização do Problema Multiobjetivo	26
3.3.2 Abordagens e Técnicas para a Otimização Multiobjetivo	31
3.3.2.1 Algoritmos Genéticos	32
3.4 Um Algoritmo Genético Epidêmico para a Calibração Multiobjetivo do IPH-II	36
3.4.1 Passos do AGE	37
3.4.2 Evolução dos Complexos no MOCOM	39
3.5 O Algoritmo NSGA-II	40
3.5.1 Um Operador de Epidemia para o NSGA-II	43
3.6 Métricas de Avaliação de Erros: Funções Objetivo	44
3.7 Estratégia OPTIS	45
4 CALIBRAÇÃO DO MODELO HIDROLÓGICO IPH-II - RE- SULTADOS	47
5 CALIBRAÇÃO DO IBIS - RESULTADOS	53
5.1 Testes com a Calibração hierárquica	55
5.2 Calibração a Partir de uma Análise de Sensibilidade Dependente da Es- cala Temporal	64

5.3	Calibração a Partir de uma Análise de Sensibilidade Dependente da Escala Temporal, usando o NSGA-II com Epidemia	71
6	CONCLUSÕES	81
	REFERÊNCIAS BIBLIOGRÁFICAS	83
	APÊNDICE - A Análise de Sensibilidade para a Calibração em Escala Temporal	89

1 INTRODUÇÃO

Nos últimos anos, a ciência tem se beneficiado de modelos computacionais como ferramentas importantes e cada vez mais confiáveis em muitos tipos de investigações. Tais modelos estabelecem uma representação de sistemas físicos, expressando-os através de equações matemáticas que são resolvidas aproximadamente por técnicas numéricas codificadas em programas de computador.

Fenômenos de dinâmica de fluidos, confinados, neutros ou ionizados, e nas áreas de meteorologia, hidrologia, de transferência de calor em estruturas de engenharia, por exemplo, são modelados com equações matemáticas, muitas vezes sob a forma de sistemas complexos de equações diferenciais. Para se garantir a aderência de um modelo à realidade do fenômeno modelado, torna-se necessário o ajuste de vários parâmetros do mesmo, procedimento este conhecido como *calibração de parâmetros*.

Com a calibração dos parâmetros, busca-se ajustar o modelo para que o mesmo simule adequadamente o fenômeno considerado. Para tal, utilizando os dados observados, buscam-se valores para os parâmetros para os quais a saída do modelo corresponda, com bastante razoabilidade, à realidade do fenômeno estudado. Está-se, portanto, adotando um procedimento inverso no estudo do fenômeno, partindo-se dos efeitos (dados) e tentando-se chegar às causas (parâmetros).

Considerando-se que um determinado fenômeno físico esteja bem compreendido, a partir de sua modelagem matemática, pode-se chegar a uma função que relaciona um conjunto de parâmetros, M , do modelo a um conjunto de dados, D , ou seja: $G(M) = D$. Chamamos de modelo direto ao problema de encontrar D , dado M , ou seja, a partir das causas, obter os efeitos. Quando, inversamente, passamos ao problema de encontrar M , dado D , o modelo é dito inverso, e se está, com isso, partindo-se dos efeitos, tentando-se determinar as causas. Neste caso, estamos diante de um problema que é denominado *problema inverso*.

Em modelos de fluidos geofísicos, há um número finito de parâmetros a serem estimados bem como um número finito de dados observados a serem considerados. Trata-se, portanto, de um *problema inverso discreto* ou *problema de estimação de parâmetros*.

No início do século XX, o matemático francês Jacques Hadamard definiu um problema matematicamente *bem posto* como sendo aquele que satisfaz às seguintes condições: (i) possui solução, (ii) sua solução é única, (iii) sua solução tem uma

dependência contínua dos dados de entrada. Se uma das condições acima não se verifica, o problema é dito *mal-posto*. Problemas inversos são tipicamente problemas mal-postos.

A calibração de parâmetros é um problema de muitas soluções igualmente possíveis, podendo ocorrer soluções que não representem a realidade do problema. Por outro lado, das soluções aceitáveis tecnicamente, pode não ser possível distinguir a mais adequada. Portanto, a calibração de parâmetros é um problema inverso mal-posto.

A calibração manual por tentativa e erro é uma das técnicas pioneiras de calibração. Consiste em um procedimento iterativo em que o usuário do modelo altera os valores dos parâmetros a cada tentativa e compara os resultados obtidos com os valores medidos até se chegar a um conjunto de parâmetros para os quais, no seu entendimento, os resultados do modelo são os mais apropriados para o processo simulado. A comparação a cada tentativa é feita geralmente com alguma métrica ou estatística que avalia a medida do erro entre os valores calculados (saída do modelo) e os valores observados. Apesar de sua robustez e simplicidade, a calibração manual depende essencialmente da habilidade do usuário. Além de entender o modelo, é preciso possuir uma habilidade para se encontrar uma solução por tentativa, mesmo que, implicitamente, sejam consideradas múltiplas métricas para a avaliação dos erros. A falta de um critério mensurável de comparação e a utilização de critérios subjetivos que levam a conjuntos de parâmetros diferentes por usuários diferentes podem ser apontadas como grandes deficiências da calibração manual.

O processo de calibração torna-se mais rápido e mais eficiente com o uso de técnicas de calibração automática, baseadas no uso de computação intensiva. Nestas técnicas, o procedimento básico para o ajuste dos parâmetros de um modelo considera os elementos seguintes: uma ou mais medidas de avaliação dos erros (as chamadas medidas de desempenho ou funções objetivo), um algoritmo de otimização e um critério de parada. Assim, a calibração se torna um problema de otimização de uma única função (calibração mono objetivo) ou de mais de uma função (calibração multiobjetivo).

A calibração de um modelo complexo, envolvendo muitos parâmetros e muitas funções objetivo, pode ser um problema bastante difícil, se não impossível, de se resolver de forma exata e eficiente (NASH; SUTCLIFFE, 1970; JOHNSTON; PILGRIM, 1976; PICKUP, 1977; DUAN et al., 1993). O grau de dificuldade geralmente está associado a três fatores principais:

- 1 O número de dimensões (parâmetros) do problema. Quanto maior o número de parâmetros, maior é o espaço de busca e mais esforço é necessário para se encontrar o ótimo global;
- 2 As características da função objetivo. A superfície da função objetivo no espaço dos parâmetros pode não ser suave ou pode nem mesmo ser contínua. Além disso, pode haver diversas regiões de atração no espaço de busca, sendo que cada uma pode conter vários ótimos locais (DUAN et al., 1993);
- 3 A quantidade de funções objetivo. É comum a necessidade de se calibrar um modelo levando em consideração mais de uma variável de saída do modelo ou, mesmo considerando apenas uma variável, fazer uso de mais de uma medida de ajuste para sua avaliação. Esta característica multiobjetivo torna o processo de calibração muito mais complicado, principalmente porque, em geral, tais medidas podem ser conflitantes (objetivos conflitantes) - a melhora em uma medida pode ser acompanhada pela piora em outra.

A dificuldade apresentada em (1) pode ser amenizada através de uma análise de sensibilidade (AS) que permita identificar, entre os parâmetros a serem ajustados, os mais importantes para o modelo, restringindo, então, o espaço de busca. Para resolver o problema descrito em (2) é necessário o uso de técnicas de busca global, as chamadas meta-heurísticas, que podem fornecer boas soluções para o ajuste. Para se considerar o fator (3) é preciso usar técnicas de otimização multiobjetivo. Ainda assim, um grande número de funções objetivo é um fator considerável no aumento da dificuldade de se determinar as soluções ótimas, sendo importante buscar-se alguma forma de reduzir tal número ou dividir o problema em vários subproblemas com um número menor de funções objetivo ou, ainda, agrupar vários objetivos em uma única função a ser otimizada.

Na calibração mono objetivo, considera-se uma única medida de desempenho, entre as várias possíveis, obtendo-se como resultado um conjunto único de parâmetros que otimizam tal medida. Na calibração multiobjetivo, considera-se a otimização simultânea de mais de uma medida de desempenho, que resultará em vários conjuntos de parâmetros que irão constituir um conjunto de soluções ótimas (conjunto de Pareto). Deste conjunto de soluções ótimas, espera-se, de forma apropriada, chegar-se a uma resposta desejada para o ajuste dos parâmetros do modelo.

Dos diferentes algoritmos de otimização multiobjetivo que vêm sendo desenvolvidos,

os algoritmos evolucionários têm recebido uma maior atenção nas últimas quatro décadas. Em particular, os algoritmos genéticos têm-se mostrado bastante eficazes nesta tarefa (FONSECA; FLEMING, 1993; SRINIVAS; DEB, 1994; DEB et al., 2000; DEB, 2004).

Este trabalho trata da calibração multiobjetivo dos parâmetros de dois modelos de fluidos geofísicos: o modelo hidrológico de pequenas bacias IPH-II (TUCCI, 2005) e o modelo de superfície atmosférico IBIS (FOLEY et al., 1996). Para a calibração do modelo IPH-II, é apresentado um algoritmo genético com epidemia, o AGE (ARAÚJO et al., 2013), desenvolvido como uma variante de um algoritmo evolucionário, MOCOM (YAPO et al., 1998), que já vem sendo usado na calibração dos parâmetros do IPH-II. Para a calibração do IBIS, usa-se o Optis, um software de calibração multiobjetivo do IBIS desenvolvido por Varejão (2009), Varejão et al. (2013) que usa como técnica de calibração o algoritmo NSGA-II (DEB et al., 2000), sendo, aqui, propostas algumas mudanças nas definições das funções objetivo, a implementação de um operador de epidemia no NSGA-II e uso de uma estratégia de calibração baseada na escala temporal das variáveis (MINJIAO; XIAO, 2014).

Este trabalho está estruturado da seguinte forma:

- O Capítulo 2 descreve os modelos de dinâmica de fluidos IPH-II e IBIS, com a apresentação dos conceitos e das equações matemáticas que envolvem os parâmetros e as variáveis dos dois modelos consideradas nas calibrações;
- O Capítulo 3 trata, inicialmente, de uma breve descrição da formulação de problemas inversos como problemas de otimização. Em seguida, é feita uma abordagem teórica da análise de sensibilidade como um instrumento de grande importância na calibração de modelos de alta complexidade, apresentando o conceito, os principais tipos de análise de sensibilidade e o Método de Morris. A seguir, são apresentados os conceitos da otimização multiobjetivo, expondo-se algumas das principais técnicas para a abordagem dos problemas de otimização multiobjetivo. Uma seção deste capítulo é dedicada à apresentação de um algoritmo genético epidêmico, o AGE, que foi implementado, como parte deste trabalho de tese, para a calibração multiobjetivo do modelo IPH-II. Em outra seção, é apresentado o algoritmo multiobjetivo NSGA-II, o qual é utilizado na calibração do modelo IBIS. Na seção seguinte, é apresentado um operador de epidemia proposto para o NSGA-II, visando a uma calibração do IBIS com uma versão do NSGA-II com epidemia. Em seguida, há uma seção onde são apresentadas algumas

métricas de avaliação de erros (funções objetivo) que têm sido, comumente, usadas na calibração de modelos hidrológicos e que são utilizadas no Optis. Finalmente, uma última seção deste capítulo é dedicada à apresentação do software Optis;

- No Capítulo 4, são apresentados os resultados de calibração do modelo IPH-II;
- No Capítulo 5, são apresentados os resultados de calibração do IBIS;
- O Capítulo 6 é dedicado à apresentação das conclusões gerais resultantes deste trabalho, com a apresentação de algumas sugestões de trabalhos futuros e ideias para aprimoramento das estratégias aqui apresentadas.
- No Apêndice, encontram-se os resultados da análise de sensibilidade usadas na seção 5.2.

2 MODELOS DE DINÂMICA DE FLUIDOS GEOFÍSICOS

Os modelos de fluidos geofísicos são desenvolvidos para simular a dinâmica de fluidos na escala de forças e dimensões do planeta Terra, tais como: os fenômenos atmosféricos (estudo do clima e dinâmica das massas de ar), a dinâmica dos oceanos (ondas, vórtices e correntes), os movimentos do magma e do núcleo externo do interior da Terra (efeito dínamo do núcleo e convecção do manto terrestre).

Neste capítulo, abordaremos os modelos a serem calibrados neste trabalho, a saber: o modelo hidrológico de pequenas bacias IPH-II e o modelo de superfície para atmosfera IBIS.

2.1 O Modelo Hidrológico IPH-II

O modelo hidrológico IPH-II é um modelo chuva-vazão do tipo concentrado no qual os processos hidrológicos são representados por variáveis concentradas no espaço e a bacia hidrológica é representada por uma precipitação média (TUCCI, 2005).

Baseado no algoritmo de separação de escoamento desenvolvido por Berthelot (1970), Tucci et al. (1981) desenvolveram o modelo IPH-II, utilizando o referido algoritmo e associando-o a outros três: de perdas, de escoamento superficial e subterrâneo. Com o objetivo de tornar o modelo aplicável em diferentes problemas de recursos hídricos, com um mínimo de parâmetros a serem ajustados, (TUCCI; CAMPANA, 1993) fizeram modificações na estrutura do algoritmo de escoamento superficial. O modelo é composto, basicamente, por três algoritmos: um para as perdas por evaporação e interceptação, um para a separação dos escoamentos e um para a propagação dos escoamentos superficial e subterrâneo.

2.1.1 Evaporação e interceptação

Da precipitação que chega à superfície, parte é perdida por evaporação e parte é retida pela interceptação. Representa-se, então, o total da precipitação retida por interceptação através de um reservatório de perdas, com uma capacidade máxima R_{MAX} e uma variável de estado R_t (Figura 2.1b). O restante da precipitação será a entrada para o algoritmo de separação do escoamento.

Caso a precipitação não seja suficiente para atender a evapotranspiração potencial, considera-se a parte desta que é atendida pelo reservatório de perdas, sendo o restante da evapotranspiração retirado do solo, de acordo com o seu estado de umidade. Para se obter a retirada de água do solo, utiliza-se a seguinte relação linear entre a

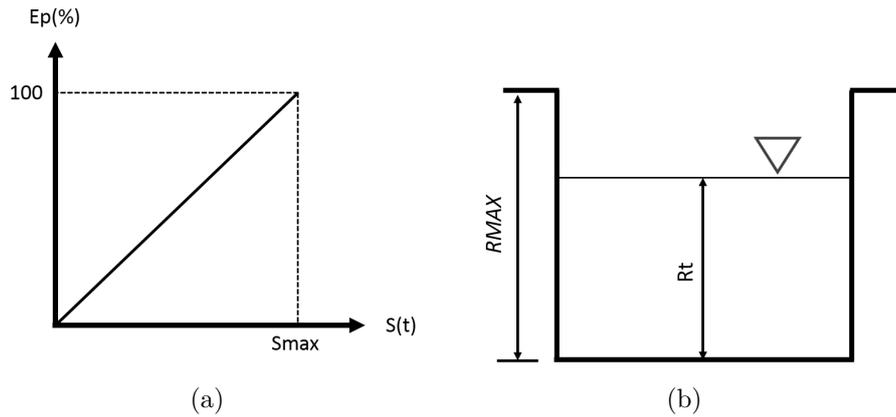


Figura 2.1 - (a) Relação entre Evapotranspiração potencial e umidade do solo; (b) Reservatório de perdas por interceptação e depressões do solo.

Fonte: Adaptado de Tucci (2005).

percentagem da evapotranspiração potencial e a umidade do solo (Figura 2.1a):

$$E(t) = \frac{E_p(t) \cdot S(t)}{S_{max}} \quad (2.1)$$

onde $E(t)$ é a evapotranspiração da superfície no tempo t ; $E_p(t)$ é a evapotranspiração potencial; $S(t)$ é o estado de umidade da camada superior do solo; S_{max} é a capacidade máxima de umidade da camada superior do solo.

O algoritmo de perdas por evaporação e interceptação utiliza um único parâmetro, $RMAX$, que descreve a capacidade máxima de armazenamento em um reservatório de perdas. Este parâmetro representa a lâmina de água interceptada pela vegetação ou armazenada em poças, e que fica facilmente disponível para a evaporação (COLLISCHONN; TUCCI, 2003).

2.1.2 Separação dos volumes

A parcela de precipitação resultante pode gerar escoamento superficial ou infiltrar no subsolo. Não há infiltração no escoamento superficial gerado pela parcela de água que precipita sobre áreas impermeáveis, como ocorre em regiões urbanas, nas quais as áreas impermeáveis estão diretamente ligadas com sistemas de coletores pluviais. Torna-se necessário, portanto, diferenciar as áreas impermeáveis que escoam diretamente para os pluviais daquelas que escoam para partes da bacia com superfície permeável. Da parcela que precipita sobre áreas permeáveis, é necessário calcular

o volume infiltrado, o volume percolado para o aquífero e o volume que gera o escoamento superficial (descrito na seção seguinte).

Consideram-se as seguintes equações para os cálculos da infiltração $I(t)$ e da percolação $T(t)$:

$$I(t) = I_b + (I_o - I_b)e^{-kt} \quad (\text{Equação de Horton}) \quad (2.2)$$

$$T(t) = I_b(1 - e^{-kt}) \quad (2.3)$$

onde I_b é a capacidade de infiltração quando solo está saturado; I_o é a capacidade de infiltração inicial; k é um parâmetro que caracteriza o decaimento da curva exponencial de infiltração $h = e^{-kt}$, dependente das características do solo.

No algoritmo de separação do escoamento, há três parâmetros, próprios da equação de Horton: os parâmetros I_o e I_b que representam as capacidades inicial e mínima de infiltração no solo, respectivamente, e o parâmetro H , que representa a função de variação da capacidade de infiltração no solo, sendo dado pela equação $H = e^{-k}$.

2.1.3 Escoamentos superficial e subterrâneo

A propagação do escoamento superficial é feita com o método de Clarke (1973) que utiliza parâmetros de tempo de concentração que podem ser mantidos fixos ou serem calibrados, dependendo da disponibilidade de informações de características físicas da bacia. O método de Clarke utiliza o histograma tempo-área para levar em conta a translação, e o modelo do reservatório linear simples para o amortecimento. A teoria do reservatório linear simples considera que o escoamento superficial sofre amortecimento equivalente ao de um reservatório, através da seguinte equação:

$$\frac{dS}{dt} = V_s - Q_s \quad (2.4)$$

onde S corresponde ao armazenamento; V_s é a vazão de entrada; Q_s é a vazão de saída.

Considerando que o armazenamento e a vazão de saída sejam lineares, tem-se que:

$$S = K_s Q_s \quad (2.5)$$

onde K_s é o tempo médio de esvaziamento do reservatório.

A partir das equações acima, e considerando que a vazão de entrada seja constante, chega-se à seguinte equação diferencial ordinária:

$$Ks \frac{dQ_s}{dt} + Q_s = V_s \quad (2.6)$$

que irá representar o reservatório.

Portanto, usando o modelo do reservatório linear simples para se considerar o efeito do armazenamento da bacia, tem-se o parâmetro Ks , que representa o tempo médio de esvaziamento do reservatório de retardo do escoamento superficial, o qual depende do deslocamento do escoamento na bacia. Deve ser definido o percentual de área impermeável da bacia.

Na propagação do escoamento subterrâneo, utiliza-se também um modelo de reservatório linear simples com o parâmetro $Ksub$, que representa o tempo médio de esvaziamento do reservatório subterrâneo de escoamento. Para a simulação de séries contínuas de longo período, introduz-se um parâmetro, α , que só é usado na simulação do processo de separação do escoamento no caso em que a capacidade de infiltração supera a precipitação. Este parâmetro representa o escoamento direto de áreas impermeáveis.

Para a simulação do processo, além dos sete parâmetros acima, devem ser informadas a área da bacia e as variáveis de entrada do modelo: precipitação e evaporação (BRAVO et al., 2007).

A vazão estimada pelo modelo, \bar{Q}_t , depende, em cada intervalo de tempo t , dos valores observados de precipitação, P_t , e de evaporação, E_t , e do vetor de parâmetros, θ , ou seja:

$$\bar{Q}_t = \mathcal{H}(P_t, E_t, \theta) \quad (2.7)$$

onde $\mathcal{H}(\cdot)$ representa o modelo hidrológico.

2.2 O Modelo de Superfície para Atmosfera IBIS

O modelo IBIS (Integrated Biosphere Simulator) foi projetado para conectar explicitamente processos de superfície terrestre e processos hidrológicos, ciclos biogeoquímicos terrestres e dinâmica de vegetação em uma única estrutura de modelagem, visando a uma melhor compreensão das consequências da atividade humana nos processos biofísicos globais (FOLEY et al., 1996; KUCHARIK et al., 2000).

Por considerar mudanças transientes na composição e na estrutura da vegetação em resposta a mudanças no meio ambiente, o IBIS é classificado como um *modelo de vegetação dinâmica global*. Além de ser capaz de simular a dinâmica da fenologia da vegetação, ele inclui processos de física da superfície terrestre, interações entre a superfície terrestre e a atmosfera, ciclos de carbono e do nitrogênio, bem como ciclagem de nutrientes, simulando também os efeitos das mudanças no balanço de carbono terrestre sobre o clima e na concentração atmosférica de CO_2 .

Em sua versão mais atual (2D), o IBIS pode ser acoplado a outros modelos de circulação geral da atmosfera de mesma resolução, representando as componentes de superfície destes. A sua estrutura básica (Figura 2.2) é constituída de diversos módulos que se comunicam entre si para fazer uma variedade de simulações com incorporações de processos do ecossistema em uma única estrutura, considerando aspectos tais como:

- Processos de superfície: transferência de energia, água, CO_2 e *momentum*;
- Fisiologia da planta: fotossíntese, condutância estomática e respiração;
- Fenologia da vegetação: aparecimento e crescimento (medidos por graus-dia);
- Dinâmica da vegetação: índice de área foliar e biomas para 12 tipos funcionais de plantas;
- Competição da vegetação: por luz, por água e por nutrientes;
- Biogeoquímica do solo: ciclos de CO_2 e de nitrogênio, nutrientes terrestres, produtividade primária líquida, respiração do solo e decomposição.

Tais módulos são interligados, seguindo uma hierarquia conceitual e organizados conforme a ocorrência dos processos (superfície, biogeoquímica do solo, dinâmica da vegetação e fenologia das plantas) em escala temporal com intervalos de integração que variam de 60 minutos a um ano (Figura 2.2).

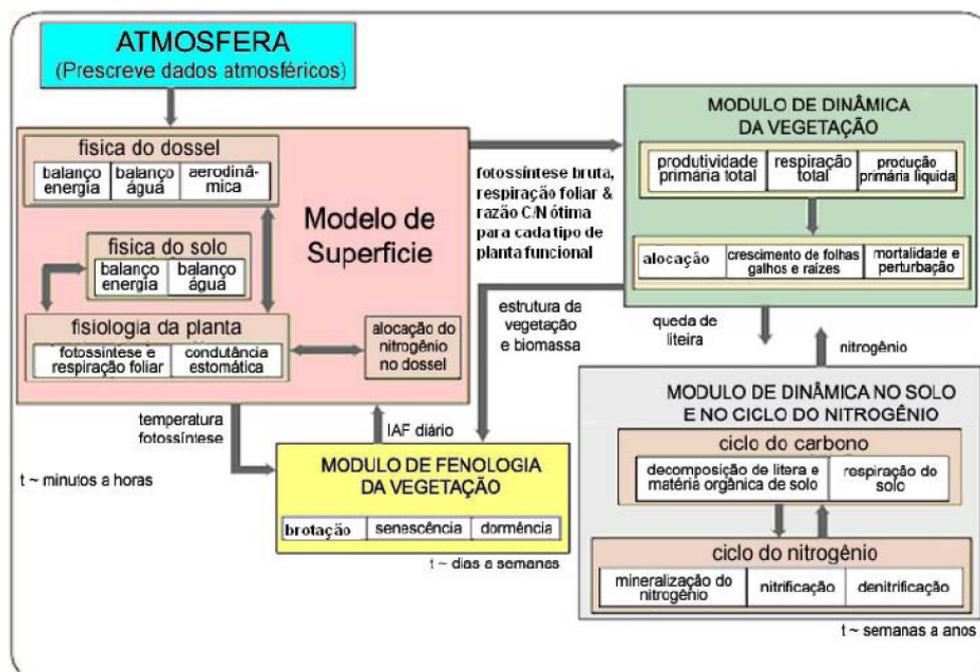


Figura 2.2 - Esquema modular do IBIS 2.6.

Fonte: Adaptado de Kucharik et al. (2000).

Dentre os diversos processos e variáveis simulados pelo IBIS, consideramos, neste trabalho, as seguintes variáveis (saídas do modelo) que foram usadas nas calibrações:

- PAR_o : radiação fotossinteticamente ativa refletida;
- f_{APAR} : fração absorvida de radiação fotossinteticamente ativa;
- R_n : saldo de radiação;
- u^* : velocidade de fricção do vento;
- HE e LE : fluxos de calor sensível e latente, respectivamente;
- NEE : troca líquida do ecossistema;
- NPP : produção primária líquida;
- LAI : índice de área foliar.

Uma descrição de cada uma das variáveis acima é apresentada a seguir, conforme Varella (2009).

A radiação fotossinteticamente ativa refletida é calculada no modelo através da equação:

$$PAR_o = (S_{dir,vis} \cdot r_{dir,vis} + S_{dif,vis} \cdot r_{dif,vis}) \quad (2.8)$$

onde $S_{dir,vis}$ e $S_{dif,vis}$ são os fluxos de radiação solar direta e difusa incidentes, na banda visível; $r_{dir,vis}$ e $r_{dif,vis}$ são a reflectância direta e difusa, respectivamente, da superfície na banda do visível.

