

Universidade Federal de Juiz de Fora
Faculdade de Engenharia
Programa de Pós Graduação em Engenharia Elétrica

Marina Borges Arantes de Souza

Técnicas de Otimização Aplicadas à Identificação de Sistemas Dinâmicos

Juiz de Fora
2023

Marina Borges Arantes de Souza

Técnicas de Otimização Aplicadas à Identificação de Sistemas Dinâmicos

Tese de doutorado apresentada ao Programa de Pós Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora, na área de concentração em Sistemas de Energia Elétrica, como requisito para obtenção do título de Doutor em Engenharia Elétrica.

Orientador: Edimar José de Oliveira

Coorientador: Leonardo de Mello Honório

Juiz de Fora

2023

Ficha catalográfica elaborada através do Modelo Latex do CDC da UFJF
com os dados fornecidos pelo(a) autor(a)

Souza, Marina Borges Arantes de.

Técnicas de Otimização Aplicadas à Identificação de Sistemas Dinâmicos / Marina Borges Arantes de Souza. – 2023.

95 f. : il.

Orientador: Edimar José de Oliveira

Coorientador: Leonardo de Mello Honório

Qualificação (Doutorado) – Universidade Federal de Juiz de Fora, Faculdade de Engenharia. Programa de Pós Graduação em Engenharia Elétrica, 2023.

1. Identificação de Sistemas Dinâmicos. 2. Estimação de Parâmetros. 3. Sistemas Não Lineares. 4. Geração de Sinal de Excitação. 5. Modelo híbrido. I. Oliveira, Edimar José de, orient. II. Mello Honório, Leonardo de, coorient. III. Título.

Marina Borges Arantes de Souza

Técnicas de Otimização Aplicadas à Identificação de Sistemas Dinâmicos

Tese apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Engenharia Elétrica. Área de concentração: Sistemas de Energia Elétrica

Aprovada em 30 de março de 2023.

BANCA EXAMINADORA

Prof. Dr. Edimar José de Oliveira - Orientador

Universidade Federal de Juiz de Fora

Prof. Dr. Leonardo de Mello Honório - Coorientador

Universidade Federal de Juiz de Fora

Prof. Dr. Raphael Paulo Braga Poubel

Centro Federal de Educação Tecnológica de Minas Gerais

Profa. Dra. Milena Faria Pinto

Centro Federal de Educação Tecnológica Celso Suckow da Fonseca

Prof. Dr. Alexandre Bessa dos Santos
Universidade Federal de Juiz de Fora

Prof. Dr. Leonardo Willer de Oliveira
Universidade Federal de Juiz de Fora

Juiz de Fora, 24/03/2023.



Documento assinado eletronicamente por **Edimar Jose de Oliveira, Professor(a)**, em 30/03/2023, às 16:57, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Milena Faria Pinto, Usuário Externo**, em 30/03/2023, às 17:04, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo Willer de Oliveira, Professor(a)**, em 30/03/2023, às 17:09, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **RAPHAEL PAULO BRAGA POUHEL, Usuário Externo**, em 30/03/2023, às 17:24, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Alexandre Bessa dos Santos, Professor(a)**, em 30/03/2023, às 18:35, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Leonardo de Mello Honorio, Professor(a)**, em 27/04/2023, às 15:24, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **1201963** e o código CRC **3A43AB96**.

AGRADECIMENTOS

Primordialmente, agradeço a Deus pelo dom da vida e por me iluminar, trazendo clareza de qual caminho seguir.

Agradeço imensamente aos meus pais, Walter e Sandryonara, que estiveram sempre ao meu lado, me incentivando e me ajudando a passar pelos momentos de alegria e de dificuldades. A toda minha família, meu muito obrigada.

Ao meu marido, Vitor, que esteve presente sempre me motivando a continuar e não desistir.

O meu agradecimento especial, ao meu orientador Prof. Edimar Oliveira, que acreditou em mim e que possibilitou aprender um pouco com seus exemplos e experiências. Sem a sua força e seus "puxões de orelha" não teria chegado ao fim dessa jornada. Ao meu coorientador Prof. Leonardo Honório por ter me dado a oportunidade de trabalhar no GRIn com pessoas tão capacitadas.

A todos os membros do GRIn pela parceria e um agradecimento especial ao Accacio que, com sua paciência e lealdade, me ajudou a conduzir esse trabalho.

Ao corpo docente do Programa de Pós Graduação em Engenharia Elétrica da UFJF por todos ensinamentos recebidos.

Aos professores membros da banca pelas contribuições apresentadas.

Ao INERGE, pelo suporte técnico e à CAPES, CNPq e FAPEMIG pelo auxílio financeiro e estrutural que possibilitou a realização desse trabalho.

Por fim, agradeço a todos que contribuíram de alguma forma.

“A suficiência dos meus méritos
está em saber que meus méritos
não são suficientes”
(Santo Agostinho)

RESUMO

A construção de um modelo confiável é um dos pilares do controle de sistemas autônomos. A identificação é um processo que extrai características de um processo para posterior análise e manipulação do sistema. Modelos matemáticos são um direcionamento do comportamento real de um processo. Entretanto, tais modelos exigem a entrada de alguns parâmetros específicos da operação do sistema. A estimação de parâmetros, é uma forma de identificação que objetiva afinar modelos para que eles se comportem de forma próxima à realidade. Para tanto, utiliza-se métodos de estimação, que aliados a um sinal de entrada eficiente, podem produzir bons resultados. O *Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation* (SOESGOPE) é um método apresentado pela literatura que combina a geração de sinais de excitação com a estimação de parâmetros. Embora a metodologia apresente resultados satisfatórios, deve-se analisá-la sob o ponto de vista de otimização, ou seja o quanto ele é confiável e robusto frente a diferentes configurações. Assim, este trabalho apresenta análises nesse aspecto e uma proposta de alteração quando for verificado uma falta de convergência. A aplicabilidade do método é potencializada em situações em que o modelo seja uma cópia razoável do comportamento real. Entretanto, em situações reais, não se pode garantir que um modelo matemático produza saídas similares ao processo real frente a uma mesma entrada. Para contornar esse problema, neste trabalho também é proposta uma técnica híbrida que é capaz de utilizar o modelo matemático como direcionamento e aprender características peculiares àquele processo. Tal modelo híbrido, composto por etapa de estimação e de aprendizado de comportamento, será capaz de corrigir certos erros do sistema real relacionados ao modelo e atingir saídas mais confiáveis através da combinação do modelo teórico com técnicas de aprendizado. Além disso, o SOESGOPE é um aliado ao processo de aprendizado visto que fornece sinais de excitação mais eficientes. A eficácia do método é comprovada através da aplicação em uma embarcação autônoma com características desafiadoras sob a ótica de identificação, reduzindo o erro de estimação de trajetória.

Palavras-chave: Identificação de Sistemas Dinâmicos. Estimação de Parâmetros. Sistemas Não Lineares. Geração de Sinal de Excitação. Modelo híbrido.

ABSTRACT

Building a reliable model is one of the pillars of controlling autonomous systems. Identification is a process that extracts characteristics of a process for further analysis and manipulation of the system. Mathematical models are a guide to the real behavior of a process. However, such models require the input of some specific parameters of the system operation. Parameter estimation is a form of identification that aims to fine-tune models so that they behave as closely as possible to reality. Therefore, estimation methods together with an efficient input signal can produce good results. SOESGOPE is a method that combines excitation signal generation with parameter estimation. Although it presents satisfactory results, it should be analyzed from an optimization point of view, that is, how reliable and robust it is against different configurations. Thus, this work presents analysis in this aspect and a proposal for alteration when a lack of convergence is verified. The applicability of the method is enhanced in situations where the model is a reasonable copy of real behavior. However, in real situations, it is not possible to guarantee that a mathematical model will produce outputs similar to the real process when faced with the same input. To circumvent this problem, this work also proposes a hybrid technique that is able to use the mathematical model as a guide and learn characteristics peculiar to that process. Such a hybrid model will be able to correct certain errors of the real system related to the model and achieve more reliable outputs. Furthermore, SOESGOPE can count as an ally to the learning process since it provides more efficient excitation signals. To confirm the applicability of the proposed method, an autonomous boat is used for application and verification of the effectiveness of the method.

Keywords: Identification of Dynamical Systems. Parameter Estimation. Nonlinear Systems. Excitation Signal Generation. Hybrid model.

LISTA DE ILUSTRAÇÕES

Figura 1 – Tipos de modelos matemáticos	22
Figura 2 – Diagrama esquemático do processo geral do método de estimação de parâmetros baseado em mínimos quadrados	25
Figura 3 – Sinal binário pseudo-randômicos com 4 estágios e $N = 15$	29
Figura 4 – Sinal binário pseudo-randômicos com amplitude modulada	30
Figura 5 – Diagrama esquemático do processo SOESGOPE	38
Figura 6 – Robô móvel diferencial: <i>Very Small Size Soccer</i>	40
Figura 7 – Valores máximos e mínimos da função objetivo com diferentes número de partículas para cada modelo	47
Figura 8 – Valores máximo mínimo da função de recuperabilidade de parâmetros - $f_{\delta}(\cdot)$ - em escala logarítmica para diferentes níveis de confiabilidade	49
Figura 9 – Gráfico de convergência dos parâmetros	53
Figura 10 – Diagrama esquemático da estrutura híbrida: modelo paramétrico e rede neural profunda (RNP)	64
Figura 11 – Diagrama esquemático do processo em série: estimação de parâmetros e correção do comportamento	65
Figura 12 – Estrutura do modelo híbrido resultante	65
Figura 13 – Processo iterativo das redes GRU e LSTM	68
Figura 14 – Catamarã autônomo com propulsão aérea	70
Figura 15 – Esquema de forças de um catamarã equipado com quatro propulsores azimutais (forças F_1 , F_2 , F_3 e F_4) com forças descompostas em x e y	75
Figura 16 – Diagrama esquemático da estrutura de refinamento, contendo o refinamento da taxa angular e velocidades lineares	79
Figura 17 – Esquema da arquitetura RNP, composta pelas camadas <i>flatten</i> e GRU	80
Figura 18 – Curvas dos sinais subótimos gerados pelo método SOESGOPE, indicando a força que deve ser aplicada em cada grau de liberdade no movimento de teste	82
Figura 19 – Trajetória estimada por modelo matemático (MM) e modelo híbrido (HM), considerando sinais subótimos 1, 2 e 3 com u [m/s], v [m/s], r [rad/s]	83
Figura 20 – Trajetória estimada por modelo matemático (MM) e modelo híbrido (HM), considerando os sinais manuais 4 (validação), 5 e 6 (teste)	84
Figura 21 – Impacto do número da amostras (n) na melhoria do modelo matemático, considerando 100 rodadas de treinamento	86

LISTA DE TABELAS

Tabela 1 – Possibilidades de funções objetivo para o processo de estimação paramétrica através de métodos iterativos, sendo que c é uma constante definida previamente.	27
Tabela 2 – Conjunto de parâmetros do modelo do robô “real”	42
Tabela 3 – Conjunto de parâmetros do modelo compacto do robô “real”	42
Tabela 4 – Estimativas inicial e perturbada do modelo completo	44
Tabela 5 – Estimativas inicial e perturbada do modelo compacto	44
Tabela 6 – Métricas relacionadas à função objetivo da camada externa	44
Tabela 7 – Estimativa a posteriori do modelo completo	45
Tabela 8 – Estimativa a posteriori do modelo compacto	45
Tabela 9 – Sensibilidade ao número de partículas - modelo completo	46
Tabela 10 – Sensibilidade ao número de partículas - modelo compacto	46
Tabela 11 – Métricas relacionadas à melhor estimação fornecida pelo SOESGOPE para cada modelo	48
Tabela 12 – Métricas relacionadas à pior estimação fornecida pelo SOESGOPE para cada modelo	48
Tabela 13 – Experimento recursivo - Métricas	54
Tabela 14 – Experimento recursivo - Parâmetros	54
Tabela 15 – Estimativa a priori definida para o experimento	54
Tabela 16 – Região de busca definida para o experimento	55
Tabela 17 – Experimento recursivo - Alterações na região de busca	55
Tabela 18 – Experimento recursivo - Parâmetros	55
Tabela 19 – Região de busca - SOESGOPE original	56
Tabela 20 – Estimativa a priori - Comparação entre método original e recursivo	56
Tabela 21 – Melhor, pior e média das estimativas a posteriori fornecidas pelo método recursivo	57
Tabela 22 – Limites operacionais do AERO4River	81
Tabela 23 – Parâmetros estimados para o VAS - Γ^+	82
Tabela 24 – RMSE relacionado ao modelo matemático (MM) e ao modelo híbrido (HM) para as faixas T1, T2 e T3	84
Tabela 25 – Raiz do erro quadrático médio relacionado ao modelo matemático (MM) e ao modelo híbrido (HM) para as trajetórias T4, T5 e T6	85
Tabela 26 – RMSE referente ao modelo matemático (MM) e ao modelo híbrido (HM) para as trilhas T1, T2 e T3, usando como conjunto de treinamento as trilhas T4, T5 e T6	86

LISTA DE ABREVIATURAS E SIGLAS

AABB	<i>Axis-Aligned Bounding Box</i>
APRBS	<i>Amplitude-modulated Pseudo-Random Binary Signal</i>
DoF	<i>Degrees of Freedom</i>
GPS	<i>Global Positioning System</i>
GRU	<i>Gated Recurrent Units</i>
IMU	<i>Inertial Measurement Unit</i>
LSTM	<i>Long-Short Term Memory</i>
MEP	Método de Estimação de Parâmetros
PNL	Programação Não Linear
PNLIM	Programação Não Linear Inteira Mista
PRBS	<i>Pseudo-Random Binary Signal</i>
PSLO	<i>Particle Swarm-Like Optimization</i>
PSO	<i>Particle Swarm Optimization</i>
RMSE	<i>Root Mean Squared Error</i>
RNA	Redes Neurais Artificiais
RNP	Redes Neurais Profundas
rSOESGOPE	<i>Robust Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation</i>
SOESGOPE	<i>Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation</i>
SVM	<i>Support Vector Machine</i>
VAS	Veículo Autônomo de Superfície
VSSS	<i>Very Small Size Soccer</i>

LISTA DE SÍMBOLOS

Γ	Parâmetros do sistema real
$\hat{\Gamma}^-$	Estimativa de parâmetros a priori
$\hat{\Gamma}^+$	Estimativa de parâmetros a posteriori
$\hat{\Gamma}^p$	Parâmetros do sistema perturbado
$\mathcal{R}(\Gamma)$	Sistema real
$\mathcal{M}(\Gamma)$	Modelo matemático associado ao sistema real
\mathbf{u}^\oplus	Sinal subótimo
$\Theta(\cdot)$	Parcela relacionada aos limites associados aos estados
$f_o(\cdot)$	Parcela relacionada à acurácia da saída
$f_\delta(\cdot)$	Parcela relacionadas à recuperabilidade dos parâmetros
$\kappa_1, \kappa_2, \kappa_3 \in \mathbb{R}_{\geq 0}$	Constantes de peso relacionadas às parcelas da função objetivo $\Theta(\cdot)$, $f_o(\cdot)$ e $f_\delta(\cdot)$, respectivamente
$f(\Xi)$	Função objetivo multiparamétrica
t_u	Tempo do experimento
$f_V(\cdot)$	Função que quantifica o uso do espaço
$\Psi(\cdot)$	Função de penalidade
α	Índice de relaxamento
$\mathcal{O}\mathbf{x}_{\min}$	Limite operacional mínimo das variáveis de estado
$\mathcal{O}\mathbf{x}_{\max}$	Limite operacional máximo das variáveis de estado
p_x	Posição do robô no eixo x no instante de tempo t (m)
p_y	Posição do robô no eixo y no instante de tempo t (m)
$\theta(t)$	Posição angular do robô no instante de tempo t (rad)
$v(t)$	Velocidade linear do robô no instante de tempo t (m/s)
$\omega(t)$	Velocidade angular do robô no instante de tempo t (rad/s)
$\dot{\phi}_r(t)$	Velocidade angular da roda direita no instante de tempo t (rad/s)

$\dot{\phi}_l(t)$	Velocidade angular da roda esquerda no instante de tempo t (rad/s)
m	Massa do sistema
I_{zz}	Momento de inércia em torno do eixo z
I_w	Momento de inércia das rodas
K_t	Constante de torque
K_w	Constante de tensão
R	Resistência de armadura (Ω)
ν	Coefficiente de atrito viscoso
r	Raio das rodas
b	Distância entre a roda e o centro de gravidade
p_1, p_2, p_3	Parâmetros do sistema compacto do VSSS
$\bar{\lambda}, \underline{\lambda}$	Vetores superior e inferior dos multiplicadores de Lagrange
η	Vetor de posição e orientação de acordo com o referencial inercial
ν	Vetor de velocidades lineares (u, v) e velocidade angular (r)
$\mathbf{R}(\psi)$	Matriz de rotação entre o referencial inercial e o referencial fixo à embarcação
\mathbf{M}	Matriz de inércia do sistema
$\mathbf{C}(\nu)$	Coriolis e matriz centrípeta
\mathbf{D}	Matriz de amortecimento linear
\mathbf{D}_n	Matriz de amortecimento não linear
ν	Vetor de velocidades lineares e angular
η	Vetor de posição e orientação
τ	Forças e momentos aplicados a embarcação
\mathbf{M}_{RB}	Matriz de inércia do sistema do corpo rígido
\mathbf{M}_A	Matriz de massa adicionada
\mathbf{C}_{RB}	Matriz de Coriolis do corpo rígido

M_A	Matriz de Coriolis hidrodinâmica
D	Matriz de amortecimento linear
D_n	Matriz de amortecimento não linear
x_g	Centro de gravidade da embarcação em relação ao eixo x
I_z	Momento de inércia em relação ao eixo z da embarcação
$T_{\dot{\nu}}$	Derivados hidrodinâmicos, sendo que $T = [X, Y, N]$ e $\nu = [u, v, r]$
$T(\alpha)$	Matriz de alocação de forças
Γ_{M_A}	Parâmetros relacionado à massa adicionada ($X_{\dot{u}}, X_{\dot{v}}, X_{\dot{r}}, Y_{\dot{u}}, Y_{\dot{v}}, Y_{\dot{r}}, N_{\dot{u}}, N_{\dot{v}}, N_{\dot{r}}$)
Γ_D	Parâmetros relacionado ao amortecimento linear ($X_u, X_v, X_r, Y_u, Y_v, Y_r, N_u, N_v, N_r$)
Γ_{D_n}	Parâmetros relacionado ao amortecimento não linear ($X_{ u u}, X_{ u v}, X_{ u r}, Y_{ u u}, Y_{ u v}, Y_{ u r}, N_{ u u}, N_{ u v}, N_{ u r}$)
F	Força aplicada

SUMÁRIO

1	INTRODUÇÃO	16
1.1	CONSIDERAÇÕES INICIAIS	16
1.2	MOTIVAÇÃO	18
1.3	DEFINIÇÃO DO PROBLEMA E OBJETIVOS	18
1.4	PUBLICAÇÕES DECORRENTES DA PESQUISA	19
1.5	ESTRUTURA DO DOCUMENTO	19
2	IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS	21
2.1	INTRODUÇÃO	21
2.2	MODELO MATEMÁTICO	22
2.2.1	Modelo não paramétrico	23
2.2.2	Modelo paramétrico	23
2.3	MÉTODOS DE ESTIMAÇÃO DE PARÂMETROS	24
2.3.1	Método dos Mínimos Quadrados	24
2.3.2	Métodos de Filtragem	26
2.3.3	Métodos de Otimização	27
2.4	SINAL DE EXCITAÇÃO	29
2.5	SOESGOPE	31
2.5.1	Camada Externa	33
2.5.1.1	<i>Função de Tempo e Estados</i>	<i>33</i>
2.5.1.2	<i>Função de Precisão das Saídas</i>	<i>35</i>
2.5.1.3	<i>Função de Recuperabilidade de Parâmetros</i>	<i>35</i>
2.5.1.4	<i>Formulação do problema de otimização</i>	<i>35</i>
2.5.1.5	<i>Algoritmo de Resolução</i>	<i>36</i>
2.5.2	Camada Interna	36
2.5.3	Estrutura SOESGOPE	37
2.5.4	Execução do Experimento	38
2.6	CONSIDERAÇÕES FINAIS	39
3	ANÁLISES REALIZADAS	40
3.1	MODELAGEM DO SISTEMA	40
3.2	ANÁLISE DA QUALIDADE DA ESTIMAÇÃO DE PARÂMETROS	43
3.2.1	Caso Tutorial	43
3.2.2	Impacto do tamanho da população	45
3.2.3	Robustez	47
3.3	ABORDAGEM RECURSIVA	50

3.3.1	SOESGOPE Recursivo	51
3.3.2	Avaliação do Caso Tutorial	52
3.3.2.1	<i>Parâmetros cobertos pela região de busca</i>	53
3.3.2.2	<i>Parâmetros Inatingíveis</i>	54
3.3.2.3	<i>Análise de Robustez</i>	57
3.4	CONSIDERAÇÕES FINAIS	57
4	MÉTODO HÍBRIDO DE IDENTIFICAÇÃO	59
4.1	INTRODUÇÃO	59
4.2	METODOLOGIAS DE IDENTIFICAÇÃO HÍBRIDA	61
4.3	SOESGOPE HÍBRIDO	63
4.3.1	Redes Neurais Recorrentes	65
4.3.1.1	<i>Memória Longa de Curto Prazo (LSTM)</i>	66
4.3.1.2	<i>Unidade Recorrente com Portas Fechadas (GRU)</i>	67
4.4	CONSIDERAÇÕES FINAIS	68
5	SISTEMA DE IDENTIFICAÇÃO HÍBRIDO - VAS	70
5.1	VEÍCULO AUTÔNOMO DE SUPERFÍCIE	70
5.1.1	Modelagem Cinemática	71
5.1.2	Modelagem Dinâmica	72
5.1.3	Forças de propulsão	74
5.1.4	Parâmetros Envolvidos	76
5.2	ARQUITETURA DE IDENTIFICAÇÃO - VAS	77
5.2.1	Aplicação do SOESGOPE	77
5.2.2	Arquitetura da RNP	78
5.3	RESULTADOS	81
5.3.1	Etapa de Refinamento	81
5.3.2	Efeito do número de amostras	85
5.3.3	Impacto do sinal de entrada	86
5.4	CONSIDERAÇÕES FINAIS	87
6	CONCLUSÕES	88
6.1	SUGESTÕES PARA DESENVOLVIMENTOS FUTUROS	89
	REFERÊNCIAS	91

1 INTRODUÇÃO

1.1 CONSIDERAÇÕES INICIAIS

Modelagem matemática é um aparato que permite determinar o comportamento de fenômenos de naturezas diversas, tais como biológicos, econômicos, químicos, elétricos, magnéticos e mecânicos. Em sistemas dinâmicos, o propósito da determinação de um modelo é criar uma função ou um conjunto de funções que descreva de maneira satisfatória a evolução de um sistema ao longo do tempo. Para fins de modelagem, tal fenômeno, por padrão, deve seguir um preceito de ligação entre os estados presente e passado. Dessa forma, estados atuais de um sistema dinâmico dependem de estados anteriores e o modelo decorrente é organizado através de equações diferenciais (SINGH; KUMAR; BALEANU, 2019).

Modelagem e controle de sistemas são campos cujos interesses em áreas como engenharia elétrica, engenharia mecânica e engenharia de processos estão em constante crescimento. O foco é, principalmente, devido ao crescente desenvolvimento de sistemas autônomos que exigem controle preciso e robusto. De modo a atingir essas características, o modelo matemático, que descreve o comportamento dinâmico de um processo, deve ser abrangente cobrindo possíveis lacunas relacionadas ao comportamento do processo (BROSILOW; JOSEPH, 2002).

Identificação é um processo pelo qual se consegue extrair um conjunto de informações que irá permitir a construção de um modelo que reflita o comportamento do sistema com maior fidelidade possível. Segundo Zadeh (1962), o termo identificação pode ser descrito como:

Determinação de um sistema, com base na observação de entradas e saídas, dentro de uma classe especificada à qual o sistema em teste é equivalente; determinação do estado inicial ou final do sistema. (ZADEH, 1962, p. 857, tradução nossa)

Seguindo essa ideia, o objetivo principal do processo de identificação é selecionar, dentro de uma classe à qual o sistema real pertence, um modelo equivalente que desvie do comportamento real o mínimo possível, isto é, um modelo que determine saídas similares frente às mesmas entradas inseridas. Uma vez identificado, um modelo matemático confiável é capaz de reproduzir de forma satisfatória o comportamento de um sistema dinâmico. A partir daí, é possível ter um conhecimento mais amplo do processo, sendo uma ferramenta essencial para validação do modelo teórico e determinação de condições extremas pelas quais o sistema pode ser submetido, através de simulações computacionais. Além de análise e simulações, a construção do modelo pode estar atrelada a outros tipos de aplicações que requerem uma maior credibilidade (BROSILOW; JOSEPH, 2002).

A determinação de um controle baseado em modelo é uma dentre as possíveis aplicações. Com a finalidade de sintonizar parâmetros de controle, é necessário o conhecimento de um modelo detalhado ou pelo menos um modelo que permita a determinação de fatores interessantes para o controle do processo. Sørensen (2005) introduziu o conceito de Modelo de Planta de Controle, do inglês, *Control Plant Model*. Tal modelo é uma descrição matemática simplificada que contém apenas as principais propriedades físicas do processo. Embora seja simplificado, sua identificação deve ser realizada. É a partir desse modelo que pode-se sintonizar parâmetros de controle, determinar controles preditivos ou adaptativos. Outra aplicação, que tem ganhado um foco nos últimos anos, devido à existência de modelos mais confiáveis, é a detecção de falhas. Nesse caso, à partir de informações obtidas pelo modelo, é possível apontar desvios de comportamento do sistema e, então, atrelá-los a possíveis falhas de funcionamento.

O conhecimento de um modelo que reproduza as características dinâmicas de um sistema é essencial para o entendimento de seu comportamento frente à diferentes configurações de atuação. Tal modelo se torna referência para estudos do sistema e também para a escolha de um controlador para sua operação. Uma diversidade de métodos pode ser aplicada a fim de identificar um processo. Entretanto, vale ressaltar que, inerente aos sistemas do mundo real, está a incerteza.

Segundo Weinmann (2012), no que diz respeito ao entendimento físico, determina-se o comportamento dinâmico do processo, ou seja, a estrutura, ordem, linearidade e multivariabilidade, sendo este comportamento não perturbado, denominado de sistema nominal. Entretanto, várias propriedades permanecem desconhecidas e é preciso analisá-las sob os seguintes aspectos:

- Incerteza da estrutura de parâmetros que é associada à determinação dos parâmetros do modelo matemático. Apenas em casos muito simples é possível obter parâmetros livres de interdependências. Sistemas complexos apresentam uma relação entre a determinação dos parâmetros e os coeficientes da dinâmica do sistema em malha fechada, impactando diretamente na controlabilidade do processo.
- Incerteza na variação prospectiva, que está associada à sensibilidade diferencial, ou seja, às mudanças nas propriedades do sistema, como desempenho e estabilidade.

Métodos experimentais de identificação são amplamente utilizados pela literatura, visto que é possível extrair, a partir de seu comportamento, peculiaridades associadas ao sistema. A estimação de parâmetros pode ser aplicada através de métodos de otimização que seleciona o conjunto de parâmetros que mais aproxima ao comportamento real. Neste caso, a atuação torna-se um aliado nesse processo. Sinais de entrada que produzam uma excitação necessária e suficiente para a estimação são importantes no experimento. Dessa

forma a categoria de técnicas de geração de um sinal ótimo são desejáveis a fim de melhorar a confiabilidade do método (HONÓRIO et al., 2018).

1.2 MOTIVAÇÃO

Outra dificuldade envolvida no processo de estimação de parâmetros pura, utilizando um sinal subótimo de excitação, é relacionada aos problemas de modelagem. Isto é, sistemas que não são cópias fiéis do modelo matemático descrito por um conjunto de equações diferenciais, cenário comum em processos reais. Os sistemas dinâmicos, embora possam seguir razoavelmente o comportamento de um modelo, dificilmente, conseguem ter comportamentos completamente similares. Os motivos podem ser devido a vários fatores, dentre eles, vícios construtivos, dispositivos de atuação que não produzem a força ou torque determinados e problemas de simetria. Nesse panorama, a simples estimação de parâmetro é incapaz de produzir comportamentos parecidos com os reais, mesmo com diferentes configurações de parâmetros, pois as equações não descrevem esse movimento. A determinação de um método que seja capaz de lidar com tal incompatibilidade é fortemente requerida caso exista a necessidade de encontrar um modelo confiável.

Geralmente, sistemas mais complexos possuem esta característica, devido ao alto nível de detalhamento do projeto, a construção não corresponde à realidade. Uma embarcação autônoma é um exemplo de aplicação que, além da possibilidade de inconsistência do modelo matemático, ainda conta com a exigência de um sistema de controle robusto.

1.3 DEFINIÇÃO DO PROBLEMA E OBJETIVOS

Face ao exposto, o presente trabalho propõe a elaboração de uma metodologia de identificação de sistemas dinâmicos baseada na associação de um modelo de aprendizagem ao processo de estimação paramétrica. Este método híbrido deve ser capaz de lidar tanto com incertezas associadas aos parâmetros do modelo matemático quanto à sensibilidade diferencial.

