

Pedro Fausto Rodrigues Leite Junior

Predição de Tempo de Vida Restante através do aprendizado de máquina para condições ambientais dinâmicas em sistemas integrados/ Pedro Fausto Rodrigues Leite Junior. – Brasil, 25 de Julho, 2017-

72 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Frank Sill Torres

Dissertação de Mestrado – Universidade Federal de Minas Gerais

Programa de Pós-Graduação em Engenharia Elétrica, 25 de Julho, 2017.

1. Reliability. 2. Robustness. 2. Remaning Useful Lifetime. I. Frank Sill Torres. II. Universidade Federal de Minas Gerais. III. Programa de Pós-Graduação em Engenharia Elétrica. IV. Predição de Tempo de Vida Restante através do aprendizado de máquina para condições ambientais dinâmicas em sistemas integrados.

Pedro Fausto Rodrigues Leite Junior

Predição de Tempo de Vida Restante através do aprendizado de máquina para condições ambientais dinâmicas em sistemas integrados

Brasil

25 de Julho, 2017

Agradecimentos

Gostaria de agradecer ao meu orientador, Professor Dr. Frank Sill Torres pelo auxílio e ensinamentos. É sempre gratificante poder aprender quando se há alguém disposto a transmitir seu conhecimento. Este trabalho não seria possível sem sua incansável capacidade de ensinar, ajuda e busca pela excelência. E sua compreensão e paciência também.

A todos os membros do *OptMA^{lab}*, agradeço de coração pelas trocas de experiência, incentivos e momentos de alegria que me ajudaram a relaxar, respirar fundo e continuar o trabalho. Mencionar individualmente cada um só diminuiria o valor intrínseco do *OptMA^{lab}*: uma equipe.

Aos demais amigos, muitos, que de uma forma ou de outra me incentivaram durante essa jornada.

À minha família pelo apoio incondicional na labuta diária, me apoiando na minha decisão de voltar à pesquisa científica, sejam quais forem as dificuldades encontradas em um trabalho tão essencial à sociedade mas que ainda há de ganhar o devido reconhecimento.

À Sabrina pelo carinho, companheirismo e por sempre me manter focado no que importa.

Resumo

O estudo da confiabilidade de circuitos integrados tem se tornado de súbita importância para o entendimento, detecção e correção de falhas dos mesmos. A compreensão de como ou sob quais condições este envelhecimento se torna prejudicial a um sistema permitirá tomar decisões que sanem ou evitem estas condições.

Este trabalho estabelece um fluxo para análise e simulação de sistemas integrados que permite entender o envelhecimento dos mesmos, em diferentes condições. Além disso, permite extrair e analisar dados que são relevantes para prever a falha dos mesmos e sirvam de entrada para sistemas de verificação, avaliação e atuação contra falhas utilizando-se de técnicas de aprendizado de máquina.

A metodologia desenvolvida permite a integração de técnicas de coleta de dados *offline* e *online* para atualização dos métodos de estimativa, além de permitir que novos sejam adicionados. O trabalho utiliza três métodos diferentes para prever o Tempo Médio para Falha (*i.e.* Mean Time To Failure, MTTF) e o Tempo de Vida Restante para alguns circuitos de teste ISCAS-85 e uma cadeia de inversores. O MTTF é estimado para cada um deles utilizando um Modelo Linear Generalizado (especificamente uma Regressão de Mínimos Quadrados Parciais), a Distância Euclidiana e a Correlação de Pearson como métodos de predição.

Os resultados obtidos indicam que a representação das condições de operação dos sistemas através de perfis dinâmicos é mais realística do que a representação através de um perfil de operação que não varia no tempo, além de mais precisa. Adicionalmente, a predição do MTTF foi de aproximadamente 90% para um modelo de Regressão de Mínimos Quadrados Parciais e de Distância Euclidiana.

Palavras-chave: Tempo de vida restante. Confiabilidade. Envelhecimento de circuitos integrados. Regressão de Mínimos Quadrados Parciais. Modelo Linear Generalizado. Distância Euclidiana. Correlação de Pearson.

Abstract

The study of integrated circuits reliability has become of sudden importance to the understanding, detection and correction of their failures. Comprehend how or under which conditions this aging becomes harmful to a system will enable decisions to be made that will mitigate or prevent these conditions.

This work establishes a flow for analysis and simulation of integrated systems that allows us to understand the aging of it in different conditions. In addition, it allows us to extract and analyze data that are relevant to predict their failure and also serve as input to verification, evaluation and fault-tolerance systems using machine learning techniques.

The methodology developed allows the integration of offline and online data collection techniques to update estimation methods, as well as allowing new ones to be added. This work uses three different methods to predict the Mean Time To Failure and the Remaining Useful Lifetime for some ISCAS-85 test circuits and an inverter chain. The MTTF is estimated for each of them using a Generalized Linear Model (specifically a Partial Least Squares Regression), Euclidean Distance and Pearson's Correlation as prediction methods.

The obtained results indicate that the representation of the operating conditions of the systems through dynamic profiles is more realistic than the representation through an operation profile that does not vary in time, and more precise. Additionally, the MTTF prediction was approximately 90% for Partial Least Squares Regression and Euclidean Distance models.

Keywords: Remaining Useful Lifetime. Reliability. Aging of Integrated Circuits. Partial Least Squares. Generalized Linear Models. Euclidean Distance. Pearson Correlation.

Lista de ilustrações

Figura 1 – Relação entre as falhas, defeitos, tempos médios de falha, tempos médios de recuperação e tempo médio entre falhas.	4
Figura 2 – Tipos de cargas associadas a um estrutura $Si - SiO_2$	8
Figura 3 – Armadilhas para elétrons e lacunas em uma estrutura de $Si - SiO_2$ [1].	8
Figura 4 – Mecanismos de injeção de portadores [2].	10
Figura 5 – Geração de ligações incompletas a partir de vacâncias [3].	14
Figura 6 – Representação da quebra de ligações Si-H e Si-O na interface Si/SiO_2	20
Figura 7 – Abordagem proposta para análise de envelhecimento.	24
Figura 8 – Alteração das entradas refletem-se como uma atividade nas portas lógicas.	25
Figura 9 – Compressão dos dados medidos e representação dos mesmos através de perfis de envelhecimento.	27
Figura 10 – Modelo linear ajustado a um conjunto de dados de exemplo, representado pela reta azul. Os pontos vermelhos representam os dados coletados e seus respectivos resíduos [4].	30
Figura 11 – Representação de componentes principais. A linha verde representa o primeiro componente e a linha azul o segundo. [4].	33
Figura 12 – Fluxo geral utilizado no envelhecimento de circuitos integrados.	38
Figura 13 – Entradas e saídas da ferramenta de envelhecimento.	39
Figura 14 – Detalhamento do fluxo de criação, modificação e extração de resultados.	43
Figura 15 – Variação no atraso de saída.	45
Figura 16 – Caracterização de uma célula inversora degradada durante 1 ano.	48
Figura 17 – Caracterização de uma célula AND2 degradada durante 1 ano.	49
Figura 18 – Caracterização de uma célula NAND2 degradada durante 1 ano.	50
Figura 19 – Caracterização de uma célula OR2 degradada durante 1 ano.	50
Figura 20 – Caracterização de uma célula NOR2 degradada durante 1 ano.	51
Figura 21 – Erro relativo do MTTF para os circuitos ISCAS-85.	56
Figura 22 – Erro relativo do MTTF para a cadeia de inversores.	57
Figura 23 – Erro Quadrático Médio Normalizado do MTTF para os circuitos ISCAS-85.	58
Figura 24 – Erro Quadrático Médio Normalizado do MTTF para a cadeia de inversores.	59
Figura 25 – Relação entre o atraso e a temperatura para diferentes V_{DD}	61

Lista de tabelas

Tabela 1 – Base de Dados de Perfis Envelhecidos	28
Tabela 2 – Excerto de um BDPE com dados exemplares.	28
Tabela 3 – Tabela de exemplo para regressão linear simples.	30
Tabela 4 – Coeficientes do modelo de RLM para o exemplo da tabela 3.	31
Tabela 5 – Tabela de exemplo utilizada para o cálculo das variações de perfis. . .	42
Tabela 6 – Coeficientes do modelo de RLM para o inversor da figura 16.	49
Tabela 7 – Coeficientes do modelo de RLM para a figura 17.	49
Tabela 8 – Coeficientes do modelo de RLM para a figura 18.	50
Tabela 9 – Coeficientes do modelo de RLM para a figura 19.	50
Tabela 10 – Coeficientes do modelo de RLM para a figura 19.	51
Tabela 11 – Características da cadeia de inversores e dos circuitos ISCAS-85 utiliza- dos [5].	52
Tabela 12 – Tamanho dos circuitos de teste após a extração dos caminhos críticos. .	52
Tabela 13 – BDPE estático utilizado no envelhecimento dos circuitos de teste. . . .	53
Tabela 14 – Excerto de um BDPE com dados de simulação.	55
Tabela 15 – Erro máximo para os circuitos de teste.	57
Tabela 16 – Erro Normalizado da Raiz do Valor Quadrático Médio para cada método e circuito em %.	59
Tabela 17 – ENRVQM para cada método e circuito submetidos à perfis estáticos em %.	60
Tabela 18 – ENRVQM para cada método e circuito submetidos à 28 perfis dinâmicos.	62

Lista de abreviaturas e siglas

CMOS	Complementary Metal-Oxide Semiconductor
PMOS	P-type Metal-Oxide Semiconductor
NMOS	N-type Metal-Oxide Semiconductor
RUL	Remaining Useful Lifetime
BTI	Bias Temperature Instability
TDDDB	Time-Dependent Dielectric Breakdown
CHE	Channel hot electron
SHE	Substrate hot electron
DAHC	Drain avalanche hot carrier
SGHE	Secondary generated hot electron
HBD	Hard breakdown
SBD	Soft breakdown
PLS-R	Partial Least Squares Regression
GLM	Generalized Linear Model
HDL	Hardware Description Language
BDM	Banco de Dados de Medições
BDPE	Banco de Dados de Perfis Envelhecidos
RLS	Regressão Linear Simples
RLM	Regressão Linear Múltipla
CP	Correlação de Pearson

Lista de definições

Dispositivos: Elementos básicos de um sistema eletrônico.

Célula padrão: Um conjunto de dispositivos organizados de forma a implementar uma função lógica, *i.e.* as portas lógicas.

Circuito: Um agrupamento de componentes que, combinados, são capazes de realizar operações lógicas mais complexas, *i.e.* somadores, multiplexadores e células de memória.

Sistema: Um grupo de circuitos que interconectam-se entre si e são capazes de executar funções de uma ordem superior, *i.e.* Somadores Ripple Carry (RCAs), processadores, roteadores, arquiteturas de memória e ULAs (Unidade Lógica e Aritimética).

Defeitos: Para este trabalho, defeitos são imperfeições nas células de uma determinada estrutura e normalmente decorrentes das variações nos processos de fabricação.

Erros: Erros são variações não esperadas no comportamento de um sistema.

Foundries: Empresas que são responsáveis pela fabricação de dispositivos semicondutores, podendo atuar desde a preparação do material semicondutor até a criação de novos *designs* (seus ou de terceiros).

Netlist: Arquivo que descreve textualmente dispositivos, transistores, fontes, estímulos elétricos individuais ou de instâncias que compõe um circuito ou sistema. Tais informações são utilizadas não somente de forma documental mas servem de entrada para ferramentas que utilizam tais informações para simular estes elementos e observar o comportamento destes.

Sumário

1	INTRODUÇÃO	1
1.1	Motivação	1
1.2	Roteiro da dissertação	1
2	BASE TEÓRICA	3
2.1	Confiabilidade	3
2.1.1	Confiabilidade em Engenharia	5
2.1.2	Técnicas de confiabilidade e robustez	6
2.1.2.1	Técnicas em hardware	6
2.1.3	Confiabilidade em circuitos integrados	7
2.1.3.1	Hot Carrier Injection	9
2.1.3.2	Bias Temperature Instability	11
2.1.3.3	Time-Dependent Dielectric Breakdown	12
2.1.4	Estado da arte	15
2.1.5	Modelos Compactos	17
2.1.5.1	Modelos para HCI	17
2.1.5.2	Modelos para BTI	18
2.1.5.3	TDDDB	20
2.1.6	Simuladores	21
3	METODOLOGIA	23
3.1	Motivação	23
3.2	Visão geral	23
3.3	Sensores	25
3.4	Geração de perfis	26
3.5	Métodos de estimativa	28
3.5.1	Regressão de Mínimos Quadrados Parciais	31
3.5.2	Distância Euclidiana	34
3.5.3	Correlação de Pearson	35
4	FLUXO DE ENVELHECIMENTO DE CIRCUITOS INTEGRADOS	37
4.1	Estratégia do fluxo	37
4.2	Implementação do fluxo e integração com as ferramentas de envelhecimento	40
4.2.1	Preparação de células lógicas e netlists	40
4.2.2	Preparação dos perfis de envelhecimento	41

4.3	Extração de caminhos críticos	43
4.4	Simulação e degradação	44
4.5	Extração de atrasos e MTTF	45
4.6	Tratamentos dos dados	46
5	RESULTADOS	47
5.1	Caracterização de células	47
5.2	Degradação de circuitos de teste	51
5.3	Métricas e estatísticas	54
5.4	Resultados das simulações dos circuitos de teste	56
5.5	Discussão dos resultados	60
5.6	Críticas ao método	62
6	CONCLUSÕES	65
6.1	Futuros trabalhos	65
	REFERÊNCIAS	67

1 Introdução

*Peace is not an absence of war,
it is a virtue, a state of mind, a
disposition for benevolence,
confidence, justice.*

-Baruch Spinoza

1.1 Motivação

Sistemas computacionais exercem hoje uma tarefa de sumária importância em diversas área de atuação. O poder computacional tem facilitado a solução de problemas antes considerados de longa ou difícil obtenção. Para este intento, a integração, e consequentemente miniaturização, em escalas cada vez mais reduzidas foi necessária, aumentando a quantidade de cálculos que poderiam ser realizados em um menor intervalo de tempo [6].

Como consequência, a quantidade de efeitos indesejados, antes irrisórios, aumentou dramaticamente. Diversos destes causam a indisponibilidade temporária ou permanente de um circuito. Outros podem interferir de forma não-destrutiva na performance deste circuito, tornando-o não confiável. Essa perda de confiabilidade pode ser atribuída, em alguns casos, ao envelhecimento dos dispositivos ao longo de seu uso, seja em condições normais de operação ou não, e que degradam sua performance.

Muitos esforços têm sido realizados na análise dos efeitos que causam este envelhecimento, bem como propostas para detecção e correção, de forma dinâmica, deste comportamento indesejado. Este trabalho de pesquisa visa descobrir através de simulações, e de seus dados obtidos, padrões de comportamento de circuitos integrados que estão gradativamente sendo degradados e propor uma metodologia que dê suporte às promissoras técnicas de verificação, avaliação e atuação contra falhas.

1.2 Roteiro da dissertação

Esta dissertação é organizada em 6 capítulos, incluindo esta introdução. Os capítulos restantes são:

Capítulo 2, Base Teórica, explicando de forma sucinta as bases sobre às quais este trabalho se sustenta. Discorre sobre o conceito geral de confiabilidade, bem como os efeitos e modelos utilizados no estudo da confiabilidade de circuitos eletrônicos.

Capítulo 3, Metodologia, apresenta o método proposto para solução do problema em mãos, além de introduzir conceitos pertinentes para o entendimento da mesma.

Capítulo 4, Fluxo de envelhecimento de circuitos integrados, descreve os passos necessários para envelhecer os dispositivos aqui empregados, desde a sua síntese até a extração dos resultados.

Capítulo 5, Resultados, discorre sobre quais saídas foram obtidas do fluxo, as métricas utilizadas para ajuizar a eficiência dos modelos, discute os resultados e realiza uma autocrítica à metodologia.

Capítulo 6, Conclusões, encerra o trabalho analisando a contribuição da metodologia, as estimativas, os resultados e quais são melhorias futuras para o mesmo, seus pontos fracos e fortes.

2 Base Teórica

*To see the world in a grain of
sand, and to see heaven in a
wild flower, hold infinity in the
palm of your hands, and
eternity in an hour.*

-William Blake

Este capítulo introduz os conceitos essenciais necessários para o entendimento deste trabalho. Na primeira parte são apresentadas as bases de confiabilidade, robustez e falhas em sistemas. A segunda parte apresenta detalhes sobre os efeitos físicos que surgem em transistores CMOS que levam ao envelhecimento dos mesmos, os modelos empregados para estimar tais efeitos.

Em adição, é apresentado o estado da arte em confiabilidade de circuitos integrados e uma introdução aos atuais simuladores que auxiliam nesta tarefa e de que forma eles atuam.

2.1 Confiabilidade

A confiabilidade de um sistema é definida como a “probabilidade de que uma porção de sistema irá durar por pelo menos um tempo previamente especificado sob a ação de condições experimentais especificadas” [2]. Dado que é esperado que um sistema obedeça suas condições de operação especificadas por um tempo mínimo sem que apresente problemas ou se mostre confiável, podemos analisar, prever e informar a sistemas supervisores quais condições levariam a interrupções no funcionamento adequado do sistema observado. Para compor esta análise, é utilizado um conjunto de métricas para eventos de falha e recuperação de falhas para um sistema, quais sejam [7][8]:

1. **MTTF - Mean time to failure:** é o tempo médio que leva até que um sistema falhe;
2. **MTTR – Mean time to repair:** é o tempo médio que o sistema necessita para retornar à operação normal;
3. **MTBF – Mean time between failures:** é o tempo médio que o sistema leva entre falhas. Neste caso é levado em consideração o tempo de recuperação, anteriormente definido;

4. **FIT – Failures in time:** é a razão de falhas que acontecem em um sistema em 1 bilhão de horas.

Já para a definição de uma falha é preciso considerar as seguintes definições [7]:

1. **Faltas:** considerada uma falha física, tais como condutores quebrados, transistores com portas danificadas, resistores queimados;
2. **Erro:** A falta acima definida pode se manifestar em um erro, que é a troca ou incorreta representação da informação que o meio físico deveria permitir, como bits trocados dentro de uma memória;
3. **Falha:** O erro por sua vez pode se manifestar visivelmente como um comportamento indesejado, apresentando-se como operações errôneas, travamentos ou até indisponibilidade de um sistema, incorrendo em uma falha.

A disponibilidade de um sistema é uma função de MTBF e MTTF (ou MTTF e MTTR), podendo ser definida pela equação 2.1 e a relação entre estes é retratada na Figura 1.

$$Disponibilidade = \frac{MTTF}{MTBF} = \frac{MTTF}{MTTF + MTTR} \quad (2.1)$$

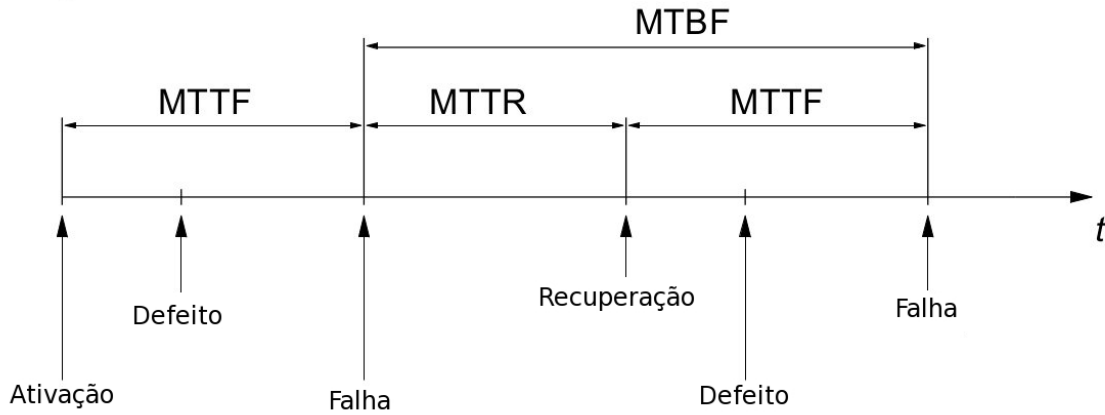


Figura 1 – Relação entre as falhas, defeitos, tempos médios de falha, tempos médios de recuperação e tempo médio entre falhas.

Ao ser definida a probabilidade de um sistema operar conforme planejado até um instante de tempo t , está sendo definido o que se compreende por *Confiabilidade*. Desta forma, a confiabilidade $R(t)$ de um sistema é normalmente descrita através de um expressão probabilística. Esta por sua vez considera que, para o período t anteriormente mencionado, existe uma taxa de falha λ que representa a probabilidade de falha deste sistema para um período t [9].

Na maioria dos casos, uma taxa de falha λ constante é assumida [10], permitindo que a confiabilidade $R(t)$ seja expressada por uma função exponencial, explicitada na equação 2.2.

$$R(t) = e^{-\lambda t} \quad (2.2)$$

Em concordância com definição de MTTF apresentada anteriormente, a mesma pode ser equacionada conforme mostrada na equação 2.3

$$MTTF = \int_0^{\infty} R(t)dt = \frac{1}{\lambda} \quad (2.3)$$

Para este trabalho o MTTF será empregado como a relação entre um acréscimo no atraso do circuito em uma saída especificada a priori, o acréscimo máximo permitido e o tempo de observação de envelhecimento do dispositivo. É importante salientar que será abordada a falha λ após o período de mortalidade infantil.

2.1.1 Confiabilidade em Engenharia

A confiabilidade em engenharia estabeleceu-se como um ramo de pesquisa nos tempos modernos, sendo o termo “confiabilidade” usado extensivamente pelo público em geral e pela comunidade técnica como sendo a capacidade de um sistema ou componente de operar normalmente sobre certas condições [2]. A base para a confiabilidade é a teoria probabilística e estatística, que permitiram o avanço deste campo de estudo tal como é conhecido atualmente. A partir daí, e com o advento da produção em massa de bens de consumo, o estudo da confiabilidade de sistemas se tornou essencial em projetos de engenharia, como foi o caso da (não)confiabilidade de tubos de vácuo [2][11].

Esta preocupação vital perdura até hoje, incluindo sistemas elétricos e eletrônicos de diferentes requisitos, e de forma pervasiva encontra-se em diferentes campos de estudo, tais como: sistemas de potência (modelagem baseada em agentes ou métodos clássicos) [12][13], rádio frequência [14][15], tempo de vida restante para transformadores de potência [16]–[17], sistemas de supervisão e intervenção de estados clínicos [18][19]).

Os métodos e técnicas utilizadas são desenvolvidos, adaptados e/ou expandidos para esses diversos campos de pesquisa levando em consideração suas particularidades e quais fenômenos contribuem para o surgimento de falhas. Apesar destas particularidades, tais metodologias podem ser estudadas em um âmbito mais generalista, sem prejuízo ao estudo de um problema específico [20].

A confiabilidade de circuitos integrados tornou-se essencial para a indústria de CI's, sendo considerada um dos maiores problemas a ser enfrentado em curto prazo, como definido no “The international technology roadmap for semiconductors”[21], explicitando efeitos como *Time Dependent Dielectric Breakdown*, *Bias Temperature Instability* e *Hot-Carrier Injection*.

2.1.2 Técnicas de confiabilidade e robustez

Diversas técnicas tem sido desenvolvidas e propostas para mitigar a redução de confiabilidade e robustez de sistemas. Tais técnicas podem abordar o problema diretamente na camada de *hardware* ou de *software*. Em adição, as abordagens pode ser divididas em dois grandes grupos [22]:

1. **Intolerante à falhas:** Também conhecida como *Prevenção de Falhas*, esta abordagem é considerada conservativa. A premissa neste caso é de que a redução da probabilidade de falha pode ser obtida através da utilização de dispositivos especialmente desenvolvidos para tal intento e/ou da utilização de métodos que propõe prolongar o tempo de vida do sistema;
2. **Tolerante à falhas:** Já nesta abordagem as falhas são toleráveis pelo sistema. Estas técnicas se utilizam de redundâncias para, ou prover tempo extra para que o sistema volte à sua operação normal, ou garantir um caminho extra que, não submetido a uma falha, propague corretamente a informação desejada.

A seção 2.1.2.1 irá sumarizar as propostas de algumas técnicas a nível de *hardware* (*i.e.* nível físico). Este trabalho apresentará posteriormente uma solução intolerante à falhas. É importante salientar que tais técnicas não são, necessariamente, mutuamente exclusivas.