A fração da radiação fotossinteticamente ativa absorvida pelo dossel, f_{APAR} , inclui a fração da radiação fotossinteticamente ativa incidente refletida pelo dossel (PAR) e a fração absorvida pelo solo. A radiação fotossinteticamente absorvida, $APAR$, é a energia solar ($400nm - 700nm$) consumida pelo dossel no processo de fotossíntese. A $APAR$ total é calculada como a soma da $APAR$ do dossel superior e inferior:

$$APAR = (PAR_{udir} \cdot a_{udir} + PAR_{udif} \cdot a_{udif}) \cdot (PAR_{ldir} \cdot a_{ldir} + PAR_{ldif} \cdot a_{ldif}) \quad (2.9)$$

onde PAR_{udir} é a PAR (radiação direta) incidente no dossel superior, a_{udir} é a fração da PAR direta absorvida pelo dossel superior, PAR_{udif} é a PAR difusa incidente no dossel superior, a_{udif} é a fração da PAR difusa absorvida pelo dossel superior, PAR_{ldir} é a PAR direta incidente no dossel inferior, a_{ldir} é a fração da PAR direta absorvida pelo dossel inferior, PAR_{ldif} é a PAR difusa incidente no dossel inferior, a_{ldif} é a fração da PAR difusa absorvida pelo dossel inferior. A f_{APAR} é calculada para cada hora do dia através da razão:

$$f_{APAR} = \frac{APAR}{PAR_{in}} \quad (2.10)$$

O saldo de radiação, R_n , é a contabilização líquida entre toda a energia radiante recebida e perdida pela superfície, sendo calculado através do saldo dos balanços de ondas curtas e de ondas longas:

$$R_n = S_{in,vis} \cdot (1 - r_{vis}) + S_{in,nir} \cdot (1 - r_{nir}) + L_{in} - L_{out} \quad (2.11)$$

onde $S_{in,banda}$ é o fluxo de radiação solar incidente, na banda indicada; r_{banda} é a reflectância da superfície em cada banda; L_{in} é a radiação de onda longa que atinge a superfície e L_{out} é a radiação de onda longa líquida emitida pela superfície, sendo:

$$L_{out} = (1 - fu) \cdot f_{up} + fu \cdot [(1 - em_u) \cdot (1 - em_s) f_{up} + em_u \cdot ste_f \cdot (tu)^4 + em_s \cdot (1 - em_u) \cdot ste_f \cdot (ts)^4] \quad (2.12)$$

$$emu = 1 - e^{-\frac{LAI}{avmuir}} \quad (2.13)$$

$$ems = 1 - e^{-\frac{SAI}{avmuir}} \quad (2.14)$$

onde emu e ems são as emissividades das folhas e galhos superiores da vegetação, respectivamente; fu é a fração da área média coberta pela vegetação; fup é o fluxo de radiação IR descendente abaixo do nível das árvores; tu e ts são as temperaturas das folhas e galhos do dossel superior, respectivamente; $stef$ é a constante de Stefan-Boltzmann; LAI é o índice de área foliar; SAI é o índice de área de galhos e $avmuir$ é a profundidade ótica difusa média.

A energia radiativa disponível é particionada nos fluxos de calor latente (LE) e sensível (HE) na superfície e no fluxo de calor no solo (G):

$$R_n = HE + LE + G \quad (2.15)$$

A velocidade de fricção do vento, u^* , é calculada pelas equações:

$$u^* = \sqrt{\left[\frac{u_a \cdot vonk}{alogu - alogu}\right]^2 \cdot stramu} \quad (2.16)$$

$$alogu = \ln[\max(0.001; alogu_coef \cdot (z1 - z2))] \quad (2.17)$$

$$alogu = \ln(za - dispu) \quad (2.18)$$

onde u_a é a velocidade do vento; $vonk$ é a constante de Von Karma; $stramu$ é um fator de correção sobre o dossel superior; $z1$ é a altura do topo do dossel superior; $z2$ é a altura da base do dossel superior; za é a altura sobre a superfície da forçante atmosférica; $dispu$ é o deslocamento do plano zero para o dossel superior.

A troca líquida do ecossistema, ou fluxo de CO_2 , é dada pela diferença entre dois termos: a captura de CO_2 pela folha durante a fotossíntese e a emissão de CO_2 pela respiração do solo. A NEE é, portanto, assim calculada:

$$NEE = R_H - NPP \quad (2.19)$$

onde R_H é a respiração e NPP é a produção primária líquida.

A produção primária líquida, NPP , representa o carbono novo armazenado como biomassa nos galhos, folhas e raízes das plantas. Ela corresponde à diferença entre o carbono assimilado durante a fotossíntese e o consumo de carbono através da respiração das plantas, sendo uma medida quantitativa de crescimento da planta e

de sequestro de carbono:

$$NPP = (1 - \eta) \int (A_g - R_{leaf} - R_{steam} - R_{root}) dt \quad (2.20)$$

onde A_g é a fotossíntese bruta do dossel; R_{leaf} , R_{steam} , R_{root} são as respirações de manutenção das folhas, dos troncos e das raízes, respectivamente; η é a fração de carbono perdido devido à respiração de crescimento.

A taxa de fotossíntese de plantas C_3 , que incluem todas as árvores e muitas herbáceas, é representada segundo as equações de Farquhar (FARQUHAR et al., 1980; COLLATZ et al., 1991). A fotossíntese bruta do dossel, A_g , é definida como o mínimo de duas taxas potenciais de fotossíntese: a taxa de fotossíntese limitada pela luz, J_E , e a taxa de fotossíntese limitada pela enzima Rubisco, J_C . Tais taxas são assim calculadas:

$$J_E = \alpha_3 \cdot Q_p \cdot \left(\frac{C_i - \Gamma_*}{C_i + 2\Gamma_*} \right) \quad (2.21)$$

$$J_C = V_m \cdot \frac{C_i - \Gamma_*}{C_i + K_C \cdot \left(1 + \frac{[O_2]}{K_O} \right)} \quad (2.22)$$

onde Q_p é a densidade do fluxo de radiação fotossinteticamente ativa (PAR) absorvida pela folha, α_3 é a eficiência quântica para absorção de plantas C_3 ; C_i é a concentração de CO_2 nos espaços de ar intracelulares da folha; Γ_* é o ponto de compensação para a fotossíntese bruta; V_m é a capacidade de carboxilase máxima da Rubisco; K_C e K_O são os coeficientes de Michaelis-Menten para CO_2 e O_2 , respectivamente.

A capacidade de carboxilase máxima da Rubisco é obtida através das seguintes fórmulas:

$$V_m = V_{max} \cdot tempvm \cdot stresstu \quad (2.23)$$

$$tempvm = \frac{e^{tempvm-coef \cdot (3.47e^3 - \frac{1}{tu})}}{[1 + e^{0.4(5-tleaf)}] \cdot [1 + e^{0.4(tleaf-50)}]} \quad (2.24)$$

onde V_{max} é o valor nominal da capacidade máxima da enzima Rubisco; $tempvm$ é a função de estresse devido à temperatura da folha em Kelvin; $tleaf$ é a temperatura da folha em Celsius; $stresstu$ é a função de estresse devido à umidade em todas as camadas do solo, e é calculada por:

$$stresstu = \sum_{j=1}^L stressu_j \quad (2.25)$$

sendo $stressu_j$ a função de estresse devido a umidade no solo em cada camada j do solo e L o número de camadas do solo. Tal função de estresse pode ser calculada por:

$$stressu = froot.max[0; min(1.0; zwilt)] \quad (2.26)$$

onde $froot$ é a fração de raiz na camada do solo e $zwilt$ é uma função de estresse hídrico, dada por:

$$zwilt = \frac{1 - e^{stressfac.awc}}{1 - e^{stressfac}} \quad (2.27)$$

sendo $stressfac$ o parâmetro para cálculo do estresse hídrico e awc o conteúdo de água disponível.

Em plantas C_4 , a taxa de fotossíntese é definida como:

$$A_g \approx \min(J_I; J_E; J_C) \quad (2.28)$$

onde $J_I = \alpha_4.Q_P$ é a taxa de fotossíntese limitada pela luz; $J_E = V_m$ é a taxa de fotossíntese limitada pela Rubisco e $J_C = k.C_i$ é a taxa de fotossíntese limitada pelo CO_2 a baixas concentrações.

A respiração da folha, R_{leaf} , é determinada por:

$$R_{leaf} = \gamma.V_m \quad (2.29)$$

onde γ é o custo de respiração da folha (KUCHARIK et al., 2000).

As taxas de respiração de manutenção da biomassa dos troncos (R_{stem}) e das raízes finas (R_{root}) são dadas por:

$$R_{stem} = \beta_{stem}.\lambda_{sapwood}.C_{stem,i}.f(T_{stem}) \quad (2.30)$$

$$R_{root} = \beta_{root}.C_{root,i}.f(T_{soil}) \quad (2.31)$$

onde C_{stem} e C_{root} são o carbono contido na biomassa do tronco e das raízes finas, respectivamente; β é o coeficiente de respiração de manutenção; $\lambda_{sapwood}$ é a fração viva da biomassa e $f(T)$ é a função de temperatura de Arrhenius, dada por:

$$f(T) = e^{E_o(\frac{1}{15-T_o} - \frac{1}{T-T_o})} \quad (2.32)$$

sendo T a temperatura (do tronco ou das raízes), E_o um fator de sensibilidade de temperatura e T_o uma temperatura de referência.

A condutância estomática é calculada pela fórmula:

$$g_{s,h_2o} = \frac{m A_n}{C_S} \cdot h_S + b \quad (2.33)$$

onde g_{s,h_2o} é a condutância estomática do vapor d'água na folha; C_S é a concentração de CO_2 na superfície da folha, h_S é a umidade relativa na superfície da folha; m e b representam, respectivamente, os coeficientes de inclinação e intercepto da relação condutância-fotossíntese.

O índice de área foliar, LAI , é a razão entre a área da superfície de folha superior total da vegetação e a área da superfície do solo no qual a vegetação cresce, sendo obtido pela soma dos índices de área foliar de cada tipo funcional de planta:

$$LAI = \sum_{j=1}^N cbiol_j \cdot specla_j \quad (2.34)$$

onde N é o número de tipos funcionais de plantas representados no modelo; $specla$ é a área de folha específica; $cbiol$ é a quantidade de carbono no reservatório de biomassa de folha, calculada por:

$$cbiol = cbiol \cdot e^{\frac{-1}{\tau aleaf}} + aleaf \cdot tauleaf \cdot \max(0; aynpp) \cdot [1 - e^{\frac{-1}{\tau aleaf}}] \quad (2.35)$$

onde $\tau aleaf$ é o tempo de retorno do carbono nas folhas; $aleaf$ é a fração de alocação de carbono para as folhas; $aynpp$ é a produção primária líquida (NPP) total anual.

As mudanças do C nos reservatórios de biomassa são dadas pelas seguintes equações:

$$outcrs = \min(decomps.krs.clitrs; clitrs) \quad (2.36)$$

$$outcws = \min(decompl.kws.clitws; clitws) \quad (2.37)$$

$$outcls = \min(decompl.kls.clitls; clitls) \quad (2.38)$$

$$outcnb = \min(decomps.knb.csoislon; csoislon) \quad (2.39)$$

$$outcpb = \min(decomps.kpb.csoislop; csoislop) \quad (2.40)$$

onde $outcrs$, $outcws$, $outcls$ representam, respectivamente, as quantidades de C saindo dos reservatórios de raízes, madeira e folhas; $outcnb$ e $outcpb$ correspondem aos fluxos para a biomassa de matéria orgânica não protegida e matéria orgânica protegida, respectivamente; krs , kws e kls são as constantes de decomposição para os reservatórios de liteira estrutural de raízes, madeira e folhas, respectivamente;

kpb , knb são constantes de decomposição para matéria orgânica protegida e não-protegida, respectivamente; $decomps$ é um fator de decomposição de matéria orgânica do solo e $decompl$ é um fator de decomposição de liteira.

O IBIS representa a distribuição vertical do sistema radicular de acordo com a equação proposta por Jackson et al. (1997)

$$Y(d) = \frac{1 - (\beta_2)^d}{1 - (\beta_2)^{d_{max}}} \quad (2.41)$$

onde $Y(d)$ é a fração de raízes finas entre a superfície e a profundidade do solo d ; d_{max} é a profundidade máxima do solo e β_2 é um parâmetro de distribuição de raízes finas.

As tabelas seguintes apresentam as variáveis e os parâmetros do IBIS considerados neste trabalho.

Tabela 2.1 - Descrição das variáveis do IBIS consideradas neste trabalho.

No.	Nome	Descrição
1	PAR_o	Radiação fotossinteticamente ativa refletida ($\mu mol m^{-2} s^{-1}$)
2	fAPAR	Fração absorvida da radiação fotossinteticamente ativa (adimensional)
3	Rn	Saldo de radiação ($W m^{-2}$)
4	u^*	Velocidade de fricção do vento ($m s^{-1}$)
5	NEE	Troca líquida de CO_2 do ecossistema ($\mu mol m^{-2} s^{-1}$)
6	HE	Fluxo de calor sensível ($W m^{-2}$)
7	LE	Fluxo de calor latente ($W m^{-2}$)
8	NPP	Produção primária líquida ($kg - C m^{-2}$)
9	LAI	Índice de área foliar (adimensional)

Tabela 2.2 - Descrição dos parâmetros do IBIS considerados neste trabalho.

No.	Nome	Descrição
1	<i>rhoveg_vis</i>	Reflectância da folha no dossel superior - visível (adimensional)
2	<i>rhoveg_NIR</i>	Reflectância da folha no dossel superior - NIR (adimensional)
3	<i>tauveg_vis</i>	Tramitância da folha no dossel superior - visível (adimensional)
4	<i>tauveg_NIR</i>	Tramitância da folha no dossel superior - NIR (adimensional)
5	<i>chifuz</i>	Fator de orientação da folha no dossel superior (-1: vertical, 0: aleatório, 1: horizontal)
6	<i>vmax_pft</i>	Capacidade de atividade máxima da enzima Rubisco ($mol - CO_2m^{-2}s^{-1}$)
7	<i>coefmub</i>	Coefficiente relacionado à condutância estomática (adimensional)
8	<i>chs</i>	Capacidade térmica dos galhos do dossel superior ($Jkg^{-1}m^{-2}$)
9	<i>chu</i>	Capacidade térmica das folhas do dossel superior ($Jkg^{-1}m^{-2}$)
10	<i>chl</i>	Capacidade térmica das folhas e galhos do dossel inferior ($Jkg^{-1}m^{-2}$)
11	<i>beta2</i>	Parâmetro relacionado à distribuição de raízes finas (adimensional)
12	<i>funca_coef</i>	Coefficiente da função de temperatura da biomassa dos galhos (adimensional)
13	<i>funcb_coef</i>	Coefficiente da função de temperatura da biomassa das raízes (adimensional)
14	<i>rroot_coef</i>	Coefficiente de manutenção de respiração das raízes (s^{-1})
15	<i>rwood_coef</i>	Coefficiente de manutenção de respiração dos troncos (s^{-1})
16	<i>rgrowth_coef</i>	Coefficiente de respiração de crescimento (adimensional)
17	<i>tempvm_coef</i>	Parâmetro de estresse térmico do Vmax (adimensional)
18	<i>stressf_coef</i>	Coefficiente relacionado ao estresse hídrico do solo (adimensional)
19	<i>clitll_coef</i>	Carbono inicial no reservatório de liteira de folha - lignina ($kg - Cm^{-2}$)
20	<i>clitlm_coef</i>	Carbono inicial no reservatório de liteira de folha - metabólico ($kg - Cm^{-2}$)
21	<i>clitls_coef</i>	Carbono inicial no reservatório de liteira de folha - estrutural ($kg - Cm^{-2}$)
22	<i>clitrl_coef</i>	Carbono inicial no reservatório de liteira de raízes finas - lignina ($kg - Cm^{-2}$)
23	<i>clitrm_coef</i>	Carbono inicial no reservatório de liteira de raízes finas - metabólico ($kg - Cm^{-2}$)
24	<i>clitrs_coef</i>	Carbono inicial no reservatório de liteira de raízes finas - estrutural ($kg - Cm^{-2}$)
25	<i>clitwl_coef</i>	Carbono inicial no reservatório de liteira de madeira - lignina ($kg - Cm^{-2}$)
26	<i>clitwm_coef</i>	Carbono inicial no reservatório de liteira de madeira - metabólico ($kg - Cm^{-2}$)
27	<i>clitws_coef</i>	Carbono inicial no reservatório de liteira de madeira - estrutural ($kg - Cm^{-2}$)
28	<i>csoipas_coef</i>	Carbono inicial no solo - húmus passivo ($kg - Cm^{-2}$)
29	<i>csoislon_coef</i>	Carbono inicial no solo - húmus não-protégido lento ($kg - Cm^{-2}$)
30	<i>csoislop_coef</i>	Carbono inicial no solo - húmus protegido lento ($kg - Cm^{-2}$)
31	<i>wsoi_coef</i>	umidade do solo inicial (adimensional)
32	<i>kfactor</i>	Fator de multiplicação das constantes de declínio dos reservatórios de C (adimensional)
33	<i>tauleaf</i>	Constante de tempo de retorno da biomassa de folhas (anos)
34	<i>tauroot</i>	Constante de tempo de retorno da biomassa de raízes finas (anos)
35	<i>tauwood</i>	Constante de tempo de retorno da biomassa de madeira (anos)
36	<i>specla</i>	Área foliar específica (m^2kg^{-1})
37	<i>aleaf</i>	Fração de alocação de carbono nas folhas (adimensional)
38	<i>aroot</i>	Fração de alocação de carbono nas raízes finas (adimensional)
39	<i>awood</i>	Fração de alocação de carbono na madeira (adimensional)
40	<i>dispu_coef</i>	Altura de deslocamento do plano zero para o dossel superior (m)
41	<i>alogl_coef</i>	Coefficiente para o cálculo da rugosidade do dossel inferior (adimensional)
42	<i>alogu_coef</i>	Coefficiente para o cálculo da rugosidade do dossel superior (adimensional)
43	<i>avmuir_coef</i>	Coefficiente para o cálculo da emissividade do dossel (adimensional)

3 CALIBRAÇÃO DE MODELOS

O processo de *calibração* consiste em identificar parâmetros de forma a tornar um modelo matemático fiel à representação de um determinado fenômeno. Contudo, tais parâmetros têm influência diferenciada sobre o modelo, isto é, o modelo tem resposta (sensibilidade) diferente para cada parâmetro. Deste modo, pode-se estruturar dois estágios distintos no processo de calibração: análise de sensibilidade e identificação de parâmetros.

A identificação de parâmetros é formulada como um problema de otimização (ver seção abaixo). Se mais de uma função objetivo é usada, o processo torna-se uma otimização multi-objetivo, onde um conjunto de soluções não dominadas deve ser determinado: o conjunto de Pareto. Estratégias de otimização multi-objetivo são descritas (MOCON e NSGA-II, incluindo a versão OPTIS). A técnica de algoritmo genético com epidemia é descrita e será incorporada às estratégias citadas.

3.1 Problema Inverso Formulado como um Problema de Otimização

Diferentemente dos *problemas diretos*: $A(u) = f$, com $u \in X$ e $f \in F$, que se caracterizam por serem problemas *bem postos*, problemas inversos são problemas *mal postos*. Hadamard (1923) definiu um problema matematicamente *bem posto* como sendo aquele que satisfaz às seguintes condições: (i) possui solução, (ii) sua solução é única, (iii) sua solução tem uma dependência contínua dos dados de entrada. Se uma das condições acima não se verifica, o problema é dito *mal-posto*.

Uma teoria geral para resolver problemas inversos foi desenvolvida pelo matemático russo Andrey Nikolayevich Tikhonov, em que seria possível encontrar um subconjunto $W \subset X$ no qual o problema inverso seja um problema bem posto. Os elementos do conjunto W são caracterizados por alguma propriedade. A propriedade em geral exige suavidade da solução, o que caracteriza a *teoria de regularização* (TIKHONOV; ARSENIN, 1977).

O processo de regularização busca soluções com regularidade global. Na formulação matemática do método, o problema inverso é formulado como um problema de otimização com restrição:

$$\min \{\|A(u) - f^\delta\|_2^2\} \quad \text{sujeito a: } \Omega[u] < \rho \quad (3.1)$$

onde Ω é o operador de regularização. O problema acima pode ser descrito como um problema de otimização sem restrição fazendo-se uso de um multiplicador de

Lagrange (α):

$$\min \{ \|A(u) - f^\delta\|_2^2 + \alpha\Omega[u] \}. \quad (3.2)$$

A literatura tem relatado que problemas inversos do tipo identificação de parâmetros, como no caso de problemas de calibração de modelos, podem ser resolvidos sem uso explícito de regularização, fazendo-se uso de meta-heurísticas. Estas técnicas estocásticas de otimização permitem sondar de forma exaustiva o espaço de solução. Deste modo, os elementos do subconjunto W podem ser identificados¹. Para o problema de estimação de funções a regularização é necessária. Uma descrição mais formal da diferença entre estimação de parâmetros e estimação de função pode ser vista na Seção 1.1 de (VELHO, 2008).

3.2 Análise de Sensibilidade: Método de Morris

A análise de sensibilidade (AS) é o estudo da relação entre as informações de entrada e as informações de saída de um modelo com o objetivo de entender que fatores de variação nos dados de entrada do modelo estarão associados a variações observadas na saída. No contexto da calibração de parâmetros de um modelo, a análise de sensibilidade procura investigar a resposta do modelo diante das variações nos seus parâmetros de modo a determinar:

- 1 Que parâmetros do modelo (ou de parte do modelo) são insignificantes para determinada saída e que, por isso, podem ser desconsiderados na calibração. A eliminação de parâmetros sem influência em determinadas saídas do modelo permite uma redução na dimensão no espaço dos parâmetros, o que facilita o processo de busca.
- 2 A região ótima no espaço de busca dos parâmetros. É preciso considerar uma análise de sensibilidade global, devendo as investigações serem feitas em todo o espaço de parâmetros e não apenas em torno de um ponto específico.
- 3 Se há, e quais são, parâmetros que interagem uns com os outros. Frequentemente, há parâmetros que possuem efeitos que são combinados e não podem ser reduzidos à soma dos efeitos individuais.

¹Isto quase nunca acontece com os métodos determinísticos, pois ficam presos em mínimos locais.

Sobre os tipos de análise de sensibilidade, [Saltelli et al. \(2009\)](#) propuseram as seguintes categorias: métodos de triagem, AS local e AS global. Enquanto as duas últimas classes são caracterizadas pela forma como os parâmetros são tratados, a primeira é caracterizada em relação ao uso (triagem) dos parâmetros.

Na *Triagem*, os experimentos são conduzidos para identificar os parâmetros que controlam a maior parte da variabilidade das saídas do modelo. Baseia-se na ideia de que a influência dos parâmetros dos modelos seguem a chamada *Lei de Pareto*, a qual afirma que, para muitos fenômenos, 80% das consequências advêm de 20% das causas, de modo que a maioria dos parâmetros pode não ser influente no modelo. Métodos de triagem podem, portanto, ser bastante úteis para modelos com alto custo computacional e com grande número de parâmetros. Embora propensos a fornecer medidas qualitativas de sensibilidade, permitindo estabelecer uma ordem de importância dos fatores de entrada do modelo, os métodos de triagem não determinam quantitativamente o quanto um parâmetro é mais importante do que outro. Neste tipo de AS há que se avaliar a relação entre o ganho com o custo computacional e perdas na informação.

As técnicas de *AS Local* procuram se concentrar nos efeitos locais dos parâmetros no modelo, sendo feitas, geralmente, através do cálculo de derivadas parciais das funções de saída do modelo em relação aos parâmetros. Para o cálculo numérico de tais derivadas, os parâmetros são variados em um pequeno intervalo em torno de um dado ponto. Os métodos de AS local são eficientes em determinar os parâmetros mais importantes do modelo apenas quando o mesmo é linear em todos os seus parâmetros, a menos que se considere algum tipo de média das derivadas calculadas por todo o espaço de parâmetros. A grande limitação ou mesmo inadequação de tais métodos se dá quando o modelo considerado possui características não-monotônicas ([SALTELLI et al., 2005](#)).

Na *AS Global*, consideram-se técnicas que atribuem as incertezas na saída dos modelos de forma proporcional. As medidas de sensibilidade de cada parâmetro consideram todo o seu intervalo de variação e são calculadas variando-se todos os outros parâmetros. Assim torna-se possível calcular, além dos efeitos individuais de cada parâmetro, os efeitos de cada parâmetro nos demais.

Ao se escolher uma técnica de AS, é preciso levar em consideração a mais importante característica de uma boa análise de sensibilidade: a sua *independência do modelo*. É preciso garantir que a acurácia do método de AS escolhido não seja influenciada pelo nível de aditividade ou linearidade do modelo ([VAREJÃO, 2009](#)).

O Método de [Morris \(1991\)](#) de análise de sensibilidade é um método global de triagem, adequado para se trabalhar com modelos que possuem um grande número de parâmetros. Ele se propõe a identificar que parâmetros do modelo possuem efeitos que são: (a) desprezíveis, (b) lineares e aditivos, ou (c) não-lineares ou envolvidos em interações com outros parâmetros.

Neste método, os parâmetros são variados um de cada vez, sendo que cada parâmetro pode assumir um número discreto de valores (níveis) que são escolhidos dentro do intervalo de variação do parâmetro. Duas medidas de sensibilidade são propostas por Morris: a medida μ que estima o efeito médio do parâmetro na saída do modelo, e a medida σ que estima, através do desvio padrão, os efeitos de segunda ordem e ordens maiores nos quais o parâmetro está envolvido (inclusive curvaturas e interações).

Considerando o vetor de parâmetros $x = (x_1, x_2, \dots, x_k)$ correspondente a uma saída y do modelo, define-se, inicialmente, uma *região de experimentação*, ω , a qual corresponde a uma grade k -dimensional com p níveis, onde cada x_i pode assumir valores do seguinte conjunto: $\{0, 1/(p-1), 2/(p-2), \dots, 1\}$. Para um dado x , define-se uma medida que corresponde ao *efeito elementar* para o i -ésimo parâmetro através da equação:

$$d_i(x) = \frac{y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - y(x)}{\Delta} \quad (3.3)$$

onde $x \in \omega$, com a restrição $x_i \leq 1 - \Delta$, sendo Δ um valor pré-determinado como múltiplo de $1/(p-1)$.

Para cada parâmetro, calcula-se um número r de efeitos elementares d_i , em pontos selecionados aleatoriamente no intervalo do i -ésimo parâmetro (r pontos). A média μ_i e o desvio padrão σ_i dos r efeitos elementares correspondem, então, às medidas de importância do i -ésimo parâmetro. A medida μ_i é usada para detectar se o i -ésimo parâmetro tem alguma influência total na saída do modelo, e a medida σ_i é usada para detectar se o i -ésimo parâmetro tem efeitos não-lineares ou está envolvido em interações com outros parâmetros.

Em modelos não monotônicos, podem ocorrer efeitos elementares de sinais opostos que, no cálculo da média, podem se anular, produzindo um baixo valor de μ e levando a uma falha do método de Morris. Para corrigir tal problema, uma média do módulo dos valores elementares, μ^* , foi sugerida por [Saltelli et al. \(2005\)](#).

No cálculo dos r efeitos elementares de um parâmetro, há a necessidade de se fazer a avaliação de y duas vezes para cada efeito d_i . Assim, o custo computacional do

cálculo de r efeitos elementares para os k parâmetros é $n = 2kr$.

Morris propôs um método de cálculo mais eficiente que usa algumas simulações do modelo para computar mais que um efeito elementar. Tal método é baseado na construção de r trajetórias, definidas como a sucessão de $k + 1$ pontos no espaço dos parâmetros. A Figura 3.1 mostra dois exemplos de trajetórias em espaços de entrada dos parâmetros para $k = 2$ e $k = 3$, respectivamente. Cada trajetória fornece k efeitos elementares, um por parâmetro, e define uma *Matriz de Orientação* B^* , com dimensão $(k + 1) \times k$, cujas linhas são pontos X^1, X^2, \dots, X^{k+1} . O custo computacional dos r efeitos elementares dos k parâmetros é, neste caso, $n = (k + 1)r$ (VAREJÃO, 2009).