Seguindo essa proposta, ainda podemos citar como objetivos específicos:

- Analisar a estimação de parâmetros com sinal subótimo sob a ótica de otimização focando em aspectos como confiabilidade e robustez;
- Analisar medida de direcionamento do resultado de estimação para um ponto viável, caso o método não convirja para um resultado satisfatório;
- Identificar problemas inerentes ao método de estimação paramétrica puro;
- Aplicar a metodologia em um sistema real, no caso, uma embarcação autônoma e testar sua eficácia;

- Reduzir a erro entre o comportamento real do sistema e o comportamento estimado do modelo.

1.4 PUBLICAÇÕES DECORRENTES DA PESQUISA

Durante a realização desta pesquisa, foram obtidos alguns resultados, na área de otimização, que podem ser destacados:

1. SOUZA, M. B. A.; HONÓRIO, L. M; OLIVEIRA, E. J.; NETO, A. F. S.; SILVA, M. F.; PANCOTI, A. A Convolutional System Identification Approach Mixing Optimal Parameter Estimation and Deep Learning. *International Journal of Control, Automation and Systems*, Springer. Aceito para publicação.
2. SOUZA, M.B.A., HONORIO, L. M. OLIVEIRA, E.J. Innovative Analysis for Parameter Estimation Quality. *International Journal of Control, Automation and Systems* 19, 363–371 (2021). <<https://doi.org/10.1007/s12555-019-0909-4>>
3. SOUZA, M. B. A.; HONÓRIO, L. M; OLIVEIRA, E. J.; MOREIRA, A. P. G. M. Recursive Approach of Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation. *International Journal of Control, Automation and Systems*, 2020, Springer. <<https://doi.org/10.1007/s12555-019-0164-8>>
4. REBELLO, G. ; DE OLIVEIRA, E. JOSE ; SOUZA, M. B. A. Two-Stage General Variable Neighborhood Search Algorithm to Solve the Static Transmission Network Expansion Planning. In: *2019 IEEE Milan PowerTech, 2019, Milan*. <<https://doi.org/10.1109/PTC.2019.8810439>>

1.5 ESTRUTURA DO DOCUMENTO

Este trabalho possui quatro capítulos, além deste, que são descritos a seguir:

O Capítulo 2 aborda as principais metodologias aplicáveis à identificação de parâmetros. Além disso, inclui uma discussão em torno da seleção de sinais de excitação que influenciam diretamente no processo de estimação. Este capítulo ainda apresenta a metodologia que será base para as análises e propostas desenvolvidas ao longo desse documento.

O Capítulo 3 apresenta análises realizadas a partir da metodologia de geração de sinal subótimo para estimação de parâmetros. Além disso, o capítulo envolve a apresentação de um algoritmo recursivo usado para contornar o problema em que os parâmetros não são compreendidos pelas limitações predeterminadas.

O Capítulo 4 apresenta a metodologia proposta na tese. Para tanto, será abordada uma adaptação da metodologia original de estimação de parâmetros *Sub-Optimal Excitation*

Signal Generation and Optimal Parameter Estimation que visa refinar a saída do modelo matemático através do acoplamento das características teóricas do modelo matemático, o benefício do sinal de excitação e o refinamento proveniente de uma rede neural profunda. O chamado, modelo híbrido é capaz de determinar saídas mais próximas, de forma percentual, às do sistema real.

O Capítulo 5 é dedicado à verificação de eficiência do método híbrido através da aplicação em uma embarcação autônoma, o ASV AERO4River, um catamarã com características desafiadoras sob o ponto de vista de identificação. Os resultados demonstram que o método é promissor e cumpre com o objetivo de reduzir o erro entre as saídas do modelo e do sistema real.

Já o Capítulo 6 apresenta as principais conclusões e contribuições da utilização da metodologia proposta, bem como propostas de trabalhos futuros.

2 IDENTIFICAÇÃO DE SISTEMAS DINÂMICOS

2.1 INTRODUÇÃO

Métodos de identificação de sistemas dinâmicos possuem suas peculiaridades dependendo da aplicação e forma de implementação. De acordo com Isermann e Münchhof (2010), os modelos matemáticos podem ser classificados em teóricos e experimentais.

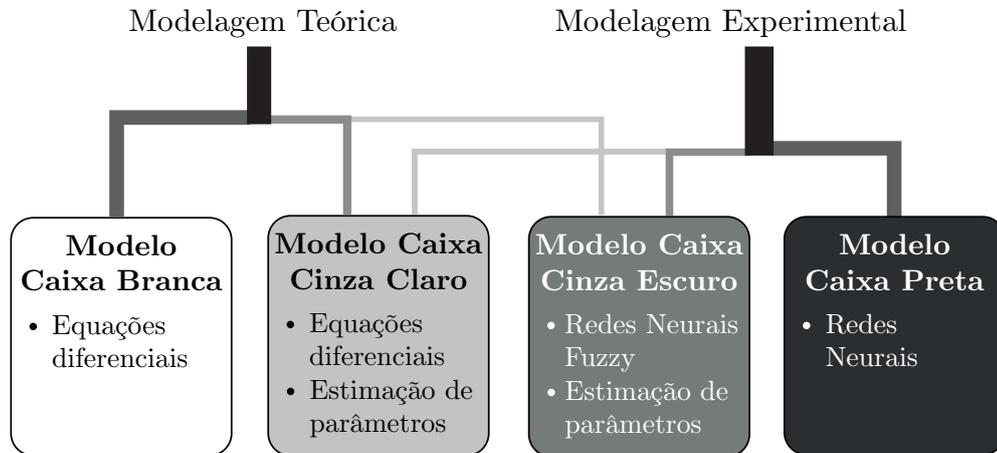
No caso da modelagem teórica, o modelo é obtido através de equações envolvidas no processo provenientes de leis e princípios da física. A partir daí, é possível selecionar um conjunto de equações que geram um comportamento aproximado do sistema. O fator que diferencia sistemas com um mesmo modelo teórico, são os parâmetros associados a cada um. Caso o conjunto de equações diferenciais, possa ser solucionado explicitamente, é possível extrair os parâmetros associados. Entretanto muitos modelos não permitem a resolução explícita, devido à sua complexidade (NELLES, 2002). Em alguns casos, o número de equações envolvidas é inferior ao número de parâmetros, sendo um desafio derivar os parâmetros envolvidos no modelo analiticamente. Uma modelagem puramente teórica leva à construção de um modelo caixa branca. Esse tipo de modelo é utilizado em âmbito de projeto, visto que é necessário determinar seu comportamento, através de simulações de forma antecipada à aplicação do processo .

Alternativamente, em uma modelagem experimental, o modelo matemático é derivado de medições em experimentos, cujos sinais de entrada e saída são condicionados a algum método de identificação. O objetivo é criar uma relação entre entradas e saídas. Isermann e Münchhof (2010) se refere à modelagem experimental como identificação. Esta categoria de técnicas é baseada na comparação de dados provenientes de experimentos realizados com o sistema, podendo lidar com uma excitação arbitrária ou previamente determinada. Uma variedade de aplicações se enquadra nesse tipo de problema, podendo determinar modelos paramétricos ou não paramétricos. Essa abordagem é bastante útil quando se deseja obter um modelo mais fiel à dinâmica real do sistema, extraindo relações de seu comportamento. O modelo resultante é chamado de modelo caixa preta. A aplicabilidade desse tipo de modelo pode ser requerida em situações de adaptação de controladores e detecção de falhas.

Embora sejam obtidas de formas distintas, pode haver uma complementariedade entre as abordagens de modelagem teórica e experimental. Caso seja possível, determinar ambas abordagens, é fundamental que haja uma similaridade entre os comportamentos obtidos através da aplicação dos modelos. Dessa forma, é possível corrigir erros de modelagem, através da comparação dos modelos. Além disso, a associação dos modelos leva a vantagens relacionadas à confiabilidade. A mistura das abordagens resulta em um modelo caixa cinza. A Figura 1 apresenta um diagrama diferenciando os diferentes tipos de modelos matemáticos, de acordo com a classificação quanto a penetração da modelagem

teórica e experimental e a interação entre as diferentes abordagens.

Figura 1 – Tipos de modelos matemáticos



Fonte: Adaptado de Isermann e Münchhof (2010)

A complexidade de alguns modelos teóricos, aliado ao fato de que uma análise puramente matemática pode não prover precisão nos coeficientes relacionados, fazem com que modelos caixa cinza sejam amplamente utilizados. Nesse caso, concilia-se as importantes informações teóricas, dadas por princípios físicos, e dados experimentais que elevam a confiabilidade do modelo.

Como vantagem do uso da modelagem experimental, pode-se citar a possibilidade de aplicação em sistemas complexos. Isso é possível tendo como base o reconhecimento do comportamento através da simples análise das medições de entradas e saídas. O processo de modelagem, portanto, se torna independente de uma investigação profunda do sistema e mais rápido. Métodos computacionais modernos permitem o processamento desses dados e posterior identificação do sistema.

Nesse capítulo, serão apresentados alguns métodos de modelagem experimental mais abordados pela literatura considerando os seguintes aspectos:

- Modelo matemático;
- Algoritmo empregado para processamento dos dados;
- Cálculo do erro entre o processo e o modelo;
- Sinal de teste empregado.

2.2 MODELO MATEMÁTICO

Quanto ao modelo matemático, as técnicas de identificação de sistemas podem ser classificadas em: modelos não paramétricos e modelos paramétricos.

2.2.1 Modelo não paramétrico

Os modelos não paramétricos determinam relações de entrada e saída que não são definidas por uma estrutura paramétrica específica, configurando um número infinito de parâmetros. O produto de uma modelagem paramétrica é uma tabela ou uma curva determinando a relação entre a entrada e a resposta correspondente. As técnicas de identificação não paramétricas pressupõem que o processo possa ser linearizado. Os modelos não paramétricos envolvidos em processos lineares são a resposta ao impulso, resposta ao degrau e resposta à frequência. Essa técnica é utilizada para validação de modelos teóricos e determinação da resposta do sistema para controladores, por exemplo (ISERMANN; MÜNCHHOF, 2011).

2.2.2 Modelo paramétrico

Alternativamente, modelos paramétricos são determinados por uma estrutura de equações com um número de parâmetros finitos. Por assumir, a priori, uma estrutura definida, é esperado, que nesse tipo de modelagem, resultados mais precisos sejam alcançados. Sistemas dinâmicos normalmente podem ser descritos por leis conhecidas da física. O modelo paramétrico, portanto, é visto como um modelo caixa cinza para o processo de identificação, por compreender uma estrutura teórica atrelada a um conhecimento experimental, cujos parâmetros serão obtidos. Nesse caso, a modelagem paramétrica providencia conjuntos de equações diferenciais ou funções de transferência que representem o comportamento dos sistema. Em muitas situações, tais modelos podem ser escritos em forma de equações diferenciais ordinárias, as chamadas equações de estado. Seu modelo é disposto da seguinte forma:

$$\frac{dx(t)}{dt} = f(x(t), u(t); \theta) \quad (2.1)$$

sendo que t representa o tempo, $x(t)$ é o vetor de estado, $u(t)$ é a entrada e θ é o vetor de parâmetros associado ao sistema dinâmico. $f(\cdot)$ é a função que depende do vetor de estado e de entrada e também do vetor de parâmetros. O objetivo principal de um processo de identificação usando um modelo paramétrico é encontrar o vetor de parâmetros θ , ou seja, um conjunto de parâmetros numéricos que determine o comportamento mais semelhante possível ao sistema real, considerando as saídas de um experimento. Esse processo é chamado de Método de Estimação de Parâmetros (MEP).

Tal método é baseado na minimização de sinais de erros entre o comportamento real e o obtido com um certo conjunto de parâmetros. Isermann e Münchhof (2011) separam os métodos de identificação paramétrica em: métodos de estimação de parâmetros; métodos baseados em otimização iterativa; redes neurais e métodos baseados em filtragem. Dessa forma, os MEP são definidos como uma metodologia baseada em regressão estatística

com o objetivo de minimizar uma função de erro. Nessa categoria, podem ser citados os algoritmos baseados em mínimos quadrados e em máxima verossimilhança. Embora a classificação defendida por Isermann e Münchhof (2011) seja válida, com propósito de simplificar o entendimento, nesta tese, um MEP poderá englobar todos os métodos aplicados para identificação paramétrica de sistemas dinâmicos, conforme é conceituado pela obra de Baillieul e Samad (2015).

2.3 MÉTODOS DE ESTIMAÇÃO DE PARÂMETROS

2.3.1 Método dos Mínimos Quadrados

Esse método é utilizado na estimação de parâmetros de sistemas dinâmicos, através das medições das saídas provenientes de um experimento. A estimação paramétrica prevê uma aproximação entre o comportamento de um sistema real com o de seu respectivo modelo, para entradas similares. A solução desse tipo de problema através de método baseado em mínimos quadrados utiliza literalmente essa definição.

A primeira citação do método de mínimos quadrados como um estimador ocorreu em 1795 por Karl Friedrich Gauss (GAUSS, 1809). Nessa obra, voltada para estudos astronômicos, Gauss determinou que o movimento dos corpos pode ser descrito por seis parâmetros, cuja estimação seria baseada em medições de dados a partir de um telescópio. Surgiu então, o método dos mínimos quadrados, do inglês *Least Squares Method*. A partir daí, várias outras áreas começaram a aplicar esse conhecimento a fim de identificar sistemas em diferentes âmbitos da engenharia, como, por exemplo, em processos elétricos, mecatrônicos e veículos em geral.

Dado um sistema real discreto com saídas $y_u(k)$, o conjunto de parâmetros que o descreve é

$$\boldsymbol{\theta}^T = (\theta_1 \ \theta_2 \ \dots \ \theta_m) \quad (2.2)$$

sendo que m é o número de parâmetros e k o tempo amostrado. A identificação paramétrica terá o compromisso de determinar um conjunto de parâmetros o mais próximo possível desses valores. A saída do sistema é afetada por distúrbios. Dessa forma, a saída medida é dada por

$$y_P(k) = y_u(k) + n(k) \quad (2.3)$$

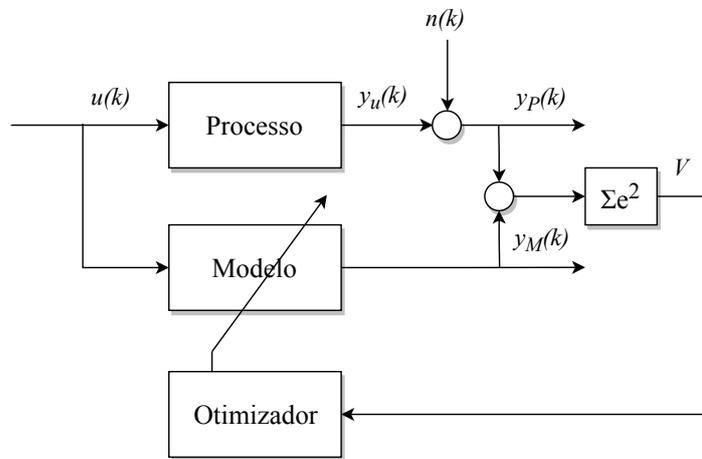
sendo que $n(k)$ representa o ruído medido. O processo segue princípios e leis físicas dependentes dos parâmetros. E dessa forma, é possível determinar um modelo, cujas saídas, simbolizadas por $y_M(k)$, simulam o comportamento do sistema real. O erro entre o sistema real e as saídas do modelo é definido por

$$e(k) = y_P - y_M . \quad (2.4)$$

O melhor dos cenários é atribuído quando esse erro é o mínimo possível. Nesse caso, o objetivo principal é minimizar a soma dos erros instantâneos ao quadrado (Figura 2), conforme a seguinte equação

$$V = e^2(1) + e^2(2) + \dots + e^2(N) = \sum_{k=1}^N (e(k))^2 . \quad (2.5)$$

Figura 2 – Diagrama esquemático do processo geral do método de estimação de parâmetros baseado em mínimos quadrados



Fonte: Adaptado de Isermann e Münchhof (2010)

A principal justificativa para o uso do erro elevado à segunda potência é devido ao fato de que é mais fácil de encontrar a solução ótima para esse tipo de função do que com relação à função de erro absoluta (ISERMANN; MÜNCHHOF, 2011). Todavia deve ser pontuado o fato de que um critério quadrático pode enfatizar erros grosseiros, os chamados *outliers*. Assim, o efeito desses *outliers* no processo de estimação pode ser aumentado, determinando parâmetros incorretos.

A simplicidade do método, aliado ao fato de demandar um pequeno custo computacional frente a outros, faz com que muitos autores utilizem o método dos mínimos quadrados em seus trabalhos. É importante mencionar que a maior parte de teorias envolvendo métodos de estimativa são modificações ou desmembramentos dessa estratégia (CRASSIDIS; JUNKINS, 2011). Isso é devido ao objetivo principal do problema, que é atendido através do uso dessa função. Dessa forma, os métodos de otimização e redes neurais, aplicados na identificação, geralmente, apresentam esse propósito, o de reduzir o erro quadrático médio entre o observado e o estimado.

2.3.2 Métodos de Filtragem

Métodos de filtragem aplicados à identificação de sistemas tem o propósito principal de estimar os estados instantâneos de um sistema baseando em medidas dos sinais de entrada e saída. Existem vários filtros utilizados na literatura, dentre estes destaca-se o Filtro de Kalman (GELB, 1974), que é um algoritmo recorrentemente utilizado em problemas de estimação linear. A sua principal vantagem é a facilidade de implementação e a viabilidade de aplicações em tempo real. De forma geral, o algoritmo computa a melhor estimativa de um vetor de estados de um sistema dinâmico linear dadas as medidas do vetor de estados adicionado de ruídos gaussianos. Para tanto, os parâmetros devem ser conhecidos.

Embora esse método de filtragem não tenha como objetivo primordial ser um estimador ótimo, é possível empregar a versão extendida do Filtro de Kalman (GELB, 1974) para estimação de parâmetros de forma concomitante à estimação de estados. Para realizar essa tarefa é necessário que os estados sejam modelados e influenciados por um distúrbio estocástico. Apesar de contar como vantagem essa possibilidade, o gasto computacional se torna elevado, uma vez que o processo de estimação conta com mais variáveis.

A complexidade computacional de um Filtro de Kalman Extendido é em torno do cubo da dimensão do vetor de estados (BAILLIEUL; SAMAD, 2015). No caso da adição de parâmetros a serem estimados, a complexidade envolve, ainda, o vetor de parâmetros. Em situações que não exigem uma taxa alta de amostragem, é possível aplicar a ferramenta para estimação de parâmetros. É o caso de uma grande embarcação, por exemplo, cujo controle não exige respostas extremamente rápidas. Revestido et al. (2011) aplicou o Filtro de Kalman Extendido juntamente com uma técnica de predição de erros para estimar estados e parâmetros relacionado a um navio-patrolha. Aliada a possibilidade de não ter um sistema de estimação extremamente rápido, os autores consideraram apenas os derivativos hidrodinâmicos como variáveis a serem estimadas, no caso, oito parâmetros. Quadrotores, por exemplo, requerem uma maior taxa de amostragem e, nesse caso, deve-se pensar em técnicas com respostas rápidas.

Abas, Legowo e Akmeliawati (2011) utilizaram Filtro de Kalman modificado a fim de estimar estados e parâmetros de um quadrotor. Entretanto, considerou-se apenas quatro parâmetros relacionados ao momento de inércia, exigindo um esforço computacional reduzido. Além de permitir trabalhar com tal dupla estimação, a ferramenta permite lidar com não linearidades. Entretanto, durante sua implementação, deve-se realizar uma linearização em torno do ponto de operação.

2.3.3 Métodos de Otimização

No campo de estimação de parâmetros, algoritmos iterativos de otimização são utilizados permitindo a minimização de funções de custo previamente determinadas. A principal vantagem desse método é a possibilidade de trabalhar facilmente com sistemas, funções objetivos e restrições não lineares. Dessa forma métodos iterativos garantem que sejam projetadas funções objetivos eficientes e condizentes com o processo de estimação. Além disso restrições podem ser associadas com o objetivo de restringir a escolha dos parâmetros às limitações físicas do sistema (HONÓRIO et al., 2018).

A função de custo possui o propósito principal de reduzir o erro que o parâmetro proporciona no comportamento do modelo em relação à saída real. Como, para esse tipo de método, existe uma certa liberdade de determinação da função objetivo, não é obrigatória a escolha da função de mínimos quadrados para o processo de estimação. Como já mencionado, o método de mínimos quadrados proporciona efeitos prejudiciais na presença de erros grosseiros. Dessa forma, outras funções podem ser indicadas por reduzir o efeitos de *outliers*, conforme pode ser observado pela Tabela 1.

Nome	Função Objetivo
Mínimos Quadrados	e^2
Huber	$\begin{cases} e^2/2; & e \leq c \\ c(e - c/2); & e > c \end{cases}$
Biquadrada	$\begin{cases} c^2/6(1 - (1 - (e/c)^2)^3); & e \leq c \\ c^2/6; & e > c \end{cases}$
L1-L2	$2(\sqrt{1 + e^2/2} + 1)$
Valor Absoluto	$ e $

Tabela 1 – Possibilidades de funções objetivo para o processo de estimação paramétrica através de métodos iterativos, sendo que c é uma constante definida previamente.

Aliado ao fato de facilidade de determinação da função objetivo, processos iterativos possuem como grande vantagem a imposição de restrições. Embora em muitos casos a função objetivo seja minimizada de forma satisfatória, caso a modelagem matemática não seja identificada corretamente, o conjunto de parâmetros pode configurar uma instabilidade no modelo, o que não é desejado para a modelagem final. Além disso, parâmetros como massa ou rigidez, por exemplo possuem limitações devidos a razões físicas. Uma vantagem da inserção de restrições é a redução no espaço de busca, facilitando a estimação dos parâmetros.

Diversos métodos de otimização podem ser aplicados nesse intuito. Métodos clássicos normalmente exigem a determinação de parâmetros iniciais para inicialização do algoritmo. Tais parâmetros podem ser obtidos por métodos não recursivos como o método dos mínimos quadrados. Como são métodos baseados em derivada, dependendo da função objetivo, podem facilmente ficar preso em ótimos locais, não refletindo a realidade.

Portanto, sua eficiência está diretamente associada ao ponto inicial indicado. A partir desse ponto de partida, os métodos usam diferentes abordagens para chegar na melhor associação de parâmetros. É o caso do Método de Pontos Interiores (PÓLIK; TERLAKY, 2010), que será amplamente utilizado nos estudos de caso dessa tese no âmbito da estimação paramétrica.

Uma variedade de trabalhos apresentados na literatura especializada atuam na área de estimação de parâmetros utilizando, como ferramenta, métodos iterativos. Para obter uma compreensão mais abrangente do modelo matemático de um Veículo Autônomo de Superfície (VAS), os autores, Erunsal et al. (2017), elaboraram uma rotina de otimização para minimizar os custos entre o movimento do veículo com os parâmetros previstos e os dados experimentais. Para o teste em questão e com parâmetros a priori bem definidos, foi possível validar a metodologia.

O método de pontos interiores foi utilizado por Honório et al. (2018) para estimar parâmetros abordando uma técnica que envolve o projeto dos sinais de entrada. Embora esse tipo de método permita a manipulação dos resultados em domínio do tempo, é possível combiná-la com resultados no domínio da frequência com objetivo de atingir maior acurácia de estimação. A vantagem de métodos de programação linear é a possibilidade de envolver esse tipo de análise sem modificações na forma de resolução. Xu (2015) abordou o problema dessa forma, utilizando o método de Newton iterativo. De forma a reduzir a ação de ruído no processo, também é possível associar técnicas de filtragem ao método de otimização (LI; LIU; DING, 2017).

Além de métodos clássicos, podem ser aplicados métodos heurísticos ou bioinspirados. Nesse tipo de algoritmo, não há necessidade de configurar um ponto inicial. De acordo com uma certa regra, o algoritmo realiza a exploração de possíveis soluções dentro do espaço de busca. É importante ressaltar, que neste caso, apesar de não abordar o problema através de gradiente, não é garantido que o ótimo global seja encontrado. Entretanto, métodos exploratórios podem ser uma excelente maneira de encontrar um ponto de partida para um método clássico. Um exemplo desse tipo de técnica é o algoritmo de otimização baseado em enxame de partículas. O algoritmo foi proposto por Eberhart e Kennedy (1995) e seleciona a solução ótima de acordo com resultados de velocidade e posição associados a cada partícula. A partir de um conjunto de soluções candidatas, ao longo das iterações o algoritmo direciona a solução àquela que possui maior função objetivo associada. Luo, Soares e Zou (2016) relaciona o método exploratório, *Particle Swarm Optimization* (PSO) com a metodologia de máquina de vetores de suporte, do inglês, *Support Vector Machine* (SVM), a fim de identificar o modelo de uma embarcação, mais especificadamente os coeficientes hidrodinâmicos. A principal vantagem associada ao SVM é que ele não requer uma estimativa inicial dos coeficientes. Nessa aplicação, o PSO é aplicado de forma a estimar parâmetros estruturais da embarcação. Posteriormente, o

método SVM é aplicado a fim de estimar os coeficientes hidrodinâmicos que são variáveis de acordo com a manobra realizada.

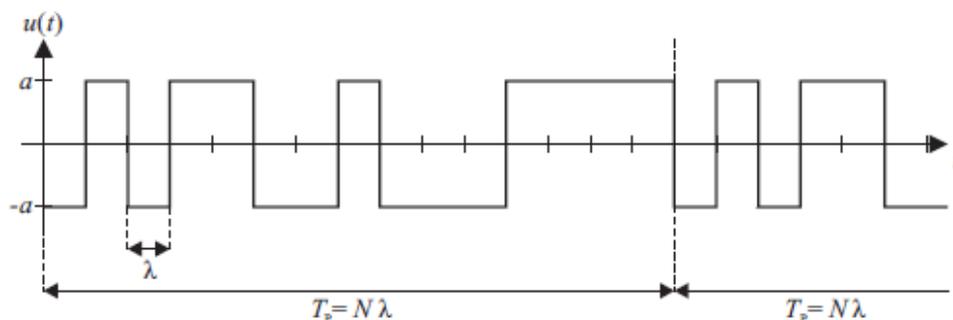
2.4 SINAL DE EXCITAÇÃO

O sinal de excitação do sistema dinâmico, com a finalidade de identificação, pode ser aleatório ou previamente determinado. A utilização de um sinal conhecido pode ser justificada pela falta de excitação necessária de um sinal de operação normal para a extração de parâmetros ou mesmo identificação de um modelo caixa preta. Fato que é extremamente comum, já que numa operação normal, os sinais aplicados podem apresentar baixas amplitudes ou características não estacionárias, por exemplo. De acordo com Isermann e Münchhof (2010), um sinal considerado favorável ao processo de identificação, normalmente satisfaz os seguintes critérios:

- Geração simples e reproduzível;
- Descrição matemática simples;
- Viabilidade de aplicação, considerando limitações físicas do sistema;
- Boa excitação das características dinâmicas relevantes.

Seguindo esses princípios básicos o sinal binário pseudo-aleatório, do inglês, *Pseudo-Random Binary Signal* (PRBS) é um exemplo de sinal de teste amplamente utilizado. Além das características descritas, esse tipo de sinal permite a identificação através de respostas ao impulso diretamente. De forma geral, o sinal é aplicado com uma periodicidade relacionada à sequência de pulsos aleatórios transmitidos através da atuação. A Figura 3 exemplifica a característica gráfica do sinal.

Figura 3 – Sinal binário pseudo-aleatório com 4 estágios e $N = 15$

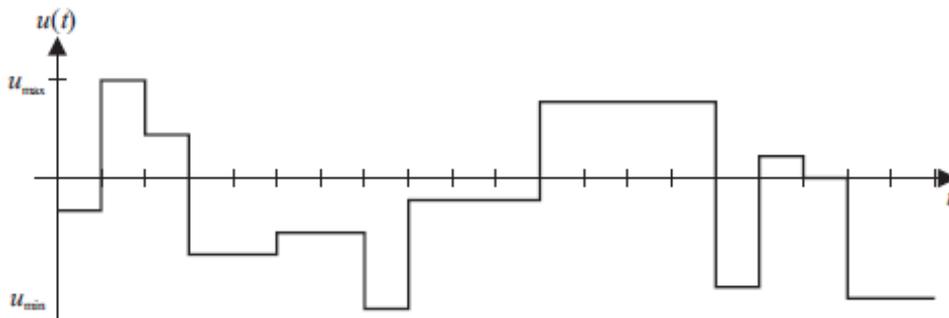


Fonte: Isermann e Münchhof (2010)

Embora seja um método altamente recomendável para excitação de sistemas lineares, o PRBS não é aconselhável para a categoria não linear. Devido à excitação

ser dependente de apenas um par de magnitudes, ou seja o sinal de teste só consegue assumir dois valores ($u(t) = -a$ ou $u(t) = a$), a extração da dinâmica do sistema torna-se deficitária. Assim, a combinação de frequência variável com amplitude variável ao longo do tempo pode ser interessante do ponto de vista da identificação. Além da definição do tempo total do sinal e do tempo dos pulsos, determina-se também a amplitude de cada pulso. Esse é o princípio utilizado pelo sinal binário pseudo-aleatório com amplitude modulada, do inglês, *Amplitude-modulated Pseudo-Random Binary Signal* (APRBS), cujo comportamento gráfico é apresentado pela Figura 4. Devido à variedade de amplitudes do sinal de entrada, é possível excitar o sistema de forma satisfatória, fato que é altamente requerido para processos de identificação.

