

CLAUDECI GOMES VAREJÃO JÚNIOR

CALIBRAÇÃO HIERÁRQUICA MULTIOBJETIVO DE UM  
MODELO DE MICROMETEOROLOGIA E DE DINÂMICA DE  
ECOSSISTEMAS TERRESTRES

Dissertação apresentada à Universidade Federal  
de Viçosa, como parte das exigências do Programa  
de Pós-Graduação em Meteorologia Agrícola, para  
obtenção do título de *Magister Scientiae*.

VIÇOSA  
MINAS GERAIS - BRASIL  
2009

Ficha catalográfica preparada pela Seção de Catalogação e  
Classificação da Biblioteca Central da UFV

T

V292c  
2009

Varejão Júnior, Claudeci Gomes, 1981-  
Calibração hierárquica multiobjetivo de um modelo de  
micrometeorologia e de dinâmica de ecossistemas  
terrestres / Claudeci Gomes Varejão Júnior. – Viçosa, MG,  
2009.  
xx, 96f.: il. ; 29cm.

Inclui apêndice.

Orientador: Marcos Heil Costa.

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 88-92

1. Meteorologia agrícola. 2. Calibração. 3. Climatologia -  
Modelos matemáticos. 4. Ecossistemas. 5. Climatologia  
agrícola I. Universidade Federal de Viçosa. II. Título.


CDD 22.ed. 630.2515


CLAUDECI GOMES VAREJÃO JÚNIOR

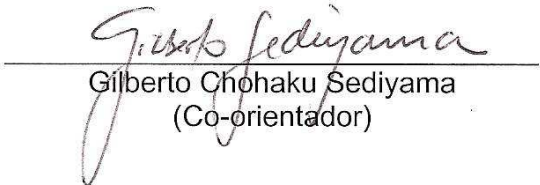
**CALIBRAÇÃO HIERÁRQUICA MULTIOBJETIVO DE UM MODELO DE  
MICROMETEOROLOGIA E DE DINÂMICA DE ECOSISTEMAS  
TERRESTRES**

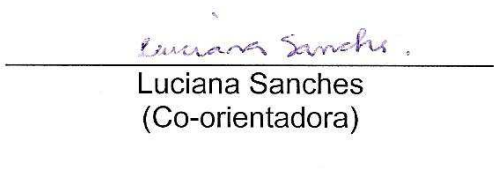
Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Meteorologia Agrícola, para obtenção do título de *Magister Scientiae*.

APROVADA: 07 de abril de 2009.

  
Flávio Barbosa Justino

  
Gilvan Sampaio de Oliveira

  
Gilberto Chohaku Sedyama  
(Co-orientador)

  
Luciana Sanches  
(Co-orientadora)

  
Marcos Heil Costa  
(Orientador)

*Aos meus pais e ao meu irmão, que foram  
o prêmio grande que a loteria da vida me deu.*

## AGRADECIMENTOS

À Universidade Federal de Viçosa (UFV) pela oportunidade de realizar o curso.

Ao Departamento de Engenharia Agrícola pelo suporte e pela oportunidade de realizar o curso.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos.

Ao professor José Luis Braga, pelo incentivo e apoio para ingressar nesta jornada.

Ao professor Vladimir Oliveira Di Iório, por estar sempre a disposição para esclarecer minhas dúvidas computacionais.

Ao meu colega de graduação, Gleidson Charles Botelho Baleeiro, por abrir as portas.

Ao meu orientador, Marcos Heil Costa, pela oportunidade de desenvolver este projeto, por acreditar em minha capacidade de vencer este desafio, por todo o conhecimento compartilhado e pela amizade.

Aos colegas do Grupo de Pesquisas, Silvia, Christiane, Lucía, Santiago, Edson, Mônica, Francisca, Cleverson, Márcia, Luciana, Marcos Paulo, Leydmere, Gabrielle, Fabrício e Thomé, pelo companheirismo e pelas contribuições no desenvolvimento deste trabalho.

À Hewlley Maria Acioli Imbuzeiro, por todo conhecimento e informações compartilhados.

À Letícia Braidá, pelo carinho e apoio imprescindível.

Ao projeto LBA, pelo fornecimento dos dados experimentais dos sítios estabelecidos na Floresta Nacional de Tapajós e na Fazenda Maracaí - Sinop.

À professora Luciana Sanches, pelo apoio na disponibilização dos dados do sítio de Sinop.

Ao professor Gilberto Sedyama, pelas contribuições e pelos valiosos ensinamentos em estatística.

Às secretárias Kelly Portugal e Graça Freitas, pelo suporte e dedicação.

Aos colegas da república, David Moreira Abreu e Rafael Aldighiere Moraes, pelo apoio e pelos momentos de descontração.

A Vicente e Maria, pelo carinho e amizade.

Ao amigo Levi Henrique Santana de Lelis, pelo apoio em todos os momentos e pelo incentivo para seguir em frente sempre.

Aos meus pais, Claudeci Gomes Varejão e Geralda Maria Vieira Varejão, pelo apoio e amor incondicional, e por terem me ensinado o que realmente importa na vida.

Ao meu grande amigo e irmão, Eduardo Vinícius Vieira Varejão, pelo incentivo, pelos preciosos conselhos e pelo exemplo de competência e dedicação.

À minha cunhada, Jodieh Santana Varejão, por todo o apoio e pela adorável convivência.

Ao meu amor, Luana Cristina Lima da Fonseca, por dar sentido a tudo.

## **BIOGRAFIA**

CLAUDECI GOMES VAREJÃO JÚNIOR, filho de Geralda Maria Vieira Varejão e Claudeci Gomes Varejão, nasceu em 13 de março de 1981, na cidade de Muriaé, Minas Gerais, Brasil.

Em outubro de 2006 concluiu o curso de graduação em Ciência da Computação pela Universidade Federal de Viçosa (UFV).

Em outubro de 2006 iniciou o curso de pós-graduação, nível mestrado, em Meteorologia Agrícola na Universidade Federal de Viçosa (UFV).

## SUMÁRIO

<b>LISTA DE TABELAS</b>	<b>viii</b>
<b>LISTA DE FIGURAS</b>	<b>x</b>
<b>LISTA DE SÍMBOLOS</b>	<b>xv</b>
<b>RESUMO</b>	<b>xvii</b>
<b>ABSTRACT</b>	<b>xix</b>
<b>1 INTRODUÇÃO</b>	<b>1</b>
<b>2 REVISÃO DE LITERATURA</b>	<b>6</b>
2.1 Análise de sensibilidade . . . . .	6
2.1.1 Tipos de análises de sensibilidade . . . . .	7
2.1.2 Método de Morris . . . . .	9
2.2 Avaliação de modelos . . . . .	10
2.2.1 Estatísticas e medidas de ajuste de modelos . . . . .	10
2.3 Otimização multiobjetivo . . . . .	13
2.4 Algoritmo genético multiobjetivo . . . . .	14
<b>3 METODOLOGIA</b>	<b>18</b>
3.1 O modelo IBIS . . . . .	18
3.2 Descrição dos sítios . . . . .	26
3.3 Dados observados . . . . .	28
3.3.1 Preenchimento de falhas dos dados . . . . .	29
3.3.2 Filtragem dos dados de avaliação do modelo . . . . .	30
3.4 Análise de sensibilidade do IBIS . . . . .	31
3.4.1 Implementação da AS . . . . .	33
3.4.2 Validação da AS . . . . .	36
3.5 Calibração multiobjetivo . . . . .	37



3.5.1	Estrutura do Optis . . . . .	39
3.5.2	Paralelização do algoritmo . . . . .	40
3.5.3	Teste do algoritmo de calibração . . . . .	40
3.6	Calibração hierárquica multiobjetivo . . . . .	42
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>46</b>
4.1	Análise de sensibilidade do IBIS . . . . .	46
4.1.1	Validação da AS . . . . .	51
4.2	Calibração multiobjetivo . . . . .	53
4.2.1	Paralelização do algoritmo . . . . .	53
4.2.2	Teste do algoritmo de calibração . . . . .	53
4.3	Calibração mono-objetivo – Flona do Tapajós km 67 . . . . .	56
4.4	Calibração hierárquica multiobjetivo – Flona do Tapajós km 67 . . . . .	61
4.5	Discussão – Flona do Tapajós km 67 . . . . .	68
4.6	Calibração mono-objetivo – Sinop . . . . .	69
4.7	Calibração hierárquica multiobjetivo – Sinop . . . . .	76
4.8	Discussão – Sinop . . . . .	83
<b>5</b>	<b>CONCLUSÕES E RECOMENDAÇÕES</b>	<b>84</b>
	<b>REFERÊNCIAS</b>	<b>88</b>
	<b>APÊNDICES</b>	<b>93</b>
	<b>APÊNDICE A</b>	<b>94</b>
	<b>APÊNDICE B</b>	<b>96</b>

## LISTA DE TABELAS

1	Variáveis simuladas pelo IBIS analisadas. . . . .	32
2	Parâmetros do IBIS analisados. . . . .	33
3	Parâmetros do modelo calibrados nos testes dos grupos A e B. . .	37
4	Funções objetivo e os parâmetros calibrados em cada nível hierárquico, nos sítios da Flona do Tapajós e de Sinop. Para o sítio de Sinop, no 5º nível hierárquico apenas o LAI é calibrado. . . . .	44
5	Resultados da validação do método de AS proposto, com o número de parâmetros calibrados e o MAE de cada calibração. . . . .	51
6	Resultados da paralelização do algoritmo de otimização. Comparação entre os tempos gastos por simulação do modelo e da eficiência da paralelização. . . . .	53
7	MAE e <i>Bmax</i> dos dados de LE simulados e observados e a razão mono-objetivo/multiobjetivo utilizada para o cálculo do índice D. Valores de MAE calculados com o modelo não calibrado, após as calibrações mono-objetivo e após a calibração multiobjetivo. . . . .	54
8	MAE dos dados simulados e observados de H e LE, e a razão mono-objetivo/multiobjetivo utilizada para o cálculo do índice D. Valores de MAE calculados com o modelo não calibrado, após as calibrações mono-objetivo e após a calibração multiobjetivo. . . . .	54
9	Resultados da calibração hierárquica, para o sítio da Flona do Tapajós km 67. Valores de MAE e <i>Bmax</i> obtidos no procedimento hierárquico multiobjetivo comparados com os respectivos valores de referência. . . . .	62
10	Valores dos parâmetros obtidos no procedimento de calibração hierárquica para o sítio da Flona do Tapajós km 67. Os objetivos foram otimizados em ordem crescente do nível hierárquico, dos processos mais rápidos para os mais lentos. . . . .	63

11	Resultados da calibração hierárquica para o sítio de Sinop. Valores de MAE e <i>Bmax</i> obtidos no procedimento hierárquico multiobjetivo comparados com os respectivos valores de referência. . . . .	76
12	Valores dos parâmetros obtidos no procedimento de calibração hierárquica, para o sítio de Sinop. Os objetivos foram otimizados em ordem crescente do nível hierárquico, dos processos mais rápidos para os mais lentos. . . . .	77

## LISTA DE FIGURAS

1	Gráficos cumulativos de LE de duas simulações distintas (a) e (b) do modelo IBIS ( <i>Imbuzeiro, 2005</i> ) . . . . .	12
2	Mapeamento dos parâmetros em um problema de otimização multiobjetivo. Adaptado de <i>Veldhuizen e Lamont (2000)</i> . . . . .	14
3	Problema de otimização multiobjetivo: (a) funções objetivos f e g e (b) fronteira de Pareto (linha em negrito) no espaço dos objetivos	15
4	Esquema do funcionamento dos operadores genéticos de um AG. .	16
5	Aproximação da fronteira de Pareto por um AG multiobjetivo com dois objetivos. Cada gráfico representa uma geração, cada eixo do gráfico representa um objetivo e cada ponto nos gráficos representa uma solução potencial para o problema (ou um indivíduo). . . . .	17
6	Estrutura em módulos do modelo IBIS. . . . .	19
7	Localização dos sítios micrometeorológicos da Flona do Tapajós e Sinop. Adaptado de <i>Costa et al. (2007a)</i> . . . . .	27
8	Dois exemplos de trajetórias no espaço dos parâmetros quando (a) $k=2$ e (b) $k=3$ . Adaptado de <i>Saltelli et al. (2004)</i> . . . . .	34
9	Exemplo da escolha da solução ótima de um problema de otimização com 2 objetivos. Nesta tese, a solução com menor $d$ é considerada a melhor solução. . . . .	39
10	Análise de sensibilidade da $PAR_o$ aos 43 parâmetros do modelo. .	47
11	Análise de sensibilidade da $f_{APAR}$ aos 43 parâmetros do modelo. .	47
12	Análise de sensibilidade do $R_n$ aos 43 parâmetros do modelo. . . .	48
13	Análise de sensibilidade do $u^*$ aos 43 parâmetros do modelo. . . .	48
14	Análise de sensibilidade do H aos 43 parâmetros do modelo. . . .	49
15	Análise de sensibilidade do LE aos 43 parâmetros do modelo. . . .	49
16	Análise de sensibilidade do NEE aos 43 parâmetros do modelo. . .	50
17	Análise de sensibilidade da NPP aos 43 parâmetros do modelo. . .	50
18	Análise de sensibilidade do LAI aos 43 parâmetros do modelo. . .	51

19	Quantidade de parâmetros calibrados antes (grupo A) e após (grupo B) a análise de sensibilidade. . . . .	52
20	Resultados de a) MAE e b) $Bmax$ , da calibração multiobjetivo $\{MAE_{LE}, Bmax_{LE}\}$ . . . . .	55
21	Resultados de a) $MAE_{LE}$ e b) $MAE_H$ da calibração multiobjetivo $\{MAE_{LE}, MAE_H\}$ . . . . .	55
22	Resultados de $PAR_o$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	57
23	Resultados de $f_{APAR}$ após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão. . . . .	57
24	Resultados de $Rn$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	58
25	Resultados de $u^*$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão. . . . .	58
26	Resultados de $H$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	59
27	Resultados de $LE$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	60

28	Resultados de NEE após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	60
29	Resultados de LAI após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.	61
30	Resultados de $PAR_o$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	64
31	Resultados de $f_{APAR}$ após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.	65
32	Resultados de $R_n$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	65
33	Resultados de $u^*$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (d) gráfico de dispersão. . . . .	66
34	Resultados de $H$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	66
35	Resultados de $LE$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	67
36	Resultados de NEE após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	67
37	Resultados de LAI após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão. .	68

38	Resultados de $PAR_o$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	72
39	Resultados de $f_{APAR}$ após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.	72
40	Resultados de $R_n$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	73
41	Resultados de $u^*$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão. . . . .	73
42	Resultados de $H$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	74
43	Resultados de $LE$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	74
44	Resultados de $NEE$ após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	75
45	Resultados de $LAI$ após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.	75

46	Resultados de $PAR_o$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão. . . . .	79
47	Resultados de $f_{APAR}$ após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.	79
48	Resultados de $R_n$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	80
49	Resultados de $u^*$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão. . . . .	80
50	Resultados de $H$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	81
51	Resultados de $LE$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	81
52	Resultados de $NEE$ após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.	82
53	Resultados de $LAI$ após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão. .	82



## LISTA DE SÍMBOLOS

$a$	Fração da PAR absorvida pelo dossel
$A_g$	Fotossíntese bruta do dossel
AG	Algoritmo genético
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
AS	Análise de sensibilidade
beta2	Parâmetro relacionado a distribuição de raízes finas
Bmax	Viés máximo da soma cumulativa
chifuz	Fator de orientação da folha no dossel superior
chl	Capacidade térmica das folhas e galhos do dossel inferior
chs	Capacidade térmica dos galhos do dossel superior
chu	Capacidade térmica das folhas do dossel superior
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 m relacionado a condutância estomática
csoipas_coef	Carbono inicial no solo - húmus passivo
csoislon_coef	Carbono inicial no solo - húmus não-protégido lento
csoislop_coef	Carbono inicial no solo - húmus protégido lento
D	Índice de desempenho relativo
$d_i$	Efeito elementar do parâmetro
dispu_coef	Altura de deslocamento do plano zero para o dossel superior
$f_{APAR}$	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
G	Fluxo de calor no solo
H	Fluxo de calor sensível
I	Índice de sensibilidade do parâmetro
IBIS	Integrated Biosphere Simulator
kfactor	Fator de multiplicação das constantes de declínio dos reservatórios de C
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
LSM	Modelos de superfície terrestre
MAE	Erro absoluto médio
NEE	Troca líquida do ecossistema
NIR	Infravermelho próximo

NPP	Produção primária líquida
NSGA-II	Algoritmo genético de classificação não dominada (versão 2)
O	dados observados em campo
P	dados simulados pelo modelo
PAR <sub>o</sub>	Radiação fotossinteticamente ativa refletida
<i>r</i>	refletância
rhoveg_NIR	Refletância da folha no dossel superior (NIR)
rhoveg_vis	Refletância da folha no dossel superior (Visível)
RMSE	Raiz do erro quadrático médio
rgrowth_coef	Coefficiente de respiração de crescimento
Rn	Saldo de radiação
rroot_coef	Coefficiente de manutenção de respiração das raízes
rwood_coef	Coefficiente de manutenção de respiração dos troncos
SAI	Índice de área de galhos
<i>S<sub>in</sub></i>	Fluxo de radiação solar incidente
specla	Área foliar específica
stressf_coef	Coefficiente relacionado ao stress hídrico do solo
tauleaf	Constante de tempo de retorno da biomassa de folhas
tauroot	Constante de tempo de retorno da biomassa de raízes finas
tauwood0	Constante de tempo de retorno da biomassa de madeira
tauveg_NIR	Transmitância da folha no dossel superior (NIR)
tauveg_vis	Transmitância da folha no dossel superior (Visível)
tempvm_coef	Parâmetro de estresse térmico do Vmax
<i>u</i> *	Velocidade de fricção do vento
vmax_pft	Capacidade de atividade máxima da enzima Rubisco
wsoi_coef	Umidade do solo inicial
$\mu$	Efeito médio do parâmetro na saída do modelo
$\mu^*$	Valor absoluto de $\mu$
$\sigma$	Efeito não-linear ou de interações dos parâmetros na saída do modelo

## RESUMO

VAREJÃO JÚNIOR, Claudeci Gomes, M.Sc., Universidade Federal de Viçosa, abril de 2009. **Calibração hierárquica multiobjetivo de um modelo de micrometeorologia e de dinâmica de ecossistemas terrestres.** Orientador: Marcos Heil Costa. Co-orientadores: Gilberto Chohaku Sedyama e Luciana Sanches.

Os ecossistemas terrestres são um componente chave no estudo do clima do planeta. Existem fortes evidências de que a superfície terrestre influencia o tempo e o clima em várias escalas de tempo, de segundos a milhões de anos. Os modelos de ecossistemas terrestres (LSMs, do inglês *Land Surface Models*) são ferramentas importantes para a pesquisa com simulações climáticas e para previsão do tempo, funcionando geralmente acoplados a modelos de circulação geral atmosférica, atuando como submodelos que calculam os fluxos através da interface biosfera-atmosfera. Para se obter boas simulações com os LSMs para determinada região, é preciso otimizar seus parâmetros e calibrar o modelo de acordo com os dados observados na região de estudo. O principal objetivo deste trabalho é desenvolver uma metodologia que permita calibrar todos os processos simulados pelo modelo Integrated Biosphere Simulator (IBIS). A metodologia se baseia na hierarquia dos sistemas ambientais para calibrar sucessivamente todas as variáveis simuladas pelo IBIS. A calibração ocorre em ordem crescente de hierarquia temporal, dos processos mais rápidos (fluxos radiativos) até os processos mais lentos (alocação de carbono). Foi proposto um índice de desempenho relativo,  $D$ , para avaliar as calibrações multiobjetivo e hierárquica do modelo em relação às calibrações mono-objetivo independentes de cada variável. As calibrações mono-objetivo são consideradas as melhores calibrações possíveis, e as comparações entre as duas calibrações fornecem um bom indicador da eficiência do método de calibração proposto. Através da calibração hierárquica de todas as variáveis simuladas, espera-se obter um modelo que simule de forma consistente todos os processos nele representados. Em primeiro lugar, foi desenvolvida uma análise de sensibilidade (AS) para selecionar, de forma eficiente, os parâmetros

mais importantes para cada saída do modelo. A AS é baseada no método de Morris, com modificações no cálculo do índice de sensibilidade para adequar o método ao uso com LSMs. Em seguida, foi desenvolvido um programa para calibração automática multiobjetivo do modelo IBIS, chamado Optis, baseado no algoritmo genético multiobjetivo NSGA-II (Nondominated Sorted Genetic Algorithm II). Esse algoritmo permite ajustar múltiplas estatísticas e várias saídas do modelo simultaneamente. Finalmente, foi proposta uma metodologia de calibração hierárquica para LSMs, que executa a calibração de todos os processos simulados pelo modelo. A AS foi eficiente ao selecionar os parâmetros mais importantes do modelo a serem considerados na calibração, eliminando do processo de busca aqueles parâmetros com influência desprezível nas saídas do modelo. O programa Optis foi eficaz em solucionar os problemas de calibração multiobjetivo nos testes desenvolvidos, atingindo mais de 90% do desempenho obtido na calibração mono-objetivo. O Optis também se mostrou eficiente no uso de recursos computacionais, com 100% de eficiência na paralelização do código. O procedimento hierárquico proposto foi testado em dois sítios da floresta Amazônia, Flona de Tapajós e Sinop. Em ambos os sítios, o procedimento hierárquico se mostrou eficiente em fazer uma calibração geral do modelo, atingindo um D maior que 0,8 em ambos os sítios. Essa metodologia de calibração permitiu uma calibração mais realista do modelo, pois todas as variáveis simuladas foram otimizadas.

## ABSTRACT

VAREJÃO JÚNIOR, Claudeci Gomes, M.Sc., Universidade Federal de Viçosa, April, 2009. **Multiobjective hierarchical calibration of a micrometeorological and dynamic ecosystem model.** Adviser: Marcos Heil Costa. Co-advisers: Gilberto Chohaku Sedyama and Luciana Sanches.

Terrestrial ecosystems are a key component for the study of Earth climate. There are evidences that land surface processes affect climate and weather in a range of time scales, from seconds to million of years. Land surface models (LSMs) are important tools for research on climatic simulations and weather prediction, generally working coupled to general circulation models, and acting as sub models calculating the fluxes throughout the biosphere-atmosphere interface. To obtain good LSMs simulations for a given region, it is necessary to optimize the parameters and calibrate the model against observed data for the region of study. The main objective of this work is to develop a calibration methodology that allow to calibrate all processes simulated by the model Integrated Biosphere Simulator (IBIS). The study is based on systems hierarchy and on the calibration of all IBIS simulated variables. The calibration is performed at temporal hierarchy order, from the fastest process (radiative fluxes) to the slowest process (carbon allocation). We proposed a relative efficiency index,  $D$ , to evaluate the multiobjective and hierarchical calibrations, based on single objective calibrations of each simulated variable. The single objective calibrations are considered the best possible ones, and are used as a reference to evaluate the proposed method. Initially, we develop a sensitivity analysis (SA) that efficiently select the most important parameters for each model output. The SA is based on the Morris method, with modifications in the sensitivity calculations, to adjust the method to work with LSMs. Then, we develop a software for multiobjective automatic calibration of the IBIS model, called Optis. Optis is based on the multiobjective genetic algorithm NSGA-II (Nondominated Sorted Genetic Algorithm II). This software allows the adjustment of multiple statistics and many model outputs

simultaneously. Finally, we proposed a hierarchical methodology for LSM calibration, that performs the calibration of all processes simulated by the model. The SA is efficient to select only the most important model parameters to be considered in calibration, thus eliminating out of the search process those parameters with low influence on model output. Optis was effective to solve all the multiobjective calibration tests performed, reaching more than 90% of the efficiency obtained by the single objective calibration. Optis also is efficient in the use of computational resources, with total efficiency in the code parallelization. The hierarchical procedure proposed was tested in two experimental sites in the Amazon rainforest, Flona Tapajós and Sinop. In both sites, the procedure demonstrated to be efficient in calibrating all the model processes, with D greater than 0.80 in both sites. This calibration methodology allows a more realistic calibration of the model, as all simulated variables are optimized.

## 1 INTRODUÇÃO

Os ecossistemas terrestres são um componente chave no estudo do clima do planeta. Existem fortes evidências de que a superfície terrestre influencia o tempo e o clima em várias escalas de tempo, de segundos a milhões de anos (*Sellers et al.*, 1997; *Pitman*, 2003; *Costa et al.*, 2007b; *Sampaio et al.*, 2007).

Os modelos de superfície terrestre (LSMs, do inglês *Land Surface Models*) simulam os diversos processos dos ecossistemas terrestres, e são ferramentas importantes para a pesquisa com simulações climáticas e para previsão do tempo. Esses modelos funcionam geralmente acoplados a modelos de circulação geral da atmosfera, atuando como submodelos que calculam processos micrometeorológicos como os fluxos de radiação, calor, vapor d'água e momentum através da interface biosfera-atmosfera.

Os LSMs tendem a se tornarem mais complexos à medida que aumentam o conhecimento sobre o funcionamento dos ecossistemas terrestres e a capacidade de processamento dos computadores. Esse aumento da base de conhecimento e da capacidade de processamento permite que novas funcionalidades sejam incorporadas ao modelo, além de permitir que haja menos simplificação dos processos modelados. Nesse sentido, os modelos estão evoluindo para incorporar, além dos

processos micrometeorológicos, também os processos de dinâmica de ecossistemas terrestres.

Nos modelos de ecossistemas terrestres, muitos processos são representados fazendo uso de vários parâmetros. Alguns desses parâmetros devem ser calibrados por meio de procedimentos de otimização, de modo que as diferenças entre os resultados do modelo e os dados observados sejam minimizadas. Cada variável simulada pelo modelo representa pelo menos um objetivo do problema de otimização. Dessa forma, o problema de calibração de um modelo com várias variáveis de saída se resume a encontrar um conjunto de parâmetros que minimize simultaneamente um vetor de funções objetivos, i.e,

$$\min_{\theta \in \Theta} F(\theta) = \{f_1(\theta), \dots, f_m(\theta)\} \quad (1)$$

onde  $f_1(\theta), \dots, f_m(\theta)$  são as  $m$  diferentes funções objetivo a serem simultaneamente minimizadas (geralmente são medidas de diferença entre observado (O) e simulado (P)),  $\theta$  são os parâmetros do modelo e  $\Theta$  é o conjunto de restrições, que pode ser fornecido explicitamente ou definido implicitamente (*Fu et al.*, 2005).

A escolha da medida de ajuste entre P e O para a calibração de um modelo se mostra muito importante, aparentemente até mais importante do que a escolha do método de otimização (*Trudinger et al.*, 2007). É a medida de ajuste que definirá uma função objetivo do problema e deverá fornecer um valor escalar único indicando a qualidade dos dados simulados em relação aos dados observados. Para obter um resultado confiável na calibração, pode ser necessário otimizar mais de uma medida de ajuste simultaneamente. Nesse caso, a otimização de uma única variável de saída do modelo já define um problema de otimização multiobjetivo, com cada medida de ajuste representando uma função objetivo.

A calibração de um modelo complexo, envolvendo muitos parâmetros e múltiplas funções objetivos, pode ser particularmente difícil de resolver exata e eficientemente (*Nash e Sutcliffe*, 1970; *Johnston e Pilgrim*, 1976; *Pickup*, 1977; *Duan et al.*, 1993). O grau de dificuldade de um problema de otimização 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, podem haver diversas regiões de atração, sendo que cada uma pode conter vários mínimos locais (*Duan et al.*, 1992, 1993).
3. A quantidade de funções objetivos. É comum a necessidade de se calibrar um modelo levando em consideração mais de uma variável de saída ou, mesmo considerando apenas uma variável, utilizando mais de uma estatística para avaliação dos resultados. Essa característica multiobjetivo torna o processo de calibração muito mais complicado, principalmente porque muitas vezes os objetivos são conflitantes – a diminuição de um objetivo conduz a um aumento de outro (*Gupta e Sorooshian*, 1998).

A dificuldade apresentada no item (1) pode ser amenizada através de uma análise de sensibilidade (AS) efetiva. Utilizar apenas os parâmetros mais importantes detectados na AS reduz o espaço de busca e permite uma otimização mais eficiente. Para resolver o problema descrito no item (2) é necessário utilizar técnicas de busca globais, chamadas metaheurísticas, que podem fornecer boas soluções para esses problemas. Para lidar com múltiplas funções objetivo (item 3) é necessário utilizar algoritmos de otimização multiobjetivos, capazes de fornecer conjuntos de soluções que satisfaçam todos os objetivos simultaneamente. Ainda assim, um número muito grande de objetivos pode tornar o problema de otimização inviável. Nesse caso deve-se buscar uma forma de reduzir o número de objetivos do problema, como por exemplo, dividir o problema em vários subproblemas com um número menor de objetivos, ou agrupar vários objetivos em uma única função a ser minimizada.

Diversas metodologias já foram propostas para se calibrar os LSMs de forma mais precisa e realista. Entretanto, nenhum método de calibração atualmente utilizado permite uma calibração satisfatória de todas as saídas produzidas pelos modelos mais complexos. A maioria das calibrações desenvolvidas otimizam no máximo três saídas do modelo simultaneamente (*Gupta e Sorooshian, 1998; Yapo et al., 1998; Madsen, 2000; Xia et al., 2002; Madsen, 2003*).

Algoritmos evolutivos são frequentemente aplicados em problemas de otimização de modelos complexos (*Yapo et al., 1998; Xia et al., 2002; Vrugt et al., 2003; Tang et al., 2006; Bekele e Nicklow, 2007*). Esse tipo de algoritmo permite uma busca global no espaço dos parâmetros, evitando o problema de convergência para ótimos locais. Além disso, esses algoritmos permitem uma otimização multiobjetivo baseada no conceito de fronteira de Pareto, que gera um conjunto de soluções ótimas possíveis que representam trocas entre os objetivos conflitantes. Isso permite que o analista escolha a melhor solução dentre o conjunto retornado pelo algoritmo, baseado em critérios previamente definidos.

O principal objetivo deste trabalho é desenvolver uma metodologia de calibração que permita fazer a calibração de todos os processos simulados por um modelo de micrometeorologia e de dinâmica de ecossistemas terrestres – o Integrated Biosphere Simulator (IBIS) (*Foley et al., 1996*).

A metodologia proposta se baseia na hierarquia dos sistemas ambientais para calibrar sucessivamente todos as variáveis simuladas pelo IBIS. A calibração ocorre em ordem crescente de hierarquia temporal, dos processos mais rápidos (fluxos radiativos) até os processos mais lentos (alocação de carbono). A calibração hierárquica do modelo é comparada com calibrações mono-objetivo independentes de cada variável. Essas calibrações mono-objetivo são consideradas as melhores calibrações possíveis, e a comparação entre as duas calibrações fornecem um indicador de desempenho relativo do método de calibração proposto. Através da calibração hierárquica de todas as variáveis simuladas, espera-se obter um modelo que simule de forma consistente todos os processos nele representados.

Em primeiro lugar, foi desenvolvida uma análise de sensibilidade para selecionar, de forma eficiente, os parâmetros mais importantes para cada saída do modelo. A AS é baseada no método de Morris (1991), com modificações no cálculo do índice de sensibilidade para adequar o método ao uso com LSMs. Em seguida, foi desenvolvido um programa para calibração automática multiobjetivo do modelo IBIS, chamado Optis, baseado no algoritmo genético multiobjetivo NSGA-II (Nondominated Sorted Genetic Algorithm II) (*Deb et al.*, 2002). Esse algoritmo permite ajustar múltiplas estatísticas e várias saídas do modelo simultaneamente. Finalmente, foi utilizada a metodologia de calibração hierárquica para LSMs, que executa a calibração de todos os processos simulados pelo modelo.

O desenvolvimento de um sistema de calibração automático, que minimize múltiplas medidas de erro simultaneamente e calibre todas as variáveis simuladas pelo modelo, permitirá uma simulação mais precisa dos processos e um estudo mais confiável dos fenômenos naturais relacionados com tais processos.

## 2 REVISÃO DE LITERATURA

### 2.1 Análise de sensibilidade

Análise de sensibilidade (AS) é o estudo da relação entre as informações de entrada e de saída de um modelo. Através da AS podemos entender como a variação na saída de um modelo pode ser dividida e/ou atribuída a diferentes fontes de variação na entrada, e como um dado modelo depende dos dados fornecidos a ele. Especificamente, no contexto de calibração de modelos, a AS tem o objetivo de investigar como um certo modelo computacional responde a variações nos seus parâmetros, determinando:

1. Os parâmetros do modelo (ou partes do próprio modelo) que são insignificantes para determinado tipo de saída e podem ser desconsiderados no processo de calibração.
2. A região ótima no espaço de parâmetros para ser usada no estudo subsequente de calibração.
3. Se, e quais, parâmetros interagem uns com os outros.

