



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS REITOR JOÃO DAVID FERREIRA LIMA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL

Leonardo Mazzaferro

**Metamodelagem dinâmica aplicada à carga térmica de refrigeração de escritórios
condicionados artificialmente em climas brasileiros**

Florianópolis

2021

Leonardo Mazzaferro

**Metamodelagem dinâmica aplicada à carga térmica de refrigeração de escritórios
condicionados artificialmente em climas brasileiros**

Tese submetida ao Programa de Pós-Graduação em
Engenharia Civil da Universidade Federal de Santa
Catarina para a obtenção do título de doutor em
Engenharia Civil.

Orientador: Prof. Roberto Lamberts, PhD.

Florianópolis

2021

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Mazzaferro, Leonardo

Metamodelagem dinâmica aplicada à carga térmica de
refrigeração de escritórios condicionados artificialmente em
climas brasileiros / Leonardo Mazzaferro ; orientador,
Roberto Lamberts, 2021.

115 p.

Tese (doutorado) - Universidade Federal de Santa
Catarina, Centro Tecnológico, Programa de Pós-Graduação em
Engenharia Civil, Florianópolis, 2021.

Inclui referências.

1. Engenharia Civil. 2. Metamodelagem dinâmica. 3. Dados
climáticos. 4. Amostragem adaptativa. I. Lamberts,
Roberto. II. Universidade Federal de Santa Catarina.
Programa de Pós-Graduação em Engenharia Civil. III. Título.

Leonardo Mazzaferro

**Metamodelagem dinâmica aplicada à carga térmica de refrigeração de escritórios
condicionados artificialmente em climas brasileiros**

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof.^a Ana Paula Melo, Dr.^a

Universidade Federal de Santa Catarina

Prof. Mauro Roisenberg, Dr.

Universidade Federal de Santa Catarina

Prof. Paulo Smith Schneider, Dr.

Universidade Federal do Rio Grande do Sul

Prof. Aldomar Pedrini, PhD.

Universidade Federal do Rio Grande do Norte

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Engenharia Civil.

Coordenação do Programa de Pós-Graduação

Prof. Roberto Lamberts, PhD.

Orientador

Florianópolis, 2021

Este trabalho é dedicado aos meus pais, à minha namorada, aos meus professores e colegas.

AGRADECIMENTOS

Agradeço aos meus pais pelo carinho, pelo suporte, e por terem proporcionado condições para a exploração e o aprendizado.

Agradeço à Isabela de Paulo, pelo amor, pela compreensão, pelos incentivos e pela confiança depositada em mim.

Agradeço ao Prof. Roberto Lamberts pela orientação, pela grande paciência, pelos sábios conselhos e pela parceria na jornada.

Agradeço à Prof.^a Ana Paula Melo pela dedicação, pela energia, pela atenção e por manter vivo o espírito de equipe no LabEEE.

Agradeço aos professores, que gentilmente, dedicaram seu tempo e suas energias para contribuir com o desenvolvimento desta tese. Muito obrigado, Prof.^a Ana Paula Melo, Prof. Mauro Roisenberg, Prof. Paulo Smith Schneider, e Prof. Aldomar Pedrini.

Agradeço à direção, aos funcionários e aos demais professores do Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Santa Catarina pela estrutura e apoio fornecido para a realização do doutorado.

Agradeço à CAPES e à Eletrobras, pelas bolsas de estudo concedidas. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Finalmente, agradeço a todos os meus queridos colegas e amigos do LabEEE, pela aliança forjada ao longo dos desafios enfrentados nas pesquisas. Meu muito obrigado!

RESUMO

Esta tese desenvolveu uma metamodelagem dinâmica, aplicada à carga térmica de refrigeração de escritórios condicionados artificialmente, em climas brasileiros. A metamodelagem dinâmica consiste em um processo iterativo de desenvolvimento do metamodelo, realizado a partir de amostragens adaptativas, que ocorre até que o critério de parada seja alcançado. No processo de metamodelagem foram incluídas variações em diversos parâmetros da geometria, da envoltória, das cargas internas, do entorno, e uma ampla faixa de temperaturas de controle para condicionamento das zonas térmicas da edificação. À medida que aplicações baseadas em metamodelos necessitam abranger numerosas combinações de edificações e localidades, o processo de metamodelagem deve priorizar uma amostragem eficiente e vasta aplicabilidade climática. A amostragem visa a maior diversidade de casos, com o menor número de simulações, enquanto a vasta aplicabilidade permite a utilização do metamodelo em múltiplas localidades. A discrepância das amostras definiu qual técnica de amostragem seria utilizada no processo de metamodelagem. Dentre as técnicas de amostragem analisadas, a sequência de Sobol apresentou os menores valores de discrepância, cobrindo mais uniformemente o espaço amostral. A adoção da temperatura de bulbo seco, da temperatura de ponto de orvalho e da radiação horizontal global, em valores médios mensais forneceu alta diversidade de dados climáticos a partir de poucas simulações, tornando o processo de metamodelagem mais eficiente. Estas variáveis climáticas foram selecionadas a partir de análises exploratórias entre a carga térmica de refrigeração em edificações de escritório e dados climáticos de 411 localidades brasileiras. A característica adaptativa da amostragem possibilitou desenvolver a base de dados de forma incremental, evitando o descarte das simulações anteriores. O critério para interromper a metamodelagem dinâmica foi a raiz do erro quadrático médio (RMSE) da rede neural, estabelecido em 2 kWh/m², de forma a concentrar as análises da pesquisa no início do processo de metamodelagem. As redes neurais artificiais alcançaram este valor após consumir dados de apenas 400 simulações e foram validadas em baterias de teste, através de milhares de combinações nunca vistas, de zonas térmicas e climas. Na validação, os valores normalizados da raiz do erro quadrático médio (NRMSE) obtidos para a rede neural artificial da última iteração foram de apenas 2,5%. A simplificação da arquitetura das redes neurais, com camadas ocultas de 32 e 16 neurônios, visou evitar a ocorrência de *overfitting* nas primeiras iterações da metamodelagem dinâmica. Ao realizar a amostragem com a técnica que obteve a maior discrepância (Halton), os valores de NRMSE obtidos foram de 70% a 100% superiores àqueles obtidos com a técnica de menor discrepância (Sobol), o que reforça a importância da amostragem para o desenvolvimento de metamodelos a partir de simulações. A maior eficiência em preencher o espaço amostral também reduziu a quantidade de simulações fornecidas às redes neurais artificiais para atingir a acurácia desejada. Caso a aplicação do metamodelo demande uma acurácia mais elevada, esta pode ser aumentada progressivamente, dando continuidade à metamodelagem dinâmica.

Palavras-chave: Metamodelagem dinâmica. Dados climáticos. Amostragem adaptativa.

ABSTRACT

This thesis developed a dynamic metamodeling, applied to thermal cooling loads of artificially conditioned offices, in Brazilian climates. The dynamic metamodeling consists of an iterative process of developing the metamodel, carried out from adaptive sampling, which occurs until the stopping criterion is reached. The metamodeling process included variations in several parameters of geometry, envelope, internal loads, surroundings, and a wide range of control temperatures for conditioning the building's thermal zones. As metamodel-based applications need to cover numerous combinations of buildings and locations, the metamodeling process must prioritize efficient sampling and wide climatic applicability. Sampling aims at the greatest diversity of cases, with the least number of simulations, while the wide applicability allows the use of the metamodel in multiple locations. The sample discrepancy defined which sampling technique would be used in the metamodeling process. Among the sampling techniques analyzed, the Sobol sequence showed the lowest discrepancy values, covering the sample space more evenly. The adoption of dry bulb temperature, dew point temperature and global horizontal radiation, in monthly average values, provided a high diversity of climatic data from a few simulations, making the metamodeling process more efficient. These climatic variables were selected from exploratory analysis between the thermal cooling loads in office buildings and climatic data from 411 Brazilian locations. The adaptive sampling feature made it possible to develop the database incrementally, avoiding the disposal of previous simulations. The criterion for interrupting the dynamic metamodeling was the root mean squared error (RMSE) of the neural network, established at 2 kWh / m², in order to concentrate the research analysis at the beginning of the metamodeling process. Artificial neural networks reached this value after consuming data from only 400 simulations and were validated in test batteries, through thousands of unseen combinations, of thermal zones and climates. In the validation, the normalized root mean squared error (NRMSE) obtained in the artificial neural network of the last iteration were only 2.5%. The simplification of neural networks architecture, with hidden layers of 32 and 16 neurons, aimed to avoid the occurrence of overfitting in the first iterations of the dynamic metamodeling. When carrying out the sampling with the technique that obtained the greatest discrepancy (Halton), the NRMSE values obtained were 70% to 100% higher than those obtained with the technique with the least discrepancy (Sobol), which reinforces the importance of sampling for the development of metamodels from simulations. Greater efficiency in filling the sample space also reduced the amount of simulations provided to artificial neural networks to achieve the desired accuracy. If the metamodel application requires a higher accuracy, it can be increased progressively, by continuing the dynamic metamodeling.

Keywords: Dynamic metamodeling. Weather data. Adaptive sampling.

LISTA DE FIGURAS

Figura 2.1 Divisão proposta pelo método “Decision tree method” para caracterizar os climas brasileiros.	29
Figura 2.2. Classificações das amostragens no DOE moderno.	32
Figura 2.3. Representação dos benefícios da transferência de aprendizado.	38
Figura 3.1. Organização do método.	42
Figura 3.2. Fluxograma da metamodelagem dinâmica.	45
Figura 3.3. Baixa densidade horizontal da grade urbana (valor = 0,03).	47
Figura 3.4. Elevada densidade horizontal da grade urbana (valor = 0,46).	48
Figura 3.5. Zoneamento térmico da edificação com uma zona interna e quatro zonas perimetrais.	51
Figura 4.1. Matriz de correlação entre variáveis climáticas e cargas térmicas anuais de refrigeração das simulações preliminares.	56
Figura 4.2. Gráfico de dispersão entre os valores mensais de temperatura de bulbo seco e de carga térmica de refrigeração sensível mensal, por zona térmica.	58
Figura 4.3. Gráficos de dispersão entre os valores mensais de temperatura do ponto de orvalho e a carga térmica de refrigeração latente mensal, por zona térmica.	59
Figura 4.4. Incidência de valores horários das variáveis climáticas selecionadas.	60
Figura 4.5. Incidência de valores médios anuais das variáveis selecionadas.	62
Figura 4.6. Incidência de valores médios mensais das variáveis selecionadas.	62
Figura 4.7. Discrepância em função da técnica e do tamanho da amostra.	64
Figura 4.8. Amostragem de 100 pontos de acordo com a técnica de Sobol.	67
Figura 4.9. Procedimento de amostragem dos arquivos climáticos.	69
Figura 4.10. Matriz de correlação entre as 26 variáveis preditoras e a variável dependente. ...	72
Figura 4.11. Representação da arquitetura da rede neural artificial.	73
Figura 4.12. Conjuntos de dados de treinamento (80%) e avaliação (20%), da última iteração.	74
Figura 4.13. Predições para escritório com 15 zonas térmicas ao longo de 12 meses, para as cidades de Curitiba e São Paulo.	75
Figura 4.14. Predições para escritório com 15 zonas térmicas ao longo de 12 meses, para as cidades de Rio de Janeiro e Belém.	76

Figura 4.15. Conjunto de dados de teste aplicado nas redes neurais artificiais da primeira e da última iteração.	77
Figura 4.16. Curva de evolução das previsões por iteração, ao utilizar Sobol na amostragem, para casos vistos (treinamento e avaliação) à esquerda, e casos nunca vistos (teste) à direita.	78
Figura 4.17. Curva de evolução das previsões por iteração, ao utilizar Halton na amostragem, para casos vistos (treinamento e avaliação) à esquerda, e casos nunca vistos (teste) à direita.	79
Figura 1.B. Matriz de dispersão dos dados de arquivos EPW - versão INMET 2016.	95
Figura 2.B. Número de cidades descritas na RNA-C por grupo climático.	96
Figura 3.C. Amostragem de 100 pontos de acordo com a técnica de Sobol.	98
Figura 4.C. Amostragem de 100 pontos de acordo com a técnica de Torus.	99
Figura 5.C. Amostragem de 100 pontos de acordo com a técnica de Halton.....	100
Figura 6.C. Amostragem de 100 pontos de acordo com a técnica de LHS melhorada.	101
Figura 7.C. Amostragem de 100 pontos de acordo com a técnica de LHS otimizada.	102
Figura 8.C. Amostragem de 100 pontos de acordo com a técnica de LHS randômica.	103

LISTA DE TABELAS

Tabela 2.1. Artigos de revisão sobre metamodelagem em função dos temas listados.....	21
Tabela 2.2. Palavras-chave utilizadas nas buscas por publicações.....	22
Tabela 2.3. Estudos que desenvolveram metamodelos para mais de um clima.	27
Tabela 3.1. Parâmetros da edificação das simulações preliminares.	43
Tabela 3.2. Variações por parâmetro da edificação.....	49
Tabela 3.3. Premissas válidas para todas as simulações.....	50
Tabela 4.1. Número de dados de saída obtidos a partir de uma simulação.	61
Tabela 4.2. Valores de discrepância para 20 dimensões.	65
Tabela 4.3. Lista das localidades amostradas por iteração.	69
Tabela 4.4. Tratamento efetuado para adequar os dados às granularidades escolhidas.	71
Tabela 1.A. Códigos e descrição.	93
Tabela 2.A. Bases de dados climáticas e descrição.	93
Tabela 3.B. Limites inferiores e superiores em cada variável adotada	94
Tabela 4.D. Valores médios das médias mensais de temperatura de bulbo seco, e os valores de desvio padrão das médias mensais de temperatura de bulbo seco, de acordo com o clima. ...	104

LISTA DE ABREVIATURAS E SIGLAS

ALC - *Annual latent cooling load*

ANN - *Artificial neural network*

ASC - *Annual sensible cooling load*

ATC - *Annual total cooling load*

CDD - *Cooling degree days*

CgTR - Carga térmica de refrigeração anual integrada

HDD - *Heating degree days*

IDF - *Input Data File*

INI-C - Instrução Normativa Inmetro para a Classificação de Eficiência Energética de Edificações Comerciais, de Serviços e Públicas

LabEEE - Laboratório de Eficiência Energética em Edificações

MVS - Máquina de vetores de suporte

NRMSE - Raiz do erro quadrático médio normalizado

PBE - Programa Brasileiro de Etiquetagem

PROCEL - Programa Nacional de Conservação de Energia Elétrica

PROCEL EDIFICA - Programa Nacional de Eficiência Energética em Edificações

R² - Coeficiente de determinação

RMSE - Raiz do erro quadrático médio

RLM - Regressão linear múltipla

RNA - Rede neural artificial

RNA-C - Rede neural artificial implementada na INI-C

RTQ-C - Regulamento Técnico da Qualidade do Nível de Eficiência Energética de Edifícios Comerciais, de Serviço e Públicas

RTQ-R - Regulamento Técnico da Qualidade do Nível de Eficiência Energética de Edifícios Residenciais

UFSC - Universidade Federal de Santa Catarina

WWR - *Window-to-wall Ratio*

SUMÁRIO

1. INTRODUÇÃO	14
1.1. APRESENTAÇÃO DO PROBLEMA	14
1.2. OBJETIVOS	19
1.2.1. Objetivo Geral	19
1.2.2. Objetivos Específicos	19
1.3. ESTRUTURA DA TESE	19
2. REVISÃO BIBLIOGRÁFICA	21
2.1. DADOS CLIMÁTICOS EM METAMODELOS	23
2.1.1. Clima único	23
2.1.2. Múltiplos climas	25
2.1.3. Zoneamentos climáticos	28
2.1.4. Reflexão da seção	30
2.2. TÉCNICAS DE AMOSTRAGEM PARA SIMULAÇÕES	30
2.2.1. Discrepância	34
2.2.2. Reflexão da seção	34
2.3. DESENVOLVIMENTO DE REDES NEURAIAS ARTIFICIAIS	35
2.3.1. Tratamento dos dados	35
2.3.2. Treinamento	36
2.3.3. Validação	39
2.3.4. Reflexão da seção e considerações finais da revisão	41
3. MÉTODO	42
3.1. ANÁLISES EXPLORATÓRIAS	42
3.1.1. Dados climáticos	43
3.1.2. Preenchimento do espaço amostral	44
3.2. METAMODELAGEM DINÂMICA	45
3.2.1. Simulação	46
3.2.1.1. Geometria	46
3.2.1.2. Envoltória	46
3.2.1.3. Cargas internas	47
3.2.1.4. Entorno	47
3.2.1.5. Sistema	48

3.2.2. Amostragem adaptativa	51
3.2.3 Tratamento de dados	52
3.2.4. Treinamento e avaliação	54
3.3. TESTES (VALIDAÇÕES)	55
4. RESULTADOS	56
4.1. DADOS CLIMÁTICOS.....	56
4.2. PREENCHIMENTO DO ESPAÇO AMOSTRAL	63
4.3. METAMODELAGEM DINÂMICA	68
4.4. TESTES (VALIDAÇÕES).....	75
5. CONCLUSÃO	81
5.1. LIMITAÇÕES.....	83
5.2. SUGESTÕES PARA TRABALHOS FUTUROS.....	84
REFERÊNCIAS	85
APÊNDICE A - Códigos desenvolvidos e bases de dados	93
APÊNDICE B - Diversidade climática na RNA-C	94
APÊNDICE C - Preenchimento do espaço amostral	98
APÊNDICE D - Valores médios das médias mensais de temperatura de bulbo seco	104

1. INTRODUÇÃO

1.1. APRESENTAÇÃO DO PROBLEMA

As edificações são grandes consumidores de energia na sociedade moderna. No mundo, o consumo de energia voltado para a climatização de edificações triplicou em menos de trinta anos. No Brasil, as edificações representam 51% do consumo final de eletricidade, subdivididos no setor residencial com 26%; no setor comercial com 17%; e no setor público com 8%. Entre 30% e 40% do consumo de eletricidade das edificações é atribuído à climatização artificial dos ambientes (EPE, 2020). A demanda de energia para refrigeração segue em aumento, reforçada pelo crescimento demográfico e econômico de regiões com climas quentes. Esta conjuntura de fatores desperta preocupações relacionadas aos impactos ambientais e fomenta pesquisas com foco na redução das demandas energéticas de edificações (IEA, 2018).

Neste contexto, o objetivo das políticas de eficiência energética é exatamente a desaceleração do crescimento das demandas energéticas. No Brasil, as políticas iniciaram com o Programa Nacional de Conservação de Energia Elétrica (PROCEL) em 1985, promovendo ações em diversos setores da economia. Focado no uso racional de energia elétrica em edificações, o Programa Nacional de Eficiência Energética em Edificações (PROCEL EDIFICA) foi instituído em 2003. No âmbito do Programa Brasileiro de Etiquetagem (PBE), lançou-se o Regulamento Técnico da Qualidade do Nível de Eficiência Energética de Edifícios Comerciais, de Serviços e Públicos (RTQ-C) em 2009, e o seu equivalente para edifícios residenciais (RTQ-R) em 2010.

Originalmente, o RTQ-C e o RTQ-R permitiram avaliar a classificação das edificações por equações de regressão linear múltipla, válidas de acordo com a zona bioclimática da localidade, e elaboradas a partir de simulações energéticas. Entretanto, o sucesso de regulamentos de eficiência energética está relacionado à aplicabilidade e robustez dos métodos utilizados para avaliar a economia de energia em edificações (WEI et al., 2018; BRE et al., 2020). Por este motivo, ao longo da última década, foi dada continuidade ao desenvolvimento de metamodelos direcionados para as políticas de eficiência energética em edificações no Brasil. Melo (2012) simulou variações em edificações de escritórios condicionados artificialmente, e propôs redes neurais artificiais para substituir as regressões lineares múltiplas presentes no RTQ-C. Versage (2015) optou por desenvolver redes neurais

artificiais a partir de simulações simplificadas, de modo a criar maior diversidade de casos, e representar as zonas térmicas de edificações comerciais. Ambas as pesquisas definiram a carga térmica de refrigeração anual integrada como variável dependente, mas limitaram-se ao clima de Florianópolis. A consideração de múltiplos climas em um único metamodelo foi sugerida como contribuição importante em eventuais trabalhos futuros (MELO, 2012; VERSAGE, 2015; MELO, 2016).

A evolução dos estudos nestes temas possibilitou o desenvolvimento de uma rede neural artificial aplicável em diferentes realidades climáticas. Para facilitar sua identificação neste documento, a referida rede neural artificial foi denominada RNA-C. A Instrução Normativa Inmetro para a Classificação de Eficiência Energética de Edificações Comerciais, de Serviços e Públicas (INI-C) foi desenvolvida para substituir o RTQ-C. A INI-C adota a RNA-C para estimar a carga térmica de refrigeração anual integrada (CgTR) de zonas térmicas de edificações comerciais condicionadas artificialmente (INMETRO, 2020).

A RNA-C foi desenvolvida a partir de simulações energéticas em localidades que representam os grupos climáticos brasileiros (RORIZ, 2014; BAVARESCO et al., 2017). Apesar de ser aplicável em diferentes grupos climáticos brasileiros, a RNA-C não deve ser aplicada em localidades específicas, com exceção das localidades utilizadas em seu desenvolvimento. Para maior compreensão das limitações de um determinado metamodelo, é necessário conhecer a sua modelagem, ou seja, o seu processo de desenvolvimento.

Em estudos de eficiência energética em edificações, por falta de dados reais, o desenvolvimento de metamodelos depende da geração de uma base de dados através de simulações energéticas. As simulações energéticas utilizam equações físicas, definidas segundo os princípios da conservação de energia, e buscam reproduzir o comportamento da edificação em um determinado clima (HENSEN & LAMBERTS, 2019). Apesar da alta confiabilidade atribuída às simulações, os profissionais de arquitetura e engenharia relatam dificuldades na modelagem e na exploração do universo de possibilidades através das simulações (ALSAADANI, 2016; BLEIL, 2016; HONG et al., 2018). Em diversos casos, os custos temporais e/ou financeiros das simulações energéticas são incompatíveis com as demandas das aplicações, fomentando simplificações e métodos alternativos (NEGENDAHL, 2015; OSTERGARD et al., 2018).

Nos últimos anos, verificou-se um forte crescimento no número de estudos que desenvolveram metamodelos para avaliar a eficiência energética em edificações (WESTERMANN & EVINS, 2019; BRUNETTI, 2020). As redes neurais artificiais representaram mais de um terço destes estudos, e aumentam em proporção a cada ano, em

função de seu desempenho e aplicabilidade, especialmente em fenômenos não-lineares. Por reduzirem drasticamente o custo computacional, as redes neurais artificiais são utilizadas em ferramentas de auxílio ao projeto; em estudos de sensibilidade e otimização; e até em programas de etiquetagem de edificações (OSTERGARD et al., 2016; ROMAN et al., 2020).

Independentemente de sua finalidade, os metamodelos necessitam primordialmente de dados. As bases de dados, em estudos que adotam a metamodelagem para avaliar a eficiência energética de edificações, podem ser compostas por variáveis relacionadas à edificação e variáveis relacionadas ao clima. Por serem definidas diretamente na modelagem das simulações, as variáveis relacionadas à edificação são descritas facilmente para o metamodelo. Já as variáveis relacionadas ao clima provêm de arquivos climáticos com valores horários, e tendem a ser descritas por valores médios. Contudo, a descrição por valores médios pode resultar em perda de informações valiosas, especialmente ao longo de um ano, devido à sazonalidade climática. O tipo de descrição climática adotada e a diversidade de valores fornecida ao metamodelo permitem ou impedem a aplicação do metamodelo em localidades diferentes àquelas utilizadas em seu treinamento.

A temperatura de bulbo seco, a radiação horizontal global e a umidade relativa do ar são as variáveis mais utilizadas para caracterizar os climas, e costumam ser descritas em granularidade temporal anual nos metamodelos (AMASYALI & EL-GOHARY, 2018). A granularidade temporal define o intervalo de tempo no qual a variável climática e a variável de saída são consideradas, podendo variar entre anual, mensal, diário, horário ou sub-horário. Em metamodelos desenvolvidos para considerar somente uma única localidade, a incidência de variáveis climáticas é associada a granularidades temporais inferiores à anual (YILDIZ et al., 2017; BOURDEAU et al., 2019). Metamodelos desenvolvidos com simulações em mais de uma localidade são incomuns, e geralmente consideram ao menos um clima frio, um intermediário e um clima quente (YAN & YAO, 2010; CATALINA et al., 2013; SINGARAVEL et al., 2017). Todavia, ao restringir a variedade de climas simulados, limita-se a aplicabilidade do metamodelo em outras localidades.

Os estudos que efetivamente consideram múltiplas localidades em um único metamodelo são ainda mais raros. Lopes e Lamberts (2018) simplificaram a descrição climática ao adotar a entalpia como único indicador climático para descrever 18 arquivos climáticos brasileiros, mas encontraram limitações ao extrair predições para localidades com temperaturas amenas. Rackes et al. (2016) adotaram uma abordagem quase oposta, ao utilizar treze variáveis climáticas, calculadas a partir de 427 arquivos climáticos de localidades

brasileiras. Westermann et al. (2020) evitaram a representação numérica ao implementar o reconhecimento de imagens sobre as séries temporais de 569 arquivos climáticos canadenses.

Contudo, não é computacionalmente eficiente utilizar a totalidade dos arquivos climáticos disponíveis para compor a base de dados de simulações a ser fornecida como insumo ao metamodelo, pois diversas localidades apresentam climas semelhantes. Ao considerar a restrição de recursos computacionais, as variações (possivelmente desnecessárias) de climas reduzem as variações que poderiam ser efetuadas sobre a edificação. Além disso, do ponto de vista descritivo, a adoção de numerosas variáveis climáticas também pode ser ineficiente, uma vez que as variáveis climáticas apresentam correlações entre si, o que acarreta em repetições entre os dados que descrevem os climas. A descrição climática poderia simplesmente focar nas variáveis que apresentam as maiores influências sobre a variável dependente.

Supõe-se que a limitada quantidade de estudos sobre o tema esteja relacionada à ausência de aplicações que demandam numerosas combinações de edificações e realidades climáticas, concomitantemente. À medida que surgem aplicações que se beneficiam de tamanha diversidade de dados, como é o caso da INI-C, torna-se essencial aperfeiçoar o desenvolvimento de metamodelos, maximizando o valor dos dados provenientes das simulações.

Em estudos de eficiência energética em edificações, o desenvolvimento de metamodelos tende a ser realizado através de abordagens estáticas, com separações claras entre o começo e o fim das etapas de metamodelagem. A metamodelagem, quando realizada a partir de simulações energéticas de edificações, é tipicamente composta por cinco etapas: a amostragem, a simulação, o tratamento de dados, o treinamento e a validação. É essencial zelar pela robustez das etapas, pois afetam a capacidade de predição do metamodelo (AHMAD et al., 2018; MOHANDÉS et al., 2019).

A amostragem é a etapa que determina a diversidade de combinações inseridas nas simulações. Apesar de ser a etapa inicial, a amostragem é explorada e documentada de forma limitada por estudos que desenvolvem metamodelos a partir de simulações energéticas em edificações. O critério de definição do tamanho da amostra não é especificado em mais da metade dos estudos analisados, que optam por adotar valores sem justificativa ou fórmulas questionáveis (WESTERMANN & EVINS, 2019; FATHI et al., 2020; ROMAN et al., 2020).

A eficiência das técnicas de amostragem, em preencher o espaço amostral, varia de acordo com o tamanho da amostra e com o número de dimensões consideradas. Entretanto, a eficiência de preenchimento do espaço amostral pode ser quantificada pela discrepância.