Figura 4 – Sinal binário pseudo-aleatório com amplitude modulada



Fonte: Isermann e Münchhof (2010)

A prévia determinação do sinal de teste é totalmente relacionada ao grau de confiabilidade da identificação dinâmica de um sistema. Deflorian e Zaglauer (2011) aplicou sinais de teste baseados em APRBS para identificar a dinâmica de um motor à combustão. Em todos os testes, foi possível determinar modelos com dinâmicas similares à do processo real.

Um estudo apresentado por Gringard e Kroll (2016) analisou o impacto do nível de amplitude do sinal de entrada na qualidade da estimação de parâmetros. Utilizando como base para os experimentos um acelerador mecânico, os autores conseguiram extrair algumas informações sobre o uso de um sinal baseado em APRBS:

- i) O nível de amplitude utilizado impacta diretamente na excitação do sistema, ou seja, dependendo do sinal APRBS utilizado, é possível estimar melhor ou pior os parâmetros do sistema;
- ii) É importante, no design do sinal de entrada, escolher valores próximos aos fornecidos em uma operação normal do sistema.

Conforme foi observado no estudo mencionado, é possível extrair diferentes informações do sistema dinâmico, utilizando sinais de entrada diferentes. Embora a utilização de sinais de teste baseados em APRBS seja consagrada como uma boa alternativa, a

manipulação desse sinal pode ser mais benéfica ao processo de estimação de parâmetros. Baseado nessa ideia, métodos de otimização podem ser aplicados na escolha do sinal de teste.

De forma a obter dados experimentais significativos, Licitra et al. (2018) determinaram manobras ótimas capazes de superar os resultados de entradas de teste convencionais. Algoritmos heurísticos também podem ser utilizados. É o caso de Lichota, Sibilski e Ohme (2017), que apresenta uma técnica baseada em algoritmos genéticos para projetar uma manobra de excitação de identificação de sistema de aeronave com deflexões simultâneas de controle de voo.

2.5 SOESGOPE

Em algumas aplicações, considerar restrições é extremamente importante. Tais limitações podem ser relacionadas principalmente às limitações físicas do sistema e às limitações experimentais. Com o objetivo de poder considerar limitações no processo de escolha do sinal de teste, Honório et al. (2018) desenvolveram uma técnica de design de entrada, a Geração de Sinal de Excitação Subótimo e Estimação Ótima de Parâmetros, do inglês SOESGOPE. O princípio básico desse método é realizar a escolha do sinal através de um algoritmo de otimização que consiga desviar de ótimos locais, uma vez que existem um conjunto de possibilidades de atuações que atendem as restrições e determinam um valor de função objetivo satisfatória.

O método SOESGOPE será descrito em detalhes tendo em vista que as principais contribuições do trabalho visam o seu aprofundamento e a proposição de adaptação frente às lacunas observadas em sua aplicação.

Dado o sistema real $\mathcal{R}(\mathbf{\Gamma})$, o objetivo primário é identificar a dinâmica do sistema a fim de obter um modelo $\mathcal{M}(\mathbf{\Gamma})$ que tenha o comportamento mais similar possível ao real. É importante ressaltar que os parâmetros do modelo real são desconhecidos e podem ser descritos, de uma forma puramente teórica, por $\mathbf{\Gamma} \in \mathbb{R}^r \mid \mathbf{\Gamma} = [\gamma_1, \gamma_2, \dots, \gamma_r]^T$. Embora os parâmetros sejam desconhecidos, é importante que haja uma estimativa prévia desses parâmetros, chamada de estimativa a priori, $\hat{\mathbf{\Gamma}}^-$, que pode ser extraída através de análises matemáticas, ou mesmo por um experimento simples prévio. No caso de veículos autônomos, por exemplo, é possível aplicar movimentos triviais, tais como deslocamentos em linha reta horizontalmente e verticalmente, em zigue-zague, em espiral ou circulares. A partir de experimentos como esses uma estimativa inicial dos parâmetros envolvidos no processo pode ser obtida.

Apesar de estratégia de deslocamento pré-determinadas serem amplamente utilizadas pela literatura para extração do comportamento dinâmico de veículos (FOSSEN, 1994; NAEEM; SUTTON; CHUDLEY, 2006; WIRTENSOHN et al., 2013; WIRTENSOHN et

al., 2015), tais sinais de entrada podem não ser suficientes para excitar o sistema de uma forma generalista. Com um movimento de linha reta seguindo o eixo vertical, a título de exemplo, parâmetros relacionados a movimentos em deriva serão negligenciados, podendo não configurar um modelo razoável. Dessa forma, o principal objetivo do método é gerar um sinal que excite o sistema da melhor forma possível, que será simbolizado por \mathbf{u}^\oplus . O sinal \mathbf{u}^\oplus deve ser persistente o suficiente para conseguir extrair o conjunto de parâmetros relacionados, que será chamado de estimativa a posteriori $\hat{\Gamma}^+$. O modelo deve ser capaz de aproximar o comportamento do sistema real ($\mathcal{M}(\hat{\Gamma}^+) \approx \mathcal{R}(\Gamma)$). Além desse objetivo, o sinal deve atender a outros requisitos e propósitos.

Uma vez que o método trabalha com sistemas cujos parâmetros são desconhecidos, a garantia de um sinal com persistência suficiente para encontrar os parâmetros exatos é impraticável, mesmo em situações cujos modelos produzam o mesmo comportamento dinâmico do sistema real. De forma a determinar o sinal mais adequado não se utiliza, portanto, os parâmetros reais, e sim um outro conjunto de parâmetros definido a partir da perturbação da estimativa a priori, $\tilde{\Gamma}^p$. Esse novo sistema obtido, $\mathcal{M}(\tilde{\Gamma}^p)$, será denominado como perturbado e possui como objetivo ser uma orientação na escolha do sinal de entrada para o experimento. No processo de escolha do sinal, $\tilde{\Gamma}^p$ é usado como emulação do sistema real gerado da seguinte forma:

$$\tilde{\Gamma}^p = \hat{\Gamma}^- + \tilde{\delta}^p \quad (2.6)$$

sendo que $\tilde{\delta}^p \in \mathbb{R}^r \mid \tilde{\delta}^p = [\tilde{\delta}_1^p, \tilde{\delta}_2^p, \dots, \tilde{\delta}_r^p]^T$ é um conjunto de perturbações gaussianas com média zero e variância q que é determinado empiricamente e proporcionalmente ao inverso do conhecimento do parâmetro correspondente.

A partir de $\mathcal{M}(\tilde{\Gamma}^p)$, o objetivo é, dado o sinal \mathbf{u}^\oplus , encontrar $\mathcal{M}(\hat{\Gamma}^+)$ em que $\hat{\Gamma}^+$ é a estimativa a posteriori que minimiza

$$\Sigma_\varepsilon := \sum_{k=0}^{t_u} (\|\hat{\mathbf{x}}^+(k) - \tilde{\mathbf{x}}^p(k)\| + \|\hat{\mathbf{y}}^+(k) - \tilde{\mathbf{y}}^p(k)\|) \quad (2.7)$$

sendo que:

$\tilde{\mathbf{x}}^p, \tilde{\mathbf{y}}^p$ são os estados e saídas do modelo perturbado $\mathcal{M}(\tilde{\Gamma}^p)$;

$\hat{\mathbf{x}}^+, \hat{\mathbf{y}}^+$ são os estados e saídas do modelo a posteriori $\mathcal{M}(\hat{\Gamma}^+)$.

A métrica $\Sigma_\varepsilon \in \mathbb{R}_{\geq 0}$ quantifica a coincidência entre as saídas do modelo perturbado e estimado. É determinante mencionar que por mais que $\Sigma_\varepsilon = 0$, não é possível garantir que os parâmetros foram corretamente estimados. Entretanto, com esse resultado, pode-se concluir que o sinal tem persistência de excitação suficiente para parâmetros que reproduzam o mesmo comportamento.

Dessa forma, a tarefa de estimação deve levar em consideração tanto a determinação dos parâmetros quanto a geração de um sinal de excitação subótimo, configurando um problema de otimização sequencial multi-paramétrica. Para atingir resultados satisfatórios, o problema pode ser dividido em duas camadas com funções objetivo diferentes e pertinentes em cada caso. O processo de otimização deve abranger tanto variáveis discretas, quanto variáveis contínuas. A camada externa, responsável pela geração do sinal, consiste em um problema de Programação Não Linear Inteira Mista (PNLIM). Já a camada interna, cuja finalidade é a estimação de parâmetros propriamente dita, configura-se como um problema de natureza não linear, chamado de Programação Não Linear (PNL).

2.5.1 Camada Externa

O algoritmo de otimização da camada externa é responsável por atingir o objetivo final do processo, ou seja, determinar o sinal de entrada que possua boas características de excitação. Embora o propósito principal seja este, tal sinal deve obedecer algumas restrições impostas por limitações físicas e temporais relacionadas ao experimento. O processo de otimização, portanto, trabalha com a minimização de uma função multiparamétrica $f(\Xi) \in \mathbb{R}_{\geq 0}$ para a escolha do sinal, que é dada por:

$$f(\Xi) = \min_{\Xi} (\kappa_1 \Theta(\cdot|\cdot, \Xi) + \kappa_2 f_o(\cdot|\Theta) + \kappa_3 f_{\hat{\delta}}(\cdot|f_o)) \quad (2.8)$$

sendo que:

$\Theta(\cdot)$ Parcela relacionada aos limites associados aos estados;

$f_o(\cdot)$ Parcela relacionada à acurácia da saída;

$f_{\hat{\delta}}(\cdot)$ Parcela relacionadas à recuperabilidade dos parâmetros;

$\kappa_1, \kappa_2, \kappa_3 \in \mathbb{R}_{\geq 0}$ Constantes de peso relacionadas às parcelas da função objetivo $\Theta(\cdot)$, $f_o(\cdot)$ e $f_{\hat{\delta}}(\cdot)$, respectivamente.

2.5.1.1 Função de Tempo e Estados

Essa parcela da função objetivo é responsável por manter as saídas e os estados dentro de um limite permissível. Para evitar violações de restrições impostas, adota-se que $\kappa_1 \gg \kappa_2, \kappa_3$. Nessa parcela, inclui-se, adicionalmente, quesitos temporais e espaciais com o propósito de tornar o experimento mais rápido e com menos deslocamentos possíveis. Dessa forma, a escolha do sinal é realizada considerando o menor tempo t_u e espaço f_V de experimento. Entretanto, essa otimização do tempo e do espaço somente ocorrerá caso as restrições impostas sejam satisfeitas, como pode-se observar na seguinte expressão:

$$\Theta(\tilde{\mathbf{x}}^p, \tilde{\mathbf{y}}^p, t_u|u) = \begin{cases} \frac{1}{\kappa_1} t_u + \frac{\kappa_t}{\kappa_1} f_V(\tilde{\mathbf{x}}^p) + \frac{\kappa_t}{\kappa_1} f_V(\tilde{\mathbf{y}}^p), & \text{if } \sum \Psi(\cdot) \leq \alpha \\ \Psi(\tilde{\mathbf{x}}^p) + \Psi(\tilde{\mathbf{y}}^p) + \Psi(t_u) + 1, & \text{otherwise} \end{cases} \quad (2.9)$$

onde:

$$\begin{aligned} \Psi(t_u) &= |t_u - T_{\max}|, & \text{if } t_u > T_{\max} \\ \Psi(t_u) &= 0, & \text{if } t_u \leq T_{\max} \\ \Psi(\tilde{\mathbf{x}}^p) &= \begin{cases} \sum_{k=0}^{t_u} \sum_{i=1}^n |\tilde{x}_{p,i}(k) - \mathcal{O}x_{i,\max}|, & \forall \tilde{x}_{p,i}(k) > \mathcal{O}x_{i,\max} \\ + \\ \sum_{k=0}^{t_u} \sum_{i=1}^n |\tilde{x}_{p,i}(k) - \mathcal{O}x_{i,\min}|, & \forall \tilde{x}_{p,i}(k) < \mathcal{O}x_{i,\min} \\ + \\ 0, & \mathcal{O}x_{i,\min} \leq \forall \tilde{x}_{p,i}(k) \leq \mathcal{O}x_{i,\max} \end{cases} \end{aligned} \quad (2.10)$$

sendo que:

t_u Tempo do experimento;

$f_V(\cdot)$ Função que quantifica o uso do espaço;

$\Psi(\cdot)$ Função de penalidade;

α Índice de relaxamento;

$\mathcal{O}x_{\min}$ Limite operacional mínimo das variáveis de estado;

$\mathcal{O}x_{\max}$ Limite operacional máximo das variáveis de estado.

Esse tipo de função é baseada em α -*constrain* (HONORIO et al., 2010), cuja função é limitar $\Theta(\cdot)$ de acordo com o resultado da função de penalidade. O índice de relaxamento inicia com um certo valor e ao longo das iterações tem seu valor reduzido até 0 ao fim do processo de otimização.

A função de quantificação das saídas, f_V , é calculada pelo volume da *Axis-Aligned Bounding Box* (AABB) hiper-dimensional (KLOSOWSKI et al., 1998), definido como:

$$\mathbf{f}_V(\tilde{\mathbf{x}}^p) = (\bar{\mathbf{x}} - \underline{\mathbf{x}})^T \mathbf{W} (\bar{\mathbf{x}} - \underline{\mathbf{x}}) \quad (2.11)$$

sendo que a matriz de pesos diagonal $\mathbf{W} \in \mathbb{R}^{n \times n} \mid \mathbf{W} \geq 0$ é definida como $\mathbf{W} = \text{diag}(w_1, w_2, \dots, w_n)$ e $\bar{\mathbf{x}}$ e $\underline{\mathbf{x}}$ são dados por:

$$\begin{aligned} \bar{\mathbf{x}} &= [\max(\tilde{x}_1^p(k)), \max(\tilde{x}_2^p(k)), \dots, \max(\tilde{x}_n^p(k))]^T \\ \underline{\mathbf{x}} &= [\min(\tilde{x}_1^p(k)), \min(\tilde{x}_2^p(k)), \dots, \min(\tilde{x}_n^p(k))]^T \end{aligned} \quad (2.12)$$

2.5.1.2 Função de Precisão das Saídas

A função de acurácia das saídas é responsável por determinar um sinal que favoreça a coincidência entre as saídas do modelo perturbado e do modelo a posteriori. Em outras palavras, essa função mensura a diferença entre as saídas do sistema considerado real e o estimado, conforme Equação 2.7. O valor final de $f_o(\cdot)$ é calculado após a determinação dos parâmetros estimados no processo de otimização interna, que será explicado posteriormente.

2.5.1.3 Função de Recuperabilidade de Parâmetros

Embora a similaridade entre saídas seja essencial para determinação do sinal de entrada é importante mencionar que saídas iguais podem não ser um parâmetro exclusivo para estimação dos parâmetros. Um sinal pode determinar saídas iguais, porém parâmetros completamente diferentes, fato que configuraria uma má escolha da entrada. Por isso, torna-se essencial a análise da recuperabilidade dos parâmetros. Como na camada externa, consegue-se extrair os parâmetros do sistema perturbado, é fácil saber se os parâmetros estimados foram satisfatórios. Dessa forma, percebe-se que um sinal é rico (com relação à excitação do sistema) caso os parâmetros $\hat{\Gamma}^+$ convirjam para $\tilde{\Gamma}^p$. Para atingir tal feito, é necessário minimizar a função de recuperabilidade, dada por:

$$f_{\delta}(\hat{\Gamma}^+ | f_o) = \sum_{i=1}^r \left(\frac{\hat{\gamma}_i^+ - \tilde{\gamma}_i^p}{\hat{\gamma}_i^-} \right)^2 \quad (2.13)$$

sendo que cada $\hat{\gamma}$ é um elemento no vetor de parâmetros $\mathbf{\Gamma} = [\hat{\gamma}_1, \hat{\gamma}_2, \dots, \hat{\gamma}_r]^T$. A normalização é necessária para manter cada medida dos parâmetros em sua própria escala.

2.5.1.4 Formulação do problema de otimização

A partir da formulação do problema de otimização supracitada (Equação 2.8), deve-se definir a forma de resolução. Embora o problema tenha restrições associadas, tais limitações não entram como restrições propriamente ditas, mas sim como penalidades caso o sinal escolhido configure saídas que extrapolem os limites físicos e temporais impostos. Portanto o problema é caracterizado como irrestrito. O objetivo final é encontrar o sinal de entrada, u , que minimize a função objetivo $f(\Xi) \in \mathbb{R}^{2pk}$.

Conforme analisado na seção 2.4, o sinal APRBS é uma boa alternativa para utilização em um experimento dedicado à identificação. De forma a representar matematicamente o sinal u como APRBS, utiliza-se o vetor Ξ , que compõe os p componentes de u divididos em k estágios da seguinte forma:

$$\Xi := \begin{bmatrix} t_1^{u_1}, & \dots, & t_k^{u_1}, & t_1^{u_2}, & \dots, & t_k^{u_2}, & \dots, & t_1^{u_p}, & \dots, & t_k^{u_p}, \\ a_1^{u_1}, & \dots, & a_k^{u_1}, & a_1^{u_2}, & \dots, & a_k^{u_2}, & \dots, & a_1^{u_p}, & \dots, & a_k^{u_p} \end{bmatrix} \quad (2.14)$$

onde:

$$\begin{aligned} (t_{\min})_i^{u_j} \leq \Xi_d = t_i^{u_j} \leq (t_{\max})_i^{u_j}, \quad 1 \leq d = (j-1) * k + i \leq pk \\ (a_{\min})_i^{u_j} \leq \Xi_d = a_i^{u_j} \leq (a_{\max})_i^{u_j}, \quad pk < d = (j-1) * k + i + pk \leq 2pk \end{aligned} \quad (2.15)$$

sendo que os seguintes componentes representam:

$t_k^{u_p}$	Tempo de execução do estágio k para o sinal de entrada u_p ;
$t_k^{u_p}$	Amplitude do sinal no estágio k para o sinal de entrada u_p ;
$(t_{\min})_i^{u_j}, (t_{\max})_i^{u_j}$	Tempos mínimo e máximo de execução do estágio i para o sinal de entrada j , respectivamente, com $1 \leq i \leq k$ e $1 \leq j \leq p$;
$(a_{\min})_i^{u_j}, (a_{\max})_i^{u_j}$	Amplitudes mínima e máxima do sinal de entrada j no estágio i , respectivamente, com $1 \leq i \leq k$ e $1 \leq j \leq p$;

Por motivos de simplificação e coerência, o tempo de execução do experimento é o mesmo para todos os componentes de entrada.

2.5.1.5 Algoritmo de Resolução

Devido às infinitas possibilidades de combinações de tempo e amplitude dos estágios de cada sinal de entrada, demanda-se um algoritmo de otimização que consiga obter a minimização da função, explorando diferentes áreas de procura. O uso de algoritmos heurísticos e bio-inspirados é uma opção viável para o processo de procura. Embora tais algoritmos sejam eficientes na determinação de um sinal viável, atendendo às restrições, não se pode garantir que o resultado encontrado seja o ótimo global da função. Isso porque, problemas como esse apresentam altas não-linearidades configurando inúmeros ótimos locais. A opção utilizada para composição do sinal de entrada no algoritmo original (HONÓRIO et al., 2018) é o método baseado no Algoritmo de Enxame de Partículas (PSO), o *Particle Swarm-Like Optimization* (PSLO), proposto por Honório et al. (2018).

2.5.2 Camada Interna

A camada interna é responsável pela estimação de parâmetros propriamente dita, a partir do sinal de entrada fornecido pela camada externa. Nessa etapa consegue-se extrair o conjunto de parâmetros de um modelo, cujas saídas coincida melhor com as saídas do sistema perturbado. A partir dos parâmetros estimados, é possível calcular as métricas relacionadas à precisão das saídas, f_o , e à recuperabilidade de parâmetros, f_δ . Dessa forma, a camada interna é acionada sempre quando o algoritmo da camada externa calcula o valor dessas parcelas da função objetivo.

É importante ressaltar, que na camada interna, um novo processo de otimização é caracterizado. Dessa forma, a formulação do problema de otimização é diferente. O objetivo principal, nesse caso, é simplesmente determinar parâmetros que mais coincidam com as saídas performadas pelo sistema perturbado. Assim, o problema, nesse caso, possui apenas uma métrica a ser minimizada conforme a fórmula a seguir:

$$\text{Minimizar } f_o := \sum_{k=0}^{t_u} (\|\hat{\mathbf{x}}^+(k) - \tilde{\mathbf{x}}^P(k)\| + \|\hat{\mathbf{y}}^+(k) - \tilde{\mathbf{y}}^P(k)\|) \quad (2.16)$$

Como, nesse caso, é possível obter uma estimativa a priori dos parâmetros, pode-se utilizar algoritmos que tenham a característica de demandar uma estimativa inicial. O Método de Pontos Interiores é um algoritmo eficiente de solução de problemas de otimização não linear é implementado pela *toolbox* denominada *fmincon* do programa MATLAB, sendo essa ferramenta utilizada para a realização da etapa de otimização vislumbrada a partir da formulação proposta para coordenação dos movimentos de múltiplos robôs. A *fmincon* utiliza o método de pontos interiores valendo-se do método Barreira Logarítmica (BYRD; HRIBAR; NOCEDAL, 1999), assim como é explicitado na própria página do programa (MATHWORKS, 2020).

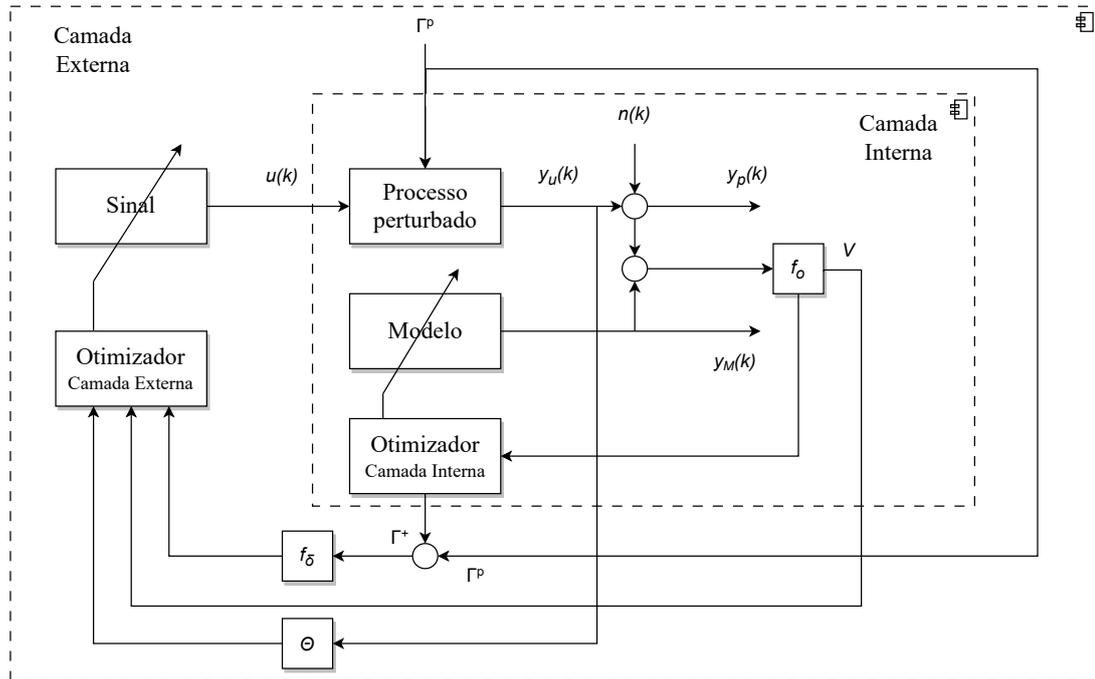
2.5.3 Estrutura SOESGOPE

O método SOESGOPE objetiva a determinação de um sinal de excitação que seja o mais eficiente possível para a estimação dos parâmetros envolvidos no modelo. A Figura 5 ilustra o diagrama esquemático do processo, que é dividido em duas camadas:

(i) A externa, responsável pela determinação do sinal subótimo. A função objetivo da camada externa avalia o tempo de execução do experimento, o uso do espaço disponibilizado, bem como restringe possíveis violações de limitações físicas, através da métrica θ . Além disso, avalia-se também a similaridade dos parâmetros, f_δ e das saídas dos sistemas perturbado e estimado, f_o . A partir desses dados o otimizador é capaz de selecionar o sinal que apresente a melhor combinação das métricas supracitadas. Entretanto, essa camada é dependente da camada interna que é responsável pela estimação de parâmetros propriamente dita.

(ii) A camada interna tem como entradas os parâmetros perturbados Γ^P e o sinal determinado na camada externa, $u(k)$. A partir da inserção do sinal no processo perturbado, com saída $y_u(k)$, influenciada por um ruído $n(k)$, a função objetivo avalia tão somente a similaridade entre as saídas para atingir o melhor conjunto de parâmetros dado o sinal fornecido pela etapa anterior, f_o . Para o cálculo de f_o , utiliza-se a saída do processo perturbado com a influência do ruído, $y_p(k)$ e a saída do modelo, cujos parâmetros são selecionados pelo otimizador da camada interna.

Figura 5 – Diagrama esquemático do processo SOESGOPE



Fonte: Elaborado pelo autor

2.5.4 Execução do Experimento

A partir da execução do SOESGOPE apenas com modelos numéricos, inclusive com modelo referência sendo uma perturbação de uma estimativa a priori, é possível aplicar o sinal no sistema real em um ambiente controlado. Sendo essa a etapa de experimento. O sinal foi gerado considerando as limitações físicas tanto do espaço e tempo disponíveis, quanto do sistema propriamente dito. Assim, garante-se a possibilidade de implementação. Além disso, a geração do sinal deve se enquadrar nos sinais de entrada suportados pelo sistema.

Originalmente, o método SOESGOPE foi desenvolvido para trabalhar com o sistema em malha aberta, inserindo o sinal conforme ele foi gerado. De posse do sinal, ele pode ser inserido no sistema a ser identificado. As saídas devem ser coletadas através de um sistema de sensoriamento. Posteriormente, de modo off-line os dados coletados são manipulados, permitindo a aplicação de um algoritmo de otimização que encontre o conjunto de parâmetros ótimo. Esse procedimento é similar ao que ocorre na camada interna do SOESGOPE, sendo que o processo perturbado é, neste caso, o sistema real. A avaliação da qualidade dos parâmetros encontrados nessa etapa se vale apenas da função de precisão das saídas que mede apenas a similaridade entre as saídas real e do modelo a posteriori.

2.6 CONSIDERAÇÕES FINAIS

Esse capítulo foi destinado à compreensão de conceitos e aspectos relacionados à identificação de sistemas dinâmicos. A literatura aborda diversas formas de apurar um modelo matemático. De forma geral, algoritmos de otimização são considerados altamente relevantes nessa área, visto que tais métodos permitem a identificação de sistemas não lineares sem grandes complexidades. Além disso é possível atingir a otimização de uma função objetivo configurada pelo usuário.

Embora consagrados, os métodos de estimação de parâmetros dependem diretamente da inserção de um sinal de entrada. Através dele, consegue-se excitar de forma suficiente ou não o sistema para que se consiga extrair os parâmetros. Assim, o design de um sinal de excitação através de um processo de otimização pode ser considerado muito importante. Seguindo essa linha, o SOESGOPE é uma alternativa eficiente para a geração de um sinal de excitação subótimo com propósito de estimação. Este método é eficiente para o objetivo proposto e servirá como base para as análises nos capítulos subsequentes.

3 ANÁLISES REALIZADAS

A aplicabilidade da metodologia SOESGOPE se estende a sistemas dinâmicos em geral, contanto que apresentem alguma forma de modelar matematicamente através de equações diferenciais que contenham incógnitas, relacionadas às características físicas, que serão descritas pelo método. Nesse capítulo, serão propostas análises mais profundas da técnica, envolvendo confiabilidade e robustez. Adicionalmente, uma abordagem recursiva baseada na análise dos multiplicadores de Lagrange provenientes do processo de otimização será abordada. Para servir como base para essas análises, o modelo de um robô diferencial será utilizado.

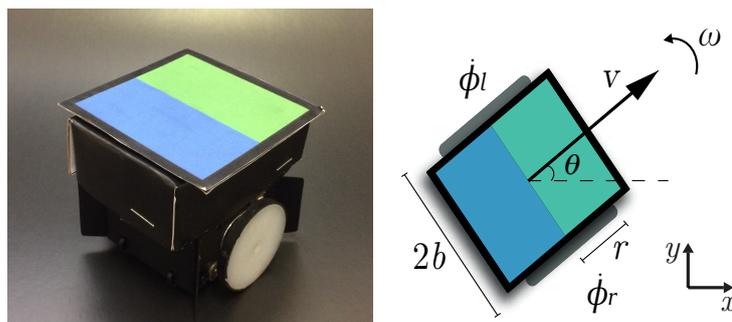
As análises foram foco de dois artigos da revista *International Journal of Control, Automation, and Systems* (IJCAS) da editora *Springer*:

- O primeiro, intitulado *Innovative Analysis for Parameter Estimation Quality*, teve como intuito demonstrar que a redução no espaço de soluções pode ser benéfica para o processo de estimação de parâmetros. O artigo aceito está em processo final de edição para posterior publicação.
- O segundo, *Recursive Approach of Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation*, analisou a influência dos multiplicadores de Lagrange em um processo recursivo, cujo objetivo é aumentar a confiabilidade do processo de estimação frente a diferentes concepções do sistema (SOUZA et al., 2020).

3.1 MODELAGEM DO SISTEMA

A discussão abordada nesse capítulo será baseada na estimativa de parâmetros de um robô diferencial de pequena dimensão, o *Very Small Size Soccer* (VSSS). A Figura 6 apresenta uma imagem do equipamento (esquerda) e sua representação esquemática (direita).