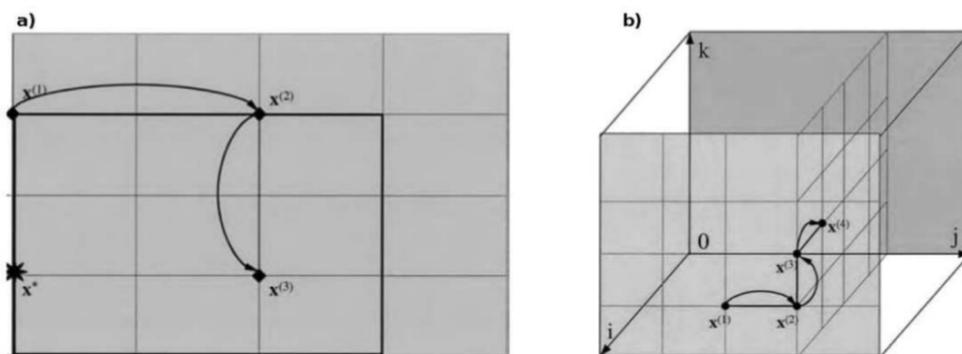


Figura 3.1 - Dois exemplos de trajetórias no espaço de parâmetros para (a) $k = 2$ e (b) $k = 3$.

Fonte: Varejão (2009).

A *Matriz de Orientação* B^* é construída através dos seguintes passos:

- 1 Defina uma matriz B , de dimensão $(k + 1) \times k$, como uma matriz triangular estritamente inferior de 1s.
- 2 Construa uma matriz D^* , de dimensão $k \times k$, diagonal na qual cada elemento é 1 ou -1 , com igual probabilidade.
- 3 Tome uma matriz P^* , de dimensão $k \times k$, de permutação aleatória, na qual cada coluna contém um elemento igual a 1 e todos os outros iguais a 0 e não existem duas colunas com 1s na mesma posição.

4 Construa a matriz de orientação aleatória B^* assim:

$$B^* = (J_{k+1,1}x^* + \frac{\Delta}{2}[(2B - J_{k+1,k})D^* + J_{k+1,k}])P^* \quad (3.4)$$

onde $J_{k+1,k}$, de dimensão $(k+1) \times k$, e $J_{k+1,1}$, de dimensão $(k+1) \times 1$, são matrizes de 1s e x^* é um valor base de X escolhido aleatoriamente.

A matriz B^* fornece um efeito elementar, aleatoriamente selecionado, por parâmetro.

3.3 Otimização Multiobjetivo

Muitos problemas do mundo real apresentam uma coleção de objetivos a serem considerados que são, na maioria das vezes, conflitantes entre si, isto é, a melhoria de algum(uns) objetivo (s) causa(m) a deterioração de outro(s). Podemos citar como um exemplo de um problema com objetivos conflitantes a tarefa de aquisição de um computador. A aquisição ótima é aquela que apresenta o custo mínimo enquanto maximiza o desempenho do equipamento. Tais objetivos são conflitantes entre si uma vez que existirão desde computadores com elevado custo e elevado desempenho até aqueles com baixo custo e baixo desempenho. Um computador pelo menor custo com o mais alto desempenho, embora ideal, não existe no mundo real. Assim, nenhuma solução que tenha menor custo e menor desempenho pode ser considerada melhor do que outra com maior custo e maior desempenho. Contudo, dentre todas as configurações de equipamentos, existem algumas que são superiores a outras, isto é, apresentam desempenho maior ou equivalente por um preço menor ou igual. Dizemos, então, que tais configurações (soluções) são não dominadas pelas demais, constituindo, assim, as chamadas soluções não dominadas do problema multiobjetivo em questão. Elas evidenciam o fato de que o problema não tem uma solução ótima única, mas, sim, um conjunto de soluções. Tais soluções são ótimas porque não existem outras soluções no espaço de busca melhores do que elas quando todos os objetivos são simultaneamente considerados, e são chamadas de soluções ótimas de Pareto.

3.3.1 Formalização do Problema Multiobjetivo

A formulação geral de um problema de otimização multiobjetivo pode ser assim expressa:

$$\text{Otimizar: } Z = (f_1(x), f_2(x), \dots, f_k(x)) \text{ para } k = 1, \dots, l, \text{ com } l \geq 2 \quad (3.5)$$

sujeito a:

$$g_j \leq 0, \text{ para } j = 1, \dots, q \quad (3.6)$$

$$h_j = 0, \text{ para } j = q + 1, \dots, m \quad (3.7)$$

$$v_i \leq x_i \leq u_i, \text{ para } i = 1, \dots, n \quad (3.8)$$

onde $x = (x_1, x_2, \dots, x_n)$.

O que se deseja ao resolver o problema é encontrar $x = (x_1, x_2, \dots, x_n)$ que satisfaça às restrições 3.6 (restrições de desigualdade), 3.7 (restrições de igualdade) e 3.8 (restrições laterais) e que otimize simultaneamente as funções objetivo de 3.5. As restrições acima estabelecem um espaço de soluções factíveis, $\Omega \subset R^n$, das variáveis de decisão $x = (x_1, x_2, \dots, x_n)$. Este espaço pode ser mapeado com o chamado espaço objetivo $Z = (f_1(x), f_2(x), \dots, f_k(x))$. A Figura 3.2 ilustra o mapeamento acima para um espaço das variáveis de decisão bidimensional e um espaço objetivo também bidimensional. A relação entre as soluções mapeadas no espaço objetivo irá guiar a escolha de pontos desejáveis para a solução do problema no espaço das variáveis de decisão.

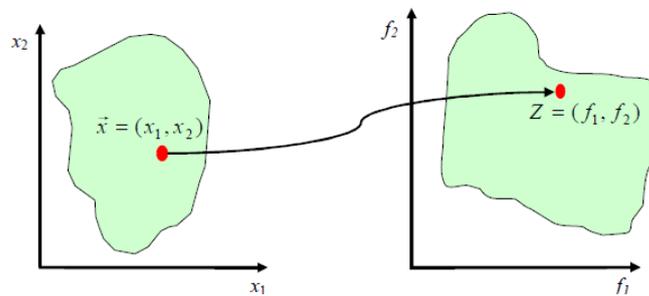


Figura 3.2 - Mapeamento do espaço de decisão no espaço objetivo.

Fonte: Adaptado de Zini (2009).

A principal característica de um problema de otimização multiobjetivo é que a sua solução, em geral, não é única. É comum existirem várias soluções tais que, ao se mudar de uma solução para outra, percebe-se a melhora obtida em uma função objetivo em detrimento de pelo menos uma das outras funções objetivo. É preciso, pois, saber identificar dentre as soluções aquelas que são eficientes para a resolução do problema. Para tal, torna-se importante o conceito de dominância, já mencionado

na introdução desta seção e que será formalizado a seguir (ZINI, 2009):

Definição1: Dizemos que uma solução \tilde{x} domina outra solução \bar{x} quando as duas condições seguintes são satisfeitas:

- a) a solução \tilde{x} não é pior que \bar{x} em todos os objetivos;
- b) a solução \tilde{x} é estritamente melhor do que \bar{x} em pelo menos um objetivo.

Considerando que o problema seja de minimização de todas as funções objetivo, a definição acima pode ser matematicamente formalizada assim:

Definição2: Dizemos que uma solução \tilde{x} domina outra solução \bar{x} se para todo $k \in \{1, 2, \dots, l\}$, tal que $f_k(\tilde{x}) \leq f_k(\bar{x})$, existe $j \in \{1, 2, \dots, l\}$ tal que $f_j(\tilde{x}) < f_j(\bar{x})$.

Quando uma solução \tilde{x} domina outra solução \bar{x} , o mais comum é dizer que \tilde{x} é não dominada por \bar{x} .

A Figura 3.3 a seguir ilustra, no espaço objetivo, a definição de dominância, considerando um problema de minimização de duas funções (ZINI, 2009).

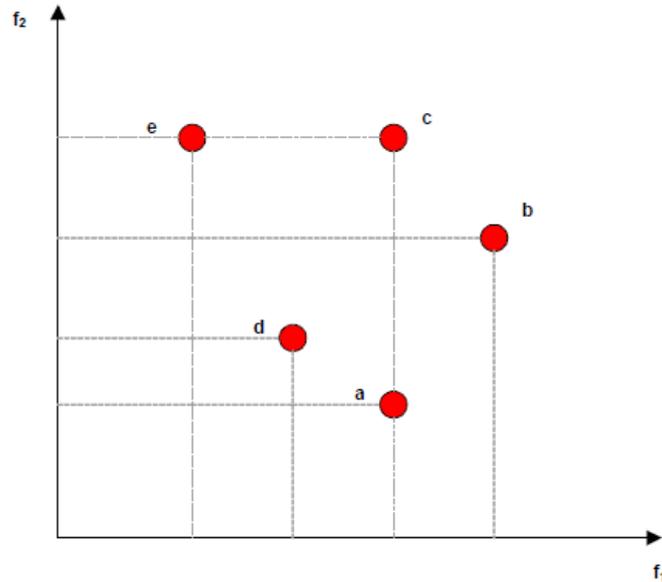


Figura 3.3 - Dominância em uma minimização de duas funções.

Fonte: Adaptado de Zini (2009).

Pelas definições de dominância, a solução a domina as soluções b e c , as soluções a e d são não dominadas entre si e o conjunto $\{a, d, e\}$ é um conjunto de soluções não dominadas que dominam as soluções b e c , ou seja, pelo menos um elemento do conjunto $\{a, d, e\}$ domina os elementos b e c .

De modo geral, dado um conjunto de soluções P do problema de otimização, o conjunto das soluções não dominadas P' é o conjunto que contém todas as soluções não dominadas por qualquer elemento do conjunto P . Assim quaisquer duas soluções de P' são não dominadas entre si, e qualquer solução das demais do conjunto P são dominadas por pelo menos um elemento de P' . Quando o conjunto P é o próprio espaço de busca, então o conjunto P' é chamado de *conjunto de Pareto* (ZINI, 2009).

A Figura 3.4 ilustra um problema simples de minimizar as funções $g(x) = x^2$ e $h(x) = (x-2)^2$ em relação a um parâmetro real x (SCHAFFER, 1985). As soluções do problema consistem de todos os pontos no intervalo $[0, 2]$ no espaço dos parâmetros, no qual se pode observar que, para $0 \leq x < 1$, tem-se $h(x) > g(x)$, ao passo que para $1 < x \leq 2$, tem-se $g(x) > h(x)$. Fica claro, portanto, que o intervalo $[0, 2]$ corresponde ao conjunto de soluções não dominadas do problema, ou seja, é o conjunto de Pareto.

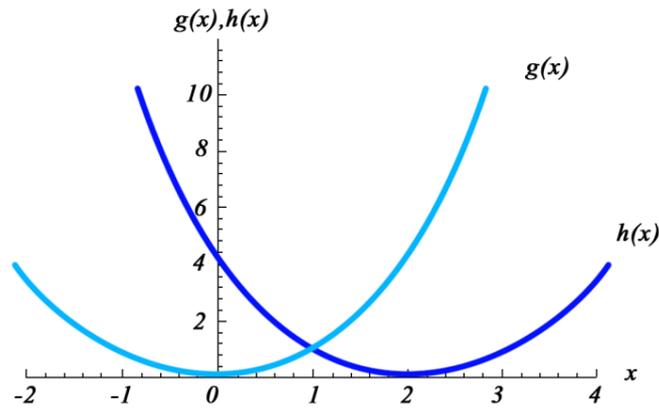


Figura 3.4 - Minimização de duas funções exemplos g e h .

A formulação seguinte, feita por Yapo et al. (1998), estabelece a divisão do espaço das variáveis de decisão em duas partes: o das soluções não dominadas (conjunto de Pareto) e o das soluções dominadas, considerando a minimização da função $Z = (f_1(x), f_2(x), \dots, f_k(x))$. Por definição, toda solução x_P pertencente ao conjunto de Pareto deve satisfazer às seguintes propriedades:

- i) $f_j(x_P) < f_j(x_D)$, para todo $j = 1, 2, \dots, k$, para toda solução x_D não contida no conjunto de Pareto;
- ii) não existe uma solução x_P^* no conjunto de Pareto tal que $f_j(x_P^*) < f_j(x_P)$, para todo $j = 1, 2, \dots, k$.

O espaço factível de soluções fica, então, dividido em dois conjuntos, um de “boas” soluções (soluções Pareto) e um de soluções “ruins”, sendo impossível distinguir a melhor entre as soluções “boas”. O conjunto das soluções não dominadas é, portanto, o alvo da otimização multiobjetivo. Este conjunto, quando mapeado no espaço objetivo, forma uma superfície conhecida como *frente de Pareto*, ou seja, a frente de Pareto é o resultado (no espaço objetivo) do mapeamento do conjunto de Pareto.

Na Figura 3.5, é ilustrada a frente de Pareto para as funções g e h do exemplo anterior. Observa-se que, movendo-se ao longo da frente de Pareto, a melhora em função objetivo é acompanhada por uma piora na outra função. Neste caso, por se tratar de um problema de apenas duas funções objetivo, foi fácil identificar a frente de Pareto. Para problemas mais complexos, no entanto, não é possível fazer uma análise gráfica, sendo necessário o uso de técnicas computacionais para a obtenção de uma aproximação da frente de Pareto. É sempre desejável uma aproximação que permita identificar o maior número de soluções não dominadas.

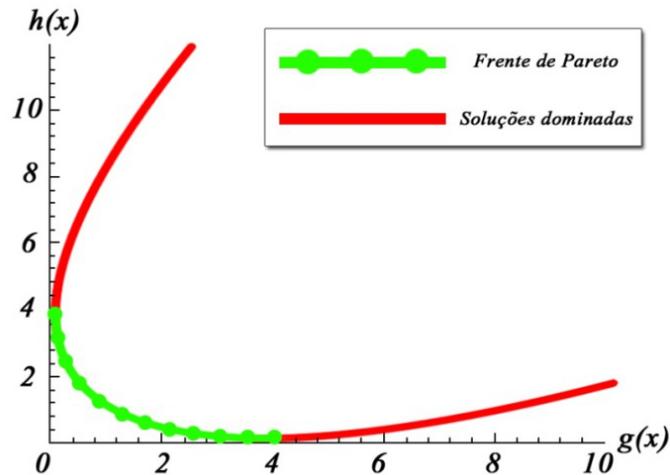


Figura 3.5 - Espaço objetivo para as funções exemplo g e h .

3.3.2 Abordagens e Técnicas para a Otimização Multiobjetivo

Os exemplos simples já considerados neste capítulo evidenciam que além da quantidade de funções objetivo, uma diferença fundamental entre a otimização mono objetivo e a otimização multiobjetivo está na quantidade de soluções ótimas obtidas. Do ponto de vista prático, no entanto, ao se resolver um problema de otimização, necessita-se, na maioria das vezes, de uma única solução como resposta, independente do tipo de otimização adotado. Assim sendo, duas abordagens podem ser consideradas na otimização multiobjetivo (CASTRO, 2001)

Abordagem 1: Pesos e prioridades são estabelecidos entre os vários objetivos de interesse, transformando o problema em um problema de otimização mono objetivo, e procura-se por uma solução única do problema;

Abordagem 2: Sem o uso de informações adicionais, procura-se encontrar o conjunto de soluções não dominadas para, posteriormente, escolher-se uma entre as soluções desse conjunto.

Na primeira abordagem, é importante ressaltar que, geralmente, pouco se conhece do problema real que dá origem ao problema de otimização a ser resolvido, o que dificulta o estabelecimento do nível relativo de importância dos objetivos e, conseqüentemente, a definição das prioridades e dos pesos dos mesmos. Sendo o problema reduzido a um problema de otimização mono objetivo, existem vários algoritmos disponíveis para esta abordagem que podem ser baseados em métodos clássicos, algoritmos de busca busca aleatória guiada como, por exemplo, o Recozimento Si-

mulado, ou algoritmos evolucionários.

Dos muitos métodos clássicos disponíveis na literatura para tratar de problemas de otimização multiobjetivo, há alguns bastante consagrados que trabalham de acordo com a primeira abordagem. Vale ressaltar que o termo *métodos clássicos* é geralmente usado para distingui-los dos algoritmos evolucionários (ZINI, 2009). Em (COELLO, 1996; DEB, 2004; ZINI, 2009) são apresentados e bem descritos, por exemplo, os seguintes métodos clássicos: método da soma ponderada, método da restrição- ϵ , métodos de programação por metas (ponderado, lexicográfico e mini-max).

Deve-se dizer que os métodos clássicos apresentam como vantagem principal a existência de provas que garantam a convergência para as soluções de Pareto, e como grande desvantagem o fato de que para transformarem um problema multiobjetivo em um problema mono objetivo, são introduzidos parâmetros adicionais que afetam diretamente nos resultados (ZINI, 2009).

Para a segunda abordagem, os algoritmos evolucionários têm tido a preferência dos pesquisadores. Os algoritmos genéticos, em particular, têm se mostrado bastante eficientes para esta abordagem e vêm ganhando cada vez mais espaço nas últimas décadas (FONSECA; FLEMING, 1993; SRINIVAS; DEB, 1994; DEB et al., 2000; DEB, 2004).

3.3.2.1 Algoritmos Genéticos

Um Algoritmo Genético (AG) (HOLLAND, 1975) é uma metaheurística inspirada na Teoria da Evolução de Darwin que simula o processo de seleção natural e a sobrevivência dos indivíduos mais aptos de uma população.

Em um problema de otimização, um AG tenta encontrar uma boa solução, gerando, inicialmente, uma população aleatória, constituída de soluções viáveis para o problema, e manipulando essas soluções através de operadores matemáticos. Tais operadores usam as soluções já existentes para produzir novas soluções que, espera-se, sejam melhores que as anteriores. Diz-se que uma população de soluções é gerada, sendo cada solução identificada como um cromossomo (ou indivíduo da população). Para cada indivíduo, determina-se um valor de aptidão escalar que representa uma medida numérica da sua capacidade em bem resolver o problema. A ideia principal é selecionar para reprodução as soluções com maior aptidão e aplicar sobre elas operações que gerem melhores soluções para o problema.

Os entes básicos de um algoritmo têm a sua denominação associada a termos da

Biologia. O conjunto de pontos a partir dos quais se deseja obter a melhor solução do problema é chamado de população, sendo tais pontos chamados de indivíduos ou cromossomos. Cada indivíduo possui uma unidade básica, denominada gene, que descreve uma certa variável do problema. Cada iteração do algoritmo é denominada geração. As operações usadas sobre os indivíduos da população para gerar indivíduos mais aptos, ou seja, melhores soluções para o problema são chamadas de operadores genéticos.

A representação ou codificação das variáveis em um AG pode ser binária, ou com números inteiros ou reais. A geração da população pode ser feita de forma aleatória ou usando alguma heurística de construção. A avaliação da população é feita usando uma ou mais funções de aptidão para avaliar a qualidade das soluções (no caso de uma otimização multiobjetivo, serão usadas as funções objetivo). A seleção dos indivíduos é um mecanismo usado para, depois de avaliados os indivíduos, permitir a hereditariedade entre as gerações e, com isso, preservar as boas características dos indivíduos. Ela pode ser feita aplicando os seguintes métodos: da roleta, do torneio e do ranking (GOLDBERG, 1989). Os operadores genéticos básicos são os de cruzamento e mutação.

O operador de cruzamento faz alterações ou combinações de sequências de informações entre indivíduos da população, gerando novos indivíduos que herdam características dos indivíduos anteriores. Para manter a diversidade das soluções, os indivíduos com características muito semelhantes, que possivelmente tenham sido gerados por este operador, devem ser eliminados. As técnicas de elaboração do operador de cruzamento levam em conta a representação computacional das variáveis com a qual se trabalha. As seguintes técnicas, por exemplo, podem ser mencionadas: cruzamento de um ponto, cruzamento multipontos, cruzamento uniforme.

O operador de mutação consiste em alterar aleatoriamente o valor de um ou mais genes de um cromossomo. Ele não aumenta o tamanho da população, somente modificando os indivíduos existentes e transformando-os em indivíduos diferentes. Caso o resultado da mutação de um indivíduo seja um indivíduo que já existe, ele é descartado. O operador de mutação permite uma maior diversidade genética e uma maior exploração do espaço de busca e evita que o algoritmo fique estacionado em mínimos locais.

Com o objetivo aumentar a velocidade de convergência do algoritmo é importante preservar e utilizar em próximas gerações as melhores soluções encontradas em uma determinada geração, o que acarreta em elitismo. Um dos aspectos mais relevantes

na estruturação de um AG é o controle de seus parâmetros: tamanho da população, taxa de cruzamento e taxa de mutação.

Além dos operadores genéticos acima mencionados, um novo operador pode ser considerado. Trata-se do Operador Epidêmico, ou, simplesmente, Epidemia (MEDEIROS, 2002; CHIWIAKOWSKY; VELHO, 2003). Tal operador é ativado sempre que seja alcançado um número pré-fixado de gerações sem que se consiga obter uma melhora desejada na população. Ele pode também ser usado para evitar ou solucionar um problema de convergência prematura. Tal operador simula a ocorrência de uma epidemia que recai sobre a população, dizimando os indivíduos menos aptos, de modo que aqueles que apresentam melhores aptidões são preservados. Os indivíduos eliminados são, então, substituídos por novos elementos, gerados pelo mesmo processo de geração da população inicial, e o processo de evolução é reiniciado.

Há uma vasta literatura sobre a teoria de algoritmos genéticos. Para uma introdução a esta teoria, bem como um bom entendimento dos vários aspectos técnicos relacionados à representação das variáveis em um AG e aos operadores genéticos, além das referências mencionadas acima, (GOLDBERG, 1989; HOLLAND, 1975), pode-se recomendar, por exemplo, os seguintes livros: Introduction to Genetic Algorithms (SIVANANDAM; DEEPA, 2008) e Genetic Algorithms + Data Structures (MICHALEWICZ, 1996).

A seguir, são apresentadas algumas vantagens e desvantagens na utilização dos algoritmos genéticos em problemas de otimização, conforme (CASTRO, 2001). Deve-se dizer que as vantagens mencionadas a seguir são próprias das metaheurísticas de modo geral.

Vantagens:

- São robustos e aplicáveis a uma grande variedade de problemas;
- Não requerem conhecimento ou informações do gradiente da superfície definida pela função objetivo, o que torna os algoritmos genéticos muito adequados para funções com descontinuidades ou para as quais seja impossível o cálculo das derivadas. Descontinuidades ou complexidades na superfície pouco ou nada afetam no seu desempenho;
- Apresentam um bom desempenho para uma grande variedade de problemas, já que não impõem muitas das limitações que são encontradas nas técnicas de busca tradicionais;

- São mais resistentes a se prenderem a ótimos locais (o que ocorre com determinados métodos clássicos);
- São de fácil implementação e proporcionam maior flexibilidade no tratamento do problema a ser resolvido.

Desvantagens:

- Dificuldade de encontrar o ótimo global exato;
- Podem requerer um grande número de avaliações da função de aptidão;
- Grandes possibilidades de configurações que podem complicar a resolução do problema.

Deve-se ressaltar que o avanço das capacidades de computação e as constantes melhorias na técnica vêm atenuando as desvantagens mencionadas acima.

Uma variedade de algoritmos genéticos multiobjetivos pode ser encontrada na literatura especializada. Entre os principais, podem ser citados, por exemplo, os seguintes (DEB, 2004):

- VEGA (Vector Evaluated Genetic Algorithm): implementado por Schaffer (1985), foi o primeiro algoritmo genético multiobjetivo feito para encontrar um conjunto de soluções não dominadas;
- MOGA (Multiple Objective Genetic Algorithm): implementado por Fonseca e Fleming (1993);
- NPGA (Niche-Pareto Genetic Algorithm), proposto por Horn et al. (1994);
- NSGA (Non-Dominated Sorting Genetic Algorithm), implementado por Srinivas e Deb (1994);
- NSGA-II (Elitist Non-Dominated Sorting Genetic Algorithm) (DEB et al., 2000) é uma melhora do NSGA.

Estes algoritmos podem ser encontrados em (DEB, 2004).

Uma característica muito comum nos algoritmos de otimização multiobjetivo evolucionários é o uso da estratégia do Ordenamento de Pareto. Tal estratégia, proposta

por Goldberg (1989), consiste em classificar os indivíduos da população de acordo com a relação de dominância entre eles, dividindo a população em M subconjuntos, $F_m, m = 1, \dots, M$, denominados *frentes de dominância*. Em cada frente de dominância F_j , nenhuma solução domina outra solução de F_j , e todas as soluções de F_j dominam as soluções de F_l para $l > j$. Os indivíduos da primeira frente, F_1 , correspondem às melhores soluções para a geração atual, enquanto os indivíduos da última frente correspondem às piores, podendo, portanto, considerar-se esta primeira frente como uma aproximação, na geração atual, para a frente de Pareto. A Figura 3.6 abaixo ilustra o ordenamento de Pareto para um caso de minimização de duas funções objetivo.

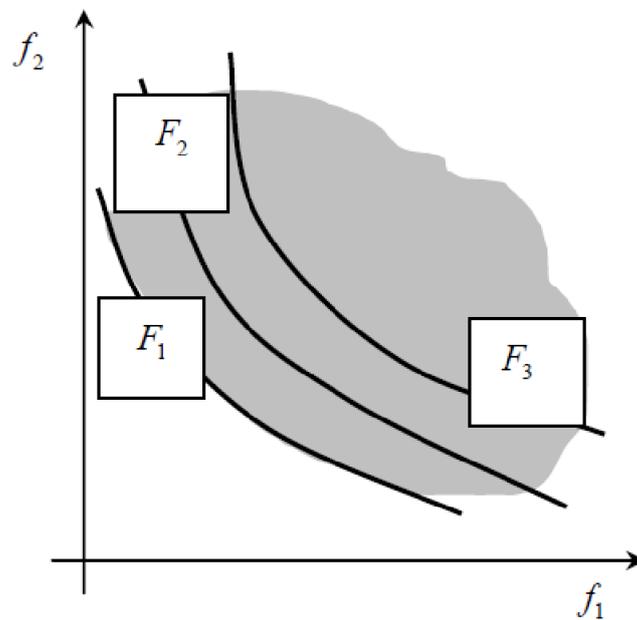


Figura 3.6 - Ordenamento de Pareto no Espaço Objetivo.

Fonte: Lopes (2010).

3.4 Um Algoritmo Genético Epidêmico para a Calibração Multiobjetivo do IPH-II

Este tópico é dedicado à apresentação de um algoritmo genético epidêmico que foi implementado, neste trabalho de tese, para a calibração multiobjetivo do modelo hidrológico IPH-II. Tal algoritmo, denotado por AGE, foi desenvolvido como uma variante do algoritmo evolucionário MOCOM (YAPO et al., 1998), o qual já vem sendo

usado na calibração do modelo hidrológico IPH-II. O MOCOM combina técnicas evolucionárias com o algoritmo simplex de [Nelder e Mead \(1965\)](#). A ideia central na elaboração do AGE foi substituir o processo de evolução do MOCOM por uma evolução genética (com o uso de operadores genéticos).