Independente da técnica de amostragem utilizada, à medida que a amostra aumenta de tamanho, a tendência é que a discrepância da amostra diminua gradativamente. Quanto mais baixa é a discrepância da amostra, mais uniforme é a distribuição dos pontos amostrados (KUCHERENKO et al., 2015; LOYOLA et al., 2016).

A adoção de técnicas de amostragem com menores valores de discrepância para um mesmo tamanho de amostra pode otimizar a metamodelagem, ao minimizar a quantidade de simulações que o metamodelo necessita consumir para atingir os níveis de acurácia desejados. Indubitavelmente, a influência do tamanho da amostra merece exploração, especialmente em estudos de espaço amostral com múltiplas dimensões (OSTERGARD et al., 2017, ROMAN et al., 2020; WESTERMANN et al., 2020).

Em contraste com a amostragem estática, adotada na vasta maioria dos estudos revisados, a amostragem adaptativa permite gerar novas combinações, a partir das combinações amostradas previamente, evitando o descarte de casos simulados e preenchendo as lacunas do espaço amostral de forma incremental (GARUD et al., 2017). Teoricamente, a amostragem adaptativa permite aumentar progressivamente a diversidade de edificações e climas na base de dados, até o ponto em que as predições do metamodelo alcançam os valores de acurácia desejados.

A combinação da amostragem adaptativa com o caráter iterativo intrínseco às redes neurais artificiais tende a viabilizar o desenvolvimento do processo de metamodelagem de forma dinâmica e incremental. Visando ganhos de eficiência computacional, pode-se testar a técnica de transferência de aprendizado entre as redes neurais artificiais. Esta abordagem permite exportar o aprendizado da rede neural artificial da iteração anterior para a iteração subsequente.

Ao consolidar as informações expostas, surgem possíveis respostas para a pergunta "Como desenvolver um metamodelo que considere efetivamente múltiplos climas brasileiros, com o mínimo de recursos computacionais e que possa ser melhorado progressivamente?". Constata-se que o desenvolvimento de um metamodelo aplicável em múltiplos climas envolve principalmente desafios atrelados à descrição climática e à diversidade de dados fornecidos aos metamodelos.

A descrição climática engloba a seleção de variáveis climáticas em função da variável dependente. A diversidade de dados é favorecida pela adoção de uma granularidade temporal adequada e por uma amostragem eficiente. Estes desafios demandaram análises exploratórias focadas nos dados climáticos e na eficiência das técnicas de amostragem em preencher o espaço amostral. Ao considerar o custo computacional das simulações e o uso eficiente de

recursos computacionais, uma metamodelagem dinâmica pode fazer parte da solução. Neste trabalho, a metamodelagem dinâmica consiste em um processo iterativo de desenvolvimento do metamodelo, realizado a partir de amostragens adaptativas, até que um determinado critério de parada seja alcançado.

1.2. OBJETIVOS

1.2.1. Objetivo Geral

O objetivo desta tese é desenvolver um processo dinâmico de metamodelagem, aplicado à carga térmica de refrigeração de escritórios condicionados artificialmente, em climas brasileiros.

1.2.2. Objetivos Específicos

A partir do objetivo geral, delimitaram-se os seguintes objetivos específicos:

- a) Identificar variáveis e granularidades adequadas para descrever os dados climáticos na metamodelagem, em função da variável dependente;
- b) Avaliar a eficiência de preenchimento do espaço amostral de técnicas de amostragem, em múltiplos tamanhos de amostra, para posterior adoção na metamodelagem;
- c) Quantificar a evolução da acurácia de predição das redes neurais artificiais à medida que a base de dados de simulação aumenta.

1.3. ESTRUTURA DA TESE

A tese está dividida em cinco seções. O primeiro capítulo contextualiza o tema, ao apresentar os desafios que motivam a realização do trabalho e os objetivos a serem alcançados com a pesquisa.

Elaborada para mapear o estado da arte nos assuntos abordados, a revisão de literatura está disposta no segundo capítulo. Foram exploradas áreas adjacentes do conhecimento a fim de encontrar inspirações para o desenvolvimento desta tese.

O terceiro capítulo fornece os procedimentos do método, adotados no desenvolvimento da pesquisa. O método está dividido em três módulos: as análises exploratórias, a metamodelagem dinâmica e os testes (validações).

Os resultados da pesquisa são apresentados pelo quarto capítulo, de acordo com os procedimentos metodológicos propostos. O quarto capítulo contém as análises, as discussões e as justificativas adotadas na pesquisa.

O quinto capítulo apresenta as conclusões da tese, as limitações, e as sugestões para trabalhos futuros.

2. REVISÃO BIBLIOGRÁFICA

O interesse crescente por metamodelos em estudos de eficiência energética em edificações é evidenciado por recentes publicações de artigos de revisão nesta área do conhecimento. Estes artigos de revisão exploram múltiplas abordagens e métodos adotados por dezenas de autores que desenvolveram metamodelos aplicados à eficiência energética em edificações. Os temas abordados pelos artigos de revisão foram expostos na Tabela 2.1, de acordo com: a abordagem de metamodelo, o tipo de condicionamento (artificial ou natural), a presença de variáveis climáticas, a granularidade temporal adotada, a consideração de múltiplos climas em metamodelos, a consideração do entorno e a amostragem (técnicas, tamanhos de amostra e eficiência).

Tabela 2.1. Artigos de revisão sobre metamodelagem em função dos temas listados.

Artigos de revisão	Autores (et al.)	Yildiz	Ahmad	Amasyali	Wei	Bourdeau	Mohandes	Westermann	Fathi	Roman
	Ano	2017	2018	2018	2018	2019	2019	2019	2020	2020
	Abordagem metamodelo	Diversos	Diversos	Diversos	Diversos	Diversos	RNA	Diversos	Diversos	Diversos
	Tipo de condicionamento	Artificial	Artificial e natural	Artificial	Artificial e natural	Artificial e natural	Artificial e natural	Artificial e natural	Artificial e natural	Artificial e natural
Dados climáticos	Variáveis climáticas	sim	não	sim	sim	sim	sim	sim	não	sim
	Granularidade temporal	sim	sim	sim	sim	sim	sim	sim	sim	sim
	Múltiplos climas	parcial	não	não	parcial	parcial	parcial	sim	não	não
	Consideração do entorno	não	não	não	não	não	não	não	não	não
Amostragem	Técnicas e tamanho	não	parcial	não	parcial	não	parcial	sim	parcial	sim
	Eficiência	não	não	não	sim	não	não	sim	não	sim
Tratamento dados	Granularidade espacial	não	sim	não	sim	sim	não	não	sim	sim
	Pré-processamento	não	sim	sim	não	sim	sim	sim	parcial	sim
	Seleção de características	não	parcial	parcial	não	sim	não	sim	sim	sim
Treinamento e validação	Otimização de hiperparâmetros	não	não	não	não	não	sim	sim	não	sim
	Transferência de aprendizado	não	não	não	não	sim	não	não	não	não
	Rigor da validação	parcial	parcial	sim	não	parcial	sim	parcial	não	sim

Fonte: elaborada pelo autor (2020).

A granularidade está relacionada à escala em que os dados são considerados no experimento. A granularidade pode ser separada em dois tipos: a temporal e a espacial. Na escala espacial, os dados podem ser compilados por blocos urbanos, edificação, pavimento ou ambiente/zona térmica. Na escala temporal, os dados podem ser analisados em períodos anuais, mensais, semanais, diários, horários ou sub-horários. As oscilações de valores das variáveis climáticas tornam-se menos bruscas em períodos de tempo maiores. A dispersão de valores é maior em intervalos temporais menores.

Há vantagens e desvantagens para cada escala, seja ela temporal ou espacial. As escalas maiores permitem trabalhar com menos dados, mas podem limitar a quantidade de informações embutidas no metamodelo. As escalas menores permitem considerar informações em níveis mais detalhados, mas podem refletir em uma quantidade desnecessária ou excessiva de dados. Os estudos que aliam a simulação energética à metamodelagem adotam diferentes variáveis climáticas e granularidades temporais, de acordo com seus respectivos objetivos. Por último, a Tabela 2.1 também indica se as respectivas publicações abordaram itens relacionados ao tratamento de dados, ao treinamento e validação de metamodelos. A cor verde indica “sim” e a cor vermelha indica “não”. A ocorrência do termo “parcial” significa que o artigo de revisão comentou o tema, mas não se aprofundou.

A análise dos artigos de revisão possibilitou confirmar os temas com potenciais de pesquisa compatíveis com esta tese. Entre as principais lacunas constam a consideração efetiva de múltiplos climas e a eficiência da amostragem. Desta forma, o capítulo da revisão bibliográfica foi organizado visando mapear e expandir os conhecimentos sobre as lacunas identificadas. As buscas por publicações foram executadas através da plataforma *ScienceDirect*, e da base de dados *Scopus*. A busca de publicações foi realizada ao combinar as palavras-chave descritas na Tabela 2.2.

Tabela 2.2. Palavras-chave utilizadas nas buscas por publicações.

Buscas (em inglês)	Tradução
“ <i>metamodeling</i> ”	metamodelagem
“ <i>data-driven</i> ”	baseado em dados
“ <i>surrogate model</i> ”	modelo substituto
“ <i>artificial neural networks</i> ”	redes neurais artificiais
“ <i>building energy performance</i> ”	desempenho energético de edificações
“ <i>cooling loads</i> ”	carga de resfriamento

<i>“sampling methods”</i>	métodos de amostragem
<i>“climate data”</i>	dados climáticos
<i>“weather files”</i>	arquivo climático
<i>“climatic variables”</i>	variáveis climáticas

Fonte: elaborada pelo autor (2020).

Publicações adicionais foram coletadas a partir das referências encontradas nos diversos artigos, à medida que as leituras avançaram. A partir da leitura dos estudos, a revisão bibliográfica foi estruturada em três seções: dados climáticos em metamodelos, técnicas de amostragem para simulações, e desenvolvimento de redes neurais artificiais.

2.1. DADOS CLIMÁTICOS EM METAMODELOS

Esta seção expõe os dados climáticos mais utilizados para descrever os climas em metamodelos. Inicialmente, foram analisados estudos que consideraram somente um clima, e posteriormente, estudos com múltiplos climas. Ao final da seção também foram adicionadas considerações relacionadas aos métodos de zoneamento dos climas brasileiros.

2.1.1. Clima único

Braun et al. (2014) combinaram a umidade relativa do ar e a temperatura de bulbo seco em regressões lineares múltiplas para estimar o consumo de edificações de supermercados, no Reino Unido. Devido à alta relevância da temperatura de bulbo seco, a remoção da umidade relativa entre as variáveis independentes acarretou em pequena redução do desempenho das regressões.

Jain et al. (2014) desenvolveram máquinas de vetor de suporte, adotando diferentes granularidades temporais e espaciais, para prever o consumo de energia de edificações. Os parâmetros utilizados foram os valores de consumo de dois intervalos de tempo anteriores, a temperatura do ar, o tipo de dia e de hora. Os autores concluíram que as granularidades impactam de forma significativa o desempenho das predições. Wang et al. (2016) também utilizaram máquinas de vetores de suporte para estimar o consumo de energia de aquecimento, adotando a temperatura do ar, umidade relativa, velocidade do vento e radiação solar, em base horária.

A umidade relativa do ar é comumente adotada por estudos que envolvem climas quentes e úmidos, pois impacta o consumo de energia em edificações condicionadas. A

radiação solar também é muito utilizada para descrever os climas, pois o efeito dos ganhos solares é significativo nas demandas energéticas das edificações, especialmente em climas intertropicais e/ou em edificações com maiores proporções de componentes translúcidos nas fachadas (BOURDEAU et al., 2019). Entretanto, os valores de temperatura e radiação solar aos quais as edificações reais estão expostas variam de acordo com a localização da edificação no ambiente urbano. A modelagem do entorno, mesmo que simples, permite ajustar parcialmente as diferenças de desempenho entre a edificação simulada e a edificação real (MIRZAEI, 2015).

De acordo com Lima et al. (2019), em climas quentes e úmidos, a consideração do entorno urbano nas simulações, acarretou em redução de aproximadamente 17% na carga térmica de edificações de escritórios condicionados. A modelagem do entorno possibilita considerar o sombreamento e as reflexões provenientes das edificações vizinhas. Segundo Machado (2019), os maiores erros causados por simplificações na modelagem do entorno ocorrem em configurações urbanas mais adensadas, e em edificações de escritórios com baixas cargas internas e/ou com maiores porcentagens de vidro na fachada. A consideração da ilha de calor urbana pode ser efetuada pela modificação do arquivo climático. Depois de modificados, os arquivos climáticos representam as realidades climáticas adaptadas ao contexto urbano.

Mena et al. (2014) desenvolveram uma rede neural artificial para estimar a demanda de energia de uma edificação bioclimática. Os dados foram organizados em valores horários e de acordo com o tipo do dia. As variáveis climáticas adotadas foram a temperatura externa e a radiação solar. Chae et al. (2016) também desenvolveram redes neurais artificiais para prever o consumo de energia elétrica em edificações, em intervalos de quinze minutos. As variáveis climáticas utilizadas foram a temperatura e a umidade relativa do ar. Além dos valores médios, Fan et al. (2014) incluíram as temperaturas de bulbo seco máximas e mínimas diárias, no desenvolvimento de metamodelos de diversos tipos para estimar o consumo diário de energia.

A temperatura de bulbo seco também pode ser representada na forma de indicadores, como os graus-hora de aquecimento (GHA) e os graus-hora de resfriamento (GHR), ou os graus-dia de aquecimento (GDA) e os graus-dia de resfriamento (GDR). Os graus-hora e graus-dia são vastamente utilizados para descrever resumidamente os climas, na literatura de eficiência energética em edificações (LAM et al., 2002; HYGH et al., 2012). Por serem calculados a partir de valores de temperatura de bulbo seco, estes indicadores são mais apropriados para descrever climas secos, ou seja, com predominância do calor sensível. Em

climas quentes e úmidos, o calor latente é responsável por uma parte considerável do consumo com condicionamento, devido ao processo de desumidificação. Na busca por um indicador que considere ambas as parcelas de calor, pesquisadores também utilizaram a entalpia (KRESE et al., 2011).

Lopes (2019) procurou determinar o consumo anual para climatização de escritórios com sistemas do tipo água gelada através de um metamodelo baseado em redes neurais artificiais. O autor critica a limitação de uso de metamodelos em somente um clima e utiliza a entalpia (*cooling enthalpy hours* - CEH) para descrever 18 climas brasileiros, relacionada com a variável de saída. Entretanto, a predição de consumo em climas amenos apresentou erros elevados e restringiu o uso do metamodelo em cidades quentes.

Um fator de atenção relacionado ao uso de indicadores, como graus-hora, ou entalpia é que são atrelados a um limiar (normalmente uma temperatura base) e podem conflitar com outras variáveis da base de dados do metamodelo. Um exemplo de conflito seria a utilização de CDD (*cooling degree days*) calculados a 18°C para descrever os climas em simulações de edificações com temperatura de controle (*setpoint*) de condicionamento a 24°C. Na margem, o conflito diminui o valor dos dados fornecidos aos metamodelos, mas pode ser corrigido ao equalizar as temperaturas utilizadas como limiares.

2.1.2. Múltiplos climas

Os estudos que desenvolveram metamodelos em mais de um clima são incomuns. Yan e Yao (2010) utilizaram redes neurais artificiais (RNA) para prever o consumo de energia em cinco zonas climáticas da China. O metamodelo foi desenvolvido a partir de 20 parâmetros, dos quais dois eram relacionados ao clima: os graus-dia de aquecimento (GDA) e os graus-dia de resfriamento (GDR).

Catalina et al. (2013) desenvolveram um modelo de regressão linear múltipla (RLM) para estimar a demanda de energia para aquecimento em edificações residenciais. O modelo é simplificado e apresenta somente três variáveis, das quais somente uma está relacionada ao clima. A variável que representa o clima é calculada a partir da diferença entre a temperatura de ajuste interna para aquecimento e a temperatura sol-ar externa. O modelo foi desenvolvido a partir de simulações em três climas: Moscou (Rússia), Bucareste (Romênia) e Nice (França). Moscou foi escolhida para representar um clima frio, enquanto Nice buscou representar um clima quente e mediterrâneo. Bucareste foi escolhida como clima de características intermediárias. O modelo de regressão múltipla linear apresentou bons resultados nos três climas e foi considerado como uma alternativa ágil para estimar a

demanda de energia de aquecimento. Entretanto, os autores não realizaram testes para climas diferentes dos três citados.

Brownlee e Wright (2015) adotaram funções de base radial (FBR) para otimizar o projeto de edificações, em função da energia consumida e do custo de construção da edificação. Os climas considerados foram Birmingham (Reino Unido), Chicago (EUA) e São Francisco (EUA). Os autores desenvolveram três metamodelos separados, ou seja, consideraram somente uma localidade por metamodelo.

Rackes et al. (2016) adotaram as máquinas de vetores de suporte para desenvolver um metamodelo que avalia edificações comerciais naturalmente ventiladas, em climas quentes e amenos. Os autores executaram 50 mil simulações distribuídas em 427 localidades brasileiras. O estudo utiliza diversas variáveis para descrever o clima: a temperatura externa, a amplitude da temperatura externa, os graus-dia, a radiação solar, a velocidade do vento e a altitude. Ao analisar as variáveis escolhidas pelos autores, nota-se certa redundância na extensa descrição climática, que pode ter sido uma estratégia para maior diferenciação entre os climas, uma vez que foi adotada a granularidade anual. Em novos trabalhos, pode-se reduzir o número de climas simulados e de variáveis descritas, ao adotar uma granularidade mensal, devido à maior diversidade de dados climáticos.

Amasyali e El-Gohary (2017) definiram 14 variáveis climáticas para treinar diversos tipos de metamodelos para prever o consumo horário de energia em escritórios em cinco climas americanos. As principais variáveis foram a temperatura de bulbo seco, a temperatura do ponto de orvalho, a umidade relativa, a pressão atmosférica, a radiação normal direta, radiação horizontal difusa, a direção e a intensidade do vento.

Singaravel et al. (2017) analisaram o consumo de energia de uma edificação simplificada, a partir das contribuições de seus componentes construtivos e da infiltração, nos climas de Amsterdã, Bruxelas e Paris. O estudo adotou a temperatura de bulbo seco, a radiação solar, a velocidade e direção do vento. No final, foram combinadas as previsões de quatro tipos de metamodelos: florestas aleatórias (*random forests*), redes neurais artificiais (*artificial neural networks*), *k*-vizinhos mais próximos (*k-nearest neighbors*), e árvores de regressão (*regression trees*).

Rossi et al. (2019) consideraram os climas de Curitiba, São Paulo e Manaus para desenvolver regressões lineares múltiplas que avaliam o desempenho térmico de residências ventiladas naturalmente. Os autores executaram dez mil simulações, amostradas pelo método de Monte Carlo, para treinar os metamodelos em cada clima.

Westermann et al. (2020) desenvolveu um metamodelo para predição das cargas térmicas de aquecimento e de resfriamento de uma edificação de escritório, com algumas variações, em 569 localidades do Canadá. Os autores desenvolveram redes neurais convolucionais para processar séries temporais de dados climáticos, em resolução horária. Apesar de considerarem somente localidades canadenses, os autores afirmaram não visualizar obstáculos para o desenvolvimento de metamodelos em escala global.

A Tabela 2.3 consolida os estudos revisados que desenvolveram metamodelos para mais de um clima, com suas respectivas informações de número de climas considerados e a motivação da escolha dos climas.

Tabela 2.3. Estudos que desenvolveram metamodelos para mais de um clima.

Artigo			Número de localidades		Escolha das localidades		Descrição climática	
Primeiro autor	Ano	Tipo	Nas simulações	Por metamodelo	Zoneamento	Caracterização	Variáveis	Granularidade
Jaffal	2009	RLM	3	1	-	X	-	anual
Yan	2010	RNA	5	5	X	-	graus-dia	anual
Lam	2010	RLM	5	1	X	-	-	anual
Hygh	2012	RLM	4	1	X	-	-	anual
Catalina	2013	RLM	3	3	-	X	T(sol-ar) - T(i)	anual
Brownlee	2015	FBR	3	1	-	X	-	anual
Romani	2015	RLM	6	1	X	-	-	anual
Rackes	2016	MVS	427	427	-	-	diversas	anual
Amasyali	2017	diversos	5	5	-	X	diversas	horária
Singaravel	2017	diversos	3	3	-	X	diversas	horária
Chen	2018	MVS	5	1	X	-	nº mês	mensal
Rossi	2019	RLM	3	1	-	X	-	anual
Lopes	2019	RNA	18	18	-	X	entalpia	anual
Westermann	2020	RNA	569	569	-	-	diversas	horária

Fonte: elaborada pelo autor (2020).

Apesar dos estudos citados envolverem simulações em localidades diferentes, somente a metade dos autores considera efetivamente mais de um clima no desenvolvimento de um mesmo metamodelo. A consideração efetiva de múltiplos climas nos metamodelos é indicada por valores maiores que 1 no número de localidades por metamodelo. Em

metamodelos desenvolvidos separadamente para cada localidade, os autores não adotaram variáveis para a descrição climática, mas o poderiam ter feito, se considerassem granularidades temporais menores. A exceção foi observada em Chen e Yang (2018), que desenvolveram máquinas de vetores de suporte (MVS) separadas por localidade, mas diferenciaram os meses enumerando-os.

Nos estudos citados, a escolha das localidades esteve relacionada à representação de zonas climáticas ou à caracterização do tipo de clima (quente ou frio; seco ou úmido). Os autores, em sua maioria, adotaram a granularidade anual nos dados climáticos, representada por valores médios. Para considerar diferentes climas nos metamodelos, também podem ser utilizadas descrições por variáveis categóricas. Entretanto, ao adotar variáveis categóricas, o uso do metamodelo gerado é limitado aos climas descritos pelas categorias. Para garantir a aplicação do metamodelo em múltiplos climas, é recomendado descrever numericamente as variáveis climáticas mais relevantes para a variável dependente (YILDIZ, 2017).

2.1.3. Zoneamentos climáticos

Por utilizarem dados climáticos e simulações energéticas, também foram explorados estudos de zoneamentos climáticos aplicados à eficiência energética em edificações. A abordagem mais comum para o zoneamento climático é o método graus-dia, pois está altamente correlacionado com o uso de energia em climas frios e amenos. Entretanto, o método dos graus-dias não considera outras variáveis climáticas que influenciam as necessidades energéticas do edifício em climas quentes ou amenos (DE ROSA et al., 2015).

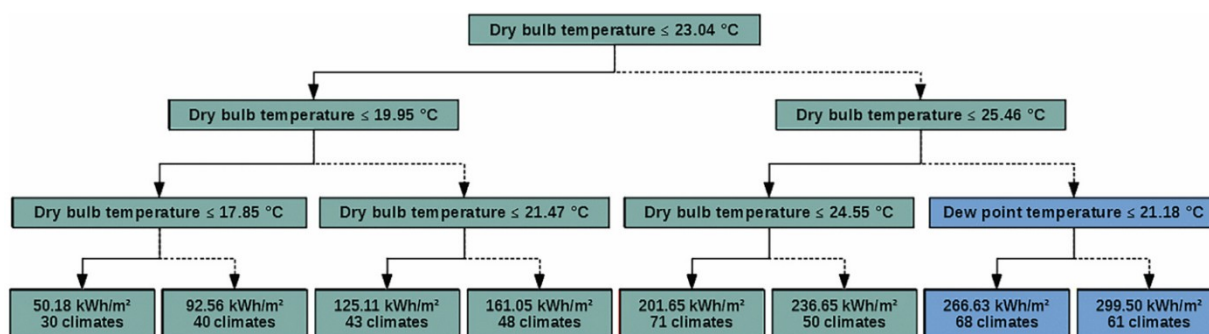
A temperatura de bulbo seco, a temperatura de bulbo úmido, a radiação horizontal global, a velocidade e a direção do vento são as principais variáveis climáticas utilizadas para estabelecer o zoneamento climático de países de dimensões continentais como Austrália, China e Índia (LAM et al., 2006). A relevância da variável climática depende fortemente da tipologia do edifício e varia para climas frios, amenos e quentes. Variáveis como a velocidade do vento, radiação solar e umidade relativa são particularmente relevantes para o desempenho de edifícios em climas quentes. Os edifícios localizados em climas tropicais são tipicamente mais sensíveis ao efeito climático, uma vez que esses edifícios geralmente apresentam menos isolamento em seu envelope que os edifícios de clima frio (WALSH et al., 2019). Portanto, a aplicação de metodologias de zoneamento climático e/ou a definição de descrições climáticas demasiadamente simplificadas tende a ser problemática em climas quentes.

Segundo a norma NBR 15220 - Parte 3 (ABNT, 2005) o território brasileiro é dividido em oito zonas bioclimáticas, estabelecidas a partir das médias mensais das

temperaturas máximas, das médias mensais das temperaturas mínimas, e das médias mensais de umidade relativa do ar. No caso do zoneamento proposto pela ASHRAE *Standard* 169 (*Climatic Data for Building Design Standards*) - 2013, o Brasil é dividido em quatro zonas climáticas, de acordo com os graus-dia de aquecimento, os graus-dia de refrigeração, e a precipitação. Por fim, quatro indicadores climáticos foram definidos por Roriz (2012) para propor um novo zoneamento climático brasileiro com 24 zonas climáticas (RORIZ, 2014). Segundo Roriz (2014), os climas brasileiros podem ser classificados de acordo com a temperatura média anual, a amplitude média anual, o desvio padrão da média mensal das temperaturas médias diárias e o desvio padrão da amplitude média mensal.

Mazzaferro et al. (2020) avaliaram quatro métodos de zoneamento climático existentes para o Brasil (ASHRAE 169, Koppen-Geiger, NBR 15220, e Roriz) e mais três métodos de zoneamento climático propostos pelos autores (*“Data-driven method”*, *“Enhanced degree-day method”* e *“Decision tree method”*). Ao comparar a unicidade e a dispersão que as simulações apresentaram nos sete métodos de zoneamento, o melhor resultado foi obtido no *“Decision tree method”*. Este método foi elaborado a partir de uma árvore de decisão com algoritmo de clusterização não supervisionado, que escolheu justamente a temperatura de bulbo seco e a temperatura do ponto de orvalho para caracterizar e clusterizar os 411 climas brasileiros, como apresentado na Figura 2.1.

Figura 2.1 Divisão proposta pelo método *“Decision tree method”* para caracterizar os climas brasileiros.



Fonte: adaptado de Mazzaferro et al. (2020)

Apesar de não ser adotada por tantos estudos, como é o caso da umidade relativa, a temperatura do ponto de orvalho apresenta forte potencial para representar a parcela latente da carga térmica de refrigeração em edificações condicionadas artificialmente.

2.1.4. Reflexão da seção

Nesta primeira parte da revisão foram mapeados os conhecimentos relacionados aos dados climáticos, úteis para formular a descrição climática que pode ser utilizada nos metamodelos. A maioria dos estudos que adotam dados climáticos entre as variáveis preditoras do metamodelo considera somente uma localidade.

Os dados climáticos costumam ser descritos por valores médios, mesmo ao longo de um ano inteiro. A temperatura de bulbo seco é a variável climática mais utilizada para descrever o clima entre os estudos, especialmente entre aqueles que estimam as demandas energéticas para condicionamento.

A adoção de diferentes variáveis climáticas varia de acordo com o objetivo dos estudos, a variável dependente e os climas considerados nas simulações. Adicionalmente, verificou-se que as descrições mais utilizadas em metamodelos coincidem, em grande parte, com as variáveis climáticas adotadas para elaborar zoneamentos climáticos.