A eliminação de parâmetros sem influência em determinadas saídas permite uma redução dimensional no espaço dos parâmetros, facilitando o processo de busca. O ítem (2) salienta a necessidade de uma AS global, sendo necessário investigar todo o espaço dos parâmetros, e não apenas em torno de um ponto específico. O ítem (3) aponta para uma questão técnica importante: frequentemente parâmetros possuem efeitos combinados que não podem ser reduzidos à soma dos efeitos individuais. Isso é relevante, já que a presença de interação tem implicações para todos os demais ítems.

### 2.1.1 Tipos de análises de sensibilidade

*Saltelli et al.* (2000) propõem uma divisão dos métodos de AS em três classes: métodos de triagem, métodos de AS local e métodos de AS global. A primeira classe é caracterizada em relação ao seu uso (triagem), enquanto as outras duas se caracterizam pelo modo como tratam os parâmetros.

**Triagem** - Os experimentos de triagem são usados para identificar o conjunto de parâmetros que controlam a maior parte da variabilidade da saída do modelo. Por serem métodos econômicos computacionalmente, são geralmente apropriados quando se trabalha com modelos com alto custo computacional e com grande número de parâmetros. Eles se baseiam na idéia de que a influência dos parâmetros nos modelos seguem a *Lei de Pareto*<sup>1</sup>, com poucos parâmetros muito influentes e a maioria dos parâmetros não influentes. Em compensação, esses métodos tendem a fornecer medidas qualitativas de sensibilidade, isto é, eles ordenam os fatores de entrada do modelo em ordem de importância, mas não quantificam o quanto um parâmetro é mais importante que outro. Em contrapartida, um método de AS quantitativo pode informar, por exemplo, a porcentagem exata da variância total da saída por que cada parâmetro (ou grupo de parâmetros) é responsável. Nos métodos de triagem há claramente uma troca

---

<sup>1</sup>A Lei de Pareto (também conhecido como princípio 80-20), afirma que para muitos fenômenos, 80% das consequências advém de 20% das causas.

entre custo computacional e informação.

**AS Local** - Análises de sensibilidade locais se concentram no impacto local dos parâmetros no modelo. Essas análises são normalmente feitas através do cálculo de derivadas parciais das funções de saída do modelo em relação aos parâmetros de entrada. Para calcular essas derivadas numericamente, os parâmetros são variados em um pequeno intervalo, em torno de um valor nominal. A AS local pode ser vista como um caso particular da abordagem um-fator-por-vez (OAT, do inglês one-factor-at-a-time), já que quando um fator é variado, todos os outros são mantidos constantes. Esses métodos de AS local são efetivos em selecionar os parâmetros mais importantes apenas quando o modelo é linear em todos os seus parâmetros, ou a menos que seja calculado algum tipo de média das derivadas por todo o espaço de entrada dos parâmetros. Esses métodos também têm seu uso limitado ou inadequado quando lidamos com modelos com características não-monotônicas e não-aditivas (*Saltelli et al., 2005*).

**AS Global** - Técnicas de AS globais atribuem, proporcionalmente, as incertezas na saída do modelo às incertezas nos parâmetros de entrada. As estimativas de sensibilidade de cada parâmetro consideram todo o seu intervalo de variação, e são desenvolvidas variando-se também todos os outros parâmetros. Dessa forma, é possível calcular, além dos efeitos individuais de cada parâmetro, os efeitos de interações entre os parâmetros ou grupos de parâmetros.

Uma característica importante, que deve ser considerada ao escolher a técnica de AS mais adequada para determinado problema, é a *independência de modelo*. Essa característica garante que a acurácia do método de AS não seja influenciada pelo nível de aditividade ou linearidade do modelo.

### 2.1.2 Método de Morris

O método de análise de sensibilidade proposto por *Morris* (1991) é um método de triagem, desenvolvido para trabalhar com modelos que possuem um grande número de parâmetros ou com um alto custo computacional de execução. O objetivo deste tipo de método é descobrir quais parâmetros, dentre os vários potencialmente importantes, são realmente relevantes. O método de Morris se propõe a identificar quais parâmetros possuem efeitos que são (a) desprezíveis, (b) lineares e aditivos, ou (c) não-lineares ou envolvidos em interações com outros parâmetros.

Nesse método os parâmetros são variados um de cada vez. Cada parâmetro pode assumir um número discreto de valores, chamados níveis, que são escolhidos dentro do intervalo de variação do parâmetro. Duas medidas de sensibilidade são propostas por Morris: a medida  $\mu$ , que estima o efeito médio do parâmetro na saída do modelo, e a medida  $\sigma$ , que estima, através do desvio padrão, os efeitos de segunda ordem e ordens maiores nos quais o parâmetro está envolvido (incluindo curvaturas e interações). A medida de Morris,  $\mu$ , é obtida calculando-se um número  $r$  de variações incrementais do parâmetro em diferentes pontos do seu espaço e, então, calculando a média das variações na saída do modelo devido a essas variações na entrada. Como a medida de sensibilidade proposta por Morris, chamada de *efeito elementar*, usa variações incrementais nos parâmetros, ela aparenta ser uma AS local. Entretanto, a medida final  $\mu$  é obtida através da média de vários efeitos elementares calculados em diferentes pontos do espaço dos parâmetros. Assim, ela perde a dependência dos pontos específicos nos quais os efeitos elementares são calculados. Portanto, o método tende a explorar diversas regiões do espaço dos parâmetros, e pode ser considerado global.

O uso de uma terceira medida,  $\mu^*$ , é sugerida por *Saltelli et al.* (2005). Ela é a versão em módulo da medida proposta por Morris, ou seja, é a média do módulo dos efeitos elementares. Essa medida evita uma falha do método de Morris quando utilizado em modelos não-monotônicos. Nesse caso, podem ocorrer

efeitos elementares de sinais opostos que, quando calculada a média, poderiam se anular, produzindo um baixo  $\mu$ . Esse problema é eliminado com a medida  $\mu^*$ .

Detalhes sobre a implementação do método são apresentados na seção 3.4.

## 2.2 Avaliação de modelos

Com o crescente aumento do número e tipos de modelos meteorológicos e ambientais, aumenta também a necessidade de técnicas eficazes para avaliar e validar esses modelos, já que a validação é uma etapa fundamental no processo de modelagem. Muitos trabalhos discutem sobre quais formulações e medidas de erros são mais adequadas para verificar a acurácia e eficácia dos modelos (*Fox*, 1980; *Willmott*, 1981, 1982; *Willmott et al.*, 1985; *Legates e McCabe Jr.*, 1999; *Willmott e Matsuura*, 2005).

A avaliação de um modelo é feita comparando-se os resultados preditos com os valores observados da variável em questão. Essa comparação geralmente é feita através de medidas que retornam valores numéricos representando o quanto os dados simulados se ajustam aos dados observados. Fica evidente, portanto, a importância da escolha de uma estatística adequada para se avaliar corretamente um modelo.

Assim como para a avaliação de modelos, a escolha das medidas de ajuste é também um fator crítico para o problema de calibração de modelos, talvez até mais importante do que a escolha do método de otimização (*Trudinger et al.*, 2007). De certa forma, a avaliação do modelo pode ser vista como uma das etapas do procedimento de calibração, já que durante a busca pelos valores ótimos dos parâmetros são realizadas diversas avaliações dos resultados do modelo.

### 2.2.1 Estatísticas e medidas de ajuste de modelos

Diversas estatísticas e técnicas são citadas na literatura e utilizadas para determinar o erro entre os dados simulados por um modelo (P) e os dados observados (O). Dentre elas destacam-se o coeficiente de correlação de Pearson ( $r$ ), o



coeficiente de determinação ( $r^2$ ), o erro absoluto médio (MAE), a raiz do erro quadrático médio (RMSE), o coeficiente de eficiência de Nash-Sutcliffe (E), o índice de concordância de Willmott (d) e os coeficientes m e b da reta de regressão (declividade e intercepto) dos dados simulados versus os observados (Fox, 1980; Willmott et al., 1985; Yapo et al., 1998; Legates e McCabe Jr., 1999; Tang et al., 2006; Duan et al., 2006).

**Medidas de erro médio** - As medidas de erro médio (e.g. MSE, RMSE, MBE e MAE) contém informações apropriadas e compreensíveis sobre o ajuste do modelo aos dados observados. Dentre essas medidas, o RMSE e o MAE são particularmente interessantes por serem dimensionais, ou seja, expressam o erro do modelo nas unidades da variável em questão. É recomendado que um procedimento de avaliação de modelo reporte pelo menos uma dessas medidas (Willmott, 1982).

Apesar de expressarem um erro nas mesmas unidades dos dados analisados, estas duas estatísticas possuem algumas características que as diferenciam. Como mostrado por Willmott e Matsuura (2005), o RMSE depende da variabilidade das magnitudes dos erros (ou erros-quadrados), bem como com a magnitude do erro médio e de  $n^{\frac{1}{2}}$ . Portanto, sem o auxílio de outra informação (e.g. MAE), é impossível distinguir o quanto do RMSE representa a tendência central (erro médio) e o quanto dele representa a variabilidade na distribuição dos erros quadrados ou  $n^{\frac{1}{2}}$  (o limite do RMSE é dado por  $MAE \leq RMSE \leq MAE * n^{\frac{1}{2}}$ ) (Willmott e Matsuura, 2005).

O MAE é uma medida mais natural de erro médio, pois não possui potências quadradas ou raízes na sua fórmula. Diferentemente do RMSE, o MAE não é ambíguo - ele expressa somente o erro médio em valores absolutos. Por ser de interpretação mais clara, é mais apropriado utilizar o MAE para avaliação de modelos, ao invés do RMSE ou medidas relacionadas ao RMSE.

As fórmulas do RMSE e do MAE são dadas por:

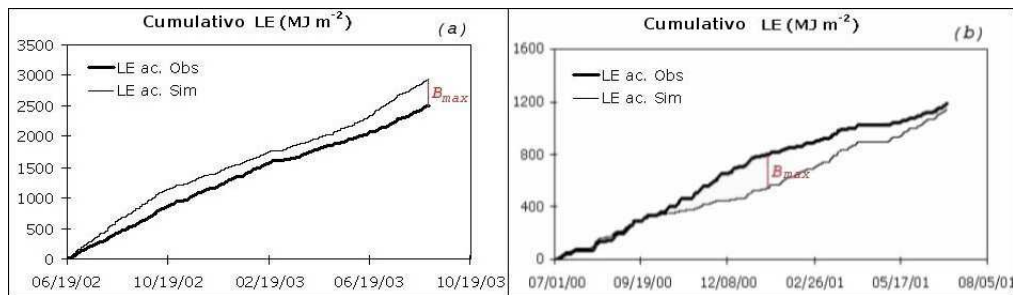
$$MAE = n^{-1} \sum_{i=1}^n |P_i - O_i| \quad (2)$$

$$RMSE = \left[ n^{-1} \sum_{i=1}^n (P_i - O_i)^2 \right]^{\frac{1}{2}} \quad (3)$$

**Gráfico de soma cumulativa -** A soma cumulativa de uma série de dados é calculada tomando-se o valor anterior acumulado e somando-se com o valor atual da série. Dessa forma, é gerada uma nova série de dados acumulados.

O gráfico cumulativo dos dados simulados e observados fornecem informações interessantes sobre o comportamento dos dados ao longo do período analisado. Mesmo que o viés final entre os dados seja pequeno, ou nulo, o gráfico cumulativo consegue detectar o viés entre os dados em períodos intermediários das séries (Figura 1-b).

Na Figura 1-a, a maior distância entre as duas curvas está no final da série, enquanto na Figura 1-b a máxima distância se localiza no meio da série, com uma distância muito pequena no final. Nesse segundo caso, o cálculo simples do viés não detectaria a diferença que ocorre entre as médias dos dados no período intermediário da série.



**Figura 1:** Gráficos cumulativos de LE de duas simulações distintas (a) e (b) do modelo IBIS (Imbuzeiro, 2005)

Uma forma de utilizar a informação da soma cumulativa para calibração de modelos é tentar minimizar a máxima distância entre as curvas cumulativo dos dados observados e simulados. Esse tipo de calibração tenderia a produzir uma

simulação não viesada do modelo durante todo o período de dados.

### 2.3 Otimização multiobjetivo

O termo otimização se refere ao estudo de problemas nos quais se busca minimizar ou maximizar uma certa função objetivo, escolhendo sistematicamente os valores das variáveis dessa função, e geralmente satisfazendo determinadas restrições.

A função objetivo de um problema de otimização é a função matemática cujo máximo ou mínimo deseja-se determinar. Por exemplo, no ajuste de um modelo computacional, poderíamos ter o objetivo de minimizar o erro médio entre os dados observados e as predições do modelo.

As variáveis são um conjunto de incógnitas que afetam o valor da função objetivo. Nos problemas de modelagem, as variáveis normalmente são os parâmetros do modelo.

As restrições são os limites impostos ao sistema, os quais definem os valores que as variáveis podem assumir. No problema de modelagem, as restrições poderiam ser, por exemplo, conjuntos de valores inválidos para determinados parâmetros.

Os problemas em que existem mais de uma função objetivo concorrentemente são chamados de problemas de otimização multiobjetivo. Enquanto os problemas de otimização mono-objetivo devem ter uma solução ótima única, problemas de otimização multiobjetivo geralmente apresentam um conjunto de soluções possíveis, cujos componentes representam trocas no espaço dos objetivos. (*Veldhuizen e Lamont, 2000; Deb et al., 2002; Vrugt et al., 2003*)

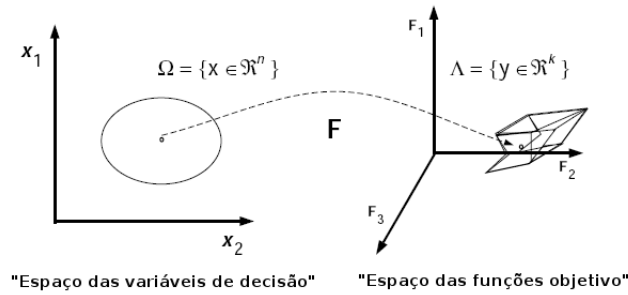
Os problemas de otimização multiobjetivo podem ser definidos da seguinte forma:

$$\min_{\vec{x} \in \Omega} F(\vec{x}) = \{f_1(\vec{x}), \dots, f_k(\vec{x})\}, \text{ sujeito a } g_i(\vec{x}) \leq 0, i = 1, \dots, m \quad (4)$$

onde  $f_1(\vec{x}), \dots, f_k(\vec{x})$  são as  $k$  diferentes funções objetivo a serem simulta-

neamente minimizadas,  $g_i(\vec{x})$  são as  $m$  restrições do problema,  $\vec{x}$  é um vetor de  $n$  dimensões ( $\vec{x} = x_1, \dots, x_n$ ) e  $\Omega$  é o espaço dos parâmetros.

Um problema de otimização multiobjetivo, portanto, consiste de  $n$  variáveis de decisão,  $m$  restrições, e  $k$  objetivos. A função de avaliação do problema,  $F : \Omega \rightarrow \Lambda$ , mapeia variáveis de decisão ( $\vec{x} = x_1, \dots, x_n$ ) em vetores no espaço dos objetivos ( $\vec{y} = y_1, \dots, y_n$ ) (Veldhuizen e Lamont, 2000). Esse mapeamento é representado na Figura 2, para o caso de  $n = 2$ ,  $m = 0$  e  $k = 3$ .

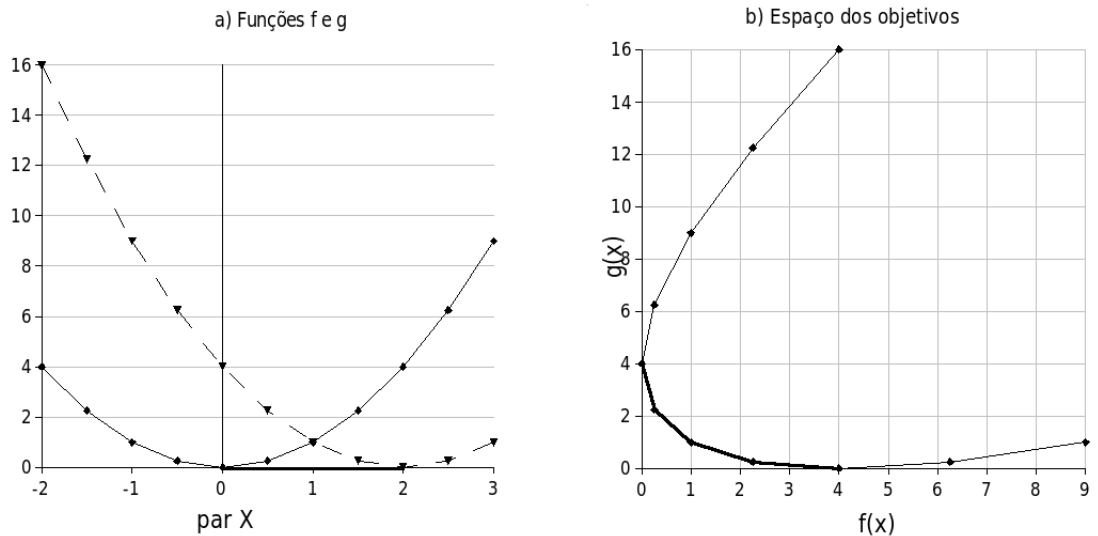


**Figura 2:** Mapeamento dos parâmetros em um problema de otimização multiobjetivo. Adaptado de Veldhuizen e Lamont (2000)

Um exemplo simples de otimização multiobjetivo é o problema de minimizar duas funções,  $f = x^2$  e  $g = (x + 2)^2$  em relação a um parâmetro  $x \in \mathbb{R}$ . Como mostra a Figura 3-a, o ponto  $x = 0$  indica o valor do parâmetro que minimiza a função  $f$ , enquanto o ponto  $x = 2$  corresponde ao valor do parâmetro que minimiza a função  $g$ . A solução do problema consiste de todos os pontos entre  $x = 0$  e  $x = 2$ , já que entre esses dois pontos acontece uma troca entre as funções objetivo, isto é, enquanto uma aumenta a outra diminui o valor. O mapeamento dos parâmetros ótimos no espaço dos objetivos, formando a Fronteira de Pareto, é mostrado na Figura 3-b.

## 2.4 Algoritmo genético multiobjetivo

Segundo Reeves (1995), uma heurística é uma técnica que busca boas soluções para um determinado problema complexo. Pode-se dizer que uma heurística é uma regra ou um algoritmo de aproximação, que reduz ou limita a busca por



**Figura 3:** Problema de otimização multiobjetivo: (a) funções objetivos  $f$  e  $g$  e (b) fronteira de Pareto (linha em negrito) no espaço dos objetivos

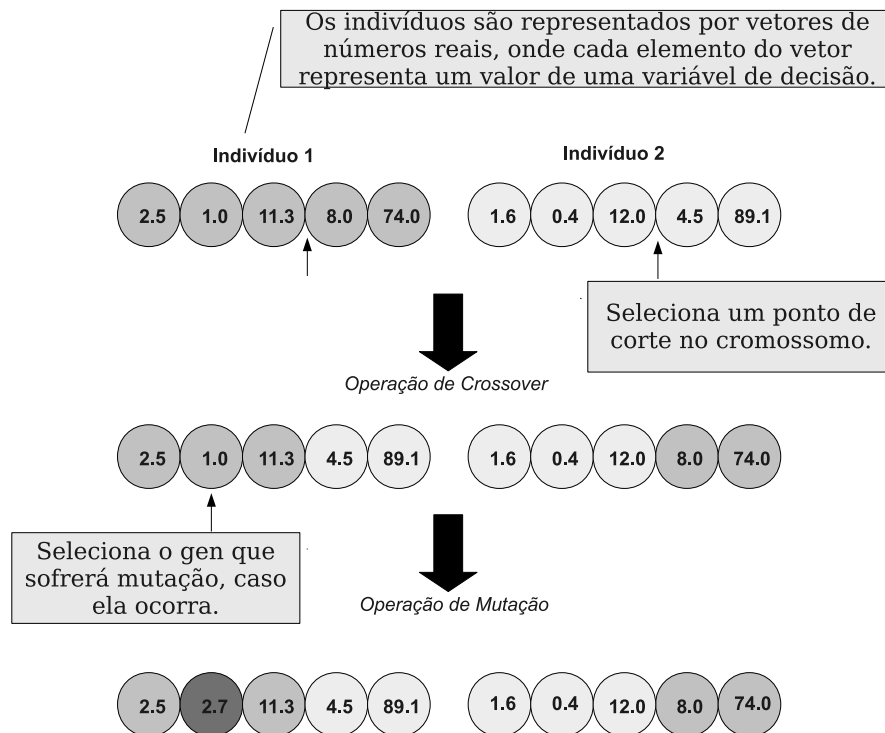
soluções em domínios que são difíceis e pouco compreendidos. Um algoritmo heurístico fornece soluções sem um limite formal de qualidade, e é tipicamente avaliado empiricamente em termos de qualidade das soluções.

As metaheurísticas vão além desse conceito, se baseando em idéias e conceitos de outras disciplinas para ajudar a resolver os problemas (*Jones et al.*, 2002).

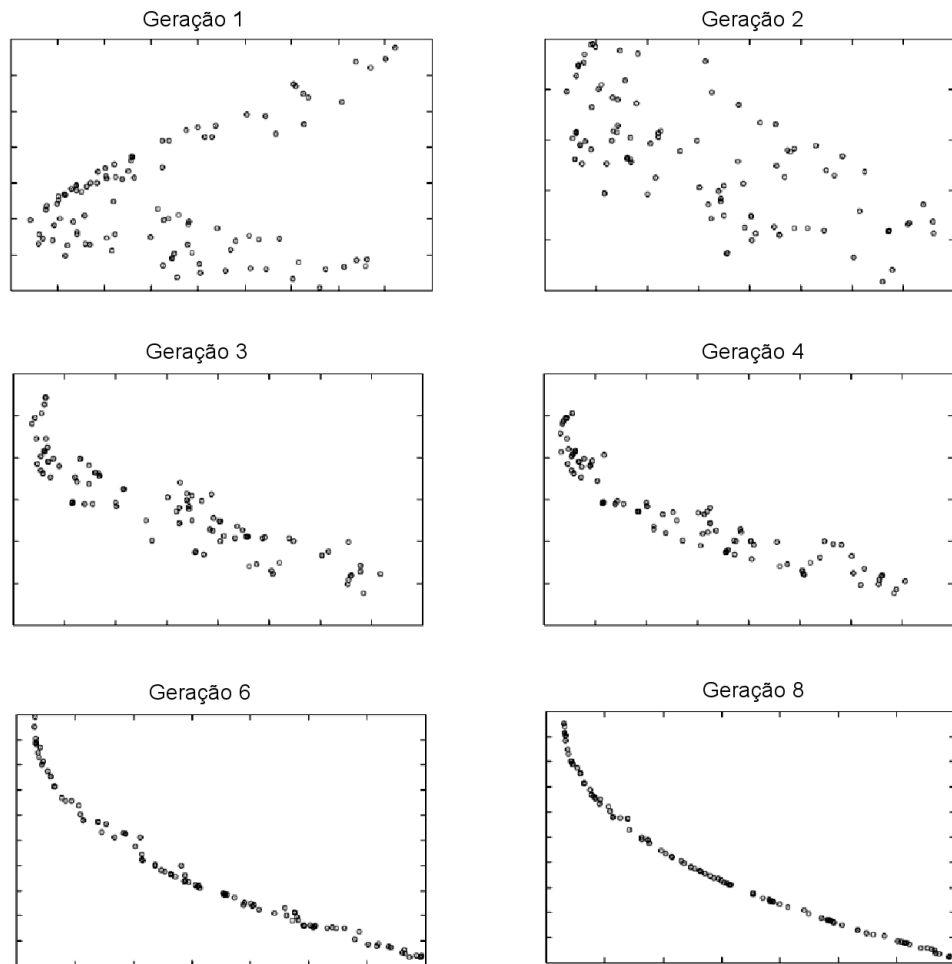
Algoritmos Genéticos (AGs) (*Goldberg*, 1989) são metaheurísticas baseadas na teoria da Evolução de Darwin, e simulam o processo de seleção natural e a sobrevivência dos indivíduos mais adaptados. Os AGs tentam encontrar uma boa solução para algum problema (como encontrar o mínimo de uma função) gerando aleatoriamente uma coleção de soluções potenciais para o problema, e manipulando estas soluções usando o que são chamados de operadores genéticos. Esses operadores genéticos usam as soluções existentes para produzir novas soluções, supostamente melhores que as anteriores. Na terminologia de AG, se diz que uma população de soluções é gerada e refere-se a cada solução como um cromossomo (ou indivíduo da população). Para cada indivíduo é determinado um valor de aptidão escalar, que é uma representação numérica de quão bem ele resolve o problema. A idéia principal é selecionar para reprodução as soluções com maior aptidão e aplicar os operadores genéticos a elas para gerar novos indivíduos.

Através de operações de mutação e recombinação (crossover), espera-se gerar indivíduos melhores fora do atual conjunto de soluções potenciais. Esse processo continua até que uma solução aceitável seja encontrada. O funcionamento de um AG é ilustrado na Figura 4.

Em um AG multiobjetivo, os indivíduos possuem múltiplas funções de aptidão, que devem ser avaliadas simultaneamente durante o processo de seleção dos melhores indivíduos. Os indivíduos são selecionados buscando a formação de uma curva ou uma superfície no espaço dos objetivos, denominada fronteira de Pareto. Nessa fronteira todos os indivíduos são igualmente aptos, e a escolha de um solução ótima única deve ser feita com base em critérios subjetivos ou empíricos, ou de acordo com prioridades previamente definidas (*Veldhuizen e Lamont, 2000*). A evolução de um algoritmo genético multiobjetivo e a formação da fronteira de Pareto é ilustrada na Figura 5.



**Figura 4:** Esquema do funcionamento dos operadores genéticos de um AG.



**Figura 5:** Aproximação da fronteira de Pareto por um AG multiobjetivo com dois objetivos. Cada gráfico representa uma geração, cada eixo do gráfico representa um objetivo e cada ponto nos gráficos representa uma solução potencial para o problema (ou um indivíduo).

## 3 METODOLOGIA

### 3.1 O modelo IBIS

O modelo IBIS (Integrated Biosphere Simulator) foi projetado para conectar explicitamente processos de superfície terrestre e hidrológicos, ciclos biogeoquímicos terrestres e dinâmica da vegetação em uma única estrutura de modelagem (Foley et al., 1996; Kucharik et al., 2000). O modelo inclui representações de processos físicos da superfície (incluindo transferência de radiação solar e terrestre através do dossel, processos turbulentos, interceptação de água pelo dossel, transferência de massa e energia pelo dossel), fisiologia das plantas, fenologia do dossel, diferenças de tipos funcionais de plantas, alocação de carbono e biogeoquímica do solo. Todos os processos são organizados em módulos de forma hierárquica (Figura 6) e operam em diferentes intervalos de tempo, variando de minutos a anos. Tal abordagem permite o acoplamento explícito entre processos ecológicos, biofísicos e fisiológicos ocorrendo em diferentes escalas de tempo.

Por considerar mudanças transientes na composição e estrutura da vegetação em resposta a mudanças no meio ambiente, o IBIS é classificado como um Modelo de Vegetação Dinâmica Global.



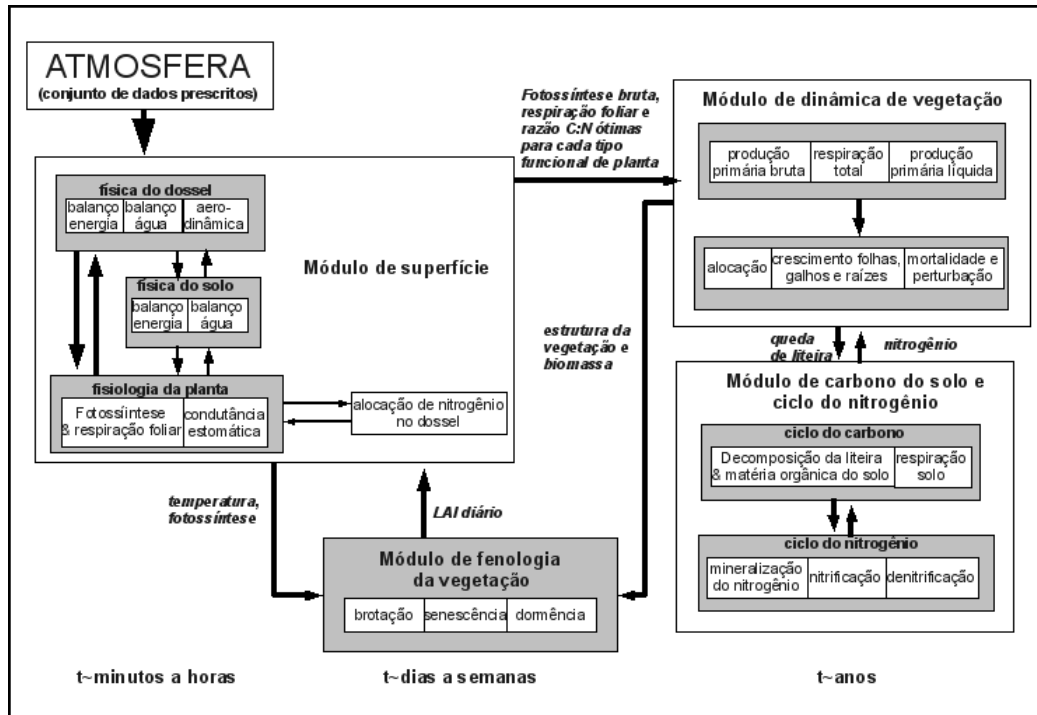


Figura 6: Estrutura em módulos do modelo IBIS.

Dentre os diversos processos e variáveis simulados pelo IBIS, destacam-se nove variáveis do modelo que foram consideradas neste trabalho durante a calibração: radiação fotossinteticamente ativa refletida ( $PAR_o$ ), fração absorvida da radiação fotossinteticamente ativa ( $f_{APAR}$ ), saldo de radiação ( $R_n$ ), velocidade de fricção do vento ( $u^*$ ), fluxos de calor sensível e de calor latente ( $H$  e  $LE$ ), troca líquida do ecossistema ( $NEE$ ), produção primária líquida ( $NPP$ ) e índice de área foliar ( $LAI$ ). Essas variáveis foram escolhidas de acordo com disponibilidade de dados, selecionando para calibração todas as variáveis que possuíam medições de campo para validação dos dados simulados.

A  $PAR_o$  é calculada pelo modelo por meio da seguinte fórmula:

$$PAR_o = (S_{dir,vis} \cdot r_{dir,vis} + S_{dif,vis} \cdot r_{dif,vis}) \cdot 4,59 \quad (5)$$

em que  $S_{dir,vis}$  e  $S_{dif,vis}$  são os fluxos de radiação solar direta e difusa incidentes, na banda do visível.  $r_{dir,vis}$  e  $r_{dif,vis}$  são a refletância da radiação direta e difusa da superfície na banda do visível.

A  $f_{APAR}$  é definida como a fração da radiação fotossinteticamente ativa absorvida pelo dossel. Ela exclui a fração da PAR incidente refletida pelo dossel e a fração absorvida pela superfície do solo. A radiação fotossinteticamente ativa 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:

$$APAR = (PAR_{udir} \cdot a_{udir} + PAR_{udif} \cdot a_{udif}) + (PAR_{ldir} \cdot a_{ldir} + PAR_{ldif} \cdot a_{ldif}) \quad (6)$$

onde  $PAR_{udir}$  é a PAR 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 e  $a_{udif}$  é a fração da PAR difusa absorvida pelo dossel superior.  $PAR_{ldir}$  é a PAR direta incidente no dossel inferior e  $a_{ldir}$  é a fração dessa radiação absorvida pelo dossel inferior.  $PAR_{ldif}$  e  $a_{ldif}$  têm o mesmo significado, porém para a PAR difusa.