Figura 6 – Robô móvel diferencial: *Very Small Size Soccer*



Fonte: Elaborado pelo autor

Considerando a modelagem tradicional de um robô móvel apresentada por Fahimi (2008), com a dinâmica do motor relacionada, as variáveis de estado podem ser representadas por:

$$x(t) = [p_x(t), p_y(t), \theta(t), v(t), \omega(t), \dot{\phi}_r(t), \dot{\phi}_l(t)]^T \quad (3.1)$$

sendo que os seguintes estados representam:

p_x Posição do robô no eixo x no instante de tempo t (m);

p_y Posição do robô no eixo y no instante de tempo t (m);

$\theta(t)$ Posição angular do robô no instante de tempo t (rad);

$v(t)$ Velocidade linear do robô no instante de tempo t (m/s);

$\omega(t)$ Velocidade angular do robô no instante de tempo t (rad/s);

$\dot{\phi}_r(t)$ Velocidade angular da roda direita no instante de tempo t (rad/s);

$\dot{\phi}_l(t)$ Velocidade angular da roda esquerda no instante de tempo t (rad/s).

A variável de entrada é representada por $u = [E_r(t), E_l(t)]^T$, sendo que $E_r(t)$ e $E_l(t)$ são os níveis de voltagem aplicados nos motores que acionam a roda direita e esquerda, respectivamente. Como a dinâmica do robô e dos motores é acoplada, as variáveis de entrada são os níveis de tensão no sistema e não os torques diretos. O modelo paramétrico do robô é dado por:

$$\dot{x}(t) = \begin{bmatrix} v(t) \cos(\theta(t)) \\ v(t) \sin(\theta(t)) \\ \omega(t) \\ \frac{\ddot{\phi}_r(t) + \ddot{\phi}_l(t)}{2} \\ \frac{(\ddot{\phi}_r(t) - \ddot{\phi}_l(t))r}{b} \\ p_2 \ddot{\phi}_l(t) + p_3 \dot{\phi}_r(t) \\ p_2 \ddot{\phi}_r(t) + p_3 \dot{\phi}_l(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ p_1 & 0 \\ 0 & p_1 \end{bmatrix} u + \vartheta(t) \quad (3.2)$$

$$y(t) = x(t)$$

com

$$p_1 = \frac{K_t}{R(2a_1)}, \quad p_2 = -\frac{a_2}{2a_1}, \quad p_3 = -\frac{\frac{K_t K_w}{R} + \nu}{2a_1 + I_s} \quad (3.3)$$

e

$$a_1 = \frac{mr^2}{8} + \frac{I_{zz}r^2}{2b^2} + \frac{I_w}{2}, \quad a_2 = \frac{mr^2}{4} - \frac{I_{zz}r^2}{b^2} \quad (3.4)$$

sendo que os parâmetros envolvidos nas equações representam:

m Massa do robô (kg);

I_{zz} Momento de inércia em torno do eixo z ($kg\ m^2$);

I_w Momento de inércia das rodas ($kg\ m^2$);

K_t Constante de torque;

K_w Constante de tensão;

R Resistência de armadura (Ω);

ν Coeficiente de atrito viscoso (kg/ms);

r Raio das rodas (m);

b Distância entre a roda e o centro de gravidade (m).

As análises serão realizadas considerando os parâmetros descritos na tabela 2 como reais.

Tabela 2 – Conjunto de parâmetros do modelo do robô “real”

m	r	b	R	ν	K_t	I_{zz}	K_w
0,4500	0,0188	0,0613	1,5511	0,0027	0,0008	0,0000	5,0211

Embora p_1 , p_2 e p_3 incorporem outros parâmetros, eles não têm um significado físico individual, portanto, seu valor pode representar valores imprecisos ou mesmo irrealis. As restrições sobre essas variáveis também são mais difíceis de definir. No entanto, o espaço de solução é reduzido, facilitando o algoritmo de otimização a encontrar boas soluções. Os parâmetros do sistema compacto podem ser observados pela Tabela 3.

Tabela 3 – Conjunto de parâmetros do modelo compacto do robô “real”

p_1	p_2	p_3	r	b
0,1716	-0,1248	-1,7790	0,0195	0,0652

3.2 ANÁLISE DA QUALIDADE DA ESTIMAÇÃO DE PARÂMETROS

O efeito da modelagem sobre o resultado final de um processo de estimação é um fato que deve ser investigado. Diferentes modelos podem determinar diferentes comportamentos simulacionais (KIM; KWON, 2015). Nesse caso, a estimação de parâmetros é diretamente afetada. A complexidade envolvida na modelagem muda de acordo com o conhecimento e a precisão desejada. Todavia, como qualquer modelo é a simplificação de um sistema real, não é garantido que o uso de um maior número de parâmetros garantirá uma boa aproximação.

Uma forma de avaliar a pertinência dos parâmetros é avaliando a qualidade da estimação. A partir de informações obtidas pela Matriz de Fisher, é possível selecionar os parâmetros mais importantes para o modelo (RODRIGUEZ-FERNANDEZ; MENDES; BANGA, 2006; WIRTESENH et al., 2015). Em algumas situações, o parâmetro pode ser redundante ou até não acrescentar informação importante ao modelo. Como exemplo, alguns modelos podem apresentar uma redundância associada aos parâmetros, significando que duas ou mais variáveis podem ser combinadas em um único parâmetro ou até mesmo descartada. Esse conjunto redundante pode, desnecessariamente, aumentar o espaço de soluções, levando a situações em que o método de otimização não consiga atingir o ótimo global. O uso de um modelo paramétrico completo, portanto, pode ser custoso computacionalmente e prejudicial ao processo de identificação.

Face ao exposto, nessa seção será avaliada, através de simulações no sistema não linear descrito, a influência do espaço de soluções na aplicação da técnica SOESGOPE. Três avaliações serão realizadas:

- Caso tutorial, em que apenas um cenário será analisado, a fim de reconhecer a aplicabilidade do uso de parâmetros distintos;
- Impacto do tamanho da população, nessa análise é possível reconhecer vantagens e desvantagens dos modelos em diferentes configurações do método;
- Robustez, a partir desses dados verifica-se o quão robusto é o processo frente à incertezas associadas a estimativa inicial.

De posse dessa avaliação, é possível estabelecer essas análises como capazes de identificar uma modelagem incorreta de um sistema que pode impactar diretamente na identificação e causar danos ao processo.

3.2.1 Caso Tutorial

Para essa análise, será assumido que os parâmetros mostrados pela Tabela 2 e 3 representam precisamente a dinâmica do robô e serão tratados como o sistema real $\mathcal{R}(\Gamma)$.

Pretende-se estimar tal sistema usando como base os dois tipos de modelo. Para efeitos de comparação, assumir-se-á as mesmas estimativas iniciais e perturbadas dos sistemas completo e compacto, presentes nas Tabelas 4 e 5, respectivamente.

Tabela 4 – Estimativas inicial e perturbada do modelo completo

	m	r	b	R	ν	K_t	I_{zz}	K_w
$\hat{\Gamma}^-$	0,4500	0,0210	0,0780	1,5100	0,0050	0,0005	0,0005	5,0000
$\tilde{\Gamma}^p$	0,4250	0,0226	0,0734	1,5437	0,0053	0,0005	0,0005	4,9884

Tabela 5 – Estimativas inicial e perturbada do modelo compacto

	p_1	p_2	p_3	r	b
$\hat{\Gamma}^-$	0,0707	-0,0029	-1,4204	0,0210	0,0780
$\tilde{\Gamma}^p$	0,0689	-0,0015	-1,4709	0,0226	0,0734

Dados esses dois sistemas, a camada externa é executada para gerar o sinal de excitação \mathbf{u}^\oplus , o que torna possível encontrar uma estimativa $\hat{\delta}$ que minimiza a função objetivo.

Em relação à qualidade do sinal, ambos os sistemas retornam valores da métrica global f_Ξ com a mesma magnitude: o sistema com 8 parâmetros apresentou $f_\Xi = 194,50$, contra $f_\Xi = 194,78$ do sistema configurado com 5. O sistema completo convergiu com 20 iterações enquanto o modelo compacto gastou 34 iterações. Métricas individuais relacionadas ao processo de determinação de sinal podem ser observadas na Tabela 6.

Tabela 6 – Métricas relacionadas à função objetivo da camada externa

Modelo	$\theta(\cdot)$	$f_o(\cdot)$	$f_\delta(\cdot)$
8	$3,1025 \cdot 10^{-4}$	1,6340	$8,1399 \cdot 10^{-5}$
5	$3,2969 \cdot 10^{-4}$	1,6174	$7,6581 \cdot 10^{-5}$

Analisando apenas as métricas relacionadas à qualidade do sinal, é possível inferir algumas informações. As métricas $\theta(\cdot)$ e $f_\delta(\cdot)$ são semelhantes nos dois casos. Dessa maneira, a proximidade do parâmetro não é relevante para a determinação do sistema mais vantajoso. Por outro lado, a métrica $f_o(\cdot)$ relacionada à semelhança dos estados revela que o sistema compacto causaria um comportamento mais correlacionado ao sistema perturbado. Para esse sistema e este caso de tutorial específico, percebe-se que, embora as métricas de recuperabilidade dos parâmetros sejam semelhantes, a métrica que avalia os estados para o sistema completo é maior que a métrica para o modelo compacto.

A análise, agora, focará na qualidade do processo de estimativa dos parâmetros a partir da inserção do sinal de entrada no sistema real. Assim, o sinal de entrada \mathbf{u}^\oplus é aplicado a $\mathcal{R}(\Gamma)$ para encontrar as estimativas a posteriori. As Tabelas 7 e 8 mostram os parâmetros encontrados para os sistemas completos e compactos, respectivamente, bem

como a diferença para o sistema real, sendo que $\Delta\Gamma = |\Gamma - \hat{\Gamma}^+|$. Tal métrica avalia o desvio de $\hat{\Gamma}^+$ à Γ em relação à magnitude de cada componente de Γ .

Tabela 7 – Estimativa a posteriori do modelo completo

	m	r	b	R	ν	K_t	I_{zz}	K_w
$\hat{\Gamma}^+$	0,4815	0,0194	0,0621	1,0005	0,0028	0,0005	0,0000	5,1566
$\Delta\Gamma$	0,0315	0,0006	0,0008	0,5506	0,0001	0,0003	0,0000	0,1355

Tabela 8 – Estimativa a posteriori do modelo compacto

	p_1	p_2	p_3	r	b
$\hat{\Gamma}^+$	0,1764	-0,1529	-1,8280	0,0194	0,0656
$\Delta\Gamma$	0,0048	0,0281	0,0490	0,0001	0,0004

Analisando o erro absoluto, não é possível perceber qual modelo é mais eficaz para identificar o sistema devido à magnitude dos parâmetros, entretanto, o erro relativo é capaz de fornecer informações importantes sobre a qualidade da estimação. O erro relativo total para a estimativa considerando o sistema completo foi de 10,12% contra 3,81% do sistema compacto. Nessa simulação, portanto, esse fato conta como a maior vantagem do modelo usando apenas 5 parâmetros.

Comparando os estados aproximados por cada um dos sistemas, os seguintes valores da função de precisão das saídas foram encontrados: $f_o(\cdot) = 1,6341$ para o sistema completo e $f_o(\cdot) = 1,6167$ para o compacto. Assim como nos resultados relacionados à métrica f_δ , nessa análise o sistema completo apresentou um valor maior de diferença entre os estados real e aproximado.

A partir dessas análises ligadas à qualidade do sinal e da estimação, nesse caso, o modelo compacto é mais vantajoso para estimar os parâmetros do veículo. Entretanto, é precipitado chegar a alguma conclusão, visto que essa análise é baseada em um caso isolado e, dessa forma, os resultados não podem ser usados para uma comparação confiável e justa. No sentido de subsidiar uma decisão, as próximas subseções farão análises mais aprofundadas do método.

3.2.2 Impacto do tamanho da população

A influência do número de partículas, n_p , no algoritmo PSLO usado para obter o sinal de entrada, \mathbf{u}^\oplus , será investigado. Esse índice tem conexão direta com o tempo de simulação e esforço computacional.

Com propósito de avaliação, quatro cenários com diferentes tamanhos de população são comparados: $n_p = 5, 10, 25$ e 100 . Até 100 iterações da camada externa são realizadas em cada situação. Em todas elas, a mesma estimativa inicial e perturbada são usadas em 100 diferentes amostras. Para avaliar a eficiência do método e das modelagens variando o

número de partículas, as métricas relacionadas à qualidade do sinal serão utilizadas. Além das métricas, as tabelas apresentam o tempo de convergência $t_{conv}(s)$. Já as Tabelas 9 and 10 apresentam a média de cada métrica para os dois modelos.

Tabela 9 – Sensibilidade ao número de partículas - modelo completo

n_p	$\theta(\cdot)$	$f_o(\cdot)$	$f_\delta(\cdot)$	$f_\Xi(\cdot)$	t_{conv}
5	$3,1 \cdot 10^{-4}$	1,64	$8,5 \cdot 10^{-3}$	203,79	127
10	$2,4 \cdot 10^{-4}$	1,65	$7,0 \cdot 10^{-3}$	196,62	256
25	$2,3 \cdot 10^{-4}$	1,65	$6,9 \cdot 10^{-3}$	195,05	609
100	$1,9 \cdot 10^{-4}$	1,66	$5,5 \cdot 10^{-3}$	192,18	2193

Tabela 10 – Sensibilidade ao número de partículas - modelo compacto

n_p	$\theta(\cdot)$	$f_o(\cdot)$	$f_\delta(\cdot)$	$f_\Xi(\cdot)$	t_{conv}
5	$4,9 \cdot 10^{-4}$	1,63	$5,5 \cdot 10^{-4}$	213,78	149
10	$3,3 \cdot 10^{-4}$	1,64	$6,2 \cdot 10^{-4}$	197,63	144
25	$3,1 \cdot 10^{-4}$	1,63	$4,5 \cdot 10^{-4}$	195,53	429
100	$2,4 \cdot 10^{-4}$	1,65	$7,6 \cdot 10^{-4}$	190,72	1806

A partir das tabelas, algumas informações podem ser extraídas. A primeira está relacionada com o uso do espaço e do tempo. Quanto maior o número de partículas, menor o valor da métrica $\theta(\cdot)$. Essa melhoria é evidente, pois o número de partículas aumenta a confiabilidade de qualquer método de otimização. Embora o conjunto de partículas seja crucial, mesmo com poucas partículas, todas as simulações convergiram para experimentos viáveis. Portanto, o maior número de partículas pode não trazer vantagens significativas relacionadas a essa métrica.

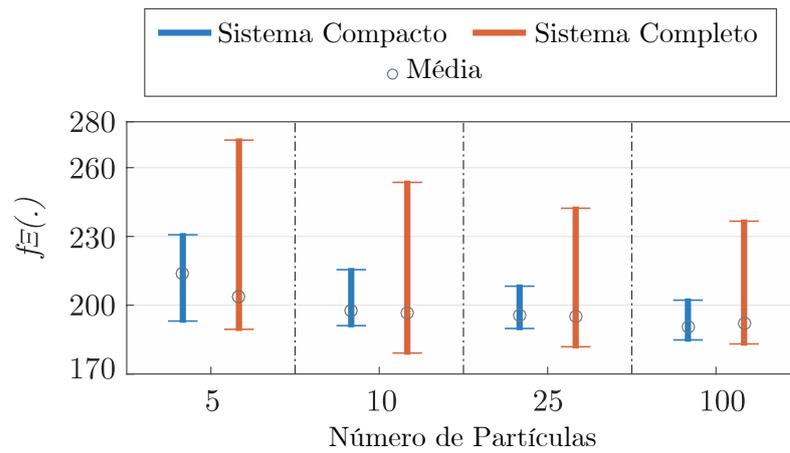
Apesar do aumento do número de partículas, as outras métricas $f_\delta(\cdot)$ e $f_o(\cdot)$ não acompanham uma lógica definida. Considerando o modelo completo e a média das métricas, $f_o(\cdot)$ aumenta enquanto $f_\delta(\cdot)$ reduz. Da mesma forma, o processo considerando o modelo compacto não determina um padrão. Nesse caso, a melhor situação de $f_\delta(\cdot)$ é observada com 25 partículas e $f_o(\cdot)$ para esse número de partículas também era menor. Portanto, essas métricas não são adequadas para definir o melhor tamanho da população.

A última informação é a soma ponderada das outras métricas $f_\Xi(\cdot)$. Aumentando o número de partículas, essa métrica é reduzida em todos os dois casos. O aumento do número de partículas é determinante para uma melhor qualidade do sinal, no entanto, o tempo de convergência também aumenta substancialmente. Essa redução de $f_\Xi(\cdot)$ é determinada principalmente por causa da métrica de estado e espaço, pois as outras métricas não contribuem efetivamente com a função objetivo final.

A partir desses resultados, a confiabilidade do método deve ser avaliada, ou seja, quanto a função objetivo pode se desviar dessa média. A figura 7 mostra a evolução de $f_\Xi(\cdot)$ de acordo com o número de partículas, para ambos os sistemas. Em geral, um

número maior de partículas leva a uma variação menor da função objetivo $f_{\Xi}(\cdot)$ para obter maior confiabilidade na determinação do sinal de excitação do sistema. Embora o modelo completo possa obter um resultado mais satisfatório, ou seja, com menor valor de $f_{\Xi}(\cdot)$ no melhor caso, ele tem a desvantagem de apresentar uma grande variação entre os melhores e os piores casos. Por outro lado, o sistema compacto apresenta maior confiabilidade, obtendo uma variação reduzida da função objetivo.

Figura 7 – Valores máximos e mínimos da função objetivo com diferentes número de partículas para cada modelo



Fonte: Elaborado pelo autor

Essas análises sugerem um limite prático para o número de partículas, uma vez que uma melhoria significativa na qualidade do sinal não é acompanhada pela redução do tempo de convergência, da recuperabilidade do parâmetro e da aproximação de estados. Assim, o número de partículas pode ser determinado empiricamente com base no conhecimento do usuário sobre o sistema.

No que se diz respeito à comparação entre os dois modelos, alguns aspectos de cada um podem ser enfatizados. Assim como o resultado obtido para o caso do tutorial (consulte Subseção 3.2.1), o aumento da recuperabilidade dos parâmetros não é acompanhada pela aproximação entre os estados estimado e real. Por outro lado, essa métrica piora. A fonte desse fenômeno pode ser a redundância de parâmetros. O maior número de parâmetros leva a um maior número de combinações que fornecem comportamentos semelhantes. O modelo compacto apresenta uma melhor qualidade de sinal quando o número de partículas aumenta. No entanto, outras métricas não seguem esse comportamento. Nesse caso, não é observado um aumento da métrica $f_o(\cdot)$ com a redução de $f_{\delta}(\cdot)$. Considerando esse fato e a confiabilidade da função objetivo global, o melhor modelo é o compacto.

3.2.3 Robustez

Como o SOESGOPE é um algoritmo dependente de uma estimativa inicial e de um sistema perturbado que será utilizado para emular o comportamento do sistema real e que

é desconhecido, deve-se mensurar o quão robusto ele é frente à incertezas dessa estimativa. O sistema do caso tutorial será utilizado para analisar a sensibilidade da convergência de saída em relação à amplitude do parâmetro incerteza. Ou seja, como a “distância” da estimativa inicial aos parâmetros do sistema real, Γ , influenciam na estimação. Um aumento percentual na incerteza será utilizado e os conjuntos de parâmetros de $\hat{\Gamma}^-$ e $\tilde{\Gamma}^p$ serão randomicamente obtidos da seguinte forma:

$$\hat{\gamma}_i^- = \gamma_i \left(1 + \frac{var}{100} \varpi_1 \right) \quad (3.5)$$

$$\tilde{\gamma}_i^p = \hat{\gamma}_i^- \left(1 + \frac{var}{100} \varpi_2 \right) \quad (3.6)$$

sendo que $i \in \{1, 2, \dots, r\}$, $var = \{10\%, 30\%, 50\%, 70\%, 90\%\}$, e $\varpi_1, \varpi_2 \in \mathbb{R} \mid \varpi_1, \varpi_2 \in [-1, 1]$ são números randômicos uniformemente distribuídos. Note que a variação depende da magnitude do parâmetro γ_i e também afeta γ_i^p . A partir dessas equações de geração serão geradas 100 amostras para cada valor de porcentagem de incerteza var . Os melhores e os piores resultados relacionados aos casos simulados são sumarizados pelas Tabelas 11 e 12.

Tabela 11 – Métricas relacionadas à melhor estimação fornecida pelo SOESGOPE para cada modelo

var	$f_o(\cdot)$		$f_\delta(\cdot)$		t_u		$f_\Xi(\cdot)$	
	5	8	5	8	5	8	5	8
10%	1,66	1,66	$1,8 \cdot 10^{-3}$	$1,2 \cdot 10^{-4}$	14,6	5,3	190,71	179,75
30%	1,67	1,64	$4,0 \cdot 10^{-4}$	$1,5 \cdot 10^{-3}$	13,6	14,8	190,47	191,93
50%	1,65	1,65	$6,3 \cdot 10^{-4}$	$8,8 \cdot 10^{-3}$	14,7	14,7	190,10	198,34
70%	1,69	1,67	$1,0 \cdot 10^{-4}$	$1,3 \cdot 10^{-3}$	5,3	14,1	184,51	190,91
90%	1,70	1,65	$1,9 \cdot 10^{-4}$	$4,4 \cdot 10^{-3}$	5,4	17,0	184,17	193,91

Tabela 12 – Métricas relacionadas à pior estimação fornecida pelo SOESGOPE para cada modelo

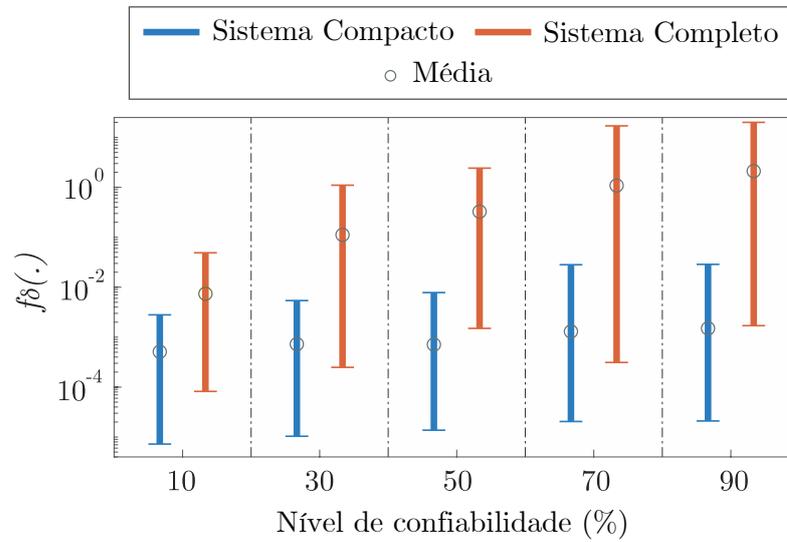
var	$f_o(\cdot)$		$f_\delta(\cdot)$		t_u		$f_\Xi(\cdot)$	
	5	8	5	8	5	8	5	8
10%	1,65	1,65	$1,1 \cdot 10^{-3}$	0,04	46,1	48,8	225,86	272,56
30%	1,64	1,62	$1,3 \cdot 10^{-3}$	1,10	43,8	24,6	220,00	$1,30 \cdot 10^3$
50%	1,64	1,65	$5,2 \cdot 10^{-3}$	2,42	73,3	46,4	251,20	$2,65 \cdot 10^3$
70%	1,65	1,63	$2,8 \cdot 10^{-2}$	17,01	49,4	33,6	250,62	$2,68 \cdot 10^5$
90%	1,65	1,66	$1,3 \cdot 10^{-2}$	20,00	50,0	53,0	245,84	$1,33 \cdot 10^7$

A primeira informação que pode ser extraída, principalmente da Tabela 12 é que quanto mais distante os parâmetros iniciais estiverem dos reais, mais tempo de experimento é requerido. Isso significa que para estimar os parâmetros, mais dados experimentais são necessários. Além disso, acompanhando esse comportamento de t_u , a função objetivo

global, $f_{\Xi}(\cdot)$, aumenta consideravelmente seu valor, especialmente para o sistema completo que em alguns casos não converge para um ponto viável.

É necessário enfatizar que o aumento da incerteza pode ser benéfico para atingir parâmetros mais confiáveis, como pode ser notado pela Tabela 11 para o modelo compacto. Entretanto, essa incerteza maior reduz a repetibilidade dos resultados, conforme pode ser observado pela Figura 8. Tal figura ilustra a dispersão da métrica $f_{\delta}(\cdot)$ de acordo com o nível de confiabilidade.

Figura 8 – Valores máximo mínimo da função de recuperabilidade de parâmetros - $f_{\delta}(\cdot)$ - em escala logarítmica para diferentes níveis de confiabilidade



Fonte: Elaborado pelo autor

Embora o uso do sistema completo apresente resultados satisfatórios no melhor caso, a análise indica parâmetros incoerentes no pior caso com o aumento da incerteza: restrições violadas e alta função objetivo $f_{\Xi}(\cdot)$. Existem duas causas principais, que podem levar à não convergência:

- (i) o conjunto inicial de partículas é insuficiente ou mal escolhido; ou
- (ii) os modelos $\mathcal{M}(\hat{\Gamma}^-)$ e $\mathcal{M}(\tilde{\Gamma}^p)$ não são compatíveis com $\mathcal{R}(\Gamma)$ nem estáveis.

A solução da situação (i) é direta: escolhendo novas partículas ou aumentando seu número.

No caso da condição (ii), é mais difícil identificar e depende de quão boas são as aproximações $\mathcal{M}(\hat{\Gamma}^-)$ e $\mathcal{M}(\tilde{\Gamma}^p)$. Destaca-se que essa situação causa um impacto significativo em $\hat{\Gamma}^+$, uma vez que afeta a qualidade do sinal de estimativa \mathbf{u}^{\oplus} , consulte Tabela 12.

Portanto, uma conclusão geral dessa análise é que, quanto maior a “distância” de uma estimativa a priori excelente, mais fraca é a qualidade do sinal de estimativa, que pode ser medido através de $f_{\delta}(\cdot)$. Focando na comparação dos modelos, alguns aspectos

de cada um devem ser pontuados. O sistema completo é totalmente ineficiente usando parâmetros iniciais muito distantes. Em alguns casos, o processo não consegue convergir. Assim, pode-se dizer que o uso de parâmetros adicionais aumenta a probabilidade de não convergência, tornando o processo de estimativa instável, independentemente do sinal de entrada aplicado. Por outro lado, embora o modelo compacto não apresente excelente resultados, quando se utiliza estimativas iniciais piores ainda assim o algoritmo atinge a convergência.

3.3 ABORDAGEM RECURSIVA

Conforme observado na Seção 3.2, o SOESGOPE é um algoritmo altamente sensível à estimação inicial. Dependendo do conjunto de parâmetros selecionado, altera-se o sinal selecionado e o resultado final. Essa situação se agrava quando os parâmetros não têm significado físico, ou quando existem incertezas em relação a alguns aspectos operacionais, ou mesmo quando não existem equações analíticas para fornecer valores iniciais (UGALDE et al., 2015; AHSAN; CHOUDHRY, 2017).

Considerando um processo de estimativa de parâmetros, uma melhor semelhança entre os comportamentos do modelo e o sistema real está diretamente relacionada ao aumento da precisão dos parâmetros a priori (ISERMANN; MÜNCHHOF, 2010). Portanto, o uso de um processo recursivo dentro do método de estimativa pode ser útil para a obtenção de um modelo mais confiável (DING; LIU; LIU, 2016; DING; LIU; GU, 2016; XU; DING, 2017; LIU et al., 2018; CHEN; JIANG; LI, 2018). Todos os trabalhos anteriores envolvendo processos recursivos visam aumentar a precisão dos parâmetros encontrados. No entanto, é importante destacar que essas aplicações não consideraram restrições de tempo e variáveis. Ao considerar essas restrições, o estado do espaço e a incerteza sobre as variáveis aumentam junto com a complexidade dos modelos matemáticos do sistema, dificultando que os algoritmos de otimização encontrem soluções razoáveis.

Um problema com maiores incertezas sobre as variáveis envolve encontrar limites operacionais corretos. É bem sabido pela literatura de otimização (OLIVEIRA et al., 2015) que a falta de limites razoáveis pode causar instabilidade numérica e problemas de convergência. No entanto, restringir severamente esses limites pode fornecer soluções inadequadas. Portanto, há uma interação entre soluções razoáveis e viabilidade que é crítica para vários sistemas. Assim, o desenvolvimento de um método que possa alterar dinamicamente os limites do espaço de pesquisa, considerando situações em que o intervalo de solução não é conhecido, é relevante. Assim, a junção entre uma abordagem iterativa baseada em um sinal de excitação subótimo e a recursividade no processo de estimativa de parâmetros pode trazer bons benefícios.

Dessa forma, esta seção apresenta um aprimoramento do método SOESGOPE de acordo com um ponto de vista de otimização recursiva. O fator primordial que permite a

integração do método ao processo recursivo é a aplicação dos multiplicadores de Lagrange a partir do processo de otimização. Esses índices direcionam a pesquisa e também podem ser usados como critérios de parada eficientes, fornecendo uma ferramenta poderosa quando a estimativa inicial não é precisa. Para demonstrar a eficácia do método, serão discutidos resultados baseados no sistema apresentado na Seção 3.1 utilizando o modelo compacto.