Os passos do AGE são apresentados a seguir, sendo os cinco primeiros passos e os dois últimos idênticos aos do MOCOM.

3.4.1 Passos do AGE

- Primeiro Passo: Inicialmente, são estabelecidos os limites máximos e mínimos dos valores que os k parâmetros a serem ajustados podem assumir, definindo, assim, uma região factível para os vetores de parâmetros.
- Segundo Passo: Usando uma distribuição uniforme, são gerados ns pontos com k coordenadas na região factível de parâmetros do modelo. Cada um desses pontos é um indivíduo da população inicial, a partir da qual se buscará o conjunto de soluções não dominadas do problema de otimização multiobjetivo. Portanto ns é o parâmetro do algoritmo que representa o tamanho da população.
- Terceiro Passo: Os pontos da população são avaliados com as nf funções objetivo, gerando uma matriz de resultados $R(ns, nf)$.
- Quarto Passo: Aplica-se o seguinte Ordenamento de Pareto: Identificam-se todos os indivíduos da população que são não dominados, e a esses indivíduos atribui-se o ranking 1. Estes indivíduos com ranking igual a 1 são, temporariamente, retirados da população. Em seguida, na população restante, são identificados os indivíduos não dominados, e a esses indivíduos atribui-se o ranking 2. Retiram-se, então, temporariamente, esses indivíduos com ranking 2 dessa população. Estes passos se repetem até que a cada indivíduo da população tenha sido atribuído algum ranking. Os piores indivíduos da população são os que se encontram mais distantes da frente de Pareto, possuindo o maior valor de ranking, o qual é denotado por R_{max} .

Desta forma, são atribuídos, para cada ponto i da população, valores de ranking r_i , que variam de 1 a R_{max} , sendo $R_{max} \leq ns$. Os melhores indivíduos da população inicial (não dominados) possuem ranking igual a 1 e os piores possuem ranking igual a R_{max} . Com este ordenamento, são criadas,

no espaço objetivo, várias frentes de dominância que indicam o nível de dominância de uma solução em relação às demais.

- Quinto Passo: Criação de Complexos: Feito o ordenamento de Pareto, cada indivíduo com ranking R_{max} dará origem a um conjunto denominado *Complexo*. Um *Complexo* é formado por um indivíduo com ranking R_{max} e mais k (número de parâmetros do modelo) outros indivíduos selecionados aleatoriamente entre os indivíduos da população que não possuem ranking R_{max} , usando a seguinte equação de probabilidade associada a cada ponto i da população:

$$p_i = \frac{R_{max} - r_i + 1}{\sum_{j=1}^{ns} (R_{max} - r_j + 1)} . \quad (3.9)$$

A probabilidade de um indivíduo ser selecionado depende de seu valor de ranking, sendo favorecidos os indivíduos de menor ranking (melhores). Cada *Complexo* possui, portanto, $k + 1$ elementos. Os *Complexos* podem possuir elementos comuns.

- Sexto Passo: Aqui é feita a evolução dos *Complexos* com o objetivo de fazer com que os pontos da população converjam para a frente de Pareto. Neste algoritmo é proposta uma evolução genética aplicada a cada *Complexo*.

Dado um *Complexo*, aplicam-se operadores genéticos de mutação e cruzamento em seus indivíduos, gerando-se $k + 1$ novos elementos (filhos) no *Complexo* até que, considerado um ordenamento de Pareto (descrito no quarto passo) feito no *Complexo*, o indivíduo de maior ranking tenha um ranking menor do que o ranking máximo no *Complexo* antes das operações. Quando isto ocorre, este novo indivíduo de maior ranking substitui o indivíduo de ranking máximo do *Complexo* e diz-se, neste caso, que o *Complexo* evoluiu. Colocou-se, aqui, um parâmetro que estabelece o número máximo de tentativas de gerações em cada *Complexo* até que se alcance sua evolução. Aqui, começou a ser considerada a aplicação de um operador de epidemia. Para tal, foi introduzido um parâmetro no intervalo $(0, 1)$ de modo que, se o percentual de *Complexos* sem evolução alcança o valor deste parâmetro, ativa-se uma epidemia. E esta epidemia é aplicada em todos os indivíduos da população cujos valores de ranking sejam maiores ou iguais a um determinado valor pré-fixado (que também é um parâmetro do algoritmo), isto é, todos os indivíduos com valores de ranking maiores ou iguais a esse valor pré-fixado são eliminados da população e são gerados, do mesmo modo que a população inicial, outros tantos indivíduos

que irão substituí-los, voltando-se, então, ao terceiro passo do algoritmo. A aplicação da epidemia também foi limitada a um número máximo de iterações (outro parâmetro do algoritmo).

- Sétimo Passo: Feita a evolução de todos os *Complexos*, todos os seus indivíduos são devolvidos à população.
- Oitavo Passo: Teste de Convergência: se o critério de parada é satisfeito, o algoritmo para. O critério de parada para o algoritmo finaliza o processo quando todos os indivíduos da população são não dominados, isto é, quando $R_{max} = 1$, ou quando um número máximo de iterações é atingido.

3.4.2 Evolução dos Complexos no MOCOM

No algoritmo MOCOM, o procedimento de evolução usado para melhorar o pior ponto no complexo, isto é, o ponto com ranking igual a R_{max} , é uma extensão multiobjetivo do método simplex de descida, denominado MOSIM (NELDER; MEAD, 1965). Este método gera um novo ponto que substitui o ponto do complexo com ranking igual a R_{max} , a partir de duas operações, *reflexão* e *contração*, sobre complexo conforme o seguinte procedimento.

Considera-se a seguinte equação: $C_n = \gamma.C_g + (1 - \gamma).C_w$, onde C_n é o novo ponto, C_w é o ponto de pior ranking do complexo e C_g é a localização do centroide dos k pontos melhores ranqueados do complexo. Quando $\gamma = 2$, obtém-se um ponto de reflexão, C_{ref} , e quando $\gamma = 0,5$, um ponto de contração, C_{con} , é obtido.

A Figura 3.7 traz uma ilustração geométrica dos passos de contração e reflexão para um caso simples em que o *Complexo* é formado por três pontos x_1 , x_2 e x_3 (os índices dos pontos correspondem ao seu valor de ranking).

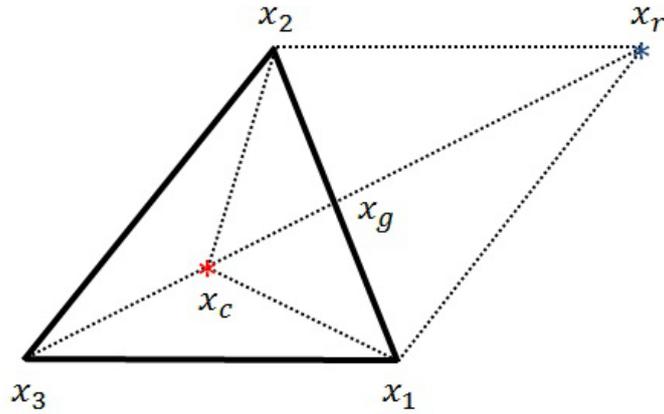


Figura 3.7 - Passos de reflexão e contração em um *Complexo* no MOCOM.

A regra para escolher, entre C_{ref} e C_{con} , o ponto que irá substituir o ponto de ranking igual a R_{max} do *Complexo* é baseada no conceito de dominância. O ponto C_{ref} é aceito se, e somente se, ele é não dominado em relação aos demais k pontos do complexo que foram usados para calcular o centróide. Caso o ponto C_{ref} seja dominado, ele é rejeitado e o ponto C_{con} é, então, aceito.

3.5 O Algoritmo NSGA-II

O NSGA-II (Elitist Non-dominated Sorted Genetic Algorithm II) é um algoritmo genético multiobjetivo proposto por Deb et al. (2000) como uma versão melhorada do NSGA (Non-dominated Sorted Genetic Algorithm) (SRINIVAS; DEB, 1994).

No NSGA-II, o Ordenamento de Pareto é feito de uma forma mais rápida (*Fast Non-dominated Sorting*) em duas etapas.

Na primeira etapa, os elementos da população P são comparados uns com os outros pelos valores das funções objetivo, sendo, então, classificados de acordo com o seu grau de dominância. O grau de dominância de um indivíduo p da população, denotado por np , corresponde ao número de indivíduos que dominam p . Ao final desta etapa, na primeira frente de dominância, estarão todos os elementos cujo grau de dominância é $np = 0$, e que serão os melhores indivíduos de toda a população.

Na segunda etapa, os indivíduos serão separados em diferentes *frentes de dominância* de acordo com os seus respectivos graus de dominância. Cada indivíduo p em uma frente é retirado totalmente do contexto do sistema, decrementando os valores dos graus de dominância dos indivíduos que são dominados por p . Isto se repete até que não sobrem mais indivíduos na população restante.

Assim, ao final do Ordenamento de Pareto, a população é classificada em M subconjuntos de frentes de dominância, $F_m, m = 1, \dots, M$, ou seja, a população é dividida em grupos de indivíduos de acordo com o grau de dominância de cada indivíduo. Em cada subconjunto, nenhuma solução domina outra. Os indivíduos da primeira frente correspondem às melhores soluções para a geração atual, enquanto os indivíduos da última frente correspondem às piores. Portanto na primeira frente estão os indivíduos que não são dominados por nenhum outro indivíduo da geração, e estão próximos à frente de Pareto.

Para estimar a densidade de soluções ao redor de um dado ponto propôs-se um operador de diversidade, denominado *crowding distance* (distância de aglomeração), usado para garantir um espalhamento uniforme das soluções ao longo da frente de Pareto. Este operador ordenará cada ponto de uma frente de acordo com uma distância média, em relação às funções objetivo, entre o ponto e seus vizinhos na mesma frente. Às soluções extremas da frente é atribuído um valor arbitrariamente grande para essa distância. Para as soluções intermediárias a distância de aglomeração corresponde à distância entre um indivíduo central i e dois indivíduos adjacentes, $i - 1$ e $i + 1$, localizados em cada lado do indivíduo central (Figura 3.8). A ideia é dar prioridade aos pontos mais distantes durante o processo de seleção a fim de espalhar os resultados ao longo da frente de Pareto. Sendo assim, uma solução com a melhor classificação numa frente é aquela que apresenta um maior valor desta medida de distância (LOPES, 2010).

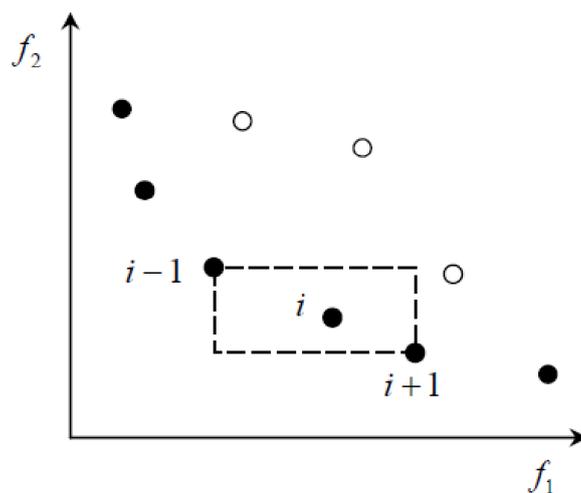


Figura 3.8 - Cálculo das distâncias de aglomeração nas soluções intermediárias.

Fonte: Adaptado de Deb et al. (2000).

Considerando m funções objetivo f_1, \dots, f_m , a *distância de aglomeração* no i -ésimo ponto de uma frente, d_i , é obtida com a seguinte fórmula:

$$d_i = \frac{f_1(i+1) - f_1(i-1)}{f_1^{max} - f_1^{min}} + \dots + \frac{f_m(i+1) - f_m(i-1)}{f_m^{max} - f_m^{min}} \quad (3.10)$$

onde f_j^{max} e f_j^{min} correspondem, respectivamente, aos valores máximo e mínimo da j -ésima função objetivo, em todo o espaço de soluções; $f_j(i+1)$ e $f_j(i-1)$ correspondem aos valores das funções objetivo nos pontos $i+1$ e $i-1$, respectivamente, para $j = 1, \dots, m$.

A distância de aglomeração d_i corresponde ao semiperímetro do hipercubo formado pelas soluções vizinhas da solução i na mesma frente de dominância.

A Figura 3.9 mostra como o NSGA-II funciona. Inicialmente, uma população $R_t = P_t \cup Q_t$ é formada, onde P_t é a população pai, de tamanho N , e Q_t é a população filha, também de tamanho N . Esta nova população R_t , de tamanho $2N$, será ordenada de acordo com o ordenamento de Pareto, sendo escolhidos desta os melhores indivíduos para formar a nova população P_{t+1} (geração seguinte da população P_t). Esta nova população P_{t+1} obtida é usada para formar uma outra população Q_{t+1} por meio dos operadores genéticos de seleção, cruzamento e mutação (LOPES, 2010).

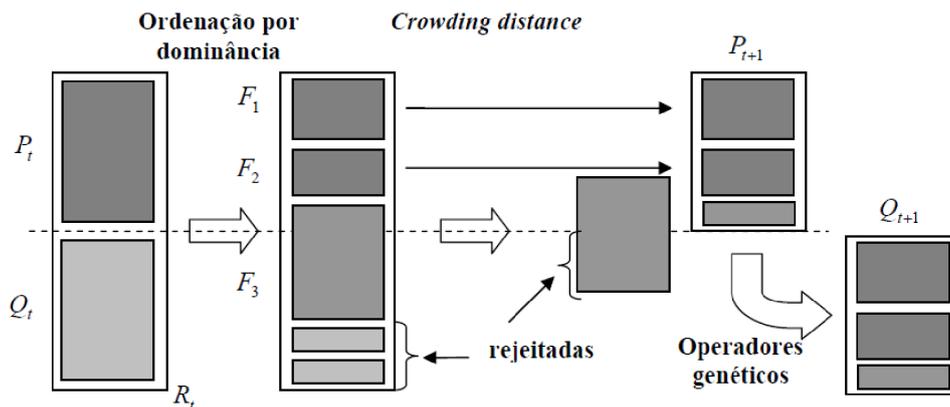


Figura 3.9 - Esquema do Algoritmo NSGA-II.

Fonte: Lopes (2010).

De acordo com o ordenamento de Pareto, o conjunto F_1 conterá as melhores soluções

da população combinada. Quando o tamanho de F_1 é menor do que N , todos os indivíduos deste conjunto são escolhidos para formar a nova população P_{t+1} , seguidos pelas frentes subsequentes dadas pela ordem de dominância (por exemplo, F_2 e F_3), até completar os N indivíduos da população. Caso o tamanho de F_1 seja maior do que N , as melhores soluções são escolhidas baseadas em um sorteio, usando um *operador de comparação de aglomeração* (*crowding comparison operator*). Este operador necessita tanto da ordenação de Pareto quanto do valor da *distância de aglomeração* de cada solução na população. Entre duas soluções, a solução escolhida é aquela que apresenta o menor valor (melhor) de ordenação. No caso das duas soluções estarem em uma mesma frente, a solução escolhida é aquela que apresenta o maior valor da *distância de aglomeração*, ou seja, a que está localizada na região menos densa de soluções (LOPES, 2010).

3.5.1 Um Operador de Epidemia para o NSGA-II

Como visto nesta seção, o algoritmo NSGA-II é fortemente baseado no ordenamento de Pareto. Depois de estabelecidas as várias frentes de dominância e calculados os valores de *crowding distance* em todos os indivíduos de cada frente, a evolução da população inicial, de tamanho N , se dá com a criação, a cada iteração, de uma população de descendentes de mesmo tamanho através das operações genéticas de mutação e cruzamento. Esta população de descendentes é, então juntada à população inicial, formando uma nova população. Nesta nova população, de tamanho $2N$, aplica-se o ordenamento de Pareto e são escolhidos para formar a população final os N indivíduos pertencentes às melhores frentes. Ao fim de um dado número de iterações do algoritmo, a frente de dominância F_1 (indivíduos que não são dominados por nenhum outro indivíduo da população) constituirá uma aproximação da frente de Pareto. O ideal é que, ao fim da evolução, a frente de dominância F_1 tenha tantos elementos quanto a população inicial, ou seja, F_1 será a única frente de dominância. Da forma como está implementado o NSGA-II, o único critério de parada do algoritmo é o número de iterações. Assim, findada a evolução, não há a garantia de que se tenha obtido uma aproximação da frente de Pareto com uma única frente de dominância (mesmo tamanho da população inicial).

Aqui, apresentamos uma estratégia de Epidemia que foi implementada no NSGA-II com o intuito de se tentar uma possível melhora na obtenção de uma aproximação da frente de Pareto. A ideia é estabelecer o objetivo de se chegar a uma aproximação da frente de Pareto com um número que corresponda a um certo percentual da população inicial.

O Operador de Epidemia pode ser assim descrito: findado o número de iterações, se o número de indivíduos da frente de dominância F_1 for menor 40% da população inicial, todos os indivíduos das outras frentes de dominância são eliminados e são substituídos por novos indivíduos, criados pelo mesmo processo de criação da população inicial.

3.6 Métricas de Avaliação de Erros: Funções Objetivo

Como mencionado na introdução deste trabalho, a calibração dos parâmetros de um modelo vai depender de métricas (estatísticas) que avaliem os erros entre os valores de entrada e de saída do modelo.

Assim para uma variável de saída (Y) que representa um processo simulado pelo modelo, teremos uma função de erro, ϵ , na dependência de um vetor θ de parâmetros a serem ajustados:

$$\epsilon(\theta)_t = Y_t - \bar{Y}_t, \quad t = 1, 2, \dots, NT \quad (3.11)$$

onde Y_t e \bar{Y}_t representam os valores observados e os valores estimados pelo modelo, respectivamente, em um número NT de intervalos de tempo.

Tais erros podem ser, então, avaliados através de funções $F(\epsilon(\theta))$, chamadas de *Funções Objetivo*, que estabelecem uma medida de desempenho dos modelos. A tarefa de ajustar os parâmetros do modelo passa a ser, então, a de encontrar o vetor de parâmetros θ que minimize uma ou mais funções objetivo.

Para os modelos hidrológicos do tipo chuva-vazão, como o IPH-II, em que a variável de saída (única) é a vazão (Q), temos, por exemplo, as seguintes funções objetivo muito comumente usadas (BRAVO et al., 2009):

$$F_1 = \sqrt{\frac{\sum_{t=1}^{NT} (Q_t - \bar{Q}_t)^2}{NT}} \quad (\text{Desvio padrão das vazões}) \quad (3.12)$$

$$F_2 = \sqrt{\frac{\sum_{t=1}^{NT} (\frac{1}{Q_t} - \frac{1}{\bar{Q}_t})^2}{NT}} \quad (\text{Desvio padrão do inverso das vazões}) \quad (3.13)$$

$$F_3 = \frac{\sum_{t=1}^{NT} |Q_t - \bar{Q}_t|}{NT} \quad (\text{Desvio absoluto médio das vazões}) \quad (3.14)$$

$$F_4 = 1 - \frac{\sum_{t=1}^{NT} (Q_t - \bar{Q}_t)^2}{\sum_{t=1}^{NT} (Q_t - \tilde{Q})^2} \quad (\text{Coeficiente de Nash e Sutcliffe}) \quad (3.15)$$

onde \tilde{Q} é a média da vazões observadas.

A função objetivo F_1 dá maior peso a erros de pico (cheias), enquanto a função F_2 , por exemplo, dá maior ênfase a erros de vale (estiagem).

No caso de modelos de fluidos geofísicos mais complexos, como o IBIS, por exemplo, ainda não há uma literatura bem estabelecida a respeito de funções objetivo comumente usadas. Varejão (2009) considerou além das funções F_1 e F_3 (raiz quadrada do erro quadrático médio e erro absoluto médio) as duas seguintes estatísticas:

$$B = \sum_{i=1}^n (P_i - O_i) \quad (3.16)$$

que avalia o viés um um período completo de simulação,

$$B_{max} = \text{Max}_{j=1}^n \sum_{i=1}^j (P_i - O_i) \quad (3.17)$$

que dá o máximo valor do viés durante todo o período de simulação.

3.7 Estratégia OPTIS

Com o objetivo de se fazer a calibração multiobjetivo automática do IBIS, Varejão (2009) desenvolveu um software em linguagem *C*, chamado Optis, no qual ao modelo IBIS é acoplado o algoritmo NSGA-II. No Optis, durante o processo de calibração, o modelo IBIS faz as simulações de forma independente do processo de otimização pelo NSGA-II. O NSGA-II interage com o modelo apenas na leitura dos dados de saída do IBIS e mudando os valores dos parâmetros de entrada para o IBIS. Desta forma, o Optis, praticamente, não depende da versão a ser considerada do IBIS ou até mesmo do modelo a ser calibrado. Pequenas alterações foram necessárias nos protocolos de leitura dos parâmetros de entrada do IBIS para simplificar os processos de leitura e escrita parâmetros de leitura tanto no IBIS como no Optis.

O Optis é absolutamente configurável e permite vários tipos de experimentos de calibração, sendo possível escolher qualquer combinação das variáveis simuladas pelo modelo, as métricas de ajuste a serem utilizadas e quais dos 43 parâmetros do modelo (ou combinação destes) serão calibrados. Além disso, é possível realizar sequencialmente calibrações independentes do IBIS, permitindo que os parâmetros possam ser separados por grupos, de modo que cada grupo possa ser calibrado separadamente e, enquanto os parâmetros de um grupo estão sendo calibrados, os dos outros grupos são mantidos fixos (VAREJÃO, 2009; VAREJÃO et al., 2013).

Como já mencionado neste capítulo, o NSGA-II apresenta como solução, ao final de

sua execução, a frente de Pareto, ou seja, um conjunto de soluções não dominadas do problema. No Optis, adota-se uma estratégia de escolha de uma das soluções da frente de Pareto, sendo esta aquela que se encontra mais próxima da origem, ou seja, aquela que otimiza todas as funções objetivo de forma simétrica. Antes de se fazer tal escolha, é feita uma normalização dos valores retornados pelo NSGA-II com o intuito de se evitar interferências das unidades ou escalas nos valores das funções objetivo.

4 CALIBRAÇÃO DO MODELO HIDROLÓGICO IPH-II - RESULTADOS

Foram efetuados testes de calibração multiobjetivo do modelo IPH-II com o algoritmo genético AGE e com o algoritmo MOCOM, descritos no Capítulo 3. O modelo direto foi executado com os valores pré-fixados dos sete parâmetros do modelo e com dados reais de precipitação e evaporação (301 registros obtidos) de uma bacia do rio Canoas, em Santa Catarina, cuja área é de 989 km^2 . O período de dados foi de 11/03/1983 a 05/01/1984 e o intervalo de tempo considerado foi de um dia.

Uma série de vazões foi, então, obtida. A esta série foi adicionado um ruído de 5%, e esta nova série sintética de vazões, foi utilizada como série de vazões observadas para o procedimento de calibração. Os valores atribuídos aos sete parâmetros para gerar a série sintética de vazões foram: $I_o = 18, 20$, $I_b = 2, 95$, $H = 0, 13$, $K_s = 5, 72$, $K_{sub} = 40, 25$, $RMAX = 0, 24$ e $\alpha = 12, 66$ (BRAVO et al., 2009).

Foram consideradas as duas funções objetivo abaixo:

$$F_1 = \sqrt{\frac{\sum_{t=1}^{NT} (Q_t - \bar{Q}_t)^2}{NT}} \quad (\text{Desvio padrão das vazões}) , \quad (4.1)$$

$$F_2 = \sqrt{\frac{\sum_{t=1}^{NT} (\frac{1}{Q_t} - \frac{1}{\bar{Q}_t})^2}{NT}} \quad (\text{Desvio padrão do inverso das vazões}) . \quad (4.2)$$

O espaço de busca dos parâmetros é apresentado na tabela abaixo que mostra os limites máximo e mínimo para cada um dos 7 parâmetros a serem ajustados no IPH-II.

Tabela 4.1 - Limites máximo e mínimo dos parâmetros do modelo IPH-II para a calibração automática.

Parâmetro	Unidade	Valor mínimo	Valor máximo
I_o	$mm\Delta t^{-1}$	10.00	300.00
I_b	$mm\Delta t^{-1}$	0.10	10.00
H	--	0.01	0.99
K_s	Δt	0.01	10.00
K_{sub}	Δt	10.00	500.00
$RMAX$	mm	0.00	9.00
α	--	0.01	20.00

Como visto na descrição do AGE, além do tamanho da população, ns , que é o único

parâmetro do algoritmo MOCOM, outros parâmetros foram inseridos no AGE. Os testes iniciais foram feitos, partindo-se, inicialmente, de uma população com um tamanho $ns = 500$, considerado como bastante satisfatório para a calibração multi-objetivo do IPH-II com o algoritmo MOCOM (BRAVO et al., 2009), e se procurando ajustar, além do tamanho da população, os demais parâmetros do algoritmo AGE. Com este tamanho da população, verificou-se uma dificuldade do AGE na convergência dos pontos da população para a frente de Pareto, embora os valores obtidos para as funções objetivo fossem muito próximos dos obtidos com o MOCOM. Um número muito grande iterações (gerações) se dava sem que se conseguisse chegar a uma população em que todos os indivíduos fossem não dominados.

A Figura 4.1 mostra a frente de Pareto, no espaço objetivo, para testes de calibração do IPH-II com os dois algoritmos, AGE e MOCOM, com uma população de tamanho $ns = 200$ para ambos. Nos dois algoritmos, considerou-se o mesmo número máximo de gerações, igual a 1000. Pode-se observar na figura que o AGE parou antes de se chegar a uma população em que todos os indivíduos fossem não dominados. Nota-se, no entanto, pelos valores das funções objetivo, que o ajuste dos parâmetros por ele é similar ao do MOCOM.

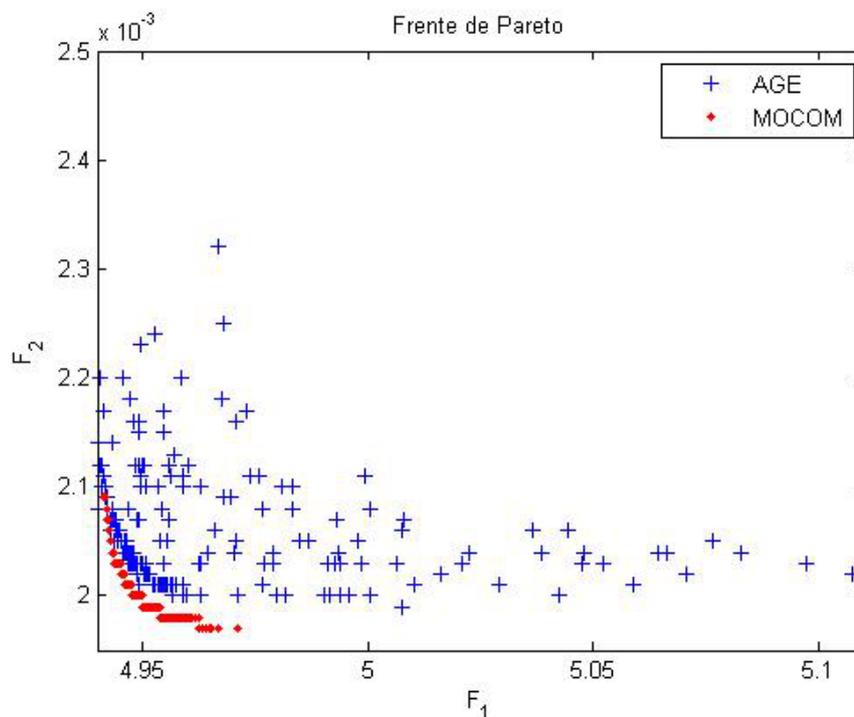


Figura 4.1 - Espaço objetivo resultante da calibração multiobjetivo do IPH-II, com $ns = 200$.