A  $f_{APAR}$  é calculada por meio da razão:

$$f_{APAR} = \frac{APAR}{PAR_{in}} \quad (7)$$

O Rn é a contabilização líquida entre toda a energia radiante recebida e perdida pela superfície. Ele é calculado através do saldo dos balanços de ondas curtas e de ondas longas:

$$Rn = S_{in,vis} \cdot (1 - r_{vis}) + S_{in,nir} \cdot (1 - r_{nir}) + L_{in} - L_{out} \quad (8)$$

onde  $S_{in,banda}$  é o fluxo de radiação solar incidente, na banda indicada.  $r_{banda}$  é a refletância 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, calculada pelas fórmulas:

$$L_{out} = (1 - fu).fup + fu.[(1 - emu).(1 - ems).fup + emu.stef.tu^4 + ems.(1 - emu).stef.ts^4] \quad (9)$$

$$emu = 1 - e^{(-LAI/avmuir)} \quad (10)$$

$$ems = 1 - e^{(-SAI/avmuir)} \quad (11)$$

onde  $emu$  e  $ems$  são as emissividades das folhas e galhos superiores da vegetação, respectivamente.  $fu$  é a fração média da área média coberta pela vegetação,  $fup$  é o fluxo de radiação IR ascendente 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 (H) na superfície e no fluxo de calor no solo (G):

$$Rn = H + LE + G \quad (12)$$

A velocidade de fricção do vento é calculada pelas equações:

$$u^* = \sqrt{\left[ \frac{u_a \cdot vonk}{(aloga - alogu)} \right]^2 \cdot stramu} \quad (13)$$

$$alogu = \ln [\max(0,01, alogu\_coef.(z1 - z2))] \quad (14)$$

$$aloga = \ln(z_a - dispu) \quad (15)$$

onde  $u_a$  é a velocidade do vento,  $vonk$  é a constante de Von Karman,  $stramu$  é um fator de correção do momentum sobre o dossel superior,  $z1$  é a altura do topo

do dossel superior,  $z_2$  é a altura da base do dossel superior,  $z_a$  é a altura sobre a superfície da forçante atmosférica,  $dispu$  é o deslocamento do plano zero para o dossel superior.

O fluxo saldo de  $CO_2$ , ou troca líquida do ecossistema, é a diferença entre dois termos – a captura de  $CO_2$  durante a fotossíntese pela folha e a emissão de  $CO_2$  pela respiração do solo. No IBIS, portanto, a NEE é calculada pela seguinte fórmula:

$$NEE = R_H - NPP \quad (16)$$

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

A NPP representa o carbono novo armazenado como biomassa nos galhos, folhas e raízes das plantas. Ela é a diferença entre o carbono assimilado durante a fotossíntese e o consumo de carbono através da respiração das plantas – é uma medida quantitativa de crescimento da planta e de sequestro de carbono. A NPP no IBIS é calculada pela fórmula:

$$NPP = (1 - \eta) \int (A_g - R_{leaf} - R_{stem} - R_{root}) dt, \quad (17)$$

onde  $A_g$  é a fotossíntese bruta do dossel,  $R$  são as respirações de manutenção da folha, troncos e raízes, e  $\eta$  é a fração de carbono perdido devido a respiração de crescimento (*Kucharik et al.*, 2000).

A taxa de fotossíntese de plantas  $C_3$ , que inclui todas as árvores e muitas herbáceas, é representada seguindo as equações de Farquhar (*Farquhar et al.*, 1980; *Collatz et al.*, 1991).  $A_g$  é definida como o mínimo de duas taxas potenciais,

$$A_g \approx \min(J_E, J_C) \quad (18)$$

$J_E$  é a taxa de fotossíntese limitada pela luz e  $J_C$  é a taxa de fotossíntese limitada

pela enzima Rubisco, definidas como

$$J_E = \alpha_3 Q_p \left( \frac{C_i - \Gamma_*}{C_i + 2\Gamma_*} \right) \quad (19)$$

onde  $Q_p$  é a densidade do fluxo de radiação fotossinteticamente ativa (PAR) absorvida pela folha,  $\alpha_3$  é a eficiência quântica para absorção de  $CO_2$  em plantas  $C_3$ ,  $C_i$  é a concentração de  $CO_2$  nos espaços de ar intracelulares da folha, e  $\Gamma_*$  é o ponto de compensação para fotossíntese bruta.

$$J_C = \frac{V_m (C_i - \Gamma_*)}{C_i + K_C \left( 1 + \frac{[O_2]}{K_O} \right)} \quad (20)$$

onde  $V_m$  é a capacidade de carboxylase máxima da Rubisco, e  $K_C$  e  $K_O$  são os coeficientes de Michaelis-Menten para  $CO_2$  e  $O_2$ , respectivamente.

$V_m$  é calculada de acordo com as seguintes fórmulas:

$$V_m = V_{max} \cdot tempvm \cdot stresstu \quad (21)$$

$$tempvm = \frac{e^{[tempvmcoef \cdot (3,47e-03 - \frac{1}{tu})]}}{(\{1, 0 + e^{[0,40(5,0 - tleaf)]}\} \cdot \{1, 0 + e^{[0,40(tleaf - 50,0)]}\})} \quad (22)$$

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

$$stresstu = \sum_{j=1}^L stressu_j \quad (23)$$

$$stressu = froot \cdot \max(0, \min(1, 0, zwilt)) \quad (24)$$

$$zwilt = \frac{[1 - e^{(stressfac.awc)}]}{[1 - e^{stressfac}]} \quad (25)$$

onde  $stressu$  é a função de estresse devido a umidade no solo em cada camada do solo,  $L$  é o número de camadas de solo,  $froot$  é a fração de raiz na camada do solo,  $zwilt$  é uma função de estresse hídrico,  $awc$  é o conteúdo de água disponível e  $stressfac$  é o parâmetro para cálculo do estresse hídrico.

A taxa de fotossíntese de plantas  $C_4$  é definida como

$$A_g \approx \min(J_I, J_E, J_C) \quad (26)$$

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  $J_C = kC_i$  é a taxa de fotossíntese limitada pelo  $CO_2$  a baixas concentrações.

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

$$R_{leaf} = \gamma V_m, \quad (27)$$

onde  $\gamma$  é o custo de respiração da folha (*Kucharik et al.*, 2000). As taxas de respiração de manutenção da biomassa do tronco e raízes finas são dadas por

$$R_{stem} = \beta_{stem} \lambda_{sapwood} C_{stem,i} f(T_{stem}) \quad (28)$$

e

$$R_{root} = \beta_{root} C_{root,i} f(T_{soil}), \quad (29)$$

onde  $C_{stem}$  e  $C_{root}$  são o carbono contido na biomassa do tronco e das raízes finas, respectivamente,  $\beta$  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:

$$f(T) = e^{E_0 \left( \frac{1}{15 - T_0} - \frac{1}{T - T_0} \right)} \quad (30)$$

$T$  é a temperatura (do tronco ou raízes),  $E_0$  é um fator de sensibilidade de temperatura, e  $T_0$  é uma temperatura de referência.

A condutância estomática é calculada pela fórmula

$$g_{s,h_2o} = \frac{mA_n}{C_S} h_S + b, \quad (31)$$

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, e  $m$  e  $b$  são o grau de inclinação e o intercepto da relação condutância-fotossíntese.

O LAI é a razão entre a área da superfície de folha superior total da vegetação e a área da superfície do solo na qual a vegetação cresce. No IBIS, o LAI é calculado pela soma dos LAIs de cada tipo funcional de planta representado pelo modelo:

$$LAI = \sum_{j=1}^N cbiol_j \cdot specla_j \quad (32)$$

onde  $N$  é o número de tipos funcionais de plantas representados no modelo,  $specla$  é área de folha específica e  $cbiol$  é o carbono no reservatório de biomassa de folha, calculado pela fórmula:

$$cbiol = cbiol \cdot e^{(-1/tauleaf)} + aleaf \cdot tauleaf \cdot \max(0, aynpp) \cdot [1 - e^{(-1/tauleaf)}] \quad (33)$$

onde  $tauleaf$  é o tempo de retorno do carbono nas folhas,  $aleaf$  é a fração de alocação de carbono para as folhas e  $aynpp$  é a NPP total anual.

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

$$outcrs = \min(decomps.krs.clitrs, clitrs) \quad (34)$$

$$outcws = \min(decompl.kws.clitws, clitws) \quad (35)$$

$$outcls = \min(decompl.kls.clitls, clitls) \quad (36)$$

$$outcnb = \min(decomps.knb.csoislon, csoislon) \quad (37)$$

$$outcpb = \min(decomps.kpb.csoislop, csoislop) \quad (38)$$

onde *outcrs*, *outcws* e *outcls* são as quantidades de C saindo dos reservatórios de carbono estrutural de raízes, madeira e folhas, respectivamente. *outcnb* é o fluxo de matéria orgânica não protegida para a biomassa e *outcpb* é o fluxo de matéria orgânica protegida para a biomassa. *krs*, *kws* e *kls* são as constantes de decomposição para os reservatórios de liteira estrutural de raízes, madeira e folhas, respectivamente. *kpb* e *knb* são as mesmas constantes, porém para a 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):

$$Y(d) = \frac{1 - \beta_2^d}{1 - \beta_2^{d_{max}}} \quad (39)$$

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  $\beta_2$  é um parâmetro de distribuição de raízes finas.

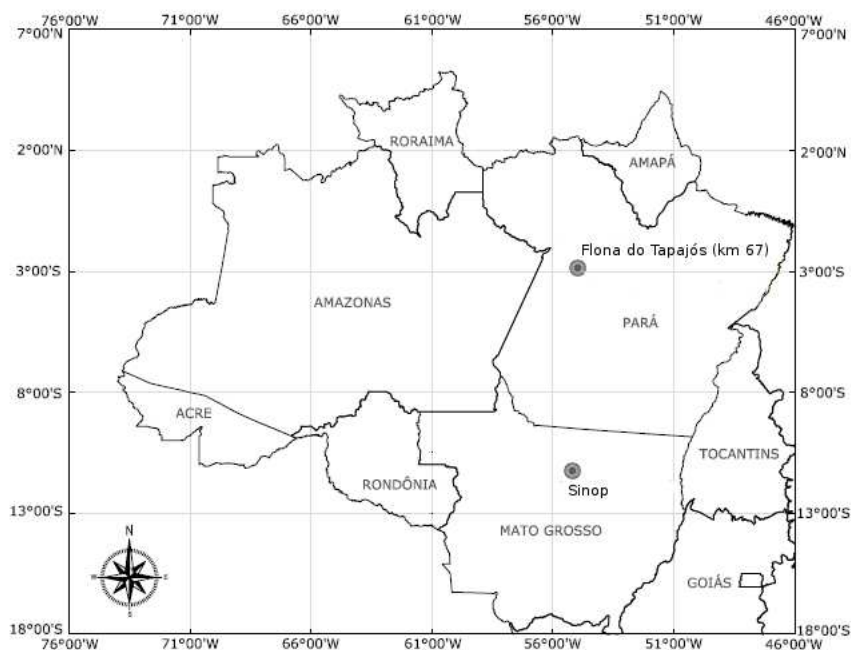
O código do IBIS é disponibilizado gratuitamente e está escrito em Fortran 77.

### 3.2 Descrição dos sítios

Para calibrar o modelo IBIS foram utilizados dados coletados em dois sítios micrometeorológicos da região Amazônica: Floresta Nacional do Tapajós km 67



e Fazenda Maracaí, Sinop, MT. A Figura 7 mostra a localização dos sítios.



**Figura 7:** Localização dos sítios micrometeorológicos da Flona do Tapajós e Sinop. Adaptado de *Costa et al.* (2007a).

Esses sítios representam diferentes zonas climáticas da Amazônia, com diferentes graus de sazonalidade e estruturas de floresta (*Costa et al.*, 2007a). Como a floresta se comporta diferentemente em cada sítio, foram necessárias calibrações distintas do modelo, uma para cada sítio.

O sítio da Floresta Nacional do Tapajós está localizado próximo ao km 67 ( $02^{\circ}51'S$ ,  $54^{\circ}58'O$ ) da rodovia Santarém-Cuiabá, ao sul de Santarém, Pará, Brasil. A vegetação é tipicamente de uma floresta tropical úmida fechada, em um terreno plano, com a maioria das espécies perenes e poucas espécies decíduas. A floresta se estende 5 km ao leste, 8 km ao sul e 40 km ao norte, antes de encontrar pastagem (*Saleska et al.*, 2003). A altura média do dossel é 40 m, com espécies emergentes atingindo 55 m (*Costa et al.*, 2007a).

O sítio da Fazenda Maracaí está situado a aproximadamente 50 km ( $11^{\circ}25'S$ ,  $55^{\circ}19'O$ ) da cidade de Sinop, norte do Mato Grosso, Brasil. Essa área pertence à bacia do Rio Teles Pires, numa região de transição entre floresta e cerrado. As florestas de transição ocupam uma faixa climática larga, que separa a floresta de

terra firme e úmida, da bacia central da Amazônia, da savana tropical mais seca (cerrado) (*Spolador et al.*, 2006). A altura média do dossel é 28 m, mas as árvores emergentes mais altas alcançam 42 m (*Priante-Filho et al.*, 2004).

### 3.3 Dados observados

Os dados observados nos sítios experimentais foram usados como dados de entrada do modelo IBIS e para validar os seus resultados. Como dados de entrada do modelo foram utilizadas as medições de temperatura e umidade específica do ar, precipitação, velocidade do vento, fluxos de radiação de onda curta e fluxos de radiação de onda longa.

Os dados da Flona do Tapajós foram coletados nos seguintes períodos: O dado de NPP é uma média anual de dados coletados no ano de 2004. Os dados de  $f_{APAR}$  e LAI são médias mensais, coletados nos anos 2002 e 2004, respectivamente (obtidos pelo projeto BIGFOOT). Os dados de  $PAR_o$ ,  $R_n$ ,  $u^*$ ,  $H$ ,  $LE$ ,  $NEE$  e os dados que alimentam o modelo são médias de uma hora, do período de janeiro de 2002 a dezembro de 2004 (fornecidos por Scott Saleska, Harvard University). O modelo foi calibrado em relação às variáveis  $PAR_o$ ,  $f_{APAR}$ ,  $R_n$ ,  $u^*$ ,  $H$ ,  $LE$ ,  $NEE$ ,  $NPP$  e  $LAI$ .

Os dados de Sinop foram coletados no período de janeiro a dezembro de 2002. Para calibração do modelo neste sítio foram utilizados, além dos dados para alimentação do modelo, dados de  $PAR_o$ ,  $f_{APAR}$ ,  $R_n$ ,  $u^*$ ,  $H$ ,  $LE$ ,  $NEE$  e  $LAI$  (fornecidos por Luciana Sanches, UFMT).

Em ambos os sítios, os fluxos de calor latente e fluxo de calor sensível são medidos por sistemas de covariância de vórtices turbulentos. O balanço de radiação na superfície, a temperatura e umidade do ar, velocidade do vento e outras variáveis foram medidos por estação meteorológica automática (*Costa et al.*, 2007a).

Não foram coletados os dados de radiação de onda longa incidente ( $L_{in}$ ) em

nenhum dos sítios experimentais. Neste caso,  $L_{in}$  foi estimada pela fórmula

$$L_{in} = Rn - S_{in} + \lambda\sigma T_a^4, \quad (40)$$

onde  $\lambda$  é a emissividade,  $\sigma$  é a Constante de Stefan-Boltzmann e  $T_a$  a temperatura absoluta do ar. Nesse caso, supõe-se que a temperatura efetiva da emissão do dossel é igual à temperatura absoluta do ar.

Devido a problemas técnicos no pluviômetro da Fazenda Maracaí, a precipitação nesse sítio foi estimada baseada em dados de uma estação meteorológica localizada a aproximadamente 17 km.

### 3.3.1 Preenchimento de falhas dos dados

As falhas dos dados observados utilizados como entrada do IBIS foram preenchidas por interpolação, de acordo com as seguintes fórmulas propostas por *Senna* (2004):

- se o número de pontos de falhas ( $f$ ) for menor ou igual a 3 horas:

$$X_i = X_p + \frac{(X_{p+f+1} - X_p)}{f + 1} \cdot (i - p) \quad (41)$$

onde  $X_i$  é o valor a ser preenchido,  $X_p$  é o antecessor da falha e  $X_{p+f+1}$  é o sucessor da falha.  $i$  é o índice da célula que será interpolada e  $p$  é o índice da célula que precede a falha.

- se a duração da falha for maior que 3 horas e menor que 24 horas:

$$X_i = X_{i-24} + \frac{[(p + f + 1) - i]}{f + 1} \cdot (X_p - X_{p-24}) + \frac{(i - p)}{f + 1} \cdot (X_{p+f+1} - X_{p+f+1-24}) \quad (42)$$

- se a duração da falha for maior ou igual a 24 horas:

$$X_i = X_{i-24m} + \frac{[(p+f+1)-i]}{f+1} \cdot (X_p - X_{p-24m}) + \frac{(i-p)}{f+1} \cdot (X_{p+f+1} - X_{p+f+1-24m}) \quad (43)$$

onde  $m$  é a parte inteira de  $(\frac{f}{24} + 1)$ .

Essas equações de interpolação fazem com que o período sem dados siga o mesmo comportamento de um período similar com dados. Dessa forma, as falhas são preenchidas de forma mais realística do que com uma interpolação linear.

### 3.3.2 Filtragem dos dados de avaliação do modelo

Neste trabalho foram utilizados três tipos de filtragem dos dados utilizados na calibração do modelo, como proposto por *Imbuzeiro* (2005): a filtragem de acordo com as falhas dos dados de entrada, a filtragem do NEE pelo limiar da velocidade de fricção, e a filtragem baseada no fechamento do balanço de energia.

A filtragem dos dados de entrada elimina do processo de avaliação do modelo os períodos onde haviam falhas nos dados de entrada. Essa filtragem visa minimizar os efeitos de possíveis inconsistências dos dados preenchidos na avaliação do modelo.

A filtragem do NEE de acordo com a velocidade de atrito elimina da avaliação do modelo os períodos de dados em que havia baixa turbulência durante a medição. Em condições de baixa turbulência o fluxo de  $CO_2$  pode não ser detectado pelo sistema de covariância de vórtices turbulentos (*Curtis et al.*, 2002; *Saleska et al.*, 2003; *Kruijt et al.*, 2004). Considerou-se condições de baixa turbulência aquelas onde o  $u^*$  é menor que  $0.1 \text{ m.s}^{-1}$ .

A filtragem de dados baseada no fechamento do balanço de energia elimina da calibração os períodos de dados em que o erro no fechamento do balanço de energia é maior que um determinado valor  $\epsilon$ . Um problema fundamental com os dados de fluxos de energia medidos com sistema de vórtices turbulentos é que os dados

violam o princípio da conservação de energia. O balanço de energia determinado usando esse sistema geralmente não "fecha" e os fluxos de calor sensível e latente, combinados, subestimam a energia disponível em 20% ou mais (*Maayar et al.*, 2008).

Como os modelos de superfície são desenvolvidos baseados no princípio da conservação de energia, a exclusão dos períodos de dados em que essa diferença do fechamento é muito grande permite uma avaliação do modelo mais consistente. Neste trabalho foi utilizado  $\epsilon=0.4$ , ou seja, apenas os períodos de dados cujo balanço de energia diário medido pelo sistema de vórtices turbulentos estava dentro de 40% do balanço de energia medido pelos instrumentos radiativos e de fluxo de calor no solo (Equação 44). Como não havia dados disponíveis de fluxo de calor no solo ( $G$ ), esses dados foram considerados desprezíveis, o que é razoável para médias diárias (*Imbuzeiro*, 2005).

$$(1 - \epsilon) \leq \left| \frac{H + LE}{Rn - G} \right| \leq (1 + \epsilon) \quad (44)$$

### 3.4 Análise de sensibilidade do IBIS

Para simplificar o procedimento de calibração, foi realizada uma análise de sensibilidade preliminar dos parâmetros do IBIS. Dessa forma, foi possível isolar parâmetros com influência insignificante nas saídas do modelo, diminuindo o espaço de busca do procedimento de otimização.

A AS foi realizada com dados coletados no sítio experimental do LBA da Flona do Tapajós (km 67), no período de janeiro de 2002 a dezembro de 2004. Foi utilizado o método de análise de sensibilidade de Morris (1991) para fazer a triagem e a classificação dos parâmetros do modelo. O método estima os efeitos principais dos parâmetros do modelo computando, para cada parâmetro, um número de medidas locais chamadas efeitos elementares. O efeito elementar é calculado variando-se um parâmetro de cada vez, através de um número discreto,  $p$ , de níveis selecionados no espaço dos parâmetros de entrada. Foi analisada a

sensibilidade do erro absoluto médio ( $MAE$ ) de nove saídas do modelo IBIS a 43 parâmetros, e os intervalos dos parâmetros foram divididos em  $p=4$  níveis. A região de experimentação foi, portanto, um grid de 43 dimensões e quatro níveis. As nove variáveis de saída do modelo e os 43 parâmetros considerados neste trabalho são listados nas Tabelas 1 e 2.

**Tabela 1:** Variáveis simuladas pelo IBIS analisadas.

Nome	Descrição
$f_{APAR}$	Fração absorvida da radiação fotossinteticamente ativa (adimensional)
H	Fluxo de calor sensível ( $Wm^{-2}$ )
LAI	Índice de área foliar ( $m^2m^{-2}$ )
LE	Fluxo de calor latente ( $Wm^{-2}$ )
NEE	Troca líquida de $CO_2$ do ecossistema ( $\mu mol m^{-2}s^{-1}$ )
NPP	Produção primária líquida ( $kg-C m^{-2}ano^{-1}$ )
$PAR_o$	Radiação fotossinteticamente ativa refletida ( $\mu mol m^{-2}s^{-1}$ )
Rn	Saldo de radiação ( $Wm^{-2}$ )
$u^*$	Velocidade de fricção do vento ( $ms^{-1}$ )

No método proposto por Morris, o efeito elementar do  $i$ -ésimo parâmetro é definido como uma variação incremental do parâmetro,  $d$ :

$$d_i(\mathbf{x}) = \frac{[y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - y(\mathbf{x})]}{\Delta} \quad (45)$$

Um número,  $r$ , de efeitos elementares,  $d_i$ , é calculado em pontos selecionados aleatoriamente do espaço de parâmetros. A média ( $\mu_i$ ) e o desvio padrão ( $\sigma_i$ ) dos  $r$  efeitos elementares foram usados por Morris como medidas da importância do parâmetro. A medida  $\mu_i$  é usada para detectar parâmetros com uma influência total na saída do modelo.  $\sigma_i$  é usado para detectar parâmetros envolvidos em interações com outro parâmetro ou aqueles cujo efeito é não-linear.

Neste trabalho foi utilizada a versão em módulo ( $\mu_i^*$ ) do método de Morris, como sugerido por *Saltelli et al.* (2004). A medida  $\mu_i^*$  evita a deficiência do método de Morris quando aplicado a modelos não-monotônicos. A importância do parâmetro foi, então, avaliada através das medidas  $\mu_i^*$  e  $\sigma_i$ .

**Tabela 2:** Parâmetros do IBIS analisados.

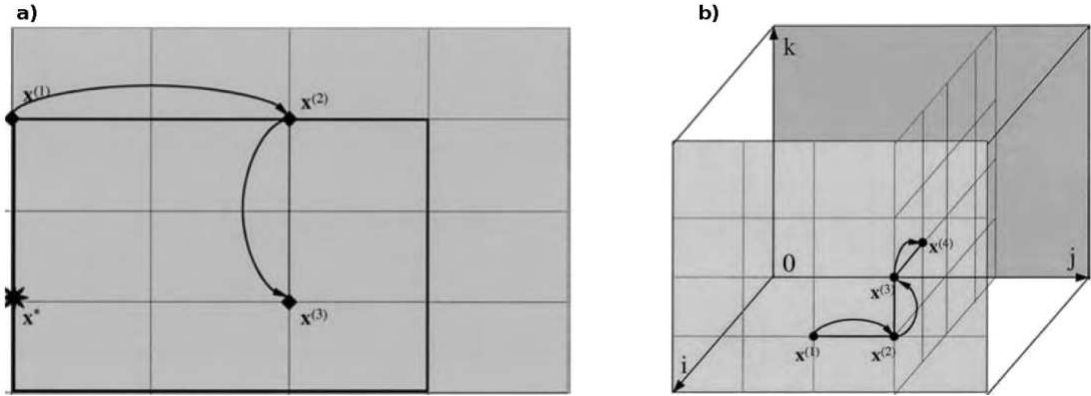
Nome	Descrição
aleaf	Fração de alocação de carbono nas folhas (adimensional)
alogl_coef	Coefficiente para o cálculo da rugosidade do dossel inferior (adimensional)
alogu_coef	Coefficiente para o cálculo da rugosidade do dossel superior (adimensional)
aroot	Fração de alocação de carbono nas raízes finas (adimensional)
awood	Fração de alocação de carbono na madeira (adimensional)
avmuir_coef	Coefficiente para o cálculo da emissividade do dossel (adimensional)
beta2	Parâmetro relacionado a distribuição de raízes finas (adimensional)
chifuz	Fator de orientação da folha no dossel superior (-1 vertical, 0 aleatório, 1 horizontal)
chl	Capacidade térmica das folhas e galhos do dossel inferior ( $J\ kg^{-1}m^{-2}$ )
chs	Capacidade térmica dos galhos do dossel superior ( $J\ kg^{-1}m^{-2}$ )
chu	Capacidade térmica das folhas do dossel superior ( $J\ kg^{-1}m^{-2}$ )
clitll_coef	Carbono inicial no reservatório de liteira de folha - lignina (kg-C $m^{-2}$ )
clitlm_coef	Carbono inicial no reservatório de liteira de folha - metabólico (kg-C $m^{-2}$ )
clitls_coef	Carbono inicial no reservatório de liteira de folha - estrutural (kg-C $m^{-2}$ )
clitrl_coef	Carbono inicial no reservatório de liteira de raízes finas - lignina (kg-C $m^{-2}$ )
clitrm_coef	Carbono inicial no reservatório de liteira de raízes finas - metabólico (kg-C $m^{-2}$ )
clitrs_coef	Carbono inicial no reservatório de liteira de raízes finas - estrutural (kg-C $m^{-2}$ )
clitwl_coef	Carbono inicial no reservatório de liteira de madeira - lignina (kg-C $m^{-2}$ )
clitwm_coef	Carbono inicial no reservatório de liteira de madeira - metabólico (kg-C $m^{-2}$ )
clitws_coef	Carbono inicial no reservatório de liteira de madeira - estrutural (kg-C $m^{-2}$ )
coefmub	Coefficiente m relacionado a condutância estomática (adimensional)
csoipas_coef	Carbono inicial no solo - húmus passivo (kg-C $m^{-2}$ )
csoislon_coef	Carbono inicial no solo - húmus não-prottegido lento (kg-C $m^{-2}$ )
csoislop_coef	Carbono inicial no solo - húmus protegido lento (kg-C $m^{-2}$ )
dispu_coef	Altura de deslocamento do plano zero para o dossel superior ( $m$ )
funca_coef	Coefficiente da função de temperatura da biomassa dos galhos (adimensional)
funcb_coef	Coefficiente da função de temperatura da biomassa das raízes (adimensional)
kfactor	Fator de multiplicação das constantes de declínio dos reservatórios de C (adimensional)
rhoveg_vis	Refletância da folha no dossel superior (Visível) (adimensional)
rhoveg_NIR	Refletância da folha no dossel superior (NIR) (adimensional)
rgrowth_coef	Coefficiente de respiração de crescimento (adimensional)
rroot_coef	Coefficiente de manutenção de respiração das raízes ( $s^{-1}$ )
rwood_coef	Coefficiente de manutenção de respiração dos troncos ( $s^{-1}$ )
specla	Área foliar específica ( $m^2kg^{-1}$ )
stressf_coef	Coefficiente relacionado ao stress hídrico do solo (adimensional)
tauleaf	Constante de tempo de retorno da biomassa de folhas ( <i>anos</i> )
tauroot	Constante de tempo de retorno da biomassa de raízes finas ( <i>anos</i> )
tauwood0	Constante de tempo de retorno da biomassa de madeira ( <i>anos</i> )
tauveg_vis	Transmitância da folha no dossel superior (Visível) (adimensional)
tauveg_NIR	Transmitância da folha no dossel superior (NIR) (adimensional)
tempvm_coef	Parâmetro de estresse térmico do Vmax (adimensional)
vmax_pft	Capacidade de atividade máxima da enzima Rubisco ( $mol-CO_2m^{-2}s^{-1}$ )
wsoi_coef	Umidade do solo inicial (adimensional)

### 3.4.1 Implementação da AS

A forma mais simples de se calcular os  $r$  efeitos elementares de um parâmetro é através da avaliação de  $y$  duas vezes para cada  $d_i$ . Assim, o custo computacional total para uma amostra aleatória de  $r$  valores  $d_i$  é  $n = 2kr$  simulações do modelo ( $k$  é o número de parâmetros).

Morris propôs um modelo de cálculo mais eficiente, que usa algumas simulações do modelo para computar mais que um efeito elementar. O método proposto por Morris é baseado na construção de  $r$  trajetórias, definidas como a

sucessão de  $(k + 1)$  pontos no espaço dos parâmetros de entrada. As Figuras 8 (a) e (b) mostram dois exemplos de trajetórias em espaços de entrada quando  $k=2$  e  $k=3$ , respectivamente. Cada trajetória fornece  $k$  efeitos elementares, um por parâmetro, e define uma Matriz de Orientação ( $B^*$ ), com dimensões  $(k + 1)k$ , cujas linhas são os pontos  $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^{k+1}$ . O custo computacional nesse caso é  $n = (k + 1)r$ . Neste trabalho, utilizou-se  $r = 4$ , como sugerido por *Saltelli et al.* (2000). Logo,  $n = (43 + 1)4 = 176$ .



**Figura 8:** Dois exemplos de trajetórias no espaço dos parâmetros quando (a)  $k=2$  e (b)  $k=3$ . Adaptado de *Saltelli et al.* (2004).

Para construir a matriz  $B^*$  foi usado o seguinte procedimento, conforme Morris (1991):

1. Definir  $B$  como uma matriz triangular estritamente inferior de 1s.
2. Assumir que  $D^*$  é uma matriz diagonal de  $k$  dimensões, na qual cada elemento é  $+1$  ou  $-1$ , com igual probabilidade.
3. Assumir que  $P^*$  é uma matriz de permutação aleatória  $k \times k$ , 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. Construir uma matriz de orientação aleatória  $B^*$  dada por

$$\mathbf{B}^* = \left( \mathbf{J}_{k+1,1} \mathbf{x}^* + \left( \frac{\Delta}{2} \right) \mid (2\mathbf{B}\mathbf{P}^* - \mathbf{J}_{k+1,k}) \mathbf{D}^* + \mathbf{J}_{k+1,k} \mid \right) \quad (46)$$



onde  $J_{k+1,k}$  é uma matriz  $(k+1) \times k$  de 1s, e  $\mathbf{x}^*$  é um valor base de X escolhido aleatoriamente. A Equação 46 representa uma pequena modificação do modelo proposto por Morris. Ela simplifica o processo de permutação de parâmetros diferentes (em termos de escalas, intervalos e unidades).

$B^*$  fornece um efeito elementar, aleatoriamente selecionado, por parâmetro. Foram construídas  $r = 4$  matrizes  $B^*$  (como sugerido por *Saltelli et al. (2000)*), geradas independentemente, correspondendo a quatro trajetórias no espaço dos parâmetros. Então, os quatro efeitos elementares de cada parâmetro foram utilizados para calcular as medidas de sensibilidade ( $\mu_i^*$  e  $\sigma_i$ ).

Para evitar o problema de se analisar parâmetros com unidades ou intervalos de escalas diferentes, e para atribuir um significado às medidas de sensibilidade geradas pelo método, foram propostas modificações no cálculo do efeito elementar da Equação 45, ficando o novo  $d$  definido da seguinte forma:

$$d_i(\mathbf{x}) = \frac{[y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_k) - y(\mathbf{x})]}{\bar{P}} \quad (47)$$

onde  $\bar{P}$  é o valor médio da saída do modelo, produzido variando-se os parâmetros no seus respectivos intervalos.

Esse novo  $d_i$  descreve a influência do parâmetro em relação ao valor médio da saída do modelo.

Para identificar os parâmetros mais importantes para a calibração, os parâmetros foram classificados de acordo com um índice de sensibilidade para calibração, calculado pela soma da média dos efeitos elementares com um desvio padrão:

$$I = \mu_i^* + 1.0\sigma_i \quad (48)$$

Esse índice deve fornecer uma boa medida do potencial do parâmetro para calibração, já que ele classifica os parâmetros levando em conta tanto os seus efeitos gerais nas saídas do modelo quanto seus níveis de não-linearidade ou interação com outros parâmetros.