3.3.1 SOESGOPE Recursivo

O método SOESGOPE é adequado para ser aplicado em situações em que os parâmetros podem ser suficientemente bem estimados previamente, ou seja, a estimativa a priori deve ser o mais próxima possível dos valores reais. Esse fato é facilmente abrangido caso os parâmetros correspondam a características físicas, como, por exemplo, massa e comprimento. No entanto, em uma situação real, alguns parâmetros não têm significado físico. Dessa forma, é difícil obter valores iniciais precisos relacionados a eles. Dessa maneira, a aplicação de um método recursivo junto com o SOESGOPE resulta em uma alternativa para lidar com esse problema.

É possível reiniciar todo o processo de otimização recursivamente usando o $\hat{\Gamma}^+$ recém-obtido como um novo $\hat{\Gamma}^-$, pois esse procedimento refina gradualmente $\hat{\Gamma}^+$ para Γ . Portanto, esta subseção apresenta um algoritmo recursivo adequado para situações em que se tem pouco conhecimento do conjunto de parâmetros, ou seja, quando $\hat{\Gamma}^-$ esteja distante Γ . A métrica que calcula essa “distância” é definida como:

$$\Sigma_{\hat{\Gamma}^-} = \sum_{i=1}^r \left| \frac{(\hat{\gamma}_i^+ - \hat{\gamma}_i^-)^2}{\hat{\gamma}_i^+} \right| \quad (3.7)$$

onde $\hat{\Gamma}^- = [\hat{\gamma}_1^-, \hat{\gamma}_2^-, \dots, \hat{\gamma}_r^-]^T$ e $\hat{\Gamma}^+ = [\hat{\gamma}_1^+, \hat{\gamma}_2^+, \dots, \hat{\gamma}_r^+]^T$. Tal métrica é um meio de determinar se a procura pelos parâmetros atingiu a convergência.

Outra condição de parada para o algoritmo recursivo proposto refere-se ao algoritmo de otimização, mais especificamente aos multiplicadores Lagrange. Essa informação é determinada como uma métrica que mensura o quanto o parâmetro encontrado tenta extrapolar os limites superior, $\bar{\lambda}$, e inferior, $\underline{\lambda}$, configurados. Caso a métrica $|\bar{\lambda} + \underline{\lambda}|$ for muito pequena, isso indica que o algoritmo de estimativa de parâmetros está longe dos limites e o ponto interior encontrou um ótimo dentro da região aceitável.

Caso os parâmetros forem desconhecidos e os limites de pesquisa não puderem ser determinados com clareza, o SOESGOPE pode não convergir satisfatoriamente. Por exemplo, se o valor real de um determinado parâmetro for 2 e o limite de pesquisa estiver entre 1 e 1,5, o valor real não poderá ser alcançado. Portanto, é necessário conduzir esses limites a valores plausíveis.

Como a estimativa a posteriori provém de um processo de otimização, é possível

extrair dos multiplicadores de Lagrange métricas relacionadas à extrapolação dos limites e, a partir daí, estender os limites predeterminados para cada parâmetro. Portanto, além de ser um critério de parada do algoritmo, os multiplicadores de Lagrange são fatores determinantes para a convergência do método.

O respectivo nível de confiabilidade de um parâmetro fornece os limites de pesquisa para cada parâmetro. Parâmetros mensuráveis, como massa e diâmetro, têm limites bem estabelecidos dentro da precisão do equipamento de medição. No entanto, alguns parâmetros podem não ter limites definidos. É o caso de parâmetros que não têm significado físico. Uma maneira de resolver isso é deixar o limite de pesquisa amplo desde o início das rodadas. Mas essa estratégia pode levar a uma solução estacionada em um mínimo local ruim.

Portanto, é essencial ter uma maneira de alterar os limites, juntamente com as iterações do algoritmo, caso o valor do parâmetro tenda a extrapolá-los. Uma maneira é a determinação inicial de um intervalo de pesquisa restrito. A partir daí, a cada iteração, relaxa-se o limite de apenas um parâmetro. Dessa forma, a pesquisa é direcionada para esse parâmetro, sem afetar outros, que podem estar dentro do intervalo predeterminado. É possível analisar o comportamento dos multiplicadores Lagrange com propósito de verificar se os limites dos parâmetros estão ou não bem definidos e mesmo quais deles estão mais distantes dos valores ideais. Ao inspecionar os vetores multiplicadores Lagrange de limite superior e inferior pode ser detectado se a região de pesquisa deve ser aumentada e se o algoritmo de otimização está preso em qualquer mínimo local.

Dada uma região de pesquisa determinada por $\delta_i^l < \delta_i < \delta_i^u$, sendo que $i = \{1, 2, \dots, r\}$, relacionado à estimativa do parâmetro r . Na primeira rodada do algoritmo recursivo, é possível encontrar Γ^+ dentro dos limites impostos. No processo de otimização, os vetores superiores $\bar{\lambda}$ e inferiores $\underline{\lambda}$ dos Multiplicadores de Lagrange podem ser obtidos. Após analisar a magnitude dos valores normalizados, $\bar{\lambda}/\|\bar{\lambda}\|$ e $\underline{\lambda}/\|\underline{\lambda}\|$, o parâmetro que obtiver a maior tentativa de violação de limite é selecionado e, portanto, esse limite é aumentado para cobrir uma área de pesquisa maior na próxima iteração. Esse processo é repetido até que um dos critérios de parada explicados anteriormente seja atingido. É importante observar que apenas uma região de pesquisa, δ_i^l ou δ_i^u é alterada em cada iteração do algoritmo recursivo. A abordagem proposta é apresentada pelo Algoritmo 1, em que $th_1, th_2 \in \mathbb{R}_{>0}$ são valores relacionados aos critérios de parada definidos pelo usuário.

3.3.2 Avaliação do Caso Tutorial

Quando a estimativa inicial $\hat{\Gamma}^-$ está longe dos valores reais, a metodologia SOESGOPE pode não conseguir encontrar parâmetros aceitáveis $\hat{\Gamma}^+$ de uma só vez. No entanto, sempre é possível usar recursivamente o $\hat{\Gamma}^+$, recentemente obtido, como um novo $\hat{\Gamma}^-$ e reiniciar

Algorithm 1 : Algoritmo Recursivo

Require: Estimativa inicial $\mathcal{M}(\hat{\Gamma}^-)$; Região de busca δ_i

while $\Sigma_{\hat{\Gamma}^-} > th_1$, or $\|\bar{\lambda} + \underline{\lambda}\| > th_2$, ou limite de iterações seja atingido **do**

Obter \mathbf{u}^\oplus usando $\hat{\Gamma}^-$ como estimativa inicial

Estimar $\hat{\Gamma}^+$ usando \mathbf{u}^\oplus e obter $\bar{\lambda}$ and $\underline{\lambda}$

$\text{ind} \leftarrow \text{index}(\max(\max(\underline{\lambda}), \max(\bar{\lambda})))$

Alterar limite da região de busca δ_{ind}

Avaliar $\Sigma_{\hat{\Gamma}^-}$ e $\|\bar{\lambda} + \underline{\lambda}\|$

$\hat{\Gamma}^- \leftarrow \hat{\Gamma}^+$ (nova estimativa inicial)

end while

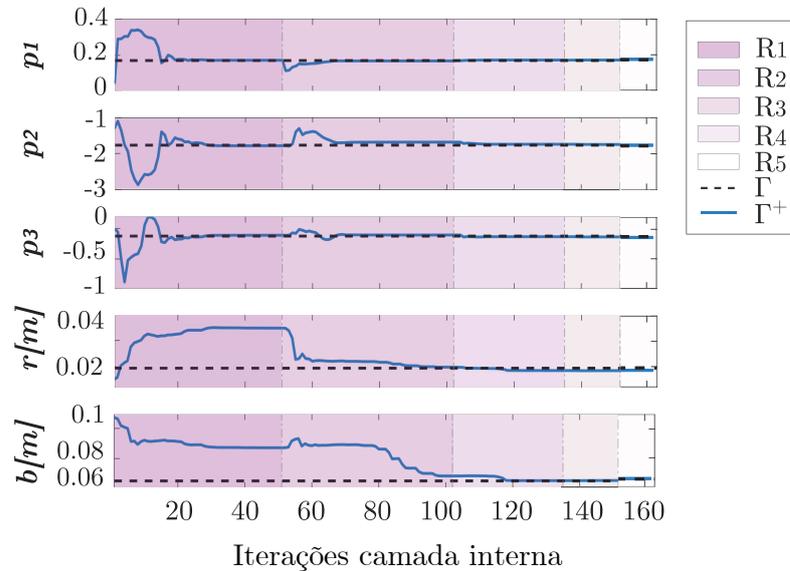
return: Conjunto representativo de parâmetros $\hat{\Gamma}^+$

todo o processo de otimização.

3.3.2.1 Parâmetros cobertos pela região de busca

Seja a estimativa inicial $\hat{\Gamma}^- = [0,0401 \ -0,0032 \ -1,3231 \ 0,0152 \ 0,1072]^T$, o número máximo de iterações do algoritmo recursivo seja 10 e os limites dos parâmetros e multiplicadores de Lagrange sejam $th_1 = 10^{-4}$ e $th_2 = 10^{-6}$. A região de pesquisa é definida como $|\delta_i| < [2 \ 2 \ 2 \ 0,2 \ 0,2]$, $i \in \{1, 2, 3, 4, 5\}$, em todas as iterações. Para esse exemplo, a Figura 9 mostra o comportamento de $\hat{\Gamma}^+$ em função das iterações, apresentando a dinâmica do algoritmo de ponto interior, através de 5 iterações do Algoritmo 1.

Figura 9 – Gráfico de convergência dos parâmetros



Fonte: Elaborado pelo autor

O processo como um todo utilizou 163 iterações do algoritmo de pontos interiores (camada interna). A Tabela 13 apresenta os valores da métrica $f_o(\cdot)$ e os critérios de parada de cada rodada. Adicionalmente, a Tabela 14 apresenta os parâmetros encontrados ao final de cada rodada.

Tabela 13 – Experimento recursivo - Métricas

Rodada	$f_o(\cdot)$	$\Sigma_{\hat{\Gamma}^-}$	$\ \bar{\lambda} + \underline{\lambda}\ $
1	1,6282	$3,47 \cdot 10^{-1}$	$1,62 \cdot 10^{-1}$
2	1,6210	$2,19 \cdot 10^{-2}$	$2,13 \cdot 10^{-4}$
3	1,6256	$7,90 \cdot 10^{-3}$	$4,59 \cdot 10^{-5}$
4	1,6251	$1,13 \cdot 10^{-4}$	$5,66 \cdot 10^{-5}$
5	1,6251	$1,07 \cdot 10^{-6}$	$5,90 \cdot 10^{-6}$

Tabela 14 – Experimento recursivo - Parâmetros

$\hat{\Gamma}^+$				
p_1	p_2	p_3	r	b
0,1729	-0,1114	-1,7945	0,0347	0,0872
0,1694	-0,1075	-1,6974	0,0199	0,0685
0,1742	-0,1351	-1,7545	0,0186	0,0654
0,1742	-0,1387	-1,7607	0,0186	0,0655
0,1742	-0,1390	-1,7614	0,0186	0,0655

Observa-se que o processo de otimização ficou preso no mínimo local em torno da 50^a iteração do algoritmo de pontos interiores. Nesta iteração, embora $f_o(\cdot)$ tenha apresentado um bom valor, as métricas de avaliação $\Sigma_{\hat{\Gamma}^-}$ e $\|\bar{\lambda} + \underline{\lambda}\|$ indicaram que alguns parâmetros ainda não conseguiram encontrar um resultado satisfatório; nesse caso, percebe-se que r e b (segunda linha da Tabela 14) estavam longe do valor real, 0,0195 e 0,0652, respectivamente. Esse fato significa que o conjunto anterior de parâmetros $\hat{\Gamma}^-$, juntamente com seu conjunto perturbado $\tilde{\Gamma}^p$, não eram considerados como uma aproximação válida do sistema real.

A cada iteração, o processo recursivo é repetido e continuado até que os critérios de parada sejam atingidos. Na quinta iteração, a métrica $\Sigma_{\hat{\Gamma}^-}$ ultrapassou o limite imposto e os parâmetros foram estabelecidos. É importante observar que, neste exemplo, os limites de pesquisa impostos cobriam os valores reais. No entanto, situações em que os parâmetros não são encontrados podem ocorrer devido à falta de cobertura dos parâmetros na região de pesquisa. Portanto, uma maneira de resolver esse problema é usar as informações provenientes do problema de otimização.

3.3.2.2 Parâmetros Inatingíveis

Considera-se a estimativa inicial $\hat{\Gamma}^-$ dada pela Tabela 15 e a região de busca definida pela Tabela 16:

Tabela 15 – Estimativa a priori definida para o experimento

	p_1	p_2	p_3	r	b
$\hat{\Gamma}^-$	0,1000	-0,1964	-1,0000	0,0283	0,0849

Tabela 16 – Região de busca definida para o experimento

	p_1	p_2	p_3	r	b
δ_i^l	0,0001	-0,2000	-1,0000	0,0100	0,0100
δ_i^u	0,1000	-0,0001	-0,0001	0,0400	0,1000

sendo que os parâmetros p_2 e p_3 são negativos, como pode ser visto pelas Equações 3.3. A precisão de medição definiu os limites de r e b . Para esta experiência, os limites dos parâmetros p_1 , p_2 e p_3 foram dados sem o conhecimento prévio. Observando os parâmetros reais fornecidos pela Tabela 3, é possível perceber que esses limites não conseguem envolver alguns valores “reais” do modelo (p_1 , p_2 e p_3). Assim, esse exemplo demonstra a aplicabilidade do método recursivo a situações em que os limites não são bem definidos previamente. Para esse problema, a cada rodada, os limites de pesquisa podem ser reduzidos ou aumentados em 0,2.

As Tabelas 17 e 18 resumem os resultados obtidos no processo de otimização, considerando diferentes regiões de pesquisa. Os multiplicadores de Lagrange são informados para analisar a resposta do sistema aos limites impostos.

Tabela 17 – Experimento recursivo - Alterações na região de busca

Rodada	Região de Busca	$\ \bar{\lambda} + \underline{\lambda}\ $
1	$\delta_i^l < \delta_i < \delta_i^u$	0,1107
2	$-1,2000 < \delta_3 < -0,0001$	0,5650
3	$0,0001 < \delta_1 < 0,3000$	0,0348
4	$-0,4000 < \delta_2 < -0,0001$	0,0575
5	$-1,4000 < \delta_3 < -0,0001$	0,0653
6	$-1,6000 < \delta_3 < -0,0001$	0,0592
7	$-1,8000 < \delta_3 < -0,0001$	$4 \cdot 10^{-5}$

Tabela 18 – Experimento recursivo - Parâmetros

$\hat{\Gamma}^+$				
p_1	p_2	p_3	r	b
0.0907	-0.2000	-1.0000	0.0216	0.0661
0.0907	-0.2000	-1.2000	0.0278	0.0708
0.1410	-0.2000	-1.2000	0.0194	0.0660
0.1488	-0.1819	-1.2000	0.0184	0.0663
0.1491	-0.1741	-1.4000	0.0186	0.0662
0.1493	-0.1250	-1.6000	0.0208	0.0667
0.1659	-0.1227	-1.7392	0.0202	0.0629

Como resultado dos estreitos limites impostos, a primeira solução encontrada não é satisfatória, fato que pode ser percebido pelo alto valor encontrado para a norma de soma dos multiplicadores de Lagrange inferior e superior.

Na primeira rodada do algoritmo recursivo, atinge-se: $\bar{\lambda}/\|\bar{\lambda}\| = [0, 7 \ 0 \ 0 \ 0, 4 \ 0, 1]$ e $\underline{\lambda}/\|\underline{\lambda}\| = [0 \ 0, 6 \ 0, 8 \ 0 \ 0]$. O valor mais alto da tentativa de violação dos limites está associado ao parâmetro p_3 . Esse fato indica que esse parâmetro excede o limite inicialmente imposto para o limite inferior. Portanto, para p_3 obter uma solução aceitável, é necessário que o limite atribuído a ele seja aumentado para reduzir $\underline{\lambda}$. Nesse caso, o limite foi estendido para $\delta_3^l = -1, 2$, conforme descrito na Rodada 2. Embora o limite para o parâmetro p_3 tenha sido alterado, a nova norma da soma dos multiplicadores de Lagrange aumentou na nova rodada do problema, indicando que ainda existem parâmetros que estão fora do limite de pesquisa. Nesse caso, o parâmetro p_1 .

Outras rodadas foram desenvolvidas da mesma maneira. O valor de $\|\bar{\lambda} + \underline{\lambda}\| = 4, 10 \cdot 10^{-6}$ menor que th_2 indica que os limites impostos ao final do processo são suficientes para encontrar uma solução viável.

O exemplo acima também foi usado para comparar o método recursivo proposto com o SOESGOPE original. Como o método original não contempla a alteração da região de pesquisa ao longo do processo, considera-se os limites relaxados de acordo com a Tabela 19.

Tabela 19 – Região de busca - SOESGOPE original

	p_1	p_2	p_3	r	b
δ_i^l	0,0001	-2,0000	-2,0000	0,0100	0,0100
δ_i^u	2,0000	-0,0001	-0,0001	0,0400	0,1000

A Tabela 20 apresenta os resultados para os dois métodos e a melhor estimativa Γ , sendo que MO representa o método original e MR, o método recursivo.

Tabela 20 – Estimativa a priori - Comparação entre método original e recursivo

	p_1	p_2	p_3	r	b
MO	0,1905	-0,0526	-1,9999	0,0197	0,0652
MR	0,1659	-0,1227	-1,7392	0,0202	0,0629
Γ	0,1716	-0,1248	-1,7790	0,0195	0,0652

Embora o SOESGOPE englobe todos os parâmetros reais, é possível perceber que o uso de uma região de pesquisa relaxada pode ser prejudicial ao processo de estimativa. Ao usar o método original, ignorando os limites, a probabilidade de encontrar um mínimo local é grande. Portanto, a metodologia proposta pode direcionar a região de pesquisa ao longo das iterações, levando a uma boa alternativa.

Embora o pior resultado não apresente alta precisão, ele pode atingir valores próximos ao real, o que significa que os limites iniciais de busca, incorretamente definidos, não afetaram o processo de estimativa. Em todas as amostras, os limites finais do

processo recursivo englobavam os parâmetros reais, verificando a eficácia da análise dos multiplicadores de Lagrange.

3.3.2.3 *Análise de Robustez*

Para analisar a robustez da metodologia recursiva, foram gerados aleatoriamente conjuntos de parâmetros $\hat{\Gamma}^-$ e $\tilde{\Gamma}^p$. O mesmo se aplica à região de pesquisa $\delta_i^l < \delta_i < \delta_i^u$. O procedimento é repetido 100 vezes configurando diferentes cenários. Ruídos são considerados para os testes. O melhor (ME) e o pior (PE) conjuntos de estimativas a posteriori e a média da estimativa (M) das 100 repetições são resumidos pela Tabela 21.

Tabela 21 – Melhor, pior e média das estimativas a posteriori fornecidas pelo método recursivo

	p_1	p_2	p_3	r	b
ME	0,1725	-0,1247	-1,7765	0,0193	0,0651
PE	0,1715	-0,2001	-1,7366	0,0191	0,0657
M	0,1719	-0,1264	-1,7790	0,0195	0,0650

Para obter a melhor estimativa, os parâmetros atingiram o valor real e a métrica relacionada à recuperabilidade do parâmetro é igual a $8,8425 \cdot 10^{-6}$. Na direção oposta, as piores estimativas obtiveram uma aproximação mais pobre do sistema real, com $f_{\hat{\delta}} = 0,01765$. Nesse caso, é essencial enfatizar o efeito do ruído inserido nos estados, afetando diretamente o processo de identificação. A média das estimativas a posteriori se aproxima de Γ para este sistema.

3.4 CONSIDERAÇÕES FINAIS

Nesse capítulo foram analisados em profundidade alguns aspectos relacionados ao método de geração de sinal de excitação subótimo e estimativa de parâmetro ideal. Podem ser indicadas algumas informações da análise.

Um número maior de parâmetros nem sempre fornece o melhor resultado. Em outras palavras, o maior número de parâmetros pode até atrapalhar o processo de estimativa. No caso do robô móvel analisado, não há ganho significativo ao usar um modelo completo ao invés do modelo compacto, considerando a análise de número de partículas e incerteza. A causa principal pode ser atribuída à redundância de conjuntos diferentes de parâmetros, que pode confundir o processo. Portanto, tentar usar menos parâmetros pode ser interessante para melhorar o desempenho, com mais confiabilidade, do processo de identificação de parâmetros em modelos não lineares.

O caso em que uma estimativa a priori dos parâmetros está longe dos valores ótimos dos parâmetros viola a hipótese fundamental em que a estratégia de otimização SOESGOPE se baseia. No entanto, aplicar um algoritmo recursivo pode ser uma alternativa suficiente para tolerar essa violação e um conjunto de valores ótimos pode ser atingido.

Mesmo quando um limite de pesquisa de parâmetros é desconhecido, existe a possibilidade de usar os multiplicadores de Lagrange no processo de otimização para aumentar essa barreira. O resultado é obter uma faixa adequada de valores e uma estimativa confiável.

4 MÉTODO HÍBRIDO DE IDENTIFICAÇÃO

Verificando as lacunas do SOESGOPE observadas pelas análises do capítulo 3, o presente capítulo tem o objetivo de introduzir e apresentar uma metodologia híbrida com a capacidade de refinar as saídas do modelo através da inclusão de uma etapa de aprendizado, que possibilita o reconhecimento de comportamentos não modelados. Tais comportamentos são provenientes, principalmente, de diferenças construtivas com relação ao modelo matemático.

Esta proposta foi foco do artigo intitulado *Convolutional System Identification Approach Mixing Optimal Parameter Estimation and Deep Learning* aceito pelo *International Journal of Control, Automation and Systems*. Além da proposição da metodologia, ela foi validada através da aplicação em uma embarcação autônoma com características suficientes para a dificuldade de estimação paramétrica. Os resultados foram promissores, reduzindo o erro entre as saídas do modelo híbrido e as saídas do sistema real.

4.1 INTRODUÇÃO

Método de estimação de parâmetros, em sistemas dinâmicos, é uma técnica eficiente quando se pode assumir uma estrutura modelo. Entretanto, em processos dinâmicos não lineares, cujo interesse está em determinar um modelo para grande parte de sua operação, tal técnica pode se apresentar de forma deficitária. Nessa linha, a utilização de estruturas que não requerem um conhecimento específico do processo pode ser uma alternativa de interesse. As redes neurais artificiais possuem essa qualidade e conseguem caracterizar um sistema através de neurônios matematicamente formulados e parâmetros determinados através de aprendizagem. Assim, é possível encontrar uma estrutura que se adapte da melhor forma possível às saídas do processo em questão sem eliminar qualquer nuance de comportamento, o que na estimação de parâmetros não é possível, uma vez que um modelo matemático é cravado, tendo alterações apenas nos parâmetros.

As análises apresentadas no Capítulo 3 corroboram com essa ideia, já que verifica que a utilização de um número maior de parâmetros não colabora com a eficiência da identificação, indicando que a complexidade afeta na qualidade da identificação. Além disso, um conhecimento a priori suficiente é exigência fundamental para a eficiência de estimação. E quando pensa-se em dificuldades e restrições de testes físicos de identificação, a limitação de geração de trajetórias para fins de comparação é outro fator que deve ser levado em consideração.

Assim, como métodos de estimação paramétrica, o processo de treinamento de uma rede neural com a finalidade de identificação leva em conta o mínimo desvio entre as saídas da rede e as saídas observadas durante um experimento. Dessa forma, para encontrar os parâmetros ou ganhos associados à rede, deve-se aplicar um método de regressão linear,

para redes lineares, obtendo uma solução direta ou algoritmos de otimização iterativos, em redes não lineares. Tanto a técnica Redes Neurais Artificiais (RNA), quanto métodos de aprendizado profundo são baseados no comportamento de aprendizado observado em neurônios humanos. Modelando-os matematicamente e configurando as conexões entre eles, a rede torna-se capaz de descrever relações entre sinais de entrada e saída, o que é possível através de um processo chamado de treinamento ou aprendizado.

Técnicas de aprendizado orientado a dados estão se tornando técnicas extremamente poderosas para aprender essas dinâmicas a partir dos dados, pois muitos sistemas a serem modelados não se é possível extrair uma lei física que descreva de modo satisfatório, seu comportamento. E, devido a essa facilidade de compreensão do comportamento, vários trabalhos na literatura especializada abordam essa técnica (RAJENDRA; BRAHMAJIRAO, 2020).

Com o objetivo de estabelecer um modelo robusto para o movimento de manobra de navios, Wang et al. (2020) propôs uma técnica baseada em regressão por vetores suporte. Através de uma função, a chamada Kernel, o método consegue reduzir o erro através da inserção de vetores que contenham o maior número de registros do experimento real. Essa técnica tem a característica de não permitir que *outliers*, ou seja pontos muito fora da curva, atrapalhem o processo. Mesmo obtendo resultados satisfatórios, os autores perceberam que devido ao conhecimento prévio da modelagem não paramétrica limitado, o processo apresenta *overfitting* (sobre-ajuste) em manobras de giro ou zigue-zague.

De forma similar, Jiang et al. (2022) identificou movimentos de manobra de um navio, entretanto, neste caso, o método foi baseado em rede neural profunda de memória longa de curto prazo, *Long-Short Term Memory* (LSTM). Essa escolha foi feita de forma a levar em consideração informações temporais do movimento, a fim de melhorar a qualidade da identificação. O estudo conseguiu verificar a potencialidade da metodologia, logrando êxito em acompanhar o comportamento definido por um modelo matemático. Além de não contemplar a identificação de um sistema real, o trabalho utiliza manobras simples de zigue-zague e $25^\circ/5^\circ$ para os testes de identificação.

O uso de sinais de excitação para testes de identificação também é investigado em modelos baseados em dados. Zhang, Wang e Zou (2022) apresenta um modelo em caixa preta, construído através de regressão por vetores suporte. Entretanto, diferentemente dos trabalhos citados, os sinais de excitação são o objeto de estudo. Verifica-se, então a superioridade de manobras aleatórias sobre a manobra zigue-zague padrão, amplamente utilizadas na literatura. Entretanto, não houve uma manipulação desse sinal a fim de maximizar a qualidade da identificação.

Para descrever movimentos não lineares complexos e acoplados da dinâmica de um VAS, Woo et al. (2018) apresentou uma estrutura de aprendizado profundo baseada em LSTM e unidades recorrentes com portas fechadas, *Gated Recurrent Units* (GRU). Embora

a metodologia seja capaz de abranger as manobras treinadas, esse tipo de modelagem apresenta deficiência de informação e pode não ser capaz de extrair movimentos aleatórios mais complexos. Além dessa desvantagem, como no caso da identificação baseada apenas na estimativa de parâmetros, ao se considerar modelos experimentais, o efeito de sinais mais ricos em termos de excitação é uma situação a ser analisada, pois o uso de sinais simples causa aprendizado limitado (MORENO-SALINAS et al., 2019).

Neste contexto, podem ser ressaltadas duas situações:

- A necessidade de um conhecimento prévio de um processo pode não ser crucial, entretanto, nota-se que sua utilização pode adicionar informações às lacunas de modelo não paramétricos, aumentando a qualidade da identificação e reduzindo a necessidade de mais dados para treinamento da rede neural.
- Como métodos de identificação baseado em aprendizado necessitam de dados para sua elaboração, é fundamental que esses dados apresentem informações que maximizem a eficiência de captação de comportamentos. Muitas vezes, manobras padrão não conseguem atingir níveis adequados para identificação de alta gama de possíveis atuações do modelo.

Sendo assim, esse capítulo será dedicado para apresentar uma técnica híbrida de identificação que contempla tanto um sub-modelo paramétrico quanto um sub-modelo não paramétrico, que aliados conseguem superar a eficiência de identificação individual de cada um. O sub-modelo paramétrico é baseado no SOESGOPE, que dispõe de uma técnica de otimização para determinação do sinal de entrada, aumentando as possibilidades de identificação de diferentes comportamentos.

4.2 METODOLOGIAS DE IDENTIFICAÇÃO HÍBRIDA

Abordagens combinadas são amplamente utilizadas pela literatura especializada. Tiumentsev e Egorchev (2019) nomeou esse tipo de combinação como modelo semiempírico. Uma qualidade associada a modelos desse tipo é a possibilidade de obter maior acurácia de identificação e conhecer parâmetros aproximados do modelo.

Apesar de um submodelo baseado em modelagem teórica dificilmente conseguir cobrir as características de um processo como um todo, ele se torna uma alternativa para refinar dinâmicas não modeladas (NELLES, 2020). Esse tipo de abordagem é benéfico para o processo de identificação, pois modelos simples podem direcionar os resultados de um segundo modelo orientado a dados. Pode-se configurar o modelo híbrido através de uma abordagem paralela ou serial. A abordagem paralela envolve a utilização de um modelo caixa branca e um modelo caixa preta que minimiza o resíduo entre a saída do modelo matemático e saída real. Já a abordagem serial, utiliza dois modelos em série, um

caixa branca e outro caixa preta. Ele é utilizado de forma a aprimorar uma parte não conhecida do sistema. Dessa forma, o modelo de caixa branca prediz uma saída para um determinado conjunto de parâmetros de entrada e o modelo de caixa preta corrige a saída prevista.