2.1.2.1 Técnicas em hardware

De forma geral, as técnicas a nível físico podem ser distinguidas por [22]:

1. **Modificação das condições ambientais:** engloba um conjunto de medidas que tendem a melhorar as condições ambientais de operação de um sistema, seja através da redução de efeitos degradantes de temperatura, dissipação de calor, radiação, entre outros;
2. **Melhoria de qualidade:** Neste caso é proposto a utilização de materiais com propriedades superiores ao já normalmente utilizados ou a melhoria da qualidade do processo de fabricação dos componentes que compõe o sistema;
3. **Duplicação:** é considerada uma técnica relativamente simples para detecção de falhas. Como o próprio nome indica, há uma duplicação de componentes e, havendo a falha em um destes, é feita uma comparação da informação. Uma diferença indica que uma falha ocorreu;
4. **Códigos para detecção de erros:** para estes casos, é averiguado qual conjunto de informações, de todas as possíveis em um sistema, realmente representa uma

informação válida. Isso significa que para um conjunto de bits, por exemplo, apenas uma determinada combinação deles é esperada. A partir daí, uma comparação é realizada para determinar a validade dos dados. Alguns exemplos básicos desta técnica são os bits de paridade, códigos de Hamming e os *checksums*;

5. **Reconfiguração:** neste caso, após detectado o erro ou falha, o sistema altera parcialmente sua estrutura ou comportamento. Isso significa que ele pode combinar as técnicas anteriores, podendo rotear o caminho percorrido pela informação desejada para uma duplicata dos componentes que estão falhos, alterar o instante de sincronização das informações na saída ou até mesmo simplesmente desligar uma porção do sistema, fazendo com que ele opere de forma degradada.

2.1.3 Confiabilidade em circuitos integrados

Na década de 70, os problemas mais comuns de falha de dispositivos estavam relacionados à problemas de empacotamento e montagem (*e.g.*, corrosão, conexão de fios e contaminação iônica). Somente após o advento de dispositivos integrados menores, efeitos como *Hot Carrier Injection* se tornaram mais preocupantes, uma vez que começaram a afetar o desempenho destes dispositivos.

Encarados inicialmente como de fácil predição (*i.e.*, determinísticos), tais efeitos passaram a ser considerados estocásticos e o primeiro deles a apresentar este comportamento foi o *Time-Dependent Dielectric Breakdown* [23]. Estas mudanças surgiram à medida que a redução do óxido do dielétrico se intensificou, atingindo a escala atômica. Com a continuada redução do tamanho dos dispositivos, entrando então na escala nanométrica, erros estocásticos e variações em nível atômico tornaram-se cada vez mais aparentes.

Variações no processo de fabricação surgem, introduzindo variabilidade nos dispositivos, e aumentando os desafios para projetistas de circuitos integrados. Para este trabalho é importante compreendermos os fenômenos básicos através dos quais cargas ficam presas em determinadas regiões do transistor e como se relacionam com os mecanismos de envelhecimento que serão descritos em seguida. Tendo isto em mente, podemos então dividir os tipos de cargas nas seguintes categorias [1]:

1. **Fixas no óxido:** proveniente de defeitos estruturais na camada de óxido que surgem durante processos de oxidação térmica, são encontradas próximas a interface $Si - SiO_2$. Esta região é normalmente descrita por $Si - SiO_x$, sendo uma representação intermediária entre o substrato e óxido. Assim sendo, são independentes das condições de operação do transistor;
2. **Íons móveis no óxido:** são associados às impurezas no próprio óxido decorrentes dos processos de fabricação da estrutura [24];

3. **Presas no óxido:** juntamente com as cargas presas na interface, tem um importante papel na degradação gradual das características do óxido. Estas cargas possuem energia cinética o suficiente para serem injetadas após vencerem a barreira de potencial da região entre o silício e óxido;
4. **Presas na interface:** semelhantemente, estas cargas possuem uma energia cinética alta o suficiente para vencer a barreira de potencial porém não adentram o óxido, mantendo-se portanto na região interfacial entre o silício e óxido.

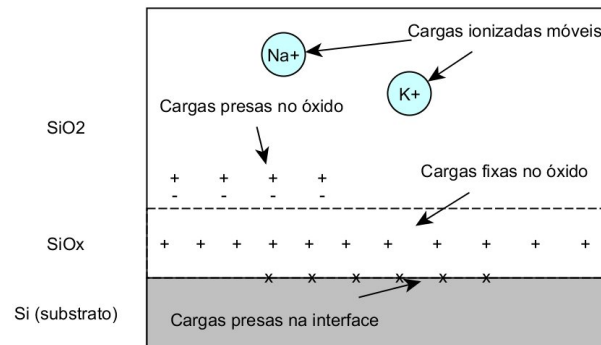
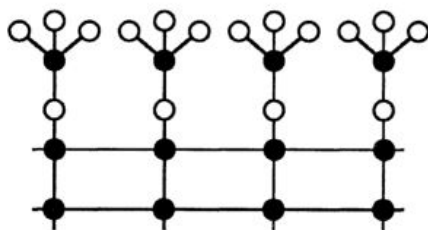
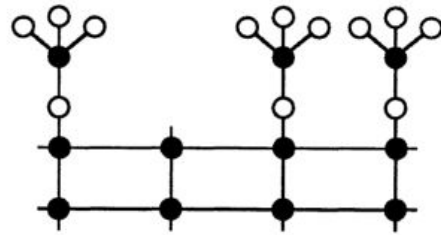


Figura 2 – Tipos de cargas associadas a um estrutura $Si - SiO_2$.

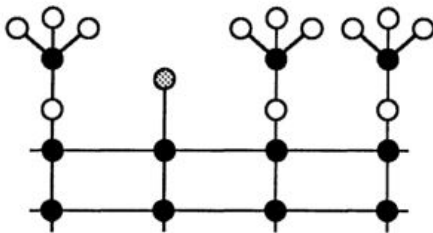
Diversos modelos atômicos foram propostos para descrever as armadilhas que capturam elétrons e lacunas. A maioria destas armadilhas são usualmente ligações incompletas, como mostradas na figura 3:



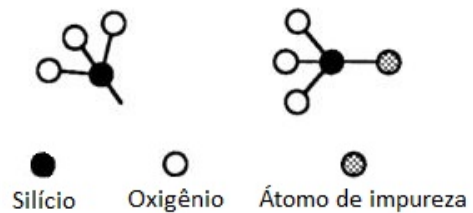
(a) Interface sem imperfeições



(b) Interface com uma ligação incompleta,



(c) Interface com uma impureza,



(d) Tipos de defeitos possíveis: ligação incompleta de silício e impureza no dióxido de silício

Figura 3 – Armadilhas para elétrons e lacunas em uma estrutura de $Si - SiO_2$ [1].

2.1.3.1 Hot Carrier Injection

Hot Carrier Injection (HCI) é um dos efeitos mais antigos e estudados, tendo sido introduzido por Shockley em 1961 para explicar fenômenos em junções $p-n$ à época [25]. O campo elétrico próximo à região de dreno de um transistor (e.g., *MOSFET*) é considerado como a principal causa deste efeito, onde alguns portadores possuem uma energia cinética muito alta em comparação à média dos demais. Tais portadores de alta energia possuem uma temperatura efetiva acima da ambiente (“mais quentes”) e são capazes de vencer o potencial da superfície e são injetados no óxido de porta.

Parte destes portadores permanecem no óxido, gradualmente acumulando cargas e alterando a tensão de limiar V_{TH} e a condutância g_o ao longo da vida do dispositivo. Na década de 80 estes efeitos começaram a chamar a atenção, como consequência da redução contínua da escala dos transistores sem que existisse uma redução da tensão de alimentação [26].

Já na década de 90, o *HCI* foi mitigado com a redução da tensão de operação, redução esta que visava primeiramente um menor consumo de energia [2]. Entretanto, a variação da tensão não tem acompanhado a miniaturização [27][28] dos dispositivos, não contribuindo significativamente com a redução desta degradação. Apesar de ser considerado menos relevante em dispositivos pMOS em comparação à dispositivos nMOS (pois as lacunas são mais pesadas que elétrons) o *HCI* pode, ainda assim, realçar outros efeitos de envelhecimento [29]. Atualmente quatro mecanismos de HCI são largamente conhecidos e estudados [2]:

1. **Channel Hot Electron (CHE):** Alguns elétrons “sortudos” (i.e., *lucky electrons*), possuindo energia o suficiente, vencem a barreira da interface Si/SiO_2 na região do canal próxima ao dreno (ver figura 4a). Caso a tensão na porta V_g seja baixa, o campo elétrico resultante não atrai suficientemente elétrons para a porta. Quando a tensão é aproximadamente igual à tensão do dreno V_d , atinge seu considerado valor máximo;
2. **Substrate Hot Electron (SHE):** Quando uma grande diferença de potencial é aplicada no corpo do transistor, seja positiva ou negativa, os portadores são direcionados para a interface Si/SiO_2 adquirindo energia cinética o suficiente para vencer a barreira ao longo de toda a interface do canal e eventualmente sendo injetado no óxido de porta (ver figura 4b) uniformemente;
3. **Drain Avalanche Hot Carrier (DAHC):** Em determinadas condições de estresse, onde a tensão V_d no dreno é alta e a tensão V_g é baixa, portadores de alta energia cinética podem colidir com elétrons, que antes estavam ligados à estrutura cristalina, cedendo a eles energia, deixando-os livres para condução e criando um par elétron-lacuna no local [30]. Estes, por sua vez, são acelerados mediante influência do campo

elétrico no canal e podem potencialmente vencer a barreira da interface Si/SiO_2 , ficando presos na região ou criando estados interfaciais (ver figura 4c). Pode existir ainda uma componente adicional de corrente no corpo do transistor (i.e., *bulk*);

4. **Secondary Generated Hot Electron (SGHE):** esse efeito compreende a criação de portadores “quentes” (i.e. hot carriers, HC) através da ionização por impacto envolvendo um segundo portador, sendo este também gerado por uma anterior ionização por impacto (e.g., *DAHC*). Os primeiros portadores podem gerar durante seu deslocamento, por influência do campo transversal, novos portadores que, também acelerados por este campo em direção à interface, adquirem energia suficiente para a superar a barreira e atravessar o material (ver figura 4d). Apesar disso, este efeito é considerado como um contribuinte inferior na degradação dos transistores se comparado aos demais efeitos [2].

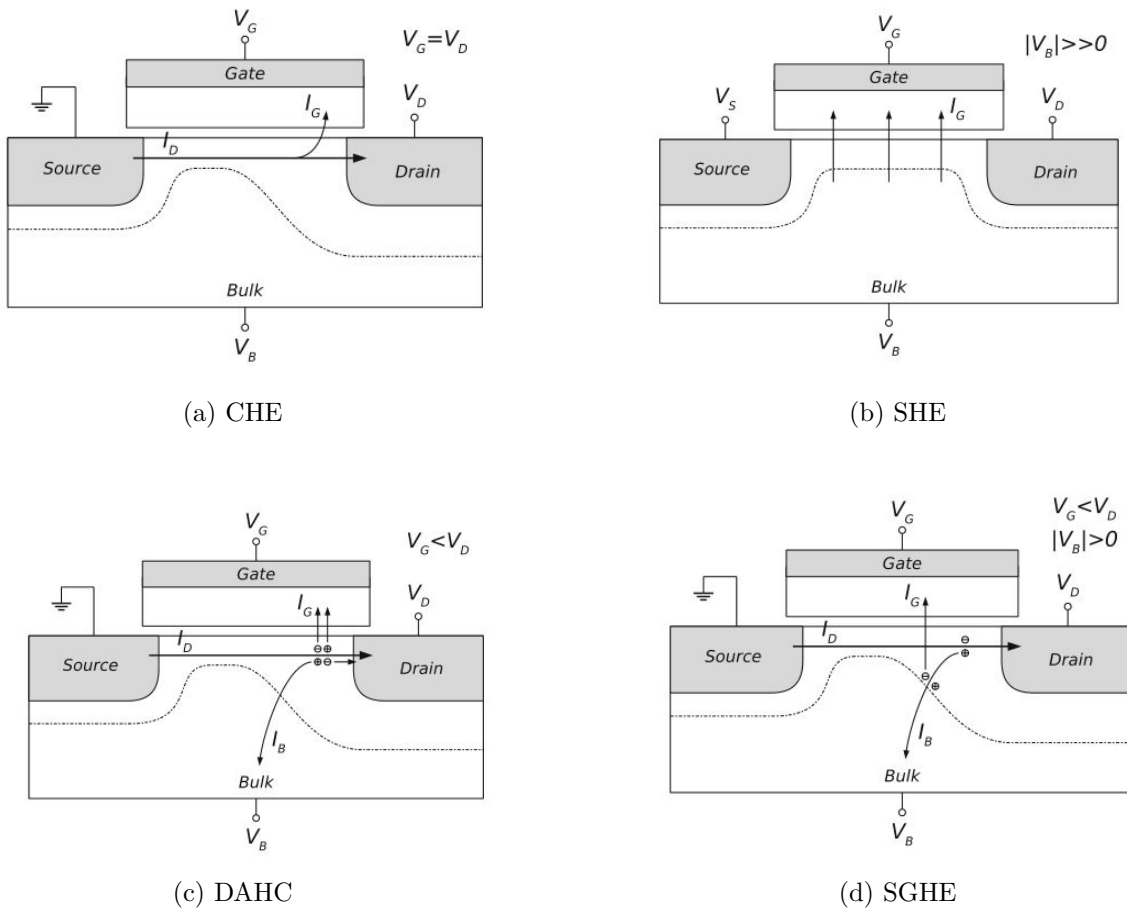


Figura 4 – Mecanismos de injeção de portadores [2].

Estes quatro mecanismos se manifestam no transistor em diferentes condições de operação, sendo o DAHC e o CHE os que mais contribuem para a degradação do dispositivo. Para transistores de $0.35\mu m$, o DAHC é mais proeminente do que os demais efeitos. Já para

dimensões inferiores, o CHE é considerado o efeito de degradação dominante [26]. A variação na tensão de limiar através de HCI foi modelada como dependente do tempo, do campo elétrico no óxido (E_{ox}), do campo elétrico lateral máximo (E_{lat}), da temperatura (T) e do comprimento do transistor (L) :

$$\Delta V_{TH} = A_{HCI} t^{n_{HCI}} \quad (2.4)$$

sendo A_{HCI} quantificado por [31][32][33]:

$$A_{HCI} \propto \frac{1}{\sqrt{L}} \exp(\alpha_{HCI,1} E_{ox}) \exp\left(-\frac{\alpha_{HCI,2}}{E_{lat}}\right) \quad (2.5)$$

onde $\alpha_{HCI,1}$ e $\alpha_{HCI,2}$ são parâmetros específicos da tecnologia.

2.1.3.2 Bias Temperature Instability

Conhecido de forma genérica por BTI (Bias Temperature Instability), é considerada uma das ameaças mais críticas em tecnologias CMOS. Apesar de conhecido há bastante tempo, este efeito tem se tornado mais proeminente em tecnologias abaixo de 90nm [34].

Este efeito é influenciado pela temperatura e se mostra presente quando uma tensão estressante é aplicada na porta do transistor, sendo percebido como uma alteração nos parâmetros de operação do transistor, tipicamente a tensão de limiar V_{TH} . No caso do NBTI, (*i.e.* Negative BTI) esse fenômeno é observado em pFETs e em nFETs para o PBTI. Ainda não há um pleno consenso sobre a origem do BTI (a nível microscópico).

Alguns pesquisadores são de opinião que este efeito surge de uma combinação do aprisionamento de portadores nas imperfeições do óxido e a geração de estados na interface do canal do óxido. Uma propriedade singular deste mecanismo se torna aparente em situações nas quais são aplicadas tensões estressantes sobre o dispositivo. Esta propriedade, conhecida como *relaxamento* ou *recuperação*, indica a redução da degradação em um momento imediatamente posterior ao da aplicação da diferença de potencial estressante.

Por este motivo, a modelagem do BTI é complexa, com indícios de degradação residual após o relaxamento das condições de estresse, sendo indicados por duas componentes: uma permanente e outra recuperável [2]. Alguns estudos modelam então este efeito como consequência de dois mecanismos [35]:

1. Geração de imperfeições próximas a interface entre silício e óxido, resultando numa componente de probabilidade de tunelamento dos átomos de Hidrogênio (H+);
2. E uma componente de recuperação, quantificando a captura e soltura de portadores.

Este comportamento é matematicamente expresso da seguinte forma:

$$\Delta V_{TH} \propto \left[\underbrace{\exp(\alpha_1 V_{GS}) t^{n_P}}_{\text{Componente permanente}} + \underbrace{V_{GS}^{\alpha_2} (C_R + n_R \log_{10}(t))}_{\text{Componente recuperável}} \right] \exp\left(-\frac{E_a}{kT}\right) \quad (2.6)$$

sendo a variação da tensão de limiar, ΔV_{TH} , uma função do campo elétrico (E_{ox}) aplicado no óxido de porta do transistor e a temperatura (T). Os parâmetros α_1 e α_2 são fatores de escalonamento da voltagem e dependentes da tecnologia empregada, E_a é a energia de ativação correspondente, C_R , n_P e n_R são expoentes de ajuste do tempo para as componentes permanente e recuperável, e k é a constante de Boltzman.

É importante salientar que a equação 2.6 só é válida para uma voltagem estressante que seja invariante, sendo necessária um modelo mais preciso para o BTI quando a tensão aplicada é variante no tempo [36]. Os efeitos de BTI são considerados determinísticos para transistores de tamanho micrométrico, com uma variação idêntica nos parâmetros dos transistores [27].

Porém, ao serem reduzidos para a escala nanométrica, estes efeitos passam a ser considerados como estocásticos. A mudança dos parâmetros dos transistores varia com o tempo. Além disso, estas mudanças são intensificadas pelo aumento do desvio padrão deste mesmos parâmetros, também no tempo.

2.1.3.3 Time-Dependent Dielectric Breakdown

A operação adequada de um transistor MOS exige que o mesmo esteja operando sob as corretas condições, incluindo aí a magnitude do campo elétrico aplicado na sua região isolante. Estes campos elétricos podem eventualmente exceder o campo elétrico máximo suportado pelo dielétrico e levá-los à ruptura. Como consequência há um acréscimo substancial da corrente na porta deste transistor.

Mas mesmo durante a operação normal de transistores, diferentes estados de ruptura podem se manifestar, sendo importante diferenciar as categorias existentes e em quais condições elas podem surgir, quais sejam [2]:

1. **Soft breakdown (SBD)**: é observada como uma perda parcial das propriedades dielétricas, em comparação com o *Hard Breakdown*, e que não impede a operação do mesmo. Apesar disso, o dispositivo passa a operar em condições não idealizadas pela sua especificação. Entretanto possuem apenas um pequeno aumento da corrente de porta e a completa quebra do dispositivo é improvável [2];
2. **Progressive breakdown (PBD)**: É um passo gradativo que surge após o SBD, sendo percebido como um pequeno incremento na corrente de porta no decorrer do tempo. É presenciado em estruturas de óxido ultra-fino, tipicamente abaixo de 2.5nm de espessura;
3. **Hard breakdowns (HBD)**: é percebida como um aumento significativo da corrente de porta e uma redução da controlabilidade da tensão de porta do dispositivo. Este

tipo de ruptura em qualquer transistor é considerada como sendo capaz de causar uma falha permanente no circuito.

O TTDB é entendido como sendo este processo de danificação gradual no óxido graças a estes fortes campos elétricos. Apesar da importância do TDDB para o estudo da confiabilidade de transistores, os modelos físicos a nível molecular existentes são, de certa forma, especulativos [3], dividindo-se em três categorias principais: *a)* modelamento pelo campo elétrico; *b)* modelamento pela corrente elétrica; *c)* e modelamento pela combinação de ambos. Apesar de não existir um modelo largamente aceito, existem algumas características que são amplamente reconhecidas [3]:

1. TDDB é fortemente dependente do campo;
2. TDDB pode ser fortemente dependente da temperatura;
3. TDDB pode ser dependente da polaridade;
4. Alguns defeitos, armadilhas e ligações incompletas podem ser formadas durante o estresse causado por TDDB;
5. A geração de falhas leva a criação de um caminho condutivo de infiltração que, eventualmente, ocasiona a quebra do dispositivo;
6. O TDDB pode ser estatisticamente descrito por distribuições de Weibull;
7. Os parâmetros de aceleração de partículas no campo crescem diretamente proporcional à constante dielétrica κ .

Assim sendo, o entendimento do mecanismo por detrás da quebra do dielétrico é de vital importância. Os defeitos gerados durante o estresse são normalmente consideradas como neutros pois não há mudança significativa do diagrama de banda de energia para um dispositivo MOS. Além disso, acredita-se que estes defeitos devem estar relacionados a um processo de geração de defeitos a nível molecular [3]. Ainda mais: a geração de defeitos no material é considerada irreversível.

Mesmo após o recozimento do dispositivo ou a inversão do campo, não há evidência de um aumento significativo do tempo de vida. O entendimento passa a ser de que o dano causado permanece na rede de silício e é praticamente irreversível. Este tipo de comportamento é então associado à mudanças na ligação dos átomos. Durante a ruptura da ligação covalente do íon de silício, por influência do campo elétrico local, há a criação de ligações incompletas (*i.e.* dangling bonds) que possuem apenas um elétron e de uma estrutura diferente (ver figura 5).

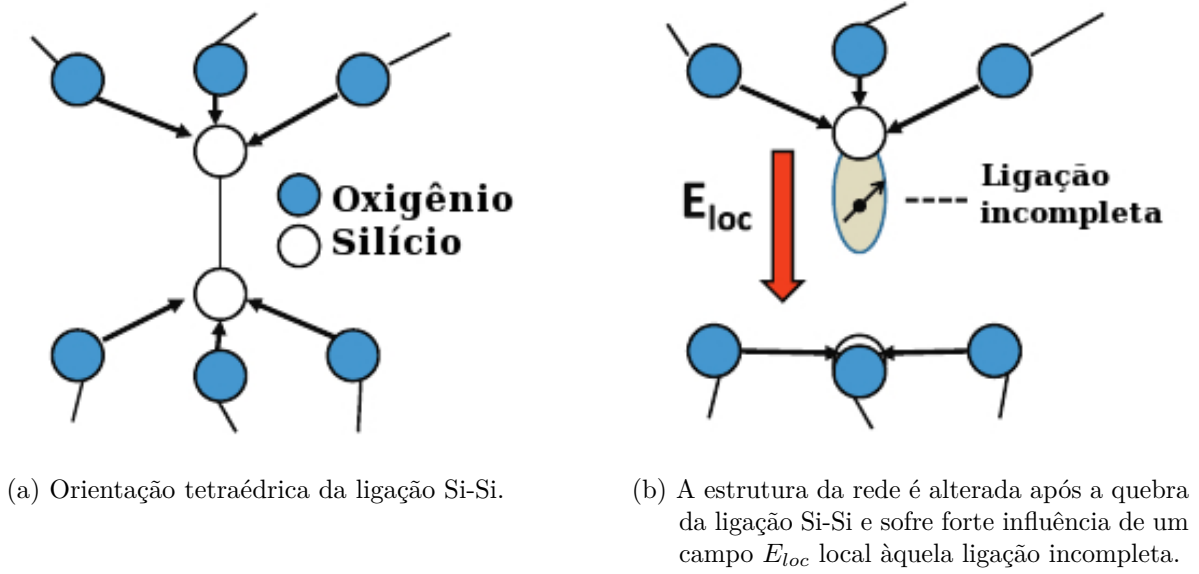


Figura 5 – Geração de ligações incompletas a partir de vacâncias [3].

Esta estrutura pode ser “relaxada” a tal ponto que mesmo após o recozimento ou alteração do sentido do campo, como mencionado anteriormente, a ligação não será reparada, sendo esta uma das observações mais significantes do estresse provocado pelo TDDB.

Diversos modelos têm sido propostos para as categoria anteriormente mencionadas. Para o *HDB*, os mais conhecidos são: modelo termoquímico (*i.e.* termochemical), modelo injeção-anodo-lacuna (*i.e.* anode-hole-injection) e o modelo direcionado à voltagem (*i.e.* voltage-driven). O modelo termoquímico é conhecido como *modelo E* (*i.e.* E model), pois considera que existe uma correlação direta entre o campo elétrico aplicado no óxido do dispositivo E_{ox} e o tempo de vida do dispositivo; sendo as ligações químicas no óxido fracas o suficiente para que este campo elétrico as rompa.