### 3.4.2 Validação da AS

A validação do método de AS proposto consistiu em verificar se o método é realmente capaz de identificar os parâmetros do modelo mais importantes para calibração.

A importância de um parâmetro para calibração é definida de acordo com o índice de sensibilidade proposto (I). Quanto maior o valor desse índice, maior a sensibilidade da saída do modelo ao parâmetro.

Para validar o método, foram desenvolvidos dois grupos, A e B, de testes de calibrações com algumas variáveis simuladas pelo modelo –  $PAR_o$ ,  $R_n$ ,  $H$ ,  $LE$  e  $NEE$ . No grupo A, as variáveis foram calibradas utilizando todos os parâmetros com alguma influência nessas variáveis. Assim, todos parâmetros que apresentaram o índice de sensibilidade não nulo ( $I > 0$ ) foram considerados para calibração.

No grupo B, foram considerados apenas os parâmetros mais importantes para as saídas do modelo, detectados na AS. Para definir os parâmetros mais importantes, foram feitas calibrações com o mínimo possível de parâmetros para se obter resultados similares aos obtidos nos experimentos do grupo A. Ou seja, para uma dada saída do modelo, foi feita a calibração apenas com o parâmetro mais importante para essa saída, de acordo com o índice I. Se o resultado não foi próximo ao obtido no grupo A (com uma diferença maior que 10% do valor obtido na calibração do grupo A), incorpora-se à calibração o próximo parâmetro mais importante e repete-se o processo. Isso é repetido até obter-se o conjunto de parâmetros que forneça um resultado de calibração similar ao do grupo A.

A Tabela 3 mostra os parâmetros considerados em cada calibração dos dois grupos. Os resultados que levaram ao grupo B são discutidos na seção 4.1.

O método de AS foi validado comparando-se o desempenho das calibrações dos dois grupos. Se a calibração que utiliza apenas os parâmetros mais importantes (grupo B) for tão eficaz quanto a calibração que utiliza todos os parâmetros (grupo A), então o método de AS é capaz de detectar os parâmetros mais influentes.

**Tabela 3:** Parâmetros do modelo calibrados nos testes dos grupos A e B.

Nome	Grupo A					Grupo B				
	PAR <sub>o</sub>	Rn	H	LE	NEE	PAR <sub>o</sub>	Rn	H	LE	NEE
rhoveg_vis	x	x	x			x		x		
rhoveg_NIR	x	x	x	x	x		x	x		
tauveg_vis	x	x	x			x				
tauveg_NIR	x	x	x	x	x		x	x		
chifuz	x	x	x	x	x	x	x	x	x	x
vmax_pft	x	x	x	x	x			x	x	
coefmub	x	x	x	x	x			x	x	x
chs	x	x	x	x	x			x	x	
chu		x								
chl										
beta2	x	x	x	x	x		x	x	x	x
funca_coef	x				x					
funcb_coef	x	x	x	x	x					
rroot_coef	x	x	x	x	x			x		
rwood_coef	x				x					
rgrowth_coef	x	x	x	x	x		x	x	x	x
tempvm_coef	x	x	x	x	x			x	x	
stressf_coef	x			x	x					
clitll_coef										
clitlm_coef					x					
clitls_coef					x					x
clitrl_coef										
clitrm_coef					x					
clitrs_coef					x					x
clitwl_coef					x					
clitwm_coef					x					
clitws_coef					x					x
csoipas_coef					x					
csoislon_coef					x					x
csoislop_coef					x					x
wsoi_coef	x	x	x	x	x			x	x	x
kfactor					x					x
tauleaf	x	x	x	x	x		x	x	x	x
tauroot	x	x	x	x	x					
tauwood0					x					
specla	x	x	x	x	x			x		
aleaf	x	x	x	x	x					
aroot	x		x		x					
awood					x					
dispu_coef		x	x	x	x			x	x	
alogl_coef										
alogu_coef			x		x					
avmuir_coef	x	x	x	x	x		x	x		
<b>Total</b>	<b>23</b>	<b>21</b>	<b>22</b>	<b>19</b>	<b>36</b>	<b>3</b>	<b>7</b>	<b>16</b>	<b>10</b>	<b>12</b>

### 3.5 Calibração multiobjetivo

O problema da calibração de LSMs pode ser visto como um problema de otimização multiobjetivo de diversas ópticas diferentes. Pode-se calibrar várias variáveis simuladas pelo modelo simultaneamente, e nesse caso, o ajuste de cada variável torna-se um objetivo do problema de otimização.

Outra forma de se tornar a calibração um problema multiobjetivo é tentar otimizar mais de uma estatística ou medida de erro simultaneamente para uma

variável simulada pelo modelo. Nesse caso, cada medida de ajuste da variável analisada torna-se um objetivo do problema de otimização.

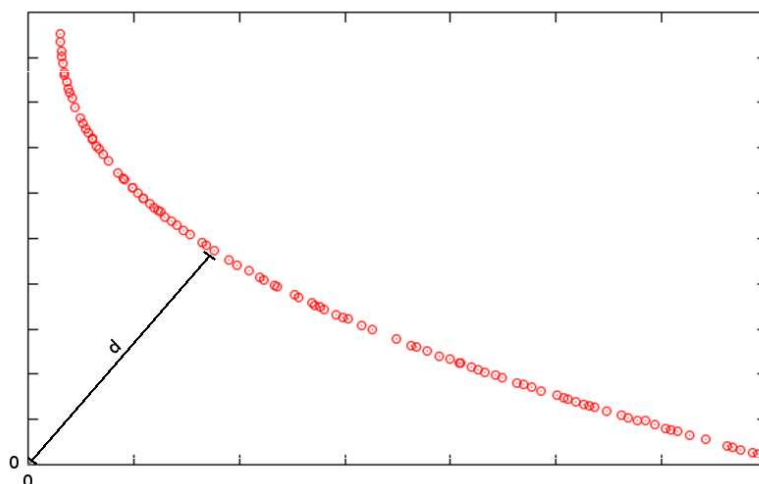
Existe ainda a possibilidade de fundir as duas abordagens de calibração, buscando calibrar várias variáveis simuladas pelo modelo através da otimização de várias estatísticas simultaneamente. Nesse caso, depara-se com um problema de otimização com um número muito grande de objetivos, possivelmente intratável.

Neste trabalho foi desenvolvido um software para a otimização multiobjetivo do modelo IBIS, chamado de Optis. O Optis pode ser utilizado para tratar qualquer um dos três tipos de calibrações citadas acima, e é baseado no algoritmo de otimização multiobjetivo NSGA-II (*Deb et al.*, 2002). NSGA-II é a sigla em inglês para Nondominated Sorting Genetic Algorithm II, que pode ser traduzido como Algoritmo Genético de Classificação Não-dominada II. Esse algoritmo é uma versão revisada e melhorada de um dos primeiros algoritmos genéticos multiobjetivos desse tipo, o NSGA (*Srinivas e Deb*, 1995).

O algoritmo NSGA-II baseia-se no conceito de fronteira de Pareto para encontrar as soluções ótimas do problema de otimização multiobjetivo. Após definido o conjunto de soluções da fronteira de Pareto, o Optis retorna como a melhor solução aquela que otimiza todos os objetivos de forma simétrica, ou seja, é selecionada a solução que se encontra mais próxima da origem do gráfico dos objetivos (Figura 9). Com isso, opta-se por atribuir igual peso a todos os objetivos.

Para evitar a interferência das unidades ou escalas dos valores das funções objetivos, é feita uma normalização dos valores retornados pelo algoritmo, antes da escolha do valor ótimo. Com esta normalização, todos os valores dos objetivos das soluções encontradas (fronteira de Pareto) estarão compreendidos entre zero (0,0) e um (1,0). Após isso, o ponto com a menor distância euclidiana da origem é selecionado como a melhor solução do problema de otimização.

O algoritmo possui uma limitação ao se trabalhar com um número grande de objetivos. Há uma perda de eficiência na otimização ao se trabalhar com mais de três objetivos. Essa é uma limitação do algoritmo NSGA-II, como apresentado



**Figura 9:** Exemplo da escolha da solução ótima de um problema de otimização com 2 objetivos. Nesta tese, a solução com menor  $d$  é considerada a melhor solução.

por Praditwong e Yao (2007). Mesmo com o aumento no tamanho da população do AG, o NSGA-II produz resultados pobres com números maiores de objetivos (Purshouse e Fleming, 2003). Dessa forma, neste trabalho as calibrações multiobjetivo foram restringidas a no máximo três objetivos simultâneos.

### 3.5.1 Estrutura do Optis

O Optis utiliza como base o código em linguagem C do NSGA-II, disponível no site do Kangal (<http://www.iitk.ac.in/kangal/>). O software foi desenvolvido para funcionar integrado com o modelo IBIS, e otimizar os parâmetros do modelo de forma totalmente automática. No processo de calibração, o modelo IBIS funciona de forma independente do Optis, e é executado externamente ao algoritmo de otimização. O algoritmo de otimização interage com o modelo apenas através da alteração dos seus parâmetros de entrada e da leitura dos seus dados de saída. Dessa forma, o Optis torna-se praticamente independente da versão do IBIS, ou mesmo independente do próprio modelo a ser calibrado. Apenas uma pequena modificação na forma de leitura dos parâmetros do IBIS foi necessária, para tornar o processo de leitura e escrita desses parâmetros mais simples, tanto

pelo Optis quanto pelo IBIS.

O Optis é totalmente configurável, e permite a realização de infinitos tipos de experimentos de calibração, possibilitando calibrar qualquer combinação de variáveis simuladas pelo modelo, a utilização de várias medidas de ajuste do modelo, bem como a calibração de qualquer combinação dos 43 parâmetros do modelo. Além disso, é possível desenvolver calibrações independentes sequenciais do IBIS, de forma que várias variáveis do modelo sejam calibradas uma após a outra, mantendo-se os valores dos parâmetros obtidos em cada uma dessas calibrações. A configuração dos experimentos de calibração é feita através do arquivo *calibration.conf* (Apêndice B). Uma descrição detalhada do Optis está disponível no Apêndice A.

### 3.5.2 Paralelização do algoritmo

Como são necessárias milhares de simulações do modelo durante o processo de calibração, foi feita uma paralelização do algoritmo do Optis. Durante a calibração, as simulações do modelo para o cálculo das funções objetivos são independentes entre si. Assim, o algoritmo foi paralelizado de forma a disparar as simulações do IBIS simultaneamente, de acordo com a quantidade de processadores disponíveis. Foram realizados testes de eficiência com quatro e oito processadores, e os resultados foram comparados com a utilização de apenas um processador (algoritmo escalar).

### 3.5.3 Teste do algoritmo de calibração

Para testar e validar o programa Optis, foram desenvolvidos dois testes de calibração multiobjetivo do IBIS. No primeiro teste, uma variável do modelo (LE) foi calibrada independentemente, considerando-se duas medidas de ajuste como objetivos ( $MAE_{LE}$  e  $Bmax_{LE}$ ). Nesse caso, a multiobjetividade do problema é relacionada à quantidade de medidas de ajuste (duas medidas de ajuste  $\rightarrow$  dois objetivos).

No segundo teste de calibração multiobjetivo, duas variáveis do modelo (H e LE) foram calibradas simultaneamente, em relação a uma única medida de ajuste de cada uma ( $MAE_H$  e  $MAE_{LE}$ ). O algoritmo otimizou os parâmetros de forma a encontrar os valores que ajustassem as duas variáveis ao mesmo tempo, com igual peso para cada uma delas. Nesse caso, a multiobjetividade do problema está na quantidade de variáveis do modelo ajustadas (duas variáveis  $\rightarrow$  dois objetivos).

Para avaliar o desempenho do algoritmo em otimizar múltiplos objetivos simultaneamente, as calibrações multiobjetivo foram comparadas com as respectivas calibrações mono-objetivos. Nas calibrações mono-objetivo cada variável é exclusivamente otimizada, ou seja, em cada calibração todos os parâmetros são calibrados de forma a otimizar uma única variável. Assim, assume-se que os resultados obtidos nas calibrações mono-objetivo são os melhores resultados possíveis, já que nesse caso cada objetivo é totalmente priorizado.

Para avaliar o ajuste geral do modelo obtido na calibração multiobjetivo, foi calculado um índice de desempenho relativo (D), que mede o desempenho da calibração multiobjetivo em relação às calibrações mono-objetivo. Para calcular o índice D, ao final da calibração multiobjetivo calcula-se a razão entre os resultados das calibrações mono-objetivo e multiobjetivo, de cada função objetivo otimizada. Esses valores individuais indicam o quanto cada objetivo está próximo do seu valor ótimo (calibração mono-objetivo). O índice D é calculado fazendo-se a média de todas as razões mono-objetivo/multiobjetivo calculadas. Quanto maior for o valor de D, maior será a eficiência da calibração multi-objetivo em relação à mono-objetivo.

O índice de desempenho relativo é, portanto, definido pela fórmula:

$$D = \frac{1}{n} \sum_{i=1}^n \frac{f_i^{mono}}{f_i^{multi}} \quad (49)$$

onde  $f^{mono}$  é a função objetivo usada para avaliar o modelo sujeito à calibração mono-objetivo, e  $f^{multi}$  é a função objetivo usada para avaliar o modelo sujeito à calibração multiobjetivo.

### 3.6 Calibração hierárquica multiobjetivo

Para obter o desempenho desejado de um LSM, é necessário calibrar o modelo de forma que todos os seus processos sejam simulados realisticamente. Porém, calibrar os vários processos simulados pelo IBIS em um único experimento de otimização multiobjetivo seria impraticável, devido à grande quantidade de funções objetivos (10) a serem consideradas simultaneamente. Esse número de funções objetivo depende das observações de campo disponíveis, e vai variar de sítio para sítio.

Os sistemas ambientais são bastante complexos e seus processos são organizados através de várias escalas de espaço e de tempo (*O'Neill, 1988; Turner et al., 1989*). Os modelos de ecossistemas terrestres comumente são organizados de forma hierárquica, e representam os processos que ocorrem nas diferentes escalas (*Luan et al., 1996*). Por exemplo, o modelo IBIS (*Foley et al., 1996*) simula desde fluxos de energia e massa (escala de tempo: segundos) até o balanço de carbono no solo e dinâmica da vegetação, incluindo competição entre espécies (escala de tempo: décadas). Os parâmetros do modelo, portanto, estão relacionados com diferentes processos de diversas escalas temporais. Cada parâmetro pode influenciar um ou vários processos de uma mesma escala, ou influenciar processos de diferentes escalas.

Nesse trabalho propôs-se usar uma abordagem hierárquica temporal para calibrar cada processo, ou grupo de processos, separadamente, de acordo com seu nível hierárquico. Dessa forma, foi possível calibrar 43 parâmetros do modelo, otimizando nove variáveis de saída.

O nível hierárquico é definido pela hierarquia temporal dos processos simulados, começando dos processos mais rápidos para os mais lentos. No primeiro nível foram calibrados os fluxos radiativos de ondas curtas, no segundo nível o saldo de radiação da superfície, no terceiro nível a turbulência, no quarto nível os fluxos turbulentos e no quinto e último nível as variáveis relacionadas a alocação de carbono. O método permite a calibração de um número ilimitado de funções objetivo



e de variáveis de saída, distribuídas em qualquer número de níveis hierárquicos, desde que não ultrapasse três funções objetivo por nível hierárquico.

Através da análise de sensibilidade e do conhecimento prévio das relações entre os parâmetros e os processos simulados, foi feito um agrupamento dos parâmetros relacionados a cada variável de saída do IBIS. Assim, foi feita a calibração de cada processo, ou grupo de processos, como um procedimento independente, diminuindo o número de funções objetivo em cada um desses subprocedimentos de calibração. Os valores dos parâmetros obtidos em cada nível de calibração são mantidos no nível seguinte, terminando com todo o conjunto de parâmetros, de todos os níveis, ajustados.

Foram feitas calibrações do modelo para os sítios da Flona do Tapajós km 67 e de Sinop. No sítio do Tapajós os níveis hierárquicos de calibração ficaram definidos da seguinte forma:

1º nível:  $PAR_o$  e  $f_{APAR}$

2º nível:  $R_n$

3º nível:  $u^*$

4º nível:  $NEE$ ,  $H$  e  $LE$

5º nível:  $LAI$  e  $NPP$

Para o sítio de Sinop, a hierarquia definida foi:

1º nível:  $PAR_o$  e  $f_{APAR}$

2º nível:  $R_n$

3º nível:  $u^*$

4º nível:  $NEE$ ,  $H$  e  $LE$

5º nível:  $LAI$

Nos níveis hierárquicos em que haviam apenas uma variável de saída do modelo e que o uso do  $Bmax$  era aplicável foi possível utilizar as duas medidas de ajuste ( $MAE$  e  $Bmax$ ). Nos demais níveis foi utilizado apenas o  $MAE$ , para

manter o máximo de três funções objetivos por nível. A Tabela 4 apresenta as funções objetivo e os parâmetros calibrados em cada nível hierárquico.

**Tabela 4:** Funções objetivo e os parâmetros calibrados em cada nível hierárquico, nos sítios da Flona do Tapajós e de Sinop. Para o sítio de Sinop, no 5<sup>o</sup> nível hierárquico apenas o LAI é calibrado.

Nível Hierárquico	Objetivo calibrado	Parâmetro
1	$\{MAE_{PAR_o}, f_{APAR}\}$	rhoveg_vis tauveg_vis chifuz
2	$\{MAE_{Rn}, Bmax_{Rn}\}$	rhoveg_NIR tauveg_NIR avmuir_coef
3	$\{MAE_{u*}\}$	dispu_coef alogl_coef alogu_coef
4	$\{MAE_{NEE,H,LE}\}$	vmax_pft coefmub chs beta2 funca_coef funcb_coef rroot_coef rwood_coef tempvm_coef stressf_coef clitls_coef clitrs_coef clitws_coef csoislon_coef csoislop_coef kfactor
5	$\{MAE_{LAI,NPP}\}$	rgrowth_coef tauleaf specla aleaf aroot awood

Para o sítio da Flona do Tapajós o modelo foi calibrado para o período de janeiro 2002 a dezembro de 2004. No sítio de Sinop, o modelo foi calibrado para o ano de 2002. Os dados de Sinop possuíam muitas falhas, tanto nos dados utilizados para alimentar o modelo quanto nos dados para validação dos resultados. Com isso, foi feito o preenchimento de falhas por longos períodos de tempo, o que certamente diminui a confiabilidade da calibração.

Os resultados das calibrações hierárquicas foram comparados com valores de calibrações de referência, para cada variável. Os valores adotados como referência

são os valores obtidos com as calibrações mono-objetivo das respectivas variáveis. Nas calibrações mono-objetivo cada variável é exclusivamente otimizada, ou seja, em cada calibração todos os parâmetros são calibrados de forma a otimizar uma única variável.

Para avaliar o ajuste geral do modelo obtido na calibração hierárquica, foi calculado o índice de desempenho relativo  $D$ , análogo ao calculado na avaliação da calibração multiobjetivo (seção 3.5.3). Assim, o desempenho da calibração hierárquica é avaliado em relação às calibrações de referência mono-objetivo.

Para a calibração hierárquica, o índice de desempenho relativo é definido da seguinte forma:

$$D = \frac{1}{n} \sum_{i=1}^n \frac{f_i^m}{f_i^h} \quad (50)$$

onde  $f^m$  é a função objetivo usada para avaliar o modelo sujeito à calibração mono-objetivo, e  $f^h$  é a função objetivo usada para avaliar o modelo sujeito à calibração hierárquica.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Análise de sensibilidade do IBIS

Dos 43 parâmetros analisados, apenas 23 apresentaram influência significativa em alguma das saídas do modelo analisadas. Isso diminui consideravelmente o espaço de busca, o que implica em calibrações mais eficientes. As sensibilidades de cada saída do modelo aos 43 parâmetros é mostrada nas Figuras 10 a 18.

A Figura 10 mostra a sensibilidade da  $PAR_o$  aos parâmetros do modelo. Dos 43 parâmetros considerados na análise, apenas 23 apresentaram alguma influência ( $I > 0$ ) no  $PAR_o$ . Desses 23 parâmetros, destaca-se o  $rhoveg\_vis$ , que apresentou maior índice de sensibilidade, seguido dos parâmetros  $chifuz$  e  $tauveg\_vis$ .

A Figura 11 apresenta os resultados da AS para a  $f_{APAR}$ . Apenas 8 parâmetros apresentaram  $I$  maior que zero, com apenas 3 parâmetros ( $rhoveg\_NIR$ ,  $tauveg\_NIR$  e  $chifuz$ ) com índice de sensibilidade significante.

Na Figura 12 são apresentados os resultados da AS do  $Rn$ . Dos 43 parâmetros do modelo, 21 apresentaram alguma influência no  $Rn$ , sendo os parâmetros  $rhoveg\_NIR$ ,  $chifuz$ ,  $avmuir\_coef$  e  $tauveg\_NIR$  sendo os com maior valor de  $I$ .

A Figura 13 mostra a sensibilidade do  $u^*$  aos 43 parâmetros do IBIS. 26

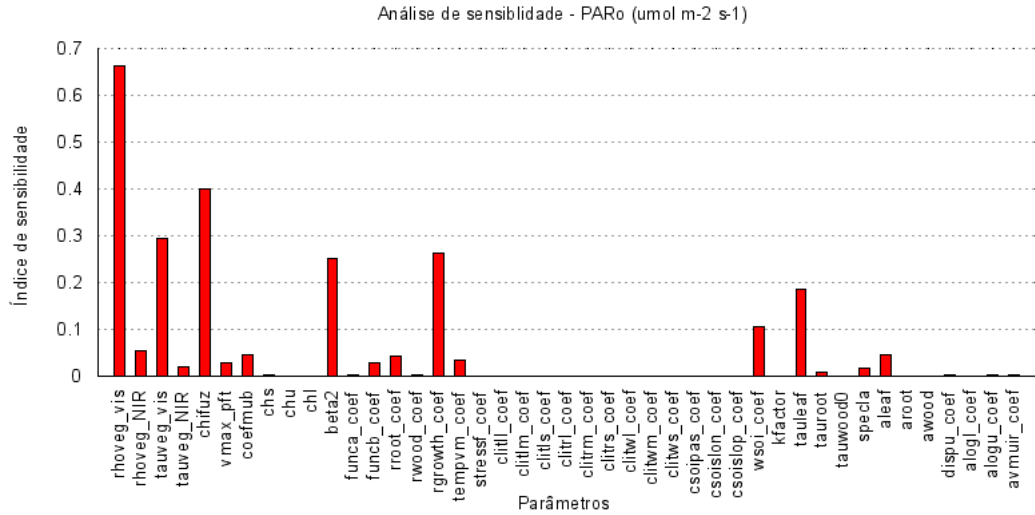


Figura 10: Análise de sensibilidade da  $PAR_o$  aos 43 parâmetros do modelo.

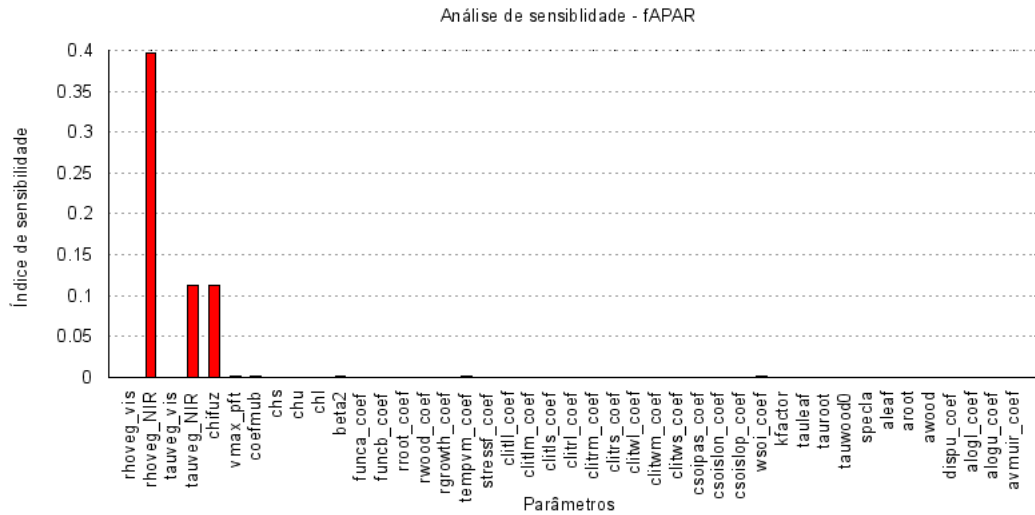
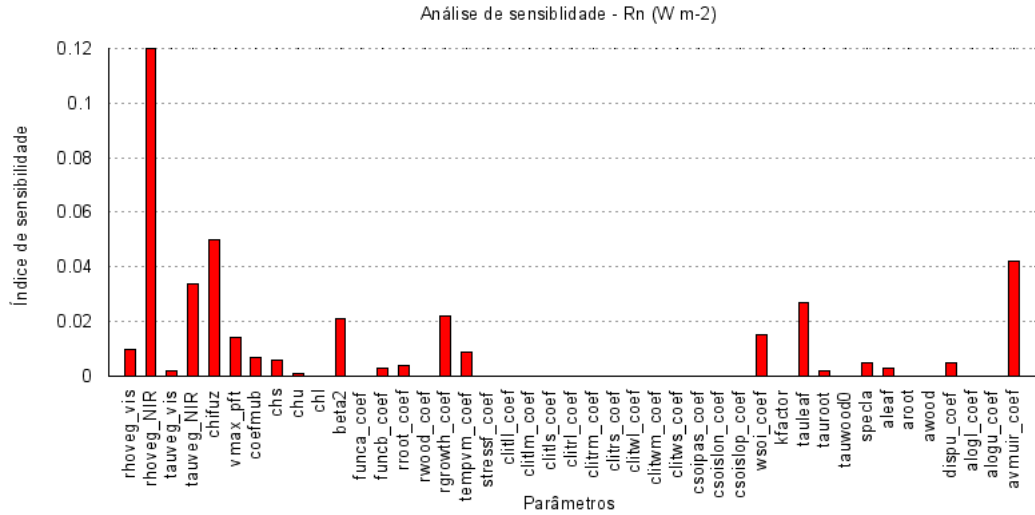


Figura 11: Análise de sensibilidade da  $f_{APAR}$  aos 43 parâmetros do modelo.

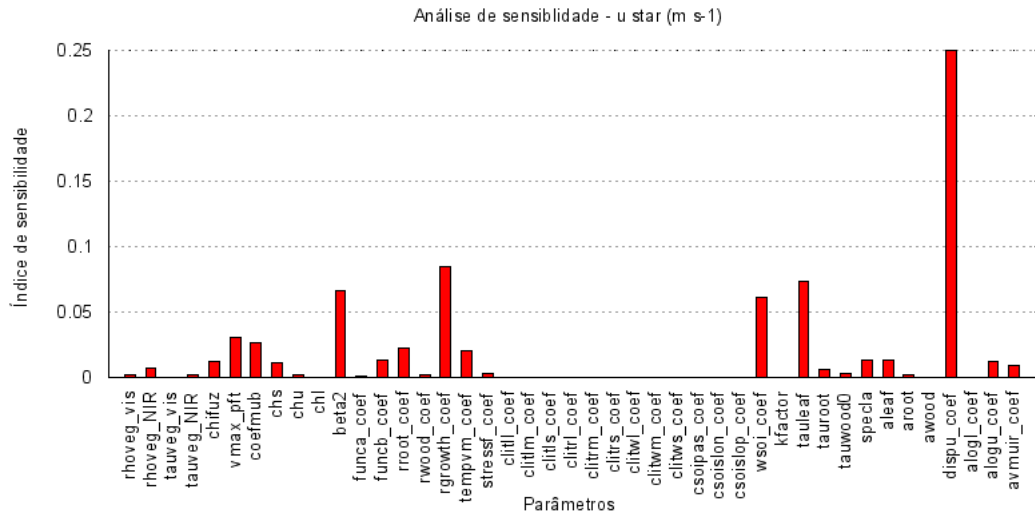
parâmetros apresentaram o índice de sensibilidade maior que zero, destacando-se o parâmetro  $dispu\_coef$  com a maior influência (0,25) no  $u^*$ .

Os resultados para o H são apresentados na Figura 14. Dos 22 parâmetros que apresentaram alguma influência no H, o mais importante é o  $rhoveg\_NIR$ , com um índice I de 0,09.

A Figura 15 apresenta os resultados da AS para o LE. 19 parâmetros apresentaram influência no LE, com os maiores índices de sensibilidade nos parâmetros  $beta2$  e  $wsoi\_coef$ .



**Figura 12:** Análise de sensibilidade do Rn aos 43 parâmetros do modelo.



**Figura 13:** Análise de sensibilidade do  $u^*$  aos 43 parâmetros do modelo.

Na Figura 16 é apresentado o resultado da AS para o NEE. Um total de 36 parâmetros, dos 43 analisados, apresentaram influência no NEE simulado. Nove parâmetros apresentaram um índice de sensibilidade maior que 0,02: *clitrs\_coef*, *clitls\_coef*, *clitws\_coef*, *wsoi\_coef*, *csoislop\_coef*, *csoislon\_coef*, *kfactor*, *rgrowth\_coef* e *beta2*.

A Figura 17 mostra a sensibilidade da NPP simulada aos parâmetros do modelo. Dos 43 parâmetros inicialmente considerados, 29 apresentaram o índice I maior que zero. O parâmetro com maior influência na saída foi o *rgrowth\_coef*,

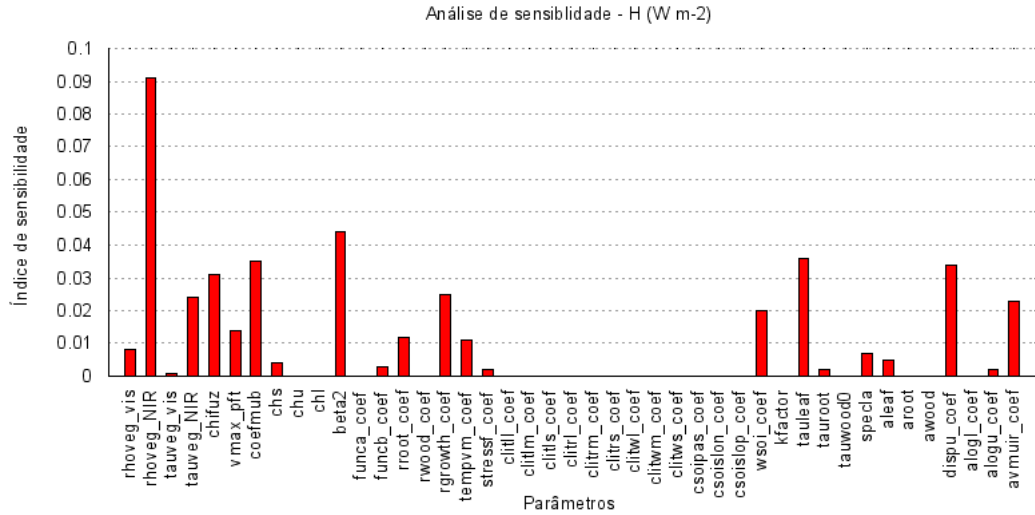


Figura 14: Análise de sensibilidade do H aos 43 parâmetros do modelo.

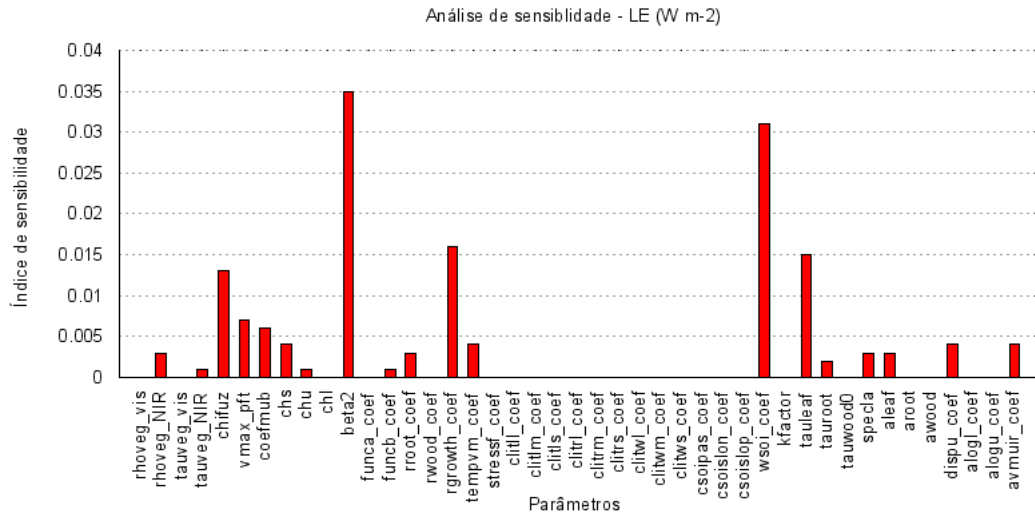


Figura 15: Análise de sensibilidade do LE aos 43 parâmetros do modelo.

que apresentou um I igual a 0,2.