A diversidade dos climas brasileiros aparenta ser suficientemente elevada para demandar, além da temperatura de bulbo seco, a adoção de variáveis como a radiação solar, e a umidade relativa ou a temperatura do ponto de orvalho. Caso seja definido um critério de exclusão de variáveis climáticas em função de altas correlações entre as mesmas, deve-se levar em consideração que as variáveis climáticas influenciam de forma diferente a edificação, como é o caso da temperatura de bulbo seco e da radiação solar.

Na próxima seção, buscaram-se soluções e inovações relacionadas à amostragem, uma vez que a simulação de múltiplas variações de edificações e climas pode elevar consideravelmente o número de combinações que devem ser simuladas para alimentar o metamodelo.

2.2. TÉCNICAS DE AMOSTRAGEM PARA SIMULAÇÕES

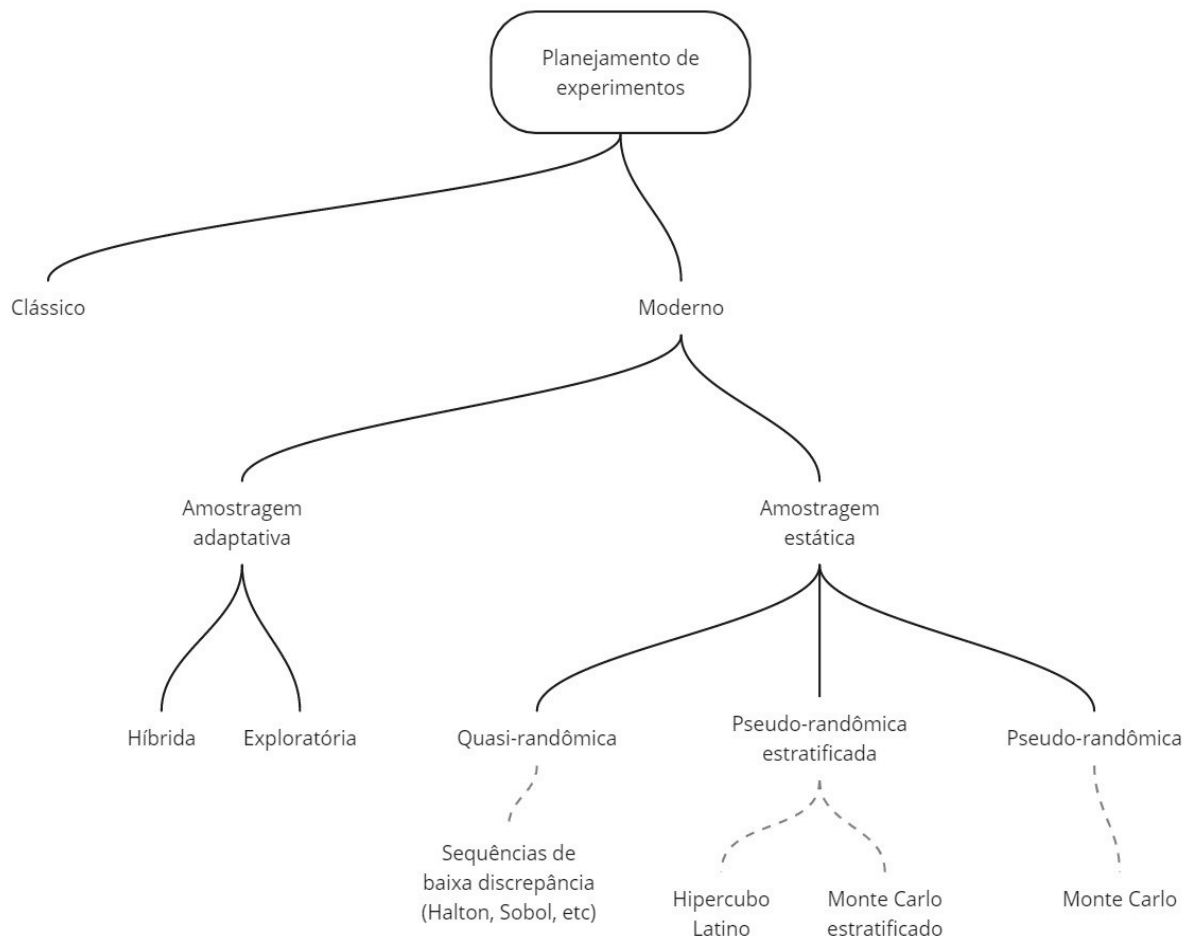
A principal sinergia entre metamodelos e estudos de simulação energética é a capacidade dos metamodelos de cobrir um amplo universo de casos a partir de um conjunto relativamente reduzido de dados, proveniente das simulações. No entanto, a performance do metamodelo depende da qualidade da amostra. Há pouco debate sobre os métodos de amostragem e a influência do tamanho da amostra sobre o desempenho do metamodelo. Segundo Roman et al. (2020), 63% dos estudos que desenvolveram redes neurais artificiais, a partir de simulações energéticas de edificações, não especificaram como definiram o tamanho da amostra.

Outros estudos definem o tamanho da amostra por meio de equações, ou optam por discretizar a amostra e simular a totalidade de combinações possíveis. A discretização da amostra possibilita encontrar as melhores combinações de dados de entrada na amostra, mas nada garante que estas coincidam com a melhor combinação no espaço amostral. Entretanto, a prioridade ao desenvolver metamodelos tende a ser a diversidade de dados e não as combinações. Técnicas de amostragem maximizam a quantidade de informação obtida a partir de um número reduzido de pontos.

Em experimentos computacionais com numerosas combinações e recursos disponíveis limitados, os métodos de amostragem surgem como alternativa viável. De forma geral, um bom método de amostragem deve criar um conjunto de dados que possa representar acuradamente a função subjacente, preservando as características estatísticas do conjunto de dados completo.

Os primeiros experimentos ocorreram por tentativa e erro, ou variavam um fator por vez. Além de impossibilitar a consideração da interação entre os fatores, estas abordagens eram altamente ineficientes. Para vencer a ineficiência, surgiu o planejamento de experimentos (“*Design of Experiments*”, ou simplesmente “DOE”). O planejamento de experimentos é uma metodologia estatística que permite analisar um dado sistema a partir de suas respostas, que por sua vez, variam de acordo com os fatores controláveis do sistema. Com o avanço dos métodos numéricos, surgem as simulações computacionais. A natureza determinística das simulações computacionais contrasta com o DOE clássico, que surgiu para lidar com fenômenos estocásticos. No planejamento de experimentos computacionais, também conhecido como DOE moderno, as técnicas são classificadas em amostragens estáticas ou adaptativas (KLEIJNEN, 2005; MONTGOMERY, 2013), como apresentado na Figura 2.2 abaixo.

Figura 2.2. Classificações das amostragens no DOE moderno.



Fonte: adaptado de Garud et al. (2017) e Westermann et al. (2020).

Na vasta maioria dos estudos, a metamodelagem é executada em blocos, ou seja, primeiro define-se a amostra com todas as simulações, para depois treinar o metamodelo. Esta forma de amostragem é denominada estática, pois as amostras são definidas a priori, e não são incrementadas ou ajustadas de acordo com a necessidade do metamodelo. As amostragens adaptativas, por outro lado, permitem obter conhecimento de forma dinâmica e sequencial. Amostras novas são adicionadas à base de dados, quando necessário, complementando as amostras anteriores. Critérios de otimização ou a própria acurácia do metamodelo pode ser utilizada para balizar a continuidade das iterações (WESTERMANN & EVINS, 2019).

A adoção de técnicas de preenchimento de espaço permite distribuir os pontos, referentes às combinações, de forma eficiente e dentro de intervalos definidos em cada variável de entrada. Entre as amostragens estáticas, encontram-se as técnicas pseudo-randômicas (com versões estratificadas) e as técnicas quasi-randômicas.

As técnicas pseudo-randômicas geram pontos com igual probabilidade de ocorrência. O método de amostragem mais comum é o de Monte Carlo clássico, também denominado amostragem aleatória simples (SRS), por sua simplicidade e aplicabilidade. No entanto, a SRS falha ao preencher eficientemente o espaço amostral, especialmente em espaços amostrais de alta dimensionalidade, pois tende a deixar regiões inexploradas e a gerar pontos muito próximos (YONDO et al, 2018).

As técnicas pseudo-randômicas estratificadas vencem esta limitação, evitando que sejam amostrados pontos repetidos. A amostragem por hipercubo latino (LHS) é um método comum no projeto de experimentos e é largamente adotada pelos estudos que avaliam o desempenho energético de edificações através de simulações e metamodelos. A LHS randômica define uma amostragem a partir de um conjunto de distribuições uniformes, de forma randômica, sem otimização. Entretanto, existem versões aprimoradas da LHS, que utilizam critérios como entropia, discrepância, e distâncias euclidianas (FANG, 2006; KUCHERENKO et al., 2015).

As técnicas quasi-randômicas, também denominadas sequências de baixa discrepância, unem a flexibilidade das amostragens randômicas à eficiência dos planejamentos uniformes. Entre as técnicas quasi-randômicas mais utilizadas, encontram-se as sequências de Halton, de Sobol, e de Kronecker (denominada Torus em certos estudos), seguidas por Niederreiter e Hammersley (LOYOLA et al., 2016). Estas sequências, além de preencherem o espaço amostral de forma eficiente, permitem a adição de novos pontos a partir dos pontos anteriores. Esta característica é muito interessante e pode fazer parte da solução para a definição do tamanho da amostra em estudos que desenvolvem metamodelos a partir de simulações energéticas em edificações (GARUD et al., 2017).

As técnicas de amostragem utilizam sementes de randomização (*random seeds*) para gerar números pseudo-aleatórios. Sem as sementes de randomização, as execuções da amostragem não poderiam ser reproduzidas de forma idêntica, pois o gerador de números pseudo-aleatórios seria inicializado com valores diferentes para cada execução da amostragem. As sementes de randomização garantem a reprodutibilidade por outros estudos e/ou a continuidade da amostragem no caso de amostragens adaptativas. A principal vantagem em adicionar novos pontos é o aumento da resolução da base de conhecimento, que é acessada pelo metamodelo durante o treinamento (YONDO et al., 2018).

Todavia, a capacidade de aumentar a amostra considerando os pontos anteriores não é exclusividade das técnicas quasi-randômicas. A LHS também permite realizar este tipo de amostragem aumentada. O trabalho de Bre et al. (2020) adota a amostragem por Hipercubo

Latino Aumentado para diminuir o número de simulações necessárias para treinar e validar as redes neurais artificiais.

2.2.1. Discrepância

Para comparar a eficiência de preenchimento do espaço amostral das técnicas, pode-se calcular a discrepância da amostra. A discrepância avalia a diferença entre a uniformidade da amostra e a uniformidade de uma amostra uniforme ideal. Existem diferentes formas de avaliar a discrepância, com fórmulas e complexidades de cálculo diferentes. As alternativas preferíveis são as discrepâncias L^2 , por apresentarem fórmulas mais compactas, conforme Equação 1 (DAMBLIN et al., 2013; LOYOLA et al., 2016).

$$D_{L^2}(X) = \left[\int_{[0,1]^{2d}} \left(\frac{A(X, J_{a,b})}{n} - Vol(J_{a,b}) \right)^2 da db \right]^{1/2} \quad (1)$$

Onde:

$D_{L^2}(X)$ é a discrepância L^2 de X ;

$Vol(J)$ é o volume de um subconjunto J ;

$A(X;J)$ é o número de pontos de X contidos em J .

Quanto menor a discrepância, para um determinado número de dimensões e de pontos considerados, maior é a eficiência do método de amostragem em preencher o espaço amostral (PERRIER et al., 2018). O preenchimento eficiente do espaço amostral favorece a eficiência do experimento, pois contribui para a diminuição do número de casos a serem simulados para atingir um determinado desempenho do metamodelo.

2.2.2. Reflexão da seção

Como apresentado, a execução da amostragem varia de acordo com a finalidade dos estudos. A discrepância pode ser uma ferramenta útil na escolha da técnica de amostragem para o desenvolvimento de metamodelos a partir de simulações energéticas, pois permite analisar a uniformidade de preenchimento do espaço amostral, em função do número de dimensões e amostras do experimento.

À medida que o número de dimensões do experimento aumenta, a sequência de Sobol e a amostragem por Hipercubo Latino apresentam vantagens sobre as outras técnicas, pois

são altamente eficientes em preencher o espaço amostral, e possibilitam realizar amostragens adaptativas.

2.3. DESENVOLVIMENTO DE REDES NEURAI ARTIFICIAIS

O avanço do conhecimento relacionado à ciência de dados e o aumento da capacidade de processamento dos computadores contribuíram para a proliferação de metamodelos em diversas áreas do conhecimento. Em estudos aplicados à eficiência energética de edificações, as redes neurais artificiais estão entre os tipos de metamodelos mais utilizados, seguidos por regressões polinomiais e máquinas de vetores de suporte. As redes neurais artificiais permanecem entre os metamodelos dominantes devido à sua acurácia, aplicabilidade e baixo custo computacional.

Por serem excelentes em processar fenômenos não lineares, as redes neurais artificiais podem ser utilizadas para resolver problemas complexos. A arquitetura das redes neurais é uma composição de neurônios e camadas. As camadas são classificadas em três categorias: de entrada, de saída e oculta(s). As camadas de entrada e de saída servem como canal de passagem para as informações, e o processamento das informações ocorre na camada oculta.

O treinamento, ou aprendizado da rede neural, é armazenado por pesos sinápticos. Um conjunto de dados é lido para produzir uma saída, que é posteriormente comparada com o padrão de saída correto ou desejado. Os pesos sinápticos são alterados iterativamente para a redução dos erros, até que os erros atinjam a tolerância admitida. A capacidade de predição da rede neural depende de sua arquitetura. Os autores de estudos aplicados à análise energética de edificações tendem a utilizar somente uma camada oculta na estrutura das redes neurais. Os tipos de redes neurais artificiais mais adotados são *multi-layer feed-forward neural networks* (MLFFNN), *general regression neural networks* (GRNN), e *recurrent neural networks* (RNN). O desenvolvimento de redes neurais a partir de simulações energéticas em edificações demanda pelo menos três etapas: o tratamento dos dados, o treinamento e a validação (MOHANDES et al., 2019).

2.3.1. Tratamento dos dados

Antes de iniciar o desenvolvimento da rede neural artificial, é essencial tratar os dados, com o objetivo de evitar erros computacionais e favorecer o aprendizado. Amasyali et al. (2017) centralizaram e escalonaram os dados antes do treinamento dos metamodelos, para maior compatibilidade com os algoritmos de aprendizado de máquina. Entretanto, o

tratamento de dados pode envolver diversos aspectos, como a plotagem das variáveis independentes para análise das distribuições, a transformação de variáveis (especialmente as categóricas), a análise de valores únicos e de covariância, a divisão de dados em partições de treinamento e avaliação (KUHN, 2008)

As variáveis categóricas, quando não ordenadas, não são significativas e devem ser transformadas. Variáveis fictícias são usualmente utilizadas para substituir variáveis categóricas. A técnica mais comum, para transformar as variáveis categóricas, é denominada “*dummy coding*”, e consiste em utilizar as categorias para criar novas variáveis e atribuir a elas valores binários de 0 ou 1. Entretanto, a técnica denominada “*effect coding*” (codificação de efeito), que atribui valores de -1 e 1, evita a saturação dos neurônios e tende a ser mais adequada para o treinamento do metamodelo (DALY et al., 2016; GROTENHUIS et al., 2017).

A engenharia de características (*feature engineering*) consiste no processo de utilizar o conhecimento de domínio sobre os dados para garantir um melhor desempenho dos modelos preditivos. O propósito da engenharia de características é a identificação, seleção e descrição otimizada das variáveis mais relevantes para o modelo preditivo. Um robusto processo de engenharia de características tende a influenciar positivamente o desempenho dos modelos preditivos, pois produz modelos mais simples e flexíveis, com um menor número de variáveis. Assim como um tratamento de dados bem realizado beneficia o aprendizado pela rede neural durante o treinamento, e permite agregar mais valor por caso simulado (BOURDEAU et al., 2019; WESTERMANN & EVINS, 2019; MILLER et al., 2020).

2.3.2. Treinamento

A configuração das redes neurais envolve a seleção de um conjunto ótimo de hiperparâmetros para atingir uma convergência rápida durante o treinamento e a acurácia desejada na predição. Os hiperparâmetros são configurações inerentes ao treinamento de modelos preditivos, e no caso de redes neurais artificiais, podem ser divididos em duas categorias: relacionados à arquitetura (número de camadas, número de neurônios por camada) e relacionados à otimização (como taxa de aprendizado, tamanho do lote, épocas, entre outros). A escolha dos hiperparâmetros varia de acordo com o tipo de problema e dos dados (ROMAN et al., 2020).

Estudos que utilizam redes neurais para prever o desempenho térmico de edificações pouco exploram os procedimentos de engenharia de características e de

otimização de hiperparâmetros, pois geralmente utilizam todas as variáveis das simulações como parâmetros de entrada da rede neural (DING et al., 2018). Valores ótimos para os hiperparâmetros podem ser identificados com o auxílio de grades de busca aleatória (BERGSTRA et al., 2012).

O processamento da informação em cada neurônio ocorre por meio da função de ativação. As funções de ativação mais comuns são a linear, a sigmoideal, a tangencial hiperbólica, e a linear retificada. A função de ativação linear retificada (ReLU) tem sido largamente adotada nos últimos anos, pois ataca o problema do desaparecimento do gradiente, evitando a lentidão no aprendizado (AMASYALI & EL-GOHARY, 2018).

Fenômenos mais complexos tendem a demandar arquiteturas mais complexas de redes neurais, e maiores tempos de treinamento. Entretanto, um número excessivo de neurônios pode piorar o desempenho da rede neural, causando *overfitting* do metamodelo. O *overfitting* ocorre quando o metamodelo se adapta excessivamente aos dados de treinamento, e conseqüentemente, não consegue realizar boas predições para novos dados de entrada. A preocupação com o *overfitting* ainda é incomum entre os estudos revisados, o que compromete a aplicabilidade dos metamodelos desenvolvidos em condições diferentes daquelas treinadas (SRIVASTAVA et al., 2014; MOHANDDES et al., 2019).

Para evitar o *overfitting*, pode-se recorrer a alguns métodos: um número máximo de iterações (épocas), uma amostra de avaliação com critérios de parada, ou implementando a validação cruzada. A amostra de avaliação permite monitorar o desempenho do metamodelo em um conjunto de dados que não está sendo usado no treinamento. Inclusive, este é o processo utilizado pela maioria dos autores. Bre et al. (2020) sugerem outro procedimento, que adota a validação cruzada e o uso de toda a amostra no treinamento, pois a representatividade da amostra original é comprometida após a divisão dos dados entre os conjuntos de treinamento e avaliação.

Na validação cruzada, uma sub-amostra dos dados é utilizada para treinar uma rede e o restante dos dados é usado para avaliar seu desempenho, repetidas vezes. Alternativamente, para aumentar a qualidade da validação, pode-se elaborar amostras independentes para treinar, validar e testar a rede neural desenvolvida. A elaboração de amostras independentes auxilia na identificação de eventuais inconsistências nas predições.

Os cuidados tomados no processo de validação da rede neural refletem na confiabilidade nas predições. Além do *overfitting*, deve-se evitar também o *underfitting*, que ocorre quando o metamodelo não se adapta bem sequer aos dados com os quais foi treinado (FAN et al., 2017; MARKOVIC & TREECK, 2018; SINGARAVEL et al., 2018;

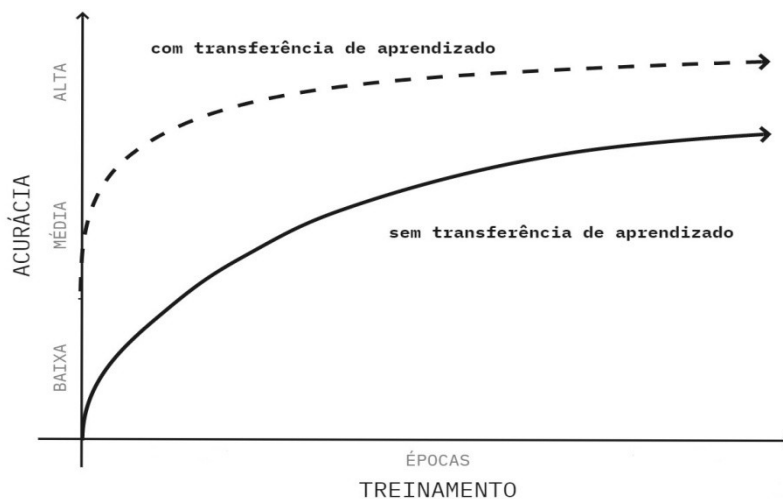
WERNCKE et al., 2021). Também é recomendado evitar tamanhos de *batch* maiores que 32, pois tendem a diminuir a capacidade de generalização do modelo (KESKAR et al., 2016).

Normalmente, o erro do conjunto de dados de treinamento é inferior ao erro do conjunto de dados de avaliação. O *underfitting* é caracterizado por erros de treinamento e avaliação altos. O *overfitting* é caracterizado por um erro de treinamento baixo e um erro de avaliação alto. Um ponto ótimo é um erro de validação baixo, ligeiramente acima do erro de treinamento.

Entretanto, também é possível que o erro de validação seja inferior ao erro de treinamento. Se o metamodelo é melhor em prever o desconhecido ao invés do conhecido, pode haver problemas na separação dos dados. Esta situação pode ocorrer quando o conjunto de treinamento recebe casos “mais difíceis” de aprender, ou se o conjunto de validação recebe muitos casos “fáceis” de prever. A adição de uma camada de regularização *dropout* também pode causar esta exceção, pois atua sobre o conjunto de dados de treinamento, mas é inerte no conjunto de dados de validação.

A técnica de transferência de aprendizado possibilita repassar o conhecimento adquirido de uma rede neural para outra. A transferência de aprendizado beneficia o treinamento em três frentes: desempenho inicial mais alto, uma taxa mais acentuada de melhora do desempenho, e um nível mais elevado para a acurácia (YOSINSKI et al., 2014; SALAKEN et al., 2019), como exposto na Figura 2.3.

Figura 2.3. Representação dos benefícios da transferência de aprendizado.



Fonte: adaptado de Torrey et al. (2009).

A técnica de transferência de aprendizado permite exportar o aprendizado entre as redes neurais artificiais e contribui para ganhos de eficiência computacional no processo. A transferência de aprendizado possibilita salvar os conhecimentos adquiridos no treinamento até a iteração ($i+0$) e utilizá-los como ponto de partida no treinamento da rede neural da iteração ($i+1$) subsequente (SINGARAVEL et al., 2018; BOURDEAU et al., 2019).

2.3.3. Validação

A confiabilidade dos metamodelos desenvolvidos deve ser sempre verificada através de rigorosos testes de validação (YONG et al., 2019). A validação de metamodelos consiste na avaliação dos erros entre as previsões e os dados observados. O MAE (“*mean absolute error*”) e o RMSE (“*rooted mean squared error*”) estão entre os dois indicadores mais utilizados para avaliar os erros entre os conjuntos de dados.

O MAE basicamente mede a acurácia de variáveis contínuas. O erro médio absoluto é utilizado como um indicador da magnitude média dos erros entre dois conjuntos de dados, entretanto, o MAE não considera a direção dos erros (positivos ou negativos). O MAE é calculado como sendo a média do valor absoluto das diferenças entre os dados previstos e os observados, conforme Equação 2 (WILLMOT & MATSUURA, 2005).

$$MAE = \frac{\sum_{i=1}^n |y_{pred,i} - y_i|}{n} \quad (2)$$

Onde, y_i é a observação;

$y_{pred,i}$ é a previsão;

n é o número total de termos da amostra;

i vai de 1 até n .

O RMSE é a raiz do erro quadrático médio da diferença entre a previsão e a observação. O RMSE é uma interessante medida de erro entre dois conjuntos de dados, pois explicita e penaliza a incidência de erros grandes. O RMSE é calculado na mesma unidade e escala dos valores da variável que está sendo avaliada, conforme Equação 3.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

Onde, y_i é a observação;

\hat{y}_i é a previsão;

n é o número total de termos;

i vai de 1 até n .

O RMSE também está entre os critérios de convergência mais utilizados para o treinamento de metamodelos. Um dos principais motivos da utilização como critério de convergência é a característica do RMSE de penalizar grandes erros, mesmo quando estes erros são relativamente poucos em quantidade. Essa característica torna o RMSE bastante recomendado em metamodelos de aplicações que não toleram grandes erros. Em contrapartida, no cálculo do MAE, os grandes erros recebem uma importância relativamente baixa, pois o MAE é um indicador linear, e pondera todas as diferenças individuais igualmente na média.

Para julgar se um valor calculado de RMSE é alto ou baixo, é necessário normalizá-lo. A normalização do RMSE facilita a comparação dos erros obtidos em metamodelos ou conjunto de dados que possuam diferentes unidades ou escalas.

Entretanto, o NRMSE (“*normalized rooted mean squared error*”) pode ser obtido de algumas formas: dividindo o RMSE pelo valor médio das observações, pela diferença entre o valor máximo e o valor mínimo das observações, pelo desvio padrão, ou pelo intervalo interquartil (Q1-Q3) das observações.

Tanto a normalização pela média, quanto pela diferença entre os valores máximos e mínimos são consistentes. A escolha da normalização pelo intervalo interquartil é especialmente interessante no caso de certas observações apresentarem valores espúrios, pois torna a normalização menos suscetível aos valores espúrios (HYNDMAN & KOEHLER, 2006).

Entretanto, segundo Roman et al. (2020), somente 2%, dos estudos que desenvolveram redes neurais artificiais, adotaram o NRMSE. A maioria permanece com o coeficiente de determinação (R^2), que não é ideal por não penalizar os grandes erros, ou o RMSE, que dificulta a comparação entre os estudos, pois não é normalizado.

Em suma, o treinamento de um metamodelo baseado em rede neural artificial requer uma função de treinamento e de um critério de parada para interromper o treinamento quando certos critérios são atingidos. Assim que o treinamento é concluído, o metamodelo é testado em um conjunto independente de dados, denominado conjunto de teste, para avaliar sua acurácia em novos casos.

2.3.4. Reflexão da seção e considerações finais da revisão

Os metamodelos desenvolvidos a partir de simulações energéticas em edificações apresentam limitações na consideração efetiva de múltiplos climas. A maioria dos metamodelos ainda é elaborada a partir de simulações executadas em uma localidade específica. Por outro lado, os estudos que utilizam múltiplas localidades tendem a executar poucas alterações na edificação, e/ou a descrever em demasia os climas.