Com resultados consideráveis e superiores aos encontrados com técnicas desacopladas, (MACIEL et al., 2020) propôs a combinação de um modelo matemático com uma arquitetura de aprendizado profundo baseada em Conv3D-LSTM para a previsão de dados hidrológicos. Nesse contexto, o termo híbrido, usado para o modelo, pode ser entendido como uma modelagem que usa tanto uma característica matemática (caixa branca) quanto uma forma de caixa preta para identificar o sistema.

An et al. (2022) propôs um método para identificação do sistema dinâmico. O modelo dinâmico do sistema é composto pelo modelo linear nominal e pelo modelo não linear. Os parâmetros do modelo linear invariante no tempo do sistema são identificados pela resposta no domínio da frequência. Já com relação aos termos diferenciais de alta ordem e erros incertos da parte não linear do modelo dinâmico, uma rede LSTM é introduzida para identificar as características não lineares do sistema.

O trabalho de Egorchev e Tiumentsev (2018b) mostrou que o conhecimento prévio sobre o objeto de simulação é eficaz para o algoritmo de aprendizado sequencial. O processo de desenvolvimento para o modelo baseado em rede neural semi-empírica do sistema dinâmico consistiu nas seguintes etapas:

- Desenvolvimento de modelo teórico de tempo contínuo para o sistema dinâmico considerado bom como aquisição de dados experimentais sobre o comportamento do sistema;
- Avaliação da precisão do modelo teórico do sistema dinâmico usando dados coletados;
- Conversão do modelo original de tempo contínuo em um modelo de tempo discreto;
- Representação aprendizagem para o modelo e RNA;
- Ajuste estrutural do modelo em RNA para atender aos requisitos de precisão da modelagem.

O chamado modelo semi-híbrido mostrou-se eficiente para identificar o movimento angular de aeronaves com a previsão de algumas amostras à frente (EGORCHEV; TIUMENTSEV, 2018a). A aplicação em sistema dinâmicos complexos é altamente requerida. Entretanto o conjunto de dados de treinamento deve ser suficiente para conseguir formar uma rede que replique o comportamento final. Essa tarefa é extremamente importante para obter um modelo de sistema dinâmico autêntico. Os dados de treinamento necessários para o modelo caixa preta gerado podem ser obtidos por meio de excitações de

teste especialmente organizadas para o sistema simulado. Egorchev e Tiumentsev (2017) propôs uma técnica para melhorar a eficiência de treinamento para identificação de uma aeronave. O procedimento sintetiza ações de controle que providenciam um conjunto de dados representativo. Embora o resultado tenha se mostrado melhor do que a inserção aleatória de excitação, esse sinal é determinado através de um algoritmo que apenas insere elementos considerados novos para o treinamento, sem qualquer análise da suficiência de excitação.

4.3 SOESGOPE HÍBRIDO

Além de um grande número de parâmetros que influenciam na solução final, deve-se levar em conta também possíveis desvios do modelo matemático em relação ao projeto real. Portanto, deve-se considerar a determinação de uma metodologia que possa identificar sistemas dinâmicos mesmo em situações em que o modelo matemático não seja fiel ao comportamento real. A aprendizagem profunda mostra-se como uma importante ferramenta por possuir um caráter empírico, o que não limita a técnica a uma concepção puramente teórica. Assim, o modelo é formado pelas peculiaridades do comportamento do sistema dinâmico utilizado como dados de treinamento. No entanto, é sabido que o bom funcionamento de uma rede neural requer uma grande quantidade de dados de treinamento (TIUMENTSEV; EGORCHEV, 2019). Ao pensar em sistemas complexos, pode não ser possível obter um conjunto de dados razoável, o que dificulta o desempenho da ferramenta.

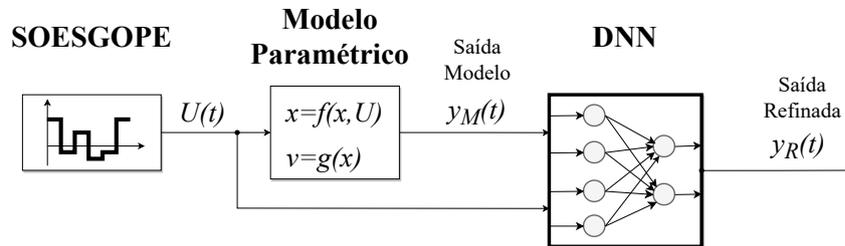
A estimação de parâmetros é um método que pode lidar com essa possível falta de dados. Por exemplo, um sinal de entrada projetado pelo SOESGOPE é capaz de excitar o sistema para otimizar o processo de extração dos parâmetros do modelo. No entanto, a estimação paramétrica tem a desvantagem de não poder detectar um possível erro de modelo. Assim, combinar os dois métodos pode ser altamente eficaz no processo de identificação final, pois é possível combinar as duas características:

- Capturar peculiaridades comportamentais por meio da aplicação de DNN em série com o modelo paramétrico;
- Direcionar o comportamento viável devido à modelagem matemática de um experimento com sinal de excitação subótimo. Este modelo adicional deve ser capaz de lidar com informações de tempos passados para determinar os estados da embarcação.

Para uma melhor compreensão, a estrutura da construção do modelo híbrido proposto pode ser dividida em três etapas principais: (i) geração do sinal e estimação de parâmetros; (ii) modelo paramétrico; (iii) correção de comportamento. A Figura 10 ilustra a relação entre essas etapas que serão descritas a seguir, a combinação entre o

modelo paramétrico e a rede neural profunda usando o sinal subótimo da metodologia SOESGOPE.

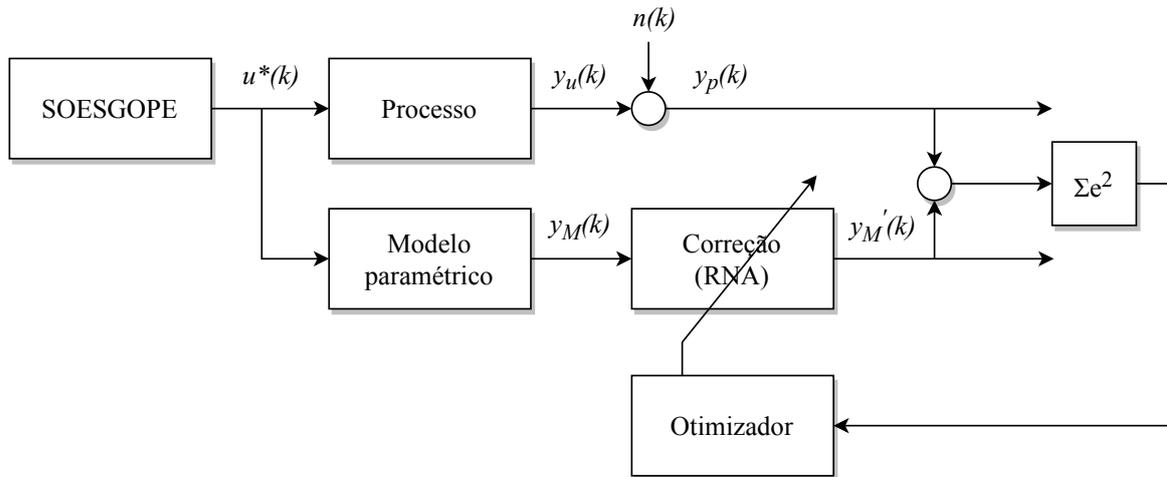
Figura 10 – Diagrama esquemático da estrutura híbrida: modelo paramétrico e rede neural profunda (RNP)



Fonte: Elaborado pelo autor

- (i) **Geração do sinal e estimação de parâmetros:** Nessa primeira etapa, o processo SOESGOPE é performed por completo, isto é, é realizado o processo de geração do sinal de excitação, baseado em APRBS, e a estimação propriamente dita. No processo de geração do sinal de entrada, a comparação é realizada com um modelo perturbado, conforme explicações anteriores. E a estimação é realizada após um experimento que utiliza o sinal como entrada para o sistema real. Ao fim dessa etapa, são fornecidos para a continuidade do processo, os parâmetros do modelo, bem como o sinal de excitação.
- (ii) **Modelo paramétrico:** Após a execução da etapa anterior, o modelo paramétrico pode ser formado através dos parâmetros encontrados. Entretanto, devido à alguma descontinuidade do modelo com relação ao sistema real, ele não descreve fielmente o comportamento real. Entretanto, tal modelo é uma referência válida, visto que seu conjunto de equações é uma aproximação das saídas do sistema dinâmico. Nesta etapa, portanto, será utilizado o sinal como entrada para o modelo e suas saídas serão computadas.
- (iii) **Correção de comportamento:** As saídas do modelo paramétrico direcionam o comportamento do sistema real. Assim, elas são utilizadas como entradas do processo de correção. A Figura 11 apresenta o diagrama esquemático da modelagem híbrida, enfocando o processo de correção. Nesta etapa, o objetivo principal é corrigir erros relacionados à modelagem paramétrica e, portanto, utiliza as saídas do mesmo como entradas do processo de aprendizagem. Nesse caso, a rede neural artificial detecta um padrão para correção entre as saídas. Ressalta-se que neste processo, a comparação se dá entre as saídas observadas no modelo paramétrico, $y_M(k)$, e as saídas observadas no sistema real, representado na figura como processo, $y_p(k)$, frente a uma mesma entrada, $u(k)$, o sinal de excitação fornecido pelo SOESGOPE.

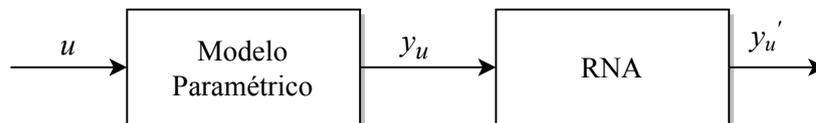
Figura 11 – Diagrama esquemático do processo em série: estimação de parâmetros e correção do comportamento



Fonte: Elaborado pelo autor

O modelo final é determinado pelo modelo paramétrico em série com a rede neural resultante do processo de treinamento. O modelo híbrido pode ser utilizado inserindo um sinal de entrada u e, será possível a obtenção da saída corrigida y'_u . A Figura 12 apresenta a estrutura do modelo híbrido (modelo paramétrico e aprendizado profundo) a ser utilizado como padrão do sistema.

Figura 12 – Estrutura do modelo híbrido resultante



Fonte: Elaborado pelo autor

Nesse processo, o sistema de aprendizado funciona como um corretor de comportamentos que o modelo identificado por apenas estimação paramétrica não tem a capacidade de identificar. Entretanto, o modelo paramétrico é um aliado do aprendizado, visto que, ele pode direcionar o resultado para um comportamento viável. É importante ressaltar que embora o comportamento não seja idêntico ao modelo matemático descrito, existe uma compatibilidade entre os estados resultantes. Dessa forma, existe um grande benefício na associação de uma técnica de estimação de parâmetros com a identificação por métodos de aprendizagem.

4.3.1 Redes Neurais Recorrentes

Usando as entradas do modelo paramétrico e o sinal subótimo, a rede neural profunda tem o objetivo de tenta prever os estados do próximo ponto no tempo. No entanto, é importante ressaltar que o uso de informações de momentos anteriores pode

criar uma dependência temporal entre os estados de saída do modelo híbrido. Assim, a utilização de uma estrutura Redes Neurais Profundas (RNP) que computa e armazena estados anteriores de séries temporais é uma alternativa interessante. A rede neural recorrente é uma classe de rede que contém iterações na camada oculta para reter as informações no intervalo de tempo anterior para prever o valor do intervalo de tempo atual.

Como na rede neural recorrente a decisão tomada no tempo $t - 1$ afeta a decisão no tempo t , a decisão de como a rede responderá a novos dados depende de duas coisas: a entrada atual e a saída do passado recente. A rede é capaz de calcular iterativamente as suas através das equações 4.1 e 4.2.

$$h_t = \mathcal{H}(\mathbf{W}_{xh}x_t + \mathbf{W}_{hh}h_{t-1} + b_h) \quad (4.1)$$

$$y_t = \mathbf{W}_{hy}h_t + b_y \quad (4.2)$$

Sendo que x são as entradas, y é a sequência de saída, h é a sequência do vetor oculto nas fatias de tempo $t = 1$ a T . \mathbf{W} representa as matrizes de peso e b representa o *bias*. \mathcal{H} é a função de ativação usada nas camadas ocultas.

4.3.1.1 Memória Longa de Curto Prazo (LSTM)

Como solução para as deficiências das RNA normais, Hochreiter e Schmidhuber (1997) criaram as redes LSTM. A arquitetura especial da célula de memória no LSTM facilita o armazenamento de informações por um longo período. Sua formulação original é baseada nas seguintes equações:

$$f_t = \sigma(\mathbf{W}_f[h_{t-1}, x_t] + b_f) \quad (4.3)$$

$$i_t = \sigma(\mathbf{W}_i[h_{t-1}, x_t] + b_i) \quad (4.4)$$

$$C'_t = \tanh(\mathbf{W}_C[h_{t-1}, x_t] + b_C) \quad (4.5)$$

$$C_t = f_t C_{t-1} + i_t C'_t \quad (4.6)$$

$$o_t = \sigma(\mathbf{W}_o[h_{t-1}, x_t] + b_o) \quad (4.7)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (4.8)$$

em que σ é a função sigmoide, que é usada para formar três portas na célula de memória. \tanh é a função tangente hiperbólica, usada para escalar a saída de uma célula de memória específica. i , f , o , C , C' são o portão de entrada, portão de esquecimento, portão de saída, conteúdo da célula de memória e novo conteúdo da célula de memória, respectivamente.

4.3.1.2 Unidade Recorrente com Portas Fechadas (GRU)

A unidade recorrente com portas fechadas, GRU, proposta por Cho et al. (2014), é semelhante à LSTM, mas possui menos parâmetros. Nesse tipo de rede há unidades fechadas como LSTMs que controlam o fluxo de informações dentro da unidade, mas sem células de memória separadas. Ao contrário do LSTM, o GRU não possui porta de saída, expondo assim todo o seu conteúdo. A formulação GRU pode ser dada pelas seguintes equações:

$$r_t = \text{sigm}(\mathbf{W}_{xr}x_t + \mathbf{W}_{hr}h_{t-1} + b_r) \quad (4.9)$$

$$z_t = \text{sigm}(\mathbf{W}_{xz}x_t + \mathbf{W}_{hz}h_{t-1} + b_z) \quad (4.10)$$

$$h'_t = \tanh(\mathbf{W}_{xz}x_t + \mathbf{W}_{hh}(r_t \odot h_{t-1} + b_h)) \quad (4.11)$$

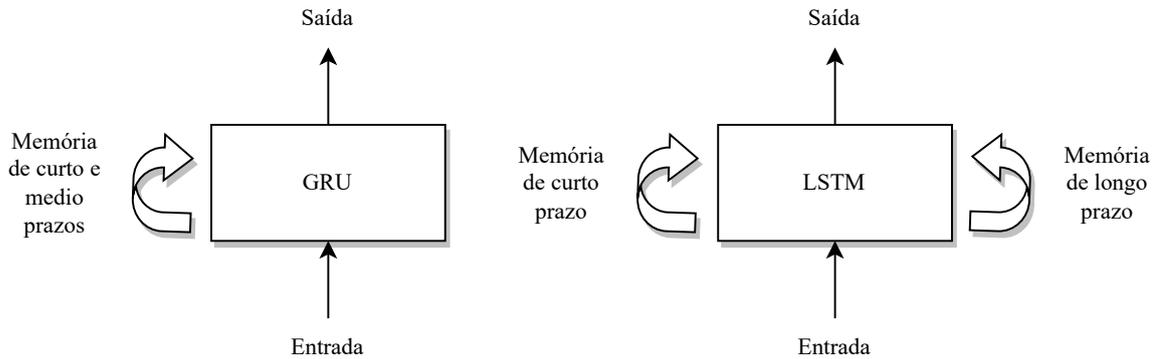
$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot h'_t \quad (4.12)$$

com r_t , z_t , x_t , h_t sendo, respectivamente, o portão de *reset*, portão de atualização, vetor de entrada e vetor de saída. Semelhante ao LSTM, W denota as matrizes de peso, b o *bias*, sigm é a ativação sigmoide e \tanh é a função de ativação da tangente hiperbólica

As redes do tipo GRU visam resolver o problema de dissipação de gradiente, comum em uma rede neural recorrente padrão. As camadas que recebem uma pequena atualização de gradiente param de aprender. Como essas camadas não aprendem, a estrutura pode esquecer o que foi visto em sequências mais longas e, assim, ter uma memória de curto prazo. Os mecanismos internos desse tipo de rede têm a capacidade de filtrar seletivamente informações irrelevantes e reter apenas o que é útil.

A LSTM tem dois estados diferentes passados entre as células, o estado da célula e o estado oculto, que carregam a memória de longo e curto prazo, respectivamente. Já as GRU têm apenas um estado oculto transferido entre as etapas do tempo. Esse estado oculto é capaz de manter as dependências de longo e curto prazo ao mesmo tempo. Fato devido aos cálculos pelos quais o estado oculto e os dados de entrada passam. A Figura 13 ilustra esse comportamento iterativo das redes recorrentes.

Figura 13 – Processo iterativo das redes GRU e LSTM



Fonte: Elaborado pelo autor

Os portões, nas redes GRU, são vetores contendo valores entre 0 e 1 que serão multiplicados com os dados de entrada ou estado oculto. Um valor 0 nos vetores indica que os dados correspondentes no estado de entrada ou oculto não são importantes e, portanto, retornarão como zero. Por outro lado, um valor 1 no vetor significa que os dados correspondentes são importantes e serão usados.

4.4 CONSIDERAÇÕES FINAIS

Em identificação de sistemas dinâmicos, a confiabilidade é um fator fundamental. Assim, a utilização de métodos que consigam refinar o modelo com o objetivo de minimizar o erro observado entre as saídas real e estimada se torna instigante. Recorrer a metodologias baseadas em estimação paramétrica é comum em processos que possuem estruturas matemáticas bem definidas, sendo necessário calibrar os parâmetros do modelo. Entretanto, essa atitude pode não ser ideal em sistema mais complexos, ou com um número elevado de parâmetros, ou com estruturas parcialmente não modeláveis, que criam condições inadequadas para a estimação paramétrica. Nesse sentido, lançar mão de outros métodos deve ser requerido.

Nesse capítulo, foi proposto um método de refinamento que é incluído em cascata à etapa de estimação de parâmetros. Baseado em aprendizagem, dispõe uma rede neural profunda, a nova etapa incluída é capaz de reduzir o erro do modelo com relação às saídas reais do sistema. Essa nova etapa, além de ser benéfica no âmbito de confiabilidade, a associação da estrutura híbrida com um sinal de excitação construída para favorecer a identificação de comportamentos leva à construção de um modelo que é capaz de lidar com diferentes perfis de sinais de controle.

A capacidade de processamento de informações passadas das redes neurais recorrentes é um fator que faz com que esse tipo de estrutura possa ser utilizada a fim de cumprir com a etapa de refinamento. Nesse caso, o modelo paramétrico pode servir, não

apenas como um ponto de partida para a RNP, responsável pelo refinamento, mas também como entradas que correlacionam uma temporalidade entre instantes anteriores de saídas. Outra fonte de informação importante para a rede é o sinal de excitação, que é aplicado um instante antes do desejado. Como a estrutura de identificação é utilizada em uma atualização de posição, ele também pode ser utilizado como entrada da rede com foco na correção da saída do modelo paramétrico.

5 SISTEMA DE IDENTIFICAÇÃO HÍBRIDO - VAS

Esse capítulo é dedicado a apresentar um sistema para validação da metodologia híbrida de identificação de sistemas dinâmicos. O sistema proposto, para tanto, possui características relevantes de complexidade e não adequação ao modelo paramétrico comum, que exige apenas uma etapa de estimação paramétrica.

O sistema de identificação proposto baseado em etapas matemática e aprendizado, respectivamente é promissor para a aplicação em sistemas complexos. O seu potencial está, principalmente concentrado, quando o sistema dinâmico não consegue ser completamente modelado através de equações matemáticas. Dessa forma, é necessário que se valide a metodologia em processos que possuem tal característica. Nesse contexto, barcos autônomos são uma fonte de pesquisa que contempla a complexidade requerida para análise mais profunda das metodologias. A modelagem de um VAS será apresentada como forma de avaliar a aplicabilidade da pesquisa.

5.1 VEÍCULO AUTÔNOMO DE SUPERFÍCIE

Um grande desafio para a identificação de sistemas dinâmicos é lidar com sistemas cuja modelagem é complexa e onde diversos focos de possíveis inconsistências construtivas podem ser detectados. O catamarã apresentado por Silva et al. (2020) e Regina et al. (2021) apresenta diversas características que fogem de um modelo simplificado (Figura 14).

Figura 14 – Catamarã autônomo com propulsão aérea



Fonte: Elaborada pelo autor

A embarcação foi desenvolvida pelo Grupo de Robótica Inteligente (GRIn) da Universidade Federal de Juiz de Fora no âmbito do projeto intitulado "Levantamento automatizado e análise de dados inteligente para medição de descargas líquidas e sólidas". Dentre os objetivos finais do projeto, um deles seria a construção de dois veículos autônomos,

sendo um responsável pela aquisição das descargas líquidas, sólidas e monitoramento da sedimentação do rio e outro para a coleta de dados referente à qualidade da água. No caso em que se deve monitorar as descargas líquidas, um fator que se deve ser levado em consideração é o impacto do casco e motores em contato com a superfície da água. Para mitigar esse efeito e obter uma leitura mais confiável, a embarcação foi construída no formato de catamarã, que possui uma estabilidade maior e os motores escolhidos foram aéreo, a fim de evitar vórtices de água que impactariam na aquisição de dados pelos sensores.

Dessa forma, seu design com modelagem desafiadora foi proposto devido à sua função: permitir inspeções em áreas críticas. O chamado ASV AERO4River possui características especiais: (i) Propulsão aérea com atuadores acima da linha d'água; (ii) Hiperatuação: 8 atuadores, sendo 4 servos direcionais e 4 motores de propulsão; (iii) Controle em três graus de liberdade: *surge*, *sway* e *yaw*; (iv) Projeto do casco no formato de catamarã para permitir maior estabilidade.

Para avaliar o desempenho de um VAS, é desejável o conhecimento de um modelo completo que descreva seu comportamento frente aos diversos sinais de entrada. Esse comportamento, por sua vez, requer uma investigação prévia de um modelo matemático preciso que contemple parâmetros de sistema razoáveis. Geralmente, um modelo padrão consiste na descrição cinemática e dinâmica da embarcação. O direcionamento da modelagem é fornecido por Fossen (1994). Entretanto, tal modelo foi consolidado para embarcações com casco simples. Assim, deve-se adaptá-lo. Tal adaptação será realizada pelo modelo híbrido que possui a capacidade de corrigir algum tipo de erro de modelagem matemática.

Com relação ao controle, o objetivo é apenas garantir que o barco acompanhe o caminho desejado com a maior precisão possível. Dessa forma, o modelo geral de seis graus de liberdade pode ser reduzido para considerar somente o movimento nas categorias de avanço (*surge*), deriva (*sway*) e guinada (*yaw*), enquanto a dinâmica associada aos outros movimentos de afundamento (*heave*), arfagem (*pitch*) e jogo (*roll*) serão negligenciados para manter a simplicidade do modelo.

5.1.1 Modelagem Cinemática

O modelo cinemático de um VAS típico de três graus de liberdade sem a presença de distúrbios externos pode ser expresso como:

$$\dot{\boldsymbol{\eta}} = \mathbf{R}(\psi)\boldsymbol{\nu} \quad (5.1)$$

onde $\boldsymbol{\eta} = [x, y, \psi]^T$ representa o vetor de posição e orientação de acordo com o referencial inercial e $\boldsymbol{\nu} = [u, v, r]^T$ denota o vetor de velocidade linear e angular de acordo com o

referencial inercial. $\mathbf{R}(\psi)$ representa a matriz de rotação entre o referencial inercial e o referencial fixo à embarcação, dada por:

$$\mathbf{R}(\psi) = \begin{bmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5.2)$$

5.1.2 Modelagem Dinâmica

O modelo cinemático não é suficiente para o desenvolvimento de um controle para o barco, pois nesse modelo apenas é descrito a movimentação da embarcação de acordo com a velocidade atingida. O modelo dinâmico descreve como o veículo se comporta frente a aplicação de forças externas e pelos atuadores. Considerando apenas as forças aplicadas diretamente à embarcação e ignorando as ações de ondas, ventos e outros distúrbios que possam acometer a embarcação, o modelo pode ser descrito como a seguir.

$$\mathbf{M}\dot{\boldsymbol{\nu}} + \mathbf{C}(\boldsymbol{\nu})\boldsymbol{\nu} + \mathbf{D}\boldsymbol{\nu} + \mathbf{D}_n\boldsymbol{\nu}(\boldsymbol{\nu}) = \boldsymbol{\tau} \quad (5.3)$$

sendo que os seguinte símbolos representam:

- \mathbf{M} Matriz de inércia do sistema;
- $\mathbf{C}(\boldsymbol{\nu})$ Coriolis e matriz centrípeta;
- \mathbf{D} Matriz de amortecimento linear;
- \mathbf{D}_n Matriz de amortecimento não linear;
- $\boldsymbol{\nu}$ Vetor de velocidades lineares e angular;
- $\boldsymbol{\eta}$ Vetor de posição e orientação;
- $\boldsymbol{\tau}$ Forças e momentos aplicados a embarcação.

A matriz de inércia dos sistema é a soma da matriz de inércia do sistema do corpo rígido, \mathbf{M}_{RB} , com a matriz de massa adicionada, \mathbf{M}_A .

$$\mathbf{M} = \mathbf{M}_{RB} + \mathbf{M}_A \quad (5.4)$$

sendo que:

$$\mathbf{M}_{RB} = \begin{bmatrix} m & 0 & 0 \\ 0 & m & mx_g \\ 0 & mx_g & Iz \end{bmatrix} \quad (5.5)$$

$$\mathbf{M}_A = \begin{bmatrix} -X_{\dot{u}} & -X_{\dot{v}} & -X_{\dot{r}} \\ -Y_{\dot{u}} & -Y_{\dot{v}} & -Y_{\dot{r}} \\ -N_{\dot{u}} & -N_{\dot{v}} & -N_{\dot{r}} \end{bmatrix} \quad (5.6)$$

e os seguintes símbolos representam:

m Massa total da embarcação;

x_g Centro de gravidade da embarcação em relação ao eixo x ;

I_z Momento de inércia em relação ao eixo z da embarcação;

$T_{\dot{\nu}}$ Derivados hidrodinâmicos, sendo que $T = [X, Y, N]$ e $\nu = [u, v, r]$.

A matriz centrípeta Coriolis é dada pela soma da matriz de Coriolis do corpo rígido, $\mathbf{C}_{RB}(\nu)$, com a matriz de Coriolis hidrodinâmica, $\mathbf{C}_A(\nu)$:

$$\mathbf{C}(\nu) = \mathbf{C}_{RB}(\nu) + \mathbf{C}_A(\nu) \quad (5.7)$$

sendo que:

$$\mathbf{C}_{RB}(\nu) = \begin{bmatrix} 0 & 0 & -m(x_g r + v) \\ 0 & 0 & m u \\ m(x_g r + v) & -m u & 0 \end{bmatrix} \quad (5.8)$$

$$\mathbf{C}_A(\nu) = \begin{bmatrix} 0 & 0 & Y_{\dot{v}} v + Y_{\dot{r}} r \\ 0 & 0 & -X_{\dot{u}} u \\ -Y_{\dot{v}} v - Y_{\dot{r}} r & X_{\dot{u}} u & 0 \end{bmatrix} \quad (5.9)$$

As matrizes 5.6 e 5.9 são relacionadas à ação das forças hidrodinâmicas na embarcação.

A matriz de amortecimento é dada pela matriz de amortecimento linear, \mathbf{D} e não linear \mathbf{D}_n conforme descrito abaixo:

$$\mathbf{D} = \begin{bmatrix} -X_u & -X_v & -X_r \\ -Y_u & -Y_v & -Y_r \\ -N_u & -N_v & -N_r \end{bmatrix} \quad (5.10)$$

$$\mathbf{D}_n(\nu) = \begin{bmatrix} -X_{|u|u}|u| & -X_{|u|v}|u| & -X_{|u|r}|u| \\ -Y_{|u|u}|v| & -Y_{|u|v}|v| & -Y_{|u|r}|v| \\ -N_{|u|u}|r| & -N_{|u|v}|r| & -N_{|u|r}|r| \end{bmatrix} \quad (5.11)$$

sendo que $X_u, X_v, X_r, Y_u, Y_v, Y_r, N_u, N_v$ e N_r representam os parâmetros relacionados ao amortecimento linear e os outros ao não linear. O amortecimento hidrodinâmico surge devido às forças lineares e não lineares, relacionadas com: atrito superficial linear e não linear e circulação linear de água ao redor do casco e fluxo cruzado. Cada parâmetro está relacionado a esses fenômenos em seus respectivos graus de liberdade.