Finalmente, a Figura 18 mostra os resultados da AS do LAI. 29 parâmetros apresentaram alguma influência nessa saída, destando-se os parâmetros beta2 e rgrowth\_coef com os maiores valores de sensibilidade.

Como se observa, várias variáveis simuladas pelo modelo podem ser sensíveis a um mesmo parâmetro. Isso geralmente ocorre porque os sistemas naturais, bem como o modelo IBIS, estão organizados de forma hierárquica, com processos ocorrendo em diferentes escalas temporais e espaciais, e com cada processo influ-

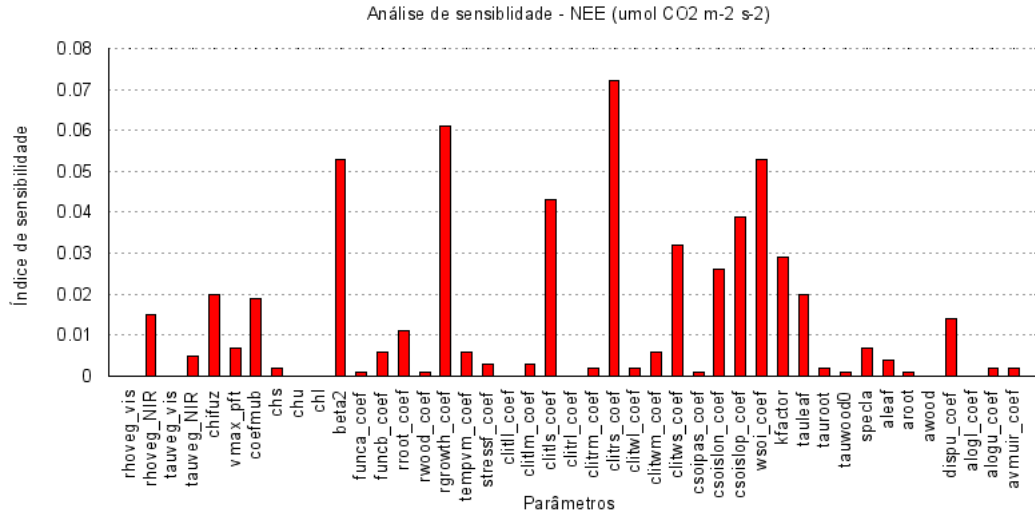


Figura 16: Análise de sensibilidade do NEE aos 43 parâmetros do modelo.

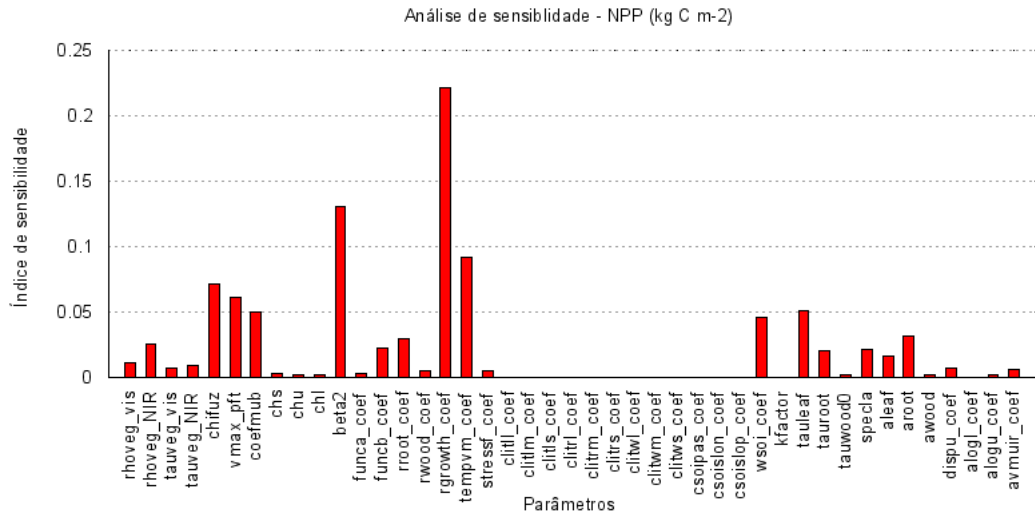


Figura 17: Análise de sensibilidade da NPP aos 43 parâmetros do modelo.

enciando um ao outro. Portanto, um parâmetro pode influenciar outras variáveis simuladas de forma indireta. Em um procedimento de calibração de várias saídas do modelo, a escolha final de qual parâmetro calibrar para cada saída deverá se basear também no conhecimento teórico dos processos simulados, além da análise de sensibilidade.

Para o cálculo da média  $\bar{P}$  (Equação 47) dos valores dos objetivos simulados pelo modelo, foram usados os valores produzidos durante o cálculo de  $y$  em cada ponto das trajetórias no espaço dos parâmetros. Portanto, o novo cálculo pro-



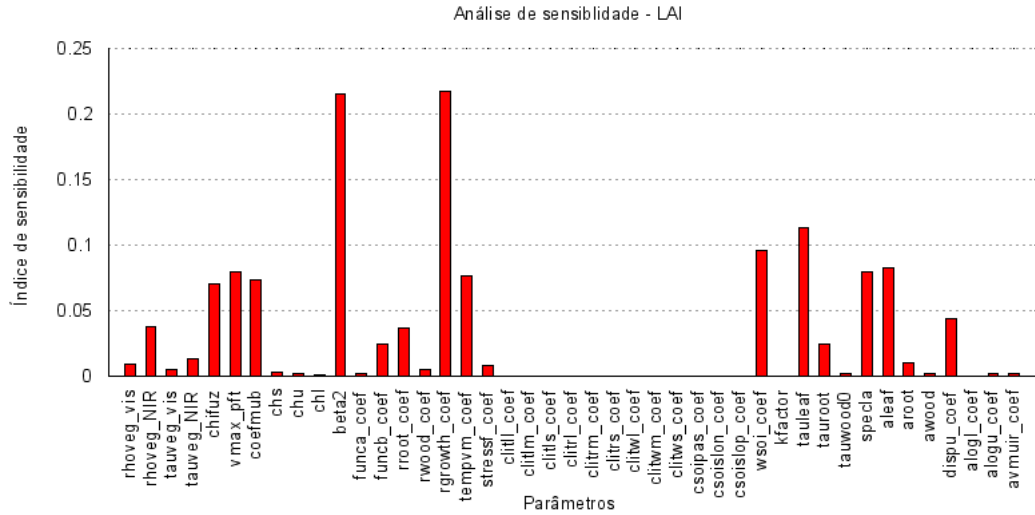


Figura 18: Análise de sensibilidade do LAI aos 43 parâmetros do modelo.

posto para o efeito elementar não diminui a eficiência computacional do método originalmente proposto por Morris, já que ele não requer nenhuma execução adicional do modelo.

#### 4.1.1 Validação da AS

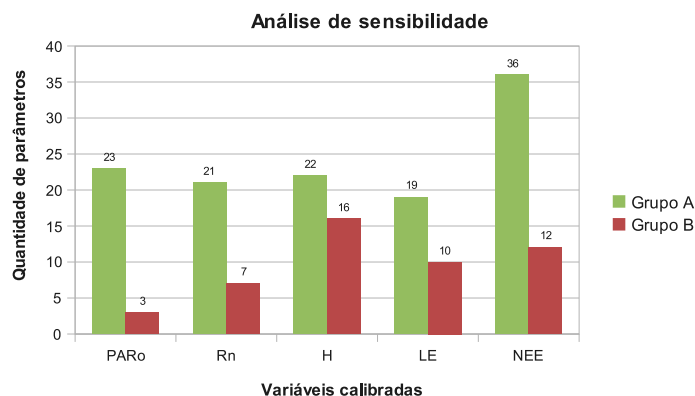
Os resultados da validação do método de AS estão apresentados na Figura 19 e na Tabela 5. A Figura 19 compara a quantidade de parâmetros utilizados nas calibrações do grupo A e B que forneceram resultados similares de calibração.

Tabela 5: Resultados da validação do método de AS proposto, com o número de parâmetros calibrados e o MAE de cada calibração.

	Grupo A		Grupo B		Diferença		
	Param.	MAE	Param.	MAE	Param.	MAE	Percentual
PAR <sub>o</sub> ( $\mu\text{molm}^{-2}\text{s}^{-1}$ )	23	1.05	3	1.05	-20	0.00	0%
Rn ( $W\text{m}^{-2}$ )	21	17.43	7	19.13	-14	1.71	10%
H ( $W\text{m}^{-2}$ )	22	22.52	16	22.29	-6	-0.23	-1%
LE ( $W\text{m}^{-2}$ )	19	30.49	10	31.11	-9	0.62	2%
NEE ( $\mu\text{molm}^{-2}\text{s}^{-1}$ )	36	3.71	12	3.70	-24	-0.01	0%

A calibração do PAR<sub>o</sub> ajustando apenas os três parâmetros mais importantes fornece o mesmo valor de MAE,  $1,05 \mu\text{molm}^{-2}\text{s}^{-1}$ , obtido quando se utiliza os 23 parâmetros com alguma influência, detectadas na AS.

A calibração do Rn com apenas sete parâmetros não foi capaz de otimizar o



**Figura 19:** Quantidade de parâmetros calibrados antes (grupo A) e após (grupo B) a análise de sensibilidade.

modelo tão bem quanto a calibração com todos os 21 parâmetros. Entretanto, o aumento no erro de  $1,71 \text{ W/m}^2$  (10%) está no limite do acréscimo do MAE para seleção das variáveis, e pode ser considerado pequeno em relação à grande diminuição no número de parâmetros ajustados. O mesmo ocorreu com os testes com o LE, onde o aumento no erro do modelo é praticamente desprezível. Esses casos ilustram bem a escolha entre complexidade e acurácia, que deve ser tomada por um analista ao lidar com calibração de modelos complexos.

Para o H, com os 16 parâmetros mais influentes foi possível obter um MAE igual a  $22,29 \text{ W/m}^2$ , enquanto que, utilizando todos os 22 parâmetros importantes detectados na AS, obteve-se um MAE de  $22,52 \text{ W/m}^2$ . Portanto, obteve-se um resultado um pouco melhor utilizando-se menos parâmetros. Isso ocorre porque a diminuição do espaço dos parâmetros permite uma busca mais eficiente do algoritmo de otimização, possibilitando uma calibração mais refinada.

Assim como com o H, o NEE foi melhor calibrado utilizando-se os 12 parâmetros mais importantes, fornecendo um MAE de  $3,70 \mu\text{molm}^{-2}\text{s}^{-1}$ . Utilizando-se todos os 36 parâmetros influentes detectados na AS, obteve-se um MAE de  $3,71 \mu\text{molm}^{-2}\text{s}^{-1}$ . A razão para essa melhora no resultado ao se diminuir o número de parâmetros, assim como no teste de calibração com o H, é uma busca mais refinada no menor espaço dos parâmetros.

Esses resultados mostram que o método de AS foi eficiente ao selecionar os

parâmetros mais importantes para a calibração do modelo.

## 4.2 Calibração multiobjetivo

### 4.2.1 Paralelização do algoritmo

A Tabela 6 apresenta os resultados dos testes da eficiência da paralelização do código do Optis. A eficiência obtida foi máxima, de 100%, já que o ponto crítico do algoritmo é justamente o cálculo da aptidão dos indivíduos da população do algoritmo genético. Como cada indivíduo da população corresponde a uma simulação do modelo, e cada cópia do modelo roda de forma independente do algoritmo de calibração, a paralelização desse trecho do algoritmo produziu um efeito bastante satisfatório.

**Tabela 6:** Resultados da paralelização do algoritmo de otimização. Comparação entre os tempos gastos por simulação do modelo e da eficiência da paralelização.

# de CPUs	# de simulações	Tempo (s)	Tempo/simulação (s)	Eficiência
1	40,000	98705	2.47	100%
4	40,000	24664	0.62	100%
8	40,000	12325	0.31	100%

### 4.2.2 Teste do algoritmo de calibração

Os gráficos da Figura 20 e a Tabela 7 comparam os resultados da calibração multiobjetivo, com múltiplas estatísticas, da variável LE. Como era de se esperar, observa-se que o melhor valor do MAE é obtido quando se minimiza o próprio objetivo  $\{MAE_{LE}\}$  (o nome envolvido entre  $\{\}$  indica uma referência ao objetivo do problema de otimização, e não ao valor da variável). Da mesma forma, o melhor valor de  $Bmax$  é obtido quando se otimiza o objetivo  $\{Bmax_{LE}\}$ . Ao desenvolver a otimização multiobjetivo  $\{MAE_{LE}, Bmax_{LE}\}$ , foi obtido um valor de MAE e  $Bmax$  intermediários aos encontrados nas duas otimizações mono-objetivo. Esse resultado segue o comportamento da curva de Pareto, ou seja, na otimização multiobjetivo houve uma troca entre os valores ótimos das funções objetivos.

De acordo com a razão mono-objetivo/multiobjetivo, observa-se que o valor final do MAE e do  $Bmax$  na calibração multiobjetivo é bastante próximo do valor obtido nas otimizações mono-objetivo, com um D de 0,92. Portanto, o algoritmo de otimização multiobjetivo foi bem eficiente ao calibrar múltiplas medidas de erro.

**Tabela 7:** MAE e  $Bmax$  dos dados de LE simulados e observados e a razão mono-objetivo/multiobjetivo utilizada para o cálculo do índice D. Valores de MAE calculados com o modelo não calibrado, após as calibrações mono-objetivo e após a calibração multiobjetivo.

	Não calibrado	{MAE <sub>LE</sub> }	{Bmax <sub>LE</sub> }	{MAE <sub>LE</sub> , Bmax <sub>LE</sub> }	Razão
MAE <sub>LE</sub> (Wm <sup>-2</sup> )	58,70	29,45	44,29	30,17	0,98
Bmax <sub>LE</sub> (MJm <sup>-2</sup> )	1.677,79	190,269	118,64	136,71	0,87
D					0,925

**Tabela 8:** MAE dos dados simulados e observados de H e LE, e a razão mono-objetivo/multiobjetivo utilizada para o cálculo do índice D. Valores de MAE calculados com o modelo não calibrado, após as calibrações mono-objetivo e após a calibração multiobjetivo.

	Não calibrado	{MAE <sub>H</sub> }	{MAE <sub>LE</sub> }	{MAE <sub>H</sub> , MAE <sub>LE</sub> }	Razão
MAE <sub>H</sub> (Wm <sup>-2</sup> )	67,75	22,85	29,58	24,66	0,93
MAE <sub>LE</sub> (Wm <sup>-2</sup> )	58,70	34,75	29,45	31,84	0,92
D					0,925

A Figura 21 e a Tabela 8 mostram os resultados do segundo teste de calibração multiobjetivo  $\{MAE_H, MAE_{LE}\}$ . Assim como no primeiro teste do algoritmo, os objetivos  $\{MAE_{LE}\}$  e  $\{MAE_H\}$  são conflitantes, ou seja, quando se minimiza um, ocorre um aumento no outro. O algoritmo de otimização também foi eficiente em encontrar um valor ótimo intermediário que otimiza ambas as variáveis simultaneamente, obtendo um D de 0,93.

Esses resultados mostram que o Optis é eficiente em solucionar problemas de calibração multiobjetivo, atingindo mais de 90% do desempenho obtido na calibração mono-objetivo.

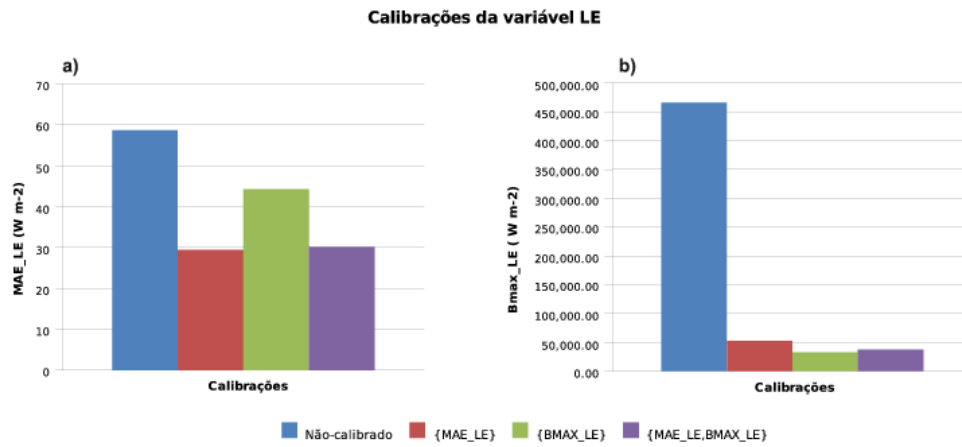


Figura 20: Resultados de a) MAE e b)  $Bmax$ , da calibração multiobjetivo  $\{MAE_{LE}, Bmax_{LE}\}$

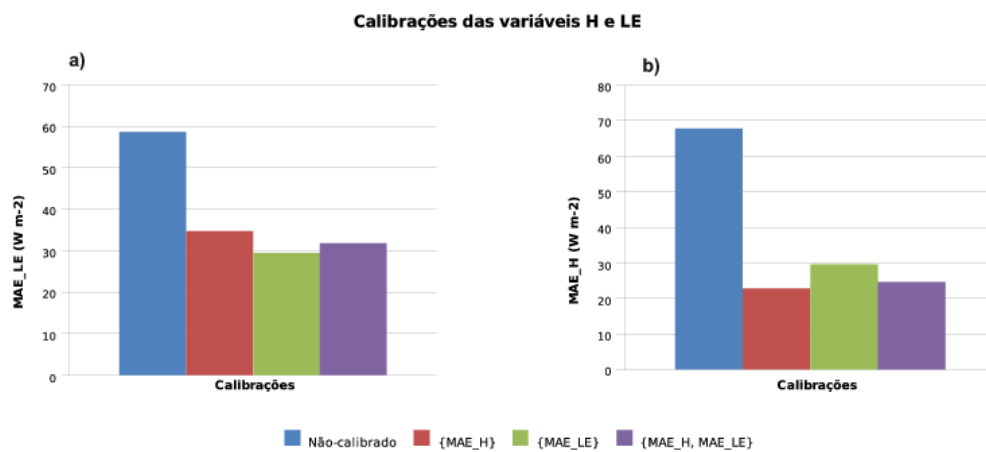


Figura 21: Resultados de a)  $MAE_{LE}$  e b)  $MAE_H$  da calibração multiobjetivo  $\{MAE_{LE}, MAE_H\}$

### 4.3 Calibração mono-objetivo – Flona do Tapajós km 67

Os resultados das calibrações mono-objetivo do MAE e  $B_{max}$  são apresentados na Tabela 9. As Figuras 22 a 29 mostram os resultados de simulações do modelo após a calibração mono-objetivo de cada variável, e compara as simulações com os dados observados.

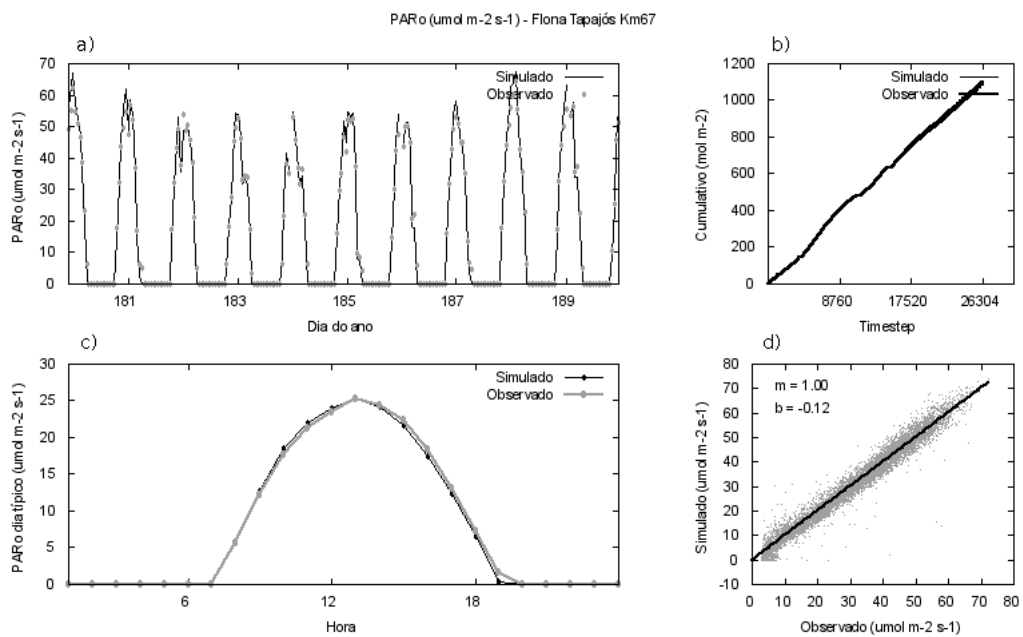
A Figura 22 mostra a  $PAR_o$  calibrada de forma mono-objetivo. Os gráficos indicam um ajuste praticamente perfeito entre os dados horários de  $PAR_o$ , com os coeficientes da reta de regressão linear  $m$  e  $b$  iguais a 1,00 e -0,12, respectivamente.

A Figura 23 mostra a  $f_{APAR}$  simulada após a calibração mono-objetivo. Apesar de os dados observados e simulados mensais estarem bem ajustados, os coeficientes da reta de regressão não apresentaram bons valores, influenciados pela pequena variação nos dados e pouca quantidade de pontos.

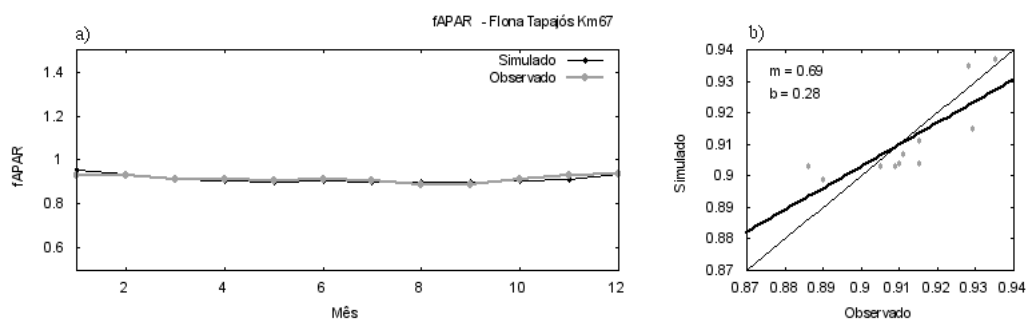
Na Figura 24 é apresentado o resultado do  $R_n$  calibrado pelo método mono-objetivo. Os dados de  $R_n$  simulados estão bem ajustados aos observados, com uma reta de regressão quase perfeita. Observando o gráfico do dia típico (Figura 24-c), percebe-se uma superestimativa do  $R_n$  simulado no período noturno, o que é evidenciado também pelo pequeno viés do gráfico da soma cumulativa (Figura 24-b). É possível que esse erro seja causado pela metodologia de estimativa de  $L_{in}$  (Equação 40).

A Figura 25 mostra  $u^*$  após a calibração mono-objetivo. Observa-se que houve uma subestimativa dos picos máximos e uma superestimativa dos valores mínimos, apesar do comportamento simétrico entre as duas séries. Isso pode indicar uma deficiência na simulação das correções da turbulência para condições atmosféricas estáveis e instáveis, que não foram calibradas. Essa deficiência afetará a capacidade do modelo em simular os fluxos turbulentos, descritos a seguir.

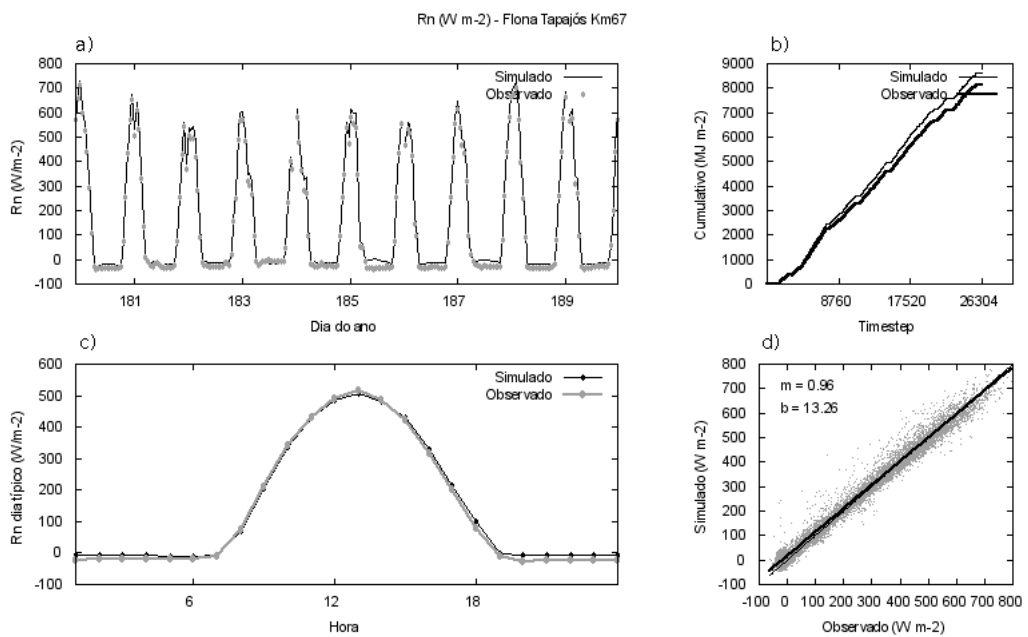
Os resultados do  $H$  após a calibração mono-objetivo são apresentados na Figura 26. Observa-se um viés acentuado entre os dados simulados e observados (Figura 26-b), causado pela superestimativa do fluxo no período noturno, quando o modelo não conseguiu simular adequadamente os valores negativos de  $H$ . Apesar



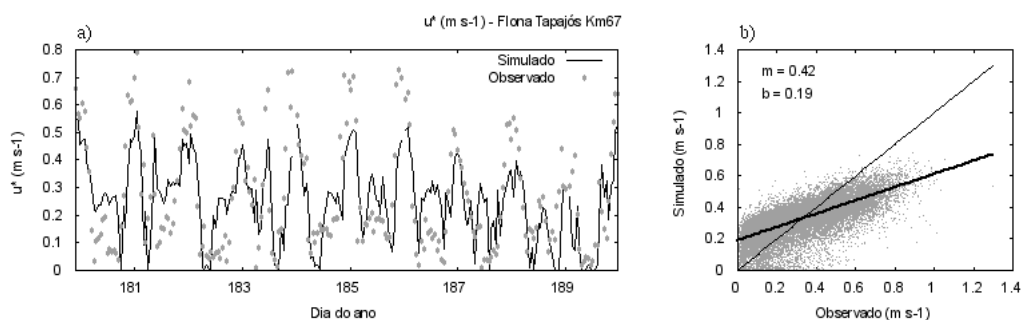
**Figura 22:** Resultados de  $PAR_o$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 23:** Resultados de  $f_{APAR}$  após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.



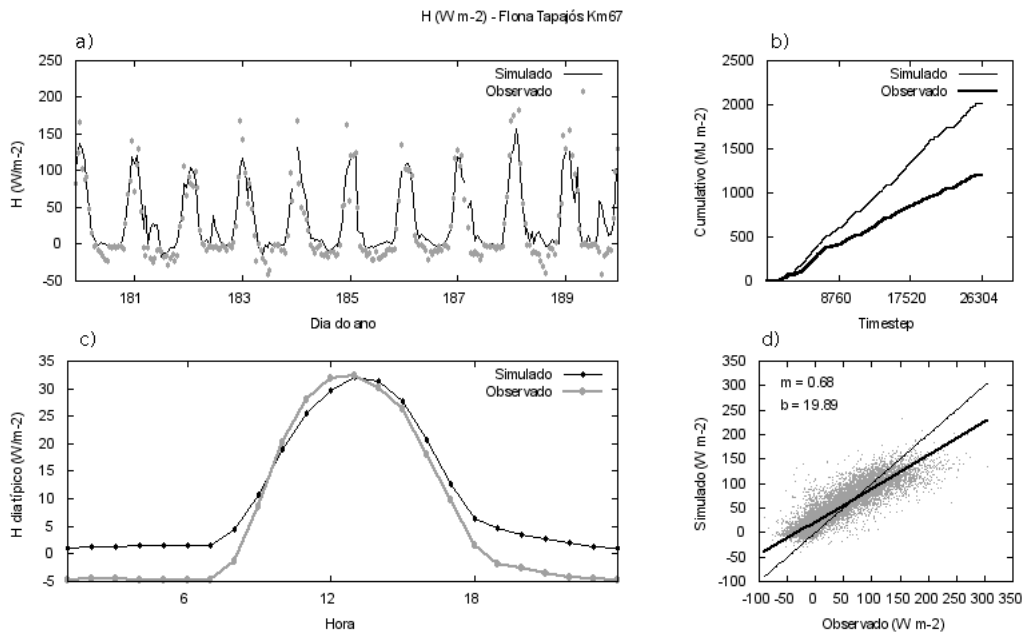
**Figura 24:** Resultados de  $R_n$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 25:** Resultados de  $u^*$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão.



da dificuldade em atingir alguns picos observados (Figura 26-a), os valores mais altos do H foram, em média, bem simulados pelo modelo ao longo do período analisado (Figura 26-c). Esse comportamento pode ter sido causado pelas deficiências na simulação da turbulência.



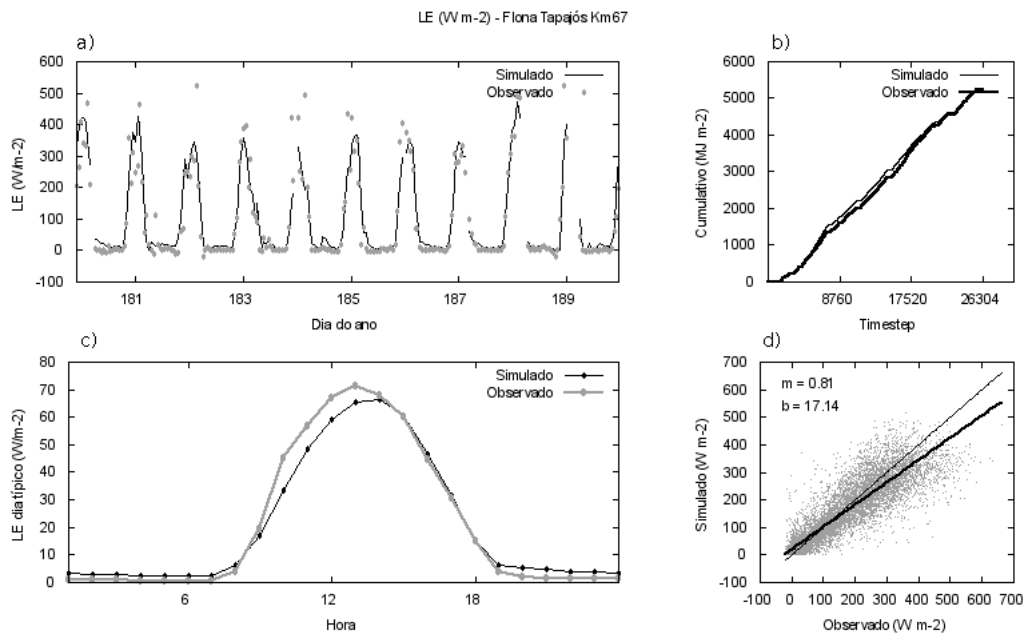
**Figura 26:** Resultados de H após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

A Figura 27 apresenta o LE calibrado pelo método mono-objetivo. O LE simulado ficou bem ajustado ao observado, com um viés muito pequeno entre as curvas da soma cumulativa. Percebe-se uma dificuldade em simular os picos do LE (Figura 27-a), o que também pode ser observado no comportamento da reta da regressão (Figura 27-d).