Resultados bastante satisfatórios para os testes com o AGE foram obtidos quando se considerou o tamanho da população igual a 50. E, a partir de vários testes, estabeleceu-se o seguinte conjunto de parâmetros do algoritmo:

- Número máximo de iterações: $MAX(ITER) = 5000$;
- Taxa de mutação: $TAX(MUT) = 0.003$;
- Máximo de tentativas de evolução de um Complexo: $MAX(GER) = 100$;
- Taxa de ativação da epidemia (percentual de Complexos sem evolução a partir do qual a epidemia é aplicada): $TAE = 0.70$;
- Taxa de epidemia (indica em que indivíduos a epidemia será aplicada): indivíduos com valores de ranking maiores ou iguais a $\frac{R_{max}}{2} + 1$ serão eliminados;
- Até quando serão aplicadas epidemias: até que o número de iterações alcance a metade do valor de $MAX(ITER)$.

A Figura 4.2 mostra resultados de calibração do IPH-II com o MOCOM e com o AGE (este, com os parâmetros acima), considerando uma população de tamanho $ns = 50$. É importante observar que os melhores resultados de calibração se dão à medida que a frente de Pareto se aproxima da origem.

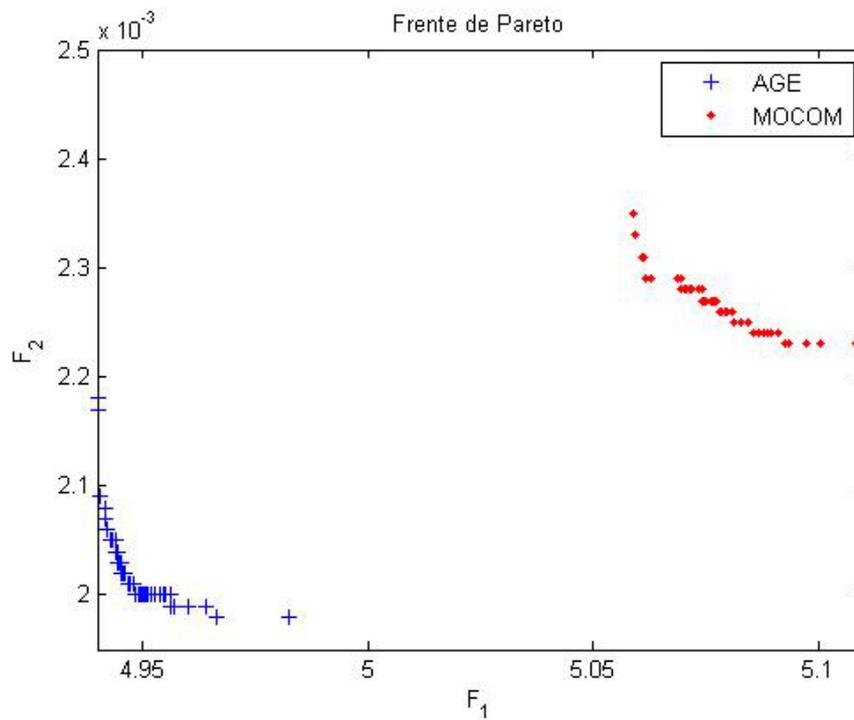


Figura 4.2 - Espaço objetivo resultante da calibração multiobjetivo do IPH-II, com $ns = 50$.

As duas estratégias, AGE e MOCOM, mostraram-se bastante satisfatórias para a recuperação dos valores pré-definidos dos parâmetros do modelo IPH-II.

Na tabela 4.2, são apresentados os valores máximo e mínimo de cada um dos parâmetros obtidos a partir de uma série de calibrações multiobjetivo do IPH-II, usando o AGE e o MOCOM, considerando em ambos uma população de mesmo tamanho, $ns = 50$, e com os parâmetros do AGE sendo os listados anteriormente. Para ambas as estratégias, a faixa de variação dos parâmetros foi relativamente estreita, o que é uma boa indicação da validade do modelo hidrológico. Nas calibrações com o AGE, os valores do parâmetro α estão bem próximos do limite superior estabelecido para o parâmetro, o que indica a necessidade de ampliar tal limite.

Tabela 4.2 - Faixas de variação dos parâmetros do IPH-II obtidas por uma série de calibrações com o AGE e com o MOCOM, usando população de tamanho $ns = 50$.

Parâmetro	Unidade	Valor exato	Varição-MOCOM	Varição-AGE
I_o	$mm\Delta t^{-1}$	18.20	11.72–23.45	22.60–27.04
I_b	$mm\Delta t^{-1}$	2.95	2.64–3.05	2.87–3.08
H	--	0.13	0.028–0.077	0.035–0.40
K_s	Δt	5.72	5.61–5.68	5.60–5.73
K_{sub}	Δt	40.25	35.89–47.53	35.76–43.29
$RMAX$	mm	0.24	0.01–1.53	0.44–0.86
α	--	12.66	4.89–9.94	17.04–20.00

A Figura 4.3 apresenta um gráfico gerado a partir de todos os pontos do conjunto de Pareto obtidos ao final de uma das calibrações do IPH-II com o AGE, da série de calibrações mencionada acima. Os parâmetros foram normalizados.

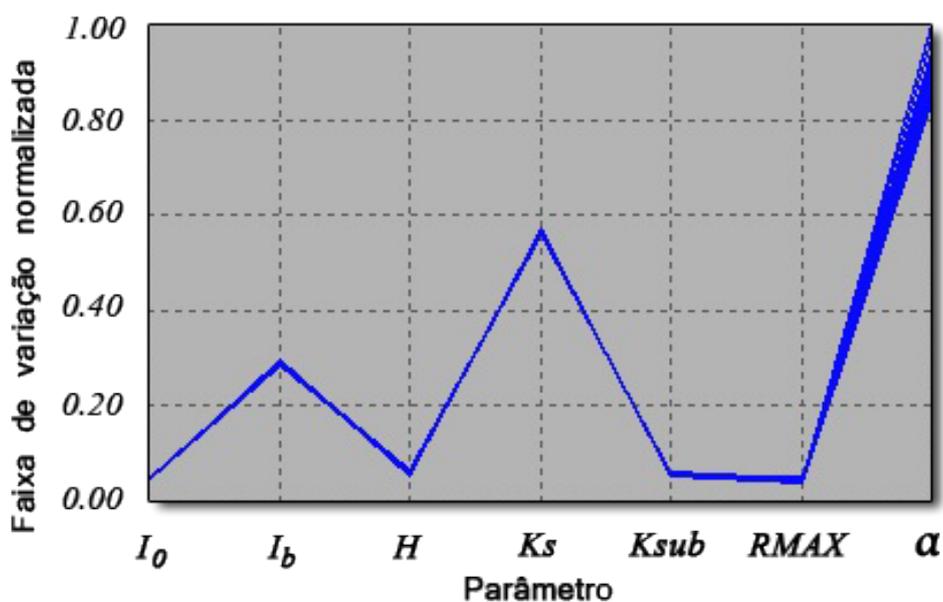


Figura 4.3 - Conjunto de Pareto resultante de uma calibração multiobjetivo do IPH-II, com o AGE.

A Figura 4.4 apresenta hidrogramas (vazões-tempo), com a vazão observada e a série de vazões calculadas a partir de todos os pontos do conjunto de Pareto representado na Figura 4.3.

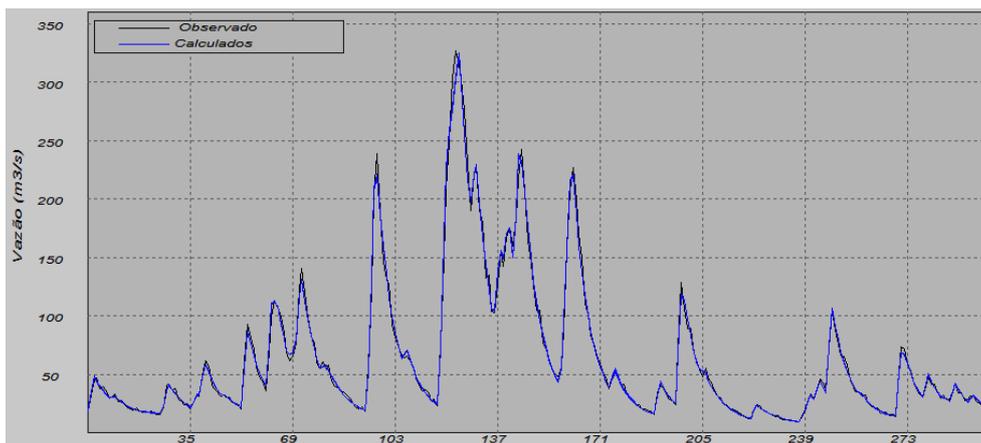


Figura 4.4 - Vazões observadas e vazões calculadas com o conjunto de Pareto.

Como se pode observar na Figura 4.3, as soluções de Pareto tendem a se aglomerar com bastante proximidade. No entanto, o parâmetro α apresenta uma maior variação, indicando uma tendência de que ele seja diferente nos dois objetivos considerados na calibração. Este parâmetro, certamente, desempenha um importante papel na forma dos hidrogramas calculados.

5 CALIBRAÇÃO DO IBIS - RESULTADOS

Utilizando o Optis, (VAREJÃO, 2009; VAREJÃO et al., 2013) propuseram uma calibração baseada em uma hierarquia temporal dos processos simulados pelo IBIS, a qual ele chamou de *calibração hierárquica*, começando dos processos mais rápidos para os mais lentos, estabelecendo 5 níveis de calibração. Com um conhecimento prévio das relações entre os parâmetros do modelo e os processos simulados e através da análise de sensibilidade dos parâmetros com o método de Morris, estabeleceram-se as variáveis a serem consideradas e o grupo de parâmetros a serem calibrados em cada nível:

- Nível 1 - Fluxos radiativos: PAR_o e $fAPAR$ (3 parâmetros);
- Nível 2 - Saldo de radiação da superfície: Rn (3 parâmetros);
- Nível 3 - Turbulência: u^* (3 parâmetros);
- Nível 4 - Fluxos turbulentos: NEE , HE , LE (16 parâmetros);
- Nível 5 - Alocação de carbono: LAI (6 parâmetros).

Assim, a calibração dos parâmetros é executada no Optis, com o algoritmo NSGA-II, começando pela calibração dos parâmetros do Nível 1, sendo que os parâmetros obtidos em cada nível são mantidos fixos no nível seguinte e, ao fim da calibração dos parâmetros do último nível, têm-se todos os parâmetros calibrados.

Para definir as funções objetivo na calibração, (VAREJÃO, 2009; VAREJÃO et al., 2013) usaram as seguintes estatísticas (já mencionadas neste trabalho):

$$MAE = \frac{\sum_{i=1}^n |P_i - O_i|}{n} \quad (5.1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}} \quad (5.2)$$

$$B = \sum_{i=1}^n (P_i - O_i) \quad (5.3)$$

$$B_{max} = Max_{j=1}^n \sum_{i=1}^j (P_i - O_i) . \quad (5.4)$$

Um dos problemas apontados por Varejão como motivadores da elaboração da estratégia de calibração hierárquica foi o fato de considerar que cada variável de saída do

modelo já consiste, por si só, uma função objetivo. Assim, se fosse considerar todas as 9 variáveis de saída do modelo, ainda que usando uma mesma estatística para avaliar os erros em cada uma delas, ter-se-iam 9 funções objetivo, o que constitui um fator de considerável dificuldade para qualquer algoritmo de otimização multiobjetivo baseado no conceito de dominância. Sua intenção, portanto, foi a de reduzir o número de variáveis nas calibrações, considerando o máximo de 3 variáveis (nível 4).

Adotando, por exemplo, a métrica MAE no nível 1, têm-se duas funções objetivo: $MAE(PAR_o)$ e $MAE(fAPAR)$; no nível 2, têm-se três funções objetivo: $MAE(NEE)$, $MAE(H)$ e $MAE(LE)$.

Neste trabalho, estamos propondo uma estratégia para definir as funções objetivo, a partir das estatísticas acima, mas considerando que a multiobjetividade vai estar relacionada não ao número de variáveis de saída envolvidas na calibração, mas, sim, ao número de métricas (estatísticas) usadas na avaliação dos erros.

Considerando que são 9 variáveis do modelo consideradas em nosso trabalho, vamos denotá-las por $Y^{(1)}, Y^{(2)}, \dots, Y^{(9)}$. Consideremos duas métricas (estatísticas) diferentes, M_1 e M_2 . Daí, construímos dois vetores:

$$V_1 = (M_1(Y^{(1)}), M_1(Y^{(2)}), \dots, M_1(Y^{(9)})) \quad (5.5)$$

$$V_2 = (M_2(Y^{(1)}), M_2(Y^{(2)}), \dots, M_2(Y^{(9)})) . \quad (5.6)$$

As funções objetivo serão, então, estabelecidas, considerando uma minimização das normas dos dois vetores acima. Com isto, definimos as seguintes funções objetivo:

$$\Phi_1 = M_1(Y^{(1)}) + M_1(Y^{(2)}) + \dots + M_1(Y^{(9)}) \quad (5.7)$$

$$\Phi_2 = M_2(Y^{(1)}) + M_2(Y^{(2)}) + \dots + M_2(Y^{(9)}) . \quad (5.8)$$

Para ilustrar um dos vetores acima, consideremos, por exemplo, a métrica $M = MAE$. Para cada variável de saída, tomando os seus n valores calculados, $P^{(1)}, P^{(2)}, \dots, P^{(9)}$, e observados, $O^{(1)}, O^{(2)}, \dots, O^{(9)}$, temos, para cada $j = 1, 2, \dots, 9$:

$$M(Y^{(j)}) = \frac{\sum_{i=1}^n |P_i^{(j)} - O_i^{(j)}|}{n} . \quad (5.9)$$

Os resultados aqui apresentados foram obtidos considerando as seguintes métricas:

$$M_1(Y^{(j)}) = \sqrt{\frac{\sum_{i=1}^n (P_i^{(j)} - O_i^{(j)})^2}{n}} \quad (5.10)$$

e

$$M_2(Y^{(j)}) = \sqrt{\frac{\sum_{i=1}^n \left(\frac{1}{P_i^{(j)}} - \frac{1}{O_i^{(j)}}\right)^2}{n}}. \quad (5.11)$$

Portanto, foram usadas as duas funções objetivo Φ_1 e Φ_2 , definidas pelas equações (5.7) e (5.8), com as métricas acima.

Como séries de dados observados para as variáveis de saída do IBIS foram construídas séries sintéticas. Foram usados dados observados disponíveis para as variáveis consideradas correspondentes à Flona de Tapajós. Este sítio é localizado próximo ao km 67 da rodovia Santarém-Cuiabá (2°51, 54°58'W). Inicialmente, a partir dos dados disponíveis, fez-se uma calibração dos 43 parâmetros do IBIS, usando o Optis. Os parâmetros calibrados foram, então, utilizados como valores pré-fixados de parâmetros para gerar a série sintética de dados. Rodou-se o modelo com os valores pré-fixados dos parâmetros e com dados de radiação solar incidente, de radiação de onda longa incidente, de temperatura do ar, de velocidade horizontal do vento, de precipitação e de umidade relativa do ar correspondentes à Flona de Tapajós. Em seguida, às variáveis de saída foi acrescentado um ruído de 5% para se chegar às séries sintéticas de dados observados. O período de dados foi de janeiro de 2002 a dezembro de 2004. Temos Rn , PAR_o , HE , LE , NEE e u^* como variáveis horárias; $fAPAR$ e LAI como variáveis mensais e a NPP como variável anual.

Foram feitas calibrações, considerando três tipos de situações que serão detalhadas na apresentação dos resultados nos tópicos seguintes.

5.1 Testes com a Calibração hierárquica

Seguindo a estratégia de calibração hierárquica (VAREJÃO, 2009; VAREJÃO et al., 2013), temos a seguinte estrutura representada na tabela abaixo, considerando, agora, as funções objetivo Φ_1 e Φ_2 .

Tabela 5.1 - Funções objetivo, variáveis e parâmetros calibrados em cada nível na calibração hierárquica.

Nível	Variáveis	Funções Objetivo	Parâmetros
1	PAR_o e $fAPAR$	Φ_1, Φ_2	<i>rhoveg_vis</i> <i>tauveg_vis</i> <i>chifuz</i>
2	Rn	Φ_1, Φ_2	<i>rhoveg_NIR</i> <i>tauveg_NIR</i> <i>avmuir_coef</i>
3	u^*	Φ_1, Φ_2	<i>dispu_coef</i> <i>alogl_coef</i> <i>alogu_coef</i>
4	NEE, HE e LE	Φ_1, Φ_2	<i>vmax_pft</i> <i>coefmub</i> <i>chs</i> <i>beta2</i> <i>funca_coef</i> <i>funcbcoef</i> <i>root_coef</i> <i>rwood_coef</i> <i>tempvm_coef</i> <i>stressf_coef</i> <i>clitls_coef</i> <i>clitrs_coef</i> <i>clitws_coef</i> <i>csoislon_coef</i> <i>csoislop_coef</i> <i>kfactor</i>
5	LAI e NPP	Φ_1, Φ_2	<i>rgrowth_coef</i> <i>tauleaf</i> <i>specla</i> <i>aleaf</i> <i>arrot</i> <i>awood</i>

A tabela seguinte mostra os valores pré-fixados dos parâmetros e os valores obtidos na calibração.

Tabela 5.2 - Resultados da calibração hierárquica.

Parâmetros	Valores exatos	Valores calibrados
<i>rhoveg_vis</i>	0.08720739	0.07782133
<i>tauveg_vis</i>	0.04984299	0.05236471
<i>chifuz</i>	-0.22492749	-0.16428874
<i>rhoveg_NIR</i>	0.29658099	0.34676619
<i>tauveg_NIR</i>	0.20376832	0.22352694
<i>avmuir_coef</i>	370.97404379	366.48724663
<i>dispu_coef</i>	0.97789433	0.98953577
<i>alogl_coef</i>	3.92890608	4.46822338
<i>alogu_coef</i>	7.2151337	6.44468483
<i>vmax_pft</i>	0.00005841	0.00004996
<i>coefmub</i>	7.54944019	7.66805367
<i>chs</i>	33448.51893589	22679.84521921
<i>beta2</i>	0.78376625	0.80806123
<i>funca_coef</i>	5477.41470401	6678.27075942
<i>funcbcoef</i>	5900.21273165	5306.04776845
<i>root_coef</i>	0.84288381	1.07025164
<i>rwood_coef</i>	0.15634128	0.0856671
<i>tempvm_coef</i>	3961.88003908	3502.32838518
<i>stressf_coef</i>	-5.31207677	-5.11142866
<i>clitls_coef</i>	1.46855317	2.17268297
<i>clitrs_coef</i>	4.51240714	4.92262684
<i>clitws_coef</i>	1.04517844	1.04517844
<i>csoislon_coef</i>	0.2501865	0.34285772
<i>csoislop_coef</i>	6.37357809	6.6973777
<i>kfactor</i>	1.47438363	1.38101766
<i>rgrowth_coef</i>	0.29653374	0.27381527
<i>tauleaf</i>	0.58080392	0.72082252
<i>specla</i>	31.29498096	32.90618069
<i>aleaf</i>	0.26536141	0.18564253
<i>aroot</i>	0.32691409	0.18695653
<i>awood</i>	0.62796179	0.47728325

Nas figuras seguintes, temos os gráficos comparando os valores observados (sintéticos) e os valores das variáveis de saída do modelo calculados com os parâmetros calibrados apresentados na tabela 5.2.

A Figura 5.1 mostra os gráficos dos valores horários observados e calculados, ao longo de um dia, da radiação fotossinteticamente ativa refletida (PAR_o). Observa-se um bom ajuste dos dados, com a superestimação nos valores calculados no período entre as 8h e as 17h, aproximadamente, compreendendo o horário em que ocorre o valor de pico das séries.

Na Figura 5.2, temos os gráficos dos valores mensais calculados e observados, ao longo de um ano, da fração absorvida da radiação fotossinteticamente ativa ($fAPAR$). Podemos observar que houve um bom ajuste entre os dados, com uma superestimação dos valores calculados ao longo de todo o ano.

A Figura 5.3 mostra o gráfico dos valores horários calculados e observados, ao longo de um dia, do saldo de radiação Rn . Observa-se que houve um excelente ajuste (praticamente perfeito) entre os dados.

A Figura 5.4 traz os gráficos dos valores horários calculados e observados da velocidade de fricção do vento (u^*) ao longo de um dia. Podemos observar que houve um ótimo ajuste entre os dados, com uma superestimativa dos valores calculados, que é mais acentuada nas proximidades do horário de pico.

A Figura 5.5 traz os gráficos dos valores horários observados e calculados da troca líquida de CO_2 do ecossistema (NEE) ao longo de um dia. Podemos observar que houve um ótimo ajuste entre o observado e o calculado, com uma superestimativa dos valores calculados, que foi maior nas 8 primeiras horas do dia.

Na Figura 5.6, temos o gráfico dos valores horários calculados e observados do fluxo de calor sensível (HE) ao longo de um dia. Os valores estão muito bem ajustados, ocorrendo alguma subestimação em alguns intervalos e uma superestimação nas proximidades do horário de pico dos valores calculados.

Na Figura 5.7, temos os gráficos dos valores horários calculados e observados do fluxo de calor latente (LE) ao longo de um dia, onde podemos observar que houve um ajuste muito bom entre os dados, com uma superestimação dos valores calculados no período entre as 9h e as 17h, aproximadamente.

Finalmente, a Figura 5.8 traz os gráficos dos valores mensais observados e calculados

do índice de área foliar (*LAI*) ao longo de um ano. Observa-se que houve um ajuste muito bom dos dados com uma superestimativa nos valores calculados.

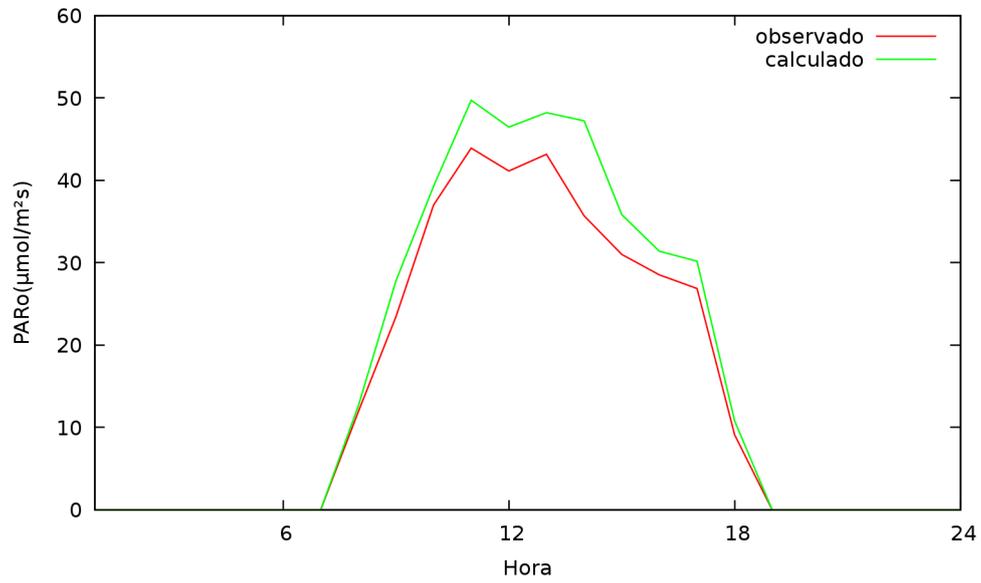


Figura 5.1 - Resultados de PARo em um dia típico - calibração hierárquica.

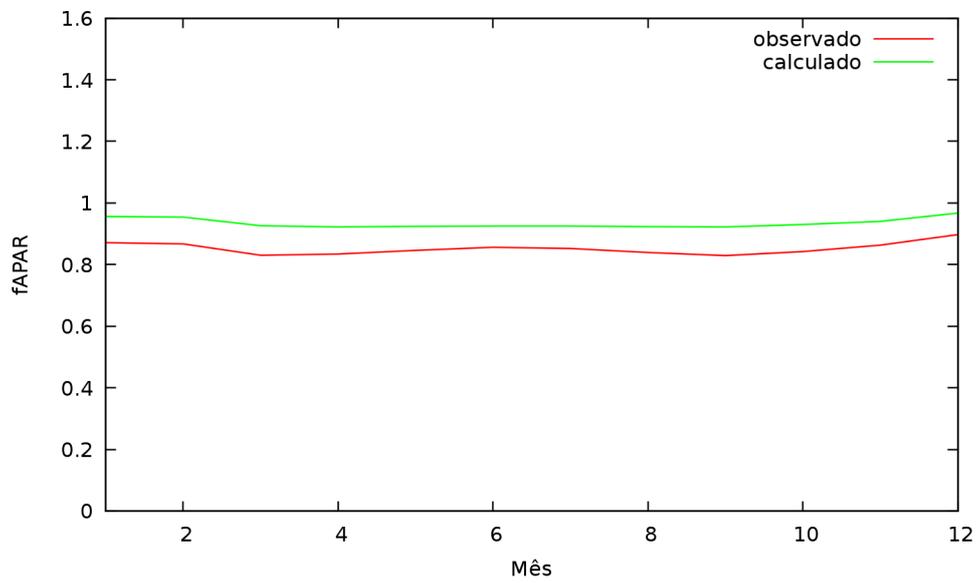


Figura 5.2 - Resultados de fAPAR com dados mensais durante um ano - calibração hierárquica.