5.1.3 Forças de propulsão

As atuações envolvidas na embarcação englobam propulsores azimutais. Esse tipo de propulsor é capaz de rotacionar em um ângulo α em torno do eixo z e produzir dois componentes de força (F_x, F_y) no plano horizontal. Dessa forma ele consegue produzir forças em diferentes direções. As forças de controle podem ser expressas por:

$$\mathbf{f} = \mathbf{K}\mathbf{u} \quad (5.12)$$

sendo que $\mathbf{u} = [u_1, \dots, u_r]$ é o vetor de entradas de controle e $\mathbf{K} \in R^{r \times r}$ é a matriz diagonal de coeficientes de força dada por:

$$K = \text{diag}\{K_1, \dots, K_r\} \quad (5.13)$$

As forças e momentos de atuação são dadas por:

$$\boldsymbol{\tau} = \mathbf{T}(\boldsymbol{\alpha})\mathbf{f} = \mathbf{T}(\boldsymbol{\alpha})\mathbf{K}\mathbf{u} \quad (5.14)$$

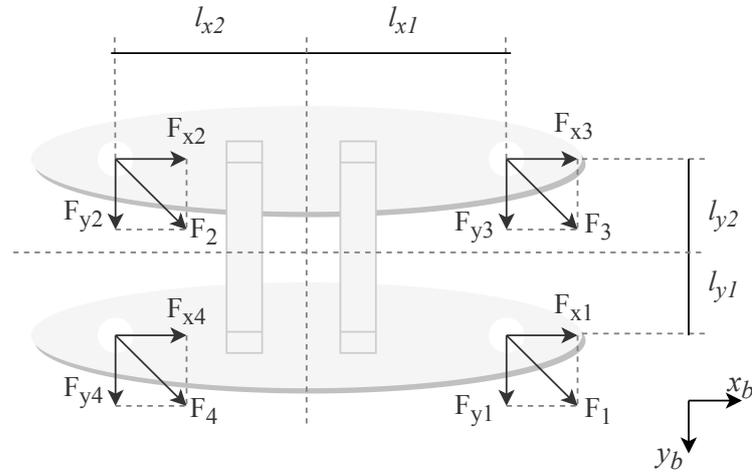
sendo que $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_p]^T \in R^p$ é o vetor de ângulos dos propulsores azimutais e $\mathbf{T}(\boldsymbol{\alpha}) \in R^{n \times r}$ é a matriz de alocação de forças.

No caso do catamarã, uma embarcação com quatro propulsores azimutais deve ser considerada, conforme a Figura 15.

As forças e momentos X, Y e N em avanço, deriva e guinada, respectivamente, para a embarcação supracitada, satisfazem a Equação 5.15.

$$\begin{bmatrix} X \\ Y \\ N \end{bmatrix} = \begin{bmatrix} \cos(\alpha_1) & \cos(\alpha_2) & \cos(\alpha_3) & \cos(\alpha_4) \\ \text{sen}(\alpha_1) & \text{sen}(\alpha_2) & \text{sen}(\alpha_3) & \text{sen}(\alpha_4) \\ l_y \cos(\alpha_1) - l_x \text{sen}(\alpha_1) & -l_y \cos(\alpha_2) + l_x \text{sen}(\alpha_2) & -l_y \cos(\alpha_3) - l_x \text{sen}(\alpha_3) & l_y \cos(\alpha_4) + l_x \text{sen}(\alpha_4) \end{bmatrix} \begin{bmatrix} K_1 & 0 & 0 & 0 \\ 0 & K_2 & 0 & 0 \\ 0 & 0 & K_3 & 0 \\ 0 & 0 & 0 & K_4 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \quad (5.15)$$

Figura 15 – Esquema de forças de um catamarã equipado com quatro propulsores azimutais (forças F_1 , F_2 , F_3 e F_4) com forças descompostas em x e y



Fonte: Elaborado pelo autor

Nesse caso, a matriz de alocação de forças $T(\alpha)$ é não linear. Isso implica em uma solução através de métodos de otimização. De forma a evitar isso, os propulsores azimutais podem ser tratados considerando duas forças decompostas em x e y . Assim, pode-se convencionar:

$$F_{xi} = F_i \cos(\alpha_i) \quad (5.16)$$

$$F_{yi} = F_i \sin(\alpha_i). \quad (5.17)$$

O vetor de forças estendido pode ser definido por:

$$f_e := K_e u_e \quad (5.18)$$

sendo que:

$$\tau = T_e K_e u_e \quad (5.19)$$

onde T_e e K_e são as matrizes de alocação e de coeficientes estendidas, respectivamente e u_e é o vetor de entradas de controle estendido, sendo que os controles azimutais são dados pela Equações 5.20 e 5.21.

$$u_{ix} = u_i \cos(\alpha_i) \quad (5.20)$$

$$u_{iy} = u_i \sin(\alpha_i) \quad (5.21)$$

Considerando o modelo estendido pode-se relacionar as forças e momento X, Y e N conforme a Equação matricial 5.22.

$$\begin{bmatrix} X \\ Y \\ N \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ l_y & -l_x & -l_y & l_x & -l_y & -l_x & l_y & l_x \end{bmatrix} \cdot \begin{bmatrix} K_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & K_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & K_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & K_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & K_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & K_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & K_4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & K_4 \end{bmatrix} \cdot \begin{bmatrix} u_{1x} \\ u_{1y} \\ u_{2x} \\ u_{2y} \\ u_{3x} \\ u_{3y} \\ u_{4x} \\ u_{4y} \end{bmatrix} \quad (5.22)$$

Considerando T_e constante, o vetor de entradas de controle u_e pode ser resolvido diretamente usando a abordagem de pseudo-inversa. Esse não é o caso de $T(\alpha)$, que representa um problema de otimização não linear. Se u é computado através da pseudo-inversa, é possível encontrar os ângulos α_i de acordo com a Equação 5.23.

$$u_i = \sqrt{u_{ix}^2 + u_{iy}^2}, \quad \alpha_i = \text{atan2}(u_{iy}, u_{ix}) \quad (5.23)$$

5.1.4 Parâmetros Envolvidos

De acordo com a modelagem abordada, existe uma variedade de parâmetros envolvidos. Com relação matriz de inércia e de Coriolis de corpo rígido, são relacionados os seguintes 3 parâmetros dimensionais:

$$\mathbf{\Gamma}_{MRB} = [m, x_g, I_z]$$

sendo que m , x_g e I_z são facilmente mensuráveis através de equipamentos de medição e fatores de projeto. Entretanto, o processo de identificação consegue fornecer valores mais confiáveis, visto que existem diferenças entre o projeto e a construção.

No âmbito das forças da massa adicionada, são relacionados os seguintes 9 parâmetros adimensionais:

$$\mathbf{\Gamma}_{MA} = [X_{\dot{u}}, X_{\dot{v}}, X_{\dot{r}}, Y_{\dot{u}}, Y_{\dot{v}}, Y_{\dot{r}}, N_{\dot{u}}, N_{\dot{v}}, N_{\dot{r}}]$$

que são os derivados hidrodinâmicos. Eles podem ser previamente calculados de acordo com características físicas do casco e sua ação no fluido (FOSSSEN, 1994). Esses cálculos são aproximados e o experimento de identificação fornece valores mais precisos, utilizando-os como parâmetros a priori. Os derivados hidrodinâmicos indicam a força da massa adicionada hidrodinâmica T em cada eixo devido à uma aceleração \dot{v} .

Ainda são relacionados outros 5 parâmetros adimensionais associados à matriz de amortecimento linear e não linear, respectivamente:

$$\mathbf{\Gamma}_D = [X_u, X_v, X_r, Y_u, Y_v, Y_r, N_u, N_v, N_r]$$

$$\mathbf{\Gamma}_{D_n} = [X_{|u|u}, X_{|u|v}, X_{|u|r}, Y_{|u|u}, Y_{|u|v}, Y_{|u|r}, N_{|u|u}, N_{|u|v}, N_{|u|r}]$$

e assim como os derivados hidrodinâmicos, os valores desses parâmetros podem ser previamente calculados de acordo com a taxa de amortecimento. Entretanto, tais parâmetros possuem uma dificuldade a mais de determinação, visto que são necessários conhecimentos prévios e análises através de programas computacionais especializados.

Além dos parâmetros associados à modelagem do barco, ainda pode-se relacionar os 4 parâmetros envolvidos na ação dos atuadores:

$$\mathbf{\Gamma}_F = [K_1, K_2, K_3, K_4]$$

que podem ser estimados através de ensaios de atuação. Entretanto é possível estimá-los também através do processo de estimação experimental.

Portanto, de uma forma simplificada, o veículo pode ser modelado atrelando 34 parâmetros. Embora seja possível calcular e mensurar alguns desses parâmetros, tais cálculos podem não contemplar fatores construtivos. Além disso, muitas vezes, tais cálculos são complexos, envolvendo alta carga de esforço computacional, impossibilitando a aferição de valores precisos. Assim, a utilização de um método de estimação baseado em experimento pode ser altamente requerida.

5.2 ARQUITETURA DE IDENTIFICAÇÃO - VAS

A embarcação mencionada possui características que fogem do padrão de modelos matemáticos definidos pela literatura. A identificação, portanto, pode ser adaptada a fim de extrair informações mais confiáveis da modelagem. A proposta de identificação híbrida possui características que vão ao encontro com os requisitos de modelagem de tipos complexos. Nesta seção, a metodologia de correção e refinamento do modelo teórico baseado em aprendizado profundo recorrente será descrita no âmbito do ASV AERO4River.

5.2.1 Aplicação do SOESGOPE

Para mitigar as desvantagens na utilização do SOESGOPE no ASV AERO4River, associada principalmente ao grande número de parâmetros e a falta de informações a priori, Neto et al. (2021) adequou a metodologia, incorporando uma estratégia robusta,

a chamada *Robust Sub-Optimal Excitation Signal Generation and Optimal Parameter Estimation* (rSOESGOPE). O principal objetivo é aumentar a faixa de excitação do sinal de entrada para aumentar a confiabilidade e robustez do método. O SOESGOPE robusto difere do método original por usar múltiplos sinais subótimos e impor restrições de penalidade apenas nos estados. Essa solução tem como foco fornecer uma maior quantidade de informações para que o processo de otimização encontre parâmetros que determinem uma alta similaridade com o sistema real. Tal abordagem reduziu o erro quadrático médio das saídas, principalmente o erro relacionado às velocidades lineares em comparação com a metodologia tradicional.

Nesse cenário, e verificando o impacto da utilização de múltiplos sinais subótimos no resultado do modelo, tal metodologia será incorporada a fim de melhorar a confiabilidade do primeiro passo do modelo híbrido, determinando entradas mais próximas da realidade, aumentando a eficiência do segundo passo, destinado ao refinamento.

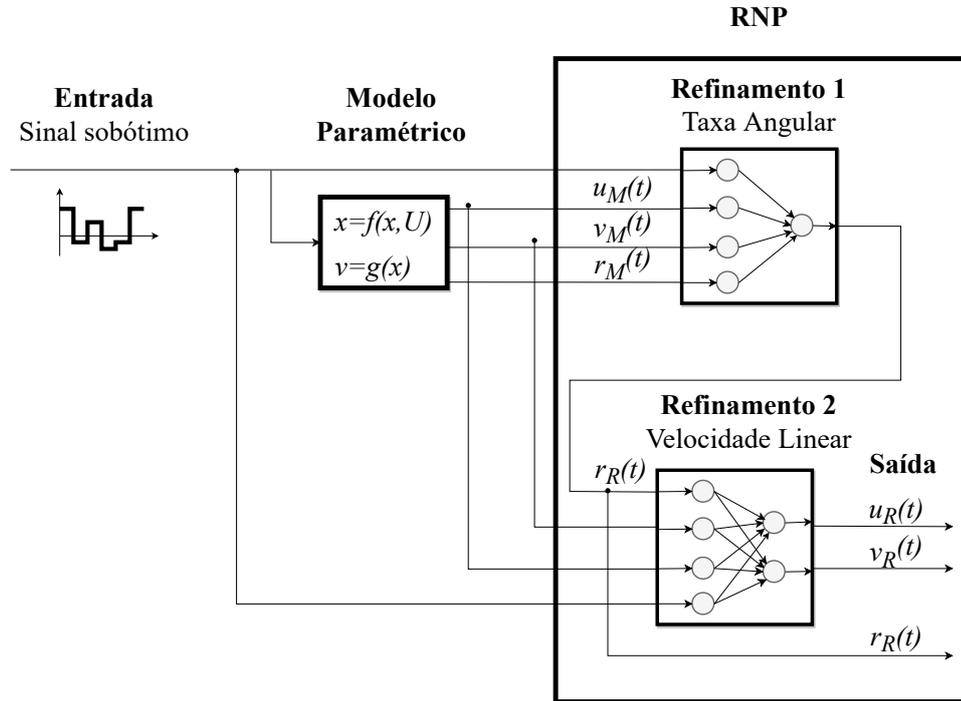
5.2.2 Arquitetura da RNP

O foco, neste passo, está no refinamento das saídas com base nos dados obtidos nas etapas anteriores. O modelo paramétrico estimado na etapa anterior é de extrema importância, pois servirá de base para o novo bloco modelo. As saídas do modelo paramétrico, diante dos sinais subótimos projetados pelo rSOESGOPE, irão configurar as entradas para o treinamento do bloco de refinamento.

As saídas reais $\nu = [u, v, r]$ são provenientes da atuação dos sensores. A velocidade linear, u e v , são as derivadas dos dados obtidos do sensor de sistema de posicionamento global, do inglês, *Global Positioning System* (GPS). As medidas relativas à velocidade angular, r , são provenientes da Unidade de Medida Inercial, do inglês, *Inertial Measurement Unit* (IMU). A IMU utilizada tem uma porcentagem menor de erros quando comparada ao GPS. Com dados mais confiáveis, é possível obter um modelo paramétrico mais próximo da realidade. Pode-se concluir que a estimativa da velocidade angular é mais precisa. Este fato é confirmado nos resultados de (NETO et al., 2021), que demonstra valores de erro quadrático médio, do inglês, *Root Mean Squared Error* (RMSE) menos significativos do que os erros das outras variáveis lineares.

Como o modelo da embarcação é altamente acoplado, ou seja, os estados interferem entre si, uma estimativa que dependa da entrada dos estados influencia fortemente as saídas do modelo. Entradas mais confiáveis para o bloco de refinamento podem definir saídas mais próximas do valor real. Levando isso em consideração, o submodelo de refinamento pode ser dividido em duas etapas em cascata: (i) estimativa da taxa angular; (ii) estimação de velocidades lineares. Considerando $\nu_M = [u_M, v_M, r_M]$ como as saídas paramétricas do modelo, e $\nu_R = [u_R, v_R, r_R]$ as saídas após o refinamento, a Figura 16 ilustra o esquema descrito.

Figura 16 – Diagrama esquemático da estrutura de refinamento, contendo o refinamento da taxa angular e velocidades lineares



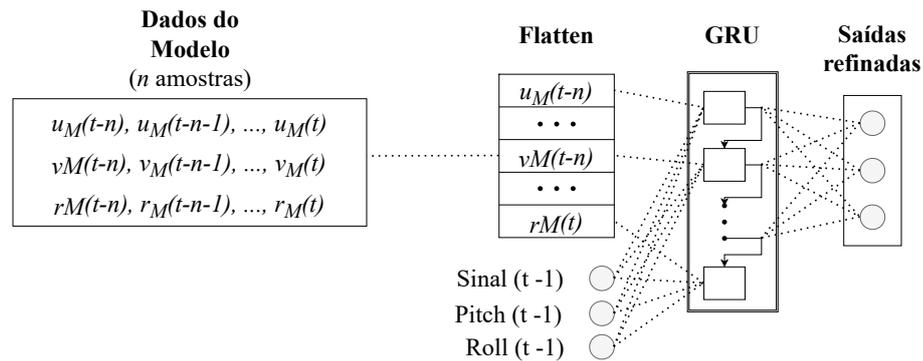
Fonte: Elaborado pelo autor

Usando as entradas do modelo paramétrico e o sinal abaixo do ideal, cada RNP tenta prever os estados do próximo ponto no tempo. No entanto, é importante ressaltar que o uso de informações de momentos anteriores pode criar uma dependência temporal entre os estados de saída do modelo híbrido. Assim, a utilização de uma estrutura RNP que computa e armazena estados anteriores de séries temporais é uma alternativa interessante. Nas redes recorrentes GRU, as camadas que recebem uma pequena atualização de gradiente param de aprender. Como essas camadas não aprendem, a estrutura pode esquecer o que foi visto em sequências mais longas e, assim, ter uma memória de curto prazo. Os mecanismos internos desse tipo de rede têm a capacidade de filtrar seletivamente informações irrelevantes e reter apenas o que é útil, tornando-se uma alternativa para o problema de refinamento das saídas. A justificativa é conseguir determinar as saídas do próximo instante através de informações do modelo paramétrico e das saídas do modelo refinado em instantes anteriores.

Portanto, o modelo paramétrico pode servir, não apenas como ponto de partida para a rede, responsável pelo refinamento, mas também como entradas que correlacionam uma temporalidade entre instantes anteriores de saídas. Outra fonte de informação importante para a rede é o sinal de excitação, que é aplicado um instante antes do desejado. Como a estrutura de identificação é utilizada em uma atualização de posição online, outras entradas podem ser incluídas.

Dado um sinal de entrada, o VAS apresenta mudanças de velocidade nos demais graus de liberdade, além dos movimentos de *sway*, *roll* e *yaw*, os comportamentos associados ao *pitch*, *roll* e *heave* podem influenciar fortemente a modelagem, embora não sejam controláveis. Em geral, todas essas informações adicionais sobre o modelo matemático são extremamente valiosas sob o ponto de vista do aprendizado, pois é possível combinar diferentes comportamentos em uma mesma rede. A Figura 17 ilustra o esquema de conexões de entrada e saída da rede.

Figura 17 – Esquema da arquitetura RNP, composta pelas camadas *flatten* e GRU



Fonte: Elaborado pelo autor

Nesta estrutura, todos os neurônios em uma camada totalmente conectada se conectam a todos os neurônios na camada anterior. Como esse tipo de arranjo é possível transmitir todas as informações da camada anterior para todas as camadas a montante, tornando-se essencial nessa aplicação, pois o VAS possui modelagem não linear com graus de liberdade totalmente acoplados. A camada GRU processa os dados do modelo considerando as n amostras anteriores. Ou seja, se o interesse é obter a saída para o tempo t_i , o processamento considera as amostras de velocidade estimadas pelo modelo paramétrico nos tempos t_{i-n} a t_i . Assim como ocorre para as entradas de sinal, *pitch* e *roll*, apenas as amostras do último instante são consideradas, t_{i-1} .

Com o objetivo de melhorar o desempenho, uma nova estrutura GRU em cascata pode ser adicionada. Já para evitar *overfitting*, é importante adicionar camadas *dropout* após as camadas GRU de cada estágio da rede (SRIVASTAVA et al., 2014). Essa estrutura descarta aleatoriamente blocos de dados para evitar *overfitting*, que ocorre quando nos dados de treino, o modelo tem um desempenho excelente, porém quando utiliza-se os dados de teste, o resultado é ruim.

A saída é representada pela camada de regressão, que calcula a função de perda através da raiz do erro quadrado médio para cada intervalo de tempo. O método do algoritmo de otimização de Adam (KINGMA; BA, 2014) é usado para treinar a arquitetura, com a métrica de avaliação da resposta ao treinamento, a raiz quadrada média do erro (RMSE).

5.3 RESULTADOS

A identificação híbrida do ASV AERO4River é realizada considerando as etapas apresentadas pela Seção 4.3. Os dois primeiros estágios definem a aplicação completa do rSOESGOPE, desde a geração de sinais subótimos até a estimativa de parâmetros. Nesta seção, os resultados da identificação híbrida serão discutidos aplicando os sinais de treinamento e os experimentos de validação e teste.

A embarcação tem as seguintes características iniciais: $m = 33,06 \text{ kg}$; $I_z = 4,12 \text{ kg} \cdot \text{m}^2$; $x_g = -0,05 \text{ m}$; e $y_g = 0 \text{ m}$.

Os limites operacionais da embarcação se encontram na Tabela 22 e foram obtidos a partir dos testes realizados antecipadamente.

Tabela 22 – Limites operacionais do AERO4River

Variável	Limite Inferior	Limite Superior	Unidade
u	0,00	2,05	m/s
v	-0,90	0,90	m/s
r	-2,87	2,87	rad/s
F_x	0,00	78,00	N
F_y	-78,00	78,00	N
F_N	-42,00	42,00	N·m

O rSOESGOPE lista um conjunto de sinais que permitem a estimação de parâmetros e, conseqüentemente, a modelagem matemática do sistema. Em geral, essa seleção é possível por meio da ação do algoritmo PSO. Seguindo orientações de Souza, Honório e Oliveira (2021), neste trabalho, são adotados os seguintes critérios nesta etapa: projeto de 3 sinais; 15 partículas; 6 estágios do sinal; ponderações da função objetivo $k_1 = 200$, $k_2 = 1$, $k_3 = 20$, relacionadas a f_o , f_δ e Θ , respectivamente. As restrições operacionais relacionadas à posição e às velocidades máximas e mínimas atingidas são: de -6π a $6\pi \text{ rad}$ para ψ ; de 0 a 2 m/s por u ; de -0.9 a 0.9 m/s para v ; de -2.8 a 2.8 rad/s por r .

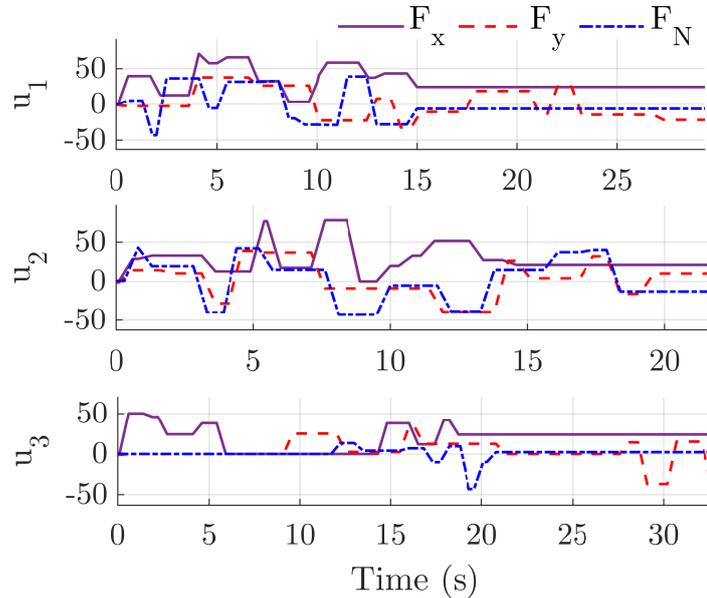
Os sinais subótimos encontrados $U = [u_1, u_2, u_3]$ são dados em termos de forças totais aplicadas a cada grau de liberdade controlável: F_x , F_y , e F_N . A Figura 18 ilustra os sinais provenientes do processo rSOESGOPE.

Tais sinais foram projetados para elevar as possibilidades de aquisição de informações acerca do comportamento. E assim, é possível estimar os parâmetros do modelo matemático através da ferramenta de otimização. A Tabela 23 mostra os parâmetros estimados para o modelo matemático.

5.3.1 Etapa de Refinamento

O estágio final da identificação híbrida é a inserção de um bloco de aprendizado profundo. Com base no modelo paramétrico, é possível fornecer suas saídas como entradas

Figura 18 – Curvas dos sinais subótimos gerados pelo método SOESGOPE, indicando a força que deve ser aplicada em cada grau de liberdade no movimento de teste



Fonte: Elaborado pelo autor

Tabela 23 – Parâmetros estimados para o VAS - Γ^+

$X_{\dot{u}}$	$X_{\dot{v}}$	$X_{\dot{r}}$	$Y_{\dot{u}}$	$Y_{\dot{v}}$	$Y_{\dot{r}}$	$N_{\dot{u}}$	$N_{\dot{v}}$	$N_{\dot{r}}$
-17.9	0.0	0.0	-6.7	-19.9	-5.4	0.0	-5.8	-8.5
X_u	X_v	X_r	Y_u	Y_v	Y_r	N_u	N_v	N_r
0.0	-39.9	-4.8	0.0	0.0	0.0	-0.9	-6.9	0.0
$X_{ u u}$	$X_{ u v}$	$X_{ u r}$	$Y_{ v v}$	$Y_{ v u}$	$Y_{ v r}$	$N_{ r r}$	$N_{ r u}$	$N_{ r v}$
-7.1	16.1	0.0	-190.0	-22.5	0.0	0.0	0.0	13.3

para a rede. Nesta fase, o interesse não está mais na estimação de parâmetros, mas sim na definição de um modelo de caixa cinza capaz de emular uma trajetória baseada em sinais de entrada.

Para o treinamento dos dois blocos de GRU (taxa angular e velocidade linear), são utilizados dados de experimentos subótimos da etapa anterior. O número de unidades ocultas foi 200 para o bloco angular e 100 para o bloco linear. Uma camada de *dropout* foi incluída em cada bloco. Como mencionado, essa camada evita o *overfitting*. Uma camada totalmente conectada foi incluída e também uma camada de regressão. Este conjunto de camadas produz o refinamento do modelo. O otimizador Adam foi usado para determinar as variáveis líquidas (ZHANG, 2018).

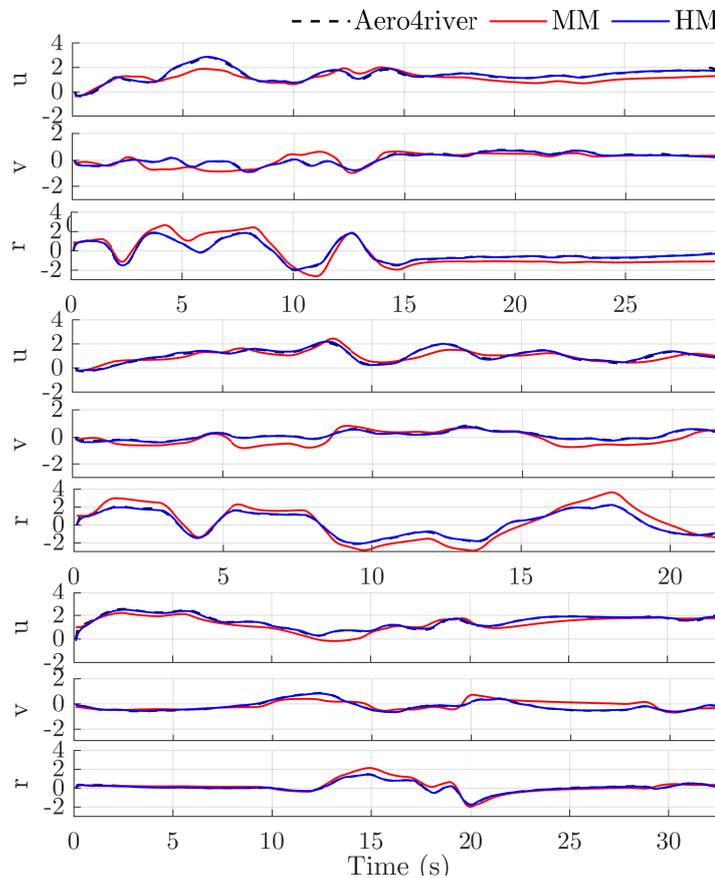
Um experimento é utilizado para validação e outros três experimentos são utilizados para teste, de acordo com os resultados apresentados a seguir.

Foram utilizadas $n = 7$ amostras de tempo anteriores, do tempo a ser estimado, considerando o tempo de amostragem de 0,1s. É importante ressaltar que devido ao

tamanho da embarcação, o acionamento tem uma sensibilidade relativamente alta, assim, os comandos são rapidamente absorvidos e executados, sendo necessária a utilização dos dados em um curto período de tempo. A avaliação é realizada analisando o RMSE entre as saídas reais e as saídas do modelo híbrido e do modelo matemático para fins de comparação.

Com relação aos dados de treinamento, os resultados mostraram uma melhora significativa, como pode ser visto na Figura 19 e Tabela 24, que descreve o RMSE para cada grau de liberdade identificável do inglês, *Degrees of Freedom* (DoF), onde MM, HM e AeroRiver representam o modelo matemático, o modelo híbrido e as saídas reais, respectivamente.

Figura 19 – Trajetória estimada por modelo matemático (MM) e modelo híbrido (HM), considerando sinais subótimos 1, 2 e 3 com u [m/s], v [m/s], r [rad/s]



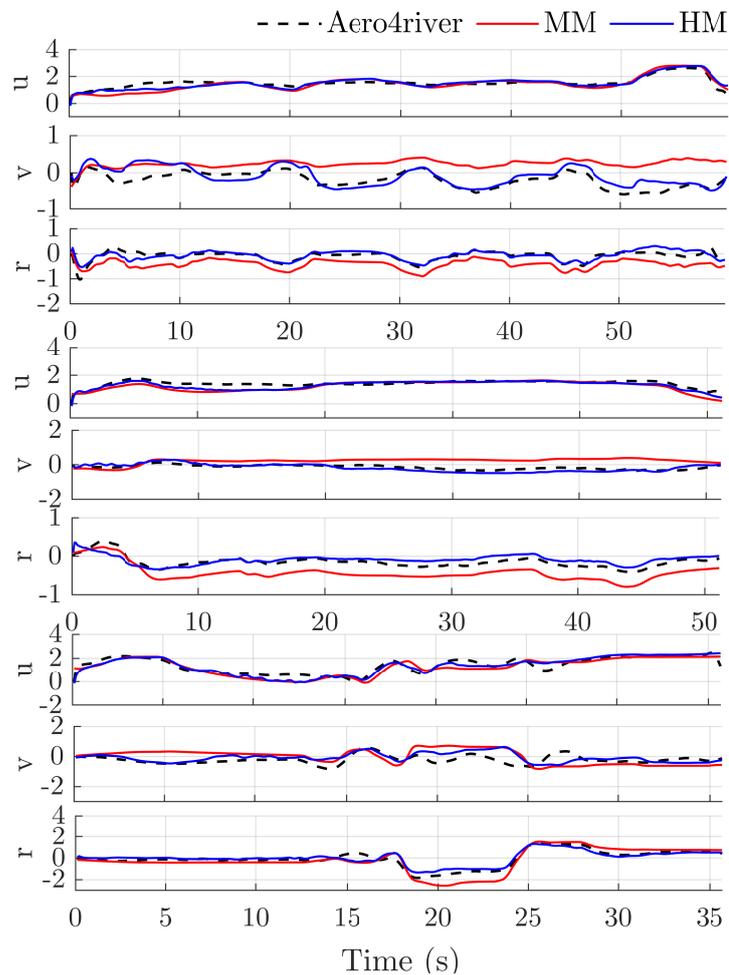
Fonte: Elaborado pelo autor

Embora a análise dos experimentos relacionados aos sinais projetados pela SOESGOPE mostre a superioridade da metodologia sobre o modelo matemático, o impacto nos experimentos de teste deve ser avaliado. Tais experimentos foram gerados manualmente. Os dados de cada trilha foram comparados com as saídas do modelo matemático e do modelo híbrido, aplicando a mesma entrada. T4 foi usado para validação e T5 e T6 para teste. A Figura 20 ilustra o comportamento do modelo matemático e a correção pelo modelo híbrido. A Tabela 25 descreve o RMSE para cada DoF identificável.