Assim sendo, o tempo para quebra (*i.e.* time to breakdown) t_{BD} é expresso por [37]:

$$t_{BD} \propto \exp(-\gamma E_{ox}) \exp\left(\frac{E_a}{kT}\right) \quad (2.7)$$

onde γ ($\gamma \approx 1.1$ década/MV/cm) é um fator de aceleração do campo e E_a ($E_a \approx 0.6-0.9$ eV) é a energia de ativação térmica.

Já o modelo injeção-anodo-lacuna é conhecido como *modelo 1/E* e prevê uma dependência entre o t_{DB} e o campo elétrico, sendo uma consequência da captura de lacunas em determinadas regiões do óxido de porta. Ela é expressa por [33]:

$$t_{BD} \propto \exp\left(\frac{\beta}{E_{ox}}\right) \exp\left(\frac{E_a}{kT}\right) \quad (2.8)$$

onde β ($\beta \approx 350\text{MV/cm}$) é um fator de aceleração do campo dependente do processo de fabricação. Existem ainda modelos que combinam os dois propostos acima. De todo caso, estes modelos vem sendo revisados para óxidos de porta inferiores à 5nm [2]. Na seção 2.1.5.3 serão apresentados modelos mais precisos e modernos.

2.1.4 Estado da arte

As diversas técnicas de confiabilidade de circuitos integrados utilizadas para reduzir o impacto do envelhecimento se concentram em contramedidas tais como: blindagem de circuitos, faixas de segurança (*i.e.* guard-banding), balanceamento de cargas, voltagem dinâmica e escalonamento de frequência [38][39], se encaixando em uma outra das categorias definidas na seção 2.1.2.1.

Para suportar algumas destas estratégias deve-se prever a falha de um sistema com antecedência, usando tal informação para ajudar na tomada de decisões destas contramedidas antes mesmo que qualquer erro se manifeste. Esta estratégia se diferencia da abordagem clássica, que atua somente após o surgimento de erros. Uma abordagem que vêm sendo explorada é a de utilização de sensores variados em diversos locais do sistema, coletando informações sobre diversos parâmetros de operação ao longo do tempo.

Eles podem ser analisados internamente ou externamente na procura de anomalias que indiquem que o sistema irá falhar. A frequência desta coleta pode ser ajustada a nível de design ou ser adaptativa durante a operação, permitindo que esta frequência seja adequada ao longo da operação para melhor capturar comportamentos que sejam anômalos ou que levem a uma degradação do sistema [40]. Exemplos de dados a serem coletados são: temperatura, tensão de alimentação, atrasos, grau de atividades e outros que o projetista julgar importante.

Estes sensores podem ser *in-situ*, adicionados a caminhos críticos específicos para detectar o não cumprimento de restrições de tempo; *online self-testing*, onde um sistema entra em modo de teste e executa rotinas de *Built-On Self Test* (BIST) que auxiliam na determinação da degradação corrente; *sensores de envelhecimento*, que utilizam circuitos de referência espalhados pelo sistema que informam alteração em parâmetros de operação; e *monitores de estresse*, que coletam informações de temperatura, alimentação e carga do sistema.

Técnicas como *shadow transistors* incluem transistores paralelos entre aqueles que, previamente identificados, são considerados mais vulneráveis. Desta forma, não apenas o tempo de vida é melhorado como também são compensadas as falhas inseridas durante a fase de fabricação (*e.g.* circuitos abertos). Esta medida não reduz a taxa de falha λ de um determinado transistor mas inclui um caminho adicional em sua *netlist*, reduzindo a probabilidade de falha, já que ambos os caminhos, tanto o original como a duplicata,

devem falhar para que haja o sistema falhe [41].

Esta técnica pode ser realizada através da inserção guiada ou aleatória de *shadow transistors*, conduzindo a resultados promissores. A técnica pode ser combinada com transistores de espessuras de óxido diferentes, aumentando ainda mais a confiabilidade de um sistema. Entretanto, tais técnicas têm como consequência o aumento da área, acréscimo do consumo dinâmico e do atraso médios.

Outra forma de aumentar a confiabilidade de um sistema é através da utilização de *sleep transistors*. Estes transistores de alta tensão de limiar V_{TH} , sejam pMOS ou nMOS, são utilizados para interromper a alimentação de partes do sistema que eventualmente esteja em modo de espera. Entretanto, esta técnica necessita considerar diversos aspectos, tais como: ativação por nMOS ou pMOS, qual será a distribuição espacial dos *sleep transistors*, largura e profundidade da porta, otimização área ocupada pelo corpo do transistor, fugas e eficiência [42].

Esses, por sua vez, podem ser utilizados para alternar a ativação de módulos do sistema. Ao desconectar a alimentação de um módulo ele não irá, idealmente, apresentar tensões ou correntes e consequentemente nenhum campo elétrico. A consequência é uma redução da temperatura e ausência de atividade. Isto é importante na redução de efeitos indesejados, tais como: eletromigração [43], ruptura do óxido de porta e NBTI.

Então, durante a fase ociosa das portas estes efeitos serão eliminados ou reduzidos drasticamente. Como consequência desejada, o MTTF é prolongado [44]. Esta técnica é conhecida como *Alternating Module Activation* (AMA).

Em abordagens mais recentes, a melhoria da confiabilidade de um sistema está sendo obtida através da adoção de técnicas de aprendizado de máquina (*i.e.* machine learning) que permitem representar e prever qual é a degradação que um sistema sofre através da observação de caminhos críticos. Ao observar por um período de tempo t qual é o acréscimo no atraso das saídas destes caminhos, estas informações são relacionadas à carga de trabalho a qual o sistema está submetido. Esta carga e este tempo t são determinados durante o design e a predição é *online*.

Tais informações de design (tempo de observação, caminhos críticos e saídas observáveis) são decididas pelo projetista e alteram a precisão da predição *online*. Porém, durante a operação, as informações utilizadas pelo modelo podem ser atualizadas à conveniência do projetista. Esta decisão impacta a área ocupada pelo sistema, pois acarretará na inclusão de dispositivos à estrutura, tornando-a maior. Parte desta tarefa de monitoramento e predição pode ser dividida entre componentes de hardware e software, mitigando uma porção deste acréscimo de área e permitindo a integração de modelos mais complexos.

Trabalhos realizados nesta área relatam que esta estratégia apresenta uma boa

precisão e é considerada promissora [45]. Entretanto, estes trabalhos consideram somente alguns dos fenômenos de envelhecimento, tais como NBTI, e apenas parte das condições ambientais que contribuem para o envelhecimento de transistores (*e.g.* temperatura, tensão de alimentação e atividade).

Essas diversas abordagens endereçam o mesmo problema mas sugerem ações diferentes, porém complementares. Isto é importante pois permite que tecnologias mais robustas de verificação, avaliação e relatórios se tornem viáveis. Ao atuarem em diversas camadas de abstração, que têm se tornado cada vez mais complexas, novas tecnologias podem integrar essas soluções, mitigando ou corrigindo erros e falhas nessas diferentes camadas. Assim sendo, a confiabilidade de um sistema é melhorada como um todo.

Em [46] temos um exemplo da empregabilidade de tais abordagens em uma tecnologia que, sendo uma proposta universal, integra e arbitra a comunicação entre estes elementos.

2.1.5 Modelos Compactos

Na seção 2.1.3 foram apresentados os principais efeitos físicos responsáveis pelo envelhecimento de circuitos integrados. Foram apresentados ainda alguns modelos que se propõem a explicar as relações entre os parâmetros de operação e as degradações observadas. Esta seção discute o desenvolvimento de modelos compactos para transistores MOS.

Será discutido primeiramente os modelos existentes para HCI e em seguida para um enfoque nos modelos existentes para *baia*s temperature instability e TDDB.

2.1.5.1 Modelos para HCI

Considerado no início como um efeito de grande importância para explicar a degradação de transistores, o HCI se tornou um efeito menos dominante à medida que a tensão de alimentação foi sendo reduzida. Entretanto ele ainda é considerado um problema para tecnologias atuais visto que o campo elétrico vem aumentando gradativamente com a redução do tamanho dos transistores.

Diversos modelos foram publicados ao longo dos anos para explicar o HCI. Entre eles o modelo do elétron “sortudo” (*i.e.* *lucky electron model*, LEM) é um dos mais conhecidos. Além disso, muito dos outros modelos que surgem posteriormente são derivados do LEM. Em 1985 [33] foi um dos primeiros a introduzir o modelo baseado no LEM. As propostas que se seguem utilizam a mesma teoria mas propõem uma formulação analítica diferente ou propõem uma acomodação do modelo para tecnologias CMOS mais avançadas.

No trabalho apresentado em [32] os efeitos de HCI e NBTI foram unificados em um mesmo modelo, levando em consideração a geração de armadilhas interfaciais

devido a desassociação das ligações de Si-H. Posteriormente esta proposta é utilizada para desenvolver outros modelos para tecnologias mais avançadas. Um modelo semi-empírico foi proposto e pode ser expressado de acordo com a quantidade de estados interfaciais gerados:

$$\Delta N_{IT} = C_1 \left[\frac{I_{DS}}{W} \exp \left(-\frac{\phi_{IT,e}}{q\lambda_e E_{lat}} \right) T_{str} \right]^n \quad (2.9)$$

onde W é a largura de canal do transistor, I_{DS} é corrente dreno-fonte, E_{lat} é o campo elétrico lateral máximo no dreno, T_{str} é o tempo de estresse ao qual o transistor é submetido, $\phi_{IT,e}$ é a energia crítica necessária para que os elétrons criem uma armadilha interfacial [2].

Apesar desses modelos serão baseados no LEM, eles incluíam apenas uma fonte de portadores “quentes”. Entretanto, como mostrado anteriormente na seção 2.1.3.1, existem outros três mecanismos que foram identificados ao longo dos anos de pesquisa e que podem potencialmente resultar em uma mudança no comportamento do transistor ao longo do tempo.

Posteriormente, [2] propôs um modelo baseado em uma abordagem conhecida como reação-difusão (*i.e. reaction-diffusion*, RD) proposta em [32] e também baseado no LEM. O modelo RD possui alguns pontos negativos ao tentar modelar o BTI. Um desses é a incapacidade de considerar o aprisionamento de óxidos.

Entretanto esse efeito está relacionado ao fenômeno de recuperação do BTI, que ocorre após o estresse aplicado no dispositivo cessar. Entretanto, esse modelo ainda pode ser utilizado para modelar o efeito de HC, visto que as armadilhas interfaciais são geradas próximas ao fim da região do dreno, onde efeito de recuperação é negligenciável.

É utilizado então um conjunto de duas equações diferenciais que descrevem a geração das partículas de hidrogênio próximas à interface óxido canal e qual é a sua difusão em direção ao contato da porta. Essas equações são então resolvidas numericamente ao sinalizar um transistor. Porém essa análise não é apropriada para o estudo de um circuito inteiro.

Em [2], é proposta uma expansão deste modelo, válido somente para tensões de estresse contínuas (modelo DC). Entretanto circuitos integrados atuais são submetidos a tensões que variam no tempo. Para incluir essa variação uma extensão ao modelo DC deve ser incluída, um modelo AC.

Ainda em [2], ambos os modelos, AC e DC, são combinados para avaliar o impacto do HCI no comportamento do transistor graças à sinais estressantes variantes no tempo.

2.1.5.2 Modelos para BTI

Para transistores cuja tecnologia está abaixo de 90 nanômetros, o NBTI é considerado uma das maiores ameaças a confiabilidade de dispositivos. Assim sendo, é considerado

essencial prever quais são os impactos desse efeito na performance de circuitos integrados. Apesar de também presente, o PBTI recebe um enfoque menor.

Um dos principais desafios enfrentados ao se analisar este efeito é explicar o comportamento dele a nível microscópico. Na década de 70, algumas pesquisas já sugeriam que a criação de estados na interface estava relacionada ao mecanismo de difusão do hidrogênio. Esse mecanismo foi posteriormente chamado de reação-difusão (RD). Em 2006 foi sugerido que o NBTI consiste de uma geração de estados na interface mais ou menos constante combinada a um componente de recuperação relacionada ao aprisionamento de lacunas [29].

Modelo de Reação-Difusão

Este modelo propõe que uma reação ativada termicamente entre lacunas e ligações de Si-H na interface dielétrico/substrato do MOSFET resulta na geração de ligações incompletas. Estas por sua vez levam ao surgimento de estados interfaciais. As partículas de hidrogênio se afastam da interface por difusão e vão em direção ao dielétrico de porta.

O modelo mais conhecido foi proposto em [47] e posteriormente aperfeiçoado por outras pesquisas [48][49]. De acordo com o modelo RD. Esta reação eletroquímica é então definida como depende da temperatura e do campo; e a taxa da geração de armadilhas é descrita pela seguinte equação diferencial:

$$\frac{dN_{IT}}{dt} = k_F (N_0 - N_{IT}) - k_R N_H(0) N_{IT} \quad (2.10)$$

sendo N_{IT} uma variável para representar a quantidade de armadilhas interfaciais, N_0 é a quantidade inicial de ligações incompletas de Si-H e k_F é uma constante que representa a taxa de desassociação depende do campo no óxido. k_R é a taxa de recozimento das ligações incompletas associada aos átomos de hidrogênio. $N_H(0)$ é a concentração de hidrogênio na interface. Entretanto, os átomos de hidrogênio podem difundir da interface para dentro do óxido, sendo este processo descrito por:

$$\frac{dN_H}{dt} = D_H \frac{d^2 N_H}{dx^2} \quad (2.11)$$

onde N_H é a concentração total de hidrogênio no óxido e D_H é uma constante de difusão. Estas ligações incompletas atuam como armadilhas “doadoras” que contribuem para variação na tensão de limiar V_{TH} , que pode ser representada por:

$$\Delta V_{TH} = \frac{q N_{IT}}{C_{ox}} \quad (2.12)$$

sendo q a carga de um elétron e C_{ox} a capacitância do óxido de porta.

A figura 6 mostra de forma simplificada onde estes fenômenos ocorrem.

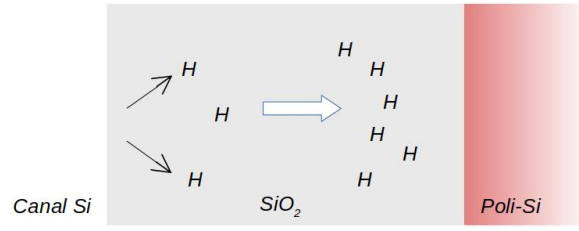


Figura 6 – Representação da quebra de ligações Si-H e Si-O na interface Si/SiO_2

2.1.5.3 TDDDB

Como explicado anteriormente na seção 2.1.3.3, a quebra do óxido de porta pode se dar em etapas, indo de um Soft Breakdown (SBD) a um Hard Breakdown (HBD); sendo necessário então descrever quais mecanismos estão envolvidos nestas duas etapas e como podemos modelar tal comportamento.

Modelo para Hard Breakdown

Existem diversos modelos que se propõem a explicar este fenômeno, sendo um dos mais conhecidos o modelo *termoquímico*. Entretanto ele foi gradativamente modificado para atender a resultados obtidos em experimentos posteriores, que indicavam que as formulações descritas nas equações 2.7 e 2.8 não eram mais válidas para óxidos de porta com espessuras menores do que 5nm [50].

O modelo foi expandido posteriormente para incluir a dependência da área e a distribuição do *tempo para ruptura* [51], visto que anteriormente só era prevista o tempo para quebra característico. Entretanto, alguns dispositivos falhavam antes desse tempo.

Considerando estas mudanças, um modelo completo para tempo de quebra em óxidos ultrafinos (*i.e.* $t_{ok} < 5nm$) pode ser descrito por:

$$t_{BD} \propto \left(\frac{1}{WL} \right)^{1/\beta} F_{BD}^{1/\beta} V_{GS}^{a+bT} \exp \left(\frac{c}{T} + \frac{d}{T^2} \right) \quad (2.13)$$

sendo W e L a largura e o tamanho do transistor, respectivamente. As constantes β , a , b , c e d são dependentes da dimensão do transistor CMOS [52].

O tempo de quebra segue uma distribuição de Weibull e sua função de probabilidade de falha cumulativa F_{BD} é descrita por:

$$F_{BD} = 1 - \exp \left(- \frac{t}{t_{BD,63}} \right)^\beta \quad (2.14)$$

Modelo para Soft Breakdown

Passando a ser importante com o advento de dielétricos ultrafinos, o SBD não se traduz na perda de controlabilidade da tensão de porta do transistor mas sim em um

pequeno acréscimo da corrente de porta e na criação de um caminho de infiltração no óxido. Porém, assim como o HBD, o SBD também segue uma distribuição de Weibull.

Essa infiltração no óxido é gradual e em concordância com o pequeno incremento da corrente de porta encontrado experimentalmente [53]. Isso significa também que um único ponto de infiltração tem uma baixíssima probabilidade de impedir o funcionamento de um circuito. Porém um conjunto deles ao longo do tempo aumenta essa chance, sendo necessário modelar a quantidade de pontos de infiltração após um período de estresse.

A probabilidade de existirem n defeitos de SBD em um instante de tempo χ pode ser descrita por uma distribuição de Poisson [51]:

$$P_n(t) = \frac{\chi^n}{n!} \exp(-\chi) \quad (2.15)$$

$$\chi = \left(\frac{t}{t_{SBD,63}} \right)^\beta \quad (2.16)$$

$$t_{SBD,63} = t_{SBD,ref} \left(\frac{WL}{A_{ref}} \right)^{1/\beta} \left(\frac{V_{GS}}{V_{ref}} \right)^\gamma \quad (2.17)$$

onde $t_{SBD,63}$ está associado ao tempo de quebra no 63º percentil para um transistor de referência de área A_{ref} estressado a uma tensão V_{ref} . β e γ são parâmetros dependentes do processo do transistor [53].

Este modelo entretanto não é dinâmico, pois não considera mudanças na tensão de estresse aplicada, mudanças estas que são consequência da alteração dos parâmetros de operação do transistor; e esta alteração é induzida justamente pelo envelhecimento.

Em [2] o modelo é alterado para considerar este dinamismo e descrever a probabilidade $P_n(t_2)$ de encontrar n pontos de SBD em um instante de tempo t_2 . Ela está relacionada à probabilidade $P_{n'}(t_1)$ de encontrar n' pontos de SBD no instante de tempo t_1 :

$$P_n(t_2) = \sum_{n'=0}^{\infty} \left[P_{n'}(t_1) \frac{\Delta\chi^{n-n'}}{(n-n')!} \exp(-\Delta\chi) \right] \quad (2.18)$$

$$\Delta\chi = \left(\frac{t_2 - t_1}{t_{SBD|V_{GS1}=V_{GS2}}} \right)^\beta \quad (2.19)$$

sendo V_{GS1} a tensão de estresse em t_1 e V_{GS2} em t_2 .

2.1.6 Simuladores

Com os efeitos de envelhecimento de transistores possuindo um impacto cada vez maior, a simulação da confiabilidade de circuitos eletrônicos se tornou uma parte importante em um fluxo moderno de design. Simulações precisas da confiabilidade permitem que o

projetista aumente significativamente o design, satisfaça especificações mais complexas e ainda assim garanta uma operação confiável do sistema.

Diversas ferramentas surgiram na década de 70 após o advento do simulador *SPICE*. Apartir daí, a complexidade das ferramentas seguiu o avanço dos métodos utilizados para caracterizar o envelhecimento de transistores e são construídos ao redor do *SPICE* [2].

Este trabalho utiliza o simulador Relxpert[©] (Cadence Systems[©]), baseado no *BERT*[©], que fornece um modelo analítico descrevendo cada efeito de envelhecimento. Normalmente, os modelos devem ser fornecidos pelos fabricantes, mas podem ser desenvolvidos pelo próprio pesquisador. Esses modelos devem descrever as mudanças nos parâmetros dos dispositivos como função da idade do transistor. Apesar de utilizar esta ferramenta, a metodologia que será apresentada independente da escolha feita pelo pesquisador.

3 Metodologia

*How do you know but ev'ry Bird
that cuts the airy way, Is an
immense world of delight, clos'd
by your senses five?*

-William Blake

3.1 Motivação

Como exposto anteriormente na seção 2.1.1, a confiabilidade de um sistema pode ser quantificada por seu MTTF. Existem diversos fatores que podem afetá-lo e estes podem se manifestar em diversas fases da concepção e operação de um sistema: na variação do processo de fabricação, no design, durante a interação com fatores externos estranhos ao sistema (*p.ex.* descargas elétricas, radiação).

Desta forma, diversos métodos e técnicas são viáveis para mitigar a degradação de transistores nestas fases, combinados ou não. Entretanto, todas estas abordagens alteram, eventualmente, a tensão de operação, a atividade do sistema e indiretamente a temperatura.

Partindo dessa premissa, este trabalho se propõe a averiguar como os parâmetros de tensão e temperatura estão variando ao longo da operação do sistema e investigar a relação destes com a degradação.

3.2 Visão geral

Neste trabalho, os dados de condições ambientais (*p.ex.* tensão e temperatura) auxiliam na criação de um *perfil de operação* para um sistema observado ao longo do tempo. Esse perfil é representado por um conjunto de variações das condições ambientais que, empregadas sobre os dispositivos, contribuem para a degradação e, consequentemente, envelhecimento dos mesmos. Isso nos permite associar de forma única esse perfil a um MTTF.

O MTTF pode ser associado a diferentes eventos dentro do ciclo de vida de um sistema, tais como: tempo médio até o surgimento de falhas permanentes em porções específicas dos circuitos, falhas transitórias e intermitentes, tempo necessário até a falha completa ou parcial de um sistema. Neste trabalho, o MTTF é considerado como sendo o tempo necessário para que um sistema sofra uma degradação alta o suficiente que o

force a violar suas especificações de projeto. Em outras palavras, os circuitos não são mais capazes de cumprir suas especificações de temporização.

A partir disso, o *Tempo de Vida Restante* (i.e. *Remaining Useful Lifetime*, *RUL*) é definido como o tempo remanescente do sistema durante o qual ainda funcionará, antes de falhar [54]. O RUL é, usualmente, determinado pela diferença entre seu tempo de vida atual e seu MTTF. Ao assumir esta proposição, a predição do MTTF através da observação da variação de temperatura T , tensão de operação V e da atividade α permite prever também o RUL.

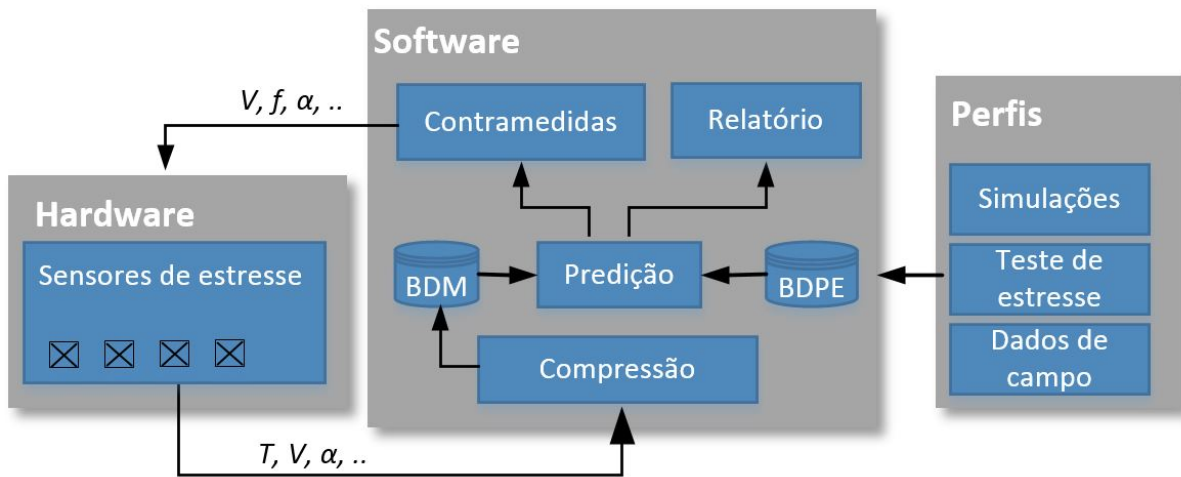


Figura 7 – Abordagem proposta para análise de envelhecimento.

Quais perfis serão definidos e como serão utilizados são decisões que podem ser tomadas pelo projetista durante o projeto. De forma geral, a unidade de *Predição*, ilustrada na figura 7, comunica-se com duas bases de dados: uma responsável por armazenar os perfis *online* medidos pelos sensores e eventualmente comprimidos pela unidade de *Compressão*; e a outra por armazenar perfis pré-existentes que foram criados *offline*, durante o projeto ou atualizados posteriormente, durante a operação do sistema. Essas bases são denominadas, para este trabalho, de Banco de Dados de Medições (BDM) e Banco de Dados de Perfis Envelhecidos (BDPE), respectivamente.