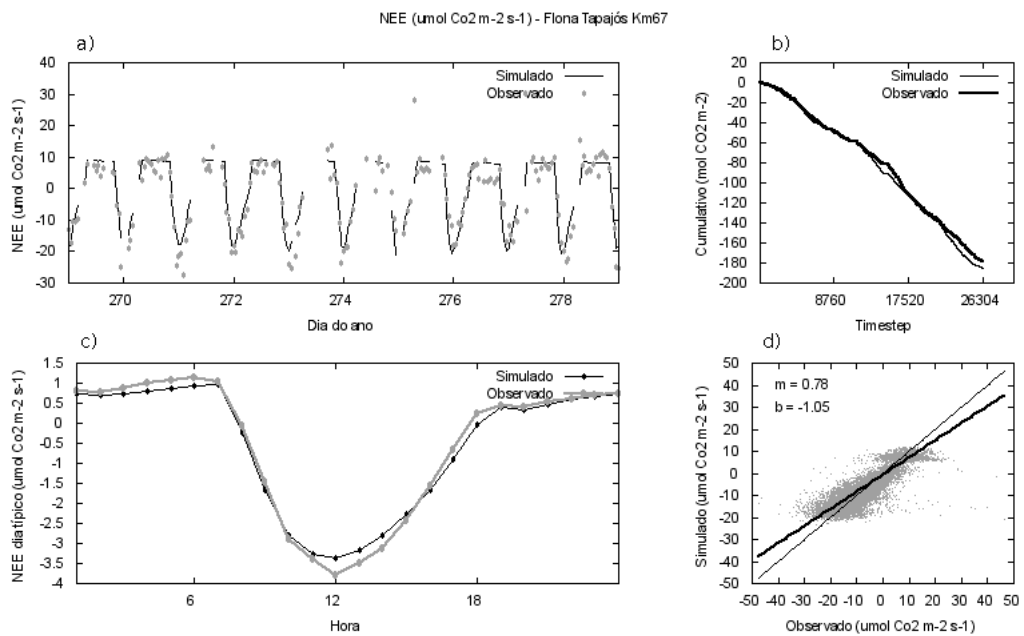
Na Figura 28 é apresentado o NEE após a calibração mono-objetivo. O NEE simulado ficou bem ajustado ao observado, com um ajuste de longo prazo bastante satisfatório (Figura 28-b) e um comportamento horário médio bem ajustado (Figura 28-c). Entretanto, a inclinação da reta de regressão ainda é menor que a unidade, provavelmente uma consequência da deficiência na simulação de  $u^*$ .

O valor simulado da NPP anual foi idêntico ao observado ( $1,055 \text{ kg-Cm}^{-2}\text{ano}^{-1}$ ).

A Figura 29 mostra o LAI calibrado pelo método mono-objetivo. Os coefici-

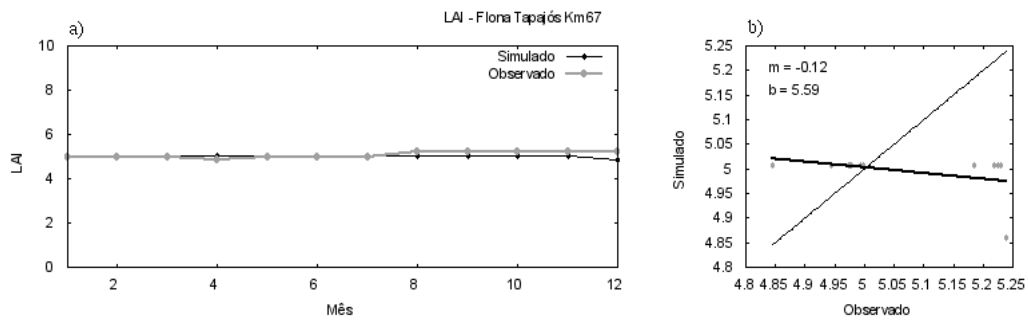


**Figura 27:** Resultados de LE após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 28:** Resultados de NEE após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

entes  $m$  e  $b$  da reta de regressão do gráfico de dispersão são ruins, apesar do bom ajuste observado na Figura 29-a. Isso mostra que a utilização dos coeficientes da reta da regressão linear para avaliação e calibração de modelos deve ser feita com cautela. Esses coeficientes devem ser utilizados como uma análise complementar a outras medidas de ajuste.



**Figura 29:** Resultados de LAI após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.

#### 4.4 Calibração hierárquica multiobjetivo – Flona do Tapajós km 67

A Tabela 9 apresenta os resultados da calibração hierárquica do IBIS, para o sítio da Flona do Tapajós. Os valores dos objetivos otimizados na calibração hierárquica são comparados com os valores obtidos nas calibrações mono-objetivo de cada variável.

O índice de desempenho relativo obtido foi de 0,721, levando em conta as 10 funções objetivo otimizadas, de um total de nove saídas do modelo calibradas. Esse resultado do D foi contaminado pelo valor muito ruim obtido para a razão mono-objetivo/hierárquica da NPP. O valor zero obtido para essa razão não reflete o bom ajuste obtido para a NPP calibrada hierarquicamente. Mesmo obtendo um MAE muito pequeno na calibração hierárquica, o ajuste perfeito obtido na calibração mono-objetivo ( $MAE = 0,000$ ) reduz a razão mono-objetivo/hierárquica da NPP a zero, forçando para baixo a razão média. Esse ajuste perfeito da NPP calibrada pelo modo mono-objetivo se deve ao fato de se possuir um único valor anual de NPP observado, o que facilita o ajuste do modelo

**Tabela 9:** Resultados da calibração hierárquica, para o sítio da Flona do Tapajós km 67. Valores de MAE e  $Bmax$  obtidos no procedimento hierárquico multiobjetivo comparados com os respectivos valores de referência.

Variáveis	Hierárquica		Mono-objetivo		Razão	
	MAE*	$Bmax^{**}$	MAE*	$Bmax^{**}$	MAE	$Bmax$
$PAR_o$ $*$ $(\frac{\mu mol}{m^2 s})$ $**$ $(\frac{mol}{m^2})$	1,115	12,315	1,055	118,812	0,946	0,965
$f_{APAR}$ (adimensional)	0,017	–	0,009	–	0,529	–
$Rn$ $*$ $(\frac{W}{m^2})$ $**$ $(\frac{MJ}{m^2})$	21,539	635,761	18,888	524,840	0,877	0,826
$u^*$ $(\frac{m}{s})$	0,122	–	0,115	–	0,943	–
$NEE$ $*$ $(\frac{\mu mol}{m^2 s})$ $**$ $(\frac{mol}{m^2})$	4,695	22,651	3,701	10,184	0,788	0,450
$H$ $*$ $(\frac{W}{m^2})$ $**$ $(\frac{MJ}{m^2})$	40,567	2.172,025	22,463	816,833	0,554	0,376
$LE$ $*$ $(\frac{W}{m^2})$ $**$ $(\frac{MJ}{m^2})$	36,524	364,327	29,495	239,749	0,808	0,658
$NPP$ $(\frac{kgC}{m^2 \cdot ano})$	0,038	–	0,000	–	0,000	–
$LAI$ $(\frac{m^2}{m^2})$	0,137	–	0,129	–	0,942	–
<b>D (com NPP)</b>					0,721	
<b>D (sem NPP)</b>					0,801	

a esse valor único, para um único ano.

Excluindo-se a NPP do procedimento, obtém-se um valor para o D de 0,801. Esse valor é mais realista, pois não é influenciado pelo falso desempenho ruim da NPP, apontado pela razão mono-objetivo/hierárquica.

As Figuras 30 a 37 mostram os resultados das simulações das variáveis calibradas pelo método hierárquico, e compara as simulações com os dados observados.

A Figura 30 mostra a  $PAR_o$  calibrada pelo método hierárquico. O ajuste entre os dados horários de  $PAR_o$  foi praticamente perfeito, apresentando uma simulação não viesada (Figura 30-b) e com os coeficientes da reta de regressão linear m e b iguais a 1,00 e 0,03 (Figura 30-d).

A Figura 31 mostra a  $f_{APAR}$  calibrada pelo método hierárquico. Os dados observados e simulados mensais estão bem ajustados. Os coeficientes da reta de regressão obtidos foram m=0,77 e b=0,19.

Na Figura 32 é apresentado o  $Rn$  calibrado pelo método hierárquico. Os dados de  $Rn$  simulados ficaram bem ajustados aos dados observados, com uma reta de regressão quase coincidente com a reta 1:1. Observando o gráfico do dia típico

**Tabela 10:** Valores dos parâmetros obtidos no procedimento de calibração hierárquica para o sítio da Flona do Tapajós km 67. Os objetivos foram otimizados em ordem crescente do nível hierárquico, dos processos mais rápidos para os mais lentos.

Nível Hierárquico	Objetivo calibrado	Parâmetro	Valor calibrado
1	$\{MAE_{PAR_o, f_{APAR}}\}$	rhoveg_vis	0,124
		tauveg_vis	0,030
		chifuz	-0,052
2	$\{MAE_{Rn}, Bmax_{Rn}\}$	rhoveg_NIR	0,200
		tauveg_NIR	0,202
		avmuir_coef	45,900
3	$\{MAE_{u^*}\}$	dispu_coef	0,958
		alogl_coef	0,018
		alogu_coef	2,497
4	$\{MAE_{NEE, H, LE}\}$	vmax_pft	$7,050 \cdot 10^{-5}$
		coefmub	7,820
		chs	$1,050 \cdot 10^4$
		beta2	0,996
		funca_coef	2280,000
		funcb_coef	4770,000
		rroot_coef	0,775
		rwood_coef	0,013
		tempvm_coef	$6,800 \cdot 10^3$
		stressf_coef	-6,490
		clitls_coef	0,328
		clitrs_coef	0,210
		clitws_coef	0,404
		csoislon_coef	1,630
		csoislop_coef	1,540
kfactor	1,200		
5	$\{MAE_{LAI, NPP}\}$	rgrowth_coef	0,036
		tauleaf	0,559
		specla	29,500
		aleaf	0,447
		aroot	0,132
		awood	0,423

(Figura 32-c) e o viés do gráfico da soma cumulativa (Figura 32-b), percebe-se uma pequena superestimativa do  $R_n$  simulado no período noturno, como na calibração mono-objetivo.

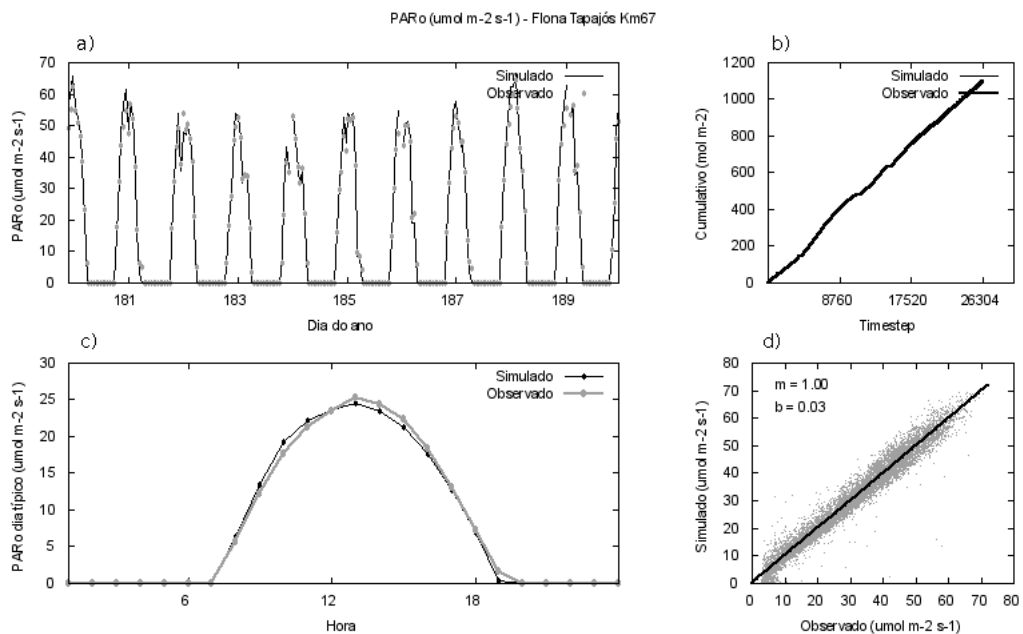
A Figura 33 mostra  $u^*$  calibrado pelo método hierárquico. Apesar do comportamento simétrico entre as duas séries, percebe-se uma subestimativa dos picos máximos e uma superestimativa dos valores mínimos de  $u^*$ , semelhante ao obtido na calibração mono-objetivo.

Os resultados do H calibrado pelo método hierárquico são apresentados na Figura 34. Observa-se um viés bem acentuado entre os dados simulados e observados (Figura 34-b), causado pela superestimativa do fluxo durante todo o

período do dia, além de um deslocamento do pico diário do H (Figura 34-c).

A Figura 35 apresenta o LE calibrado pelo método hierárquico. O viés entre o LE simulado e o observado é bem pequeno, com as curvas da soma cumulativa bastante próximas entre si. Observa-se uma pequena superestimativa do LE noturno e uma subestimativa do LE diurno ao longo da série (Figura 35-c).

Na Figura 36 é apresentado o NEE calibrado pelo método hierárquico. O NEE simulado ficou bem ajustado ao observado, com um ajuste de longo prazo bastante satisfatório (Figura 36-b) e um comportamento diário bem ajustado (Figura 34-a). Na Figura 36-c pode ser observado um atraso de uma hora no valor mínimo do NEE diurno.

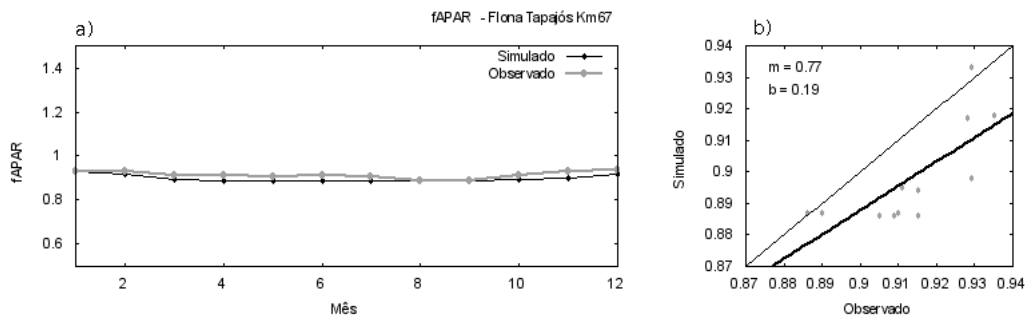


**Figura 30:** Resultados de  $PAR_o$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

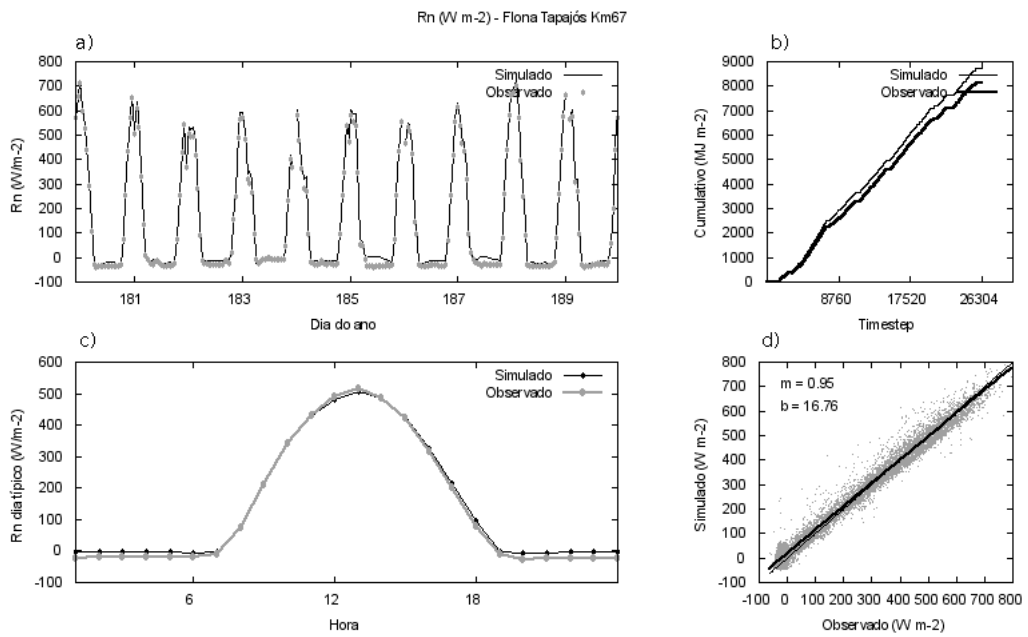
O valor simulado obtido para a NPP anual foi  $1.017 \text{ kg-Cm}^{-2}\text{ano}^{-1}$ , e o valor correspondente, medido em campo, foi  $1.055 \text{ kg-Cm}^{-2}\text{ano}^{-1}$ .

A Figura 37 mostra o LAI calibrado pelo método hierárquico. Assim como na calibração mono-objetivo, os dados simulados e observados de LAI ficaram bem ajustados, apesar do resultado ruim obtido na regressão linear.

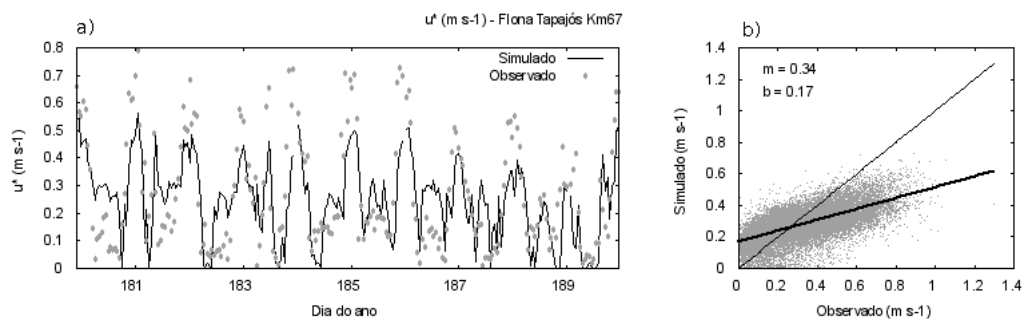
Os valores dos parâmetros obtidos após a calibração hierárquica estão apresentados na Tabela 10, agrupados pelos níveis hierárquicos em que foram calibrados.



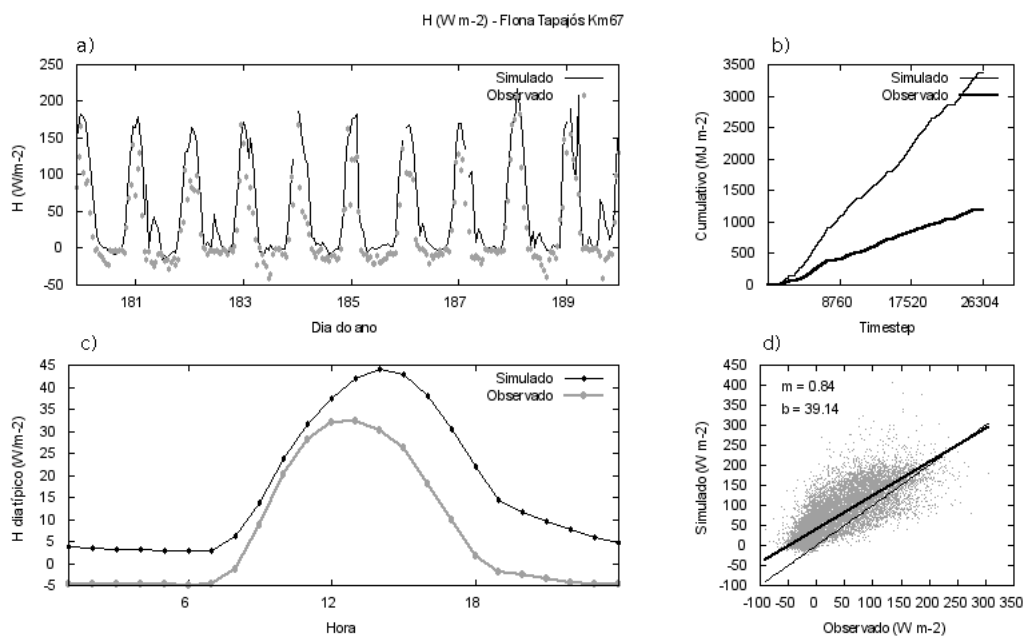
**Figura 31:** Resultados de  $f_{APAR}$  após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.



**Figura 32:** Resultados de  $R_n$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

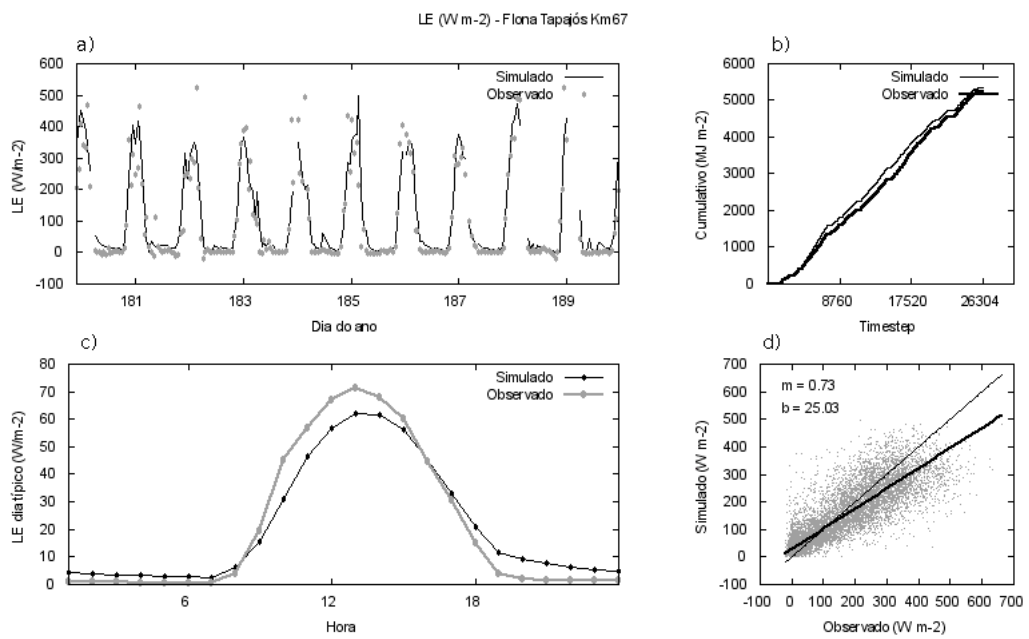


**Figura 33:** Resultados de  $u^*$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (d) gráfico de dispersão.

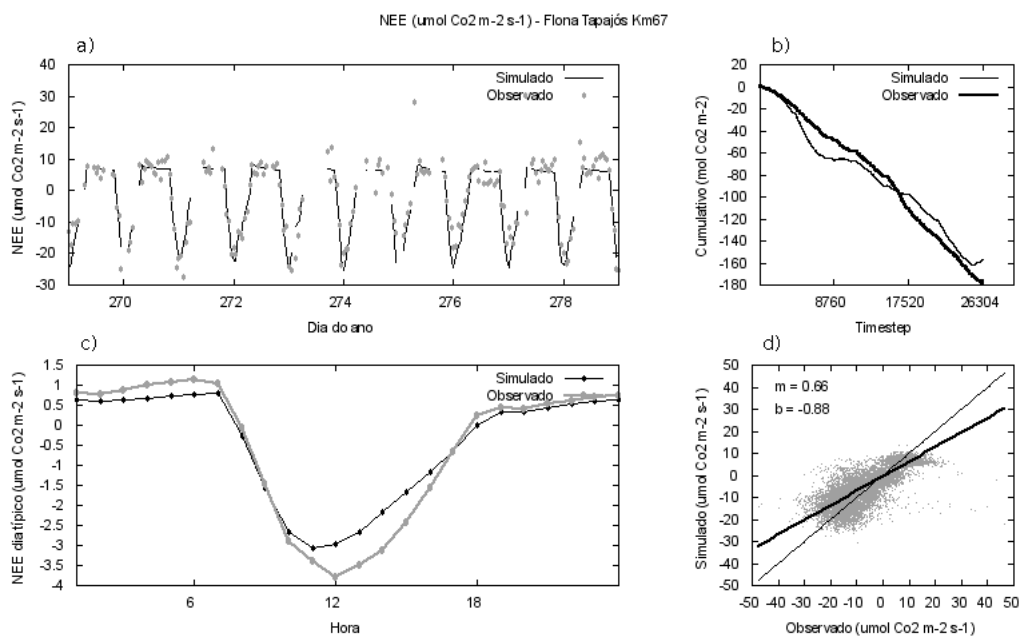


**Figura 34:** Resultados de  $H$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

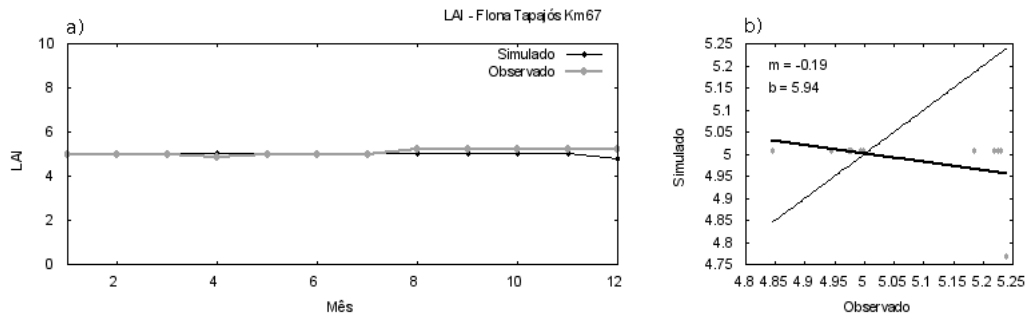




**Figura 35:** Resultados de LE após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 36:** Resultados de NEE após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 37:** Resultados de LAI após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.

#### 4.5 Discussão – Flona do Tapajós km 67

Os resultados individuais obtidos na calibração hierárquica tendem a ser sempre piores que os obtidos nas calibrações mono-objetivos. Isso é esperado, já que nas calibrações mono-objetivo cada variável foi otimizada exclusivamente, desconsiderando-se todas as demais.

A calibração hierárquica busca uma calibração global do modelo, otimizando todas as variáveis simuladas da melhor forma possível. Durante a calibração, pode haver uma perda na qualidade do objetivo de uma determinada variável ao se otimizar outra variável separadamente, em outro nível da hierarquia do processo de calibração.

O índice de desempenho relativo,  $D$ , apresentou um valor de 0,801 (excluindo do seu cálculo a NPP). Apesar da perda de qualidade em relação à calibração mono-objetivo, o procedimento hierárquico fornece uma calibração do modelo mais realista, pois todas as variáveis simuladas pelo modelo são otimizadas para simular corretamente os processos representados.

Para todos os fluxos turbulentos, em todas as calibrações, o coeficiente de inclinação  $m$  da reta de regressão é menor que 1,0. Esse comportamento dos fluxos turbulentos está relacionado com a dificuldade do modelo em simular a turbulência para condições atmosféricas estáveis e instáveis. Como o modelo superestima os valores mais baixos e subestima os valores de  $u^*$ , esse comporta-

mento tende a se repetir nos fluxos turbulentos.

As variáveis simuladas apresentaram, de modo geral, bons resultados após calibradas. Percebe-se um viés maior nos dados de H, provocado devido a uma dificuldade de simular corretamente o fluxo noturno. Mesmo na calibração mono-objetivo não se conseguiu ajustar o H para o período noturno. Esse excesso de H à noite certamente está relacionado ao excesso de Rn nesse mesmo período, que é propagado para a simulação do H. O excesso de Rn durante a noite, por sua vez, provavelmente está associado à metodologia de estimativa de  $L_{in}$  (Equação 40), onde assume-se que a temperatura efetiva de emissão do dossel é igual a temperatura do ar. Essa simplificação também vai afetar o cálculo do H, que é proporcional à diferença entre as temperaturas da folha e do ar.

Uma outra possível causa para a deficiência na simulação do H seriam incertezas na medição dos dados de campo. Um problema fundamental com o sistema de covariância de vórtices turbulentos é que o balanço de energia geralmente não fecha, com um erro de 20% ou mais (*Wilson et al., 2002; Maayar et al., 2008*). Isso viola o princípio da conservação de energia, com base no qual os modelos são construídos. O erro no fechamento dos fluxos pode ter influenciado, pelo menos em parte, a maior dificuldade de calibração do H.

#### 4.6 Calibração mono-objetivo – Sinop

Os resultados das calibrações mono-objetivo do MAE e  $B_{max}$  de cada variável, para o sítio de Sinop, são apresentados na Tabela 11. Esses resultados são utilizados como referência para avaliar a calibração hierárquica multiobjetivo.

As Figuras 38 a 45 mostram os resultados de simulações do modelo após a calibração mono-objetivo de cada variável, e compara as simulações com os dados observados.

A Figura 38 mostra a  $PAR_o$  simulada após a calibração mono-objetivo. A séries de dados simulados e observados estão bem ajustadas, conforme a Figura 38-a. Entretanto, no gráfico do dia típico (Figura 38-c), observa-se uma aparente

subestimativa da  $PAR_o$  pelo modelo. O que ocorre na realidade é uma contaminação da análise por dados espúrios, que fazem com que a média horária dos dados observados, no cálculo do dia típico, fique mais alta. A influência de dados observados ruins fica evidente no gráfico de dispersão (Figura 38-d), onde a maioria dos dados estão agrupados ao longo da reta 1:1, e alguns pontos do gráfico, cujos valores observados são exageradamente altos, deslocam a reta de regressão. No gráfico da soma cumulativa (Figura 38-b), percebe-se a grande quantidade de dados faltantes na série, correspondentes aos trechos onde a curva fica totalmente horizontal.

A Figura 39 mostra a  $f_{APAR}$  calibrada pelo método mono-objetivo. Alguns pontos da série mensal de dados estão bem ajustados. Porém, os valores do início da série não se ajustaram adequadamente. A pequena variação dos dados e a pouca quantidade de pontos prejudica a análise da regressão, que pode ser muito influenciada por valores espúrios.

Na Figura 40 é apresentado o  $R_n$  simulado após a calibração mono-objetivo. O ajuste dos dados simulados e observados é praticamente perfeito (Figura 40-a), com as curvas do gráfico cumulativo coincidentes (Figura 40-b), os dados horários médios muito bem ajustados (Figura 40-c) e uma reta de regressão com coeficientes  $m$  igual a 0,98 e  $b$  igual a 1,06 (Figura 40-d). Entretanto, pode ser observado no gráfico de dispersão, na Figura 40-d, uma quantidade de dados espúrios, onde o  $R_n$  é tanto superestimado quanto subestimado pelo modelo. Esses dados, entretanto, não aparecem nas demais análises, provavelmente porque há uma compensação entre eles.

A Figura 41 mostra  $u^*$  calibrado através do método mono-objetivo. Nota-se que há uma subestimativa dos picos máximos e uma superestimativa dos valores mínimos, apesar da simetria entre as duas séries.

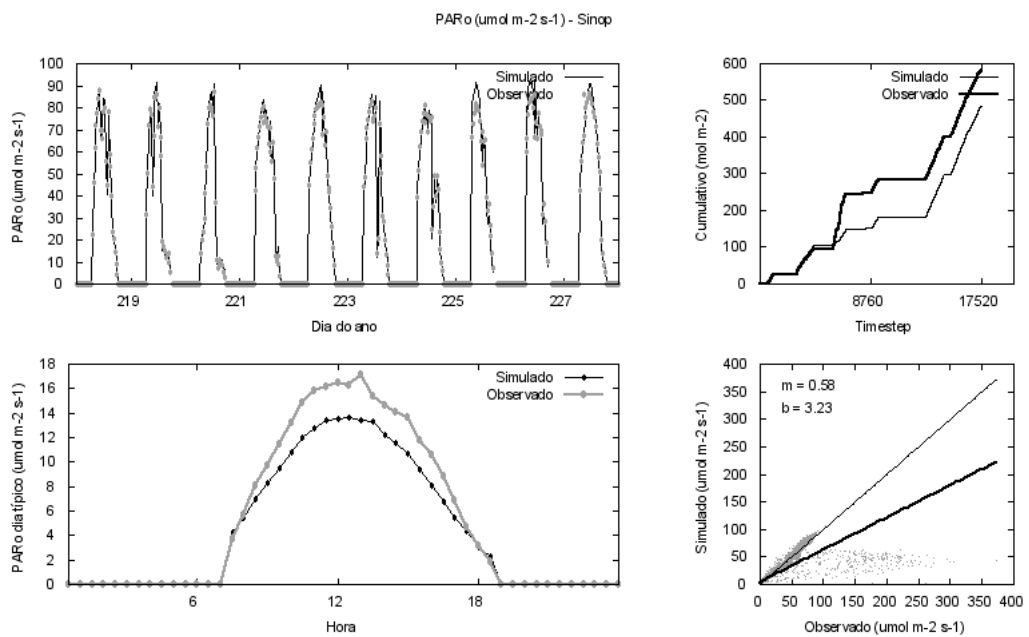
Os resultados do  $H$  simulado, após a calibração mono-objetivo, são mostrados na Figura 42. Os valores mais altos de  $H$  não são simulados corretamente, enquanto o  $H$  no período noturno apresenta um bom ajuste (Figura 42-a e c). Essa

subestimativa do H no período diurno evidencia-se no viés do gráfico cumulativo (Figura 42-b) e no baixo valor do coeficiente m da reta de regressão (Figura 42-d).