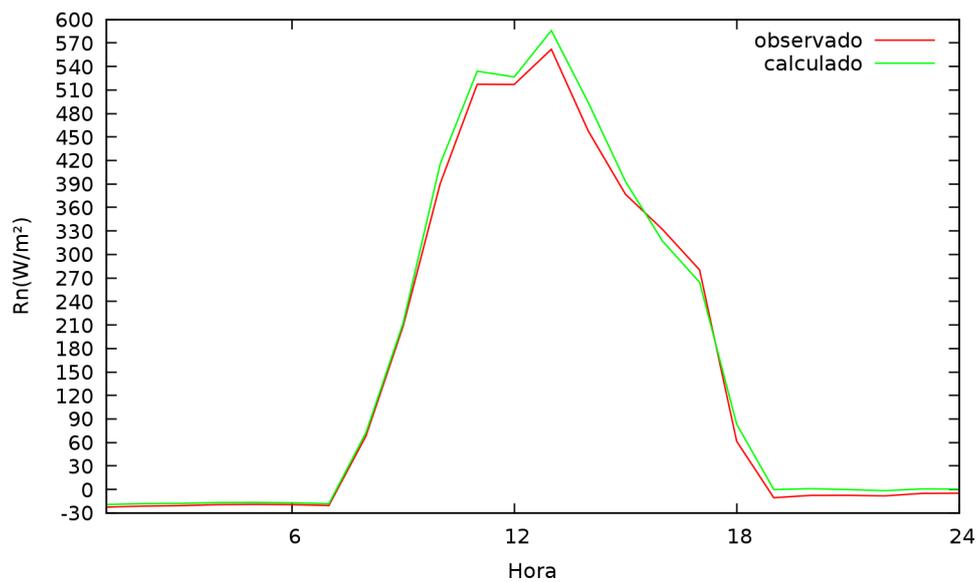


Figura 5.3 - Resultados de Rn em um dia típico - calibração hierárquica.

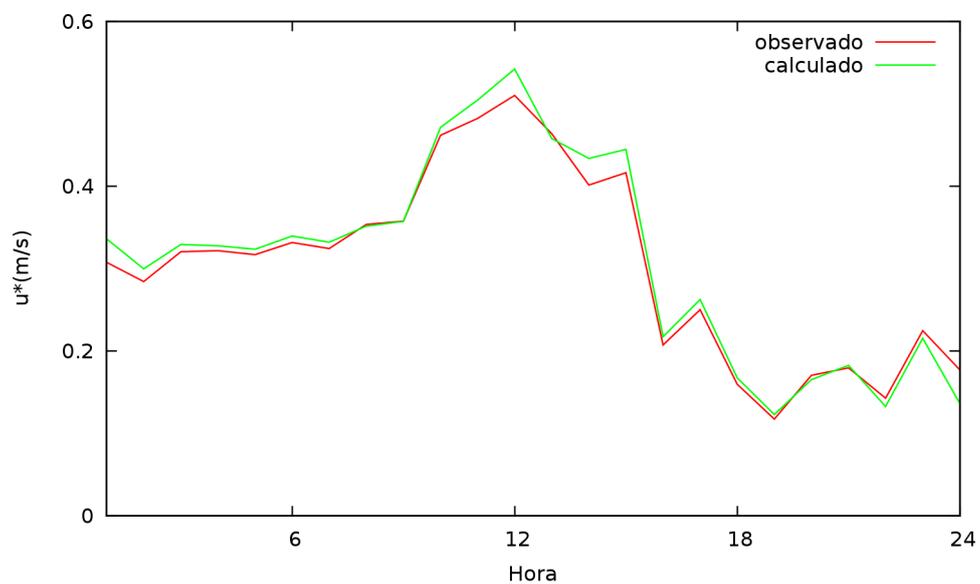


Figura 5.4 - Resultados de u^* em um dia típico - calibração hierárquica.

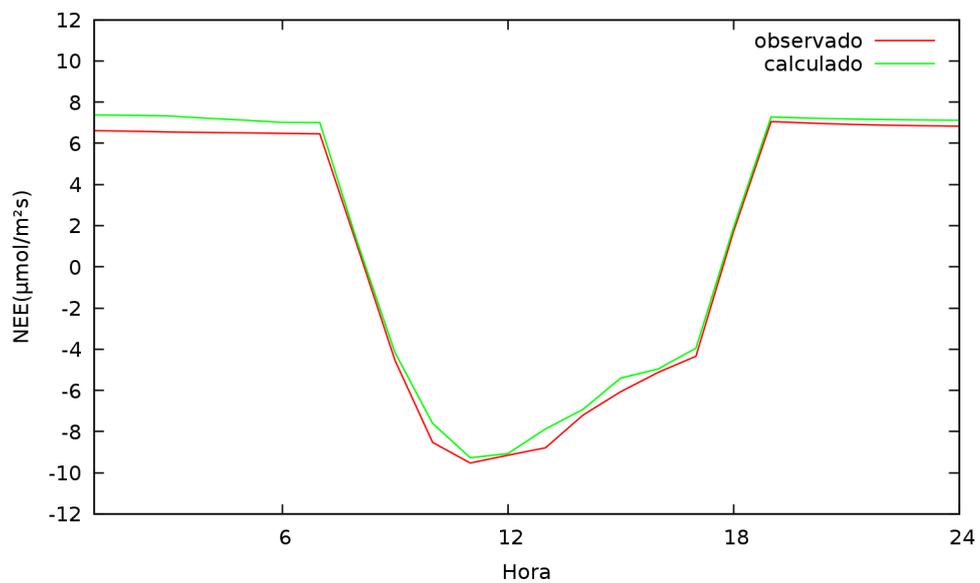


Figura 5.5 - Resultados de NEE em um dia típico - calibração hierárquica.

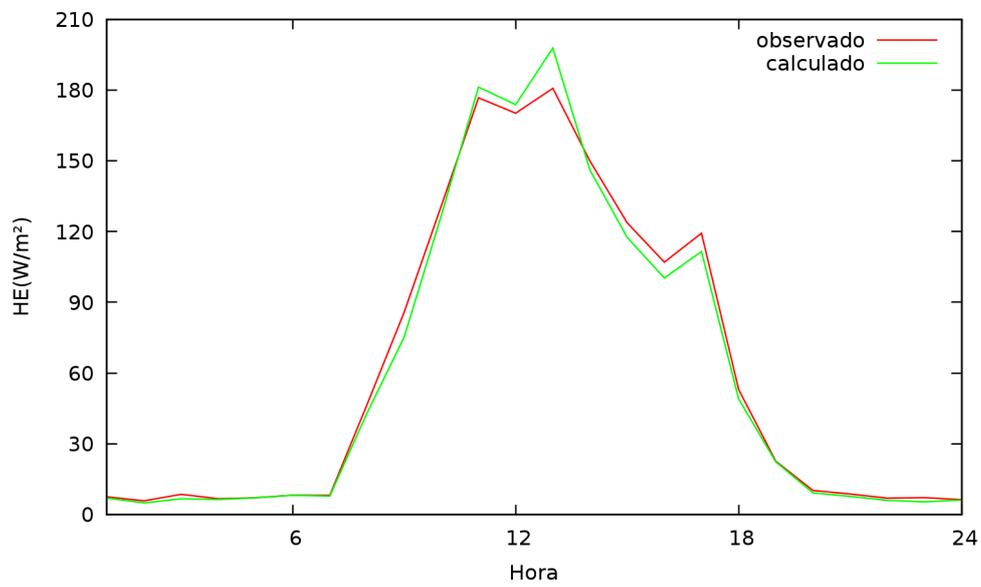


Figura 5.6 - Resultados de HE em um dia típico - calibração hierárquica.

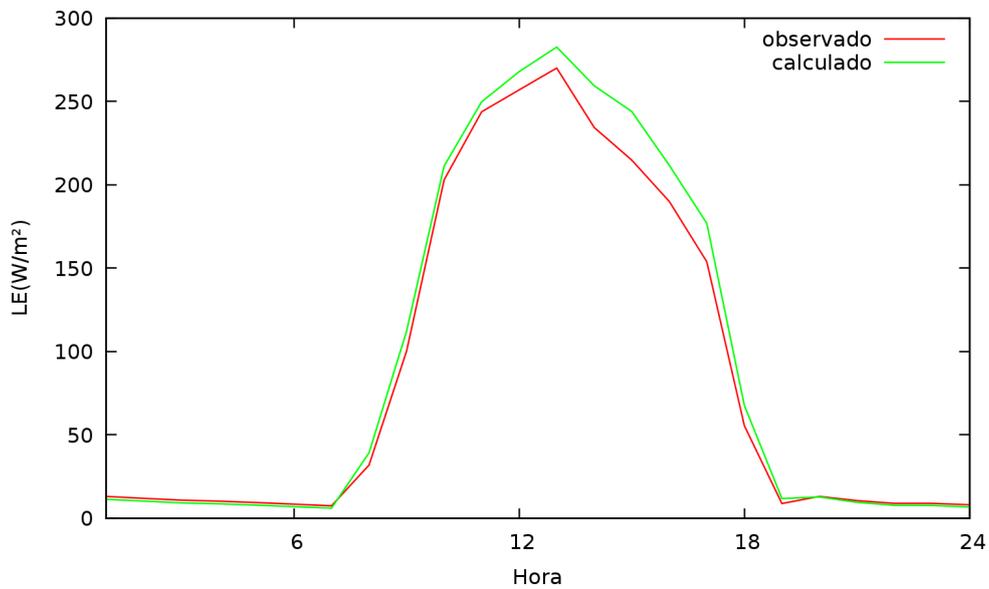


Figura 5.7 - Resultados de LE em um dia típico - calibração hierárquica.

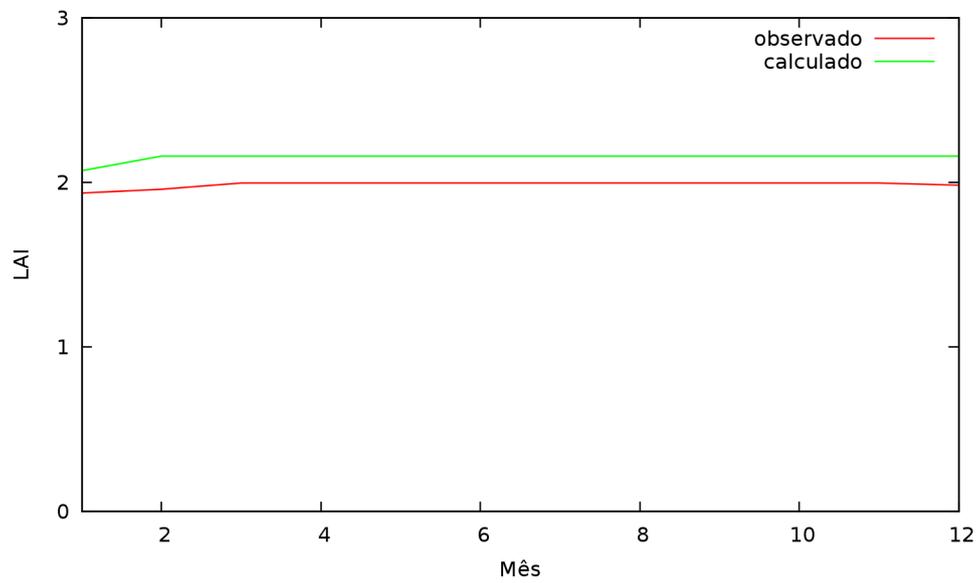


Figura 5.8 - Resultados de LAI com dados mensais durante um ano - calibração hierárquica.

Para a variável anual NPP , têm-se os valores $NPP_{observado} = 0.427770$ e $NPP_{calculado} = 0.493276$, correspondentes a um mesmo ano.

5.2 Calibração a Partir de uma Análise de Sensibilidade Dependente da Escala Temporal

Os resultados de calibração que serão apresentados aqui foram obtidos a partir de uma estratégia de calibração inspirada em um trabalho de [Minjiao e Xiao \(2014\)](#), no qual eles apresentam a calibração de um modelo hidrológico baseada em uma análise de sensibilidade dos parâmetros do modelo, através do método de Morris, considerando uma dependência das escalas temporais nos dados de vazão.

Foram consideradas as seguintes variáveis do IBIS: PAR_o , Rn , HE , LE , NEE e u^* . A partir dos dados sintéticos horários das variáveis acima, foram construídos, através de valores médios, os dados diários e mensais correspondentes às mesmas variáveis. Fez-se, então, uma análise de sensibilidade, com o método de Morris, para os dados mensais, diários e horários, obtendo-se os seguintes grupos de parâmetros:

- Grupo 1: parâmetros sensíveis para a escala mensal
- Grupo 2: parâmetros sensíveis para a escala diária
- Grupo 3: parâmetros sensíveis para a escala horária

Os resultados da análise de sensibilidade e o procedimento para a composição dos grupos acima são apresentados no Apêndice A.

Como no caso da calibração hierárquica, um esquema de calibração em etapas é adotado, com os três passos seguintes:

- (1) Assume-se um conjunto fixo de parâmetros nos grupos 2 e 3 e os parâmetros do Grupo 1 são ajustados;
- (2) Com os parâmetros do Grupo obtidos no passo 1 e com os mesmos parâmetros fixados para o Grupo 3, ajustam-se os parâmetros do Grupo 2;
- (3) Com os parâmetros dos grupos 1 e 2 ajustados nos passos anteriores, ajustam-se os parâmetros do Grupo 3.

Nas calibrações dos parâmetros em cada grupo, foram utilizadas as funções objetivo Φ_1 e Φ_2 aplicadas em todas as saídas do modelo consideradas neste caso (6 variáveis).

A tabela seguinte mostra a composição destes grupos.

Tabela 5.3 - Grupos de calibração em escala temporal.

Grupos	Escala	Parâmetros
<i>Grupo1</i>	Mensal	<i>funcb_coef</i> <i>rroot_coef</i> <i>rwood_coef</i> <i>rgrowth_coef</i> <i>tempvm_coef</i> <i>wsoi_coef</i>
<i>Grupo2</i>	Diária	<i>rhoveg_NIR</i> <i>chu</i> <i>tauleaf</i> <i>tauroot</i> <i>specla</i> <i>aleaf</i>
<i>Grupo3</i>	Horária	<i>rhoveg_vis</i> <i>tauveg_vis</i> <i>tauveg_NIR</i> <i>chifuz</i> <i>vmax_pft</i> <i>coefmub</i> <i>chs</i> <i>beta2</i> <i>stressf_coef</i> <i>clitlm_coef</i> <i>clitls_coef</i> <i>clitrs_coef</i> <i>clitwl_coef</i> <i>clitwm_coef</i> <i>clitws_coef</i> <i>kfactor</i> <i>dispu_coef</i> <i>avmuir_coef</i>

A tabela seguinte mostra o resultado da calibração em escala temporal.

Tabela 5.4 - Resultados da calibração em escala temporal.

Parâmetros	Valores exatos	Valores calibrados
<i>funcb_coef</i>	5900.21273165	4750.70310429
<i>rroot_coef</i>	0.84288381	0.65975361
<i>rwood_coef</i>	0.15634128	0.01343852
<i>rgrowth_coef</i>	0.29653374	0.37905181
<i>tempvm_coef</i>	3961.88003908	2836.74136227
<i>wsoi_coef</i>	0.91612359	0.7895921
<i>rhoveg_NIR</i>	0.29658099	0.32490211
<i>chu</i>	1302.5611543	1072.74237625
<i>tauleaf</i>	0.58080392	0.81010919
<i>tauroot</i>	1.09496897	0.08207593
<i>specla</i>	31.29498096	27.22459435
<i>aleaf</i>	0.26536141	0.28746241
<i>rhoveg_vis</i>	0.08720739	0.07704209
<i>tauveg_vis</i>	0.04984299	0.06530168
<i>tauveg_NIR</i>	0.20376832	0.16359820
<i>chifuz</i>	-0.22492749	-0.39364458
<i>vmax_pft</i>	0.00005841	0.00005482
<i>coefmub</i>	7.54944019	7.98714334
<i>chs</i>	33448.51893589	35740.95259935
<i>beta2</i>	0.78376625	0.81753497
<i>stressf_coef</i>	-5.31207677	-4.84245661
<i>clitlm_coef</i>	0.00024668	0.00019952
<i>clitls_coef</i>	1.46855317	0.62144355
<i>clitrs_coef</i>	4.51240714	4.36535523
<i>clitwl_coef</i>	0.00537632	0.00498947
<i>clitwm_coef</i>	0.00233071	0.00279589
<i>clitws_coef</i>	1.04517844	0.85122613
<i>kfactor</i>	1.47438363	0.93607121
<i>dispu_coef</i>	0.97789433	0.97739616
<i>avmuir_coef</i>	370.97404379	376.24703313

Nas figuras seguintes, temos os gráficos comparando os valores observados (sintéticos) e os valores das variáveis de saída do modelo calculados com os parâmetros calibrados apresentados na tabela 5.4.

A Figura 5.9 mostra os gráficos dos valores horários observados e calculados, ao longo de um dia, da radiação fotossinteticamente ativa refletida (PAR_o). Observa-se um bom ajuste dos dados, com a subestimação nos valores calculados no período entre as 9h e as 16h, aproximadamente, compreendendo o horário em que ocorre o valor de pico dos dados.

A Figura 5.10 mostra o gráfico dos valores horários calculados e observados, ao longo de um dia, do saldo de radiação Rn . Observa-se que houve um excelente ajuste (praticamente perfeito) entre os dados.

A Figura 5.11 traz os gráficos dos valores horários calculados e observados da velocidade de fricção do vento (u^*) ao longo de um dia. Podemos observar que houve um excelente ajuste (praticamente perfeito) entre os dados.

A Figura 5.12 traz os gráficos dos valores horários observados e calculados da troca líquida de CO_2 do ecossistema (NEE) ao longo de um dia. Podemos observar que houve um bom ajuste entre o observado e o calculado, ocorrendo uma superestimativa dos valores calculados no período de 8h às 18h, aproximadamente, que foi mais acentuada nas proximidades do horário de pico.

Na Figura 5.13, temos o gráfico dos valores horários calculados e observados do fluxo de calor sensível (HE) ao longo de um dia. Os valores estão muito bem ajustados, ocorrendo alguma superestimação dos valores calculados nas proximidades dos valores de pico dos dados.

Na Figura 5.14, temos os gráficos dos valores horários calculados e observados do fluxo de calor latente (LE) ao longo de um dia, onde podemos observar que houve um excelente ajuste entre os dados, com uma subestimação dos valores calculados nas proximidades dos valores de pico dos dados.

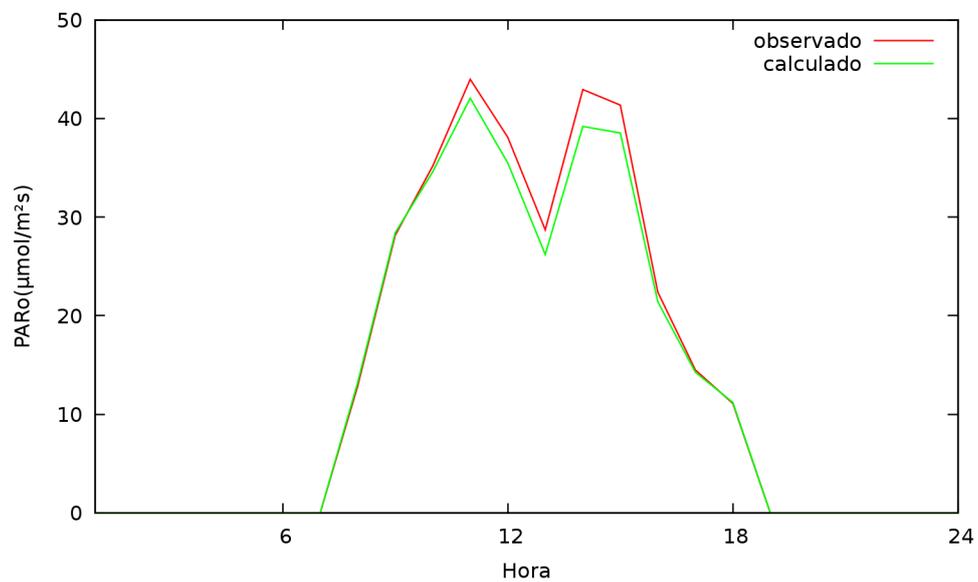


Figura 5.9 - Resultados de PARo em um dia típico - calibração em escala temporal.

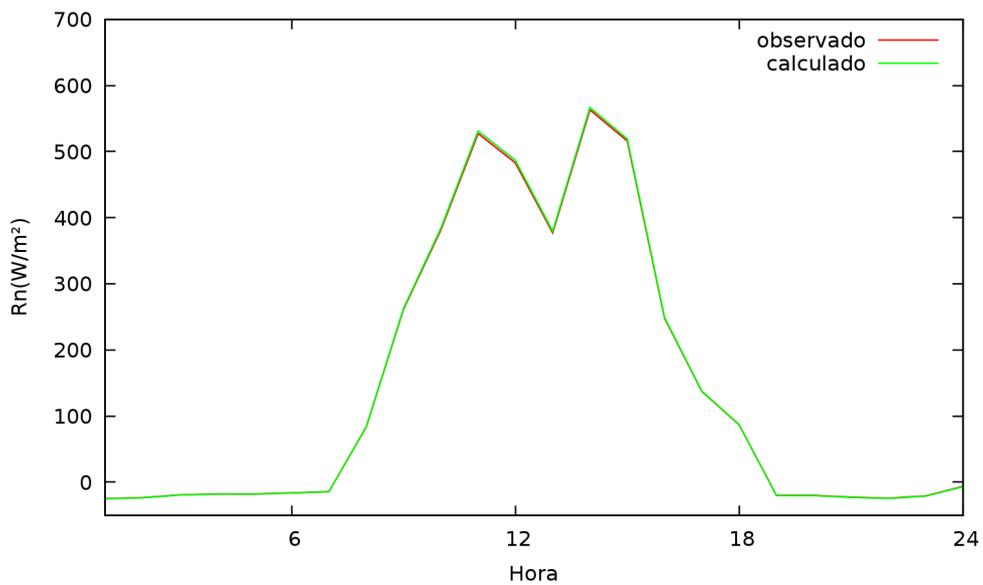


Figura 5.10 - Resultados de Rn em um dia típico - calibração em escala temporal.

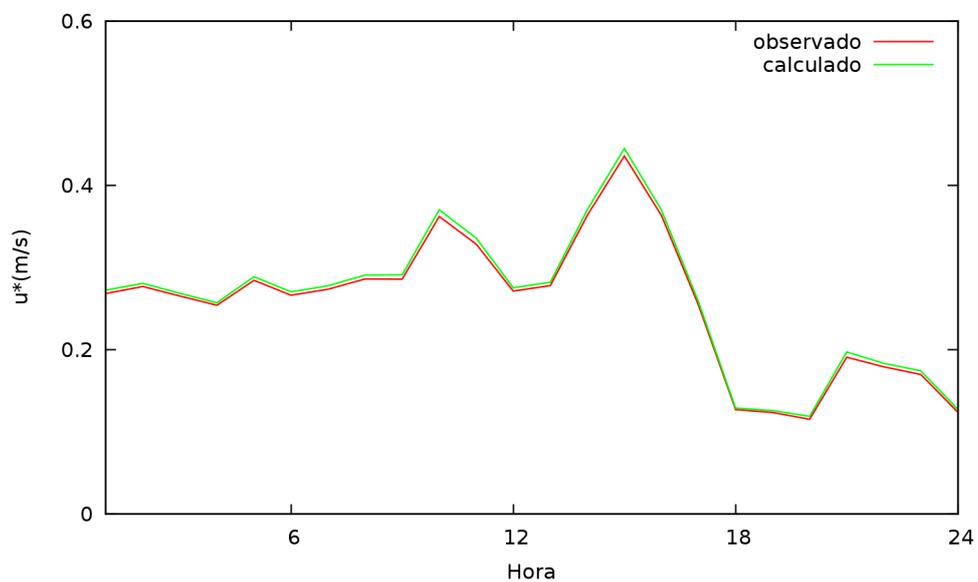


Figura 5.11 - Resultados de u^* em um dia típico - calibração em escala temporal.

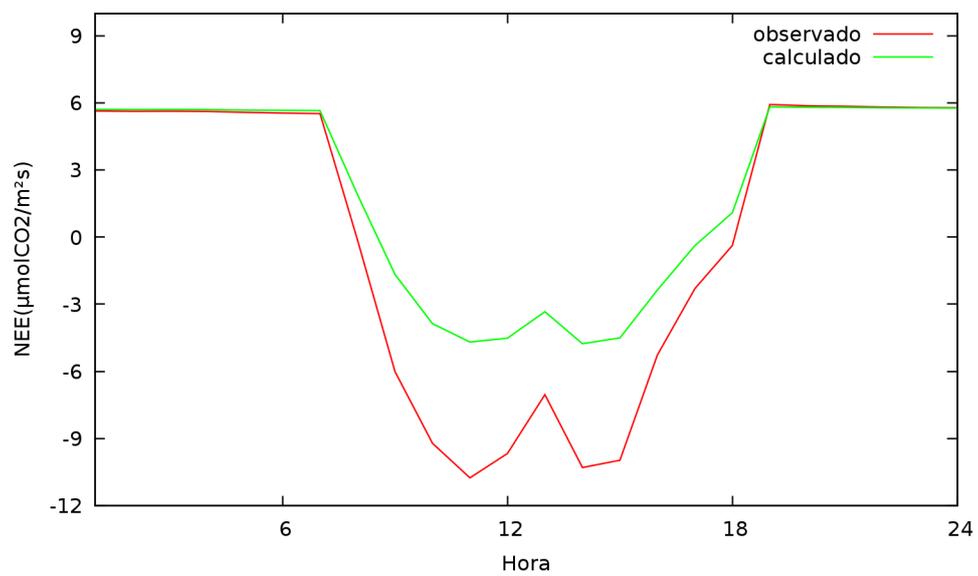


Figura 5.12 - Resultados de NEE em um dia típico - calibração em escala temporal.

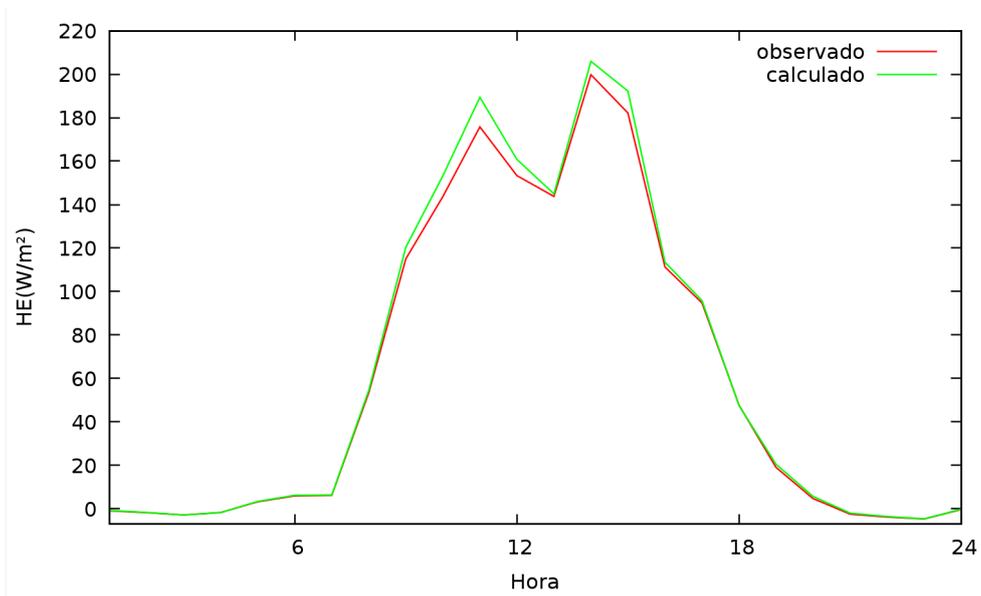


Figura 5.13 - Resultados de HE em um dia típico - calibração em escala temporal.