Através de um método de estimativa (que fica a critério do projetista), a unidade de *Predição* estima qual é o tempo de vida restante do sistema e em seguida envia esta informação à unidade de *Contramedidas*. Esta, por sua vez, realiza alguma ação objetivando mitigar a degradação do sistema, quando for o caso. Concomitantemente, essas informações são enviadas a uma unidade de *Relatório*, onde serão armazenadas.

Caso decida por mitigar a degradação, a unidade de Contramedida deve enviar sinais de controle que permitam ao Hardware alterar os parâmetros que estão contribuindo com o envelhecimento do sistema (*p. ex.* tensão de operação, grau de atividade).

A abordagem apresentada acima é generalista e flexível. Isso significa que ela

independe de quais parâmetros se deseja obter através dos sensores, de quais métodos sejam utilizados para predição dos MTTFs ou de onde se originam os perfis do BDPE: se são provenientes de simulações, testes de estresse ou dados de campo.

Outra vantagem desta abordagem é a capacidade de popular o banco de dados de perfis com novas observações durante a atual operação do sistema. Um BDPE que, por exemplo, possuía dados provenientes de simulações *offline* pode, à medida que o sistema envelhece, ter seu conteúdo atualizado com dados de campo, melhorando as estimativas posteriores. Ainda mais: é possível que os modelos utilizados sejam atualizados com melhorias provenientes de otimizações ou até mesmo substituídos por modelos diferentes e melhores.

Posto isso, nas seções subsequentes serão descritas as técnicas e métodos utilizados neste trabalho e utilizará a figura 7 como referência.

3.3 Sensores

O uso de sensores para observação de grandezas como temperatura, variações de tensão, grau de atividade e corrente de fuga é bem conhecido e objetivo de inúmeras pesquisas. Diversos deles são projetados especificamente para medir desgastes provenientes dos mecanismos de HCI, BTI e TDDB [55][56][57][58], descritos na seção 2.1.3.

Entretanto, neste trabalho é considerado o uso de sensores de estresse para obtenção dos perfis de operação, medindo parâmetros como temperatura, tensão de alimentação ou grau de atividade. Um sensor de atividade realiza o monitoramento das entradas primárias de um sistema ou de uma parte delas para auxiliar na estimativa de estresse. Como exemplo de um sensor de atividade, pode-se utilizar circuitos que observem a variação das entradas e associe a cada transistor a existência ou não de degradação [45]. A figura 8 mostra um exemplo genérico que auxilia o entendimento desta abordagem, onde $E1$, $E2$ e $E3$ são sinais de entrada que podem alterar o sinal de saída S :

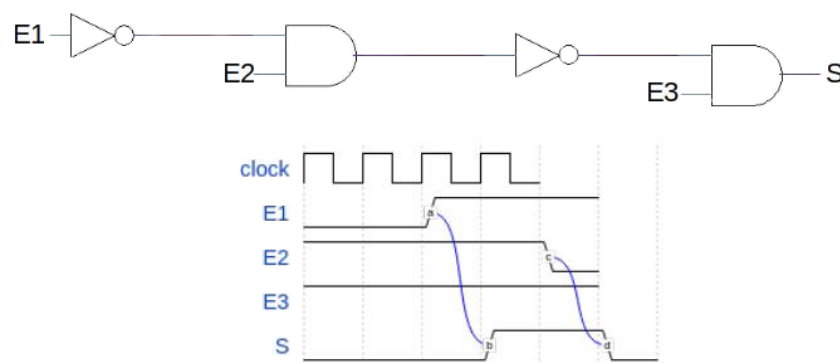


Figura 8 – Alteração das entradas refletem-se como uma atividade nas portas lógicas.

No exemplo, o sinal da entrada $E1$ é alterado após um período de tempo, mudando o sinal da saída S . O mesmo ocorre para a entrada $E2$. Na técnica de monitoramento do grau de atividade, essas alterações podem resultar em uma degradação das portas lógicas que estão presentes no caminho percorrido pelo sinal que vai desde $E1$ até a S : duas portas NOT e duas AND2.

Isso significa que, a mudança dos sinais $E1$, $E2$ e $E3$ pode alterar como os transistores de cada porta lógica se comportam (*p.ex.* número de transistores pMOS ativos das portas AND2 do exemplo). Esse comportamento, por sua vez, pode estar associado a um envelhecimento maior ou menor dos transistores decorrentes da degradação por NBTI, por exemplo.

Uma vez determinada essa atividade e a degradação, elas são associadas entre si para que futuras degradações sejam estimadas. O aumento da área total do sistema, decorrente da inserção destes circuitos de monitoramento, é dependente da quantidade de circuitos do sistema a serem monitorados e da precisão desejada na estimativa da degradação [45].

Já para o caso de sensores de temperatura e tensão de alimentação, a frequência com que essas grandezas variam ao longo do tempo impõem restrições no que se refere ao seu armazenamento, visto que armazenar cada transição de temperatura, por exemplo, exigiria dos sensores ou da unidade de Compressão uma capacidade de armazenamento restritiva. Isto por sua vez aumentaria a área total do sistema, o que pode ser ainda mais restritivo em muitos casos.

Fica então a cargo do projetista escolher qual será a frequência de coleta dos dados ou se será usada alguma métrica que represente a grandeza desejada para um intervalo de medição (*p.ex.* a temperatura média em um período de tempo δt). É necessário salientar que esta abordagem incorre na redução de precisão na aferição realizada pelos sensores.

Em contrapartida, a memória necessária para armazenar estas informações é consideravelmente menor. Para este trabalho será considerado que uma métrica (como a média) foi utilizada para descrever as informações de operação do sistema, que posteriormente serão inseridas tanto o BDM quanto no BDPE.

3.4 Geração de perfis

No contexto exposto acima, onde um conjunto de sensores amostram condições ambientais, é preciso estabelecer quais informações se deseja representar ao armazenar estas condições. O intuito desta representação é permitir que seja estimado o MTTF, e consequentemente o RUL, de um sistema.

Para este trabalho são utilizadas a tensão de alimentação V e temperatura T como

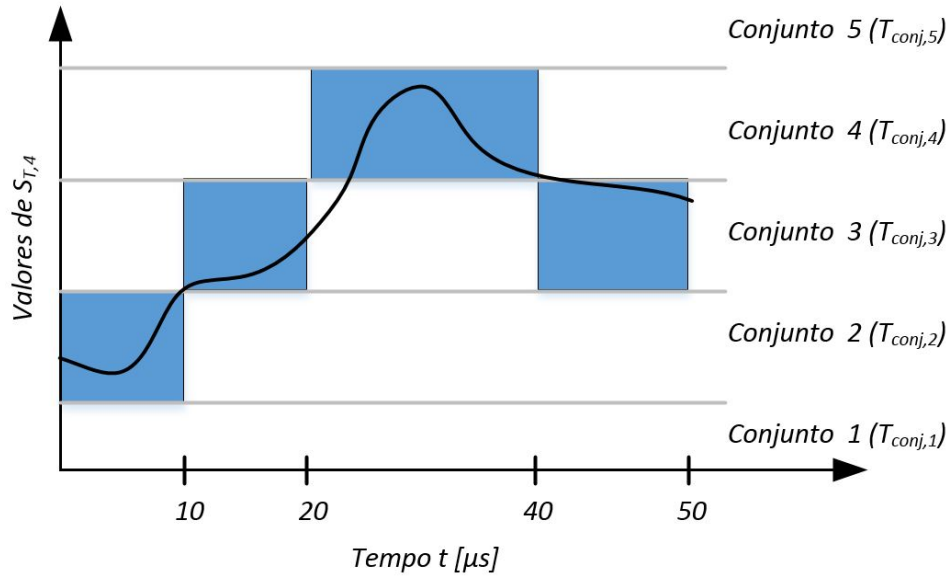


Figura 9 – Compressão dos dados medidos e representação dos mesmos através de perfis de envelhecimento.

condições ambientais a serem representadas no BDM e BDPE. A literatura e pesquisas exploradas na seção 2.1.4 evidenciam uma forte relação entre V e T com os efeitos de degradação explanados na seção 2.1.3.

Considerando que as condições ambientais impostas ao sistema variam com sua operação, podemos exemplificar os valores de temperatura obtidos por um sensor $S_{T,4}$ ao longo de t como mostrado na figura 9. Neste período o sistema está operando sob diversas faixas de temperatura T , denominadas de *Conjunto 1*, *Conjunto 2*, *Conjunto 3*, *Conjunto 4* e *Conjunto 5*. Os sensores serão agrupados pelo seu tipo k (*p.ex.* Temperatura T , tensão de alimentação V ou atividade α) e denotados por $s_{k,i}$, sendo i o índice de um sensor específico (*p.ex.* 1, 2, 3, ...).

Exemplo: Um sistema possui 6 sensores, sendo dois de temperatura e quatro de tensão de alimentação. Os dois sensores serão denotados por $s_{T,1}$ e $s_{T,2}$. Já os de tensão são denotados como $s_{V,1}$, $s_{V,2}$, $s_{V,3}$ e $s_{V,4}$.

Isso significa que, para cada sensor $s_{k,i}$ existe um vetor $S_{k,i}[\]$ com N conjuntos para cada sensor. Um elemento n do vetor $S_{k,i}[\]$ quantifica o tempo de permanência do sensor $s_{k,i}$ em uma determinada faixa de valor.

No exemplo mostrado na figura 9, para uma medição realizada em um período $\Delta t = 10\mu s$ e um vetor $S_{T,4}[\]$ inicialmente vazio, o sensor $s_{T,4}[\]$ permanece no conjunto 2 de $t = 0\mu s$ a $t = 10\mu s$. Consequentemente, $S_{T,4}[2]$ recebe o valor 10. Seguindo esta lógica o vetor resultante após $50\mu s$ é $S_{T,4} = [0, 10, 20, 20, 0]$.

A criação dos conjuntos de perfis permite quantificar o estresse aplicado ao sistema e a relação entre os parâmetros dos perfis e o tempo de vida esperado dos circuitos

monitorados pelos sensores. Cada perfil é inserido no BDPE conforme exposto na tabela 1.

Cada linha w contém o percentual de tempo $p_{T(w,x)}$ e $p_{V(w,y)}$ sob os quais um ou mais circuitos monitorados estão submetidos a uma temperatura pertencente ao conjunto $T_{conj,x}$ e uma tensão de alimentação que pertence ao conjunto $V_{conj,y}$; sendo x e y os índices dos conjuntos utilizados para descrever o sistema (*p. ex.* $x = 1, 2, \dots, X$ e $y = 1, 2, \dots, Y$).

Como exemplo, a primeira linha ($w = 1$) informa que os circuitos monitorados experimentam durante um percentual $p_{T(1,1)}$ de seu tempo de operação t_{oper} uma temperatura pertencente ao conjunto $T_{conj,1}$, um percentual $p_{V(1,1)}$ para uma tensão de alimentação pertencente ao conjunto $V_{conj,1}$ e assim por diante para as colunas subsequentes. Ao final, é esperado que após operar sob estas condições durante t_{oper} seu tempo de vida seja de $MTTF_1$.

Tabela 1 – Base de Dados de Perfis Envelhecidos

$T_{conj,1}$...	$T_{conj,X}$	$V_{conj,1}$...	$V_{conj,Y}$	MTTF
$p_{T(1,1)}$...	$p_{T(1,X)}$	$p_{V(1,1)}$...	$p_{V(1,Y)}$	$MTTF_1$
$p_{T(2,1)}$...	$p_{T(2,X)}$	$p_{V(2,1)}$...	$p_{V(2,Y)}$	$MTTF_2$
...
$p_{T(M,1)}$...	$p_{T(M,X)}$	$p_{V(M,1)}$...	$p_{V(M,Y)}$	$MTTF_M$

Aplicando a abordagem anterior à figura 9 obtemos o excerto representado na tabela 2:

Tabela 2 – Excerto de um BDPE com dados exemplares.

$T_{conj,1}$	$T_{conj,2}$	$T_{conj,3}$	$T_{conj,4}$	$T_{conj,5}$	$V_{conj,1}$	$V_{conj,2}$	$V_{conj,3}$	$V_{conj,4}$	$V_{conj,5}$	MTTF
0	10	20	20	0	0	10	20	20	0	$MTTF_1$

Este exemplo enfatiza a vantagem de perfis granulares (*i.e* que consideram variação nos perfis durante a operação) em comparação a modelos estáticos na obtenção do RUL de um sistema. Estes perfis podem ser obtidos através de simulações de envelhecimento, dados de campo ou testes acelerados de estresse. Este trabalho propõe, no capítulo 4, um fluxo de envelhecimento que simule diversos circuitos e extraia esses perfis e utiliza-se de simuladores como os descritos na seção 2.1.6.

3.5 Métodos de estimativa

Utilizando ferramentas de simulação e envelhecimento de circuitos aliados a esta abordagem de perfis, é possível estimar o tempo de vida restante de um sistema conhecido.

Sendo assim, é necessário determinar a relação entre uma entrada do BDM e o BDPE; e consequentemente o tempo de vida restante. Estas informações podem ser usadas nos métodos de estimativa que serão apresentados a seguir para determinar, aproximadamente, em qual estado o circuito está e o que acontecerá se ele continuar a operar nestas condições.

Um modelo linear geral [59] pode ser utilizado, inicialmente, para estimar a relação entre o atraso de saída e as temperaturas e tensões de entrada; mais especificamente um modelo de regressão linear simples (RLS). Uma RLS ajuda a responder às seguintes perguntas sobre o dado a ser trabalhado [4]:

1. Existe uma relação entre um conjunto de temperatura $T_{conj,x}$ ou um conjunto de tensão $V_{conj,y}$ e o MTTF?
2. A relação pode ser descrita linearmente?

A RLS descreve uma variável de resposta Y como dependente de um conjunto de variáveis explicativas x (também chamado de *preditor*) da seguinte forma:

$$Y \approx \beta_0 + \beta_1 X + \epsilon \quad (3.1)$$

sendo β_0 e β_1 duas constantes desconhecidas que representam a interceptação da reta com o eixo vertical e a inclinação da mesma, respectivamente. Para representar uma variável ou constante desconhecida e que foi estimada através do modelo, será utilizado o símbolo $\hat{\cdot}$. Isso significa que \hat{y} indica uma predição da variável de resposta Y para $X = x$.

Para se estimar $\hat{\beta}_0$ e $\hat{\beta}_1$ é necessário utilizar-se de dados pré-existentes do sistema a ser analisado. Considerando que o projetista possua n observações desse sistema na forma de pares representados por $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, que consistem de uma medida de X e uma de Y , o objetivo da RLS é obter os coeficientes de $\hat{\beta}_0$ e $\hat{\beta}_1$ e um modelo linear que melhor se ajuste ao dados. Ele é descrito como:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_i \quad (3.2)$$

sendo i o índice do i -ésimo preditor.

O método mais comum de se obter este modelo é através do critério de *mínimos quadrados* [4]. Como exemplo, consideremos um conjunto de dados como os mostrados na figura 10 e um modelo representado pela equação 3.1. Um i -ésimo preditor X , representado por x_i , tem como resultado um i -ésimo dado observado \hat{y} para esta equação.

Entretanto, ao inserir na equação 3.1 um valor de preditor $X = x_i$ cuja resposta Y seja conhecida, o \hat{y} calculado não necessariamente será igual ao Y esperado. A diferença é conhecida como *resíduo* e é representada por ϵ . Assim sendo, o i -ésimo resíduo entre um Y esperado e um \hat{y} estimado é dado por:

$$\epsilon_i = Y_i - \hat{y}_i \quad (3.3)$$

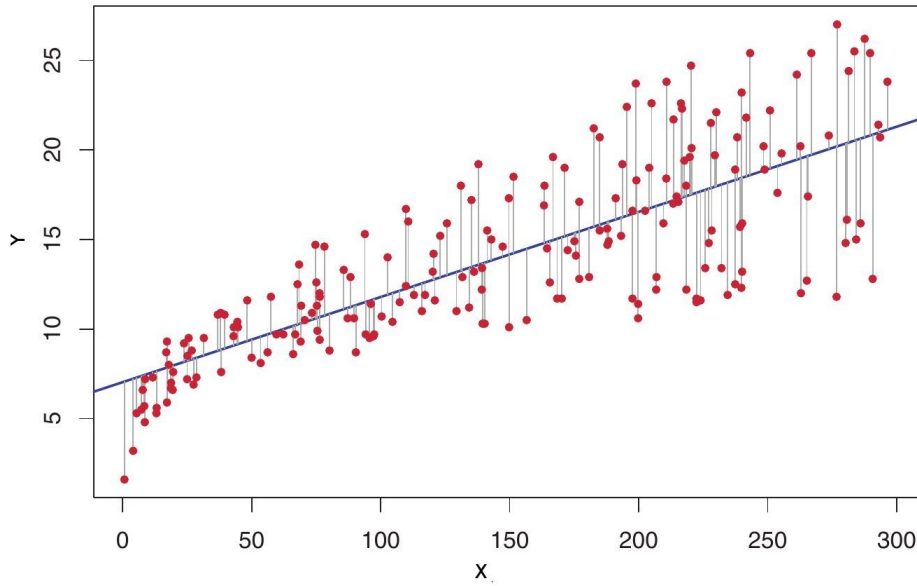


Figura 10 – Modelo linear ajustado a um conjunto de dados de exemplo, representado pela reta azul. Os pontos vermelhos representam os dados coletados e seus respectivos resíduos [4].

O critério de *mínimos quadrados* estima $\hat{\beta}_0$ e $\hat{\beta}_1$ de forma a minimizar a *soma dos quadrados dos resíduos* (SQR) que é equivalente a:

$$SQR = (y_1 - \hat{\beta}_0 - \hat{\beta}_1 x_1)^2 + (y_2 - \hat{\beta}_0 - \hat{\beta}_1 x_2)^2 + \cdots + (y_n - \hat{\beta}_0 - \hat{\beta}_1 x_n)^2 \quad (3.4)$$

Rearranjando as equações 3.2 e 3.4:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (3.5)$$

onde \bar{y} e \bar{x} são as médias amostrais de y e x , respectivamente. Um modelo de RSE descreve de forma simples a relação entre uma grandeza, que funciona como uma variável explicativa, e outra que se deseja estimar, a variável observada. Entretanto, para o problema que este trabalho tenta solucionar, não é suficiente relacionar somente um preditor a uma observação. Considerando um BDPE maior, representada pela tabela 3, não é possível utilizar somente uma variável explicativa pois o MTTF será função de quatro preditores: $T_{conj,1} = 30$, $T_{conj,2} = 50$, $V_{conj,1} = 1.0$ e $V_{conj,2} = 1.1$.

Tabela 3 – Tabela de exemplo para regressão linear simples.

30	50	1.0	1.1	MTTF
13	1	12	2	1.2
7	8	3	12	2.7
1	1	2	2	5.3
5	5	4	7	3.5
7	8	4	5	2.0

Para o exemplo da tabela 3, uma RLS não é capaz de responder os seguintes questionamentos:

1. Quão forte é a relação entre conjuntos de $T_{conj,x}$, $V_{conj,y}$ e o MTTF?
2. Qual dos conjuntos mais contribui para a redução do MTTF?

Considerando o exemplo da figura 9 e a subsequente tabela 2, existem dez variáveis explicativas para uma saída desejada, *p. ex.* MTTF. Realizar uma regressão para cada um dos preditores não é satisfatório. Uma melhor abordagem é expandir a RLS para uma Regressão Linear Múltipla (RLM), dando a cada preditor seu próprio coeficiente de inclinação β . Seu modelo é expresso então por:

$$Y \approx \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p + \epsilon \quad (3.6)$$

Assim como foi estimado para uma RLS, os coeficientes serão calculados usando o critério de *mínimos quadrados*. Para os valores disponíveis na tabela 3, os coeficientes são:

Tabela 4 – Coeficientes do modelo de RLM para o exemplo da tabela 3.

	Interceptação	$T_{conj,1}$	$T_{conj,2}$	$V_{conj,1}$	$V_{conj,2}$
Coeficiente	5.765	-0.295	-0.242	-0.056	0.092

Entretanto uma regressão linear múltipla exige uma independência entre essas variáveis explicativas [60], premissa que não podemos garantir dado que os percentuais $p_{T(w,x)}$ e $p_{V(w,y)}$ podem estar correlacionadas. Em adição, a degradação causada pelos conjuntos $T_{conj,x}$ e $V_{conj,x}$ serão dependentes da degradação subsequente dos conjuntos $T_{conj,x-1}$ e $V_{conj,x-1}$.

Ao invés disso, três métodos são propostos: Regressão de mínimos quadrados parciais (Partial Least Square Regression, PLS-R), Distância Euclidiana (DE) e Correlação (COR).

3.5.1 Regressão de Mínimos Quadrados Parciais

A Regressão de Mínimos Quadrados Parciais (PLS-R) é uma técnica que pertence à categoria dos modelos lineares generalizados e que reúne um grupo de algoritmos. Ao contrário da RLM, é menos exigente quanto a existência de correlação entre as variáveis. A PLS-R é classificada como uma abordagem que transforma os preditores e os ajusta a um modelo de mínimos quadrados utilizando-se das variáveis que foram transformadas.

Considere uma combinação linear dos preditores originais p e representada por Z_1, Z_2, \dots, Z_M para $M < p$, onde:

$$Z_m = \sum_{j=1}^p \alpha_{jm} X_j \quad (3.7)$$

sendo $\alpha_{1m}, \alpha_{2m}, \dots, \alpha_{pm}$ constantes e $m = 1, \dots, M$.

Um modelo de regressão linear pode então ser ajustado por mínimos quadrados e equacionado como:

$$y_i = \theta_0 + \sum_{m=1}^M \theta_m z_{im} + \epsilon_i, i = 1, \dots, n \quad (3.8)$$

Neste caso, os coeficientes de regressão são dados pela constante θ . É possível então definirmos valores de α tais que o modelo para as variáveis transformadas se mostre melhor do que o modelo original.

Por reduzido entende-se que, ao invés de estimarmos $p + 1$ coeficientes $\beta_0, \beta_1, \dots, \beta_p$, são estimados $M + 1$ coeficientes $\theta_0, \theta_1, \dots, \theta_M$. Dado que $M < p$, como enunciando anteriormente, a dimensão do modelo foi reduzida.

Esta abordagem é compartilhada entre os métodos de compressão. Sumarizadamente, eles possuem duas etapas:

1. Obtenção dos preditores transformados Z_1, Z_2, \dots, Z_M .
2. Ajuste do modelo utilizando M preditores.

Entretanto, quais e como os preditores Z_1, Z_2, \dots, Z_M são obtidos varia entre métodos. A proposto em comum é de que, geralmente, uma pequena quantidade deles é suficiente para explicar a variabilidade dos dados originais e sua relação com a resposta. É assumido que a direção para as quais as entradas X_1, X_2, \dots, X_p mostram uma maior variação são as mais associadas à saída Y .

Observando o exemplo da figura 11, é perceptível que a linha verde segue a direção para a qual as observações mais variam. Esta linha é denominada de *primeiro componente principal*, representado por Z_1 e expresso na equação 3.7.

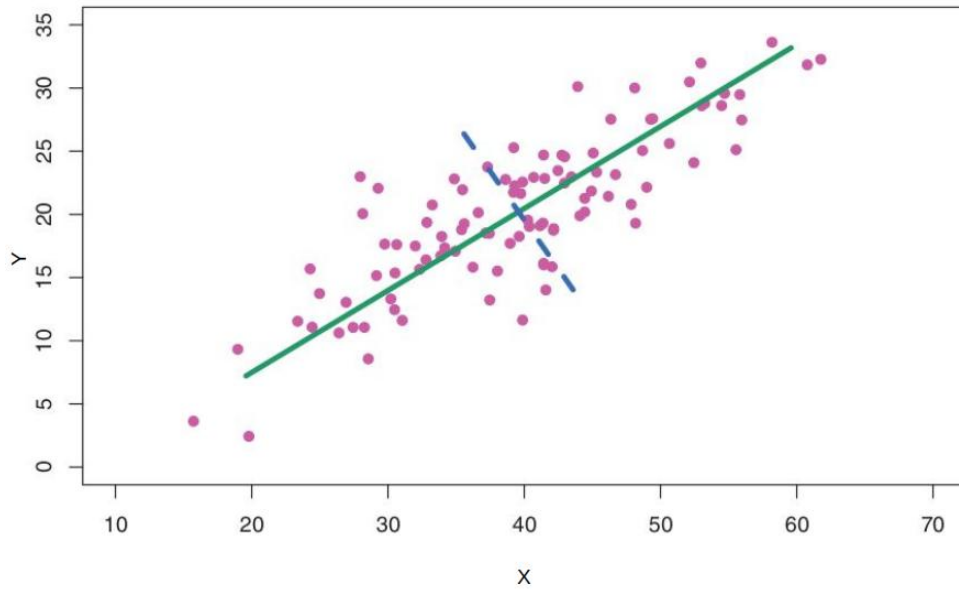


Figura 11 – Representação de componentes principais. A linha verde representa o primeiro componente e a linha azul o segundo. [4].