Tabela 24 – RMSE relacionado ao modelo matemático (MM) e ao modelo híbrido (HM) para as faixas T1, T2 e T3

		u	v	r
T1	<i>MM</i>	0.4275	0.3218	0.6259
	<i>HM</i>	0.0699	0.0310	0.0385
T2	<i>MM</i>	0.2575	0.3623	0.7072
	<i>HM</i>	0.0602	0.0320	0.0570
T3	<i>MM</i>	0.3242	0.2955	0.2553
	<i>HM</i>	0.0739	0.0253	0.0344

Figura 20 – Trajetória estimada por modelo matemático (MM) e modelo híbrido (HM), considerando os sinais manuais 4 (validação), 5 e 6 (teste)



Fonte: Elaborado pelo autor

Para dados de treinamento T1, T2 e T3, o modelo híbrido fornece uma aproximação percentual média para o caminho real de cerca de 89,29%, 86,58% e 85,05%.

Para dados de teste T4, T5 e T6, o modelo híbrido produz uma aproximação percentual média para o caminho real de cerca de 50,86%, 57,97% e 32,63%, respectivamente. Embora não sejam obtidos valores tão significativos quanto os experimentos de

Tabela 25 – Raiz do erro quadrático médio relacionado ao modelo matemático (MM) e ao modelo híbrido (HM) para as trajetórias T4, T5 e T6

		u	v	r
T4	<i>MM</i>	0.2818	0.4795	0.3488
	<i>HM</i>	0.1640	0.2437	0.1347
T5	<i>MM</i>	0.2914	0.4285	0.2917
	<i>HM</i>	0.1606	0.1489	0.1064
T6	<i>MM</i>	0.3613	0.5574	0.4398
	<i>HM</i>	0.2601	0.3574	0.2299

treinamento, a redução é significativa e contribui para um maior emparelhamento dos dados e, conseqüentemente, para a robustez do modelo.

Uma das razões para a utilização de um modelo adicional baseado em aprendizado é a possibilidade de direcionar o modelo matemático, que já possui um comportamento semelhante ao real, para contemplar características não modeladas do sistema. A partir dos resultados obtidos, é possível calcular uma redução significativa do erro, de, pelo menos 32%, entre o observado e o estimado pelo modelo. A redução evidencia a superioridade do modelo híbrido ao uso puro e simples de uma modelagem matemática, ainda que aprimorada pelo uso de um sinal de entrada ótimo.

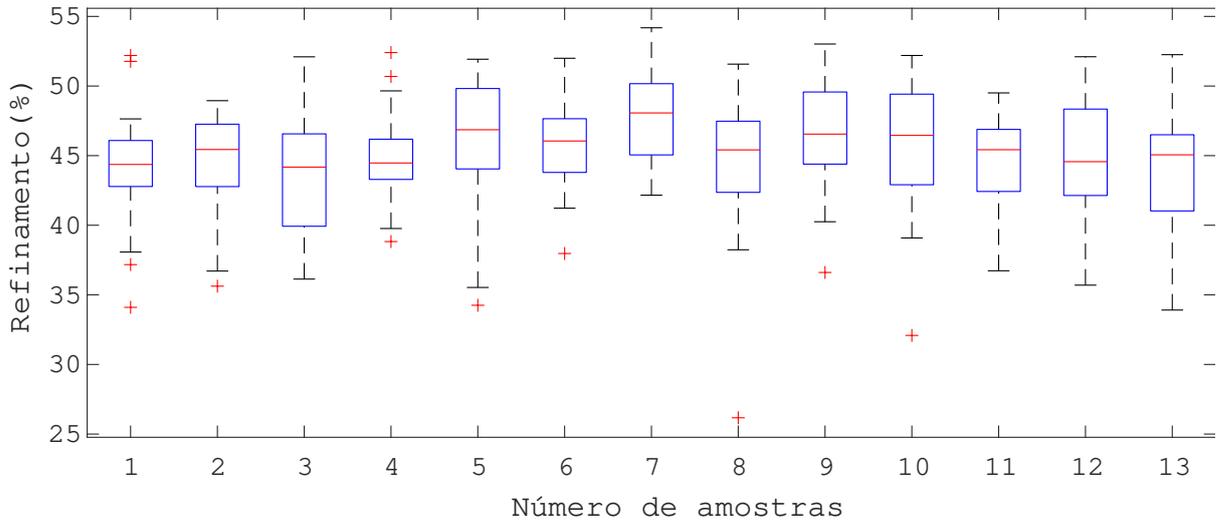
5.3.2 Efeito do número de amostras

Qual a importância do uso de amostras anteriores para treinar a rede combinada? Esta é a pergunta a ser respondida nesta subseção. No caso da embarcação citada, o tempo de amostragem dos dados adquiridos é de 100ms. O uso das últimas 10 amostras definiria os dados dos 1s anteriores. Embora o volume de dados seja crucial para o treinamento, é importante equalizar o número de amostras utilizadas, pois dados antigos podem confundir a rede. Para realizar esta análise, o comportamento do modelo foi analisado usando o número de amostras de n de 1 a 13, ou seja, de 0,1s a 1,3s. O objetivo do estudo foi a faixa 4 (T4). Para isso, foram realizados 100 treinamentos de rede para cada n . Para fins de comparação, foi selecionada a métrica de aumento percentual de similaridade com o comportamento real, a partir do modelo matemático.

O gráfico, no formato de boxplot, apresentado pela Figura 21 permite a visualização das estatísticas resumidas. A informação estatística corresponde à redução do erro quadrático médio em relação ao modelo puramente otimizado. Ou seja, para cada tipo de inicialização de amostras de entrada, é apresentada a quantificação da melhoria que é possível através da composição e aprendizado do modelo paramétrico.

Analisando o diagrama, é possível extrair algumas informações: (i) o uso de dados de amostras anteriores é válido, pois a eficiência percentual máxima aumenta até certo ponto; (ii) neste caso, a utilização de $n = 7$ amostras de tempo anteriores é mais interessante,

Figura 21 – Impacto do número da amostras (n) na melhoria do modelo matemático, considerando 100 rodadas de treinamento



Fonte: Elaborado pelo autor

pois aumenta a confiabilidade, ou seja, a diferença entre o valor mínimo e máximo diminui e aumenta o valor da mediana; (iii) aumentando n , a partir de 7, não é possível notar melhora, pelo contrário, reduz a confiabilidade e também a mediana.

5.3.3 Impacto do sinal de entrada

Apesar da significativa correção de trajetória, surge a seguinte questão: qual a importância do sinal subótimo para o modelo híbrido? Para responder a esta questão, a situação apresentada na subseção anterior pode ser invertida, ou seja, os sinais de treinamento são utilizados como testes e os últimos como treinamento. Em seguida, deve-se avaliar se sinais manuais aleatórios são capazes de fornecer informações suficientes para o treinamento de uma rede de refinamento de saída. A tabela 26 mostra o resultado desta análise, com T4, T5 e T6 como treinamento e T1, T2 e T3 como testes.

Tabela 26 – RMSE referente ao modelo matemático (MM) e ao modelo híbrido (HM) para as trilhas T1, T2 e T3, usando como conjunto de treinamento as trilhas T4, T5 e T6

		u	v	r
T1	MM	0.4275	0.3218	0.6259
	HM	0.4919	0.4565	0.4373
T2	MM	0.2575	0.3623	0.7072
	HM	0.4015	0.4615	0.5374
T3	MM	0.3242	0.2955	0.2553
	HM	0.4384	0.2442	0.2688

Os valores de RMSE revelam que os sinais manuais não foram capazes de fornecer as informações necessárias para o aprendizado do refinamento do modelo e, ao invés de

reduzir o erro, ele aumenta. Este resultado está de acordo com o propósito do projeto de sinal subótimo, pois possui características favoráveis para a excitação do sistema, que são importantes para o processo de estimação de parâmetros e também para o bloco de refinamento.

5.4 CONSIDERAÇÕES FINAIS

Este capítulo focou em apresentar uma nova abordagem do problema de identificação de sistemas dinâmicos. Duas situações foram levadas em conta a fim de motivar tal método. Primeiramente, a estimação de parâmetro, embora seja extremamente positiva em sistemas ideais, ou seja cujo modelo matemático equivale ao comportamento real, ela fica a desejar em casos que o sistema não é uma réplica fiel do modelo. E, quando se pensa no processo construtivo de qualquer sistema, é possível identificar vários equívocos na construção, tanto pelo uso de equipamentos que não correspondem ao especificado, quando na própria construção, em si, que apresenta vícios dificilmente modeláveis. Por outro lado, uma modelagem em caixa preta, não utilizando uma análise teórica do sistema, pode não conseguir resultados satisfatórios, em vista da pouca informação para treinamento.

Desta forma, um processo híbrido, que conta com a atuação de um método de estimação de parâmetros seguido por um de aprendizagem foi proposto. O método SOESGOPE, amplamente discutido nos capítulos anteriores, torna-se uma valiosa fonte de informações para o treinamento de um rede neural, que fornece, assim, uma correção de erros atribuídos à quaisquer transtornos não modeláveis. Um exemplo simulado comprovou a eficácia da nova metodologia.

Além de propor o novo método de identificação, foi proposta uma aplicação, que apresenta grandes desafios em seu modelo. O catamarã ASV AERO4River apresenta inúmeras possibilidades de inconsistências em sua construção, o que o torna interessante para verificação de eficiência do método híbrido. Além disso, ele possui uma maior quantidade de parâmetros, deixando o problema mais complexo e de difícil convergência.

Aplicando a metodologia fundamentada em uma estrutura convolucional baseada em redes neurais recorrentes, houve um aumento na aproximação do modelo à realidade, dada pelas trajetórias da embarcação em testes reais. Testes para verificar o impacto do número de amostras foi realizado, indicando que tal parametrização tem um efeito no processo de identificação e deve ser levado em consideração. Já quanto ao sinal, o projeto de entradas subótimas tiveram performance superior quando se compara a trajetórias aleatórias, indicando que essa metodologia é válida tanto no bloco de estimação de parâmetros quanto no bloco de refinamento.

6 CONCLUSÕES

O trabalho teve como objetivo avaliar o SOESGOPE de forma efetiva possibilitando o entendimento de lacunas de atuação do método. A partir daí, foi possível propor uma modificação baseada na associação do modelo matemático proveniente do método com uma etapa de rede neural profunda, sendo possível elevar a eficiência, providenciando saídas mais próximas às observadas na realidade.

No Capítulo 2 foram apresentadas as dificuldades e soluções encontradas na literatura para estimação de parâmetros de sistemas não lineares, onde observou-se observar uma área promissora para pesquisa. O uso de sinais de excitação aleatórios pode não fornecer modelagem robusta. Uma alternativa é a entrada ideal. Ao considerar experimentos com altos custos monetários e dificuldade temporária, é altamente necessário que um conjunto de dados, mesmo que limitado, possa fornecer as informações necessárias para a construção do modelo. O SOESGOPE é um método que trabalha com esse requisito. É possível criar um sinal eficiente com requisitos de tempo e espaço reduzidos. O SOESGOPE gera um sinal de excitação suficiente para um experimento de estimação de parâmetros de um modelo de sistema dinâmico. No entanto, esta metodologia é projetada para resolver o problema de estimação de parâmetros. Um grande desafio neste tipo de problema é a competência para englobar a dinâmica do sistema, que ainda não é alcançada com sucesso. A utilização de ferramentas de aprendizagem é capaz de preencher essa lacuna e proporcionar resultados mais próximos da realidade

No Capítulo 3, foram apresentadas duas modificações do método SOESGOPE que resultaram em duas publicações. Nessas análises, pôde-se concluir que um número mais significativo de parâmetros, quando se é possível, nem sempre fornece o melhor resultado, muitas vezes pode, inclusive atrapalhar a convergência, principalmente quando a estimativa inicial não se enquadra na realidade. Dessa forma, concluiu-se que o uso de um sistema compacto pode trazer benefícios como mais precisão e confiabilidade. Outra análise importante foi a condução de um algoritmo recursivo a fim de aumentar o desempenho da estimação, visto que a metodologia original não é capaz de lidar com parâmetros cuja faixa de valores permitidos é desconhecida. Através dessa abordagem, o algoritmo se torna robusto, conseguindo obter resultados satisfatórios mesmo quando os parâmetro são desconhecidos. Além disso, é possível que os limites de busca impostos não contemplem os parâmetros a serem encontrados. Através da recursividade é possível obter os multiplicadores de Lagrange provenientes do processo de otimização, aumentar as barreiras e direcionar os parâmetros. Esta abordagem é extremamente aplicável em sistemas complexos com parâmetros imensuráveis.

Não obstante a grande aplicabilidade e eficácia do SOESGOPE, percebe-se que ele não consegue lidar com inconsistências no modelo. Caso o modelo não contemple o

comportamento real, será impossível obter parâmetros que combinados consigam descrever fielmente o seu comportamento real. Por isso, no Capítulo 4 foi proposta a associação entre SOESGOPE e um bloco da Rede Neural Profunda, sendo uma alternativa para refinar as saídas de um modelo matemático. O modelo torna-se não mais a saída final, mas sim, a entrada de uma rede, que computará a atuação juntamente com esses dados provenientes dos modelo matemático. Essa informação a mais é essencial para o treinamento, que cria uma lógica a partir de dados matemáticos e experimentais. Testes em um sistema simulado demonstram a superioridade do método híbrido com relação às opções individuais (tanto a estimação de parâmetros por si só, quanto a rede neural).

Através da utilização de dados reais de navegação do ASV AERO4River, no Capítulo 5 foi possível verificar a eficácia do método híbrido, baseado na associação de modelo matemático e estrutura GRU. Um barco autônomo é uma interessante plataforma de testes para detectar a viabilidade e eficácia quando se considera um sistema altamente passível a vícios construtivos. Os resultados demonstraram que a metodologia é promissora e reduz significativamente os erros fornecidos por um modelo puramente matemático.

Os seguintes pontos da metodologia devem ser destacados:

- Redução proporcional do erro substancial em relação ao método rSOESGOPE;
- A estrutura GRU mostrou-se eficiente para lidar com o refinamento do modelo matemático. O método consegue lidar com as informações anteriores e extrair apenas as informações necessárias para determinar o estado futuro;
- Implementar RNP usando sinais do SOESGOPE não só é promissor como também cria condições favoráveis para o aprendizado já que o projeto de entrada ótimo seleciona experimentos com mais chances de excitação;
- Um sinal subótimo pode ser determinante para situações cujo experimento de identificação é complicado ou caro, pois poucos experimentos são capazes de criar condições de identificação.

6.1 SUGESTÕES PARA DESENVOLVIMENTOS FUTUROS

Baseado nos desenvolvimentos e contribuições da presente Tese, os seguintes pontos podem ser investigados para melhoria e aperfeiçoamento do modelo proposto:

- Verificando que o sinal de entrada interfere fortemente na modelagem híbrida, entender a relação entre o sinal ótimo e a robustez do modelo híbrido seria extremamente importante;
- Gerar o sinal subótimo utilizando como base o modelo híbrido de modo a tentar captar entradas que facilitem a identificação do movimento;

- Agrupar os domínios de ação do modelo matemático sendo capaz de detectar áreas de ação do modelo híbrido e evitar situações em que o refinamento aumente a discrepância entre o real e o simulado;
- Verificar a possibilidade de compactar o modelo matemático da embarcação de forma a melhorar a performance do modelo matemático e, conseqüentemente, do modelo híbrido;
- Avaliar a comparação de tempo computacional das técnicas;
- Aplicar a análise de variância a fim de comparar desempenho de diferentes arquiteturas ou algoritmos de treinamento analisando a variação do erro ou precisão dos modelos treinados. Tal análise pode ajudar a identificar quais parâmetros ou técnicas de regularização podem ter um impacto no desempenho dos modelos.

REFERÊNCIAS

- ABAS, N.; LEGOWO, A.; AKMELIAWATI, R. Parameter identification of an autonomous quadrotor. In: IEEE. *2011 4th International Conference on Mechatronics (ICOM)*. [S.l.], 2011. p. 1–8.
- AHSAN, M.; CHOUDHRY, M. A. System identification of an airship using trust region reflective least squares algorithm. *International Journal of Control, Automation and Systems*, Springer, v. 15, n. 3, p. 1384–1393, 2017.
- AN, H. et al. Dynamic model identification for adaptive polishing system. *International Journal of Control, Automation and Systems*, Springer, v. 20, n. 9, p. 3110–3120, 2022.
- BAILLIEUL, J.; SAMAD, T. *Encyclopedia of systems and control*. [S.l.]: Springer Publishing Company, Incorporated, 2015.
- BROSILOW, C.; JOSEPH, B. *Techniques of model-based control*. [S.l.]: Prentice Hall Professional, 2002.
- BYRD, R. H.; HRIBAR, M. E.; NOCEDAL, J. An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, SIAM, v. 9, n. 4, p. 877–900, 1999.
- CHEN, J.; JIANG, B.; LI, J. Missing output identification model based recursive least squares algorithm for a distributed parameter system. *International Journal of Control, Automation and Systems*, Springer, v. 16, n. 1, p. 150–157, 2018.
- CHO, K. et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.
- CRASSIDIS, J. L.; JUNKINS, J. L. *Optimal estimation of dynamic systems*. [S.l.]: Chapman and Hall/CRC, 2011.
- DEFLORIAN, M.; ZAGLAUER, S. Design of experiments for nonlinear dynamic system identification. *IFAC Proceedings Volumes*, Elsevier, v. 44, n. 1, p. 13179–13184, 2011.
- DING, F.; LIU, X.; GU, Y. An auxiliary model based least squares algorithm for a dual-rate state space system with time-delay using the data filtering. *Journal of the Franklin Institute*, Elsevier, v. 353, n. 2, p. 398–408, 2016.
- DING, F.; LIU, X.; LIU, M. The recursive least squares identification algorithm for a class of wiener nonlinear systems. *Journal of the Franklin Institute*, Elsevier, v. 353, n. 7, p. 1518–1526, 2016.
- EBERHART, R.; KENNEDY, J. Particle swarm optimization. In: CITESEER. *Proceedings of the IEEE international conference on neural networks*. [S.l.], 1995. v. 4, p. 1942–1948.
- EGORCHEV, M.; TIUMENTSEV, Y. Neural network semi-empirical modeling of the longitudinal motion for maneuverable aircraft and identification of its aerodynamic characteristics. In: SPRINGER. *International Conference on Neuroinformatics*. [S.l.], 2017. p. 65–71.

- EGORCHEV, M.; TIUMENTSEV, Y. Homotopy continuation training method for semi-empirical continuous-time state-space neural network models. In: SPRINGER. *International Conference on Neuroinformatics*. [S.l.], 2018. p. 115–120.
- EGORCHEV, M. V.; TIUMENTSEV, Y. V. Semi-empirical neural network based approach to modelling and simulation of controlled dynamical systems. *Procedia computer science*, Elsevier, v. 123, p. 134–139, 2018.
- ERUNSAL, I. K. et al. An approach for system identification in unmanned surface vehicles. In: IEEE. *2017 17th International Conference on Control, Automation and Systems (ICCAS)*. [S.l.], 2017. p. 194–199.
- FAHIMI, F. *Autonomous robots: modeling, path planning, and control*. [S.l.]: Springer Science & Business Media, 2008. v. 107.
- FOSSSEN, T. I. *Guidance and control of ocean vehicles*. [S.l.]: John Wiley & Sons Inc, 1994.
- GAUSS, C. F. *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*. [S.l.]: Perthes et Besser, 1809. v. 7.
- GELB, A. *Applied optimal estimation*. [S.l.]: MIT press, 1974.
- GRINGARD, M.; KROLL, A. On the parametrization of aprbs and multisine test signals for the identification of nonlinear dynamic ts-models. In: IEEE. *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. [S.l.], 2016. p. 1–8.
- HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural computation*, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.
- HONORIO, L. et al. Solving optimal power flow problems using a probabilistic α -constrained evolutionary approach. *IET generation, transmission & distribution*, IET, v. 4, n. 6, p. 674–682, 2010.
- HONÓRIO, L. M. et al. Persistently-exciting signal generation for optimal parameter estimation of constrained nonlinear dynamical systems. *ISA transactions*, Elsevier, v. 77, p. 231–241, 2018.
- ISERMANN, R.; MÜNCHHOF, M. *Identification of dynamic systems: an introduction with applications*. [S.l.]: Springer Science & Business Media, 2010.
- ISERMANN, R.; MÜNCHHOF, M. *Identification of dynamic systems: an introduction with applications*. [S.l.]: Springer, 2011. v. 85.
- JIANG, Y. et al. Identification modeling and prediction of ship maneuvering motion based on lstm deep neural network. *Journal of Marine Science and Technology*, Springer, v. 27, n. 1, p. 125–137, 2022.
- KIM, S.; KWON, S. Dynamic modeling of a two-wheeled inverted pendulum balancing mobile robot. *International Journal of Control, Automation and Systems*, Springer, v. 13, n. 4, p. 926–933, 2015.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

- KLOSOWSKI, J. T. et al. Efficient collision detection using bounding volume hierarchies of k-dops. *IEEE transactions on Visualization and Computer Graphics*, IEEE, v. 4, n. 1, p. 21–36, 1998.
- LI, M.; LIU, X.; DING, F. The gradient-based iterative estimation algorithms for bilinear systems with autoregressive noise. *Circuits, Systems, and Signal Processing*, Springer, v. 36, n. 11, p. 4541–4568, 2017.
- LICHOTA, P.; SIBILSKI, K.; OHME, P. D-optimal simultaneous multistep excitations for aircraft parameter estimation. *Journal of Aircraft*, American Institute of Aeronautics and Astronautics, v. 54, n. 2, p. 747–758, 2017.
- LICITRA, G. et al. Optimal input design for autonomous aircraft. *Control Engineering Practice*, Elsevier, v. 77, p. 15–27, 2018.
- LIU, Q. et al. Recursive identification methods for multivariate output-error moving average systems using the auxiliary model. *International Journal of Control, Automation and Systems*, Springer, v. 16, n. 3, p. 1070–1079, 2018.
- LUO, W.; SOARES, C. G.; ZOU, Z. Parameter identification of ship maneuvering model based on support vector machines and particle swarm optimization. *Journal of Offshore Mechanics and Arctic Engineering*, American Society of Mechanical Engineers, v. 138, n. 3, p. 031101, 2016.
- MACIEL, G. M. et al. Daily water flow forecasting via coupling between snmp and deep learning. *IEEE Access*, IEEE, v. 8, p. 204660–204675, 2020.
- MATHWORKS, I. T. *fmincon Interior-Point Algorithm with Analytic Hessian*. 2020. Disponível em: <<https://www.mathworks.com/help/optim/ug/fmincon-interior-point-algorithm-with-analytic-hessian.html>>.
- MORENO-SALINAS, D. et al. Modelling of a surface marine vehicle with kernel ridge regression confidence machine. *Applied Soft Computing*, Elsevier, v. 76, p. 237–250, 2019.
- NAEEM, W.; SUTTON, R.; CHUDLEY, J. Modelling and control of an unmanned surface vehicle for environmental monitoring. In: *UKACC International Control Conference*. [S.l.: s.n.], 2006. p. 1–6.
- NELLES, O. *Nonlinear system identification*. [S.l.]: IOP Publishing, 2002.
- NELLES, O. Nonlinear dynamic system identification. In: _____. *Nonlinear System Identification: From Classical Approaches to Neural Networks, Fuzzy Models, and Gaussian Processes*. Cham: Springer International Publishing, 2020. p. 831–891. ISBN 978-3-030-47439-3. Disponível em: <https://doi.org/10.1007/978-3-030-47439-3_19>.
- NETO, A. F. D. S. et al. Development of optimal parameter estimation methodologies applied to a 3dof autonomous surface vessel. *IEEE Access*, IEEE, v. 9, p. 50035–50049, 2021.
- OLIVEIRA, E. J. et al. An optimal power flow based on safety barrier interior point method. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 64, p. 977–985, 2015.

- PÓLIK, I.; TERLAKY, T. Interior point methods for nonlinear optimization. In: *Nonlinear optimization*. [S.l.]: Springer, 2010. p. 215–276.
- RAJENDRA, P.; BRAHMAJIRAO, V. Modeling of dynamical systems through deep learning. *Biophysical Reviews*, Springer, v. 12, n. 6, p. 1311–1320, 2020.
- REGINA, B. A. et al. Hull and aerial holonomic propulsion system design for optimal underwater sensor positioning in autonomous surface vessels. *Sensors*, Multidisciplinary Digital Publishing Institute, v. 21, n. 2, p. 571, 2021.
- REVESTIDO, E. et al. Parameter estimation of ship linear maneuvering models. In: IEEE. *OCEANS 2011 IEEE-Spain*. [S.l.], 2011. p. 1–8.
- RODRIGUEZ-FERNANDEZ, M.; MENDES, P.; BANGA, J. R. A hybrid approach for efficient and robust parameter estimation in biochemical pathways. *Biosystems*, Elsevier, v. 83, n. 2-3, p. 248–265, 2006.
- SILVA, M. F. da et al. Project and control allocation of a 3 dof autonomous surface vessel with aerial azimuth propulsion system. *IEEE Access*, IEEE, 2020.
- SINGH, H.; KUMAR, D.; BALEANU, D. *Methods of mathematical modelling: fractional differential equations*. [S.l.]: CRC Press, 2019.
- SØRENSEN, A. J. Structural issues in the design and operation of marine control systems. *Annual Reviews in Control*, Elsevier, v. 29, n. 1, p. 125–149, 2005.
- SOUZA, M. B.; HONÓRIO, L. de M.; OLIVEIRA, E. J. de. Innovative analysis for parameter estimation quality. *International Journal of Control, Automation and Systems*, Springer, v. 19, n. 1, p. 363–371, 2021.
- SOUZA, M. B. et al. Recursive approach of sub-optimal excitation signal generation and optimal parameter estimation. *International Journal of Control, Automation and Systems*, Springer, p. 1–10, 2020.
- SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014.
- TIUMENTSEV, Y.; EGORCHEV, M. *Neural Network Modeling and Identification of Dynamical Systems*. [S.l.]: Academic Press, 2019.
- UGALDE, H. M. R. et al. Computational cost improvement of neural network models in black box nonlinear system identification. *Neurocomputing*, Elsevier, v. 166, p. 96–108, 2015.
- WANG, Z. et al. Kernel-based support vector regression for nonparametric modeling of ship maneuvering motion. *Ocean Engineering*, Elsevier, v. 216, p. 107994, 2020.
- WEINMANN, A. *Uncertain models and robust control*. [S.l.]: Springer Science & Business Media, 2012.
- WIRTENSOHN, S. et al. Modelling and identification of a twin hull-based autonomous surface craft. In: IEEE. *2013 18th International Conference on Methods & Models in Automation & Robotics (MMAR)*. [S.l.], 2013. p. 121–126.

WIRTENSOHN, S. et al. Parameter identification and validation analysis for a small usv. In: IEEE. *2015 20th International Conference on Methods and Models in Automation and Robotics (MMAR)*. [S.l.], 2015. p. 701–706.

WOO, J. et al. Dynamic model identification of unmanned surface vehicles using deep learning network. *Applied Ocean Research*, Elsevier, v. 78, p. 123–133, 2018.

XU, L. Application of the newton iteration algorithm to the parameter estimation for dynamical systems. *Journal of Computational and Applied Mathematics*, Elsevier, v. 288, p. 33–43, 2015.

XU, L.; DING, F. Recursive least squares and multi-innovation stochastic gradient parameter estimation methods for signal modeling. *Circuits, Systems, and Signal Processing*, Springer, v. 36, n. 4, p. 1735–1753, 2017.

ZADEH, L. A. From circuit theory to system theory. *Proceedings of the IRE, IEEE*, v. 50, n. 5, p. 856–865, 1962.

ZHANG, Y.-Y.; WANG, Z.-H.; ZOU, Z.-J. Black-box modeling of ship maneuvering motion based on multi-output nu-support vector regression with random excitation signal. *Ocean Engineering*, Elsevier, v. 257, p. 111279, 2022.

ZHANG, Z. Improved adam optimizer for deep neural networks. In: IEEE. *2018 IEEE/ACM 26th international symposium on quality of service (IWQoS)*. [S.l.], 2018. p. 1–2.