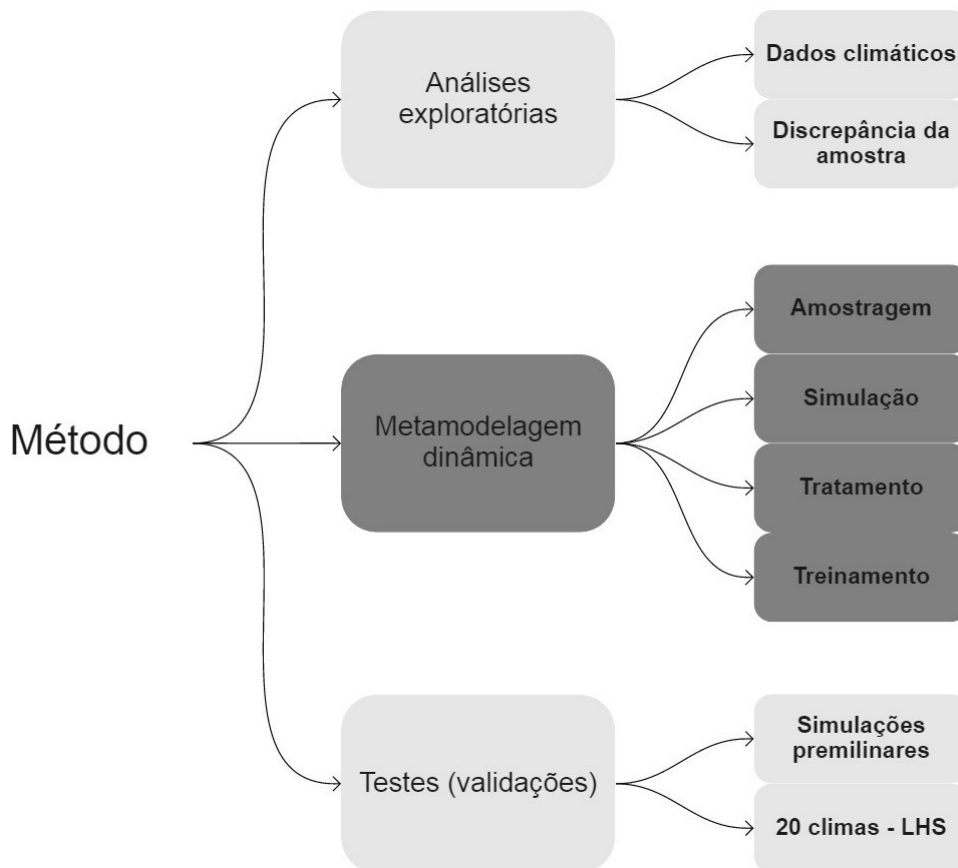
A análise dos estudos revisados expõe potenciais adicionais de melhorias: na eficiência da amostragem, na diversidade climática considerada, na inclusão do entorno da edificação, na preparação dos dados, e no desenvolvimento das redes neurais artificiais. A combinação da amostragem adaptativa com a característica iterativa das redes neurais artificiais abre a possibilidade de desenvolver metamodelos de forma dinâmica.

A inclusão de múltiplos climas na metamodelagem tende a se tornar mais comum, à medida que os metamodelos aplicados à eficiência energética em edificações necessitam cobrir diversas realidades climáticas. Os métodos adotados nesta pesquisa consistiram em formas de atacar as lacunas mapeadas pela revisão.

3. MÉTODO

O método descreve os procedimentos da pesquisa e está organizado em análises exploratórias e metamodelagem dinâmica, como apresentado pela Figura 3.1.

Figura 3.1. Organização do método.



Fonte: elaborada pelo autor (2020).

3.1. ANÁLISES EXPLORATÓRIAS

As análises exploratórias permitiram explorar:

- a) A relação entre os dados climáticos e a carga térmica de refrigeração, em diferentes granularidades;
- b) A discrepância da amostra, de acordo com o número de pontos considerados e a técnica de amostragem utilizada.

3.1.1. Dados climáticos

Para iniciar as análises exploratórias, coletaram-se os arquivos climáticos de 411 diferentes localidades, elaboradas pelo Instituto Nacional de Meteorologia - INMET (2019). Os arquivos climáticos forneceram os valores mensais de temperatura média do solo ($^{\circ}\text{C}$), e os valores horários para as seguintes variáveis climáticas: temperatura de bulbo seco ($^{\circ}\text{C}$); temperatura do ponto de orvalho ($^{\circ}\text{C}$); umidade relativa (%); pressão atmosférica (Pa); intensidade de radiação horizontal de onda longa (Wh/m^2); radiação horizontal global (Wh/m^2); radiação normal direta (Wh/m^2); radiação horizontal difusa (Wh/m^2); direção do vento ($^{\circ}$); e velocidade do vento (m/s).

A relação entre os dados climáticos e as cargas térmicas de refrigeração foi explorada através de uma bateria de simulações preliminares, modeladas a partir de uma pequena edificação de escritório, condicionada artificialmente. Os parâmetros da edificação de escritório foram definidos segundo os valores sintetizados na Tabela 3.1.

Tabela 3.1. Parâmetros da edificação das simulações preliminares.

Parâmetro	Valor
Comprimento da fachada (m)	15
Número de pavimentos	3
Pé direito (m)	3
Densidade de pessoas (W/m^2)	12
Densidade de potência de iluminação (W/m^2)	10
Densidade de potência de equipamentos (W/m^2)	10
Horas de ocupação (h)	8
Infiltração (trocas de ar/h)	0,3
Taxa de renovação do ar mínima ($\text{m}^3/\text{hora}/\text{pessoa}$)	27
Absortância solar da cobertura	0,8
Absortância solar da parede	0,5
Capacidade térmica da cobertura ($\text{kJ}/\text{m}^2.\text{K}$)	200
Capacidade térmica da parede ($\text{kJ}/\text{m}^2.\text{K}$)	180
Transmitância térmica da cobertura ($\text{W}/\text{m}^2.\text{K}$)	2,2
Transmitância térmica da parede ($\text{W}/\text{m}^2.\text{K}$)	2,3
Transmitância térmica do vidro ($\text{W}/\text{m}^2.\text{K}$)	5,7

Porcentagem de área de vidro na fachada	0,5
Fator solar do vidro	0,82
Temperatura de controle (°C)	24

Fonte: elaborada pelo autor (2020).

A escolha de uma edificação relativamente pequena foi motivada pela maior exposição a fatores externos à edificação, de modo a evidenciar as correlações entre as variáveis climáticas e as cargas térmicas de refrigeração da edificação, para zonas térmicas térreas, intermediárias e coberturas.

A diversidade de dados climáticos foi analisada por gráficos de dispersão e gráficos de densidade de frequência, em função da granularidade, a fim de avaliar a descrição dos dados climáticos a ser adotada na metamodelagem subsequente. Ao final da análise, foi consolidada uma base de dados climáticos na granularidade escolhida.

3.1.2. Preenchimento do espaço amostral

A segunda análise exploratória teve como objetivo avaliar a eficiência de preenchimento do espaço amostral (“*space filling sampling*”) de seis técnicas de amostragem compatíveis com amostragens adaptativas. As técnicas de amostragem foram escolhidas de acordo com a revisão bibliográfica, e foram implementadas através de códigos em linguagem de programação *R*, e das bibliotecas *lhs* (CARNELL, 2020) e *DiceDesign* (FRANCO et al., 2020).

Dentre as técnicas de amostragem analisadas, três são variações da amostragem por hipercubo latino (LHS) e outras três são sequências quasi-randômicas de baixa discrepância. As variações da amostragem por hipercubo latino (LHS) foram: a randômica (*Random LHS*), a melhorada (*Improved LHS*) e a otimizada (*Optimized LHS*). Dentre as sequências quasi-randômicas, foram analisadas as sequências de Halton, Sobol e Kronecker (denominada Torus).

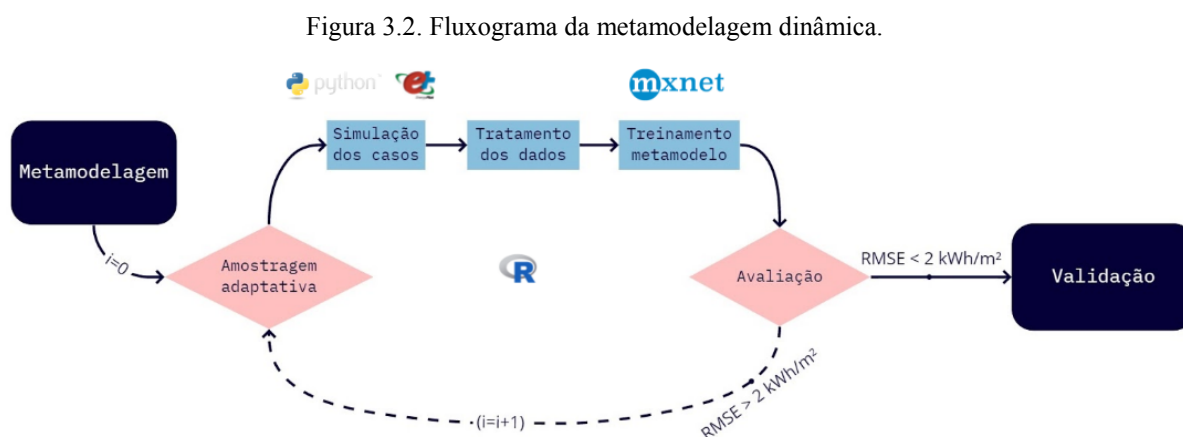
O tamanho da amostra de simulações influencia a acurácia dos metamodelos, pois impacta a representação do espaço amostral. De modo geral, tamanhos de amostra maiores favorecem a acurácia dos metamodelos. Contudo, a partir de certo tamanho de amostra, a amostragem de novos casos tende a beneficiar marginalmente o metamodelo, indicando que o espaço amostral foi preenchido satisfatoriamente.

A eficiência de preenchimento de espaço pode ser explorada, anteriormente à qualquer simulação, através do cálculo da discrepância da amostra, para um espaço amostral

com número de dimensões igual ao número de variáveis do experimento. Devido ao caráter adaptativo da amostragem, foi essencial analisar a discrepância, em cenários crescentes de tamanho de amostra, para cada técnica de amostragem. Os valores de discrepância, obtidos em diferentes tamanhos de amostra, foram representados em curvas de eficiência de preenchimento de espaço amostral. A metamodelagem, realizada na sequência, adotou a técnica de amostragem que apresentou a melhor curva de eficiência de preenchimento de espaço.

3.2. METAMODELAGEM DINÂMICA

A metamodelagem foi desenhada de modo a transferir progressivamente o conhecimento contido nas simulações energéticas para o metamodelo. Os procedimentos da metamodelagem ocorreram de forma iterativa até alcançar a acurácia desejada. O fluxograma, apresentado pela Figura 3.2, expõe as conexões entre as etapas da metamodelagem dinâmica.



Fonte: elaborada pelo autor (2020).

A amostragem adaptativa marca o início do ciclo descrito pelo fluxograma, e estabelece as combinações a serem simuladas em cada iteração. A simulação dos casos compõe a segunda etapa, seguida pelo tratamento dos dados. O treinamento do metamodelo ocorre até que um dos mecanismos de parada do treinamento seja acionado. Caso seja alcançado um RMSE inferior a 2 kWh/m² (acurácia desejada), a metamodelagem é finalizada e se dá início às baterias de validação. As etapas da metamodelagem são detalhadas nas subseções abaixo.

3.2.1. Simulação

Apesar de não ser a primeira etapa da metamodelagem dinâmica, os procedimentos de simulação são descritos por primeiro no método, a fim de facilitar a leitura da etapa de amostragem. Para modelar e executar as simulações, foi utilizado o programa de simulação computacional *EnergyPlus*, na versão 8.9. A modelagem das simulações foi executada por códigos elaborados em linguagem de programação *Python* (Apêndice A).

A utilização da biblioteca *pyIDF* permitiu configurar as classes, os objetos e os valores inseridos nos arquivos de simulação. O arquivo semente (“*seed.idf*”), foi utilizado como suporte para as modificações, pois contém as rotinas e os elementos construtivos que podem ser utilizados nas simulações. As simulações ocorrem em baterias de, no mínimo, dez arquivos IDF por arquivo climático, devido às condições implementadas no código elaborado e ao número de núcleos de processamento da máquina em que as simulações foram executadas. Os códigos desenvolvidos possibilitaram a criação de modelos com variações na geometria, nas cargas internas, na envoltória, no entorno, e na temperatura de controle do sistema de condicionamento de ar da edificação. As variações sobre os parâmetros citados estão detalhadas abaixo.

3.2.1.1. Geometria

Os parâmetros relacionados à geometria afetam diretamente o fator de forma, ou seja, a razão entre a área de envoltória e o volume da edificação, e por isso abrangem um intervalo considerável de variação. A exceção foi o número de pavimentos que não foi majorado além de três, para evitar desequilíbrios na proporção de dados gerados para as exposições das zonas térmicas (térreas, intermediárias e coberturas). A geometria dos modelos de escritórios adotados é simétrica nos eixos horizontais, ou seja, as fachadas opostas apresentam os mesmos comprimentos.

3.2.1.2. Envoltória

Destaca-se que os limites estabelecidos para os parâmetros da envoltória podem se estender além dos limites encontrados no mercado. A inclusão de valores que ultrapassam os valores mínimos e máximos encontrados na realidade, mas aceitos pela simulação, é uma medida para reforçar a incidência de casos nas “bordas” do espaço amostral para estas variáveis. Os elementos de sombreamento foram considerados em formato de brises horizontais, posicionados na altura do pé direito dos pavimentos das edificações simuladas. As aberturas da edificação foram modeladas como superfícies translúcidas de vidro, contidas

em superfícies opacas (paredes). Foram consideradas 9 construções de paredes, visando representar as combinações da capacidade térmica (baixa, média e alta) com a transmitância térmica (baixa, média e alta). O mesmo procedimento é válido para as coberturas. Além do vidro comum de 6 mm, foram adicionados vidros de controle solar nas simulações.

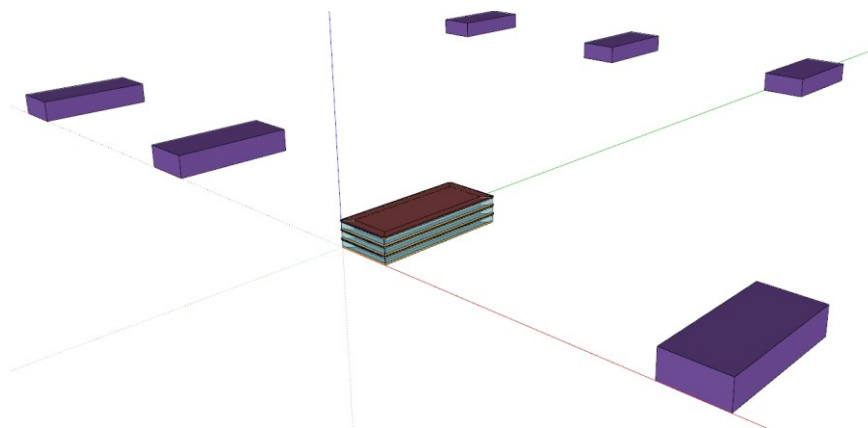
3.2.1.3. Cargas internas

Os ganhos internos provenientes de pessoas, iluminação e equipamentos são relevantes no balanço térmico do ar dos ambientes da edificação. Os padrões de ocupação possíveis foram definidos em 8,10,12,14,16 (horas/dia). A atividade metabólica adotada foi de 120 W/pessoa, valor típico para atividades de escritório. Os ganhos de calor radiante referentes às pessoas foram fixados em 30%. A fração radiante considerada para os equipamentos foi de 50% e para a iluminação foi de 70%. Para gerar resultados abrangentes de carga térmica de refrigeração, foram definidas faixas amplas para os parâmetros de cargas internas: a densidade de pessoas varia entre 0 e 0,5 (pessoa/m²); a densidade de potência de iluminação oscila entre 0 e 20 (W/m²); e a densidade de potência de equipamentos entre 0 e 20 (W/m²).

3.2.1.4. Entorno

O entorno da edificação foi modelado de forma a representar uma grade urbana de edificações. A densidade horizontal da grade urbana define quão próximas estão as edificações da grade urbana, e pode variar consideravelmente entre as simulações. Valores baixos e valores elevados de densidade horizontal da grade urbana foram representados nas Figura 3.3 e a Figura 3.4, respectivamente.

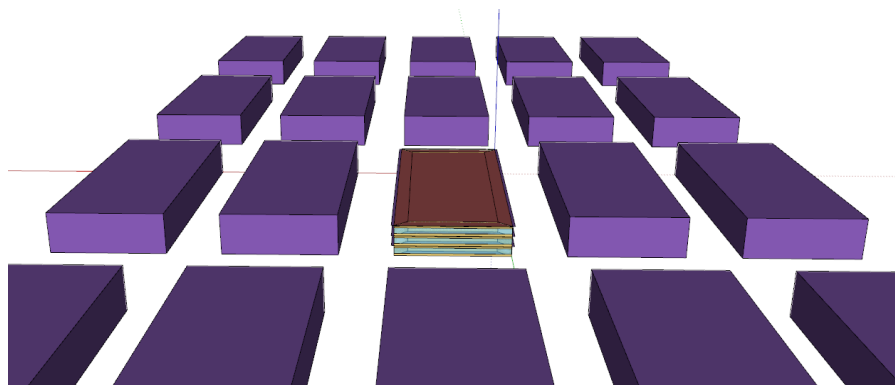
Figura 3.3. Baixa densidade horizontal da grade urbana (valor = 0,03).



Fonte: elaborada pelo autor (2020).

A edificação simulada situa-se sempre ao centro da grade urbana. Quando a densidade horizontal da grade urbana é inferior a 0,05, a mesma emula edificações distantes. Quando a densidade horizontal da grade urbana é igual a 0,25, o espaçamento entre as edificações é exatamente igual ao comprimento da fachada da edificação. Quando acima de 0,45, a densidade horizontal da grade urbana permite emular edificações vizinhas muito próximas.

Figura 3.4. Elevada densidade horizontal da grade urbana (valor = 0,46).



Fonte: elaborada pelo autor (2020).

O número de pavimentos das edificações do entorno variou de 1 até 4, novamente para evitar desequilíbrios na proporção de condições geradas para as zonas térmicas térreas, intermediárias e coberturas. Caso fossem adotados valores muito superiores para os pavimentos das edificações do entorno, uma parcela relevante das zonas térreas estaria sempre obstruída. O fator mais importante é a combinação resultante do número de pavimentos do entorno com a densidade horizontal da grade urbana.

As edificações do entorno foram modeladas como componentes de sombreamento, que interagem ativamente com a edificação simulada. Contudo, as superfícies de sombreamento modeladas no *EnergyPlus* consideram somente os fenômenos de reflexão de onda curta. A análise dos algoritmos de cálculo do *EnergyPlus* confirmou que a temperatura da superfície modelada como sombreamento é igual à temperatura do solo e à temperatura do ar, caracterizando uma limitação do programa (*United States Department of Energy, 2018*).

3.2.1.5. Sistema

O condicionamento de ar dos ambientes é realizado a partir de sistemas de condicionamento ideal (*Ideal Loads*), que suprem as necessidades de refrigeração de todas as zonas térmicas da edificação. A carga térmica de refrigeração representa a quantidade de energia térmica integrada em um determinado intervalo de tempo, que deve ser retirada do ar

da zona térmica para manter a temperatura do ar abaixo do valor de ajuste de temperatura (*setpoint*).

Optou-se por não modelar o controle de desumidificação ao condicionar os ambientes, ou seja, a umidade do ambiente não atuou no acionamento dos sistemas de carga ideal. Portanto, os sistemas de carga ideal atuam somente no controle da temperatura, como é usual nas edificações de escritórios no Brasil. Entretanto, o condicionamento do ar pode promover a desumidificação do ar, por ser uma característica inerente ao processo de climatização. As variações por parâmetro da edificação estão detalhadas na Tabela 3.2.

Tabela 3.2. Variações por parâmetro da edificação.

Classe	Parâmetro	Unidade	Tipo de variável	Faixa de variação	Categorias
Geometria	Comprimento da fachada	(m)	Numérica	14 a 50	-
	Proporção entre maior e menor fachada	-	Numérica	1 a 2	-
	Pé direito	(m)	Numérica	2,4 a 4,8	-
	Número de pavimentos	-	Numérica	1 a 3	-
Envoltória	Tipos de cobertura	-	Categórica	-	9
	Absortância solar da cobertura	-	Numérica	0,10 a 0,99	-
	Tipos de parede	-	Categórica	-	9
	Absortância solar da parede	-	Numérica	0,10 a 0,99	-
	Porcentagem de área de vidro na fachada	-	Numérica	0,01 a 0,99	-
	Tipos de vidro	-	Categórica	-	2
	Comprimento do brise horizontal	(m)	Numérica	0 a 2	-
Cargas Internas	Padrão de ocupação	(h/dia)	Categórica	-	5
	Densidade de pessoas	(pessoa/m ²)	Numérica	0 a 0,5	-
	Densidade de potência de iluminação	(W/m ²)	Numérica	0 a 20	-
	Densidade de potência de equipamentos	(W/m ²)	Numérica	0 a 20	-
Entorno	Densidade horizontal da grade urbana	-	Numérica	0,01 a 0,49	-
	Número de pavimentos	-	Numérica	1 a 4	-
	Absortância solar do entorno opaco	-	Numérica	0,01 a 0,90	-
	Porcentagem de área de vidro na fachada	-	Numérica	0,01 a 0,99	-
Sistema	Temperatura de controle	(°C)	Numérica	17 a 27	-

Fonte: elaborada pelo autor (2020).

Para as variáveis numéricas, as variações dos parâmetros ocorreram dentro de uma faixa com limites inferiores e superiores. A faixa ampla de valores possíveis para o termostato, assim como em outras variáveis, aumenta a aplicabilidade do metamodelo em situações reais. No caso de variáveis categóricas, as variações ocorreram de acordo com o número de categorias possíveis. Apesar das variações descritas anteriormente, determinadas premissas são válidas para todas as simulações e estão apresentadas na Tabela 3.3.

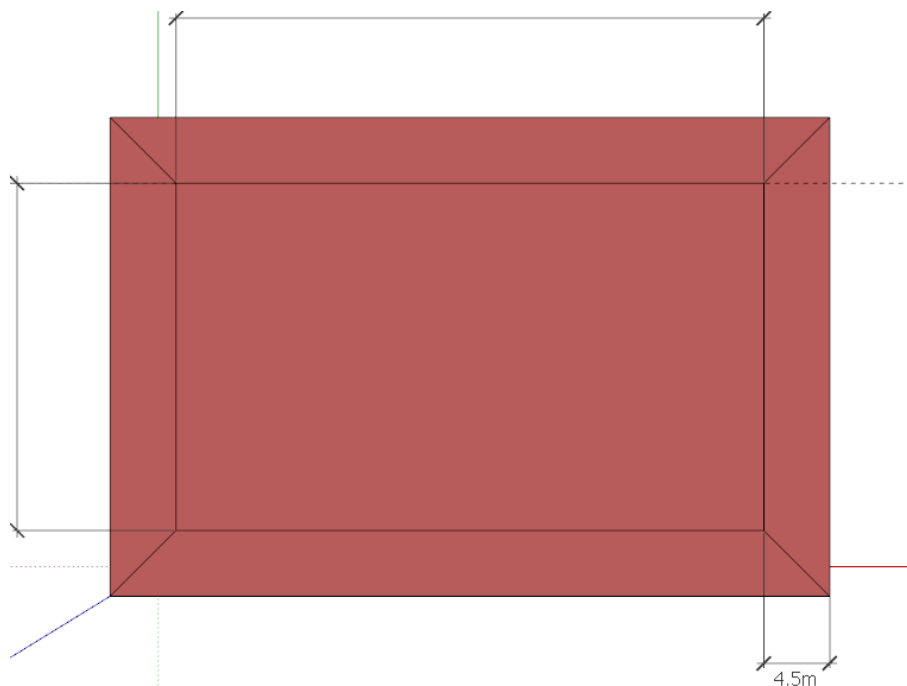
Tabela 3.3. Premissas válidas para todas as simulações.

Parâmetro	Definição
Período da simulação	Annual (365 dias)
Distribuição solar	Interna e externa com reflexões
Cálculo de sombreamento	<i>AverageOverDaysInFrequency</i>
Algoritmo de convecção em superfícies internas	TARP
Algoritmo de convecção em superfícies externas	DOE-2
Algoritmo do balanço de calor nos componentes	<i>ConductionTransferFunction</i>
Algoritmo do balanço de calor no ar da zona	<i>ThirdOrderBackwardDifference</i>
Valor de tolerância para convergência da carga	0,04
Valor de tolerância para convergência da temperatura	0,4 °C
Intervalo de cálculo base (<i>timestep</i>)	6/hora
Profundidade das zonas perimetrais	4,5 m
Infiltração (trocas de ar/h)	0,3
Taxa de renovação do ar mínima (m ³ /hora/pessoa)	27
Zoneamento da edificação	Zonas internas e perimetrais

Fonte: elaborada pelo autor (2020).

A definição destas premissas foi realizada de acordo com as necessidades do experimento e as documentações do *EnergyPlus*. O período das simulações compreendeu os 365 dias do ano. A distribuição solar foi definida como interna e externa com reflexões, para considerar a incidência de raios solares em cada superfície das zonas térmicas, assim como o efeito de eventuais superfícies externas e de dispositivos de sombreamento. A profundidade das zonas perimetrais das edificações foi mantida fixa em 4,50 metros, independentemente dos comprimentos das fachadas da edificação, como ilustrado pela Figura 3.5.

Figura 3.5. Zoneamento térmico da edificação com uma zona interna e quatro zonas periféricas.



Fonte: elaborada pelo autor (2020).

Os algoritmos de convecção e de balanço de calor foram definidos de forma a otimizar os cálculos e a duração das simulações. O intervalo de cálculo base foi de 6 vezes por hora. Especificou-se uma taxa de infiltração constante de 0,3 trocas de ar por hora. A infiltração é o fluxo de ar indesejado proveniente do ambiente externo, geralmente causado por frestas em portas e janelas ou pela operação de abertura e fechamento de portas de acesso ao ambiente. A taxa de renovação do ar mínima para os ambientes foi definida em 27 m³/hora/pessoa, respeitando as normas vigentes (ANVISA, Resolução N^o9 de 16 de Janeiro de 2003).

Para modelar e executar as simulações de cada iteração, os valores dos parâmetros da edificação tiveram de ser definidos a partir da etapa de amostragem.

3.2.2. Amostragem adaptativa

A amostragem adaptativa é realizada através da técnica de amostragem que apresentou os menores valores de discrepância na análise exploratória. A amostragem inicia sempre com os valores brutos, que variam no intervalo entre 0 e 1.

Para serem adaptados à realidade das simulações, os valores brutos foram convertidos em valores compatíveis com as variáveis numéricas ou categóricas, de acordo com a necessidade de modelagem da edificação.

Para as variáveis numéricas, a conversão ocorreu dentro do intervalo de valores estabelecidos para cada variável. A lógica de conversão de valores brutos em valores de variáveis numéricas foi exposta na Equação 4.

$$Vn = Lif + (Lsf - Lif) * x \quad (4)$$

Onde:

Vn é o valor da variável numérica;

x é valor bruto, que varia entre 0 e 1;

Lif é igual ao limite inferior da faixa de valores para a variável numérica;

Lsf é igual ao limite superior da faixa de valores para a variável numérica.

A amostragem dos arquivos climáticos (EPW) seguiu o mesmo raciocínio aplicado em variáveis numéricas. Entretanto, a variável climática que orientou a amostragem dos arquivos climáticos foi definida somente após a análise exploratória sobre os dados climáticos.

Para as variáveis categóricas da edificação, utilizou-se uma lógica similar àquela das variáveis numéricas. Para as variáveis categóricas, o intervalo de 0 a 1 é dividido em faixas equidistantes, em função do número de categorias da variável. A conversão do valor bruto em categoria é realizada de acordo com a faixa que o valor bruto incidiu.

A conversão dos valores, exposta acima, habilita a modelagem da edificação no respectivo arquivo IDF e a seleção do arquivo climático. A cada iteração, novos pontos são adicionados, complementando as amostras anteriores. Para garantir amostragens reproduzíveis, foram utilizadas sementes de randomização (*random seeds*) durante o todo o processo da amostragem adaptativa.

3.2.3 Tratamento de dados

Após o término das simulações da respectiva iteração, os resultados são coletados e organizados em valores de carga térmica de refrigeração, por zona térmica da edificação. Os lotes de dados de saída das simulações são importados para a base de dados, e combinados

aos respectivos dados de entrada. O tratamento dos dados envolveu a criação de variáveis que descrevem as zonas térmicas, e possibilita fornecer valores em formatos mais adequados.