Isso significa que, se os dados observados Y forem projetados nessa reta verde, as projeções apresentarão como resultado a maior variância possível se comparadas com uma projeção em qualquer outra linha que siga a variação dos dados. Em outras palavras, é o vetor que mais próximo está das observações.

Entretanto podemos obter mais componentes que ajudem a descrever os dados. Na figura 11 um *segundo componente principal* Z_2 é representado pela linha azul tracejada. Ele é uma combinação linear das variáveis que não estão correlacionadas à Z_1 e que, dada esta condição de não-correlação, apresente a maior variância possível.

Como os dados de exemplo da figura 11 são bidimensionais, podemos obter no máximo dois componentes principais. Caso contrário, mais podem ser obtidos e sempre seguindo a premissa de que devem maximizar a variância mas sem estar correlacionados aos componentes anteriores, ou seja, são ortogonais aos mesmos.

Na técnica de *Mínimos Quadrados Parciais* (Partial Least Squares, PLS), o primeiro componente Z_1 é obtido através do ajuste de cada α_{j1} da equação 3.7 para que sejam iguais aos coeficientes da regressão linear simples de Y sobre X_j [4]. Em seguida, o segundo componente Z_2 é computado de forma similar ao Z_1 , porém utiliza os resíduos obtidos da regressão de Z_1 .

Os dados residuais são interpretados como sendo a informação remanescente que não foi explicada pelo primeiro componente principal. Os resíduos são então projetados ortogonalmente e o Z_2 obtido através de uma regressão linear destes dados.

Essa técnica implica que os dados do BDPE e BDM podem ser reduzidos para uma representação de menor dimensão [61]. O PLS tenta maximizar a variância dos estimadores

e auxilia a identificar quantos deles são necessários para explicar os dados do BDM e BDPE. É possível então descrever o MTTF de um sistema como uma RLM cujos coeficientes são iguais aos componentes obtidos.

Assim sendo, o MTTF será uma função dessa regressão e pode ser expressada da seguinte forma:

$$MTTF = \hat{y} = \hat{\beta}_0 x_0 + \hat{\beta}_1 x_1 + \cdots + \hat{\beta}_n x_n \quad (3.9)$$

onde β_i são o ponto de interceptação da reta e coeficientes da RLM e que ajustam a contribuição dos estimadores x_i . A quantidade i de estimadores pode ser determinada pelo projetista, que deve procurar a quantidade de estimadores que melhor explique seus dados.

3.5.2 Distância Euclidiana

A Distância Euclideana é definida como a representação da distância entre dois pontos representados no espaço euclidiano [62]. Esta métrica é particularmente útil pois permite uma representação trivial da distância entre dois elementos que se deseja analisar.

Na PLS-R é feita uma estimativa a partir de dados pré-existentes e um modelo linear é obtido. Entretanto, na DE é possível se estimar qual dos perfis $p_{T(w,x)}$ e $p_{V(w,y)}$ pré-existentes no BDPE mais se aproxima do perfil medido pelos sensores. Essa proximidade é dada pela distância euclidiana.

A distância $d(a, b)$ entre duas séries de pontos $a = (a_1, a_2, \dots, a_N)$ e $b = (b_1, b_2, \dots, b_N)$ com N valores cada pode ser representada pela expressão seguinte:

$$d(a, b) = \sqrt{\sum_{i=1}^N (a_i - b_i)^2} \quad (3.10)$$

Para a estimativa do MTTF, é necessário calcular a distância entre cada entrada do BDPE e o perfil obtido pelos sensores. Usando como exemplo o BDPE representado na tabela 3, que possui 4 dimensões para suas variáveis explicativas, as distâncias euclidianas para um perfil dado, por exemplo, por $S_{T,4} = [3, 8]$ e $S_{V,4} = [9, 6]$ são dadas como:

$$d(a_1, b_1) = \sqrt{(13 - 3)^2 + (1 - 8)^2 + (12 - 9)^2 + (2 - 6)^2} = 13.19 \quad (3.11)$$

$$d(a_2, b_2) = \sqrt{(7 - 3)^2 + (8 - 8)^2 + (3 - 9)^2 + (12 - 6)^2} = 9.38 \quad (3.12)$$

$$d(a_3, b_3) = \sqrt{(1 - 3)^2 + (1 - 8)^2 + (2 - 9)^2 + (2 - 6)^2} = 10.86 \quad (3.13)$$

$$d(a_4, b_4) = \sqrt{(5 - 3)^2 + (5 - 8)^2 + (4 - 9)^2 + (7 - 6)^2} = 6.24 \quad (3.14)$$

$$d(a_5, b_5) = \sqrt{(7 - 3)^2 + (8 - 8)^2 + (4 - 9)^2 + (5 - 6)^2} = 6.48 \quad (3.15)$$

Logo, o perfil de menor distância euclidiana dentre os armazenados no BDPE é o quarto. Consultando a tabela 3, o MTTF pertencente à entrada do BDPE que possui a

menor distância euclideana é utilizado então para mensurar o tempo de vida restante do sistema. No exemplo fictício desta tabela é estimado que o $MTTF = 3.5$.

3.5.3 Correlação de Pearson

Um terceiro método proposto foi a utilização da Correlação, que essencialmente mostra a dependência e associação entre duas ou mais variáveis [63]. Para identificar uma possível relação entre os dados a serem analisados é definido um *coeficiente de correlação*. É de suma importância salientar que esta métrica não necessariamente implica em causalidade.

Existem diferentes métodos para análise de correlação. Entre eles existe a *Correlação r de Pearson* (CP), também conhecida como *correlação paramétrica*. Ela calcula a dependência linear entre duas variáveis X e Y , um coeficiente cuja representação simbólica é r , e que varia entre $+1$ e -1 . Um valor r que se aproxima cada vez mais de ± 1 é interpretado como um grau de associação “perfeita” entre duas variáveis. Caso r se aproxime de 0, é dito que não há correlação ou ela é fraca.

A sinalização \pm indica apenas a direção da correlação, sendo positiva para “+” e negativa para “-”. Isso significa que em uma correlação positiva, quando uma das variáveis cresce ou decresce em valor, a outra variável acompanha o mesmo sentido de acréscimo ou decréscimo. O contrário pode ser dito da correlação negativa. A correlação entre dois vetores x e y é quantificada pelo seu coeficiente $r_{x,y}$ e pode ser descrito por:

$$r_{x,y} = \frac{\sum_{i=1}^n (x - \bar{x})(y - \bar{y})}{\sqrt{\sum_{i=1}^n (x - \bar{x})^2 \sum_{i=1}^n (y - \bar{y})^2}} \quad (3.16)$$

$$r_{x,y} = \frac{cov(x, y)}{\sigma_x, \sigma_y} \quad (3.17)$$

onde cov é a covariância; \bar{x} e \bar{y} são as médias de x e y ; n é o tamanho desses vetores e σ é o desvio-padrão da série [60].

Considerando a tabela 3, um perfil de interesse dado por $S_{T,4} = [3, 8]$ e $S_{V,4} = [9, 6]$, os coeficientes de correlação r entre cada perfil da tabela e o perfil de interesse é:

$$r = [-0.2963478, -0.3748790, 0.4364358, -0.3504383, -0.3450328] \quad (3.18)$$

Estamos interessados no resultado com maior correlação dentre eles, no caso $r_4 = 0.4364358$. Esse coeficiente corresponde à correlação do 3º perfil da nossa tabela de exemplo, cujos valores das variáveis estimativas são $S_{T,4} = [3, 8]$, $S_{V,4} = [9, 6]$ e seu $MTTF = 5.3$.

Em suma, é realizado o cálculo da correlação entre o perfil medido e a as entradas do BDPE, encontrando-se aí a entrada de menor correlação e seu respectivo MTTF.

4 Fluxo de envelhecimento de circuitos integrados

Open your mind to what I shall disclose, and hold it fast within you; he who hears, but does not hold what he has heard, learns nothing.

-Dante Alighieri

Um dos pontos mais importantes para a simulação e obtenção de dados que deem apoio às futuras pesquisas e projetos é a decisão de quais soluções técnicas, das disponíveis, será utilizada para obter resultados coerentes com experimentos reais. Este é um ponto crucial deste trabalho pois os dados obtidos e o comportamento observado devem dar apoio à tomada de decisões por parte de sistemas que atuam na recuperação de falhas.

Para garantir esta fidelidade, foi escolhida a ferramenta Relxpert[©]. Este simulador é desenvolvido especialmente para cálculo de envelhecimento de circuitos integrados e considera efeitos como HCI, TDDB e NBTI/PBTI. Dada à experiência e proximidade da Cadence com o mercado de semicondutores, temos disponível uma ferramenta confiável para simulação e extração de resultados.

Outro fator importante é o uso de parâmetros de envelhecimento coerentes com os dispositivos reais. Ao evitar a modelagem destes parâmetros (e da criação de modelos para os efeitos de HCI, TDDB e BTI) é possível integrar a solução aqui apresentada a qualquer modelo. Como resultado, é proposta uma solução independente das ferramentas escolhidas, da tecnologia empregada e da metodologia utilizada no cálculo dos parâmetros.

Para nossa conveniência, utilizamos parâmetros disponibilizados pela própria ferramenta e que se assemelham às tecnologias de 90nm empregada pelas *foundries* de dispositivos semicondutores.

4.1 Estratégia do fluxo

O fluxo de envelhecimento para extração é composto de seis etapas principais sumarizadas abaixo e representadas na figura 12:

1. Criação de uma tabela com condições ambientais pré-estabelecidas;

2. Identificação de um caminho crítico para análise e criação de uma *netlist* primitiva;
3. Envelhecimento da *netlist* primitiva e criação de uma envelhecida;
4. Simulação da *netlist* envelhecida;
5. Extração dos atrasos relativos e do MTTF;
6. Criar e popular um BDPE.

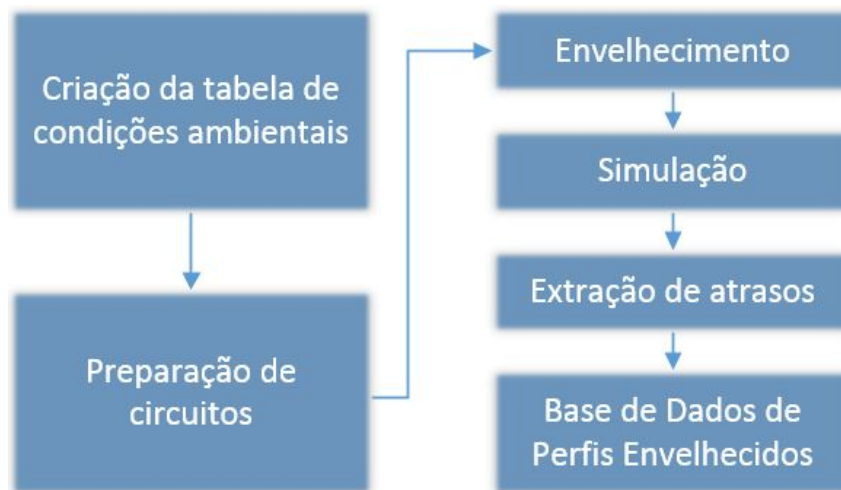


Figura 12 – Fluxo geral utilizado no envelhecimento de circuitos integrados.

É importante salientar que esta é apenas uma das formas de se popular a base de dados de perfis envelhecidos, conforme mencionado na seção 3.2, e pode ser utilizada em conjunto com as demais técnicas mencionadas na seção. Para este trabalho, esta tabela de condições ambientais servirá um duplo propósito: representar o BDM e o BDPE. No capítulo 5 é detalhado como esta tabela é utilizada na criação dos modelos de predição de MTTF.

A simulação *offline* do sistema a ser observado (ou de parte dele) exige algumas *entradas* a serem utilizadas no fluxo de envelhecimento:

1. Um conjunto de *netlists* a serem envelhecidas. Neste trabalho são utilizadas *netlists* cuja sintaxe é *Spectre*;
2. Os arquivos de modelos para os dispositivos do tipo nMOS e pMOS. Estes arquivos contém não somente os parâmetros de operação como também os de envelhecimento;
3. Uma tabela de condições ambientais a serem aplicadas ao sistema. No trabalho foram utilizados os modelos BSIM4 [58].

Como *saída* são obtidos:

1. Uma netlist com cada parâmetro de envelhecimento que será usado na simulação envelhecida;
2. Dados de simulação da netlist envelhecida e atrasos relativos à uma saída pré-determinada.

Durante este processo de envelhecimento, etapas intermediárias são necessárias, conforme mostrado na figura 13:

1. Cálculos de envelhecimento: etapa onde são obtidas as variações de parâmetros dos dispositivos empregados e como degradam com o tempo. Além disso, uma netlist “pré-processada” é criada. Isso significa que ela já foi modificada com as condições ambientais necessárias para a simulação envelhecida mas sem as informações de degradação dos dispositivos que a compõe. Estas mudanças são retiradas dos arquivos de modelo e informam como a tensão de limiar V_{TH} de um transistor nMOS, por exemplo, aumenta com o decorrer do envelhecimento.
2. Degradação de transistores: a ferramenta envelhece os dispositivos e anexa estas informações em um ou mais arquivos de modelos. Esses arquivos são atualizações baseadas nas informações originais e possuem novos valores para alguns parâmetros (*p.ex.* tensão de limiar). Estes dispositivos são atualizados de forma a refletir quais serão os valores destes parâmetros após um certo período de tempo. Exemplo: considere que, em seu estado original, um dispositivo nMOS possui um $V_{TH} = 1.1V$. Após simular dois anos de envelhecimento para sistema, seus transistores nMOS passam a precisar de $V_{TH} = 1.23V$. Este incremento de $0.13V$ é consequência da degradação do sistema ao longo desses 2 anos.

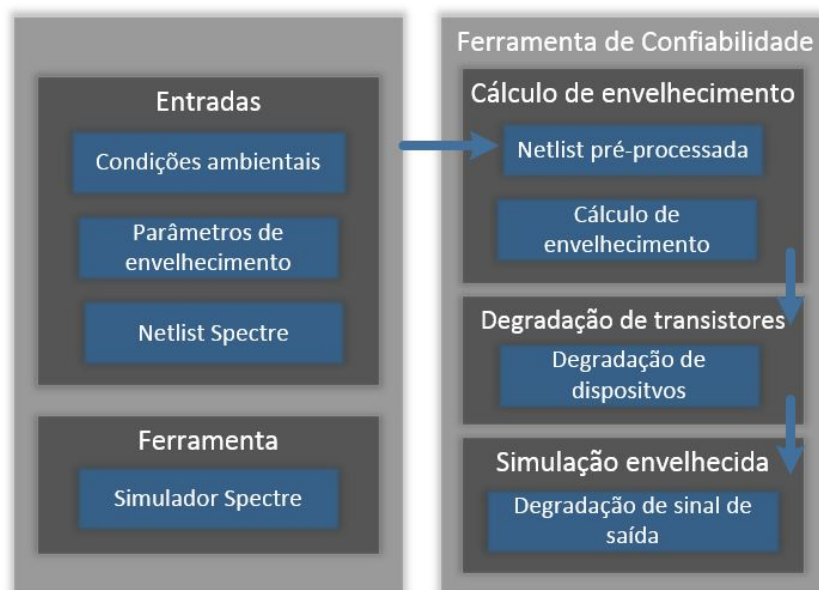


Figura 13 – Entradas e saídas da ferramenta de envelhecimento.

Estas etapas de cálculo são realizadas pela ferramenta de envelhecimento e são “transparentes”. Isso significa que não há, normalmente, uma intervenção por parte do projetista sobre estes cálculos, que seguem os modelos da própria ferramenta e os parâmetros de tecnologia fornecidos pela fabricante. É possível, entretanto, analisar quais mudanças e resultados são inseridos nos diversos arquivos utilizados pelos simuladores.

4.2 Implementação do fluxo e integração com as ferramentas de envelhecimento

As seções seguintes irão descrever as etapas necessárias para a criação e integração do fluxo com as ferramentas e a metodologia apresentada no capítulo 3. São dados mais detalhes da preparação das *netlists*, dos perfis de envelhecimento e da análise dos caminhos críticos.

4.2.1 Preparação de células lógicas e netlists

Para representar o comportamento de um sistema são, por vezes, utilizadas *netlists*, que são arquivos cuja função é descrever como um sistema ou diversos componentes eletrônicos se interconectam. A netlist pode conter fios, pinos, parâmetros elétricos, regimes de operação (*p. ex. temperatura, tempo de envelhecimento*), diretivas de simulação, referência a modelos de transistores, fontes de tensão e corrente, lógica booleana, entre outras informações.

Entretanto é preciso diferenciar as netlists geralmente disponíveis em um projeto de circuitos digitais:

1. descrição comportamental: utilizando-se de uma linguagem de descrição de hardware (HDL), especificam o comportamento de um sistema. A descrição é de *alto nível*, não contendo detalhes de quais portas lógicas são utilizadas para implementar a lógica desejada;
2. descrição estrutural: ainda utilizando HDL, descreve uma lógica através da interconexão entre portas lógicas;
3. descrição esquemática: detalha a interconexão entre os transistores além de possuir informações sobre fontes de alimentação, fios, elementos capacitivos, modelos de transistores, temperatura de operação e qualquer outra informação necessária para a simulação dos circuitos descritos.

Para o fluxo aqui apresentado é proposto a utilização de uma netlist esquemática pois, por possuir a descrição dos transistores, pode conter também os parâmetros de

degradação dos mesmos. Entretanto, caso o projetista não possua o esquemático, é possível obtê-lo através de uma descrição comportamental ou estrutural. Para isso, um processo de síntese ou tradução é realizado com o auxílio de softwares de *projeto de circuitos integrados* (*i.e.* Electronic Aided Design, EAD). Por síntese entende-se o processo de interpretação da descrição de alto nível de um circuito ou conjunto de circuitos e tradução para uma representação de portas lógicas. Define-se por tradução como sendo o processo de representar uma netlist estrutural em uma esquemática.

Caso o ponto de partida seja uma descrição comportamental, uma netlist estruturada é obtida por síntese e em seguida uma netlist esquemática. Em uma descrição comportamental não existe informação de quais portas lógicas são utilizadas, apenas as funcionalidades do sistema. Como as portas são desconhecidas, não é possível determinar quais transistores serão utilizados e, conseqüentemente, como os efeitos de degradação afetam os mesmos. Posteriormente, a descrição estrutural será traduzida para uma esquemática contendo todos os transistores necessários para implementar as funcionalidades projetadas.

Essa tradução necessita de uma biblioteca de portas lógicas que, por sua vez, possui informações detalhadas de como elas são implementadas para uma determinada tecnologia de um fabricante específico. Esta biblioteca é fornecida pelas *foundries* e sua descrição detalhada foge do escopo deste trabalho. O fluxo proposto assume que a implementação destas portas já está disponível e que é possível sintetizá-las, separadamente ou em conjunto, para que uma netlist seja gerada.

Para a utilização de um sistema no fluxo que está sendo proposto é mister identificar como ele está descrito (descrição comportamental, estrutural ou esquemática) e obter gradativamente, quando necessário, a netlist esquemática.

4.2.2 Preparação dos perfis de envelhecimento

O circuitos integrados atuais trabalham sobre condições de operação que oscilam constantemente. Podemos abordar estas variações de inúmeras maneiras. Entretanto, esse trabalho considera duas abordagens complementares que propõem interpretar as variações de duas maneiras:

1. Para um intervalo de tempo t , as variações são sumarizadas como as médias das tensões aplicadas e temperaturas medidas no dispositivo ao longo desse tempo. Isso significa que, se um sensor $S_{V,1}$ registrou uma média de $1.1V$ e um outro $S_{T,1}$ registrou uma média de $45^{\circ}C$, aquele intervalo será caracterizado por $V_{conj,1} = 1.1$ e $T_{conj,1} = 45$;
2. Um perfil de envelhecimento representa um conjunto de intervalos (ou faixas) sob os quais o sistema esteve submetido ao longo de sua operação, conforme mostrado na

figura 9. Ao se variar a quantidade de tempo sob qual o circuito permanece em cada elemento deste conjunto, obtemos uma tabela que possui um sumário das condições ambientais impostas ao sistema.

Assim sendo, existe uma tabela que associa variações de temperatura e tensão a um ciclo de vida de um sistema. Cada variação dos parâmetros de tensão de alimentação V_{DD} das portas lógicas e da temperatura T , é associada a um tempo de operação, conforme exemplificado na tabela 2. A definição dos valores $p_{T(w,x)}$ e $p_{V(w,y)}$ e da quantidade de conjuntos $T_{conj,X}$ e $V_{conj,Y}$, descritos na tabela 1, influencia diretamente na precisão das simulações a serem realizadas.

Para exemplificar, consideremos um sistema analisado durante 1 ano cujos perfis de operação são sumarizados pela tabela 5 abaixo:

Tabela 5 – Tabela de exemplo utilizada para o cálculo das variações de perfis.

30	50	1.0	1.1	MTTF
13	1	12	2	1.2
7	7	9	5	2.0

Para o primeiro registro, o sistema fica submetido a $T_{conj,1} = 30$ e $T_{conj,2} = 50$ durante 1 ano. Entretanto, $T_{conj,1} = 30$ influencia o sistema por mais tempo, pois possui um valor de $p_{T(w,x)}$ maior. Somando-se os valores de $p_{T(w,x)}$ para ambos os conjuntos de temperatura, o sistema é descrito como tendo sido submetido a uma temperatura média de $30^{\circ}C$ treze vezes ao longo de 1 ano e somente uma vez à temperatura média de $50^{\circ}C$, totalizando 14 intervalos de tempo Δt .

Desta forma, para um tempo de simulação t_{sim} e N intervalos, cada Δt corresponde a um período de tempo dado por:

$$\Delta t = \frac{t_{sim}}{N} \quad (4.1)$$

No exemplo acima $\Delta t \approx 0.07$ anos. Caso o sistema seja representado pelo BDPE da tabela 2, por exemplo, seu perfil possui então $N = 50$ intervalos distribuídos entre 5 conjuntos de médias de temperatura e tensão de alimentação, tornando o perfil mais granular. Como é necessário representar o mesmo sistema em diferentes intervalos de tempo, e cada intervalo pode ter suas próprias condições ambientais, o fluxo necessita de N representações esquemáticas.

Essa granularidade altera a quantidade de etapas necessárias para o fluxo, conforme representado na figura 14:

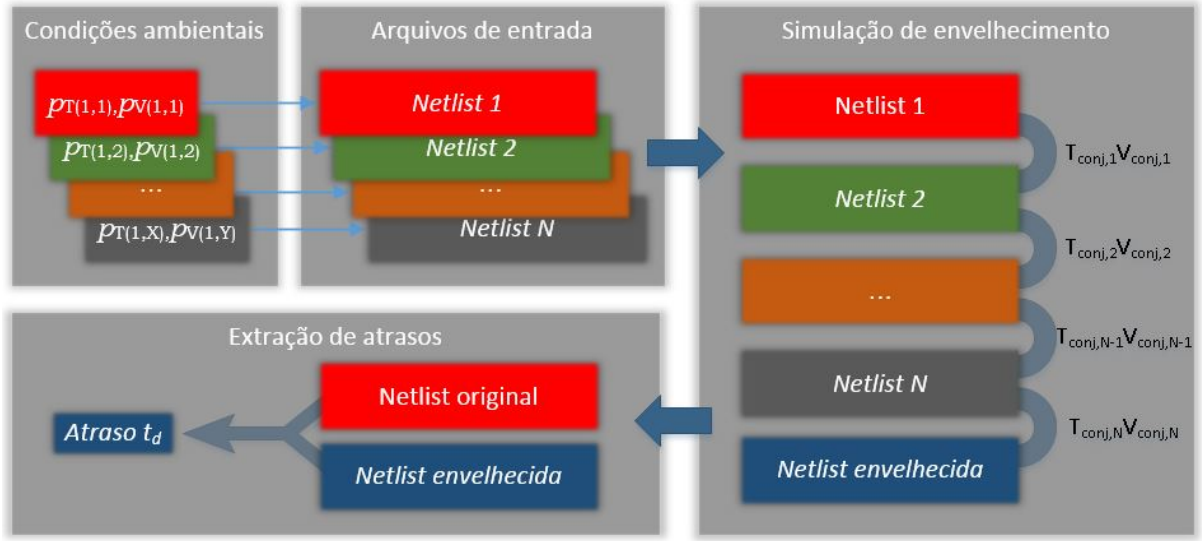


Figura 14 – Detalhamento do fluxo de criação, modificação e extração de resultados.