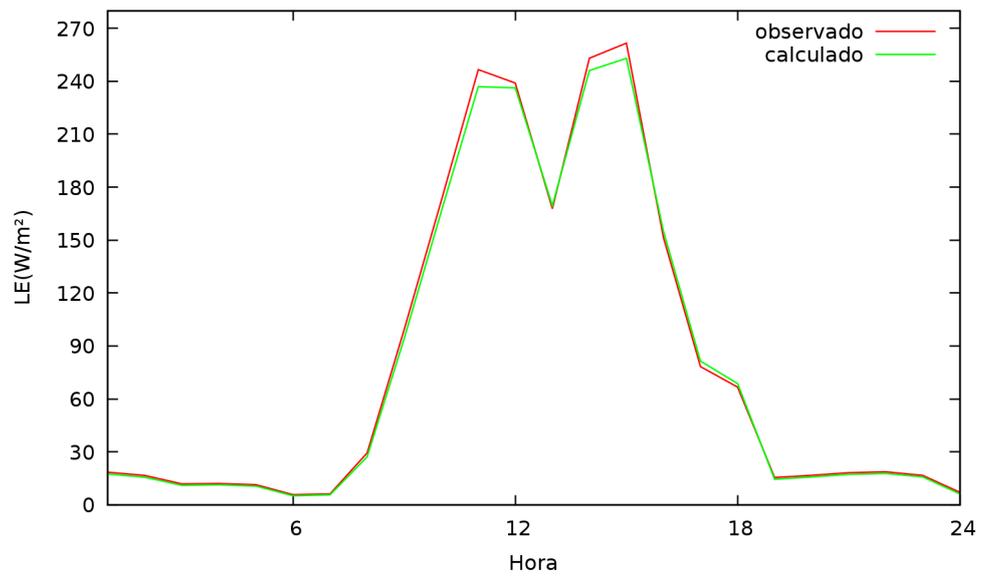


Figura 5.14 - Resultados de LE em um dia típico - calibração em escala temporal.

5.3 Calibração a Partir de uma Análise de Sensibilidade Dependente da Escala Temporal, usando o NSGA-II com Epidemia

Aqui, são apresentados resultados de testes feitos segundo a estratégia de calibração em escala temporal apresentada em 5.2, considerando o algoritmo NSGA-II com o operador de epidemia descrito em 3.5.1.

A figura abaixo mostra a frente de Pareto resultante da calibração obtida usando-se o NSGA-II (parâmetros calibrados apresentados na tabela 5.4) e a frente de Pareto resultante da calibração obtida com o NSGA-II Epidêmico. Os parâmetros calibrados são apresentados na tabela seguinte.

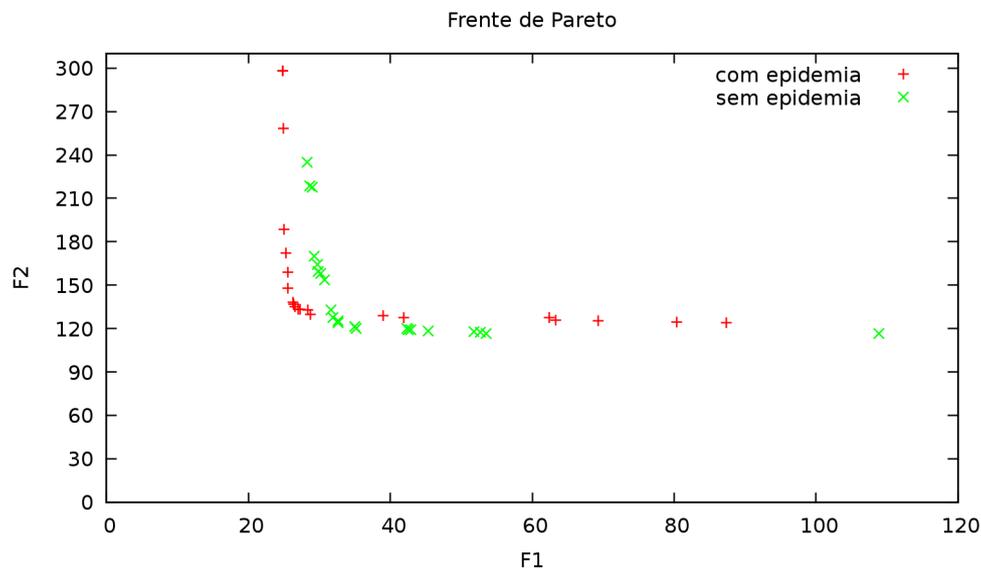


Figura 5.15 - Frentes de Pareto resultantes de calibrações do IBIS obtidas usando-se o NSGA-II com e sem epidemia - $(F_1, F_2) = (\Phi_1, \Phi_2)$.

Nestes casos, as distâncias, no espaço objetivo, das soluções escolhidas à origem foram: $d_{ep} = 0.505655$, para a calibração com epidemia; $d_{sep} = 0.546399$, para a calibração sem epidemia. Podemos observar na Figura 5.15 que as frentes de Pareto obtidas nos dois casos são muito próximas. Considerando o critério de escolha da solução ótima, o caso com epidemia levou a um melhor resultado.

Tabela 5.5 - Resultados das calibrações em escala temporal com o NSGA-II e com o NSGA-II epidêmico.

Parâmetros	Valores exatos	Valores Calibrados NSGA-II	Valores Calibrados NSGA-II Epidêmico
<i>funcb_coef</i>	5900.21273165	4750.70310429	3934.77177546
<i>rroot_coef</i>	0.84288381	0.65975361	1.03025208
<i>rwood_coef</i>	0.15634128	0.01343852	0.09358072
<i>rgrowth_coef</i>	0.29653374	0.37905181	0.39905178
<i>tempvm_coef</i>	3961.88003908	2836.74136227	3984.00792933
<i>wsoi_coef</i>	0.91612359	0.7895921	0.7096109
<i>rhoveg_NIR</i>	0.29658099	0.32490211	0.34073863
<i>chu</i>	1302.5611543	1072.74237625	1153.32436519
<i>tauleaf</i>	0.58080392	0.81010919	0.76138721
<i>tauroot</i>	1.09496897	0.08207593	0.07625137
<i>specla</i>	31.29498096	27.22459435	25.75387211
<i>aleaf</i>	0.26536141	0.28746241	0.22915312
<i>rhoveg_vis</i>	0.08720739	0.07704209	0.08112713
<i>tauveg_vis</i>	0.04984299	0.06530168	0.06121735
<i>tauveg_NIR</i>	0.20376832	0.16359820	0.14736745
<i>chifuz</i>	-0.22492749	-0.39364458	-0.27153961
<i>vmax_pft</i>	0.00005841	0.00005482	0.00004793
<i>coefmub</i>	7.54944019	7.98714334	8.17464572
<i>chs</i>	33448.51893589	35740.95259935	38495.6097157
<i>beta2</i>	0.78376625	0.81753497	0.89162825
<i>stressf_coef</i>	-5.31207677	-4.84245661	-5.31171454
<i>clitlm_coef</i>	0.00024668	0.00019952	0.000178732
<i>clitls_coef</i>	1.46855317	0.62144355	1.80555215
<i>clitrs_coef</i>	4.51240714	4.36535523	4.16392503
<i>clitwl_coef</i>	0.00537632	0.00498947	0.00517318
<i>clitwm_coef</i>	0.00233071	0.00279589	0.00191327
<i>clitws_coef</i>	1.04517844	0.85122613	2.00103714
<i>kfactor</i>	1.47438363	0.93607121	0.96205971
<i>dispu_coef</i>	0.97789433	0.97739616	0.95803133
<i>avmuir_coef</i>	370.97404379	376.24703313	352.1726548

Nas figuras seguintes, temos os gráficos comparando os valores observados (sintéticos) e os valores calculados com os parâmetros calibrados, usando o NSGA-II epidêmico, apresentados na tabela 5.5.

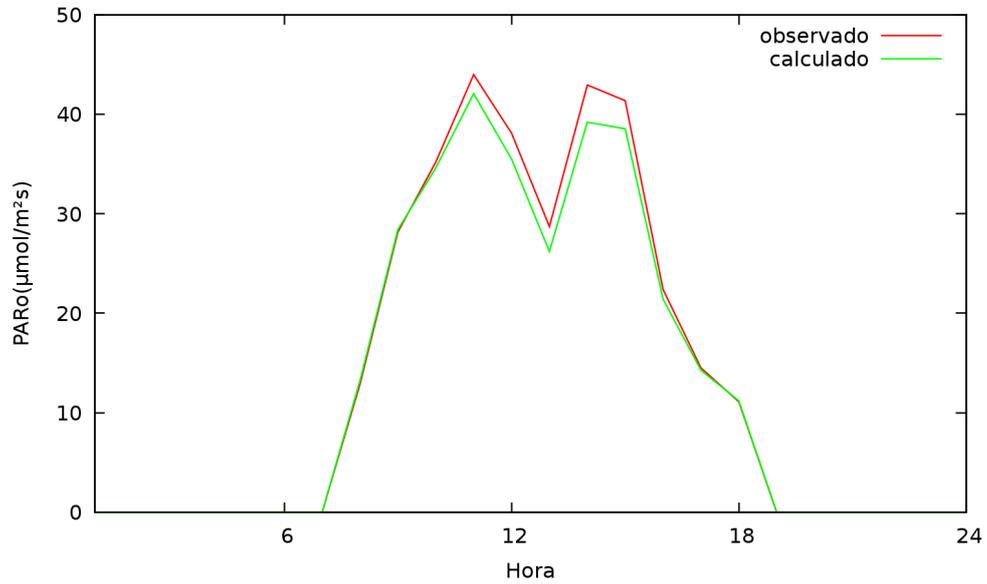


Figura 5.16 - Resultados de PARo em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

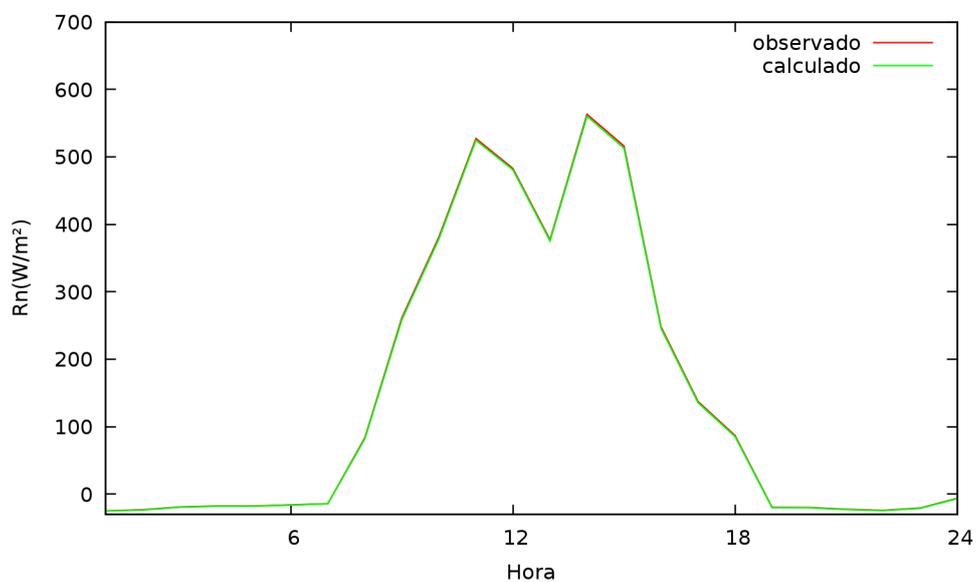


Figura 5.17 - Resultados de R_n em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

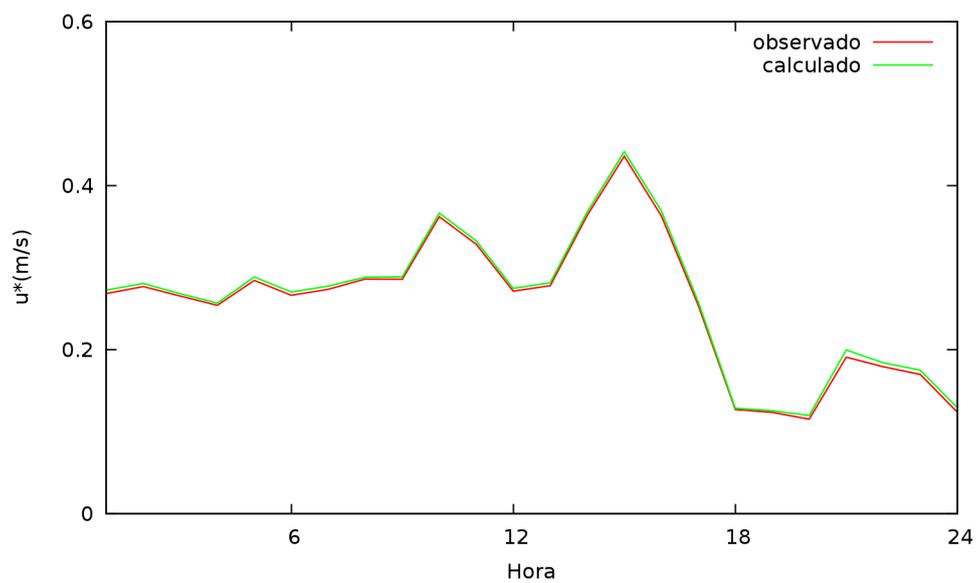


Figura 5.18 - Resultados de u^* em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

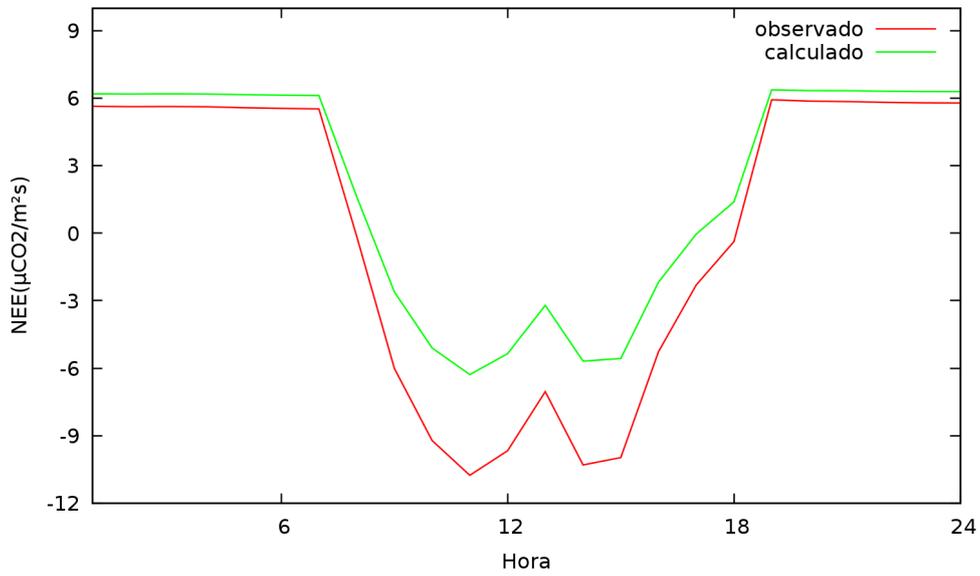


Figura 5.19 - Resultados de NEE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

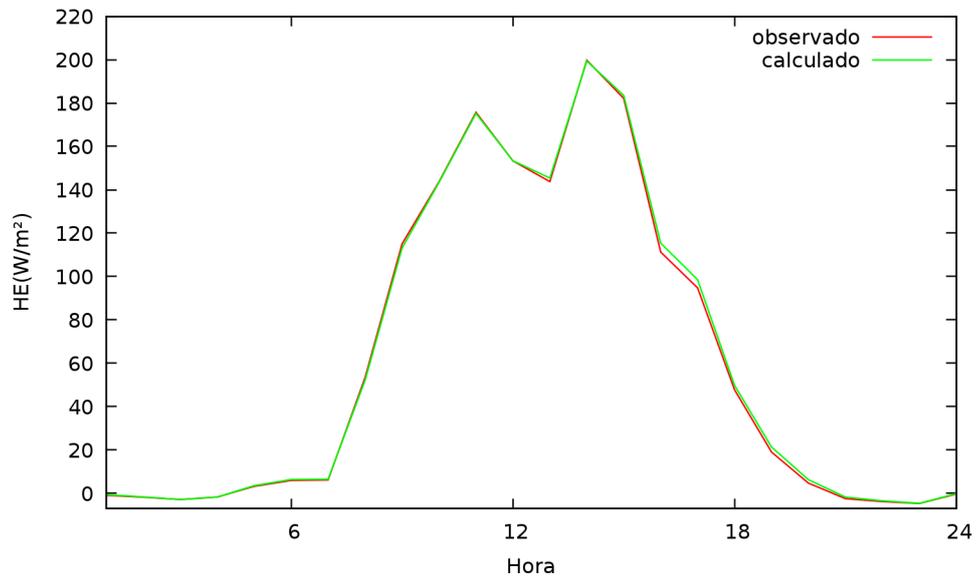


Figura 5.20 - Resultados de HE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

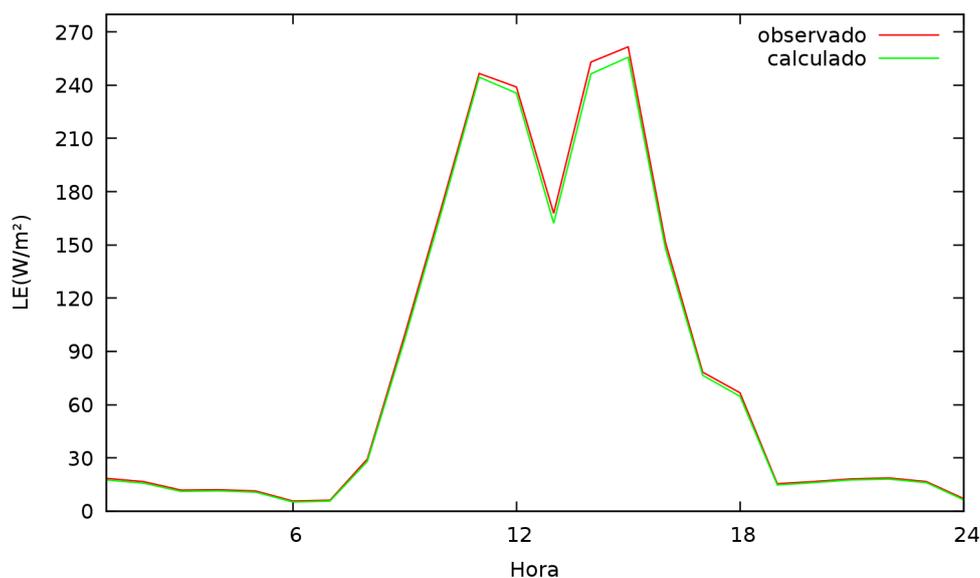


Figura 5.21 - Resultados de LE em um dia típico - calibração em escala temporal com NSGA-II epidêmico.

Comparando os gráficos acima com os obtidos com o NSGA-II sem epidemia, observamos que os resultados foram muito semelhantes nos dois casos. Podemos observar, comparando as figuras 5.19 e 5.12, por exemplo, que, para o período noturno, o ajuste com o NSGA-II sem epidemia foi melhor. Comparando as figuras 5.20 e 5.13, observamos que o ajuste com o NSGA-II epidêmico foi melhor, não apresentando a superestimativa que foi apresentada nos dados calculados com a calibração com o NSGA-II sem epidemia.

Nas figuras seguintes, os gráficos de valores observados e calculados, obtidos com o NSGA-II e com o NSGA-II epidêmico, são reapresentados, agora com os seus correspondentes gráficos das curvas de erro absoluto entre o calculado e o observado. Observa-se que os maiores valores de erro absoluto entre o calculado e o observado acontecem na calibração com o NSGA-II.

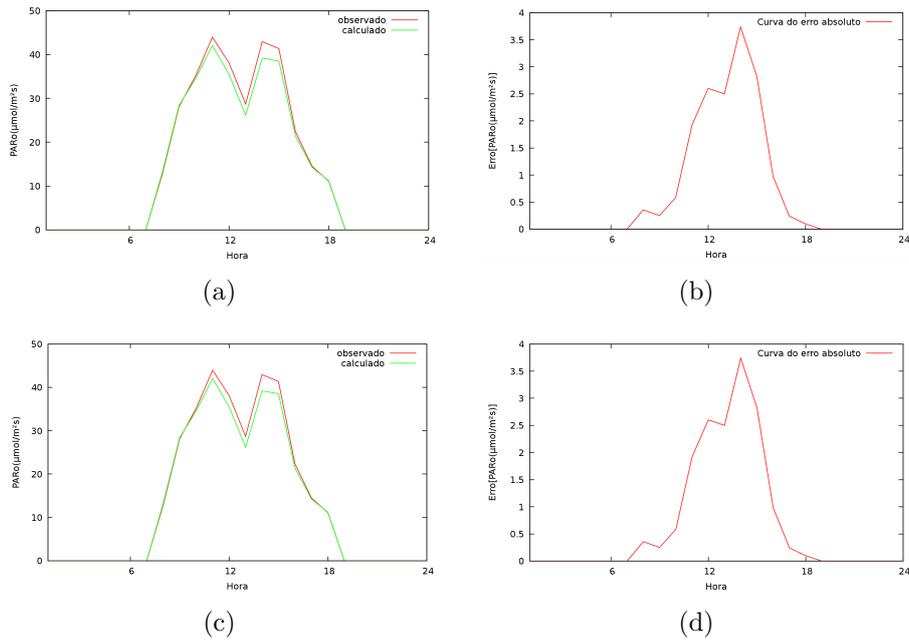


Figura 5.22 - (a) Resultados de PARo em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de PARo em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

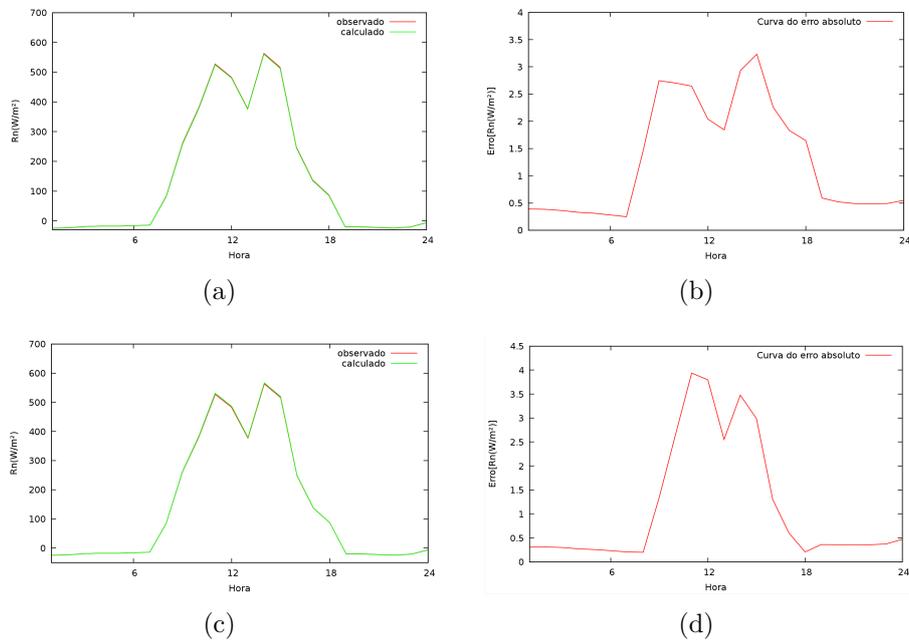


Figura 5.23 - (a) Resultados de Rn em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de Rn em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

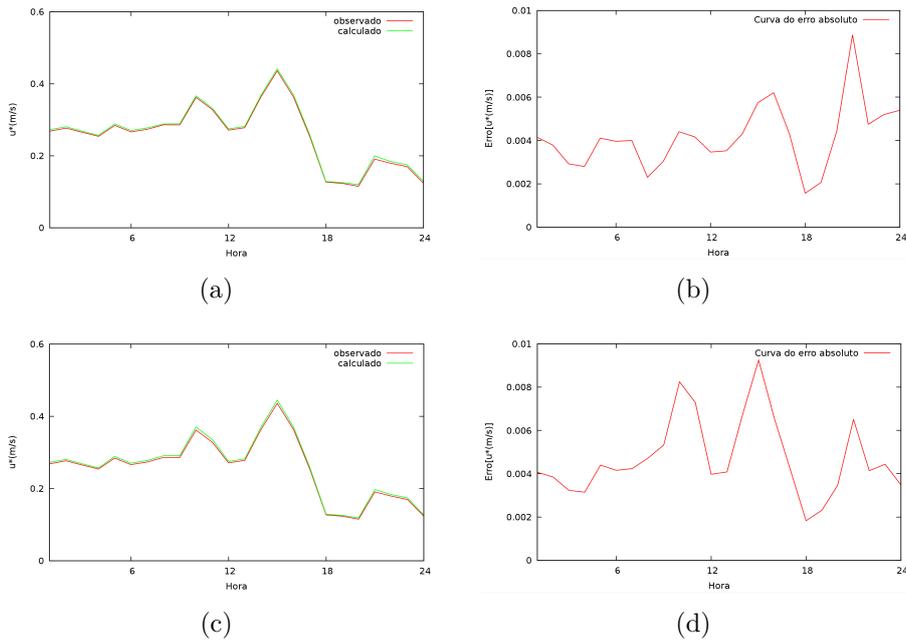


Figura 5.24 - (a) Resultados de u^* em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de u^* em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