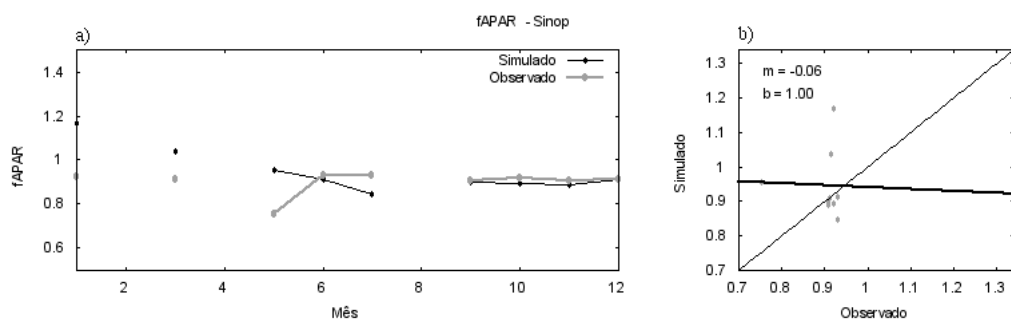
A Figura 43 apresenta o LE calibrado pelo método mono-objetivo. O LE simulado ficou bem ajustado ao observado, com um viés relativamente pequeno entre as curvas da soma cumulativa (Figura 43-b). Pelos gráficos de dispersão e cumulativo (Figura 43-d e b), percebe-se a presença de uma grande quantidade de dados espúrios, com o LE sendo tanto subestimado quanto superestimado em alguns pontos. No gráfico do dia típico (Figura 43-c), que mostrou um bom ajuste entre os dados, nota-se que há uma compensação entre as subestimativas e superestimativas ao longo da série.

Na Figura 44 é apresentado o NEE simulado após a calibração mono-objetivo. O comportamento ao longo do dia foi bem simulado, em média (Figura 44-a). O NEE simulado não representou as variações no fluxo noturno, apresentadas nos dados observados, além de ter superestimado o valor mínimo médio (Figura 44-c).

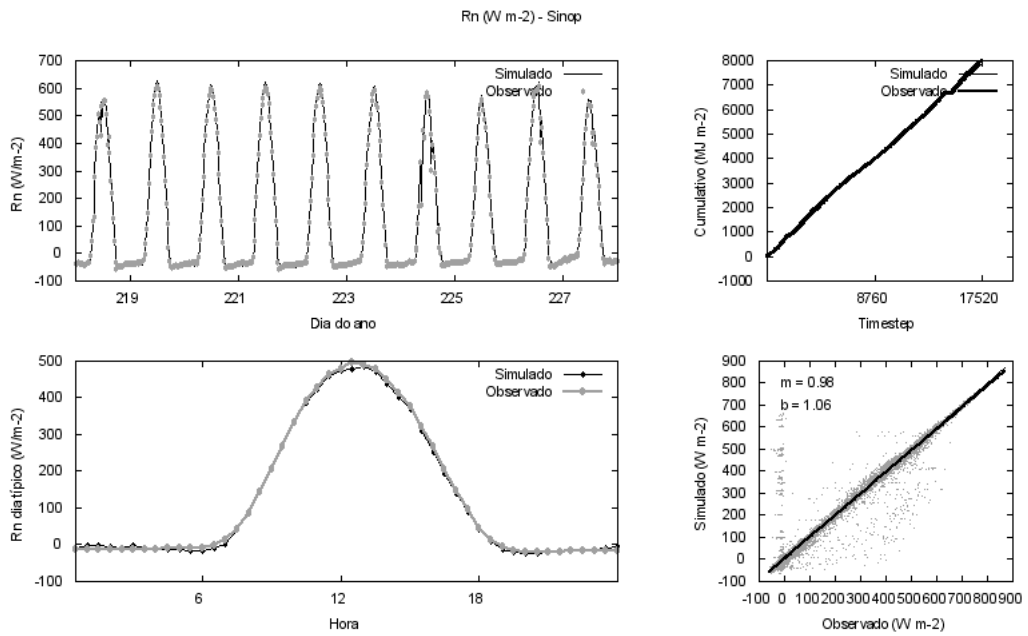
A Figura 45 mostra o LAI calibrado pelo método mono-objetivo. Mesmo após calibrado, o modelo não conseguiu simular o LAI adequadamente. Observa-se que o modelo produziu um valor constante de LAI durante praticamente todo o período simulado.



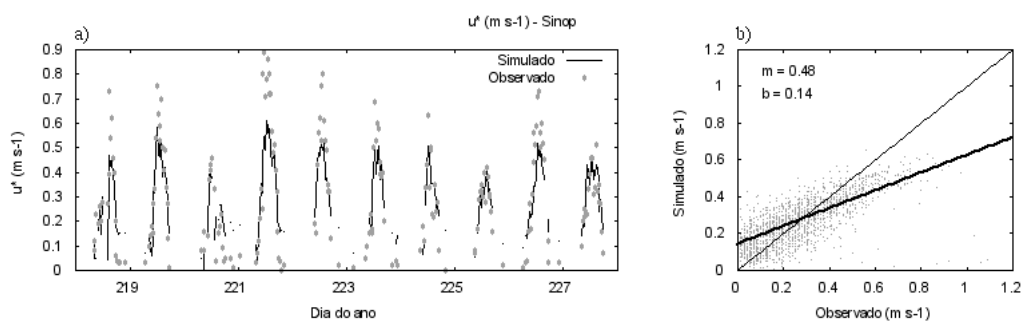
**Figura 38:** Resultados de  $PAR_o$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



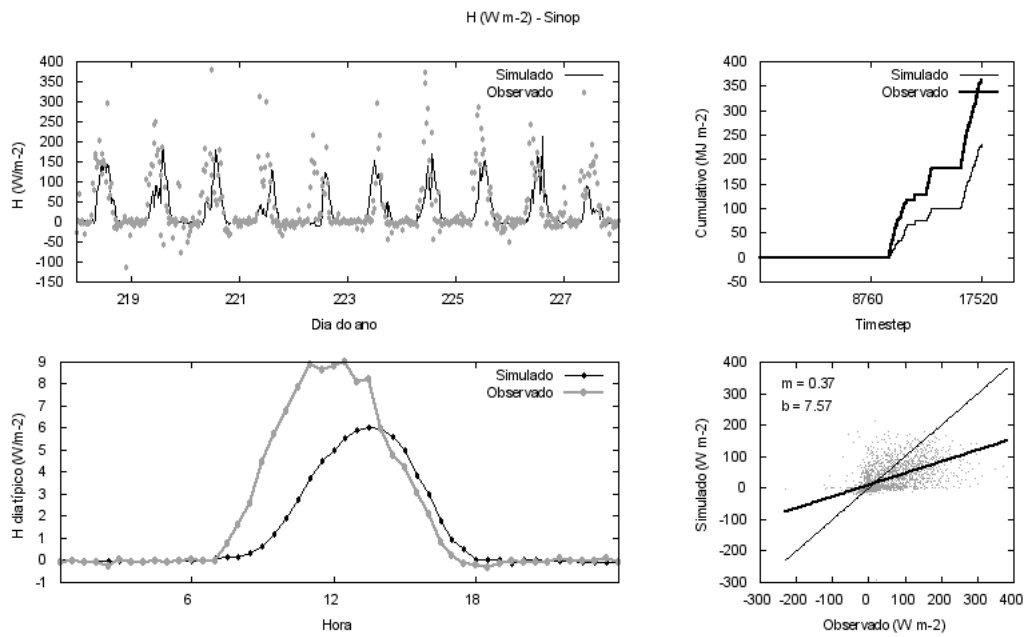
**Figura 39:** Resultados de  $f_{APAR}$  após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.



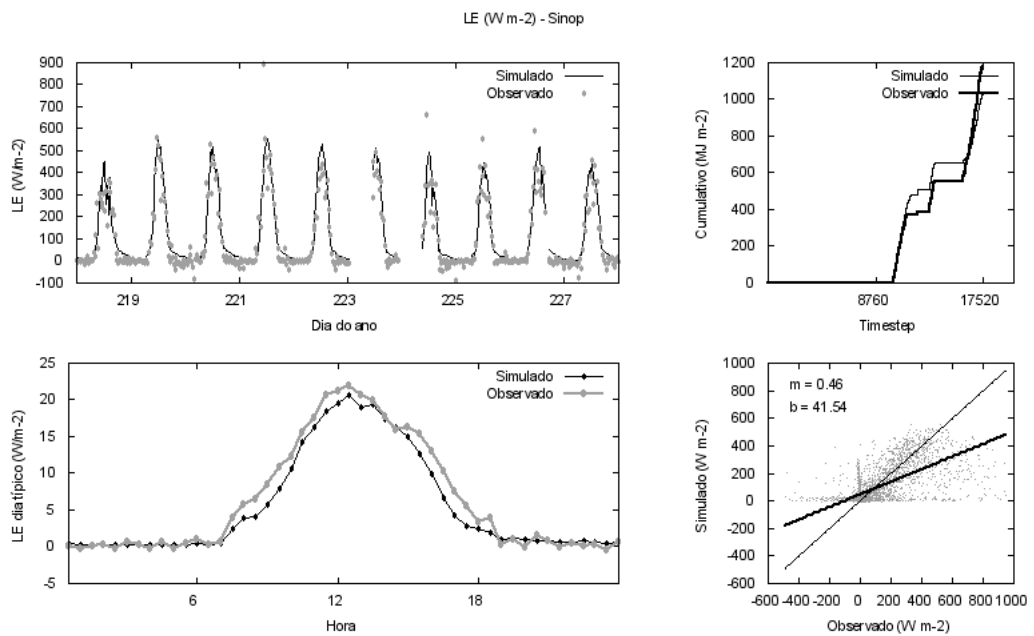
**Figura 40:** Resultados de  $R_n$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 41:** Resultados de  $u^*$  após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão.

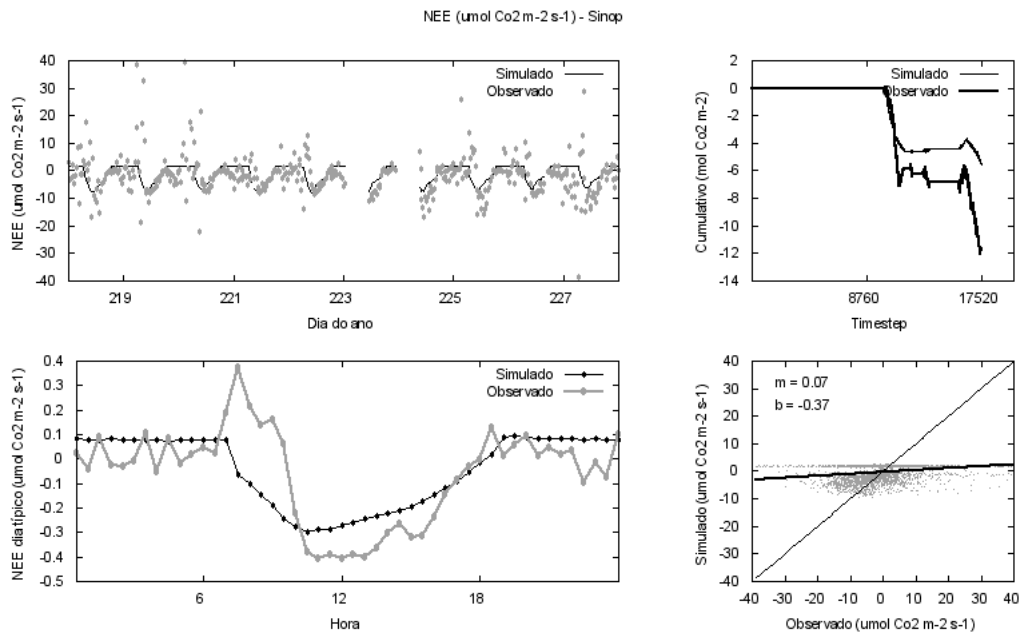


**Figura 42:** Resultados de H após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.

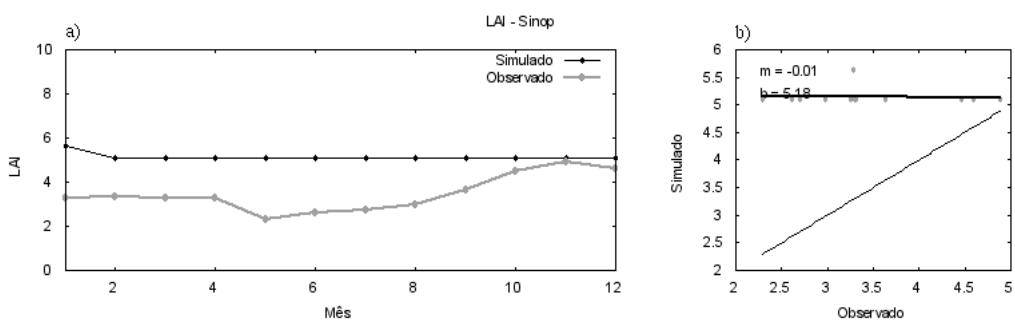


**Figura 43:** Resultados de LE após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.





**Figura 44:** Resultados de NEE após calibração mono-objetivo. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 45:** Resultados de LAI após calibração mono-objetivo. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.

## 4.7 Calibração hierárquica multiobjetivo – Sinop

A Tabela 11 apresenta os resultados da calibração hierárquica do IBIS, para o sítio de Sinop. Os valores dos objetivos otimizados na calibração hierárquica são comparados com os valores obtidos nas calibrações mono-objetivo de cada variável.

**Tabela 11:** Resultados da calibração hierárquica para o sítio de Sinop. Valores de MAE e  $Bmax$  obtidos no procedimento hierárquico multiobjetivo comparados com os respectivos valores de referência.

Variáveis	Hierárquica		Mono-objetivo		Razão		
	MAE*	$Bmax^{**}$	MAE*	$Bmax^{**}$	MAE	$Bmax$	
$PAR_o * (\frac{\mu mol}{m^2 s}) ** (\frac{mol}{m^2})$	3,848	149,876	3,170	100,967	0,824	0,674	
$f_{APAR}$ (adimensional)	0,108	–	0,081	–	0,750	–	
$Rn * (\frac{W}{m^2}) ** (\frac{MJ}{m^2})$	17,771	131,300	15,768	111,753	0,887	0,851	
$u^*$ ( $\frac{m}{s}$ )	0,108	–	0,097	–	0,898	–	
$NEE * (\frac{\mu mol}{m^2 s}) ** (\frac{mol}{m^2})$	6,192	16,503	5,774	6,834	0,932	0,414	
$H * (\frac{W}{m^2}) ** (\frac{MJ}{m^2})$	34,413	44,191	26,287	132,097	0,764	2,989	
$LE * (\frac{W}{m^2}) ** (\frac{MJ}{m^2})$	81,460	296,414	73,046	147,783	0,897	0,499	
$LAI$ ( $\frac{m^2}{m^2}$ )	2,561	–	1,705	–	0,666	–	
<b>D</b>						0,830	

O índice de desempenho relativo obtido foi de 0,830, levando em conta as nove funções objetivo otimizadas, de um total de oito saídas do modelo calibradas.

As figuras 46 a 53 mostram os resultados de simulações do modelo após a calibração hierárquica, e compara as simulações com os dados observados.

A Figura 46 mostra a  $PAR_o$  calibrada pelo método hierárquico. A séries de dados simulados e observados estão bem ajustadas, com uma pequena subestimativa dos picos do  $PAR_o$  (Figura 46-a). Essa subestimativa é mais evidente no gráfico do dia típico (Figura 46-c). Nota-se, pelo gráfico de dispersão, a existência de dados observados espúrios (Figura 46-d), onde alguns pontos com valores observados são muito elevados, deslocam a reta de regressão.

A Figura 47 mostra a  $f_{APAR}$  calibrada pelo método hierárquico. Alguns pontos da série mensal de dados estão bem ajustados. Porém, os valores do início da série

**Tabela 12:** Valores dos parâmetros obtidos no procedimento de calibração hierárquica, para o sítio de Sinop. Os objetivos foram otimizados em ordem crescente do nível hierárquico, dos processos mais rápidos para os mais lentos.

Nível Hierárquico	Objetivo calibrado	Parâmetro	Valor calibrado
1	$\{MAE_{PAR_o, f_{APAR}}\}$	rhoveg_vis	0,149
		tauveg_vis	0,068
		chifuz	-0,350
2	$\{MAE_{Rn}, Bmax_{Rn}\}$	rhoveg_NIR	0,306
		tauveg_NIR	0,141
		avmuir_coef	146,782
3	$\{MAE_{u*}\}$	dispu_coef	0,938
		alogL_coef	1,009
		alogu_coef	1,117
4	$\{MAE_{NEE, H, LE}\}$	vmax_pft	$4,584 \cdot 10^{-5}$
		coefmub	10,332
		chs	$1,010 \cdot 10^4$
		beta2	0,999
		funca_coef	4339,780
		funcb_coef	3579,086
		rroot_coef	0,879
		rwood_coef	0,066
		tempvm_coef	$5,488 \cdot 10^3$
		stressf_coef	-3,861
		clitls_coef	0,359
		clitrs_coef	0,041
		clitws_coef	0,583
		csoislon_coef	0,724
csoislop_coef	2,023		
kfactor	1,302		
5	$\{MAE_{LAI}, Bmax_{LAI}\}$	rgrowth_coef	0,604
		tauleaf	0,529
		specla	33,526
		aleaf	0,298
		aroot	0,344
		awood	0,357

não se ajustaram adequadamente. A regressão (Figura 47-d) foi prejudicada pela pouca quantidade de dados, que pode ser muito influenciada por valores espúrios.

Na Figura 48 é apresentado o  $R_n$  simulado, após sua calibração pelo método hierárquico. O ajuste dos dados simulados e observados é praticamente perfeito (Figura 48-a). As curvas do gráfico cumulativo são coincidentes (Figura 48-b) e os dados do dia típico estão muito bem ajustados (Figura 48-c). Os coeficientes  $m$  e  $b$  da reta de regressão são iguais a 0,98 e 2,86, respectivamente (Figura 48-d). Pode ser observado, no gráfico de dispersão, uma grande quantidade de dados espúrios, onde o  $R_n$  é tanto superestimado quanto subestimado pelo modelo. Esses dados, entretanto, não aparecem nos gráficos do dia típico e cumulativo,

provavelmente porque há uma compensação entre esses erros.

A Figura 49 mostra  $u^*$  calibrado pelo método hierárquico. Observa-se que ocorre uma subestimativa dos picos máximos e uma superestimativa dos valores mínimos, apesar do comportamento simétrico das duas séries.

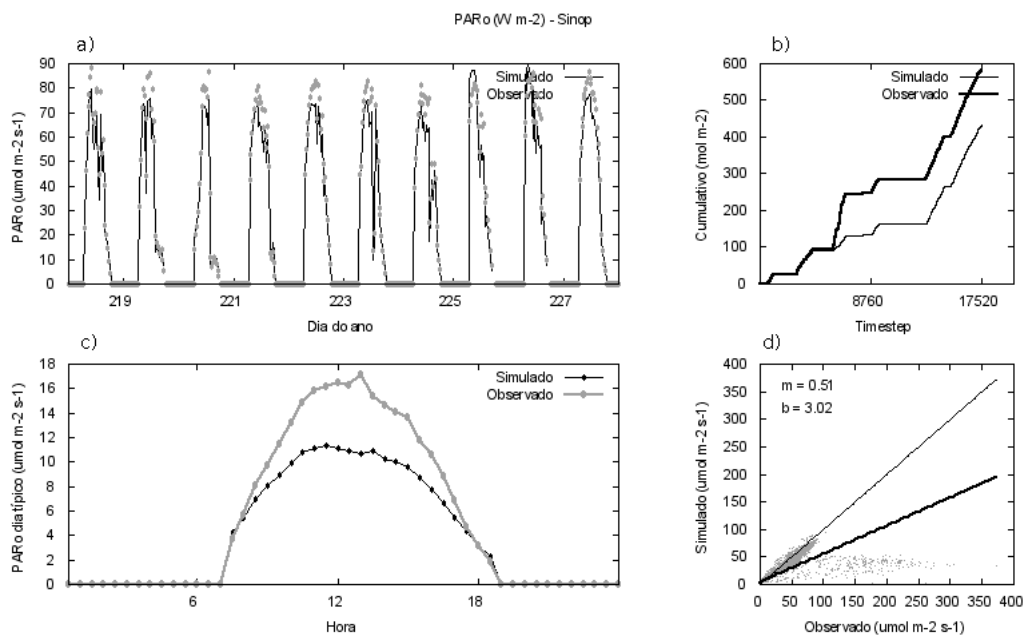
A Figura 50 mostra a H calibrado pelo método hierárquico. Os valores mais altos de H não são simulados corretamente e há um deslocamento dos dados simulados em relação aos observados (Figura 50-a e c).

O LE simulado (Figura 51), após calibrado pelo método hierárquico, ficou bem ajustado ao LE medido. Pelos gráficos de dispersão e cumulativo (Figura 51-d e b), percebe-se a presença de uma grande quantidade de dados espúrios, com o LE sendo tanto subestimado quanto superestimado em alguns pontos. O gráfico do dia típico (Figura 51-c) mostrou um bom ajuste entre os dados, o que pode indicar a ocorrência de uma compensação entre as subestimativas e superestimativas ao longo da série.

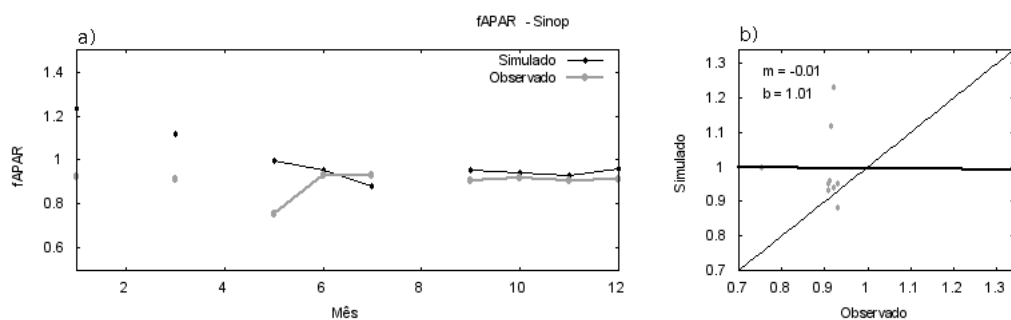
Na Figura 52 é apresentado o NEE calibrado pelo método hierárquico. Em média, o comportamento do NEE ao longo do dia foi bem simulado (Figura 52-a). O NEE simulado não representou as variações no fluxo noturno, apresentadas nos dados observados, além de ter superestimado o valores observados nesse período (Figura 52-c).

A Figura 53 mostra o LAI calibrado pelo método hierárquico. Após calibrado, o modelo não conseguiu simular o LAI adequadamente. O modelo produziu um valor constante de LAI durante todo o período simulado.

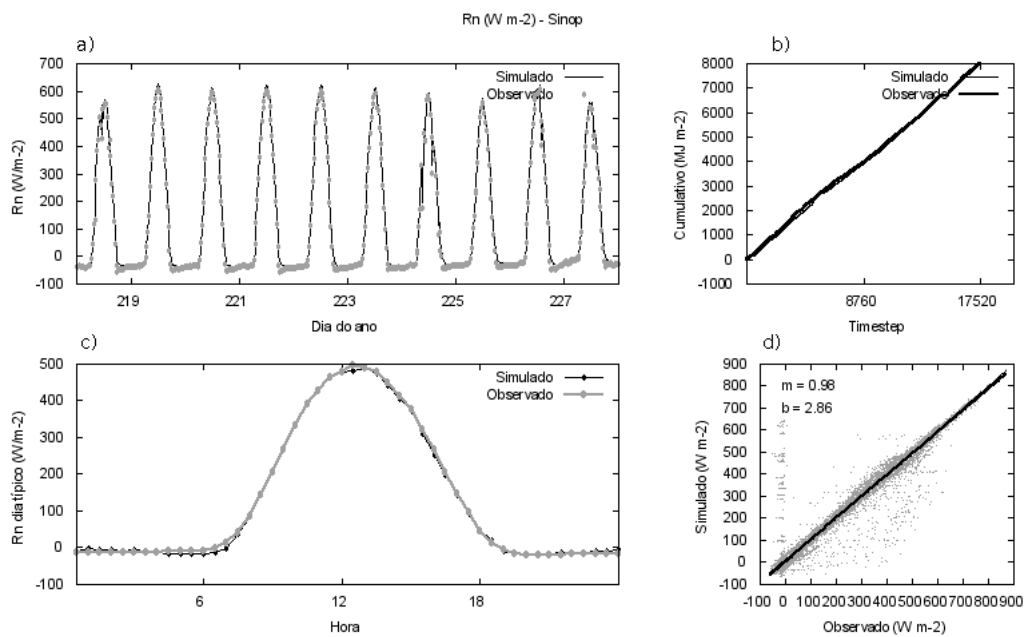
Os valores dos parâmetros obtidos após a calibração hierárquica estão apresentados na Tabela 12, agrupados pelos níveis hierárquicos em que foram calibrados.



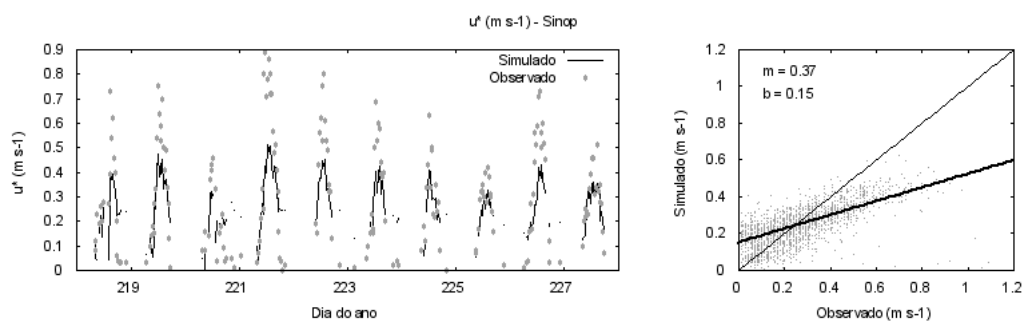
**Figura 46:** Resultados de  $PAR_o$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



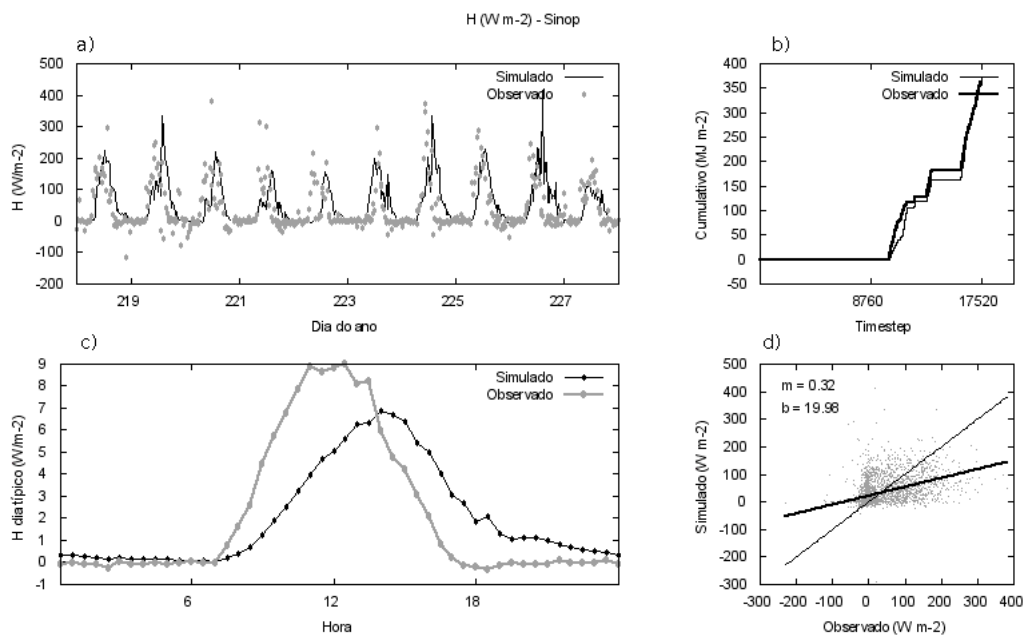
**Figura 47:** Resultados de  $f_{APAR}$  após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.



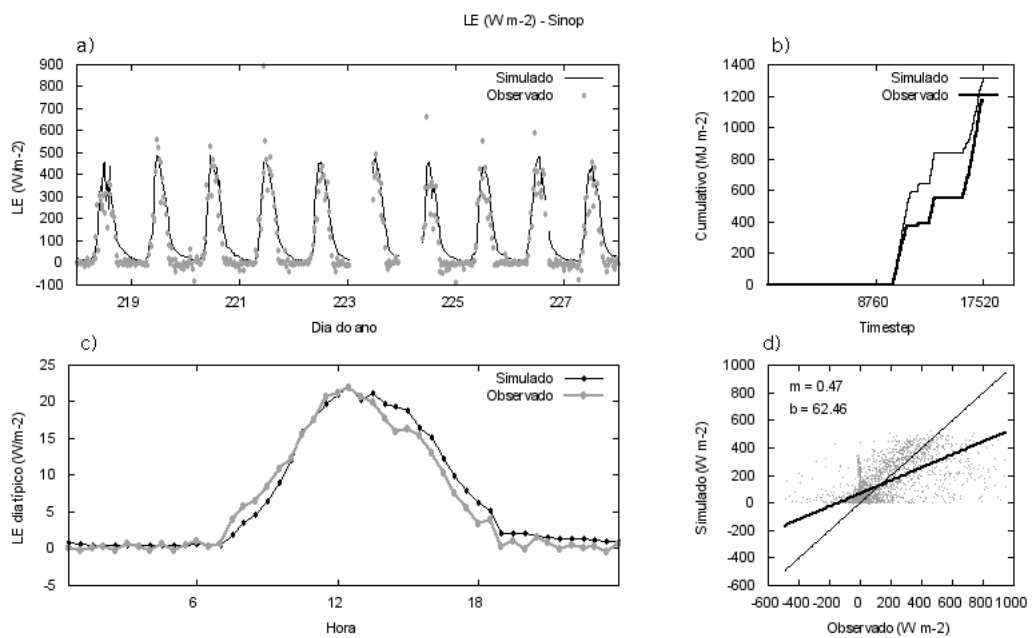
**Figura 48:** Resultados de  $R_n$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



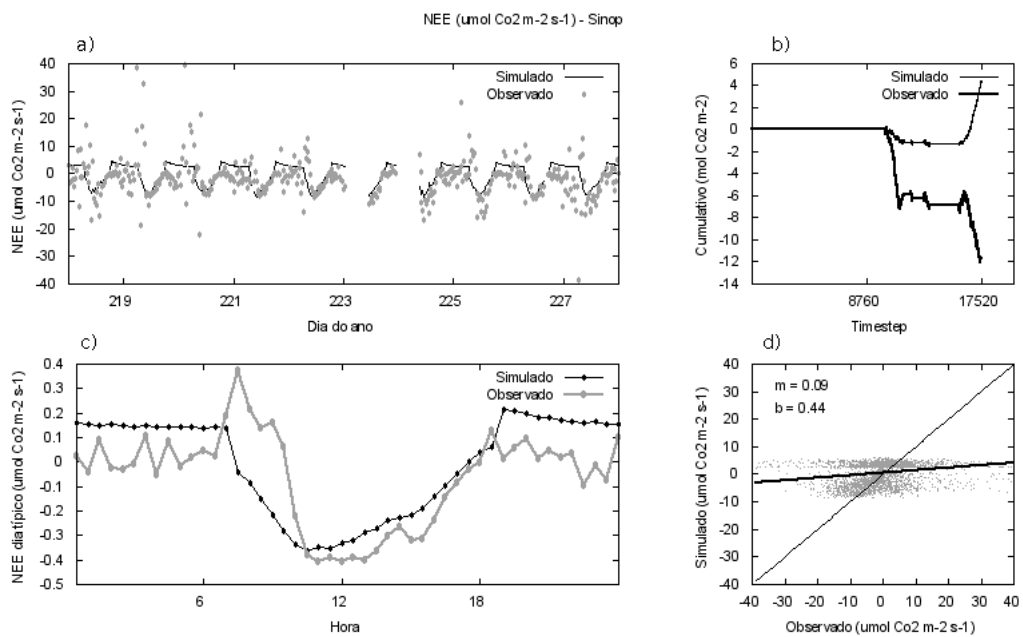
**Figura 49:** Resultados de  $u^*$  após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados e (b) gráfico de dispersão.