Cada uma das N condições ambientais, representado no bloco *condições ambientais* da figura 14, será utilizada na criação de N netlists, cada uma com seus respectivos $p_{T(w,x)}$ e $p_{V(w,y)}$. A degradação é realizada em N etapas ao longo de um tempo t_{sim} , onde cada netlist é degradada durante um intervalo Δt .

Para o exemplo da tabela 5, se o sistema estiver sendo observado durante 1 ano, a *netlist 1* estará submetida à $T_{conj,1} = 30$ e $V_{conj,1} = 1.0$ durante 0.07 anos na primeira etapa. A *netlist 2* é criada e alterada com as novas condições (que para este exemplo são as mesmas), que também prevalecerão durante 0.07 anos. Esse processo se repete até que todas as N netlists sejam criadas.

Dado o exposto anteriormente, o projetista precisa preencher o BDPE com dados que provem de simulações, testes de estresse ou extraídas de campo. Neste trabalho os BDPEs foram preenchidas com dados de simulação exclusivamente e os valores de $p_{T(w,x)}$ e $p_{V(w,y)}$ foram gerados aleatoriamente, por conveniência. Essa escolha pode obedecer uma distribuição normal, caso seja adequado. A metodologia aqui proposta é indiferente quanto à essa escolha.

4.3 Extração de caminhos críticos

Todas as descrições esquemáticas geradas na seção 4.2.2 são simuladas uma após a outra. Entretanto, representar cada transistor de um sistema pode ser impraticável. Isso se deve ao fato de que os processadores mais recentes possuem bilhões de transistores [64], tornando a etapa de degradação de todos eles inviável e computacionalmente custosa. Além disso, para uma simulação adequada, é necessária a ativação de todas as entradas do sistema, o que também pode ser impeditivo.

Para mitigar este problema, é possível se realizar a análise temporal estática do *pior caminho* de um sistema, sendo considerada uma técnica conservativa e bem estabelecida, apesar de não ser a mais moderna [65]. O *pior caminho* representa o trecho condutivo mais crítico ao sistema, onde qualquer violação de temporização torna a operação do sistema não confiável.

Dessa forma, realizar a degradação do caminho crítico é uma abordagem vantajosa e mais realística pois obtém uma quantidade reduzida de transistores e entradas a serem ativadas. Um problema desta abordagem é que não há garantia de que, após a degradação, este caminho continue sendo o mais crítico. Para mitigar este problema, mais de um caminho pode ser utilizado e posteriormente selecionado como pior aquele cuja atraso seja o maior.

A extração dos caminhos críticos é realizada pelas próprias ferramentas de EDA, que são capazes de reportar quantos caminhos forem desejados pelo projetista, gerar uma descrição esquemática e criar fontes de alimentação que ativem os piores caminhos.

4.4 Simulação e degradação

Uma vez criadas as netlists e o BDPE, a degradação é realizada de fato. Foram realizadas as etapas dos blocos de “condições ambientais” e “Arquivos de entrada” representados na figura 14. Para a simulação, um conjunto de rotinas irá, obedecendo o BDPE, usar a “Netlist 1” como argumento para o Relxpert© (software usado para a degradação).

Como saída são obtidas uma netlist degradada com seus parâmetros de operação atualizados. Isso significa que grandezas tais como tensão de limiar V_{TH} e corrente de fuga I_l são alteradas para refletir o desgaste sofrido pela primeira netlist. Em adição, parâmetros de degradação existentes nos modelos dos transistores nMOS e pMOS também podem ser atualizados para refletir o desgaste.

Esses novos parâmetros são utilizados em conjunto com a “Netlist 2”, que degrada sob efeito de suas próprias condições ambientais, determinadas conforme exposto na seção 4.2.2, assim como a “Netlist 1”. A diferença reside na reutilização dos parâmetros de degradação obtidos no primeiro passo, considerando o desgaste anterior.

Essas etapas se repetem até a criação de uma netlist envelhecida cujos parâmetros incluem a contribuição de todos os envelhecimentos anteriores. A figura 14 sintetiza essa etapa no bloco de “Simulação de envelhecimento”, onde cada netlist é exposta a uma temperatura $T_{conj,X}$ e a saída serve de entrada para o próximo passo. A contribuição de $V_{conj,Y}$ foi omitida apenas para simplificar o exemplo, mas ambas as contribuições são analisadas concomitantemente.

4.5 Extração de atrasos e MTTF

Ao fim da simulação, o atraso entre a netlist não-degradada e a degradada é calculado. Porém, ao definir o MTTF como uma métrica que considera o tempo restante de um sistema até sua falha, a metodologia aqui apresentada considera que é preciso obter primeiramente a variação do atraso de saída Δt_d .

Em seguida, a variação máxima do atraso de saída $\Delta t_{d,max}$, determinada durante a criação da especificação de projeto, e o tempo de operação t_{oper} são utilizados para a determinação do MTTF, conforme formulado na equação 4.2:

$$MTTF = \left(\frac{\Delta t_{d,max}}{\Delta t_d} \right) t_{oper} \quad (4.2)$$

A variação no atraso de saída é então definida como o acréscimo no tempo necessário para que um sinal se propague da entrada de um caminho até sua saída.

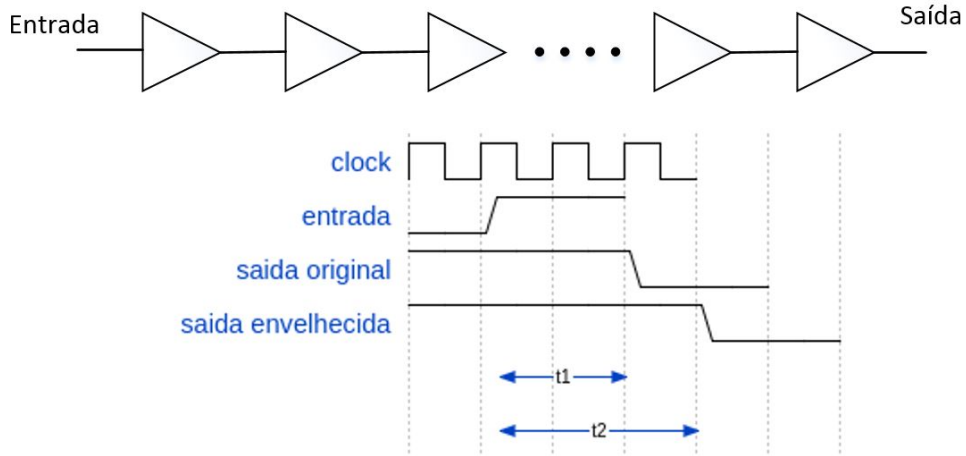


Figura 15 – Variação no atraso de saída.

Dado um tempo de propagação $t1$ como sendo o tempo necessário para que uma alteração na entrada seja percebida na saída de um circuito não-degradado e caso o tempo de propagação $t2$ medido após um tempo de operação t_{oper} seja maior do que $t1$, este circuito será considerado como envelhecido, conforme exemplo mostrado na figura 15.

Assim, a variação do atraso de saída é considerada como:

$$\Delta t_d = t2 - t1 \quad (4.3)$$

Idealmente este atraso de propagação não deveria aumentar. Entretanto, uma variação dele não é um fenômeno inesperado pelos projetistas. Por este motivo, uma salva-guarda (*i.e.* guard-banding) é planejada a nível de projeto.

Mesmo com esta medida, é possível que a degradação seja tal que extrapole essa salva-guarda. Dessa forma, a variação máxima do atraso $\Delta t_{d,max}$ não pode ultrapassar o *guard-banding*.

4.6 Tratamentos dos dados

Os dados obtidos durante a extração de atrasos, MTTF e seus perfis correspondentes são aglutinados e populam o BDPE de forma definitiva. Estes dados compõe a base a ser usada como referência para os métodos descritos na seção 3.5.

A formatação específica dessa informações a nível de software fica a critério do projetista. Além disso, o BDPE pode ser modificadp posteriormente para que novos perfis sejam adicionados.

Apesar de a metodologia não ser prejudicada pela existência de entradas duplicadas no banco de dados de perfis envelhecidos, é recomendado que as mesmas sejam únicas. Desta forma, o tamanho do BDPE é reduzido e consequentemente a quantidade de memória necessária para armazená-lo.

5 Resultados

Do or do not, there is no try.

-Master Yoda

Este capítulo apresenta os resultados das etapas do fluxo proposto, bem como a análise dos mesmos. Alguns resultados preliminares foram utilizados para verificar se o trabalho estava coerente com a base teórica anteriormente exposta. Outros auxiliaram na análise do fluxo de envelhecimento e na acuidade dos métodos de estimativa propostos.

Algumas considerações precisam ser feitas antes da análise dos dados. O ambiente de simulação criado para este trabalho utiliza ferramentas de EAD previamente disponíveis para degradação, sendo sua integração realizada através de um conjunto de *scripts* ou *wrappers*, que são responsáveis por automatizar:

1. A criação das diversas netlists utilizadas no fluxo;
2. Formatação dos arquivos utilizados como banco de dados bem como preenchimento dos mesmos;
3. A passagem de parâmetros de entrada às ferramentas de EAD;
4. A extração dos resultados de simulação;
5. A atualização do banco de dados com os resultados de estimativas de MTTF.

5.1 Caracterização de células

Para investigar a relação entre as condições ambientais e a degradação de sistemas, é preciso averiguar primeiramente se é possível reproduzir resultados que estejam coerentes com a literatura. Isso significa que é necessário garantir que, ao se degradar um sistema, as alterações nas condições ambientais (*p.ex.* temperatura) sejam refletidas no sistema conforme esperado. É bem estabelecido que uma alteração na tensão de alimentação V_{DD} e na temperatura deve influenciar não só a operação dos dispositivos mas também na contribuição dos efeitos de degradação expostos na seção 2.1.3. Esta premissa precisa então ser satisfeita.

Com este propósito foram caracterizadas as seguintes células básicas: INV (inversor), AND2, OR2, NOR2 e NAND2. Todas as células foram submetidas a BDPEs estáticos e populados exclusivamente para estas simulações. Por estático indica-se que a célula foi submetida a uma mesma temperatura e tensão ao longo de toda a simulação ao invés de

diferentes $T_{conj,X}$ e $V_{conj,Y}$. Cada BDPE possui 200 perfis e o tempo de degradação para todas as células é de 1 ano.

Foram criados 5 BDPE's diferentes, um para cada célula. Entretanto, todos os perfis foram gerados usando a mesma premissa: os valores de temperatura seguem uma distribuição normal de média $\mu = 80$ e desvio-padrão $\sigma = 10$. Os valores de tensão também seguem uma distribuição normal, porém de média $\mu = 1.1$ e desvio-padrão $\sigma = 0.2$. Foi escolhida uma média de 80°C com o intuito de garantir uma contribuição significativa da temperatura na degradação. Já a média da tensão de alimentação de 1.1V está de acordo com os parâmetros da tecnologia dos transistores utilizados na simulação.

Após cada célula ter sido degradada, o atraso de saída Δt_d para cada perfil simulado foi adicionado aos seus respectivos BDPEs e foi criado um modelo de regressão linear múltipla para cada. Os modelos foram utilizados para plotar a dependência do atraso de saída pela tensão de operação e pela temperatura. Os resultados estão representados nas figuras 16 a 19.

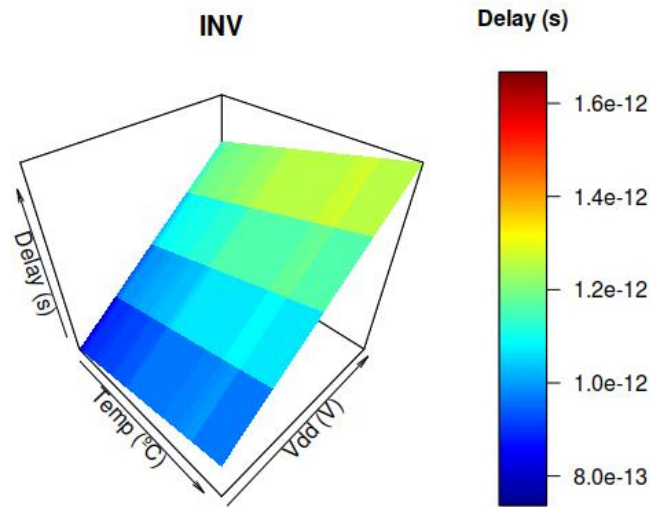


Figura 16 – Caracterização de uma célula inversora degradada durante 1 ano.

A figura 16 mostra que, à medida que a temperatura sobre o inversor (dada em $^\circ\text{C}$) aumenta, o atraso Δt_d (dado em segundos) também incrementa. Entretanto, a influência da temperatura é menor se comparada à da tensão de alimentação, o que era esperado.

Para esta análise, os coeficientes da regressão linear múltipla (e consequentemente a influência de cada grandeza) são dados pelos valores da tabela 6:

Tabela 6 – Coeficientes do modelo de RLM para o inversor da figura 16.

	Interceptação	$T_{conj,1}$	$V_{conj,1}$
Coeficiente	$-1.3049e^{-12}$	$5.0378e^{-15}$	$1.9161e^{-12}$

Do modelo fica claro que a influência da tensão sobre o envelhecimento é 3 ordens de grandeza maior que a da temperatura. Uma relação semelhante foi encontrada para as células AND2, NAND2, OR2 e NOR2.

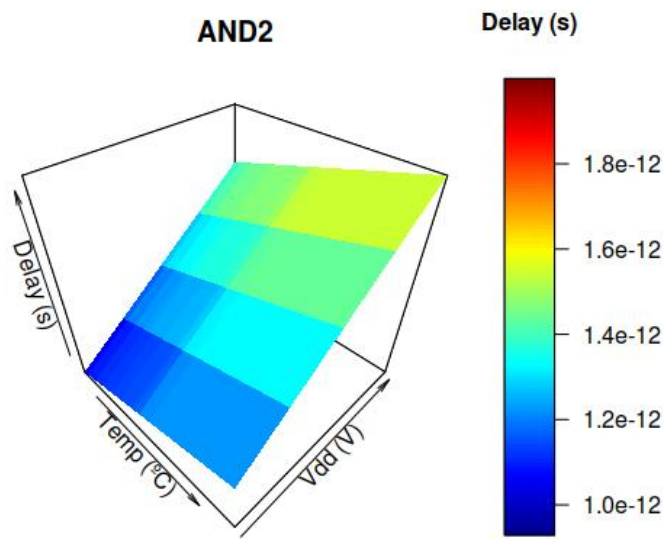


Figura 17 – Caracterização de uma célula AND2 degradada durante 1 ano.

Tabela 7 – Coeficientes do modelo de RLM para a figura 17.

	Interceptação	$T_{conj,1}$	$V_{conj,1}$
Coeficiente	$-1.4112e^{-12}$	$5.8058e^{-15}$	$2.1905e^{-12}$

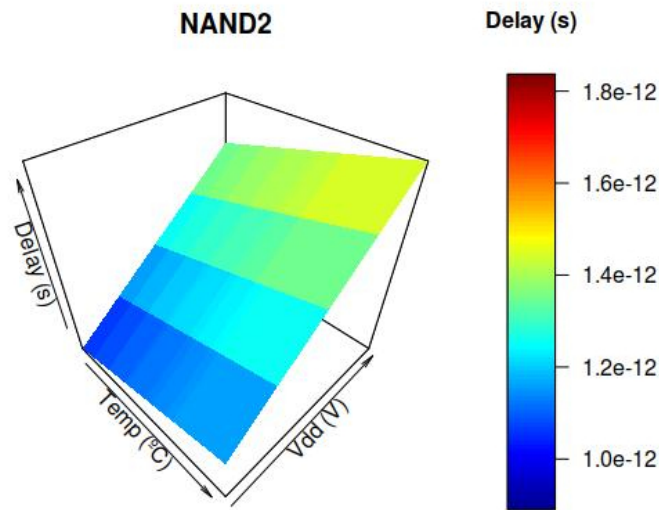


Figura 18 – Caracterização de uma célula NAND2 degradada durante 1 ano.

Tabela 8 – Coeficientes do modelo de RLM para a figura 18.

	Interceptação	$T_{conj,1}$	$V_{conj,1}$
Coeficiente	$-1.2030e^{-12}$	$5.5807e^{-15}$	$1.9283e^{-12}$

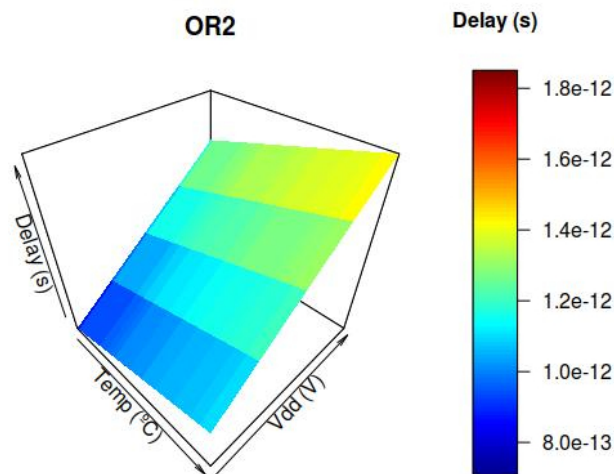


Figura 19 – Caracterização de uma célula OR2 degradada durante 1 ano.

Tabela 9 – Coeficientes do modelo de RLM para a figura 19.

	Interceptação	$T_{conj,1}$	$V_{conj,1}$
Coeficiente	$-1.5102e^{-12}$	$5.9055e^{-15}$	$2.1376e^{-12}$

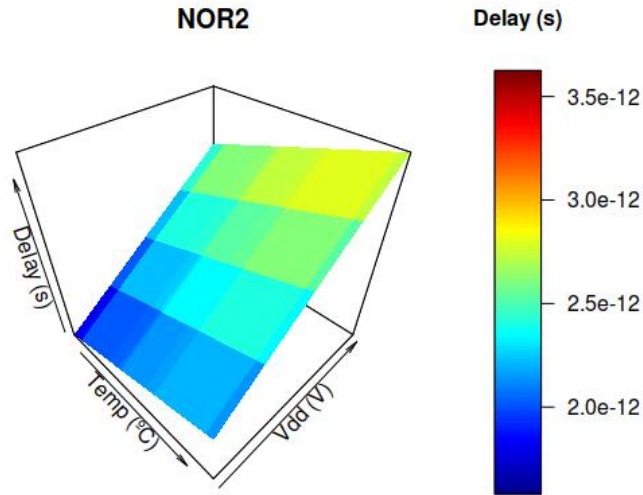


Figura 20 – Caracterização de uma célula NOR2 degradada durante 1 ano.

Tabela 10 – Coeficientes do modelo de RLM para a figura 19.

	Interceptação	$T_{conj,1}$	$V_{conj,1}$
Coeficiente	$-2.9256e^{-12}$	$1.3332e^{-14}$	$4.0494e^{-12}$

Era esperado que a relação entre tensão de alimentação e atraso fosse positiva. O mesmo pode ser dito entre a temperatura e o atraso. Estes resultados iniciais nos dão segurança de que o fluxo está pronto e que ele pode ser utilizado para sistemas maiores. Além disso, eles confirmam aquilo que já era esperado: quando da utilização do método de correlação, descrito na seção 3.5.3, devemos procurar pelo resultado de maior correlação positiva, já que o atraso cresce com o aumento da temperatura e tensão.

5.2 Degradação de circuitos de teste

Uma vez realizados testes com células básicas, mais circuitos foram utilizados para continuar investigando a viabilidade do fluxo e da metodologia. Com esse intuito, foram utilizados quatro circuitos ISCAS-85: c499, c880, c1355, c5315. Esses circuitos ISCAS são implementações de blocos combinacionais pré-determinados cujo objetivo é propor circuitos padrões para testes de diferentes propósitos [5]. A função implementada e a respectiva quantidade de portas lógicas para cada um deles está descrita na tabela 11.

Tabela 11 – Características da cadeia de inversores e dos circuitos ISCAS-85 utilizados [5].

Circuito	Função	Qtde. de portas
inv100	Cadeia de Inversores	100
c499	Single-Error-Correcting de 32-bit	474
c880	ULA de 8 bits	393
c1355	Single-Error-Correcting de 32-bit	738
c5315	ULA de 9 bits	1781

Esses circuitos podem ser utilizados para os próximos passos em sua integridade ou apenas uma porção dos mesmos. Por integridade entende-se que os circuitos, inicialmente descritos através de uma linguagem de descrição de hardware (Hardware Description Language, HDL), serão sintetizados e uma netlist esquemática no formato Spectre criada.

Por porção entende-se que somente uma parte do circuito será utilizada para criação da netlist Spectre. Esta parte deve ser capaz de, após as simulações de degradação, representar o pior caso para o restante do circuito. Esta porção pode ser representada por um ou mais caminhos críticos, conforme exposto na seção 4.3. Foram extraídos os caminhos críticos dos circuitos ISCAS da tabela 11 e netlists esquemáticas menores foram criadas a partir destes caminhos.

Para extração dessa informação, foi utilizada uma funcionalidade nativa da ferramenta de síntese capaz de relatar quais são os caminhos críticos de um circuito. Além disso, eles devem ser ativados de tal forma que uma transição na saída seja detectável. Isso significa que as portas lógicas serão ativadas ou desativadas em uma combinação específica que reflita não só o maior atraso de propagação possível mas também uma mudança lógica na saída do caminho crítico.

Após a extração destes caminhos, os circuitos foram reduzidos aos tamanhos descritos na tabela 12.

Tabela 12 – Tamanho dos circuitos de teste após a extração dos caminhos críticos.

Circuito	Qtde. de portas
inv100	100
w_c499	10
w_c880	8
w_c1355	21
w_c5315	25

O sufixo “w_” foi adicionado aos nomes dos ISCAS unicamente como forma de

diferenciar os circuitos com caminhos críticos (*p. ex.* `w_c499`) dos seus respectivos circuitos completos. Após a extração dos caminhos, um circuito de teste (*testbench*) foi criado para cada um dos circuitos da tabela 12 cujo objetivo é ativar o caminho crítico, conforme mencionado anteriormente.

Uma vez criadas, as netlists são padronizadas com valores de temperatura $T = 0$ e alimentação $V_{DD} = 0.0$. Eles são utilizados como referência pelos *scripts*, substituindo estes valores pelos respectivos $T_{conj,X}$ e $V_{conj,Y}$ durante a simulação.

Como exemplo, considere que a tabela 5 possui perfis de envelhecimento para o circuito `w_c499`. Para o primeiro perfil, serão criadas $N = 14$ netlists, das quais 13 serão submetidas a uma temperatura $T = 30^\circ C$ e uma a $T = 50^\circ C$. Já para V_{DD} , 12 estarão submetidas a uma tensão $V_{DD} = 1.0V$ e duas a $V_{DD} = 1.1V$. Dessa forma, as 12 primeiras netlists terão os valores $T = 0$ e $V_{DD} = 0.0$ substituídos por $T = 30^\circ C$ e alimentação $V_{DD} = 1.0V$. A 13ª netlist terá $T = 30^\circ C$ e $V_{DD} = 1.1V$. A restante terá $T = 50^\circ C$ e $V_{DD} = 1.1V$.

Cada circuito foi submetido a um BDPE de aproximadamente 250 entradas cujos perfis são submetidos aos conjuntos de condições ambientais $T_{conj,X} = [10, 30, 50, 70, 90]$ e $V_{conj,Y} = [0.9, 1.0, 1.1, 1.2, 1.3]$. Os valores de $p_{T(w,x)}$ e $p_{V(w,y)}$ foram gerados pseudo-aleatoriamente, sendo $\sum p_{T(w,x)} = 100$ e $\sum p_{V(w,y)} = 100$ (ou seja, $N = 100$). Assim sendo, cada entrada do BDPE corresponde a 100 degradações, totalizando 2500 simulações de envelhecimento para toda o banco de dados de perfis e que são posteriormente adicionados ao BDPE, conforme descrito no capítulo 4.