O tratamento dos dados focou na melhoria da descrição de zonas térmicas da edificação e dos climas simulados. A escolha da granularidade temporal considerou a análise exploratória e sugestões de trabalhos revisados. O tratamento de dados também foi responsável por transformar todas as variáveis categóricas das simulações em numéricas. Buscou-se minimizar a quantidade de variáveis necessárias para descrever as combinações na base de dados.

Em algumas situações, o tratamento de dados pode criar variáveis preditoras que possuem apenas um único valor, ou valores exclusivos que ocorrem com frequências muito baixas. Logo, também foram realizados testes para identificar variáveis preditoras com variância zero ou próxima de zero. A presença de variáveis preditoras com variância nula pode se transformar em fator causador de ruído, ao dificultar a convergência do treinamento e prejudicar o desempenho da rede neural.

Os dados foram normalizados (centralização e escalonamento) antes de iniciar o treinamento das redes neurais artificiais, para aumentar a estabilidade numérica do metamodelo. Para centralizar os dados, o valor médio das variáveis foi subtraído de todos os valores, induzindo as variáveis preditoras a apresentarem uma média zero. No escalonamento, os valores das variáveis preditoras foram divididos pelos seus valores de desvio padrão. A normalização dos dados facilita o aprendizado e evita erros computacionais.

Foi implementada a codificação de efeito (“*effect coding*”) para descrever certas condições de zonas térmicas, como a orientação e a exposição. As variáveis obtidas por codificação de efeito não necessitam ser normalizadas, uma vez que já estão configuradas em formatos “altamente interpretáveis” para a rede neural.

Foram realizadas análises de correlação entre as variáveis preditoras, como última etapa do tratamento de dados. Os testes de correlação tiveram a finalidade de evitar repetições de descrições, garantindo que as variáveis remanescentes fossem somente aquelas necessárias.

Após a aplicação dos tratamentos citados, as variáveis resultantes foram consolidadas em matriz bidimensional e consideradas adequadas para o treinamento das redes neurais artificiais. A estrutura matricial de dados é composta por colunas verticais que representam as variáveis preditoras e a variável dependente, e por linhas horizontais que descrevem as zonas térmicas. A cada iteração de simulações, novos dados são anexados à estrutura matricial. O processo de tratamento foi implementado totalmente em linguagem *R* (Apêndice A).

3.2.4. Treinamento e avaliação

A partir da base de dados tratados, os metamodelos foram treinados para estimar a carga térmica de refrigeração de escritórios em climas brasileiros. O treinamento do metamodelo focou na minimização do RMSE, dos conjuntos de dados de treinamento e de avaliação. Mesmo sendo um processo iterativo, o conjunto de treinamento sempre recebe 80% da base de dados, e os 20% restantes são alocados no conjunto de avaliação.

O *framework* de aprendizado de máquina *MXNet* foi utilizado para treinar e validar os metamodelos, por ser altamente escalável e eficiente, com forte potencial em computação de alto desempenho (HPC). Desde 2016, a Amazon adota e suporta o *MXNet* como seu *framework* de *deep learning*, e afirma que o *MXNET* é cerca de duas vezes mais veloz que o *TensorFlow* (*framework* da Google).

A função de ativação adotada para os neurônios da rede neural artificial foi a ReLU (*rectified linear unit*). O tamanho do lote (*batch size*) foi definido em 32, por ser um múltiplo de 2, e por respeitar as recomendações de ser menor ou igual a 32 (KESKAR et al., 2016). Em treinamentos preliminares, foram testados três otimizadores: "adam" (*Adam optimization method*), "sgd" (*Stochastic gradient descent*), e "adagrad" (*Adaptive Gradient Algorithm*). A arquitetura da rede neural foi testada para até cinco camadas ocultas, e até 512 neurônios por camada.

Os dois critérios adotados para a parada provisória do treinamento da rede neural: quando o RMSE do conjunto de avaliação aumentou (*bad steps*) por mais de 10 épocas seguidas; ou após 700 épocas. A combinação dos critérios de parada diminui o risco de *overfitting* e de *underfitting* durante o treinamento da rede neural artificial. Quando um dos critérios de parada é atingido, o treinamento é interrompido e os parâmetros da rede neural artificial são salvos, para dar início a uma nova iteração. Para executar a transferência de aprendizado, os parâmetros de treinamento salvos ao término da iteração, são importados para o treinamento da iteração subsequente.

A comparação de valores simulados e preditos permitiu avaliar a variação dos resultados de R^2 , RMSE e NRMSE para os conjuntos de treinamento e de avaliação. O processo de metamodelagem dinâmica é interrompido, por completo, somente quando o RMSE do conjunto de avaliação alcança valores abaixo de 2 kWh/m². Este valor foi definido de modo a manter o foco da análise no início da metamodelagem, quando o número de casos ainda é relativamente baixo. Após o RMSE do conjunto de avaliação alcançar este valor, procede-se para as baterias de teste do metamodelo.

3.3. TESTES (VALIDAÇÕES)

As baterias de teste foram executadas com a finalidade de validar a rede neural com casos nunca vistos. A primeira bateria de testes da rede neural artificial foi realizada para emular a utilização do metamodelo em uma situação real, na qual o usuário procura avaliar uma edificação de escritório como um todo.

Esta bateria de testes utilizou uma parcela do conjunto de dados das simulações preliminares, executadas durante as análises exploratórias. As predições do metamodelo foram validadas contra os resultados das simulações para os climas de Curitiba, São Paulo, Rio de Janeiro e Belém. Estas localidades foram selecionadas para compor a primeira bateria de testes por serem centros urbanos relevantes, expostos a condições climáticas bastante distintas em um contexto que considera somente localidades brasileiras.

Adicionalmente, com a finalidade de validar a rede neural artificial em múltiplas combinações de parâmetros, foi executada uma bateria de simulações independentes composta por 100 simulações amostradas a partir da técnica de LHS randômica. Estas 100 simulações foram distribuídas em 20 climas brasileiros.

Ao final do experimento, elaboraram-se curvas que representam a evolução das predições de redes neurais artificiais treinadas ao longo das iterações, testando-as com a base de dados consolidada ao final da última iteração.

Com a finalidade de verificar o impacto da técnica de amostragem sobre o processo de metamodelagem, foi efetuada a substituição da técnica de amostragem que obteve a menor discrepância na análise exploratória, por aquela que obteve os maiores valores de discrepância. Esta última comparação buscou avaliar a importância da análise prévia das técnicas de amostragem.

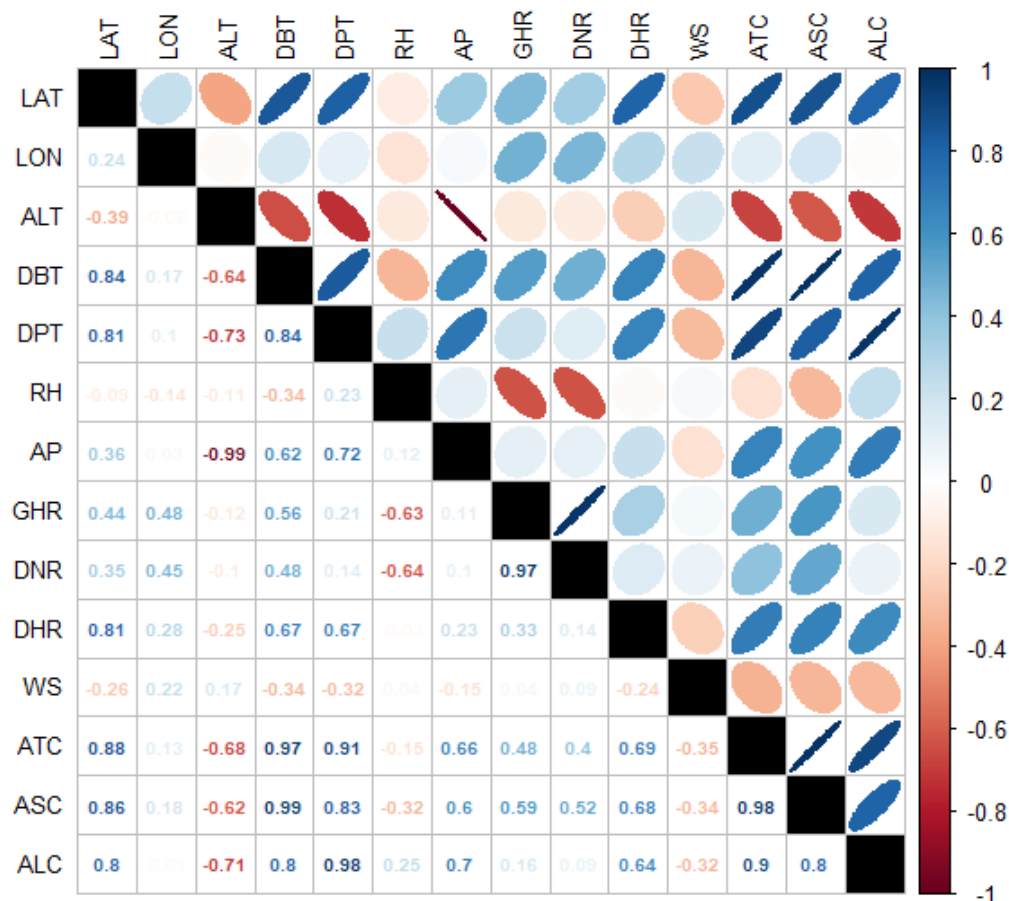
4. RESULTADOS

Este capítulo contém as análises e as reflexões da pesquisa. Os resultados foram apresentados, de acordo com os procedimentos metodológicos propostos.

4.1. DADOS CLIMÁTICOS

A análise exploratória dos dados climáticos buscou definir a descrição climática a ser adotada na metamodelagem posterior. Para analisar as relações entre as variáveis climáticas e as parcelas da carga térmica de refrigeração integrada anual, foi elaborada uma matriz de correlação, exposta na Figura 4.1. A matriz de correlação foi construída a partir de resultados de simulações preliminares distribuídas em 411 arquivos climáticos brasileiros.

Figura 4.1. Matriz de correlação entre variáveis climáticas e cargas térmicas anuais de refrigeração das simulações preliminares.



Fonte: elaborada pelo autor (2020).

A matriz de correlação apresenta as variáveis climáticas: a latitude (LAT), a longitude (LON) e a altitude (ALT) das localidades; seguidas pelas médias anuais de temperatura de

bulbo seco (DBT), temperatura do ponto de orvalho (DPT), umidade relativa (RH), pressão atmosférica (AP), radiação horizontal global (GHR), radiação normal direta (DNR), radiação horizontal difusa (DHR), e velocidade do vento (WS). A matriz também apresenta a carga térmica total de refrigeração anual (ATC), a carga térmica sensível de refrigeração anual (ASC) e a carga térmica latente de refrigeração anual (ALC).

A carga térmica total de refrigeração é resultante da soma da parcela sensível com a parcela latente. Complementarmente aos valores numéricos expostos na diagonal inferior da matriz de correlação, as altas correlações entre variáveis foram representadas por elipses alongadas, de cor azul no caso de correlação positiva, e de cor vermelha no caso de correlação negativa.

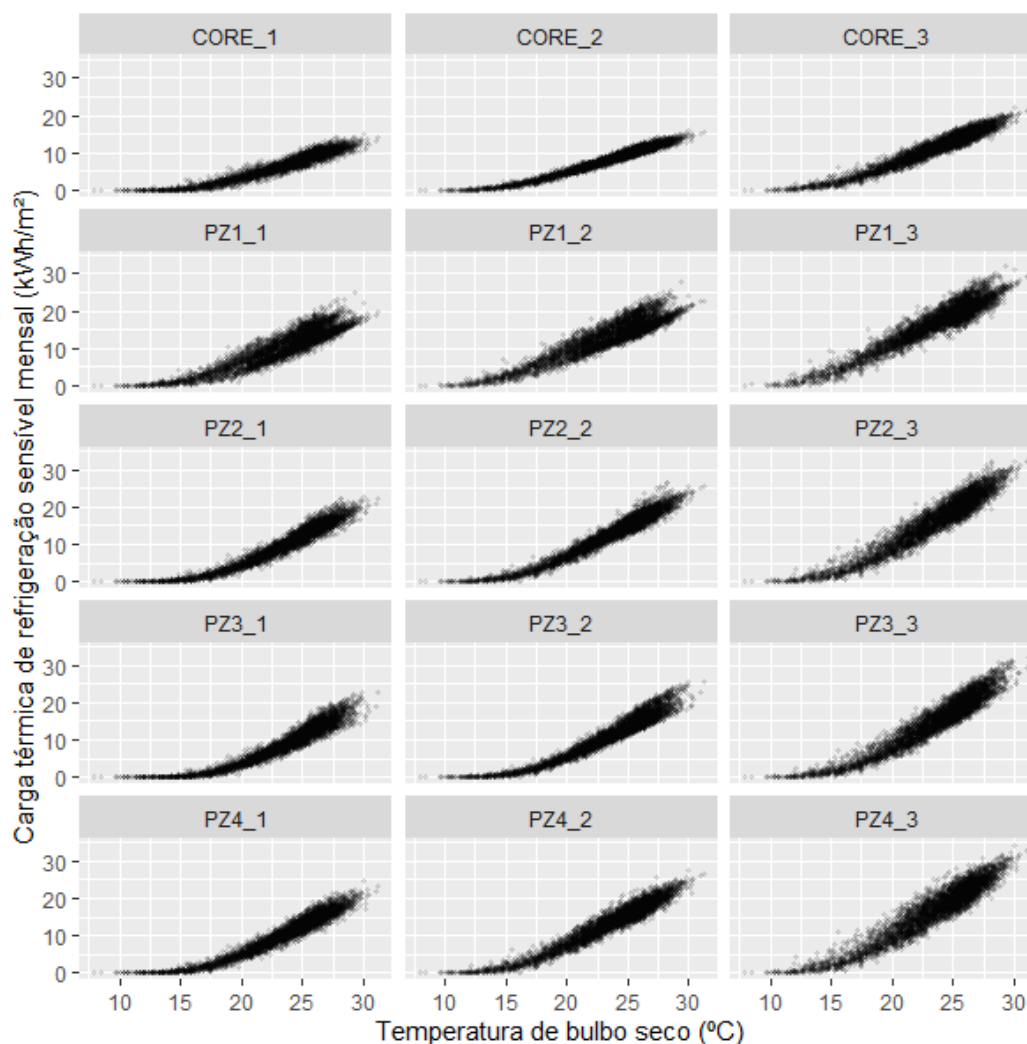
A maior correlação entre variáveis climáticas e cargas térmicas de refrigeração foi obtida entre a temperatura de bulbo seco e a carga térmica de refrigeração sensível. É importante lembrar que a temperatura do bulbo seco, a umidade relativa e a radiação horizontal global foram as variáveis climáticas mais adotadas pelas publicações analisadas na revisão bibliográfica.

A umidade relativa e a radiação horizontal global apresentaram correlações menos acentuadas com as cargas térmicas de resfriamento. Entretanto, na matriz apresentada, a correlação entre a temperatura do ponto de orvalho e a carga térmica latente foi muito alta. Esta correlação elevada levantou a possibilidade de substituição da umidade relativa pela temperatura do ponto de orvalho, na descrição climática a ser implementada na metamodelagem.

Apesar de apresentarem valores altos de correlação com as cargas térmicas de refrigeração, a latitude e a altitude não variam para uma mesma localidade, e por isso não foram consideradas como possibilidades. A temperatura de bulbo seco apresentou múltiplas correlações positivas com as outras variáveis climáticas, já a altitude apresentou o maior número de correlações negativas com as outras variáveis climáticas. Neste sentido, a verificação de correlações auxilia na escolha de variáveis, e evita repetições desnecessárias de dados.

Após analisar as correlações com os valores anuais, optou-se por adentrar nos resultados das simulações por zona térmica, em intervalo mensal, e por parcela de carga térmica de refrigeração (sensível e latente). A carga térmica de refrigeração sensível mensal de zonas internas e perimetrais, separada em zonas térmicas térreas, intermediárias e coberturas foi exposta na Figura 4.2 abaixo.

Figura 4.2. Gráfico de dispersão entre os valores mensais de temperatura de bulbo seco e de carga térmica de refrigeração sensível mensal, por zona térmica.



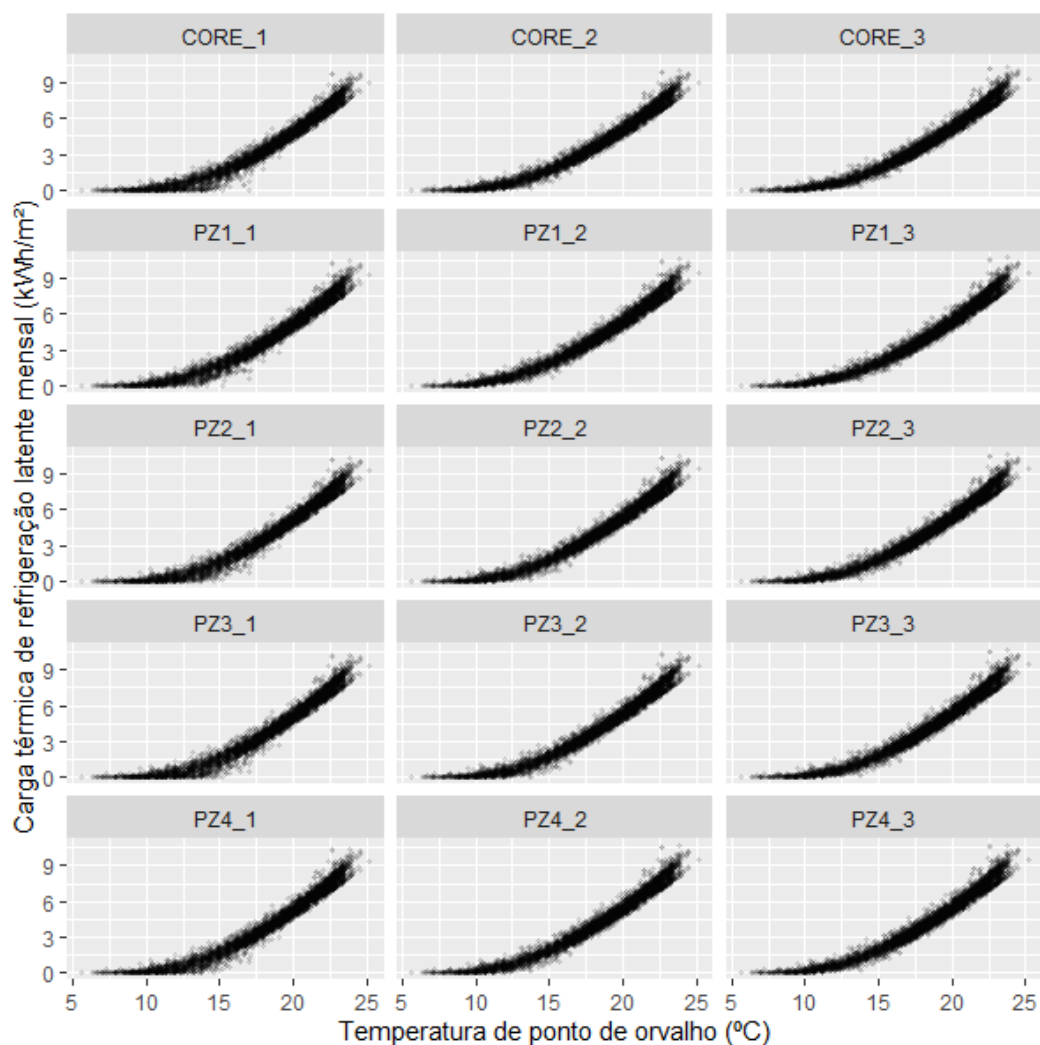
Fonte: elaborada pelo autor (2020).

A análise exploratória seguiu evidenciando a alta correlação entre a temperatura de bulbo seco e a carga térmica de refrigeração sensível. Nota-se a incidência de valores inferiores de carga térmica de refrigeração sensível nas zonas térmicas internas (início “CORE”), em relação às zonas perimetrais (início “PZ”). Com relação à orientação (Norte = PZ1, Leste = PZ2, Sul = PZ3, Oeste = PZ4) das zonas térmicas, nota-se que as zonas térmicas expostas ao norte apresentam uma faixa de valores mais ampla, impactadas pela radiação solar. As zonas de cobertura (coluna à direita) apresentaram os maiores valores de carga térmica de refrigeração sensível, por estarem consideravelmente mais expostas à radiação solar, seguidas pelas zonas intermediárias (coluna central). As zonas térreas (coluna à esquerda) apresentam as menores cargas térmicas de refrigeração, tanto por estarem em

contato com o solo, quanto por estarem menos expostas à radiação solar devido à grade urbana de edificações que representa o entorno.

Ao migrar a análise para a carga térmica de refrigeração latente mensal e a temperatura do ponto de orvalho, foram confirmadas tendências nítidas para todas as zonas térmicas da edificação, expostas na Figura 4.3.

Figura 4.3. Gráficos de dispersão entre os valores mensais de temperatura do ponto de orvalho e a carga térmica de refrigeração latente mensal, por zona térmica.



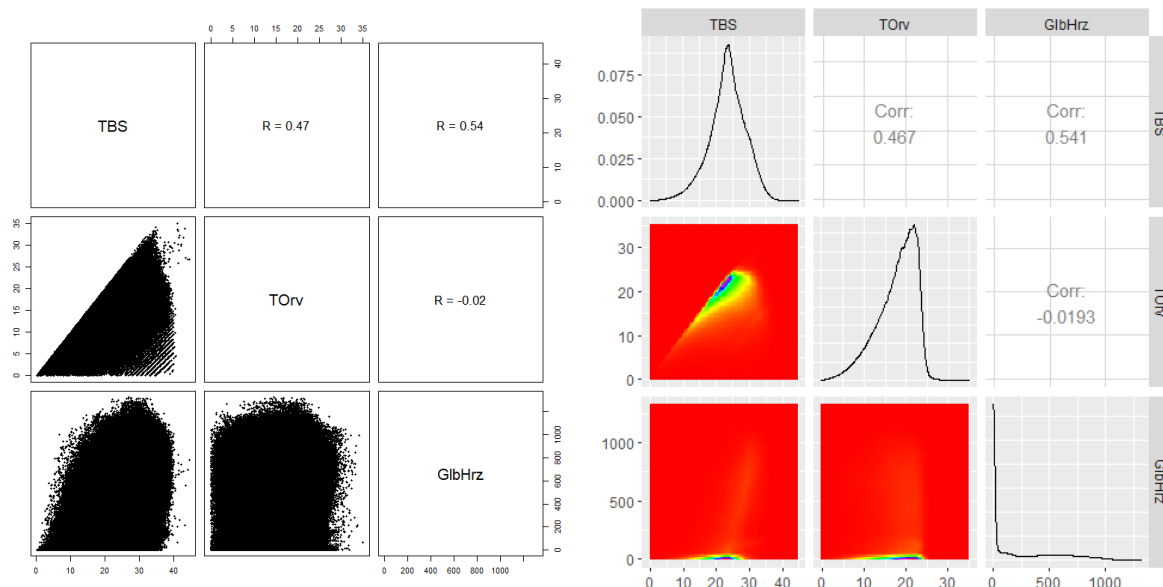
Fonte: elaborada pelo autor (2020).

Nota-se que a temperatura do ponto de orvalho descreve excepcionalmente bem a carga térmica latente. Desta forma, confirmou-se a substituição da umidade relativa pela temperatura do ponto de orvalho entre as variáveis selecionadas para descrever os climas brasileiros na metamodelagem posterior.

A radiação horizontal global não apresentou correlações com as parcelas de carga térmica na mesma intensidade que a temperatura de bulbo seco ou a temperatura do ponto de orvalho. Entretanto, a escolha da radiação horizontal global como terceira e última variável da descrição climática a ser adotada na metamodelagem foi motivada pela sua interação com todas as variáveis da envoltória. Somente para exemplificar, o valor de absorvância solar da parede ou o fator solar de vidro, sem um valor atribuível à radiação solar seria uma descrição incompleta. Por isso, a radiação horizontal global foi mantida na descrição para a metamodelagem realizada em um momento posterior.

Referente à granularidade dos dados climáticos, exploraram-se os valores da temperatura de bulbo seco, da temperatura do ponto de orvalho, e da radiação horizontal global, presentes nos 411 arquivos climáticos. Os dados de cada variável climática somaram cerca de 3,6 milhões de valores horários, e foram analisados em gráficos de dispersão na Figura 4.4. As representações de densidade de frequência foram adicionadas aos gráficos, devido à elevada sobreposição de pontos.

Figura 4.4. Incidência de valores horários das variáveis climáticas selecionadas.



Fonte: elaborada pelo autor (2020).

Apesar da sobreposição de pontos, os gráficos de dispersão à esquerda denotam a diversidade e delimitam os valores das combinações possíveis para os dados horários. À direita da Figura 4.4, observou-se que os dados horários de temperatura de bulbo seco e de temperatura do ponto de orvalho apresentaram maiores concentrações de pontos na região central dos intervalos.

A radiação horizontal global apresentou cerca de metade de seus valores próximos ou iguais a zero, devido ao período noturno. Contudo, a escolha de granularidades muito baixas, como a horária ou diária, refletem em uma quantidade crescente de dados a partir de uma única simulação, como apresentado na Tabela 4.1.

Tabela 4.1. Número de dados de saída obtidos a partir de uma simulação.

Granularidade		Espacial			
		Edificação	Zona térmica		
			<i>1 pavimento</i>	<i>2 pavimentos</i>	<i>3 pavimentos</i>
Temporal	<i>Ano</i>	1	5	10	15
	<i>Mês</i>	12	60	120	180
	<i>Semana</i>	52	260	520	780
	<i>Dia</i>	365	1825	3650	5475
	<i>Horária</i>	8760	43800	87600	131400

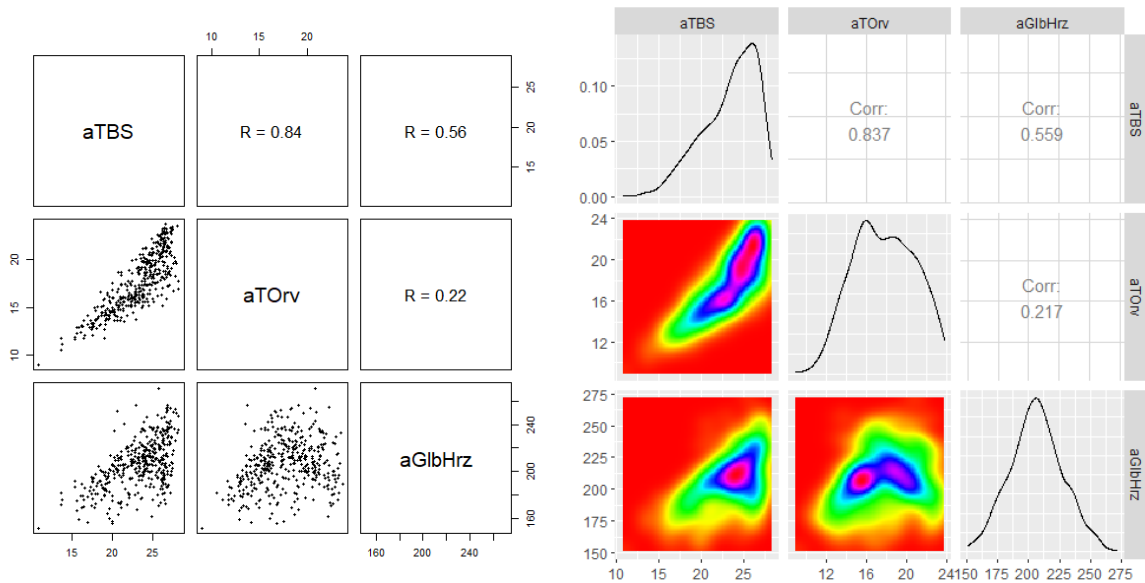
Fonte: elaborada pelo autor (2020).