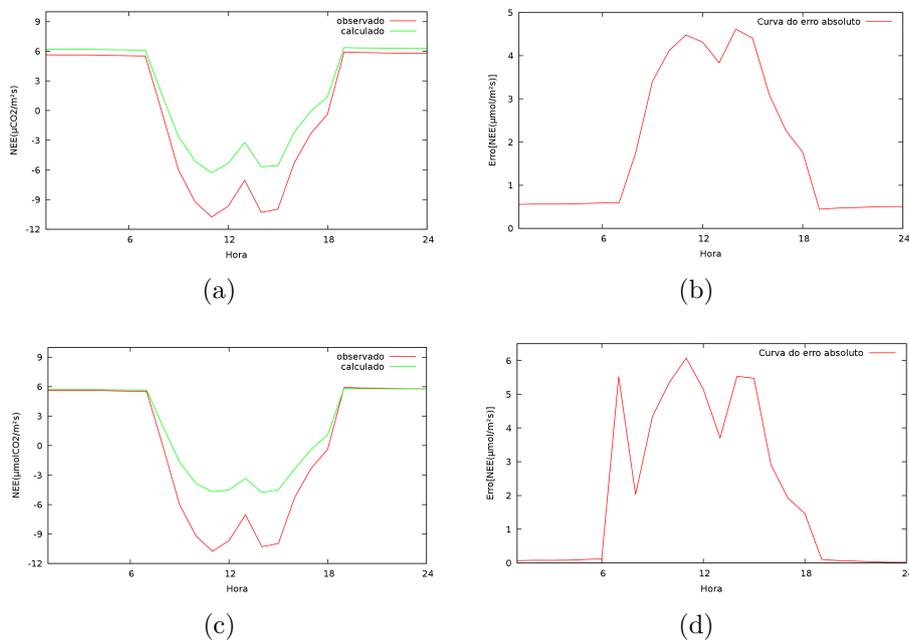


Figura 5.25 - (a) Resultados de NEE em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de NEE em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

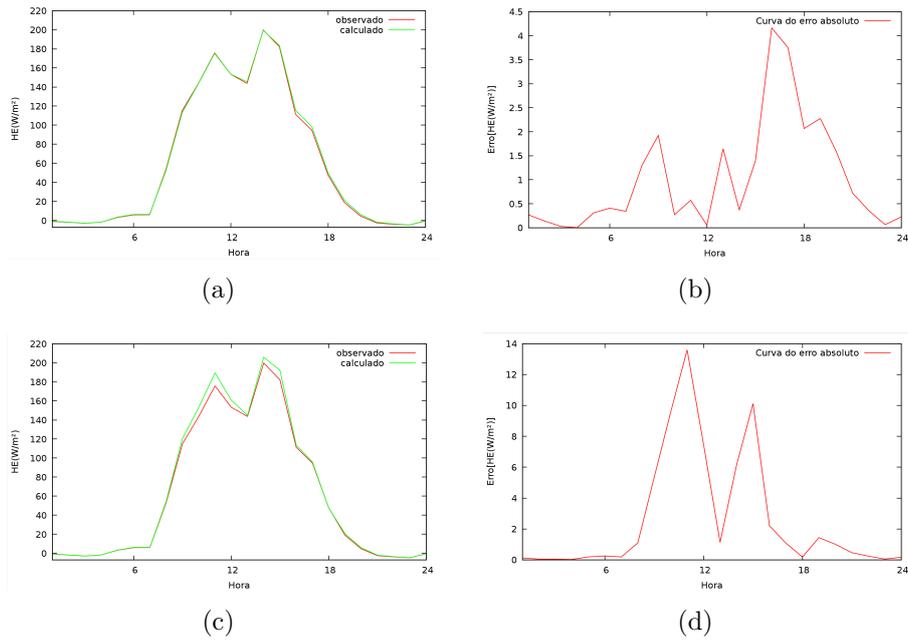


Figura 5.26 - (a) Resultados de HE em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de HE em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

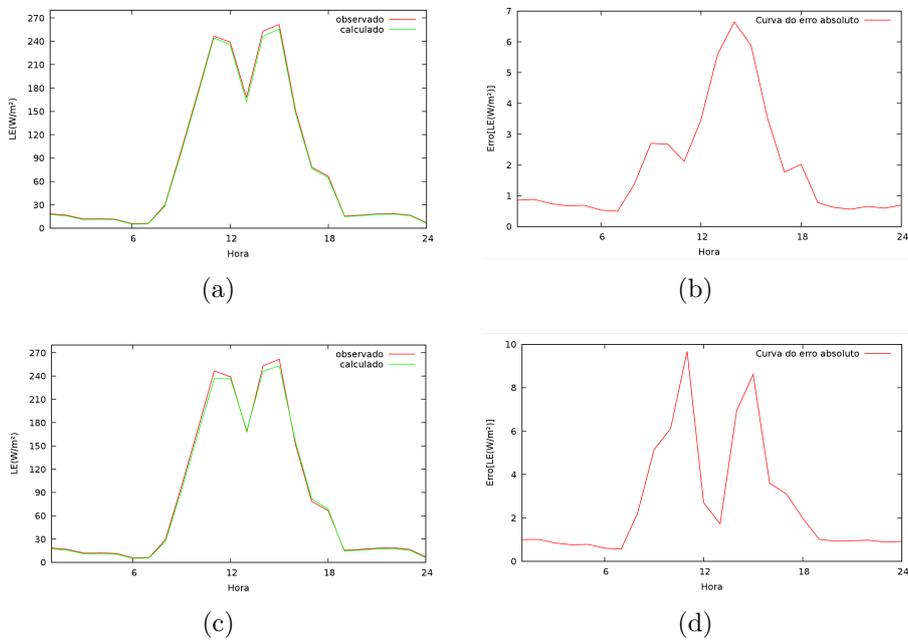


Figura 5.27 - (a) Resultados de LE em um dia típico - escala temporal (NSGA-II epidêmico); (b) Erro absoluto entre observado e calculado (NSGA-II epidêmico); (c) Resultados de LE em um dia típico - escala temporal (NSGA-II); (d) Erro absoluto entre observado e calculado (NSGA-II).

6 CONCLUSÕES

O desenvolvimento dos computadores, cada vez mais rápidos e mais potentes, tem tornado possível o surgimento e o aprimoramento de várias técnicas computacionais para se resolver um importante tipo de problema inverso: *estimação de parâmetros de alta ordem*. Tal problema surge ao se modelar matematicamente um determinado fenômeno, quando se torna necessário o ajuste de um número finito de parâmetros do modelo de modo a garantir a aderência do modelo ao fenômeno estudado. Tal ajuste é conhecido como *calibração de parâmetros*.

As experiências práticas na calibração automática de modelos (calibração com o uso intensivo de computadores) têm mostrado que o uso de mais de uma função objetivo (métrica de avaliação dos erros entre dados reais e dados calculados pelo modelo) no processo de calibração permite uma melhor avaliação das incertezas, das imperfeições do modelo e da representatividade dos parâmetros.

A complexidade na calibração de um modelo está associada a três principais fatores: o número de parâmetros, o número de funções objetivo e as características das funções objetivo.

Este trabalho teve como propósito tratar da calibração multiobjetivo dos parâmetros de dois modelos de fluidos geofísicos: o modelo hidrológico de pequenas bacias IPH-II e o modelo de superfície atmosférico IBIS.

Para a calibração multiobjetivo do modelo IPH-II, foi implementado um algoritmo genético com epidemia, denominado AGE, desenvolvido como uma variante do algoritmo evolucionário MOCOM, o qual já vem sendo usado na calibração dos parâmetros do IPH-II.

Para a calibração do modelo IBIS, foi utilizado o software de calibração multiobjetivo do IBIS, denominado Optis, que usa como técnica de calibração o algoritmo de otimização multiobjetivo NSGA-II. Foram, aqui, propostas algumas mudanças nas definições das funções objetivo, uma estratégia de calibração em uma escala temporal das variáveis do IBIS, baseada na análise de sensibilidade de seus parâmetros através do método de Morris, e a implementação de um operador de epidemia no NSGA-II.

Para os testes de calibração com os dois modelos, foram consideradas séries sintéticas de dados observados.

Na calibração do IPH-II, o AGE mostrou-se eficiente na recuperação dos valores

pré-fixados dos parâmetros usados para gerar a série sintética de vazões e também para obter uma boa aproximação da frente de Pareto.

Na calibração do IBIS, as estratégias propostas mostraram-se, também, eficazes na recuperação dos valores pré-fixados dos parâmetros e na obtenção da frente de Pareto..

De modo geral, este trabalho apresentou contribuições na calibração multiobjetivo com o uso do operador de epidemia, com aplicações no modelo hidrológico IPH-II (MOCOM+AGE) e no modelo de superfície IBIS – onde a estratégia OPTIS (VAREJÃO, 2009; VAREJÃO et al., 2013) (NSGA-II+ NSGA-II com epidemia) foi utilizada. O uso da análise de sensibilidade em escala temporal para a calibração dos parâmetros do IBIS mostrou-se uma estratégia potencial para aplicações em previsão de tempo de curto e médio prazos e previsão climática sazonal.

Como sugestões de futuros trabalhos, colocam-se:

- Considerar o uso de outras métricas na definição das funções objetivo na calibração do IBIS, testando-se, inclusive, o uso de outras normas;
- Fazer mais testes de calibração do IBIS, considerando o NSGA-II com epidemia com o objetivo de comparar o seu desempenho com o do NSGA-II;
- Implementar outros operadores de epidemia no NSGA-II;
- Usar o algoritmo AGE na calibração automática multiobjetivo de modelos hidrológicos de grandes bacias;
- Utilização de outras meta-heurísticas.
- Usar a estratégia de análise de sensibilidade em escala temporal para a calibração de modelos climáticos, com integração de longo prazo, como feito em Minjiao e Xiao (2014).

REFERÊNCIAS BIBLIOGRÁFICAS

ARAÚJO, A. S.; VELHO, H. F. C.; GOMES, V. C. F. Calibrating an hydrological model by an evolutionary strategy for multi-objective optimization. **Inverse Problems in Science and Engineering**, v. 21, n. 3, p. 438–450, 2013. Disponível em: <<http://dx.doi.org/10.1080/17415977.2012.712530>>. 4

BERTHELOT, R. **Curso de hidrologia sintética**. Porto Alegre - RS: Editora UFRGS, 1970. Apostila do Curso de Mestrado em Hidrologia Aplicada - IPH-UFRGS. 7

BRAVO, J. M.; ALLASIA, D. G.; COLLISCHONN, W.; TASSI, R.; MELLER, A.; TUCCI, C. E. M. Avaliação visual e numérica da calibração do modelo iph ii com fins educacionais. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS-ABRH, 17., 2007, São Paulo - SP. **Anais...** São Paulo - SP: ABRH, 2007. 10

BRAVO, J. M.; COLLISCHONN, W.; TUCCI, C. E. M. Verificação de eficiência de um algoritmo evolucionário multiobjetivo na calibração automática do modelo hidrológico iph ii. **Revista Brasileira de Recursos Hídricos**, v. 14, n. 3, p. 37–60, 2009. 44, 47, 48

CASTRO, R. E. **Otimização de estruturas com multiobjetivos via algoritmo genético de Pareto**. 2001. 202 p. Tese (Doutorado em Engenharia Civil) — Coordenação dos Programas de Pós-Graduação em Engenharia-COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001. 31, 34

CHIWIAKOWSKY, L.; VELHO, H. F. C. Different approaches for the solution of a backward heat conduction problem. **Inverse Problems in Science and Engineering**, v. 11, n. 6, p. 471–494, 2003. 34

CLARKE, R. T. **Mathematical models in hydrology (Irrigation and drainage paper)**. Roma: Food an Agricultural Organization, 1973. 282 p. 9

COELLO, C. A. C. **An empirical study of evolutionary techniques for multi-objective optimization in engineering design**. 1996. 488 p. Tese (Doctoral Dissertation on Computer Science) — Department of Computer Science - Tulane University, New Orleans(LA)-USA, 1996. 32

COLLATZ, G. J.; BALL, J. T.; GRIVET, C.; BERRY, J. A. Physiological and environmental regulation of stomatal conductance, photosynthesis and

transpiration: a model that includes a laminar boundary layer. **Agricultural and Forest Meteorology**, v. 54, n. 2-4, p. 107–136, 1991. ISSN 0168-1923. Disponível em:

<<http://www.sciencedirect.com/science/article/pii/0168192391900028>>.

15

COLLISCHONN, W.; TUCCI, C. E. M. Ajuste multiobjetivo dos parâmetros de um modelo hidrológico. **Revista Brasileira de Recursos Hídricos**, v. 8, n. 3, p. 27–41, 2003. 8

DEB, K. **Multi-objective optimization using evolutionary algorithms**.

Chichester-UK: John Wiley & Sons, 2004. 515 p. 4, 32, 35

DEB, K.; PRATAP, A.; AGARWAL, S.; MEYARIVAN, T. A fast elitist multi-objective genetic algorithm: Nsga-ii. **IEEE Transactions on Evolutionary Computation**, v. 6, p. 182–197, 2000. 4, 32, 35, 40, 41

DUAN, Q.; GUPTA, V.; SOROOSHIAN, S. Shuffled complex evolution approach for effective and efficient global minimization. **Journal of Optimization Theory and Applications**, Kluwer Academic Publishers-Plenum Publishers, v. 76, n. 3, p. 501–521, 1993. ISSN 0022-3239. Disponível em:

<<http://dx.doi.org/10.1007/BF00939380>>. 2, 3

FARQUHAR, G. D.; CAEMMERER, S. V.; BERRY, J. A. A biochemical model of photosynthetic co₂ assimilation in leaves of c₃ species. **Planta**, v. 149, n. 1, p. 78–90, 1980. 15

FOLEY, J.; PRENTICE, I. C.; RAMANKUTTY, N.; LEVIS, S.; POLLARD, D.; SITCH, S.; HAXELTINE, A. An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. **Global Biogeochem. Cycles**, v. 10, n. 4, p. 603–628, 1996. 4, 10

FONSECA, C. M.; FLEMING, P. J. Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In: INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 5., 1993, San Francisco, CA, USA. **Proceedings...** San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1993. p. 416 – 423. 4, 32, 35

GOLDBERG, D. E. **Genetic algorithms in search, optimization and machine learning**. Boston-USA: Addison-Wesley Pub. Co., 1989. 432 p. 33, 34, 36

HADAMARD, J. **Lectures on Cauchy's Problem in Linear Partial Differential Equations**, by Jacques Hadamard. [S.l.]: Yale University Press, 1923. 316 p. 21

HOLLAND, J. H. **Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence**. [S.l.]: The University of Michigan Press, Ann Arbor, 1975. 183 p. 32, 34

HORN, J.; NAFFLOITS, N.; GOLDBERG, D. A niched pareto genetic algorithm for multi-objective optimization. In: IEE WORLD CONGRESS ON COMPUTATIONAL INTELIGENCE, FIRST IEE CONFERENCE ON EVOLUTIONARY COMPUTATION, 1994, Orlando - USA. **Proceedings...** Orlando - USA: IEE, 1994. 35

JACKSON, R. B.; MOONEY, H. A.; SCHULZE, E. D. A global budget for fine root biomass, surface area, and nutrient contents. **The National Academy of Sciences of the USA**, v. 94, n. 14, p. 7362–7366, 1997. 18

JOHNSTON, P. R.; PILGRIM, D. H. Parameter optimization for watershed models. **Water Resources Research**, v. 12, n. 3, p. 477–486, 1976. ISSN 1944-7973. Disponível em: <<http://dx.doi.org/10.1029/WR012i003p00477>>. 2

KUCHARIK, C. J.; FOLEY, J. A.; DELIRE, C.; FISHER, V. A.; COE, M. T.; LENTERS, J. D.; YOUNG-MOLLING, C.; RAMANKUTTYA, N.; NORMAN, J. M.; GOWER, S. T. Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure. **Global Biogeochem. Cycles**, v. 14, n. 3, p. 795–825, 2000. 10, 12, 16

LOPES, P. S. **Modelagem de problema inverso de detecção de danos por técnicas de parâmetros e de otimização**. 2010. 136 p. Tese (Doutorado em Engenharia Mecânica) — Instituto de Engenharia Mecânica-Universidade Federal de Itajubá, Itajubá, 2010. 36, 41, 42, 43

MEDEIROS, F. L. L. **Algoritmo genético híbrido como um método de busca de estados estacionários de sistemas dinâmicos**. 2002. 191 p. (INPE-13913-TDI/1058). Dissertação (Mestrado em Computação Aplicada) — Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2002. 34

MICHALEWICZ, Z. **Genetic Algorithms + Data Structures = Evolution Programs**. London-UK: Springer, 1996. 338 p. 34

MINJIAO, L.; XIAO, L. Time scale dependent sensitivities of the xinjiang model parameters. **Hydrological Research Letters**, v. 8, n. 1, p. 51–56, 2014. 4, 64, 82

MORRIS, M. D. Factorial sampling plans for preliminary computational experiments. **Technometrics**, American Society for Quality Control and American Statistical Association, Alexandria, Va, USA, v. 33, n. 2, p. 161–174, abr. 1991. ISSN 0040-1706. Disponível em: <<http://dx.doi.org/10.2307/1269043>>. 24

NASH, J.; SUTCLIFFE, J. River flow forecasting through conceptual models part i - a discussion of principles. **Journal of Hydrology**, v. 10, n. 3, p. 282–290, 1970. 2

NELDER, J. A.; MEAD, R. A simplex method for function minimization. **Computer Journal**, v. 7, p. 308–313, 1965. 37, 39

PICKUP, G. Testing the efficiency of algorithms and strategies for automatic calibration of rainfall-runoff models. **Hydrological Sciences Bulletin**, v. 22, n. 2, p. 257–274, 1977. Disponível em: <<http://dx.doi.org/10.1080/02626667709491716>>. 2

SALTELLI, A.; CHAN, K.; S.; SCOTT, E. M. **Sensitivity Analysis**. Chichester-UK: John Wiley & Sons, 2009. 494 p. 23, 89

SALTELLI, A.; RATTO, M.; TARANTOLA, S.; CAMPOLONGO, F. Sensitivity analysis for chemical models. **Chemical Review**, v. 105, n. 7, p. 2811–2828, 2005. 23, 24

SCHAFFER, J. D. Multiple objective optimization with vector evaluated genetic algorithms. In: 1ST INTERNATIONAL CONFERENCE ON GENETIC ALGORITHMS, 1985. **Proceedings...** Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 1985. p. 93–100. Disponível em: <<http://dl.acm.org/citation.cfm?id=645511.657079>>. 29, 35

SIVANANDAM, S. N.; DEEPA, S. N. **Introduction to genetic algorithm**. New York-USA: Springer, 2008. 442 p. 34

SRINIVAS, N.; DEB, K. Multiobjective optimization using nondominated sorting in genetic algorithms. **Evolutionary Computation**, MIT Press, Cambridge, MA, USA, v. 2, n. 3, p. 221–248, 1994. ISSN 1063-6560. Disponível em: <<http://dx.doi.org/10.1162/evco.1994.2.3.221>>. 4, 32, 35, 40

TIKHONOV, A.; ARSENIN, V. **Solutions of ill-posed problems**. Winston, 1977. (Scripta series in mathematics). ISBN 9780470991244. Disponível em: <<http://books.google.com.br/books?id=ECrvAAAAMAAJ>>. 21

TUCCI, C. E. M. **Modelos hidrológicos**. Porto Alegre: Editora UFRGS, 2005. 669 p. 4, 7, 8

TUCCI, C. E. M.; CAMPANA, N. Simulação distribuída com o iph ii, nova versão. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS E SIMPÓSIO DE RECURSOS HÍDRICOS DO CONE SUL - ABRH, 1., 1993, Gramado - RS. **Anais...** Gramado-RS: ABRH, 1993. v. 3, p. 494–495. 7

TUCCI, C. E. M.; ORDONEZ, J. S.; SIMÕES, M. L. Modelo matemático precipitação-vazão iph ii - alguns resultados. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS-ABRH, 4., 1981, Fortaleza - CE. **Anais...** Fortaleza - CE: ABRH, 1981. 7

VAREJÃO, C. G. **Calibração hierárquica multi-objetivo de um modelo de micrometeorologia e de dinâmica de ecossistemas terrestres**. 2009. 118 p. Dissertação (Mestrado em Meteorologia Agrícola) — Departamento de Engenharia Agrícola - Universidade Federal de Viçosa, Viçosa, 2009. 4, 12, 23, 25, 45, 53, 55, 82, 89

VAREJÃO, C. M.; COSTA, M. H.; CAMARGOS, C. C. S. A multi-objective hierarchical calibration procedure for land surface-ecosystem models. **Inverse Problems in Science and Engineering**, v. 21, n. 3, p. 357–386, 2013. 4, 45, 53, 55, 82

VELHO, H. F. C. **Problemas inversos em pesquisa espacial**. São Carlos, SP: Sociedade Brasileira de Matemática Aplicada e Computacional (SBMAC), 2008. 124 p. 22

YAPO, P. O.; GUPTA, H. V.; SOROOSHIAN, S. Multi-objective global optimization for hydrologic models. **Journal of Hydrology**, v. 204, n. 1-4, p. 83 – 97, 1998. ISSN 0022-1694. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0022169497001078>>. 4, 30, 36

ZINI, E. O. C. **Algoritmo genético especializado na resolução de problemas com variáveis contínuas e altamente restritos**. 2009. 151 p. Dissertação (Mestrado em Engenharia Elétrica) — Faculdade de Engenharia de Ilha Solteira - UNESP, Ilha Solteira, 2009. 27, 28, 29, 32

APÊNDICE - A Análise de Sensibilidade para a Calibração em Escala Temporal

Para a análise de sensibilidade em escala temporal, foi utilizada uma implementação do método de Morris feita por Varejão (2009). Nesta implementação, foram consideradas as seguintes modificações.

No cálculo dos efeitos elementares, a equação 3.4 foi assim modificada:

$$d_i(x) = \frac{y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - y(x)}{\bar{P}} \quad (\text{A.1})$$

onde \bar{P} corresponde ao valor médio da saída y ao variar os parâmetros nos seus intervalos. Com este cálculo, d_i descreve a influência do i -ésimo parâmetro em relação ao valor médio da saída y .

Considerou-se também uma modificação no cálculo da matriz de orientação B^* , propondo-se uma equação pequena modificação na Equação 3.2:

$$B^* = (J_{k+1,1}x^* + \frac{\Delta}{2}[(2BP^* - J_{k+1,k})D^* + J_{k+1,k}]) \quad (\text{A.2})$$

O objetivo com esta modificação foi o de simplificar o processo de permutação dos parâmetros (em termos de escalas, intervalos e unidades) (VAREJÃO, 2009).

Considerou-se $r = 4$, ou seja, para cada parâmetro foram calculados 4 efeitos elementares, obtendo-se 4 matrizes de orientação B^* , como sugerido por Saltelli et al. (2009), correspondentes a 4 trajetórias no espaço de parâmetros.

Obtidas as medidas de sensibilidade, μ_i^* e σ_i , a identificação dos parâmetros mais importantes para cada saída do modelo foi feita, considerando-se um *índice de influência* I , definido como a soma das duas medidas de sensibilidade:

$$I_i = \mu_i^* + \sigma_i \quad (\text{A.3})$$

Este índice I leva em conta os efeitos gerais de cada parâmetro nas saídas do modelo bem como os seus efeitos não lineares ou de interação com os demais, sendo, portanto, uma boa medida da importância de cada parâmetro no processo de calibração.

As tabelas seguintes mostram os resultados dos cálculos de I em cada uma das

escalas temporais, considerando somente os parâmetros para os quais $I \neq 0$, listados em ordem crescente em relação aos valores de I .

Tabela A.1 - Parâmetros com índice de influência não-nulo na escala mensal.

Mensal	Parâmetro	I
	<i>wsoi_coef</i>	0.037
	<i>tempvm_coef</i>	0.032
	<i>rroot_coef</i>	0.019
	<i>rgrowth_coef</i>	0.06
	<i>funcb_coef</i>	0.002
	<i>tauveg_NIR</i>	0.0011
	<i>rwood_coef</i>	0.001
	<i>tauveg_NIR</i>	0.0009
	<i>chifuz</i>	0.0008
	<i>rhoveg_NIR</i>	0.00045
	<i>stressf_coef</i>	0.00036
	<i>clitwm_coef</i>	0.00035
	<i>clitwl_coef</i>	0.00035
	<i>beta2</i>	0.0003
	<i>clitlm_coef</i>	0.00021
	<i>vmax_pft</i>	0.0001

Tabela A.2 - Parâmetros com índice de influência não-nulo na escala diária.

Diária	Parâmetro	I
	<i>rhoveg_NIR</i>	0.36
	<i>rhoveg_vis</i>	0.29
	<i>tauleaf</i>	0.08
	<i>aleaf</i>	0.049
	<i>tauveg_vis</i>	0.034
	<i>dispu_coef</i>	0.021
	<i>specla</i>	0.015
	<i>tauroot</i>	0.002
	<i>chu</i>	0.0018
	<i>tauveg_NIR</i>	0.0011
	<i>chifuz</i>	0.00010
	<i>vmax_pft</i>	0.00009
	<i>coefmub</i>	0.00004
	<i>beta2</i>	0.00003
	<i>tempvm_coef</i>	0.000016

Tabela A.3 - Parâmetros com índice de influência não-nulo na escala horária.

Horária	Parâmetro	<i>I</i>
	<i>rhoveg_vis</i>	0.69
	<i>chifuz</i>	0.43
	<i>tauveg_vis</i>	0.34
	<i>beta2</i>	0.26
	<i>tauveg_NIR</i>	0.11
	<i>clitrs_coef</i>	0.061
	<i>avmuir_coef</i>	0.05
	<i>coefmub</i>	0.048
	<i>clitls_coef</i>	0.032
	<i>vmax_pft</i>	0.026
	<i>clitws_coef</i>	0.023
	<i>dispu_coef</i>	0.21
	<i>kfactor</i>	0.021
	<i>chs</i>	0.013
	<i>stressf_coef</i>	0.0041
	<i>clitwm_coef</i>	0.0035
	<i>clitwl_coef</i>	0.0035
	<i>clitlm_coef</i>	0.0021
	<i>tempvm_coef</i>	0.0016
	<i>tauleaf</i>	0.0008
	<i>alogu_coef</i>	0.0006
	<i>rhoveg_NIR</i>	0.0004
	<i>rgrowth_coef</i>	0.0005
	<i>wsoi_coef</i>	0.00017
	<i>tauroot</i>	0.00013
	<i>specla</i>	0.00011
	<i>aleaf</i>	0.00011

A partir das tabelas acima, definiram-se os 3 grupos de parâmetros, um para cada nível de escala temporal, que são mostrados na tabela seguinte.

Tabela A.4 - Grupos de calibração em escala temporal.

Grupos	Escala	Parâmetros
<i>Grupo1</i>	Mensal	<i>funcb_coef</i> <i>rroot_coef</i> <i>rwood_coef</i> <i>rgrowth_coef</i> <i>tempvm_coef</i> <i>wsoi_coef</i>
<i>Grupo2</i>	Diária	<i>rhoveg_NIR</i> <i>chu</i> <i>tauleaf</i> <i>tauroot</i> <i>specla</i> <i>aleaf</i>
<i>Grupo3</i>	Horária	<i>rhoveg_vis</i> <i>tauveg_vis</i> <i>tauveg_NIR</i> <i>chifuz</i> <i>vmax_pft</i> <i>coefmub</i> <i>chs</i> <i>beta2</i> <i>stressf_coef</i> <i>clitlm_coef</i> <i>clitls_coef</i> <i>clitrs_coef</i> <i>clitwl_coef</i> <i>clitwm_coef</i> <i>clitws_coef</i> <i>kfactor</i> <i>dispu_coef</i> <i>avmuir_coef</i>

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.