Uma segunda bateria de testes é realizada, porém com condições ambientais diferentes. Os mesmos circuitos são degradados mas desta vez submetidos a um BDPE estático. Isso significa que o perfil ao qual são submetidos permanece 100% do tempo em um único $T_{conj,X}$ e um único $V_{conj,Y}$. Esse comportamento é representado através da criação de uma tabela de perfis, representada na tabela 13.

Tabela 13 – BDPE estático utilizado no envelhecimento dos circuitos de teste.

40	70	100	1.0	1.1	1.2
3	0	0	3	0	0
0	3	0	3	0	0
0	0	3	3	0	0
3	0	0	0	3	0
0	3	0	0	3	0
0	0	3	0	3	0
3	0	0	0	0	3
0	3	0	0	0	3
0	0	3	0	0	3

Este tipo de teste auxilia a investigar melhor quais são as vantagens da metodologia proposta ao comparar os resultados anteriores com os de uma abordagem de perfis estáticos. É preciso esclarecer primeiramente que um perfil definido por $T_{conj,X} = [40, 70, 100]$ e $p_{T(1,x)} = [3, 0, 0]$, utilizando a tabela 13 como exemplo, não necessariamente indica que um sistema foi degradado sob influência de uma temperatura e tensão constantes. Para um sistema real não é possível assumir essa premissa. Entretanto, o perfil de temperatura e tensão podem, por escolha do projetista, ser representados por médias (*p.ex* temperatura média ou tensão de alimentação média).

Dessa forma, é possível avaliar se a representação estática da operação do sistema oferece uma precisão semelhante à representação dinâmica. Se for possível representar o perfil de um sistema através de uma estatística simples (como as médias) e obter um erro de predição semelhante ou melhor do que a dinâmica, então deve-se questionar a utilização de um BDPE dinâmico, já que não existiria, em tese, uma redução do erro. Esta análise também será realizada nas seções subsequentes.

5.3 Métricas e estatísticas

Antes de explorar os dados de simulação, é preciso explicar quais métricas foram utilizadas para este trabalho. Uma métrica, no contexto aqui apresentado, deve ser capaz de indicar a precisão ou assertividade dos métodos de estimativa apresentados na seção 3.5, dando uma perspectiva diferenciada dos resultados. Qual métrica se mostra mais adequada é um debate que não faz parte deste trabalho, visto que não há consenso na literatura específica da área [66][67][68]. Ao invés disso, mais de uma métrica será utilizada, cabendo ao projetista julgar a mais adequada ao entendimento de seus dados.

A análise gráfica dos resultados utilizará apenas algumas destas métricas. Entretanto serão sumarizadas, ao longo da análise, as seguintes métricas: Erro Relativo (ER), Erro Máximo (EMAX), Erro Quadrático Médio (EQM) e Erro Normalizado da Raiz do Valor Quadrático Médio (ENRVQM). Estas métricas são expressas por:

$$ER = 100 \left(1 - \frac{\hat{y}_i}{y_i} \right) \quad (5.1)$$

$$EMAX = \max \left(100 \left(1 - \frac{\hat{y}_i}{y_i} \right) \right) \quad (5.2)$$

$$EQM = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n} \quad (5.3)$$

$$ENRVQM = \frac{\sqrt{EQM}}{\bar{y}} \quad (5.4)$$

onde n representa a quantidade de observações, \bar{y} é a média dos dados observados, \hat{y}_i é o i -ésimo dado observado e y_i o i -ésimo dado estimado. Por ser a mais simples e intuitiva, o ER é a primeira métrica a ser explorada adiante.

Apesar de dispormos destas essas métricas, é interesse evitar que um modelo de estimativa não sofra com um efeito de sobreajuste, ou seja, ele tente se acomodar demasiadamente aos dados observados, reduzindo a performance do modelo e consequentemente afetando as métricas mencionadas acima.

Para evitar o sobreajuste, foi utilizado uma técnica de validação cruzada. Esta abordagem propõe dividir os dados pré-existentes em dois grupos: dados de treino e dados de teste. Os dados de treinamento são utilizados para obtenção do modelo preditivo. Já os dados de teste são utilizados para medir a precisão do modelo. A técnica de validação cruzada escolhida para a análise foi *leave-one-out cross-validation* (LOOCV), onde apenas uma das amostras é utilizada como teste e as restantes são utilizadas como dados de treino [4].

Utilizemos a tabela 14, que é um excerto de um dos BDPE utilizados neste trabalho, como exemplo.

Tabela 14 – Excerto de um BDPE com dados de simulação.

10	30	50	70	90	0.9	1.0	1.1	1.2	1.3	MTTF
37	56	1	6	0	37	56	1	6	0	1.46
70	0	7	0	23	70	0	7	0	23	1.13
47	0	0	42	11	47	0	0	42	11	1.15

Na técnica de LOOCV a primeira amostra (linha 2 da tabela 14) é retirada, por exemplo, restando somente 2 entradas para criação do modelo preditivo, seja ele PLS-R, DE ou COR. Após criado, a amostra retirada é usada como entrada no modelo e o resultado estimado \hat{y}_1 é comparado ao esperado y_1 . Nesse exemplo, considerando um tempo de operação $t_{oper} = 1$ ano, é conhecido que, para um perfil $p_{T(1,x)} = [37, 56, 1, 6, 0]$ e $p_{V(1,y)} = [37, 56, 1, 6, 0]$, o resultado esperado é $y_1 = 1.46$ (MTTF de 1.46 anos). A partir daí, as métricas de erro anteriormente mencionadas podem se utilizadas para medir a precisão dos métodos.

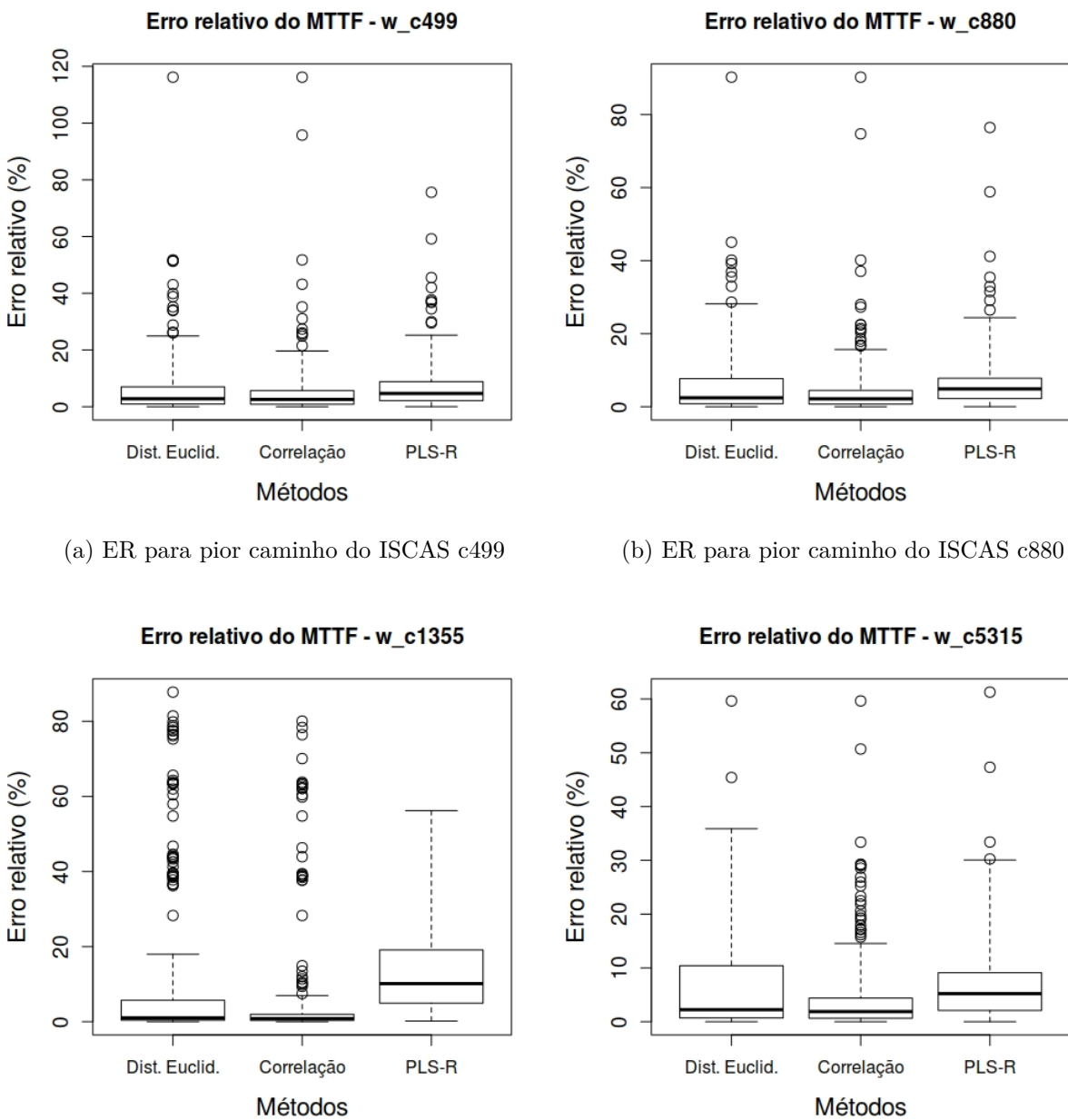
Para se obter métricas mais significativas ainda, a LOOCV foi utilizada N vezes, onde N é quantidade de amostras do BDPE. Isso significa que, após a primeira amostra ser utilizada como teste e os erros serem calculados, a LOOCV é repetida mas desta vez usando a segunda amostra como teste e assim sucessivamente. Todos os resultados são utilizados para compor as métricas mencionadas. Dessa forma não se está testando um modelo específico obtido através de um método de distância euclideana, por exemplo, mas está sendo testado “exaustivamente” a DE como método de estimativa.

É importante salientar que durante a operação real do sistema, ao contrário do ambiente de simulação, o modelo já será conhecido no momento em que a predição for necessária, pois já foi previamente treinado. O perfil que é utilizado na análise *online* como

entrada para o modelo é proveniente do BDM (e não do BDPE).

5.4 Resultados das simulações dos circuitos de teste

A primeira métrica analisada foi o erro relativo, como forma de compreender a exatidão dos métodos. Levando em consideração que o “verdadeiro” valor do MTTF é conhecido, o erro relativo em valores absolutos para o BDPE descrito na seção 5.2 é representado pelos diagramas de caixa das figuras 21 e 22.



(a) ER para pior caminho do ISCAS c499

(b) ER para pior caminho do ISCAS c880

(c) ER para pior caminho do ISCAS c1355

(d) ER para pior caminho do ISCAS c5315

Figura 21 – Erro relativo do MTTF para os circuitos ISCAS-85.

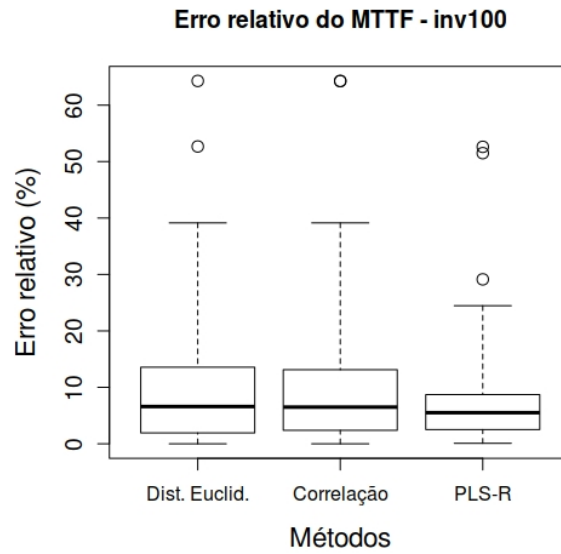


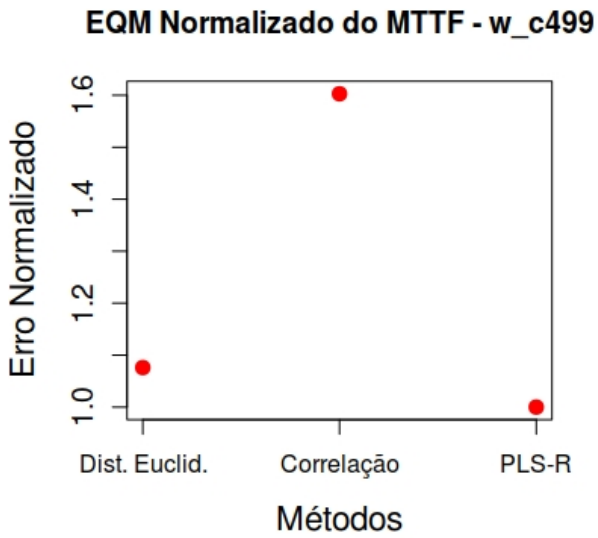
Figura 22 – Erro relativo do MTTF para a cadeia de inversores.

Extraindo o EMAX (como uma porcentagem) para cada circuito e método obtém-se:

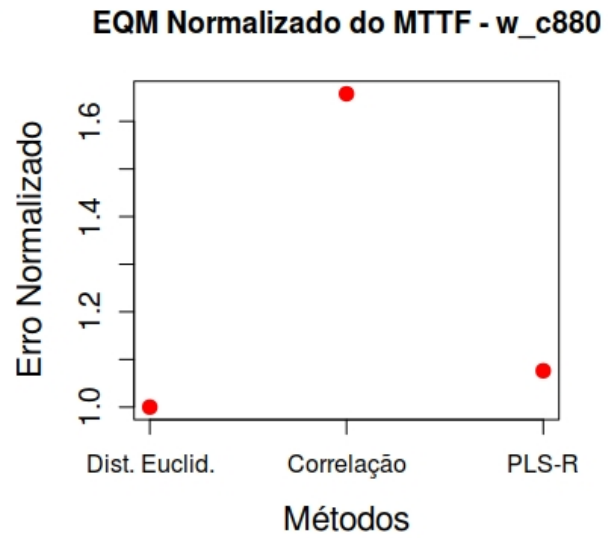
Tabela 15 – Erro máximo para os circuitos de teste.

Circuito	Dist. Eucli.	Cor.	PLS-R
inv100	64.29%	64.29%	52.61%
w_c499	116.21%	116.21%	75.61%
w_c880	90.24%	90.24%	76.45%
w_c1355	80.05%	87.78%	56.21%
w_c5315	59.64%	59.64%	61.28%

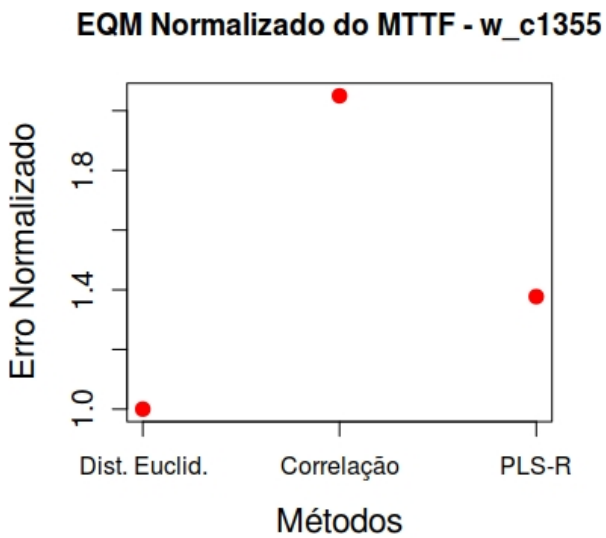
Para obter uma comparação mais intuitiva, o Erro Quadrático Médio (EQM) foi calculado para cada um dos circuitos e por fim normalizado, servindo como um rápido indicativo de qual método é mais preciso e o quão mais preciso é em comparação aos demais.



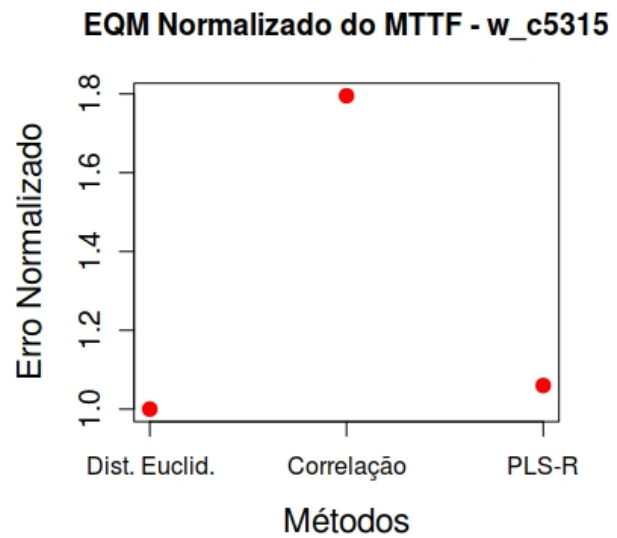
(a) EQMN para pior caminho do ISCAS c499



(b) EQMN para pior caminho do ISCAS c880



(c) EQMN para pior caminho do ISCAS c1355



(d) EQMN para pior caminho do ISCAS c5315

Figura 23 – Erro Quadrático Médio Normalizado do MTTF para os circuitos ISCAS-85.

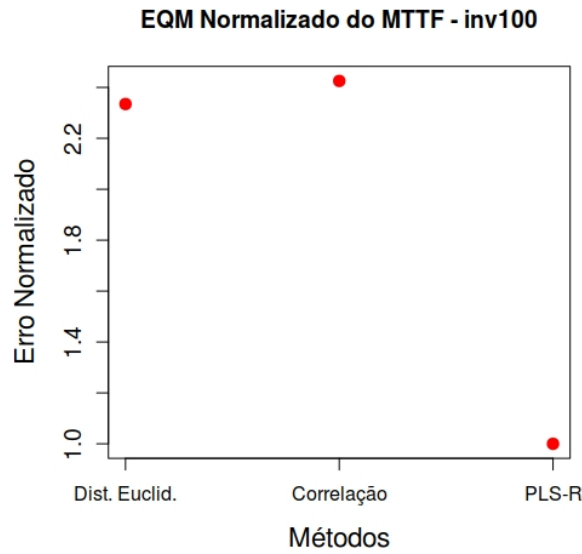


Figura 24 – Erro Quadrático Médio Normalizado do MTTF para a cadeia de inversores.

Entretanto, ao se analisar circuitos de diferentes dimensões, e conseqüentemente conjuntos de dados de diferença expressiva, algumas das métricas obtidas podem estar em escalas diferentes umas das outras, dificultando a compreensão da precisão dos métodos. Nestes casos, o Erro Normalizado da Raiz do Valor Quadrático Médio facilita a interpretação pois cada conjunto de dados é normalizado pela sua própria média, resultando em um erro percentual. Esse é o caso para a simulação da cadeia de inversores. Sendo substancialmente maior que os demais (de 4 a 10 vezes), seu atraso também é maior e a grandeza, nesse caso, maior.

O ENRVQM foi extraído para cada circuito e método e estão descritos na tabela 16.

Tabela 16 – Erro Normalizado da Raiz do Valor Quadrático Médio para cada método e circuito em %.

Circuito	Dist. Eucli.	Cor.	PLS-R
inv100	13.0%	13.1%	9.1%
w_c499	8.1%	10.1%	8.3%
w_c880	7.3%	9.6%	7.9%
w_c1355	13.5%	18.1%	14.7%
w_c5315	7.9%	10.6%	8.1%

Em seguida foram realizadas simulações de degradação para o BDPE estático mencionada na seção 5.2 e representado na tabela 13. O Erro Normalizado da Raiz do Valor Quadrático Médio para os perfis estáticos se encontram na tabela 17.

Tabela 17 – ENRVQM para cada método e circuito submetidos à perfis estáticos em %.

Circuito	Dist. Eucli.	Cor.	PLS-R
inv100	18.08%	18.08%	2.07%
w_c499	22.12%	22.12%	2.55%
w_c880	14.79%	14.79%	3.00%
w_c1355	29.76%	29.76%	4.85%
w_c5315	15.85%	15.85%	3.44%

5.5 Discussão dos resultados

Ao se analisar o Erro Relativo (ER) através do diagrama de caixas fica subentendido, inicialmente, que a COR e a DE são métodos mais apropriados. A mediana de ambos os métodos é muito próxima de 0. Isso é um indicativo de que o ER é muito pequeno em muitos casos. Além disso, a COR parece ser o melhor dos métodos, visto que 75% dos seus erros relativos estão “comprimidos” na região de $ER = 0$.

Entretanto, a quantidade de *outliers* aparenta ser alta. Se a ocorrência de outliers for considerável, isso se refletirá negativamente na precisão dos métodos. Observando a tabela 15, o EMAX dos métodos indica que, quando existe um erro na previsão, ele pode ser de aproximadamente 100%. Isso significa que o dado estimado \hat{y}_i é até duas vezes maior (ou duas vezes menor) que o y_i esperado.

Logo, é preciso averiguar se esses grandes erros de estimativa são usuais para os métodos e dados observados. Se forem, o erro quadrático médio é capaz de enfatizar estes erros. Porém, ao se trabalhar com métricas quadráticas, o resultado não mantém um significado físico plausível. Por este motivo, o resultado é normalizado (EQMN) utilizando o menor erro como referência e apresentado como mostrado na figura 23 para facilitar a compreensão.

O EQMN deixa explícito que a Distância Euclideana não é um método tão impreciso quanto indicava o diagrama de caixas do ER. Em adição, a PLS-R mostra-se como um método de precisão comparável à DE. Já a correlação está se mostrando como sendo o pior dos métodos entre os três, sendo até duas vezes mais impreciso que a DE.

Como mencionado na seção 5.3, não há consenso na literatura sobre qual métrica é a mais apropriada para análise de erros de estimativas. Por isso, além do Erro Quadrático Médio, o Erro Normalizado da Raiz do Valor Quadrático Médio também é utilizado, podendo ajudar a compreender um pouco melhor os resultados obtidos. Assim como a figura 23, a tabela 16 também indica que a DE e a PLS-R possuem uma precisão semelhante. Entretanto, o ENRVQM médio para DE é de 10% e de 9.6% para a Regressão de Mínimos Quadrados Parciais. Já a Correlação apresenta um ENRVQM médio de 12.3% e consolida-se como o pior método para estimativa.

Esses resultados mostram que uma estratégia proativa para extensão do tempo de vida de um sistema que utilize a Distância Euclidiana ou PLS-R como método de estimativa pode ter uma precisão de aproximadamente 90%. Esta performance pode ser melhorada ainda mais à medida que o BDPE for atualizado com dados de campo. A metodologia, ao propor uma modularização do sistema (conforme descrito na seção 3.2) permite que os próprios modelos sejam atualizados ou substituídos, potencialmente melhorando ainda mais a precisão.

Ao repetirmos esta análise para um BDPE estático, a tabela 17 evidencia que o Erro Normalizado da Raiz do Valor Quadrático Médio para a DE e COR piora bastante. Não é surpreendente um banco que possui poucas entradas, e um espaço amostral tão baixo, afetar estes dois métodos. Ao calcular-se a DE, por exemplo, a distância obtida será igual em todos os casos. Assim, qualquer um dos perfis pode ser enquadrado na condição de “mais próximo”, o que não é satisfatório.

Adicionalmente, o ENRVQM para a PLS-R mostra-se muito melhor. Observando-se as três primeiras linhas da tabela 13, é possível perceber que a tensão de alimentação mantém-se a mesma, enquanto a temperatura pode apresentar apenas os valores de 40, 70 e 100. O mesmo pode ser observado para o restante dos dados. Separando estas informações em grupos que possuem a mesma tensão de alimentação, a dependência do atraso de saída pela temperatura para diferentes V_{DD} é dada pela figura 25.