Apesar da maior diversidade de dados para a variável dependente e para as variáveis climáticas, a adoção de granularidades temporais menores criaria a necessidade de estabelecer variáveis adicionais para caracterizar os horários de funcionamento da edificação, as datas de recesso, entre outras descrições. Estas descrições afetariam basicamente as variáveis relacionadas às cargas internas. As variáveis que descrevem a geometria não seriam beneficiadas no processo, pois a envoltória e o entorno da edificação não variam ao longo da simulação, independente da granularidade temporal adotada.

Ao considerar o objetivo da pesquisa, julgou-se mais adequado seguir adiante com as comparações envolvendo somente as granularidades anuais e mensais. Para isso, os dados climáticos horários foram consolidados em duas bases de dados de variáveis climáticas: a primeira com os valores médios anuais e a segunda com valores médios mensais de cada clima.

Por serem extensas, as bases de dados estão disponíveis através do repositório descrito no Apêndice A. Os gráficos de dispersão com os valores médios anuais e mensais, da temperatura de bulbo seco, da temperatura do ponto de orvalho, e da radiação horizontal global foram apresentadas na Figura 4.5 e na Figura 4.6.

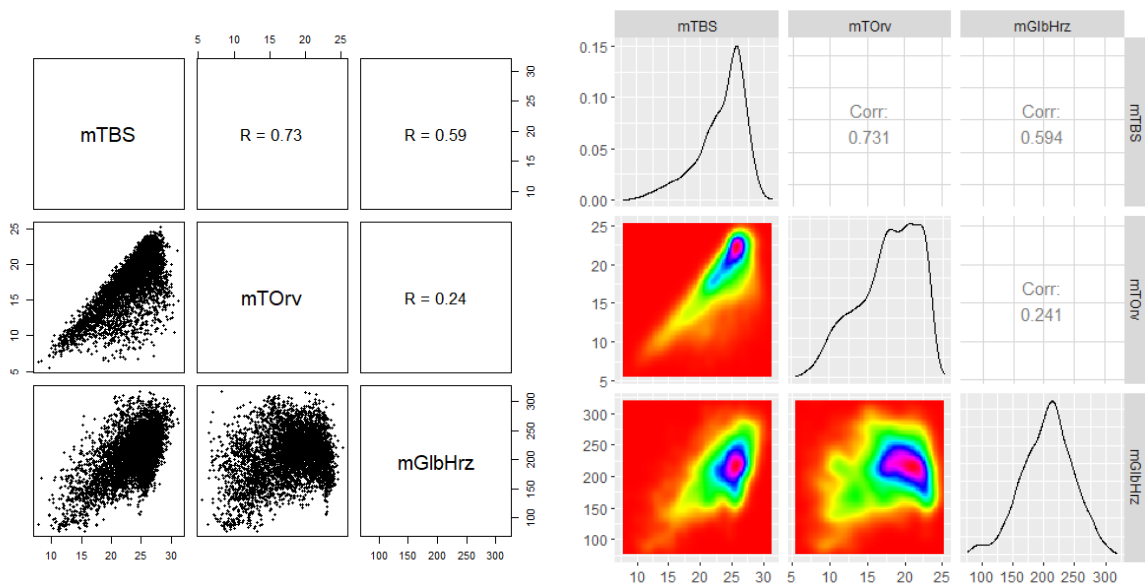
Figura 4.5. Incidência de valores médios anuais das variáveis selecionadas.



Fonte: elaborada pelo autor (2020).

Nota-se que os valores médios mensais apresentaram concentrações um pouco mais elevadas, mas cobriram uma variedade consideravelmente maior de combinações, se comparados aos 411 valores médios anuais.

Figura 4.6. Incidência de valores médios mensais das variáveis selecionadas.



Fonte: elaborada pelo autor (2020).

Além de obter doze vezes mais dados a partir de cada simulação, os valores médios mensais possibilitariam incluir combinações que não seriam possíveis através dos valores

médios anuais. Distribuições contínuas e suficientemente carregadas de pontos favorecem o treinamento do metamodelo, pois habilitam uma maior capacidade de generalização.

Após a análise exploratória voltada aos dados climáticos, definiu-se a granularidade temporal mensal para a metamodelagem. A granularidade mensal permite aumentar consideravelmente o número de descrições climáticas fornecidas ao metamodelo. Com finalidade informativa, foi efetuada uma análise similar, porém mais enxuta, para avaliar a diversidade dos dados climáticos utilizados no desenvolvimento da RNA-C (Apêndice B).

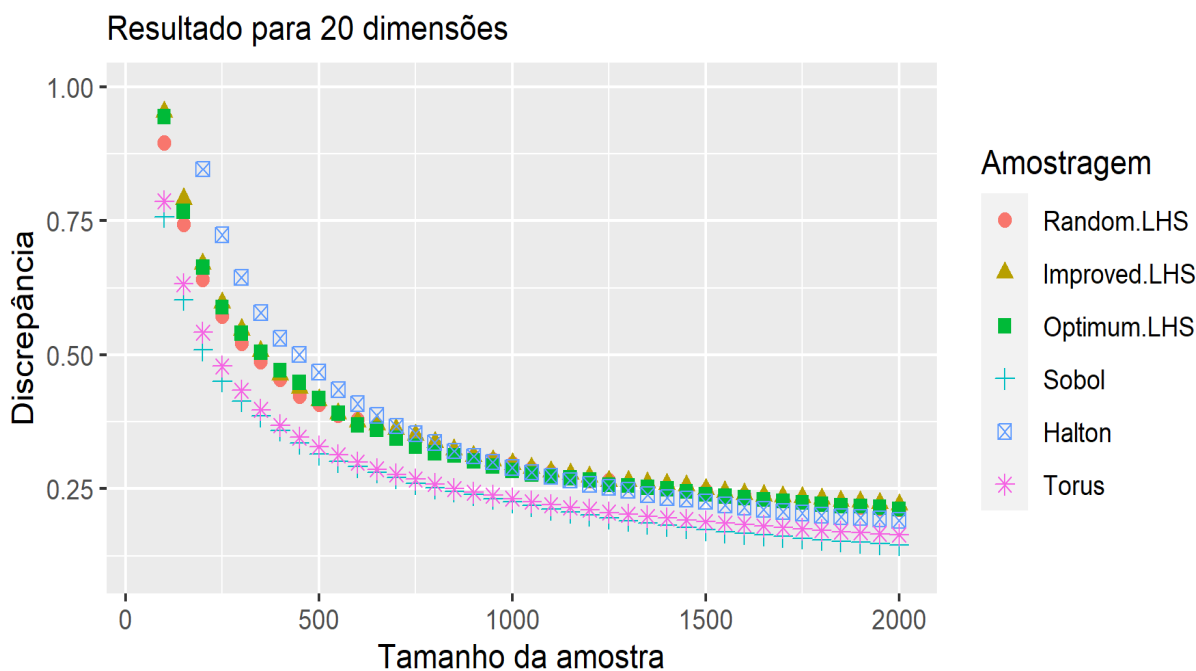
4.2. PREENCHIMENTO DO ESPAÇO AMOSTRAL

Para definir a técnica de amostragem, a segunda análise exploratória focou na discrepância das amostras. A discrepância permitiu medir quão eficientemente as técnicas preenchem o espaço amostral, com a vantagem de não requerer a simulação energética das edificações. A discrepância de cada técnica de amostragem foi calculada, para tamanhos crescentes de amostra.

Foram analisadas seis técnicas de amostragem, compatíveis com amostragens adaptativas, e suas respectivas capacidades em preencher uniformemente um espaço amostral de 20 dimensões. As técnicas de amostragem por Hipercubo Latino analisadas foram: a LHS randômica (*Random LHS*), a LHS melhorada (*Improved LHS*) e a LHS otimizada (*Optimized LHS*). Dentre as sequências quasi-randômicas analisadas, estão as sequências de Halton, Sobol e Kronecker (Torus). O número de dimensões coincide com o número de variáveis do experimento.

Para facilitar a comparação entre as técnicas de amostragem, foi elaborado um gráfico a partir dos valores de discrepância calculados, formando curvas que descrevem a eficiência de preenchimento de um espaço amostral de cada técnica. As curvas foram expostas na Figura 4.7 abaixo.

Figura 4.7. Discrepância em função da técnica e do tamanho da amostra.



Fonte: elaborada pelo autor (2020).

As curvas foram obtidas para um número de dimensões igual a 20, partindo de amostras de 50 até 2000 pontos. A discrepância calculada diminuiu à medida que o tamanho da amostra aumentou. Os valores de discrepância são consideravelmente mais elevados para amostras inferiores a 500 pontos.

A técnica de Halton apresentou valores de discrepância muito elevados em tamanhos de amostra iguais a 50 e 100 pontos, tanto que superaram o limite superior do gráfico. Entretanto, a técnica de Halton diminuiu a diferença em relação às sequências de Torus e Sobol, à medida que o tamanho da amostra aumentou. As técnicas baseadas em LHS tiveram um desempenho mediano e similar entre si, com valores de discrepância superiores, em relação às técnicas de Sobol e de Torus.

A diferença nos valores de discrepância entre a técnica de Sobol e as outras técnicas analisadas poderia ser compensada realizando amostras maiores. Comparativamente, Sobol alcançou valores de discrepância abaixo de 0,2 a partir de 1250 pontos amostrados. A sequência de Torus conseguiu valores inferiores a 0,2 com 1350 pontos, e a sequência de Halton com 1850 pontos. As técnicas baseadas em LHS não alcançaram valores abaixo de 0,2 nem mesmo com 2000 pontos amostrados. A Tabela 4.2 apresenta os valores de discrepância

calculados para 20 dimensões, de 50 até 2000 pontos, segundo as respectivas técnicas de amostragem.

Tabela 4.2. Valores de discrepância para 20 dimensões.

Iteração	Amostra	Dimensões	Random LHS	Improved LHS	Optimum LHS	Sobol	Halton	Torus
0	50	20	1,131874	1,273118	1,297401	1,089538	2,124434	1,142706
1	100	20	0,896079	0,952914	0,944137	0,757738	1,361137	0,786721
2	150	20	0,74444	0,791073	0,767487	0,603226	1,033481	0,632539
3	200	20	0,64105	0,669841	0,663423	0,509566	0,846541	0,54188
4	250	20	0,572116	0,597282	0,588876	0,450906	0,724001	0,479012
5	300	20	0,522453	0,547147	0,540591	0,41422	0,644478	0,433806
6	350	20	0,487508	0,507201	0,505167	0,385775	0,578829	0,397462
7	400	20	0,454222	0,462979	0,471101	0,359451	0,53066	0,368376
8	450	20	0,423466	0,43861	0,448245	0,335009	0,50016	0,345897
9	500	20	0,407814	0,41566	0,418217	0,314285	0,467854	0,328588
10	550	20	0,388102	0,39041	0,391211	0,301138	0,434421	0,313087
11	600	20	0,379708	0,375719	0,369484	0,291431	0,408723	0,299519
12	650	20	0,365467	0,369757	0,360372	0,280995	0,387237	0,286726
13	700	20	0,349976	0,361812	0,344783	0,270472	0,367165	0,277111
14	750	20	0,340643	0,350579	0,328958	0,260003	0,352417	0,267656
15	800	20	0,329429	0,336646	0,317444	0,251951	0,335826	0,258587
16	850	20	0,317823	0,322726	0,312192	0,245517	0,32023	0,250268
17	900	20	0,309302	0,310599	0,302038	0,238997	0,31064	0,243211
18	950	20	0,303313	0,302649	0,292507	0,231719	0,299913	0,237623
19	1000	20	0,296875	0,29646	0,283892	0,225235	0,289433	0,231759
20	1050	20	0,287202	0,288798	0,277067	0,218744	0,280758	0,22575
21	1100	20	0,280025	0,282395	0,271282	0,212376	0,273584	0,219754
22	1150	20	0,271575	0,278247	0,270536	0,206601	0,265677	0,21488
23	1200	20	0,266164	0,27272	0,265844	0,201206	0,258051	0,211371
24	1250	20	0,265122	0,26379	0,257414	<u>0,195503</u>	0,252966	0,205997
25	1300	20	0,259259	0,264281	0,256129	<u>0,190553</u>	0,246673	0,202195

26	1350	20	0,254964	0,261072	0,252112	<u>0,185919</u>	0,239336	<u>0,198427</u>
27	1400	20	0,253285	0,258201	0,249759	<u>0,181694</u>	0,233645	<u>0,195128</u>
28	1450	20	0,250763	0,255634	0,245141	<u>0,177614</u>	0,2303	<u>0,191872</u>
29	1500	20	0,245825	0,248717	0,239745	<u>0,173331</u>	0,225698	<u>0,18876</u>
30	1550	20	0,24081	0,244232	0,236519	<u>0,169337</u>	0,219551	<u>0,186235</u>
31	1600	20	0,237648	0,240137	0,23389	<u>0,166474</u>	0,215392	<u>0,183749</u>
32	1650	20	0,233433	0,237577	0,229978	<u>0,163921</u>	0,211294	<u>0,180601</u>
33	1700	20	0,22898	0,235172	0,22723	<u>0,160904</u>	0,206965	<u>0,177477</u>
34	1750	20	0,226697	0,234448	0,224463	<u>0,157791</u>	0,204338	<u>0,175013</u>
35	1800	20	0,221818	0,23075	0,221221	<u>0,15488</u>	0,20098	<u>0,17236</u>
36	1850	20	0,219587	0,226871	0,218358	<u>0,152467</u>	<u>0,198147</u>	<u>0,17033</u>
37	1900	20	0,215689	0,226172	0,217563	<u>0,150268</u>	<u>0,196755</u>	<u>0,168296</u>
38	1950	20	0,213071	0,223635	0,21552	<u>0,147849</u>	<u>0,19332</u>	<u>0,166248</u>
39	2000	20	0,212385	0,220109	0,212241	<u>0,145282</u>	<u>0,190602</u>	<u>0,164069</u>

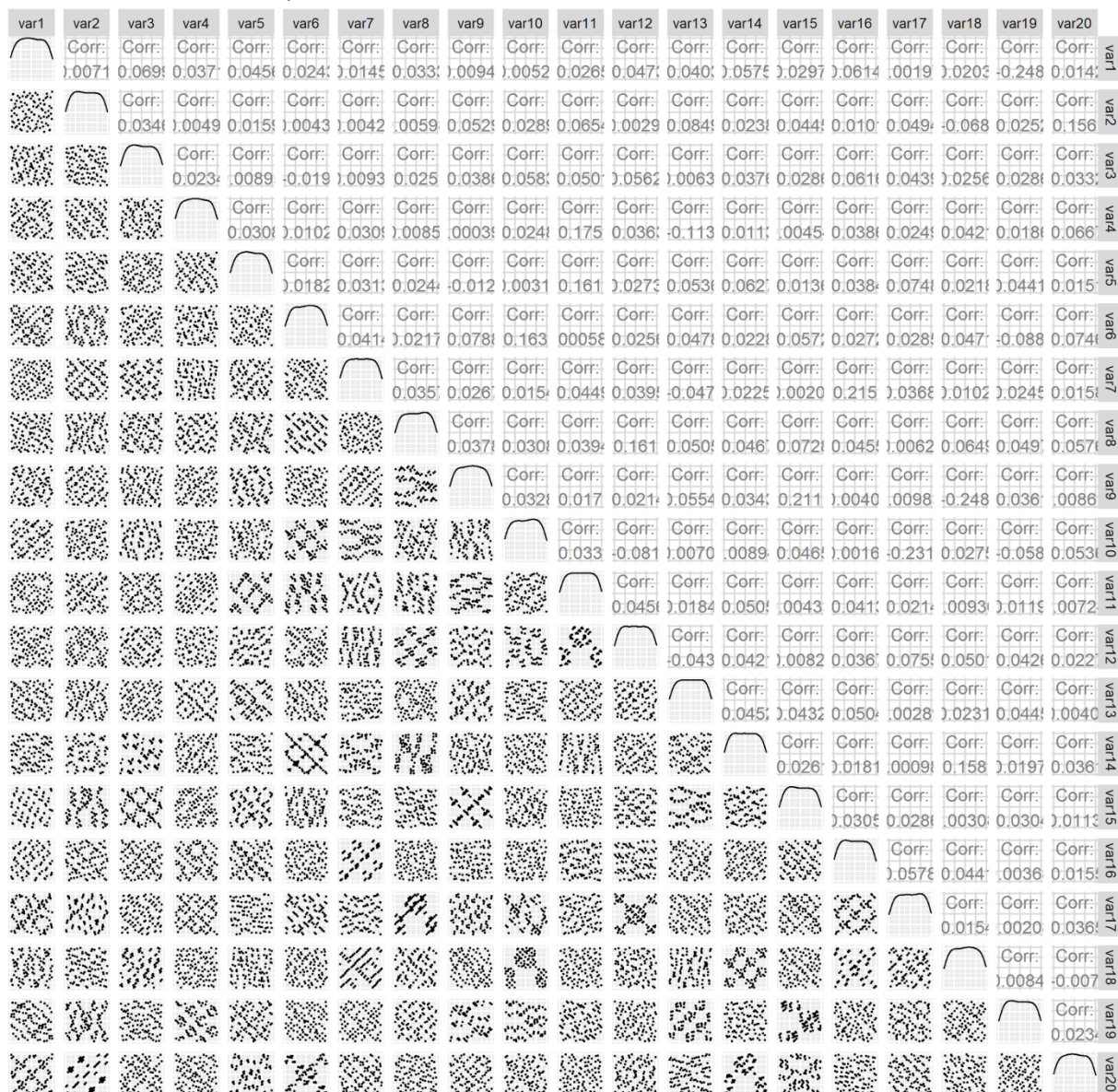
Fonte: elaborada pelo autor (2020).

Ao considerar que cada ponto representa uma simulação, e uma simulação gera resultados para 15 zonas térmicas e 12 meses, a amostra de 2000 pontos equivale a um conjunto de 360000 dados mensais de cargas térmicas de refrigeração. Teoricamente, a análise dos valores de discrepância poderia indicar o número de pontos, em função do número de dimensões e da técnica de amostragem considerada, a partir do qual a simulação de novos casos não aumenta significativamente o desempenho dos metamodelos.

Após o cálculo das discrepâncias, também foi realizada a verificação visual das amostras geradas pelas técnicas de amostragem. A Figura 4.8 abaixo representa a distribuição da amostra dos primeiros 100 pontos para a sequência de Sobol. Estes 100 pontos equivalem aos 100 primeiros arquivos IDF gerados para as simulações.

Figura 4.8. Amostragem de 100 pontos de acordo com a técnica de Sobol.

Sobol / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

A representação visual evidenciou os padrões formados pela amostragem através da sequência de Sobol, entre as dimensões do espaço amostral. Os histogramas na diagonal forneceram uma aproximação da frequência de valores, no intervalo de 0 a 1.

A parte superior dos gráficos apresenta os valores de correlação entre as 20 variáveis. Quanto menor a correlação indicada entre as variáveis, melhor tende a ser o preenchimento do espaço amostral. A análise da representação visual da amostra gerada pela técnica de Sobol, e pelas outras cinco técnicas (representadas no Apêndice C), contribuiu com as seguintes reflexões:

- a) Não foram identificados padrões nas amostras geradas pelas técnicas baseadas em LHS, mas notaram-se sobreposições e agrupamentos de pontos, em praticamente todos os cruzamentos de dimensões;
- b) As técnicas de Torus e de Sobol geraram amostras com padrões comumente encontrados em sequências quasi-randômicas, ineficientes na maioria dos cruzamentos de dimensões;
- c) Devido ao tamanho de amostra reduzido ($n=100$) em relação ao número de dimensões ($d=20$), a técnica de Halton forma padrões altamente ineficientes, da décima dimensão em diante.

Em suma, caso a discrepância da amostra fosse avaliada somente para o pior cruzamento de dimensões, as amostragens por LHS seriam praticamente sempre superiores às sequências quasi-randômicas, devido aos padrões que podem se formar. Contudo, visto que o cálculo da discrepância engloba todas as dimensões, deduz-se que as sobreposições e agrupamentos de pontos penalizam mais as técnicas baseadas em LHS, do que os padrões penalizam as sequências quasi-randômicas analisadas.

Após avaliar o cálculo da discrepância e executar a verificação visual, confirmou-se a técnica de Sobol como a mais adequada para realizar a amostragem adaptativa, ao longo da metamodelagem dinâmica. Contudo, recomenda-se avaliar a técnica de amostragem de acordo com o número de dimensões e o tamanho de amostra de cada pesquisa.

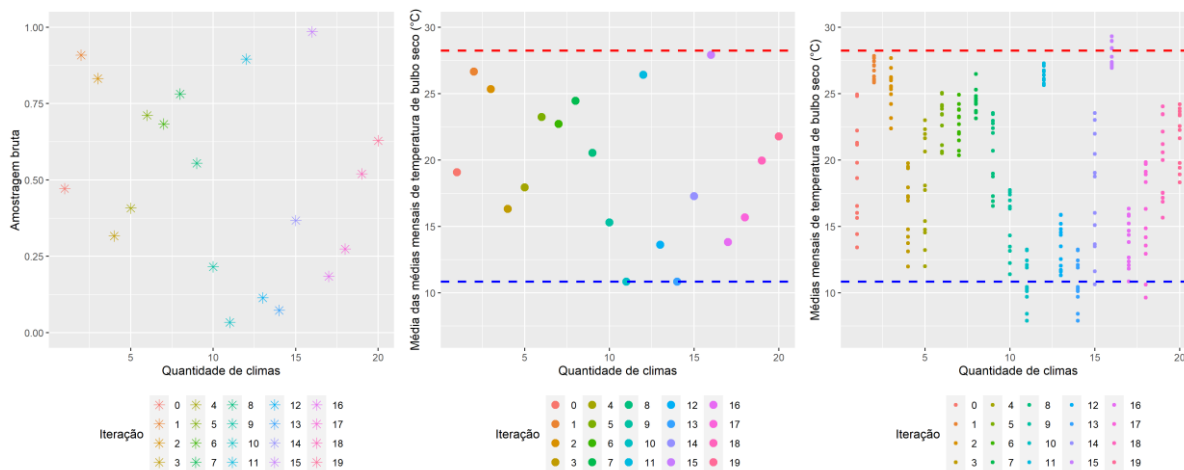
4.3. METAMODELAGEM DINÂMICA

O primeiro passo do processo dinâmico de metamodelagem é sempre a amostragem, independentemente do número da iteração em questão. A amostragem definiu as combinações de variações na edificação e a escolha dos arquivos climáticos. Para orientar a amostragem dos arquivos climáticos, foi calculada uma nova variável climática a partir da base de dados climáticos da análise exploratória: o valor médio das doze médias mensais de temperatura de bulbo seco (MMDBT). Os valores de MMDBT estão listados no Apêndice D, por arquivo climático.

A escolha do arquivo climático inicia por uma amostragem de valores no intervalo entre 0 e 1, denominada de amostragem bruta. Para a amostragem ser adaptada aos valores de MMDBT, o valor bruto igual a 0 é convertido no valor mínimo de MMDBT (linha tracejada azul), e o valor bruto igual a 1 é convertido no valor máximo de MMDBT (linha tracejada

vermelha). Dentre os 411 arquivos climáticos disponíveis, aquele que apresenta o valor de MMDBT mais próximo ao valor convertido é selecionado para a simulação. A amostragem é realizada progressivamente, a cada iteração da metamodelagem, seguindo a mesma lógica. A Figura 4.9 ilustra o procedimento de amostragem dos arquivos climáticos ao longo das iterações realizadas.

Figura 4.9. Procedimento de amostragem dos arquivos climáticos.



Fonte: elaborada pelo autor (2020).

Ao expandir cada arquivo climático em doze meses (à direita), evidenciou-se uma maior cobertura do espaço amostral. A adoção de dados mensais garante maior diversidade climática para a base de dados utilizada no treinamento do metamodelo. A lista das localidades amostradas por iteração foi apresentada na Tabela 4.3.

Tabela 4.3. Lista das localidades amostradas por iteração.

Iteração	Localidade (Estado + Cidade)
0	RS_Santa.Maria
1	CE_Fortaleza-Pinto.Martins.Intl.AP
2	MT_Pontes.de.Lacerda
3	PR_Inacio.Martins
4	RS_Camaqua
5	GO_Goiania
6	MG_Chapada.Gaucha
7	GO_Posse
8	PR_Joaquim.Tavora
9	MG_Monte.Verde.AP
10	SC_Urubici

11	AM_Itacoatiara.AP
12	RJ_Petropolis-Pico.do.Couto
13	SC_Urubici
14	RS_Chui
15	PI_Piripiri
16	RS_Sao.Jose.dos.Ausentes
17	RS_Canela.AP
18	SC_Itapoa
19	MG_Caratinga

Fonte: elaborada pelo autor (2020).

Após as simulações, o nome do arquivo climático ou localidade perde sua relevância, e a lista acima apresenta um caráter meramente informativo. Os dados de cada arquivo climático passam a ser descritos numericamente na base de dados, exclusivamente pelas médias mensais de temperatura de bulbo seco, temperatura do ponto de orvalho e radiação horizontal global.

A amostragem para as variáveis da edificação foi mais direta, pois o valor convertido é inserido diretamente no arquivo de simulação, mesmo no caso de variáveis categóricas. Após as simulações de cada iteração, as variáveis são tratadas de forma a se tornarem predictoras adequadas para o treinamento das redes neurais artificiais.

A descrição da orientação da zona térmica foi realizada por duas variáveis. A primeira variável define se a zona está orientada para o norte ou para o sul, e a segunda variável define se a zona está orientada para o leste ou orientada para o oeste. Desta forma, as zonas perimetrais da edificação sempre recebem zero em uma das variáveis e um valor diferente de zero na outra variável. As zonas internas da edificação, por sua vez, são descritas por zeros em ambas as variáveis.

Em uma lógica similar, duas variáveis foram adicionadas para descrever se a cobertura é exposta e se o piso está em contato com o solo. Também foi adicionada uma variável que indica a diferença, em número de pavimentos, entre o pavimento da zona térmica e o número de pavimentos das edificações vizinhas.

Para diferenciar as zonas internas, seus valores foram zerados nas seguintes variáveis: a absorvância solar da parede; o comprimento do componente de sombreamento; a absorvância solar do entorno opaco; a porcentagem de vidro na fachada das edificações vizinhas; a porcentagem de vidro na fachada; o fator solar do vidro; o delta pavimentos; a capacidade térmica da parede; e a transmitância térmica da parede.

Os valores da densidade urbana para zonas internas não foram zerados, pois estariam representando edificações vizinhas inexistentes, ou seja, uma exposição máxima ao meio externo. A fim de emular uma exposição externa praticamente nula, equivalente a edificações vizinhas extremamente próximas, os valores da densidade urbana foram alterados para 0,5 em todas as zonas internas. A Tabela 4.4 expõe as variações sobre os dados após o tratamento, e os dados de origem que tiveram de ser acessados para efetuá-las. Os nomes das variáveis foram abreviados, quando possível, para facilitar a visualização em representações gráficas subsequentes.

Tabela 4.4. Tratamento efetuado para adequar os dados às granularidades escolhidas.