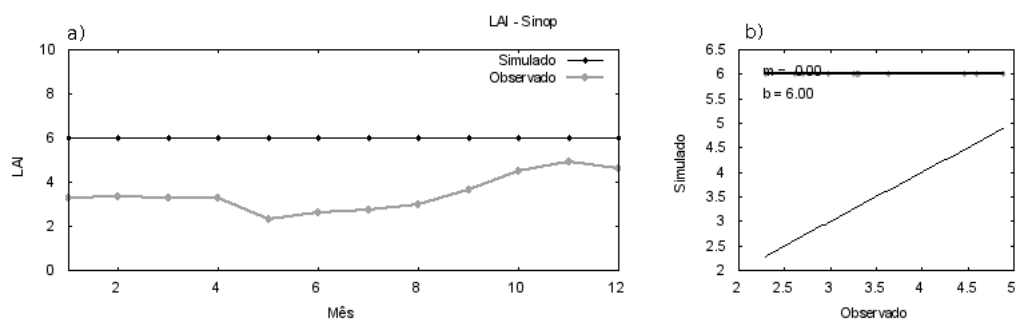
**Figura 50:** Resultados de H após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 51:** Resultados de LE após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 52:** Resultados de NEE após calibração hierárquica. Os gráficos representam (a) um trecho de 10 dias das séries de dados, (b) gráfico da soma cumulativa, (c) gráfico de dia típico e (d) gráfico de dispersão.



**Figura 53:** Resultados de LAI após calibração hierárquica. Os gráficos representam (a) um ano de dados mensais e (b) gráfico de dispersão.



## 4.8 Discussão – Sinop

Assim como para o sítio de Tapajós, os resultados individuais obtidos na calibração hierárquica para Sinop tendem a ser sempre piores que os obtidos nas calibrações mono-objetivos.

As medidas de ajuste obtidas com o modelo calibrado para Sinop foram bem piores que as obtidas para a Flona do Tapajós. Isso certamente ocorreu devido à indisponibilidade de uma série de dados mais completa para o sítio de Sinop. A pouca quantidade de dados para o ajuste do modelo provavelmente prejudicou o processo de calibração. Além disso, o preenchimento de falhas nos dados que alimentam o modelo também pode inserir erros nos dados simulados, prejudicando a comparação entre esses dados e os observados.

Mesmo com um resultado final pior em relação ao obtido no sítio de Tapajós, a calibração hierárquica para Sinop forneceu resultados satisfatórios em relação aos dados de referência (calibração mono-objetivo), com valores bastante satisfatórios das razões mono-objetivo/hierárquica. O índice de desempenho relativo,  $D$ , apresentou um valor de 0,830. Apesar da perda de qualidade em relação à calibração mono-objetivo, o procedimento hierárquico fornece uma boa calibração do modelo, já que ele permite que todas as variáveis simuladas pelo modelo sejam otimizadas.

## 5 CONCLUSÕES E RECOMENDAÇÕES

Neste trabalho foram desenvolvidos uma metodologia de análise de sensibilidade para calibração de modelos, um sistema para calibração automática multiobjetivo do modelo IBIS e uma metodologia de calibração geral de LSMs baseada na hierarquia dos sistemas. A análise de sensibilidade foi baseada no método de Morris, com modificações no cálculo do índice de sensibilidade para adequar o método a calibração de LSMs e selecionar os parâmetros mais importantes para calibração. O sistema de calibração multiobjetivo foi baseado no algoritmo genético multiobjetivo NSGA-II, e permite calibrar o modelo em função de várias estatísticas e de várias saídas do modelo, simultaneamente. A metodologia de calibração hierárquica se baseia na organização hierárquica temporal dos sistemas naturais e do próprio modelo IBIS, para otimizar os processos que ocorrem em várias escalas de tempo, fazendo uso do sistema de calibração multiobjetivo desenvolvido. A aplicação do algoritmo de otimização multiobjetivo de acordo com a metodologia hierárquica proposta, permite uma calibração geral do modelo mais confiável, ajustando todas as variáveis simuladas pelo IBIS. Essa calibração possibilita estudos de modelagem mais confiáveis envolvendo os processos simulados pelo modelo IBIS.

A análise de sensibilidade desenvolvida com base no método proposto por Morris é eficiente ao selecionar os parâmetros mais importantes do modelo a serem considerados na calibração. Neste trabalho foi proposta uma modificação no método de Morris, com o cálculo de um índice de sensibilidade adequado para selecionar os parâmetros mais importantes para calibração. Dos 43 parâmetros considerados na análise, apenas 23 apresentaram influência significativa em alguma das saídas do modelo analisadas. No melhor caso, houve uma redução do espaço de busca de 40 parâmetros, obtendo-se um resultado satisfatório de calibração do  $PAR_o$  com apenas os três parâmetros mais importantes detectados na AS. No pior caso, conseguiu-se uma redução de 27 parâmetros na calibração, calibrando-se o Rn com 16 parâmetros dos 43 inicialmente considerados. Essa diminuição do espaço de busca, através da AS, implica em calibrações mais simples e eficientes.

O sistema automático de calibração do IBIS, chamado Optis, foi desenvolvido para permitir uma otimização mais confiável e eficiente do modelo, considerando múltiplos objetivos. Para avaliar o desempenho das calibrações multiobjetivo foi proposto um índice de desempenho relativo (D), que avalia as calibrações multiobjetivo em relação às calibrações mono-objetivo. Esse índice fornece uma avaliação independente de outros erros, embutidos nos resultados, que não sejam devido ao próprio método de calibração. No teste de calibração multiobjetivo com múltiplas estatísticas, o Optis obteve um D de 0,925. No teste de calibração multiobjetivo com múltiplas saídas do modelo, o Optis obteve um D de 0,925. O Optis, portanto, é eficaz em solucionar problemas de calibração multiobjetivo.

O Optis também se mostrou eficiente no uso de recursos computacionais. O algoritmo foi paralelizado com 100% de eficiência, o que permite desenvolver grandes experimentos de calibração em tempo viável, com o uso de computadores com vários processadores.

Finalmente, foi proposto um procedimento hierárquico de calibração, utilizando o próprio programa Optis. O procedimento hierárquico foi testado em dois

sítios da floresta Amazônica, Flona do Tapajós e Sinop. Para a Flona do Tapajós, o método hierárquico permitiu otimizar nove funções objetivo, calibrando oito saídas do modelo, com um D de 0,801. Para Sinop, obteve-se um D de 0,830, calibrando oito saídas do modelo com nove funções objetivo. Apesar da perda de qualidade em relação à calibração mono-objetivo, o procedimento hierárquico fornece uma calibração do modelo mais realista, pois todas as saídas pelo modelo são otimizadas.

Com base no presente trabalho, é possível identificar diversos trabalhos para realização futura:

1. Avaliação de estatísticas para calibração de modelos, especialmente combinadas em uma calibração multiobjetivo. A calibração automatizada do IBIS permite que sejam feitas várias calibrações do modelo, com várias estatísticas combinadas em múltiplos objetivos. Assim, pode ser possível identificar o conjunto de estatísticas ou medidas de ajuste que são mais adequados para a calibração do modelo.
2. Calibração multiobjetivo considerando múltiplos sítios, utilizando o programa Optis. Com a calibração multisítio pode ser possível descobrir valores dos parâmetros que sejam representativos para toda uma determinada região.
3. Calibração multiobjetivo considerando dados 2D de sensoriamento remoto. Assim como na calibração para múltiplos sítios, pode ser possível descobrir valores dos parâmetros que sejam representativos para toda uma determinada região através de dados de sensoriamento remoto.
4. Correções no fechamento do balanço de energia dos dados observados. *Mayar et al.* (2008) argumentam que os dados de fluxo de energia devem ser corrigidos antes de serem usados para avaliação de modelos. Entretanto, não se sabe ao certo o quanto do erro no fechamento dos fluxos reside no H ou no LE. Com um sistema de calibração automático e preciso podem ser

feitos estudos no sentido de se descobrir as correções mais adequadas para cada fluxo.

5. Modificação do algoritmo base do Optis para permitir a otimização eficiente de uma quantidade grande de objetivos (maior que 3). Isso necessitaria uma mudança em todo o código do programa, já que o Optis é totalmente baseado no NSGA-II. Essa modificação permitiria calibrações considerando várias variáveis do modelo e várias estatísticas de ajuste simultaneamente, além de possibilitar a calibração multisítio, citada anteriormente, para um grande número sítios.
6. Integração da análise de sensibilidade e do algoritmo de calibração em um único framework. Com todos os passos do processo de calibração contidos em um único software, seria possível uma maior automatização de todo o processo, inclusive automatizando a geração de gráficos e análises dos resultados.

## Referências

- Bekele, E. G., e J. W. Nicklow (2007), Multi-objective automatic calibration of swat using nsga-ii, *Journal of Hydrology*, 341, 165–176.
- Collatz, G. J., J. T. Ball, C. Grivet, e J. A. Berry (1991), Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: a model that includes a laminar boundary layer, *Agricultural and Forest Meteorology*, 54, 107–136.
- Costa, M. H., M. C. Biajoli, L. Sanches, L. R. Hutyra, H. R. da Rocha, R. G. Aguiar, e A. C. de Araújo (2007a), Atmospheric versus vegetation controls of amazonian tropical rainforest evapotranspiration: Are the equatorial and tropical rainforests any different?, submetido ao Journal of Geophysical Research – Biogeosciences, LBA Special Section.
- Costa, M. H., S. N. M. Yanagi, P. J. O. Souza, A. Ribeiro, e E. J. P. Rocha (2007b), Climate change in amazonia caused by soybean cropland expansion, as compared to caused by pastureland expansion, *Geophysical Research Letters*, 34.
- Curtis, P. S., P. J. Hanson, P. Bolstad, C. Barford, J. C. Randolph, H. P. Schmid, e K. B. Wilson (2002), Biometric and eddy-covariance based estimates of annual carbon storage in five eastern north american deciduous forests, *Agricultural and Forest Meteorology*, 113, 3–19.
- Deb, K., A. Pratap, S. Agarwal, e T. Meyarivan (2002), A fast and elitist multi-objective genetic algorithm: Nsga-ii, *IEEE Transactions on Evolutionary Computation*, 6(2).
- Duan, Q., S. Sorooshian, e V. Gupta (1992), Effective and efficient global optimization for conceptual rainfall-runoff models, *Water Resources Research*, 28(4), 1015–1031.

- Duan, Q., J. Schaake, V. Andréassian, S. Franks, G. Goteti, H. V. Gupta, Y. M. Gusev, F. Habets, A. Hall, L. Hay, T. Hogue, M. Huang, G. Leavesley, X. Liang, O. N. Nasonova, J. Noilhan, L. Oudin, S. Sorooshian, T. Wagener, e E. F. Wood (2006), Model parameter estimation experiment (mopex): An overview of science strategy and major results from the second and third workshops, *Journal of Hydrology*, 320, 3–17.
- Duan, Q. Y., V. K. Gupta, e S. Sorooshian (1993), Shuffled complex evolution approach for effective and efficient global minimization, *Journal of Optimization Theory and Applications*, 76(3), 501–521.
- Farquhar, G. D., S. Von Caemmerer, e J. A. Berry (1980), A biochemical model of photosynthetic co<sub>2</sub> assimilation in leaves of c<sub>3</sub> species, *Planta*, 149, 78–90.
- Foley, J. A., I. C. Prentice, N. Ramankutty, S. Levis, D. Pollard, S. Sitch, e A. Haxeltine (1996), An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics, *Global Biogeochemical Cycles*, 10(4), 603–628.
- Fox, D. G. (1980), Judging air quality model performance, *Bulletin of American Meteorological Society*, 62, 599–609.
- Fu, M. C., F. W. Glover, e J. April (2005), Simulation optimization: A review, new developments, and applications, in *Proceedings of the 2005 Winter Simulation Conference*, editado por M. Kuhl, N. Steiger, F. Armstrong, e J. Joines, pp. 83–95, Institute of Electrical and Electronics Engineers, Piscataway, New Jersey.
- Goldberg, D. E. (1989), *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison Wesley, New York.
- Gupta, H. V., e S. Sorooshian (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resources Research*, 34(4), 751–763.
- Imbuzeiro, H. M. A. (2005), Calibração do modelo ibis na floresta amazônica usando múltiplos sítios, Tese de Mestrado, Universidade Federal de Viçosa.
- Johnston, P. R., e D. H. Pilgrim (1976), Parameter optimization for watershed models, *Water Resources Research*, 12(3), 477–486.
- Jones, D. F., S. K. Mirrazavi, e M. Tamiz (2002), Multi-objective meta-heuristics: An overview of the current state-of-the-art, *European Journal of Operational Research*, 137(1), 1–9.
- Kruijt, B., J. A. Elbers, C. von Randow, A. C. Araujo, P. J. Oliveira, A. Gulf, A. O. Manzi, A. D. Nobre, P. Kabat, e E. J. Moors (2004), The robustness of eddy correlation fluxes for amazon rain forest conditions, *Ecological applications*, 14, 101–113.

- Kucharik, C. J., J. A. Foley, C. Delire, V. A. Fisher, M. T. Coe, J. D. Lenters, C. Young-Molling, N. Ramankutty, J. M. Norman, e S. T. Gower (2000), Testing the performance of a dynamic global ecosystem model: Water balance, carbon balance, and vegetation structure, *Global Biogeochem. Cycles*, *14*(3), 795-826.
- Legates, D. R., e G. J. McCabe Jr. (1999), Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, *Water Resources Research*, *35*(1), 233-241.
- Luan, J., R. I. Muetzelfeldt, e J. Grace (1996), Hierarchical approach to forest ecosystem simulation, *Ecological Modeling*, *86*, 37-50.
- Maayar, M. E., J. M. Chena, e D. T. Price (2008), On the use of field measurements of energy fluxes to evaluate land surface models, *Ecological Modelling*, *214*, 293-304.
- Madsen, H. (2000), Automatic calibration of a conceptual rainfall-runoff model using multiple objectives, *Journal of Hydrology*, *235*(3), 276-288.
- Madsen, H. (2003), Parameter estimation in distributed hydrological catchment modelling using automatic calibration with multiple objectives, *Advances in Water Resources*, *26*(2), 205-216.
- Morris, M. D. (1991), Factorial sampling plans for preliminary computational experiments, *Technometrics*, *33*(2), 161-174.
- Nash, J. E., e J. V. Sutcliffe (1970), River flow forecasting through conceptual models part i - a discussion of principles, *Journal of Hydrology*, *10*, 282-290.
- O'Neill, R. V. (1988), *SCOPE 35 - Scales and Global Change : Spatial and Temporal Variability in Biospheric and Geospheric Processes*, cap. Hierarchy Theory and Global Change.
- Pickup, G. (1977), Testing the efficiency of algorithms and strategies for automatic calibration of rainfall-runoff models, *Hydrological Sciences Bulletin*, *22*(2), 257-274.
- Pitman, A. J. (2003), The evolution of, and revolution in, land surface schemes designed for climate models, *International Journal of Climatology*, *23*(5), 479-510.
- Praditwong, K., e X. Yao (2007), How well do multi-objective evolutionary algorithms scale to large problems, *IEEE Congress on Evolutionary Computation*, pp. 3959-3966.
- Priante-Filho, N., G. L. Vourlitis, M. M. S. Hayashi, J. de S. Nogueira, J. H. C. Jr, P. C. Nunes, L. S. Souza, E. G. Couto, W. Hoeger, F. Raiter, J. L. Trienweiler, E. J. Miranda, P. C. Priante, L. C. Pereira, M. S. Biudes, C. L. Fritzen, M. Lacerda, G. S. Suli, S. Shiraiwa, e M. Silveira (2004), Comparison of the mass and energy exchange of a pasture and a mature transitional tropical forest of the southern amazon basin during a seasonal transition, *Global Change Biology*, *10*, 863-876, doi:10.1111/j.1529-8817.2003.00775.x.



- Purshouse, R. C., e P. J. Fleming (2003), Conflict, harmony, and independence: Relationships in evolutionary multi-criterion optimization, in *EMO*, pp. 16–30.
- Reeves, C. R. (Ed.) (1995), *Modern heuristic techniques for combinatorial problems*, McGraw-Hill Publishing Co., New York.
- Saleska, S. R., S. D. Miller, D. M. Matross, M. L. Goulden, S. C. Wofsy, H. R. da Rocha, P. B. de Camargo, P. Crill†, B. C. Daube, H. C. de Freitas, L. Hutyrá, M. Keller, V. Kirchoff, M. Menton, J. WilliamMunger, E. H. Pyle, A. H. Rice, e H. Silva (2003), Carbon in amazon forests: Unexpected seasonal fluxes and disturbance-induced losses, *Science*, *302*, 1554–1557.
- Saltelli, A., K. Chan, e E. M. Scott (Eds.) (2000), *Sensitivity Analysis*, John Wiley & Sons Ltd, Chichester.
- Saltelli, A., S. Tarantola, F. Campolongo, e M. Ratto (2004), *Sensitivity analysis in practice: a guide to assessing scientific models*, John Wiley & Sons Ltd.
- Saltelli, A., M. Ratto, S. Tarantola, e F. Campolongo (2005), Sensitivity analysis for chemical models, *Chemical reviews*, *105*(7), 2811–2827.
- Sampaio, G., C. Nobre, M. H. Costa, P. Satyamurty, B. S. SoaresFilho, e M. Cardoso (2007), Regional climate change over eastern amazonia caused by pasture and soybean cropland expansion, *Geophysical Research Letters*, *34*.
- Sellers, P. J., R. E. Dickinson, D. A. Randall, A. K. Betts, F. G. Hall, J. A. Berry, G. J. Collatz, A. S. Denning, H. A. Mooney, C. A. Nobre, N. Sato, C. B. Field, e A. Henderson-Sellers (1997), Modeling the exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, *275*(5299), 502–509.
- Senna, M. C. A. (2004), Fração da radiação fotossinteticamente ativa absorvida pela floresta tropical amazônica: Uma comparação entre estimativas baseadas em modelagem, sensoriamento remoto e medições de campo., Tese de Mestrado, Universidade Federal de Viçosa.
- Spolador, J., L. Sanches, e M. H. Costa (2006), Radiação fotossinteticamente ativa em uma floresta de transição cerrado-amazônica, *Revista Brasileira de Meteorologia*, *21*(3b), 301–307.
- Srinivas, N., e K. Deb (1995), Multiobjective function optimization using nondominated sorting genetic algorithms, *Evolutionary Computation*, *2*(3), 221–248.
- Tang, Y., P. Reed, e T. Wagener (2006), How effective and efficient are multi-objective evolutionary algorithms at hydrologic model calibration?, *Hydrology and Earth System Sciences*, *10*, 289–307.
- Trudinger, C. M., M. R. Raupach, P. J. Rayner, J. Kattge, Q. Liu, B. Pak, M. Reichstein, L. Renzullo, A. D. Richardson, S. H. Roxburgh, J. Styles, Y. P. Wang, P. Briggs, D. Barrett, e S. Nikolova (2007), OptIC project: An inter-comparison of optimization techniques for parameter estimation in terrestrial biogeochemical models, *Journal of Geophysical Research*, *112*(G02027).

- Turner, M. G., V. H. Dale, e R. H. Gardner (1989), Predicting across scales: Theory development and testing, *Landscape Ecology*, 3(3/4), 245–252.
- Veldhuizen, D. A. V., e G. B. Lamont (2000), Multiobjective evolutionary algorithms: Analyzing the state-of-the-art, *Evolutionary Computation*, 8(2), 125–147.
- Vrugt, J. A., H. V. Gupta, L. A. Bastidas, W. Bouten, e S. Sorooshian (2003), Effective and efficient algorithm for multiobjective optimization of hydrologic models, *Water Resources Research*, 39(8), SWC5.1–SWC5.19.
- Willmott, C. J. (1981), On the validation of models, *Physical Geography*, 2(2), 184–194.
- Willmott, C. J. (1982), Some comments on the evaluation of model performance, *Bulletin American Meteorological Society*, 63(11), 1309–1313.
- Willmott, C. J., e K. Matsuura (2005), Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance, *Climate Research*, 30, 79–82.
- Willmott, C. J., S. G. Ackleson, R. E. Davis, H. J. Feddema, K. M. Klink, D. R. Legates, J. O’Donnell, e C. M. Rowe (1985), Statistics for the evaluation and comparison of models, *Journal of Geophysical Research*, 90(C5), 8995–9005.
- Wilson, K., A. Goldstein, E. Falge, M. Aubinet, D. Baldocchi, P. Berbigier, C. Bernhofer, R. Ceulemans, H. Dolman, C. Field, A. Grelle, A. Ibrom, B. E. Law, A. Kowalski, T. Meyers, J. Moncrieff, R. Monson, W. Oechel, J. Tenhunen, R. Valentini, e S. Verma (2002), Energy balance closure at fluxnet sites, *Agricultural and Forest Meteorology*, 113, 223–243.
- Xia, Y., A. J. Pitman, H. V. Gupta, M. Leplastrier, A. Henderson-Sellers, e L. A. Bastidas (2002), Calibrating a land surface model of varying complexity using multicriteria methods and the cabauw dataset, *Journal of Hydrometeorology*, 3, 181–194.
- Yapo, P. O., H. V. Gupta, e S. Sorooshian (1998), Multi-objective global optimization for hydrologic models, *Journal of Hydrology*, 204, 83–97.

## APÊNDICES

# APÊNDICE A

## Manual do Optis

Este é o conteúdo do arquivo Readme-Optis, que contém instruções para o uso do software de calibração hierárquica Optis (usando NSGA-II). Ele se encontra no diretório raiz do programa Optis.

### Sobre o programa

-----  
Optis: Software de calibração hierárquica multiobjetivo.  
Escrito na linguagem C. Paralelização do código com OpenMP.  
Por favor, referencie a seguinte dissertação de mestrado para detalhes sobre a metodologia de calibração hierárquica multiobjetivo:  
VAREJÃO-JR, Claudeci Gomes, M. S., Universidade Federal de Viçosa, Abril de 2009. Calibração hierárquica multiobjetivo de um modelo de micrometeorologia e de dinâmica de ecossistemas terrestres. Orientador: Marcos Heil Costa. Conselheiros: Gilberto Chohaku Sedyama e Luciana Sanches.  
-----

### Como compilar e rodar o programa

-----  
Para compilar o Optis, digite o comando make na pasta raiz do programa.  
Para executar o Optis, execute o script ./runoptis.sh na pasta raiz do programa  
-----

### Diretórios e arquivos

-----  
Dentro do diretório raiz do Optis existem três subdiretórios: optis/ibis/, optis/input\_data/ e optis/output\_data/

ibis/ : diretório do modelo IBIS

optis/input\_data/ : contém os arquivos de configuração do NSGA-II e do Optis.  
nsga.conf : é o arquivo de configuração do algoritmo genético NSGA-II.  
calibration.conf : é o arquivo de configuração do Optis, onde são configuradas as calibrações do modelo a serem executadas.

optis/output\_data/ : diretório onde são gravados os arquivos de saída do algoritmo, com os resultados das calibrações.

Os nomes dos arquivos de saída seguem o seguinte padrão:

num\_nivel.nome\_nivel.nome\_arquivo.extensão

Exemplo: '02.Rn.ibis\_params.out' ,

onde 02 é o nível hierárquico calibrado, Rn é o nome do nível, e ibis\_params.out é o arquivo de saída.

A maioria dos arquivos de saída são informações do algoritmo genético. A descrição detalhada do NSGA-II e seus arquivos pode ser encontrada no arquivo 'Readme', no diretório raiz do Optis.

Os arquivos de saída relevantes à calibração do modelo são:

ibis\_params.out : contém a lista dos parâmetros calibrados e os valores otimizados

ibis\_parvalues.csv : arquivo com os parâmetros calibrados, em um formato para ser lido pelo IBIS através da opção -r do IBIS (./ibis -r csv), onde csv podem ser 3 caracteres quaisquer, que indicam a extensão do arquivo de entrada "parvalues.xxx". A primeira linha desse arquivo contém as máscaras dos parâmetros calibrados (0 para parâmetros não calibrados e 1 para parâmetros calibrados), e a segunda linha contém os valores dos parâmetros calibrados.  
-----

Parâmetros de entrada do algoritmo NSGA-II, inseridos no arquivo input\_data/nsga.conf

-----  
popsizel: Essa variável armazena o tamanho da população (múltiplo de 4)

ngen: Número de gerações

ncon: Número de restrições

pcross\_real: Probabilidade de crossover

pmut\_real: Probabilidade de mutação

eta\_c: índice de distribuição de crossover

eta\_m: índice de distribuição de mutação  
 nbim: número de variáveis binárias  
 choice: Opção para exibir os dados em tempo real usando Gnuplot (deixe isso como 0)  
 obj1, obj2, obj3: índices dos objetivos a serem exibidos nos eixos x, y e z (se choice >> 0)  
 angle1, angle2: ângulos requeridos para exibição do gráfico (se choice >> 0)

-----  
 Parâmetros de configuração do Optis, inseridos no arquivo input\_data/calibration.conf

-----  
 ical: Índice do nível hierárquico de calibração  
 calname: Nome do nível hierárquico de calibração, usados para os nomes dos arquivos de saída  
 npars: Número de parâmetros reais usados nessa calibração  
 i min[i] max[i]: Índice do i-ésimo parâmetro (tabela abaixo), seu valor mínimo e máximo  
 nvars: Número de variáveis de saída calibradas nesse nível hierárquico  
 j: Índice da j-ésima variável de saída (veja tabela abaixo)  
 nstats: número de medidas de ajuste usadas nessa calibração  
 k: Índice da k-ésima medida de ajuste utilizada (veja tabela abaixo)

-----  
 NOTA: O número de objetivos do nível hierárquico será igual a (nvars \* nstats)

Tabela 1 - Índices dos parâmetros, usados para configurar o Optis

0 = 'rho veg_vis'	16 = 'tempvm_coef'	32 = 'tauleaf'
1 = 'rho veg_NIR'	17 = 'stressf_coef'	33 = 'tauroot'
2 = 'tau veg_vis'	18 = 'clitll_coef'	34 = 'tauwood0'
3 = 'tau veg_NIR'	19 = 'clitlm_coef'	35 = 'specla'
4 = 'chifuz'	20 = 'clitls_coef'	36 = 'aleaf' *
5 = 'vmax_pft'	21 = 'clitrl_coef'	37 = 'aroot' *
6 = 'coefmub'	22 = 'clitrm_coef'	38 = 'awood' *
7 = 'chs'	23 = 'clitrs_coef'	39 = 'dispu_coef'
8 = 'chu'	24 = 'clitwl_coef'	40 = 'alogl_coef'
9 = 'chl'	25 = 'clitwm_coef'	41 = 'alogu_coef'
10 = 'beta2'	26 = 'clitws_coef'	42 = 'avmuir_coef'
11 = 'funca_coef'	27 = 'csoipas_coef'	
12 = 'funcb_coef'	28 = 'csoislon_coef'	
13 = 'rroot_coef'	29 = 'csoislop_coef'	
14 = 'rwood_coef'	30 = 'wsoi_coef'	
15 = 'rgrowth_coef'	31 = 'kfactor'	

\* Devem ser calibrados sempre juntos.

Tabela 2 - Índices das variáveis, usados para configurar o Optis

0 = 'Rn'
1 = 'PARin'
2 = 'PARout'
3 = 'G'
4 = 'H'
5 = 'LE'
6 = 'NEE'
7 = 'ustar'
8 = 'CO2FLUX'
9 = 'FAPAR'
10 = 'LAI'
11 = 'NPP'
12 = 'LIT'
13 = 'BIOMASS'
14 = 'NEE Diurno'
15 = 'NEE Noturno'

Table 3 - Índices das estatísticas, usados para configurar o Optis

0 = 'MAE'
1 = 'Bmax'
2 = 'RMSE'

## APÊNDICE B

### Calibration.conf (Exemplo)

Exemplo de arquivo calibration.conf, para calibrar 6 saídas do modelo em 3 níveis hierárquicos.

```
-----
3          # Numero de niveis hierarquicos

1          # Indice do nivel hierarquico
PARo-fAPAR # Nome do nivel hierarquico (sem espacos, max. 10 caracteres)
3          # Numero de parametros calibrados neste nivel
0 0.05 0.15 # Indice, minimo e maximo do 1o. parametro (veja Readme-Optis)
2 0.03 0.07 # Indice, minimo e maximo do 2o. parametro
4 -1.0 1.0  # Indice, minimo e maximo do 3o. parametro
2          # Numero de variaveis a serem calibradas neste nivel
2          # Indice da 1a. variavel (veja Readme-Optis)
9          # Indice da 2a. variavel
1          # Numero de medidas de ajuste usadas neste nivel
0          # Indice da 1a. medida de ajuste (veja Readme-Optis)

2          # Indice do nivel hierarquico
Rn        # Nome do nivel hierarquico (sem espacos, max. 10 caracteres)
3          # Numero de parametros calibrados neste nivel
1 0.20 0.60 # Indice, minimo e maximo do 1o. parametro (veja Readme-Optis)
3 0.14 0.26 # Indice, minimo e maximo do 2o. parametro
42 0.5 995.0 # Indice, minimo e maximo do 3o. parametro
1          # Numero de variaveis a serem calibradas neste nivel
0          # Indice da 1a. variavel (veja Readme-Optis)
2          # Numero de medidas de ajuste usadas neste nivel
0          # Indice da 1a. medida de ajuste (veja Readme-Optis)
1          # Indice da 2a. medida de ajuste

3          # Indice do nivel hierarquico
NEE-H-LE  # Nome do nivel hierarquico (sem espacos, max. 10 caracteres)
16        # Numero de parametros calibrados neste nivel
5 4.55E-05 50.55E-05 # Indice, minimo e maximo do 1o. parametro (veja Readme-Optis)
6 5 15     # Indice, minimo e maximo do 2o. parametro
7 0.1E+05 0.4E+05  # Indice, minimo e maximo do 3o. parametro
10 0.750 0.999     # Indice, minimo e maximo do 4o. parametro
16 1750 7000      # Indice, minimo e maximo do 5o. parametro
17 -6.5 -3.5     # Indice, minimo e maximo do 6o. parametro
11 1750 15000    # Indice, minimo e maximo do 7o. parametro
12 1750 7000      # Indice, minimo e maximo do 8o. parametro
13 0.625 2.500   # Indice, minimo e maximo do 9o. parametro
14 0.00625 0.200 # Indice, minimo e maximo do 10o. parametro
20 0.0009 3.0351 # Indice, minimo e maximo do 11o. parametro
23 0.000877 5.052430 # Indice, minimo e maximo do 12o. parametro
26 0.2926 2.9999 # Indice, minimo e maximo do 13o. parametro
28 0.063860 2.255740 # Indice, minimo e maximo do 14o. parametro
29 1.000030 16.00000 # Indice, minimo e maximo do 15o. parametro
31 0.7 2.9       # Indice, minimo e maximo do 16o. parametro
3          # Numero de variaveis a serem calibradas neste nivel
4          # Indice da 1a. variavel (veja Readme-Optis)
5          # Indice da 2a. variavel
6          # Indice da 3a. variavel
1          # Numero de medidas de ajuste usadas neste nivel
0          # Indice da 1a. medida de ajuste (veja Readme-Optis)
-----
```

NOTA: Para que o arquivo calibration.conf possa ser lido corretamente pelo algoritmo, ele deve seguir exatamente o mesmo formato do exemplo: mesma ordem dos parâmetros de configuração, uma linha em branco separando cada nível hierárquico, comentário no final de cada linha de configuração.