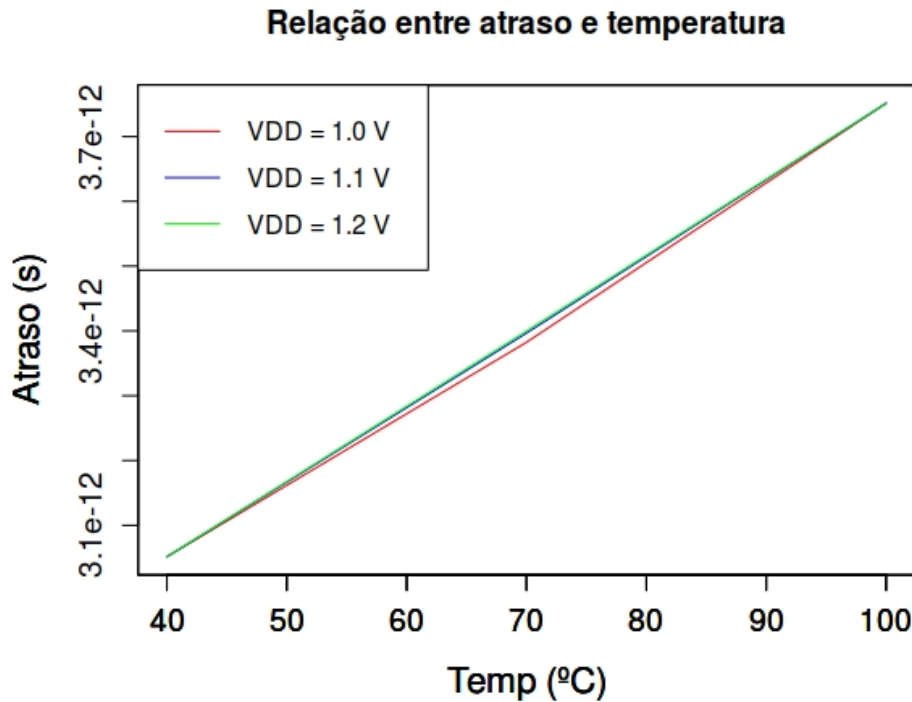


Figura 25 – Relação entre o atraso e a temperatura para diferentes V_{DD} .

A decomposição dos dados exposta na figura 25 possui uma serventia didática apenas, pois facilita visualizar a relação linear existente entre temperatura e atraso (e consequentemente MTTF). Similarmente, o mesmo pode ser feito entre a tensão de alimentação e o atraso. Se há uma relação linear tão evidente entre estas variáveis explicativas (Temperatura e Tensão) e a variável observada (Atraso ou MTTF), o modelo de PLS-R extraído será consequentemente bem mais preciso.

Este é o motivo pelo qual o ENRVQM para um BDPE estático é mais preciso do que para um BDPE dinâmico. É necessário ressaltar que o um banco de dados de perfis estático considera uma estatística (possivelmente média) para as condições ambientais, sendo bem menos fiel ao verdadeiro perfil de operação de um sistema.

5.6 Críticas ao método

O método apresentado propõe a utilização de diferentes unidades para realizar a tarefa de predição, conforme mostrado na figura 7. Apesar de essa abordagem se mostrar como uma vantagem, permitindo que suas diversas unidades sejam customizadas, melhoradas ou mesmo completamente substituídas, alguns de seus aspectos não estão bem definidos.

Não está definido, por exemplo, se existe uma dimensão recomendada para o BDPE dinâmico. Na análise realizada foram utilizados aproximadamente 250 perfis. Não há entretanto uma motivação específica. Usualmente 30 ou mais amostras são mais do que suficientes para que a distribuição amostral da média de uma população seja considerada uma distribuição *gaussiana* [69]. Entretanto, uma quantidade inferior não garante que a distribuição da estatística será “não-normal”.

No contexto desse trabalho, essa estatística pode ser o acréscimo médio no atraso t_d ou o MTTF médio. Independente de qual estatística seja escolhida pode-se realizar a análise de um BDPE dinâmico reduzido, com uma quantidade de amostras inferiores e averiguar como os métodos se comportam. Após repetir esta análise com apenas 28 perfis, evitando a garantia de normalidade da média, o ENRVQM obtido para os circuitos e métodos é apresentado na tabela 18.

Tabela 18 – ENRVQM para cada método e circuito submetidos à 28 perfis dinâmicos.

Circuito	Dist. Eucl.	Cor.	PLS-R
inv100	14.89%	15.89%	7.61%
w_c499	21.99%	26.13%	9.21%
w_c880	21.25%	22.55%	8.87%
w_c1355	14.23%	16.49%	14.86%
w_c5315	17.99%	19.41%	8.95%

O ENRVQM médio para a DE é de 18.07% e para a PLS-R é de 9.9%. Para a COR é obtido um ENRVQM médio de 20.09%. Apesar de a PLS-R se manter como o melhor método, a diferença para a DE é considerável desta vez. Então a crítica ao método se mantém e levanta alguns questionamentos:

1. Qual é a dimensão do BDPE necessária para que se tenha uma boa predição?
2. Aumentar o BDPE ainda mais garantirá que a DE se torne um método melhor que a PLS-R eventualmente?
3. Qual é a relação existente entre o aumento do BDPE e a redução dos erros de predição para cada método?

6 Conclusões

Estimar o tempo de vida de sistemas integrados é imperativo diante do atual avanço tecnológico e do consequente estresse submetido a eles. Efeitos físicos antes ignorados, agora passam a contribuir de forma preponderante para a degradação de circuitos integrados.

Nesse contexto surge a necessidade de observar, investigar e adequar a operação de um sistema afim de mitigar ou reduzir o desgaste decorrente de um perfil de operação. Inúmeras técnicas foram desenvolvidas e estão disponíveis na literatura que permitem coletar informações de diferentes origens e interpretá-las.

A metodologia aqui proposta permite integrar a coleta destes dados, ao mesmo tempo que representa estas informações na forma de perfis de operações. Não apenas isso, mas a metodologia está pronta para simular, degradar e estimar o tempo de vida restante de sistemas.

Apresentando uma abordagem que independe da origem dos dados e das técnicas já existentes, este trabalho permite que o projetista adeque o fluxo às suas necessidades; seja otimizando os métodos preditivos aqui propostos, seja utilizando outros inteiramente diferentes. Além disso, o trabalho independe da tecnologia utilizada no sensoramento dos parâmetros de operação, do nó tecnológico ou até mesmo da variação sofrida no processo de fabricação dos dispositivos.

Circuitos de teste foram degradados como prova de conceito, mostrando que modelos preditivos menos complexos são capazes de estimar o MTTF com uma precisão de aproximadamente 90% no melhor dos casos. Em adição, um banco de dados que represente estaticamente as condições ambientais de operação de um sistema não se mostra realístico.

6.1 Futuros trabalhos

Trabalhos futuros envolvem a automatização completa de todo o processo de extração de caminhos críticos, síntese e representação esquemática.

Abordagens diferentes de aprendizado de máquina (*p.ex.* redes neurais), podem ser utilizadas e comparadas aos métodos utilizados, averiguando o impacto na área, consumo e precisão das estimativas. Além disso, a metodologia pode ser melhorada para que seja possível modelar o BDPE em função do número de amostras e métodos de estimativa utilizados, permitindo ao projetista analisar o *trade-off* entre precisão, área ocupada pelo BDPE e métodos escolhidos.

Referências

- 1 LEBLEBICI, Y.; KANG, S.-M. S. *Hot-Carrier Reliability of MOS VLSI Circuits*. [S.l.]: Springer US, 1993. 212 p. ISBN 1461532507. Citado 2 vezes nas páginas 7 e 8.
- 2 MARICAU, E.; GIELEN, G. *Analog IC reliability in nanometer CMOS*. [S.l.]: Springer, 2013. ISBN 1461461634. Citado 11 vezes nas páginas 7, 3, 5, 9, 10, 11, 12, 15, 18, 21 e 22.
- 3 MCPHERSON, J. Time dependent dielectric breakdown physics – Models revisited. *Microelectronics Reliability*, v. 52, n. 9, p. 1753–1760, 2012. ISSN 00262714. Citado 3 vezes nas páginas 7, 13 e 14.
- 4 JAMES, G. et al. *An Introduction to Statistical Learning*. New York, NY: Springer New York, 2013. v. 103. (Springer Texts in Statistics, v. 103). ISBN 978-1-4614-7137-0. Citado 5 vezes nas páginas 7, 29, 30, 33 e 55.
- 5 HANSEN, M. C.; YALCIN, H.; HAYES, J. P. Unveiling the ISCAS-85 benchmarks: a case study in reverse engineering. *IEEE Design and Test of Computers*, v. 16, n. 3, p. 72–80, 1999. ISSN 07407475. Citado 3 vezes nas páginas 9, 51 e 52.
- 6 MOORE, G. E. *Cramming more components onto integrated circuits (Reprinted from Electronics, pg 114-117, April 19, 1965)*. [S.l.], 1965. v. 86, n. 1, 82–85 p. Citado na página 1.
- 7 SORIN, D. J. *Fault Tolerant Computer Architecture*. [S.l.: s.n.], 2009. v. 4. 1–104 p. ISSN 1935-3235. ISBN 9781598299540\n1935-3243 ;. Citado 2 vezes nas páginas 3 e 4.
- 8 SEYMOUR, B. MTTF, FAILRATE, RELIABILITY AND LIFE TESTING. *Burr-Brown Corporation*, p. 6, 1993. Citado na página 3.
- 9 CALABRO, S. R. *Reliability Principles and Practices*. [S.l.]: McGraw-Hill, 1962. 371 p. Citado na página 4.
- 10 KOREN, I.; KRISHNA, C. M. C. M. *Fault-tolerant systems*. [S.l.]: Elsevier/Morgan Kaufmann, 2007. 378 p. ISBN 0080492681. Citado na página 5.
- 11 SALEH, J. H.; MARAIS, K. Highlights from the early (and pre-) history of reliability engineering. In: *Reliability Engineering and System Safety*. [S.l.: s.n.], 2006. v. 91, n. 2, p. 249–256. ISBN 0951-8320. ISSN 09518320. Citado na página 5.
- 12 ENDRENYI, J. *Reliability modeling in electric power systems*. [S.l.]: Wiley, 1979. 338 p. ISBN 0471996645. Citado na página 5.
- 13 SCHLÄPFER, M.; KESSLER, T.; KRÖGER, W. Reliability Analysis of Electric Power Systems Using an Object-oriented Hybrid Modeling Approach. Citado na página 5.
- 14 STEPHENS, D.; VANHOUCKE, T.; DONKERS, J. J. T. M. RF reliability of short channel NMOS devices. In: *2009 IEEE Radio Frequency Integrated Circuits Symposium*. [S.l.]: IEEE, 2009. p. 343–346. ISBN 978-1-4244-3377-3. Citado na página 5.

- 15 LARCHER, L. et al. Oxide Breakdown After RF Stress: Experimental Analysis and Effects on Power Amplifier Operation. In: *2006 IEEE International Reliability Physics Symposium Proceedings*. [S.l.]: IEEE, 2006. p. 283–288. ISBN 0-7803-9498-4. Citado na página 5.
- 16 OSORIO, C.; SAWANT, N. Transformer Lifetime Prediction. *Intelligent Energy Projects, ee292k edition*. Citado na página 5.
- 17 JIRUTITIJAROEN, P.; SINGH, C. The effect of transformer maintenance parameters on reliability and cost: a probabilistic model. *Electric power systems Research*, 2004. Citado na página 5.
- 18 LIMOUSIN, M. et al. Value of Automatic Processing and Reliability of Stored Data in an Implanted Pacemaker: Initial Results in 59 Patients. *Pacing and Clinical Electrophysiology*, Blackwell Publishing Ltd, v. 20, n. 12, p. 2893–2898, dec 1997. ISSN 0147-8389. Citado na página 5.
- 19 ARAIN, N. et al. Comprehensive proficiency-based inanimate training for robotic surgery: reliability, feasibility, and educational benefit. *Surgical*, 2012. Citado na página 5.
- 20 NELSON, W. B. *Accelerated Testing: Statistical Models, Test Plans, and Data Analysis*. [S.l.: s.n.], 2004. 624 p. ISBN 0471697362. Citado na página 5.
- 21 SIA. The international technology roadmap for semiconductors. *Chips 2020*, 2001. Citado na página 5.
- 22 SIEWIOREK, D. P.; SWARZ, R. S. *Reliable computer systems : design and evaluation*. [S.l.]: Digital Press, 1992. 908 p. ISBN 1483297438. Citado na página 6.
- 23 SOLOMON, P. Breakdown in silicon oxide—A review. *Journal of Vacuum Science and Technology*, v. 14, n. 1977, p. 1122, 1977. ISSN 00225355. Citado na página 7.
- 24 BENTARZI, H. *Transport in metal-oxide-semiconductor structures : mobile ions effects on the oxide properties*. [S.l.]: Springer, 2011. 104 p. ISBN 3642163041. Citado na página 7.
- 25 SHOCKLEY, W. Problems related to p-n junctions in silicon. *Czechoslovak Journal of Physics*, v. 11, n. 2, p. 81–121, 1961. ISSN 00114626. Citado na página 9.
- 26 TAKEDA, E.; SUZUKI, N.; HAGIWARA, T. Device performance degradation to hot-carrier injection at energies below the Si-SiO₂ energy barrier. In: *1983 International Electron Devices Meeting*. [S.l.]: IRE, 1983. p. 396–399. Citado 2 vezes nas páginas 9 e 11.
- 27 WANG, W. W. W. et al. Compact Modeling and Simulation of Circuit Reliability for 65-nm CMOS Technology. *IEEE Transactions on Device and Materials Reliability*, v. 7, n. 4, p. 509–517, 2007. ISSN 15304388. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4359941>>. Citado 2 vezes nas páginas 9 e 12.
- 28 BRAVAIX, A. et al. Hot-carrier acceleration factors for low power management in DC-AC stressed 40nm NMOS node at high temperature. In: *IEEE International Reliability Physics Symposium Proceedings*. [S.l.: s.n.], 2009. p. 531–548. ISBN 0780388038. ISSN 15417026. Citado na página 9.

- 29 PARTHASARATHY, C. R. et al. Designing in reliability in advanced CMOS technologies. *Microelectronics Reliability*, v. 46, n. 9-11, p. 1464–1471, 2006. ISSN 00262714. Citado 2 vezes nas páginas 9 e 19.
- 30 SZE, S. M.; NG, K. K. *Physics of semiconductor devices*. [S.l.]: Wiley-India, 2007. 815 p. ISBN 8126517026. Citado na página 9.
- 31 MARICAU, E.; De Wit, P.; GIELEN, G. An analytical model for hot carrier degradation in nanoscale CMOS suitable for the simulation of degradation in analog IC applications. *Microelectronics Reliability*, v. 48, n. 8, p. 1576–1580, 2008. ISSN 00262714. Citado na página 11.
- 32 KUFLUOGLU, H.; ALAM, M. A geometrical unification of the theories of NBTI and HCI time-exponents and its implications for ultra-scaled planar and surround-gate MOSFETs. In: *IEDM Technical Digest. IEEE International Electron Devices Meeting, 2004*. [S.l.]: IEEE, 2004. p. 113–116. ISBN 0-7803-8684-1. Citado 3 vezes nas páginas 11, 17 e 18.
- 33 Chenming Hu et al. Hot-Electron-Induced MOSFET Degradation - Model, Monitor, and Improvement. *IEEE Journal of Solid-State Circuits*, v. 20, n. 1, p. 295–305, feb 1985. ISSN 0018-9200. Citado 3 vezes nas páginas 11, 14 e 17.
- 34 AMAT, E. et al. Channel Hot-Carrier degradation in short channel devices with high-k/metal gate stacks. In: *2009 Spanish Conference on Electron Devices*. [S.l.]: IEEE, 2009. p. 238–241. ISBN 978-1-4244-2838-0. Citado na página 11.
- 35 CHO, M. et al. Positive and negative bias temperature instability on sub-nanometer eot high-K MOSFETs. In: *2010 IEEE International Reliability Physics Symposium*. [S.l.]: IEEE, 2010. p. 1095–1098. ISBN 978-1-4244-5430-3. Citado na página 11.
- 36 GRASSER, T.; KACZER, B. Evidence That Two Tightly Coupled Mechanisms Are Responsible for Negative Bias Temperature Instability in Oxynitride MOSFETs. *IEEE Transactions on Electron Devices*, v. 56, n. 5, p. 1056–1062, may 2009. ISSN 0018-9383. Citado na página 12.
- 37 MCPHERSON, J.; BAGLEE, D. Acceleration Factors for Thin Gate Oxide Stressing. In: *23rd International Reliability Physics Symposium*. [S.l.]: IEEE, 1985. p. 1–5. Citado na página 14.
- 38 MINTARNO, E. et al. Optimized self-tuning for circuit aging. *Proceedings -Design, Automation and Test in Europe, DATE*, p. 586–591, 2010. ISSN 15301591. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-77953113043{&}partnerID=tZOtx>>. Citado na página 15.
- 39 MITRA, S. et al. Robust system design to overcome CMOS reliability challenges. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, v. 1, n. 1, p. 30–41, 2011. ISSN 21563357. Citado na página 15.
- 40 AGARWAL, M. et al. Circuit Failure Prediction and Its Application to Transistor Aging. In: *25th IEEE VLSI Test Symposium (VTS'07)*. IEEE, 2007. p. 277–286. ISBN 0-7695-2812-0. ISSN 1093-0167. Disponível em: <<http://ieeexplore.ieee.org/document/4209925/>>. Citado na página 15.

- 41 CORNELIUS, C. et al. Encountering gate oxide breakdown with shadow transistors to increase reliability. In: *Proceedings of the twenty-first annual symposium on Integrated circuits and system design - SBCCI '08*. New York, New York, USA: ACM Press, 2008. p. 111. ISBN 9781605582313. Disponível em: <<http://portal.acm.org/citation.cfm?doid=1404371.1404407>>. Citado na página 16.
- 42 SHI, K.; HOWARD, D. Sleep Transistor Design and Implementation – Simple Concepts Yet Challenges To Be Optimum. Citado na página 16.
- 43 SRINIVASAN, J. et al. RAMP: A Model for Reliability Aware MicroProcessor Design. *IBM Research Report*, v. 23048, 2003. Citado na página 16.
- 44 TORRES, F. S. Enhancement of System-Lifetime by Alternating Module Activation. *Embedded Systems: Design, Analysis and Verification*, p. 82–91, 2013. Citado na página 16.
- 45 BARANOWSKI, R. et al. On-Line Prediction of NBTI-induced Aging Rates. In: *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2015*. New Jersey: IEEE Conference Publications, 2015. p. 589–592. ISBN 9783981537048. Disponível em: <<http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7092455>>. Citado 3 vezes nas páginas 17, 25 e 26.
- 46 MAEDA, R. K. V.; TORRES, F. S. Clever: Cross-Layer Error Verification, Evaluation and Reporting. In: *Proceedings of the 27th Symposium on Integrated Circuits and Systems Design - SBCCI '14*. New York, New York, USA: ACM Press, 2014. p. 1–7. ISBN 9781450331562. Citado na página 17.
- 47 ALAM, M. A critical examination of the mechanics of dynamic NBTI for PMOSFETs. *IEEE International Electron Devices Meeting 2003*, p. 345–348, 2003. ISSN 01631918. Citado na página 19.
- 48 SCHRODER, D. K.; BABCOCK, J. A. Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing. *Journal of Applied Physics*, American Institute of PhysicsAIP, v. 94, n. 1, p. 1–18, jul 2003. ISSN 0021-8979. Citado na página 19.
- 49 CHAKRAVARTHI, S. et al. A comprehensive framework for predictive modeling of negative bias temperature instability. *2004 IEEE International Reliability Physics Symposium. Proceedings*, p. 273–282, 2004. ISSN 15417026. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1315337>>. Citado na página 19.
- 50 WU, E. et al. Interplay of voltage and temperature acceleration of oxide breakdown for ultra-thin oxides. *Microelectronic Engineering*, v. 59, n. 1, p. 25–31, 2001. Citado na página 20.
- 51 WU, E. Y.; SUÑÉ, J. Power-law voltage acceleration: A key element for ultra-thin gate oxide reliability. *Microelectronics Reliability*, v. 45, n. 12, p. 1809–1834, 2005. Citado 2 vezes nas páginas 20 e 21.
- 52 LI, X.; QIN, J.; BERNSTEIN, J. B. Compact modeling of MOSFET wearout mechanisms for circuit-reliability simulation. *IEEE Transactions on Device and Materials Reliability*, v. 8, n. 1, p. 98–121, 2008. ISSN 15304388. Citado na página 20.

- 53 SAHHAF, S. et al. A new TDDb reliability prediction methodology accounting for multiple SBD and wear out. *IEEE Transactions on Electron Devices*, v. 56, n. 7, p. 1424–1432, 2009. ISSN 00189383. Citado na página 21.
- 54 URMANOV, A. Electronic prognostics for computer servers. In: *2007 Proceedings - Annual Reliability and Maintainability Symposium, RAMS*. [S.l.: s.n.], 2007. p. 65–70. ISBN 0780397665. ISSN 0149144X. Citado na página 24.
- 55 KIM, K. K.; WANG, W.; CHOI, K. On-Chip Aging Sensor Circuits for Reliable Nanometer MOSFET Digital Circuits. *IEEE Transactions on Circuits and Systems II: Express Briefs*, v. 57, n. 10, p. 798–802, oct 2010. Citado na página 25.
- 56 KEANE, J. et al. An all-in-one silicon odometer for separately monitoring HCI, BTI, and TDDb. In: *IEEE Journal of Solid-State Circuits*. [S.l.: s.n.], 2010. v. 45, n. 4, p. 817–829. ISBN 978-1-4244-3307-0. ISSN 00189200. Citado na página 25.
- 57 KIM, T. H.; PERSAUD, R.; KIM, C. H. Silicon odometer: An on-chip reliability monitor for measuring frequency degradation of digital circuits. In: *IEEE Journal of Solid-State Circuits*. [S.l.: s.n.], 2008. v. 43, n. 4, p. 874–880. ISBN 9784900784048. ISSN 00189200. Citado na página 25.
- 58 KARL, E. et al. Compact In-Situ Sensors for Monitoring Negative-Bias-Temperature-Instability Effect and Oxide Degradation. In: *2008 IEEE International Solid-State Circuits Conference - Digest of Technical Papers*. [s.n.], 2008. p. 410–623. ISBN 978-1-4244-2010-0. ISSN 0193-6530. Disponível em: <<http://ieeexplore.ieee.org/document/4523231/>>. Citado 2 vezes nas páginas 25 e 38.
- 59 MCCULLAGH, P. Generalized linear models. *European Journal of Operational Research*, v. 16, n. 3, p. 285–292, jun 1984. Citado na página 29.
- 60 CHATTERJEE, S.; HADI, A. S. *Regression analysis by example*. [S.l.: s.n.]. 421 p. ISBN 1119122732. Citado 2 vezes nas páginas 31 e 35.
- 61 MAITRA, S.; YAN, J. Principle component analysis and partial least squares: Two dimension reduction techniques for regression. *Applying Multivariate Statistical Models*, 2008. Citado na página 33.
- 62 DEZA, M. M.; DEZA, E. *Encyclopedia of distances*. [S.l.: s.n.], 2009. 1–590 p. ISSN 14338351. ISBN 9783642002335. Citado na página 34.
- 63 RODGERS, J. L.; NICEWANDER, W. A. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, v. 42, n. 1, p. 59, 1988. ISSN 00031305. Disponível em: <<http://www.jstor.org/stable/2685263?origin=crossref>>. Citado na página 35.
- 64 ALCORN, P. *Intel Broadwell-EP Architecture, Models & Pricing*. 2016. 9 p. Disponível em: <<http://www.tomshardware.com/reviews/intel-xeon-e5-2600-v4-broadwell-ep,4514-2.html>>. Citado na página 43.
- 65 ORSHANSKY, M.; KEUTZER, K. A general probabilistic framework for worst case timing analysis. In: *ACM. Proceedings of the 39th annual Design Automation Conference*. [S.l.], 2002. p. 556–561. Citado na página 44.

- 66 POLI, A. A.; CIRILLO, M. C. On the use of the normalized mean square error in evaluating dispersion model performance. *Atmospheric Environment. Part A. General Topics*, v. 27, n. 15, p. 2427–2434, 1993. ISSN 09601686. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/096016869390410Z>>. Citado na página 54.
- 67 WILLMOTT, C. J.; MATSUURA, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, Inter-Research Science Center, v. 30, p. 79–82. Disponível em: <<http://www.jstor.org/stable/24869236>>. Citado na página 54.
- 68 CHAI, T.; DRAXLER, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geosci. Model Dev*, v. 7, p. 1247–1250, 2014. Disponível em: <www.geosci-model-dev.net/7/1247/2014/>. Citado na página 54.
- 69 MONTGOMERY, D. C.; RUNGER, G. C. *Applied statistics and probability for engineers*. [S.l.]: Wiley, 2011. 768 p. ISBN 0470053046. Citado na página 62.