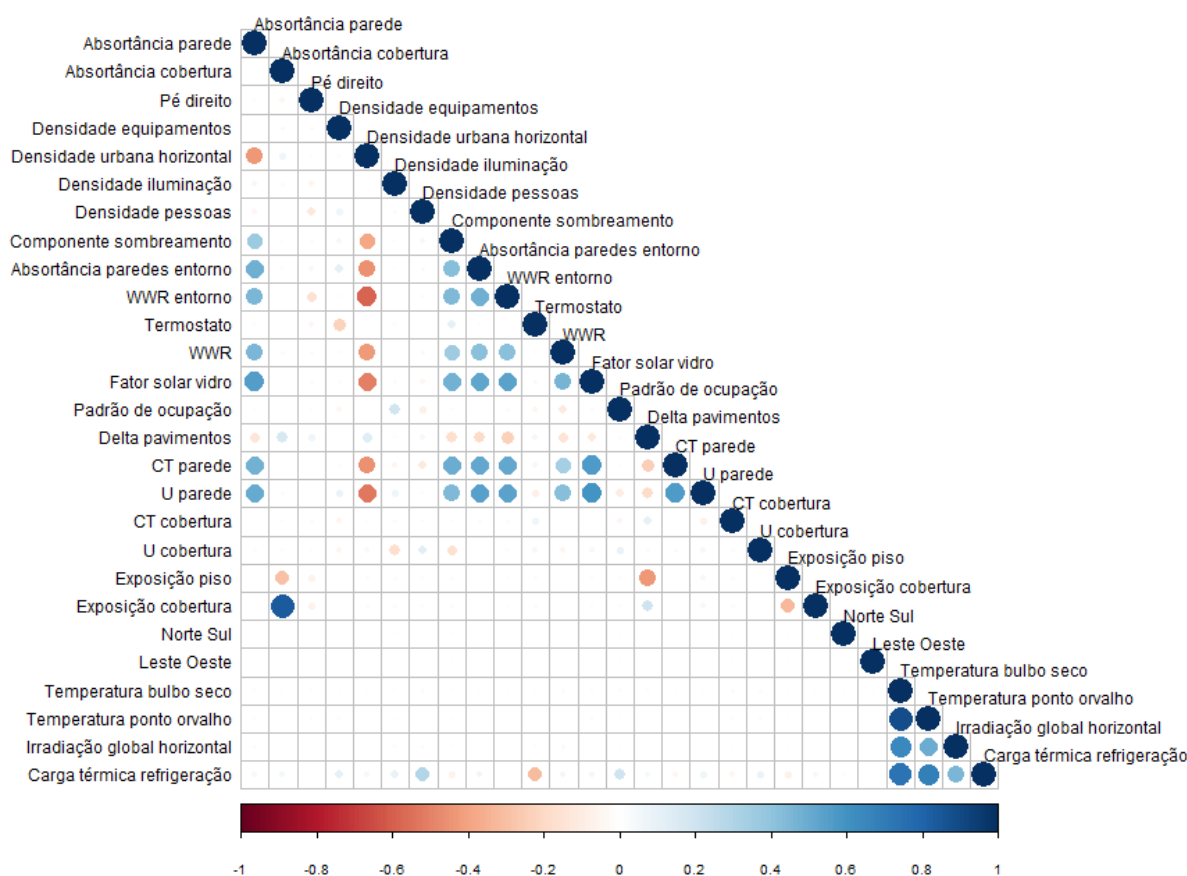
Dados de origem	Nome pós-tratamento	Variações
Tipos de cobertura	U cobertura	1, 2, 3
	CT cobertura	1, 2, 3
Tipos de parede externa	U parede	1, 2, 3 (0 interna)
	CT parede	1, 2, 3 (0 interna)
Tipos de vidro	Fator solar do vidro	0,38 a 0,82 (0 interna)
Número pavimentos das edificações vizinhas	Delta pavimentos	-3 a 2
Número de pavimentos da zona térmica		
Nome da zona térmica	Exposição da cobertura	-1 (não), 1 (sim)
Nome da zona térmica	Exposição do piso	-1 (não), 1 (sim)
Nome da zona térmica	Norte Sul	-1 (sul), 1 (norte)
Nome da zona térmica	Leste Oeste	-1 (oeste), 1 (leste)
Absortância solar da parede	Absortância parede	0 (interna)
Comprimento componente sombreamento	Componente sombreamento	0 (interna)
Absortância solar do entorno opaco	Absortância paredes entorno	0 (interna)
WWR das edificações vizinhas	WWR entorno	0 (interna)
Porcentagem de vidro na fachada	WWR	0 (interna)
Densidade de grade urbana horizontal	Densidade urbana horizontal	0,5 (interna)
Base de dados climática - média mensal	Temperatura bulbo seco (°C)	7,89 a 31,21
Base de dados climática - média mensal	Temperatura ponto de orvalho (°C)	5,56 a 25,16
Base de dados climática - média mensal	Radiação horizontal global (Wh/m ²)	77,41 a 318,24

Fonte: elaborada pelo autor (2020).

A cada iteração, elaborou-se uma matriz de correlação a partir dos dados tratados, com a finalidade de verificar inconsistências na base de dados. Uma amostragem adequada tende a dissipar as correlações entre as variáveis amostradas, entretanto, o mesmo não ocorre após um tratamento de dados que aplica diversas condições sobre as variáveis.

É importante lembrar que as condições impostas no tratamento dos dados serviram para transformar descrições da edificação inteira, provenientes da amostragem, em descrições por zonas térmicas. Após o tratamento dos dados, a correlação entre as 26 variáveis predictoras e a variável dependente foi exposta por uma matriz de correlação, apresentada pela Figura 4.10.

Figura 4.10. Matriz de correlação entre as 26 variáveis predictoras e a variável dependente.



Fonte: elaborada pelo autor (2020).

Ao analisar a matriz de correlação, observa-se que as correlações ocorrem principalmente entre as variáveis que tiveram seus valores adaptados para diferenciar as zonas perimetrais das zonas internas. As correlações são praticamente nulas entre as variáveis igualmente válidas para zonas perimetrais e zonas internas.

Ao focar somente nos dados da variável dependente, ou seja, na carga térmica de refrigeração (última linha da matriz), nota-se que as maiores correlações positivas ocorrem com as três variáveis climáticas, o que reforça a relevância do clima sobre a carga térmica.

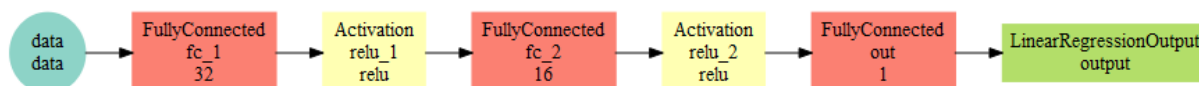
As maiores correlações positivas entre variáveis da edificação e a carga térmica de refrigeração foram encontradas na densidade de pessoas e no padrão de ocupação. Já a maior

correlação negativa ocorreu para o termostato, pois valores elevados para o termostato refletiram em cargas térmicas de resfriamento menores, e vice-versa. Após o tratamento, os dados foram utilizados para o treinamento e a avaliação das redes neurais artificiais.

Em testes preliminares, foram avaliadas configurações de redes neurais com diferentes arquiteturas. Apesar de terem sido efetuados testes com arquiteturas de até cinco camadas ocultas, e até 512 neurônios por camada, a arquitetura foi simplificada com a finalidade de evitar incompatibilização com as bases de dados menores. Nas primeiras iterações, ao combinar muitos neurônios e poucos dados, o RMSE do conjunto de dados de avaliação aumentava por mais de 10 épocas seguidas (*bad steps*), indicando provável *overfitting* e provocando a interrupção do treinamento da rede neural.

Para atender ao último objetivo específico da pesquisa, que consiste em avaliar a evolução das previsões das redes neurais artificiais à medida que a base de dados aumenta, decidiu-se simplificar a arquitetura da rede neural artificial e mantê-la intacta ao longo das iterações. Dessa forma, a arquitetura da rede neural artificial foi modelada com somente duas camadas ocultas, com número de neurônios decrescentes, de acordo com a representação da Figura 4.11.

Figura 4.11. Representação da arquitetura da rede neural artificial.



Fonte: elaborada pelo autor (2020).

À esquerda da representação da arquitetura da rede neural constam os 26 dados de entrada (*data*), seguidos pela primeira camada oculta com 32 neurônios, e pela segunda camada oculta com 16 neurônios. A função de ativação *ReLU* foi adotada nas duas camadas ocultas. Entre os três otimizadores testados, o "adam" (*Adam optimization method*) forneceu a maior estabilidade de resultados.

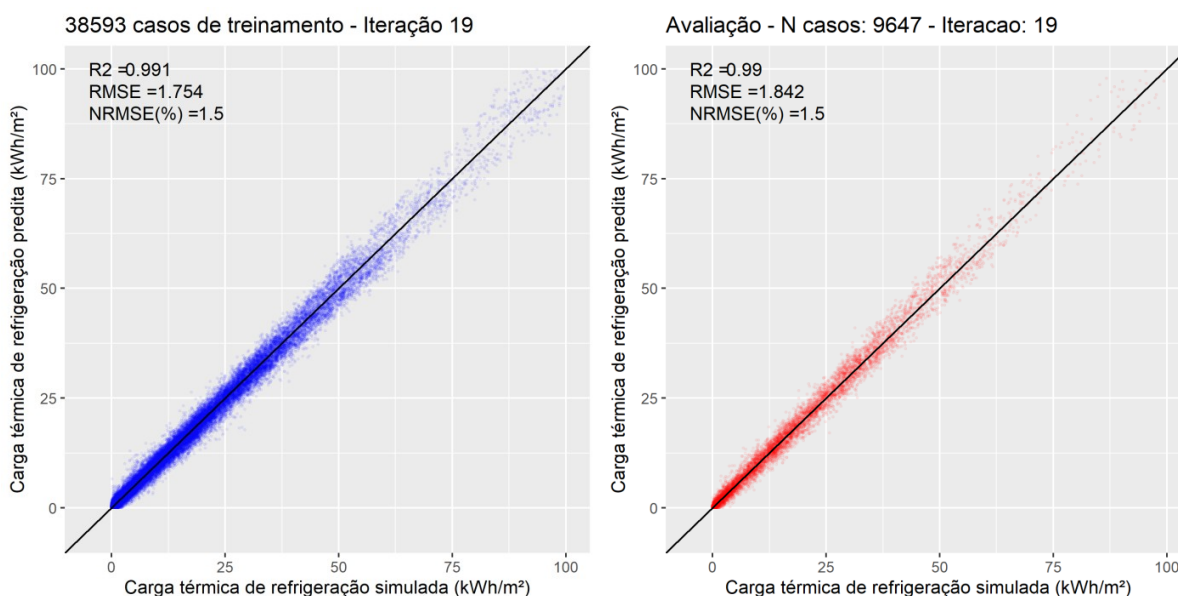
A cada iteração, ao concluir as 700 épocas, interrompeu-se o processo de treinamento e salvaram-se os pesos sinápticos da rede neural artificial. O salvamento dos pesos permitiu a continuação do treinamento na iteração subsequente e evitou que o treinamento iniciasse do zero, já a partir da segunda iteração.

A transferência de aprendizado apresentou sinergias sutis com a metamodelagem dinâmica, ao reduzir somente em poucos segundos o treinamento das últimas iterações,

provavelmente devido ao reduzido conjunto de dados utilizado. Em eventuais continuações deste ou de outros experimentos, à medida que o volume de dados aumenta, espera-se que a transferência de aprendizado possa contribuir de forma mais relevante no processo, reduzindo os tempos de treinamento ou até aumentando a acurácia dos metamodelos.

Com o objetivo de documentar e expor a acurácia dos metamodelos elaborados, ao final de cada iteração, realizou-se comparações entre o resultado predito pelos modelos preditivos e o resultado esperado, que é proveniente das simulações. A comparação de valores simulados e preditos permitiu avaliar os resultados de R^2 , RMSE e NRMSE para os conjuntos de treinamento e de avaliação, como apresentado pela Figura 4.12.

Figura 4.12. Conjuntos de dados de treinamento (80%) e avaliação (20%), da última iteração.



Fonte: elaborada pelo autor (2020).

Basicamente, nas imagens acima foram representadas as capacidades de predição da rede neural artificial da última iteração da metamodelagem dinâmica. Esta foi a última iteração, pois o RMSE do conjunto de dados de avaliação atingiu valores inferiores a 2 kWh/m², ao final do treinamento desta iteração.

Como observado, a simplificação da arquitetura da rede neural artificial não impediu a obtenção de bons resultados, pelo contrário. Os dados de treinamento e os dados de avaliação representaram 80% e 20% do total de dados simulação, para todas as iterações, por isso a diferença no número de casos indicados em cada gráfico. O RMSE do conjunto de dados de treinamento manteve-se ligeiramente inferior ao RMSE do conjunto de dados de avaliação, o que é considerado normal ou até desejado.

Entretanto, os resultados acima são válidos para os denominados dados vistos (conjuntos de treinamento e avaliação). O procedimento de teste da rede neural, com dados nunca vistos (conjuntos de teste), ocorreu somente após o término da metamodelagem dinâmica.

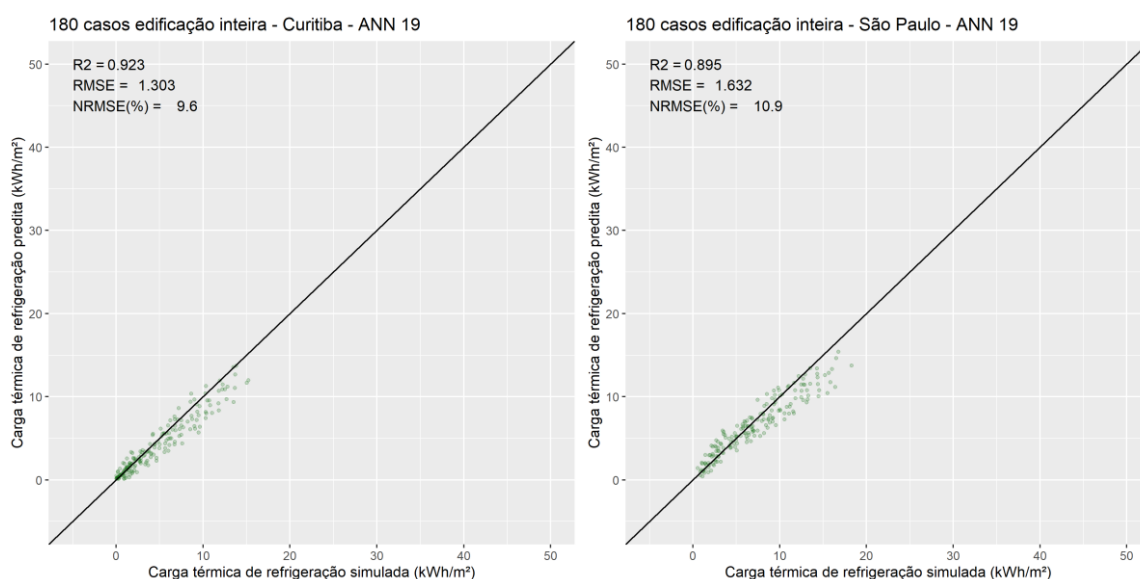
4.4. TESTES (VALIDAÇÕES)

Na primeira bateria de testes, foi utilizada a base de dados de simulações preliminares, previamente gerada durante as análises exploratórias. Esta base de dados é composta por casos nunca vistos pelos metamodelos, e serviu para testar os metamodelos ao emular uma edificação inteira de escritório, descrita segundo a Tabela 3.1.

A edificação emulada foi composta por resultados de cinco zonas térmicas térreas, cinco intermediárias e cinco coberturas. Cada pavimento incluiu uma zona interna e quatro zonas perimetrais com quatro orientações possíveis. A descrição foi complementada pela descrição de doze meses específicos das localidades selecionadas.

As previsões foram executadas inicialmente em dois climas brasileiros amenos: Curitiba e São Paulo. A Figura 4.13 apresenta as previsões (versus os valores simulados) das zonas térmicas que compõem uma edificação de escritório localizada em Curitiba e em São Paulo.

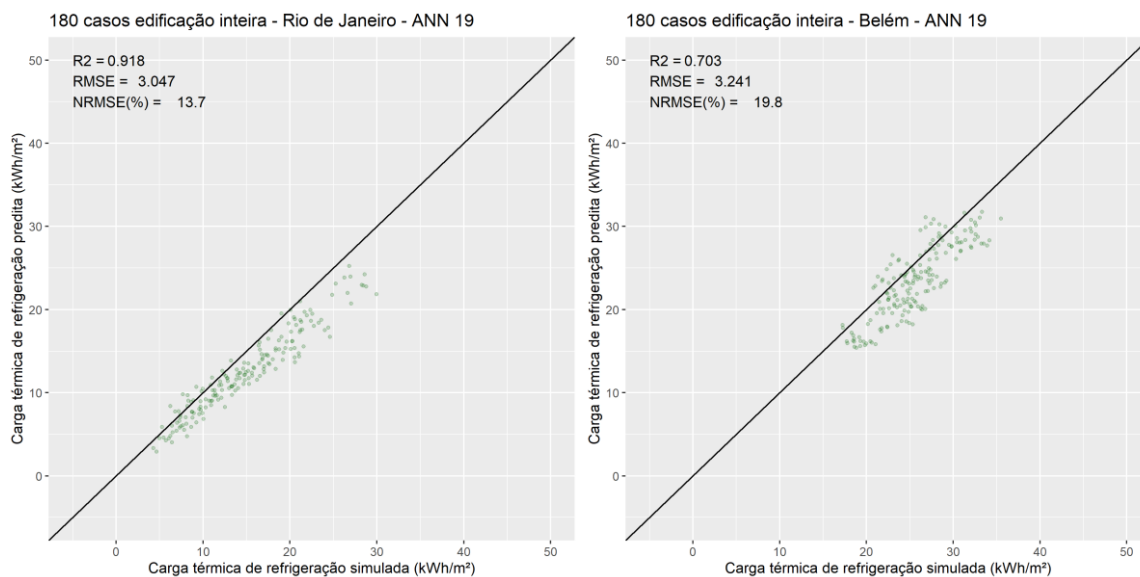
Figura 4.13. Previsões para escritório com 15 zonas térmicas ao longo de 12 meses, para as cidades de Curitiba e São Paulo.



Fonte: elaborada pelo autor (2020).

Para os climas amenos, as previsões foram particularmente boas, ao se considerar que consistem em casos nunca vistos. O RMSE obtido para ambas as localidades foi inferior a 1,7 kWh/m². Também foram executadas as previsões para dois climas brasileiros quentes: Rio de Janeiro e Belém. A Figura 4.14 apresenta as previsões das zonas térmicas que compõem uma edificação de escritório localizada no Rio de Janeiro e em Belém.

Figura 4.14. Previsões para escritório com 15 zonas térmicas ao longo de 12 meses, para as cidades de Rio de Janeiro e Belém.



Fonte: elaborada pelo autor (2020).

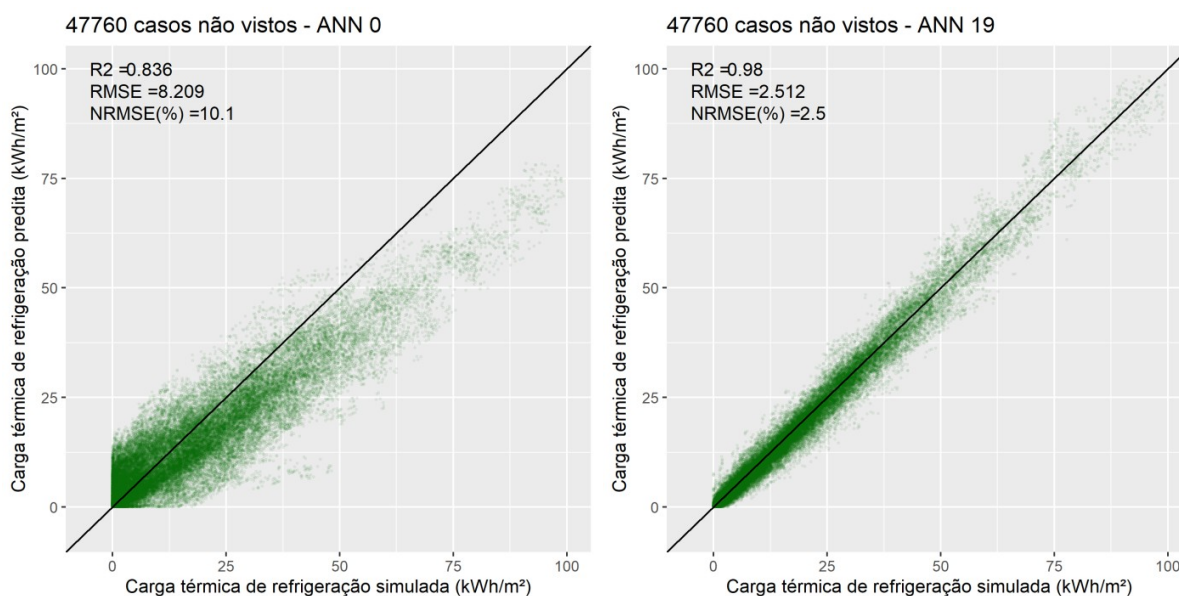
Os climas mais quentes apresentaram resultados inferiores aos climas amenos nesta análise, com valores de RMSE ligeiramente superiores a 3 kWh/m². Apesar de facilitar a comparação dos erros obtidos em conjunto de dados com diferentes unidades ou escalas, o NRMSE penalizou os climas individuais, pois é calculado a partir do intervalo entre o valor predito mínimo e o valor predito máximo. Por considerar uma amplitude de valores preditos ainda mais reduzida, os valores de NRMSE encontrados para Belém foram os mais penalizados. A baixa amplitude de valores de carga térmica obtidos para o clima de Belém é consequência das variações mais brandas entre as doze temperaturas médias mensais consideradas.

Para compor uma segunda bateria de teste, mais robusta, foi elaborado um conjunto de dados de teste a partir de novas simulações, distribuídas entre 20 climas amostrados. Com a intenção de não adotar novamente a técnica de amostragem de Sobol para definir o conjunto de dados de teste, realizou-se a amostragem do conjunto de teste pela técnica de LHS randômica. De qualquer forma, a adoção de outra técnica de amostragem é opcional,

pois bastaria modificar as sementes de randomização da técnica em questão para gerar combinações diferentes.

O conjunto de dados de teste da segunda bateria foi aplicado nas 20 redes neurais artificiais salvas ao longo das iterações, com a finalidade de verificar a evolução das predições sobre cerca de 47760 casos nunca vistos. A Figura 4.15 compara a performance da rede neural artificial da primeira (ANN 0) e da última iteração (ANN 19), ao serem submetidas aos dados de teste, provenientes de combinações nunca vistas.

Figura 4.15. Conjunto de dados de teste aplicado nas redes neurais artificiais da primeira e da última iteração.

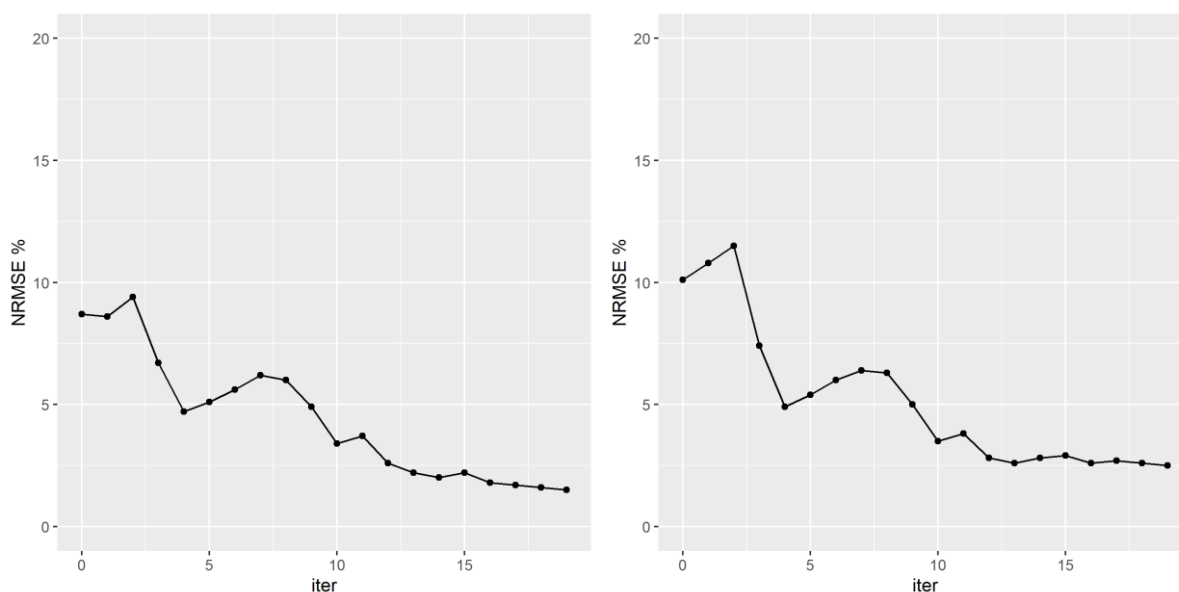


Fonte: elaborada pelo autor (2020).

A diferença de capacidade de predição entre a rede neural artificial da primeira e a da última iteração é evidente. O valor de RMSE apresentado pela rede neural da primeira iteração (ANN 0) é mais de três vezes superior ao valor de RMSE obtido pela rede neural da última iteração (ANN 19). Apesar de parecer limitado inicialmente, o desempenho da rede neural artificial da primeira iteração também surpreendeu positivamente, pois sua base de dados era composta por resultados de somente 20 simulações e um único arquivo climático. Esta é mais uma prova do potencial que a granularidade mensal fornece à metamodelagem dinâmica, ao aumentar a diversidade de dados climáticos que podem ser fornecidos à rede neural artificial. Ao comparar o valor de RMSE = 2,512 kWh/m² do conjunto de dados de teste (composto por casos nunca vistos), nota-se que é 36% superior ao RMSE = 1,842 kWh/m² do conjunto de dados de avaliação da rede neural artificial da última iteração (ANN 19).

O teste das redes neurais artificiais das 20 iterações permitiu avaliar a evolução da performance do metamodelo, à medida que mais simulações foram adicionadas à base de dados. Para evitar a exposição de múltiplos gráficos de dispersão, foram elaboradas curvas de evolução do NRMSE das redes neurais artificiais, para os dados de treinamento e dados de avaliação (casos vistos) à esquerda, e para os dados de teste (casos nunca vistos) à direita na Figura 4.16.

Figura 4.16. Curva de evolução das predições por iteração, ao utilizar Sobol na amostragem, para casos vistos (treinamento e avaliação) à esquerda, e casos nunca vistos (teste) à direita.



Fonte: elaborada pelo autor (2020).

A representação da curva de evolução foi efetuada a partir de valores de NRMSE para simplificar a comparação com eventuais estudos futuros. De qualquer forma, a tendência de RMSE e NRMSE foi praticamente a mesma nos dados explorados. Cada iteração contribuiu com 20 arquivos de simulação (IDF), e um arquivo climático (EPW).

Após a oitava iteração, observou-se uma melhoria considerável nos valores de NRMSE apresentados pelas redes neurais. A estabilização dos valores de NRMSE em casos nunca vistos ocorreu após a décima segunda iteração, indicando uma diminuição no benefício de adicionar novos casos ao treinamento do metamodelo.

A finalização da metamodelagem, após a vigésima iteração, indicou que os níveis desejados de acurácia do metamodelo foram alcançados (RMSE para o conjunto de avaliação $< 2 \text{ kWh/m}^2$). Este valor foi alcançado com o tamanho de amostra de apenas 400 casos, amostrados pela sequência de Sobol. Após a última iteração, a base de dados consolidada foi

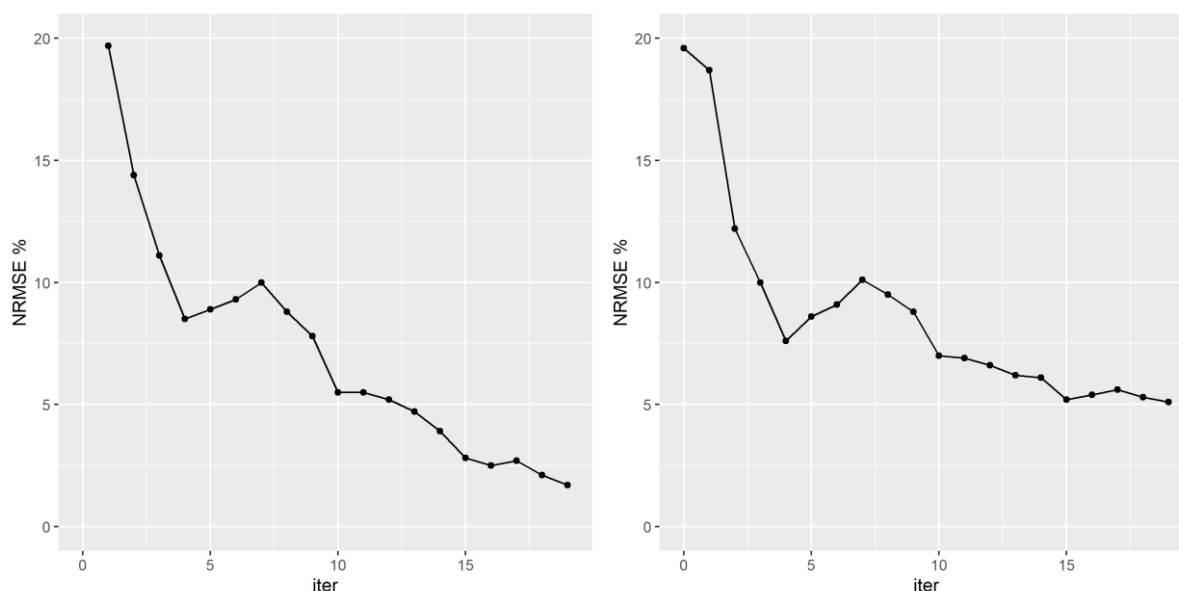
composta por 48240 zonas térmicas, resultantes da combinação de 400 casos, 12 meses por arquivo climático e, na média, 10 zonas térmicas por edificação.

Nesta avaliação, o mais importante é a comprovação da capacidade de generalização do metamodelo a partir de um conjunto reduzido de dados, elaborado a partir de 400 simulações energéticas. A continuação da metamodelagem dinâmica permitiria buscar resultados melhores, mas iterações adicionais não agregam tamanha informação quanto as iniciais. Por isso, as análises de teste mantiveram seu foco nas iterações iniciais.

É importante lembrar que foi adotada a sequência de Sobol para amostrar os casos que compuseram a base de dados das redes neurais artificiais salvas ao longo da metamodelagem dinâmica. Para fins comparativos, a mesma sequência de etapas de metamodelagem foi executada novamente, modificando apenas a técnica de amostragem de Sobol para Halton, e obrigando a parada da metamodelagem dinâmica nas mesmas vinte iterações.

Ao adotar a sequência de Halton, novas curvas de evolução do NRMSE das redes neurais artificiais foram geradas. Os casos vistos (dados de treinamento + dados de avaliação) ficaram à esquerda, e os casos nunca vistos (dados de teste) à direita na Figura 4.17.

Figura 4.17. Curva de evolução das predições por iteração, ao utilizar Halton na amostragem, para casos vistos (treinamento e avaliação) à esquerda, e casos nunca vistos (teste) à direita.



Fonte: elaborada pelo autor (2020).

A comparação permitiu verificar as diferenças de performance das predições, ao utilizar as técnicas de amostragem que apresentaram os valores mais altos e mais baixos de discrepância. Como detalhado na análise exploratória sobre a eficiência do preenchimento do

espaço amostral, a técnica de Halton apresentou os maiores valores de discrepância, para os primeiros 750 pontos amostrados (Tabela 4.2).

A dificuldade da técnica de Halton, em preencher o espaço amostral uniformemente nas condições citadas, penaliza a sua curva de evolução das previsões por iteração. Ao utilizar a técnica de Halton, os valores encontrados de NRMSE foram entre 70% e 100% superiores àqueles obtidos com a técnica de Sobol, ao longo das iterações executadas. Esta diferença de desempenho reforça a importância de análises mais aprofundadas sobre as técnicas de amostragem e sobre o tamanho de amostra para o desenvolvimento de metamodelos.

5. CONCLUSÃO

Esta tese desenvolveu um processo dinâmico de metamodelagem, aplicado à carga térmica de refrigeração de escritórios condicionados artificialmente, em climas brasileiros. Os temas que motivaram o trabalho incluíram a consideração efetiva de múltiplos climas e a eficiência da amostragem no processo de metamodelagem. Estes temas, por sua vez, estão conectados com as necessidades de normas de eficiência energética em edificações, em especial àquelas que visam desacelerar o crescimento das demandas energéticas voltadas para o condicionamento artificial de edificações.

Por meio de simulações preliminares, a primeira análise exploratória avaliou a relação entre os dados climáticos de 411 localidades brasileiras e a carga térmica de refrigeração em edificações de escritório. A temperatura de bulbo seco apresentou as correlações mais elevadas com as cargas térmicas de resfriamento, em especial com a parcela sensível. Após evidenciar sua alta correlação com a carga térmica de refrigeração latente, a temperatura do ponto de orvalho substituiu a umidade relativa, amplamente adotada na literatura. A radiação horizontal global forneceu uma descrição adicional aos climas, por exercer influência sobre todas as variáveis da envoltória e, conseqüentemente, impactar as zonas térmicas com exposição ao exterior.

A adoção de dados mensais, ao invés dos dados anuais, garantiu uma maior diversidade a partir do mesmo número de simulações e arquivos climáticos, acompanhada de maior representatividade para os valores utilizados na descrição climática. Além de cobrir o espaço amostral das variáveis climáticas de forma mais contínua, a diversidade de dados mensais permitiu obter combinações de dados não possíveis através de valores médios anuais. Granularidades inferiores ao intervalo mensal foram evitadas, pois implicariam em mais variáveis e em uma quantidade de dados desnecessários para a pesquisa.

A segunda análise exploratória avaliou a discrepância da amostra, de acordo com o tamanho de amostra e a técnica de amostragem adotada. A técnica de amostragem para a metamodelagem dinâmica foi escolhida com base na melhor curva de eficiência de preenchimento de espaço. Para o número de 20 dimensões e para amostras de até 2000 casos, a sequência de Sobol obteve os melhores resultados. A sequência de Sobol alcançou valores de discrepância abaixo de 0,2 a partir de 1250 pontos amostrados. A sequência de Torus conseguiu valores inferiores a 0,2 com 1350 pontos, e a sequência de Halton com 1850 pontos. As técnicas baseadas em LHS não alcançaram valores de discrepância abaixo de 0,2 nem mesmo com 2000 pontos amostrados.

Após as análises exploratórias, deu-se início a metamodelagem dinâmica, considerando as três variáveis climáticas definidas para balizar a descrição climática, e adotando a sequência de Sobol para realizar as amostragens adaptativas. A amostragem adaptativa permitiu aumentar progressivamente a diversidade de casos, sem descartar as combinações previamente simuladas, o que manteve o custo computacional baixo. A abordagem adaptativa pode ser considerada como alternativa eficiente à abordagem “tentativa e erro”, utilizada em estudos que envolvem simulações energéticas, para a definição de um tamanho de amostra adequado ao metamodelo.

A modelagem e a execução das simulações foram seguidas pelo tratamento de dados, que envolveu a criação e exclusão de variáveis, e a normalização dos dados antes do treinamento das redes neurais artificiais. A simplificação da arquitetura das redes neurais artificiais, para duas camadas ocultas com 32 e 16 neurônios, foi a solução adotada para evitar a ocorrência de *overfitting* nas primeiras iterações da metamodelagem dinâmica. Desta forma, foi mantido o tamanho de amostra reduzido e documentou-se o desempenho que pode ser obtido a partir de um número mínimo de simulações.

A transferência de aprendizado também foi testada, mas não foi incluída nos resultados, pois impactou muito pouco o treinamento das redes neurais, devido ao número reduzido de dados e à arquitetura simplificada da rede neural. Contudo, à medida que o número de dados aumenta e o modelo se torna mais complexo, a transferência de aprendizado tende a reduzir o tempo necessário para o treinamento e aumentar a acurácia das predições.

Desta forma, os procedimentos da metamodelagem dinâmica (amostragem, simulação, tratamento e treinamento) do metamodelo ocorreram progressivamente, até a rede neural artificial alcançar valores de RMSE inferiores a 2 kWh/m² para o conjunto de avaliação. Abaixo deste valor, interrompeu-se a metamodelagem dinâmica e iniciaram-se os testes com casos nunca vistos (validação). A interrupção da metamodelagem dinâmica após 20 iterações, indicou que os níveis desejados de acurácia do metamodelo foram alcançados (RMSE para o conjunto de avaliação < 2 kWh/m²) com o tamanho de amostra de apenas 400 casos, amostrados pela sequência de Sobol. Os conjuntos de teste foram definidos por amostragens independentes, não relacionadas com o processo de metamodelagem.

Na primeira bateria de testes, os climas amenos apresentaram valores de RMSE abaixo de 1,7 kWh/m², enquanto os climas mais quentes apresentaram resultados de RMSE ligeiramente superiores a 3 kWh/m². Na segunda bateria de testes, o RMSE obtido pela rede neural artificial da vigésima iteração foi de 2,5 kWh/m², ao predizer 47760 casos nunca vistos. O RMSE obtido pela rede neural artificial da primeira iteração foi de 8,2 kWh/m², mas

deve-se considerar que sua base de dados era composta por somente 20 arquivos de simulação e um arquivo climático.

A diversidade de valores provenientes dos dados mensais contribuiu para este desempenho. Os dados climáticos mensais proporcionaram ótimas condições de aprendizado para o treinamento de redes neurais artificiais, possibilitando a sua aplicação em qualquer clima brasileiro. Inclusive, a diversidade de valores fornecida pelos dados mensais pode ser muito útil na eventualidade de futuras modificações dos arquivos climáticos, seja por efeito de mudanças climáticas, adensamento urbano, ou simples atualização dos dados climáticos.

A última comparação verificou as diferenças de acurácia das predições, ao adotar a técnica de amostragem que apresentou os valores mais altos (Halton) e mais baixos (Sobol) de discrepância para amostras de até 400 casos. A diferença de desempenho obtido entre as duas técnicas variou entre 70% e 100% para os valores de NRMSE obtidos, reforçando a importância de análises aprofundadas sobre as técnicas de amostragem e sobre o tamanho de amostra para o desenvolvimento de metamodelos. Uma reflexão adicional, relacionada a estudos que geram dados sintéticos para desenvolver metamodelos, está associada à escassez de recursos computacionais. Ganhos de eficiência no processo de metamodelagem são importantes e devem ser explorados. Inclusive, maiores eficiências em experimentos computacionais tendem a viabilizar o desenvolvimento de metamodelos, com escopos bastante amplos em máquinas comuns, diminuindo a dependência de supercomputadores.

As melhorias implementadas nesta pesquisa viabilizaram o desenvolvimento de um metamodelo consideravelmente enxuto, que elimina a restrição de grupos climáticos, presente atualmente na INI-C. Para fins comparativos, a RNA-C foi desenvolvida a partir de 374 mil simulações, enquanto a metamodelagem dinâmica executada baseou-se em 400 simulações. Ao analisar as devidas diferenças e limitações dos métodos, vislumbram-se ganhos expressivos de eficiência computacional ao adotar e dar continuidade à metamodelagem dinâmica proposta. Espera-se que o desenvolvimento da metamodelagem dinâmica e a exploração dos temas relacionados ofereçam contribuições para o aperfeiçoamento contínuo das regulamentações de eficiência energética em edificações no Brasil, além de fornecer inspirações para as futuras pesquisas nesta área de interesse.

5.1. LIMITAÇÕES

As principais limitações deste trabalho referem-se:

- a) À utilização de dados sintéticos para a metamodelagem, ou seja, dados dependentes das limitações de simulações e premissas;
- b) Às restrições referentes aos limites das faixas de valores e/ou categorias das variáveis adotadas para a modelagem das simulações;
- c) À descrição das variáveis preditoras do metamodelo referentes às zonas térmicas, realizada no tratamento de dados;
- d) À escolha das redes neurais artificiais como único tipo de metamodelo analisado, apesar da vasta utilização das mesmas em trabalhos da área;
- e) Ao número de variáveis do metamodelo, que apesar de poder ser menor não foi reduzido, visando manter os parâmetros comumente encontrados em normas e regulamentos de eficiência energética em edificações;
- f) Às simplificações dos hiperparâmetros das redes neurais artificiais, de modo a se adequar ao volume de dados reduzido das iterações iniciais;
- g) Aos benefícios da transferência de aprendizado, ainda limitados no contexto deste trabalho, mas que se tornariam mais evidentes com o desenvolvimento de metamodelos mais aptos ao aprendizado incremental.

5.2. SUGESTÕES PARA TRABALHOS FUTUROS

Vislumbram-se as seguintes sugestões para trabalhos futuros:

- a) Predizer as cargas térmicas sensíveis e latentes separadamente;
- b) Testar a temperatura operativa na descrição climática;
- c) Desenvolver metamodelos independentes para as zonas térmicas internas e perimetrais da edificação, para facilitar a representação de múltiplos formatos de edificações, além de evitar extensivos tratamentos de dados;
- d) Criar o conjunto de treinamento e o conjunto de avaliação a partir de duas amostragens diferentes, de modo a não particionar a amostra;
- e) Modelar um sistema de *bins* (caixas de frequência), para substituir o uso de valores médios em variáveis climáticas;
- f) Explorar a interpretabilidade dos metamodelos (*explainable AI*) através das predições obtidas pelas redes neurais artificiais;
- g) Testar o aumento de dados (*data augmentation*) para criar mais dados sintéticos a partir dos dados provenientes das simulações.

REFERÊNCIAS

ABNT. Desempenho térmico de edificações Parte 3: Zoneamento bioclimático brasileiro e diretrizes construtivas para habitações unifamiliares de interesse social. **NBR 15.220-3**, 2005.

AGENCIA NACIONAL DE VIGILÂNCIA SANITÁRIA. Resolução-Re Nº 9, De 16 De Janeiro De 2003. **Diário Oficial da União**, p. seção 1 numero 14, 2003. Disponível em: <http://portal.anvisa.gov.br/documents/33880/2568070/RE_09_2003.pdf/f4af80d4-8516-4f9c-a745-cc8b4dc15727>.

AHMAD, T.; CHEN, H.; GUO, Y.; WANG, J. A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: A review. **Energy and Buildings**, v. 165, p. 301–320, 2018. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.enbuild.2018.01.017>>.

ALSAADANI, S.; DE SOUZA, C. B. Of collaboration or condemnation? Exploring the promise and pitfalls of architect-consultant collaborations for building performance simulation. **Energy Research and Social Science**, v. 19, p. 21–36, 2016. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.erss.2016.04.016>>.

AMASYALI, K.; EL-GOHARY, N. Deep learning for building energy consumption prediction. **6th CSCE-CRC International Construction Specialty Conference 2017 - Held as Part of the Canadian Society for Civil Engineering Annual Conference and General Meeting 2017**, v. 1, p. 466–474, 2017.

AMASYALI, K.; EL-GOHARY, N. M. A review of data-driven building energy consumption prediction studies. **Renewable and Sustainable Energy Reviews**, v. 81, p. 1192–1205, 2018. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2017.04.095>>.

BAVARESCO, M.; MAZZAFERRO, L.; PAULA MELO, A.; LAMBERTS, R. Classificação de climas brasileiros empregada na atualização dos Regulamentos Técnicos da Qualidade para o Nível de Eficiência Energética de Edificações. , p. 8, 2017. Disponível em: <www.cb3e.ufsc.br>.

BERGSTRA, J.; BARDENET, R.; BENGIO, Y.; KÉGL, B. Algorithms for Hyper-Parameter Optimization. **Advances in Neural Information Processing Systems (NIPS)**, p. 2546–2554, 2011.

BLEIL DE SOUZA, C. Studies into the use of building thermal physics to inform design decision making. **Automation in Construction**, v. 30, p. 81–93, 2013. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.autcon.2012.11.026>>.

BOURDEAU, M.; ZHAI, X. QIANG; NEFZAOU, E.; GUO, X.; CHATELLIER, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. **Sustainable Cities and Society**, v. 48, n. February, p. 101533, 2019. Elsevier. Disponível em: <<https://doi.org/10.1016/j.scs.2019.101533>>.

BRAUN, M. R.; ALTAN, H.; BECK, S. B. M. Using regression analysis to predict the future energy consumption of a supermarket in the UK. **Applied Energy**, v. 130, p. 305–313, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2014.05.062>>.

BRE, F.; ROMAN, N.; FACHINOTTI, V. D. An efficient metamodel-based method to carry out multi-objective building performance optimizations. **Energy & Buildings**, v. 206, p. 109576, 2020. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.enbuild.2019.109576>>.

BROWNLEE, A. E. I.; WRIGHT, J. A. Constrained, mixed-integer and multi-objective optimisation of building designs by NSGA-II with fitness approximation. **Applied Soft Computing Journal**, v. 33, p. 114–126, 2015. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.asoc.2015.04.010>>.

BRUNETTI, G. L. Increasing the efficiency of simulation-based design explorations via metamodeling. **Journal of Building Performance Simulation**, v. 1493, 2020. Disponível em: <<https://doi.org/10.1080/19401493.2019.1707875>>.

CARNELL, R. Package “lhs” in R. **Cran-R**, 2018. Disponível em: <<http://lhs.r-forge.r-project.org>>.

CATALINA, T.; IORDACHE, V.; CARACALEANU, B. Multiple regression model for fast prediction of the heating energy demand. **Energy and Buildings**, v. 57, p. 302–312, 2013. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.enbuild.2012.11.010>>.

CHAE, Y. T.; HORESH, R.; HWANG, Y.; LEE, Y. M. Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings. **Energy and Buildings**, v. 111, p. 184–194, 2016. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.enbuild.2015.11.045>>.

CHEN, X.; YANG, H. Integrated energy performance optimization of a passively designed high-rise residential building in different climatic zones of China. **Applied Energy**, v. 215, n. February, p. 145–158, 2018. Elsevier. Disponível em: <<https://doi.org/10.1016/j.apenergy.2018.01.099>>.

DALY, A.; DEKKER, T.; HESS, S. Dummy coding vs effects coding for categorical variables: Clarifications and extensions. **Journal of Choice Modelling**, v. 21, n. September, p. 36–41, 2016.

DAMBLIN, G.; COUPLET, M.; IOOSS, B. Numerical studies of space-filling designs: Optimization of Latin Hypercube Samples and subprojection properties. **Journal of Simulation**, v. 7, n. 4, p. 276–289, 2013.

DING, Y.; ZHANG, Q.; YUAN, T.; YANG, F. Effect of input variables on cooling load prediction accuracy of an office building. **Applied Thermal Engineering**, v. 128, p. 225–234, 2018. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.applthermaleng.2017.09.007>>.

EMPRESA DE PESQUISA ENERGÉTICA (BRASIL). Balanço Energético Nacional 2020: Ano base 2019. Rio de Janeiro : EPE, p. 264, 2020.

FAN, C.; XIAO, F.; WANG, S. Development of prediction models for next-day building energy consumption and peak power demand using data mining techniques. **Applied Energy**, v. 127, p. 1–10, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2014.04.016>>.

FAN, C.; XIAO, F.; ZHAO, Y. A short-term building cooling load prediction method using deep learning algorithms. **Applied Energy**, v. 195, p. 222–233, 2017. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2017.03.064>>.

FATHI, SOHEIL; SRINIVASAN, R.; FENNER, A.; FATHI, SAHAND. Machine learning applications in urban building energy performance forecasting: A systematic review. **Renewable and Sustainable Energy Reviews**, v. 133, n. July 2019, p. 110287, 2020. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.rser.2020.110287>>.

FRANCO, A. J.; DUPUY, D.; ROUSTANT, O.; et al. Package ‘DiceDesign’, 2018.

GARUD, S. S.; KARIMI, I. A.; KRAFT, M. Design of computer experiments: A review. **Computers & Chemical Engineering**, v. 106, p. 71–95, 2017. Elsevier Ltd. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135417302090>>.

GROTENHUIS, M.; PELZER, B.; EISINGA, R.; et al. When size matters: advantages of weighted effect coding in observational studies. **International Journal of Public Health**, v. 62, n. 1, p. 163–167, 2017. Springer International Publishing.

HENSEN, J. L. M.; LAMBERTS, R. Building Performance Simulation for Design and Operation, 2019.

HONG, T.; LANGEVIN, J.; SUN, K. Building Simulation : Ten Challenges Building simulation : Ten challenges. March, 2018.

HYGH, J. S.; DECAROLIS, J. F.; HILL, D. B.; RANJI RANJITHAN, S. Multivariate regression as an energy assessment tool in early building design. **Building and Environment**, v. 57, p. 165–175, 2012. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.buildenv.2012.04.021>>.

HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. **International Journal of Forecasting**, v. 22, n. 4, p. 679–688, 2006.

IEA. The Future of cooling. Opportunities for energy-efficient air conditioning. , 2018. International Energy Agency. Disponível em: <www.iea.org>.

INMETRO. Instrução Normativa Inmetro para a Classificação de Eficiência Energética de Edificações Comerciais, de Serviços e Públicas (INI-C). , p. 159, 2020. Disponível em: <https://labeee.ufsc.br/sites/default/files/documents/INIC_Dez_2020.pdf>.

JAIN, R. K.; SMITH, K. M.; CULLIGAN, P. J.; TAYLOR, J. E. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance

accuracy. **Applied Energy**, v. 123, p. 168–178, 2014. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2014.02.057>>.

KESKAR, N. S.; MUDIGERE, D.; NOCEDAL, J.; SMELYANSKIY, M.; TANG, P. T. P. On Large-Batch Training for Deep Learning: Generalization Gap and Sharp Minima. , p. 1–16, 2017. Disponível em: <<http://arxiv.org/abs/1609.04836>>.

KLEIJNEN, J. P. C.; SANCHEZ, S. M.; LUCAS, T. W.; CIOPPA, T. M. A user's guide to the brave new world of designing simulation experiments. **INFORMS Journal on Computing**, v. 17, n. 3, p. 263–289, 2005.

KRESE, G.; PREK, M.; BUTALA, V. Incorporation of latent loads into the cooling degree days concept. **Energy and Buildings**, v. 43, n. 7, p. 1757–1764, 2011.

KUHN, M. Building Predictive Models in R Using the caret Package. **Journal of Statistical Software**, v. 28, n. 5, 2008.

LAM, J. C.; CHAN, R. Y. C.; LI, D. H. W. Simple regression models for fully air-conditioned public sector office buildings in subtropical climates. **Architectural Science Review**, v. 45, n. 4, p. 361–369, 2002.

LAM, J. C.; YANG, L.; LIU, J. Development of passive design zones in China using bioclimatic approach. **Energy Conversion and Management**, v. 47, n. 6, p. 746–762, 2006.

LIMA, I.; SCALCO, V.; LAMBERTS, R. Estimating the impact of urban densification on high-rise office building cooling loads in a hot and humid climate. **Energy and Buildings**, v. 182, p. 30–44, 2019. Elsevier B.V. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/S0378778818326276>>.

LOPES, M.; LAMBERTS, R. Development of a Metamodel to Predict Cooling Energy Consumption of HVAC Systems in Office Buildings in Different Climates. **Sustainability**, v. 10, n. 12, p. 4718, 2018. Disponível em: <<http://www.mdpi.com/2071-1050/10/12/4718>>.

LOPES, M. N. Desenvolvimento de metamodelo para a determinação do consumo anual em climatização aplicado a edifícios de escritórios com sistemas do tipo água gelada em diferentes climas. Tese (Doutorado) — Universidade Federal de Santa Catarina, 2019.

LOYOLA, DIEGO G.R.; PEDERGNANA, M.; GARCÍA, S. G. Smart sampling and incremental function learning for very large high dimensional data. **Neural Networks**, v. 78, p. 75–87, 2016. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.neunet.2015.09.001>>.

MACHADO, R. M. E S. Modelagem do impacto da ilha de calor sobre o desempenho energético de escritórios condicionados artificialmente. Dissertação (Mestrado) — Universidade Federal de Santa Catarina, 2019.

MARKOVIC, R.; TREECK, C. VAN. Learning short-term past as predictor of human behavior in commercial buildings. , , n. September, 2018.

MAZZAFERRO, L.; MACHADO, R. M. S.; MELO, A. P.; LAMBERTS, R. Do we need building performance data to propose a climatic zoning for building energy efficiency regulations? **Energy and Buildings**, v. 225, p. 110303, 2020. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.enbuild.2020.110303>>.

MELO, A. P. Desenvolvimento de um método para estimar o consumo de energia de edificações comerciais através da aplicação de redes neurais. Tese (Doutorado) — Universidade Federal de Santa Catarina, 2012.

MELO, A. P.; VERSAGE, R. S.; SAWAYA, G.; LAMBERTS, R. A novel surrogate model to support building energy labelling system: A new approach to assess cooling energy demand in commercial buildings. **Energy and Buildings**, v. 131, p. 233–247, 2016. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.enbuild.2016.09.033>>.

MENA, R.; RODRÍGUEZ, F.; CASTILLA, M.; ARAHAL, M. R. A prediction model based on neural networks for the energy consumption of a bioclimatic building. **Energy and Buildings**, v. 82, p. 142–155, 2014. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.enbuild.2014.06.052>>.

MILLER, C.; ARJUNAN, P.; KATHIRGAMANATHAN, A.; et al. The ASHRAE Great Energy Predictor III competition: Overview and results. **Science and Technology for the Built Environment**, v. 0, n. 0, p. 1–21, 2020. Taylor & Francis. Disponível em: <<http://arxiv.org/abs/2007.06933>>.

MIRZAEI, P. A. Recent challenges in modeling of urban heat island. **Sustainable Cities and Society**, v. 19, p. 200–206, 2015. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.scs.2015.04.001>>.

MOHANDÉS, S. R.; ZHANG, X.; MAHDIYAR, A. A comprehensive review on the application of artificial neural networks in building energy analysis. **Neurocomputing**, v. 340, p. 55–75, 2019. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.neucom.2019.02.040>>.

MONTGOMERY, D. C. **Design and analysis of experiments**. 2013.

NEGENDAHL, K. Building performance simulation in the early design stage: An introduction to integrated dynamic models. **Automation in Construction**, v. 54, p. 39–53, 2015. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.autcon.2015.03.002>>.

ØSTERGÅRD, T.; JENSEN, R. L.; MAAGAARD, S. E. Building simulations supporting decision making in early design - A review. **Renewable and Sustainable Energy Reviews**, v. 61, p. 187–201, 2016. Elsevier. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2016.03.045>>.

ØSTERGÅRD, T.; JENSEN, R. L.; MAAGAARD, S. E. Early Building Design : Informed decision-making by exploring multidimensional design space using sensitivity analysis. **Energy & Buildings**, v. 142, p. 8–22, 2017. Elsevier B.V. Disponível em: <<http://dx.doi.org/10.1016/j.enbuild.2017.02.059>>.

ØSTERGÅRD, T.; JENSEN, R. L.; MAAGAARD, S. E. A comparison of six metamodeling techniques applied to building performance simulations. **Applied Energy**, v. 211, n. October 2017, p. 89–103, 2018.

PERRIER, H.; COEURJOLLY, D.; XIE, F.; et al. Sequences with low-discrepancy blue-noise 2-D projections. **Computer Graphics Forum**, v. 37, n. 2, p. 339–353, 2018.

RACKES, A.; PAULA, A.; LAMBERTS, R.; MELO, A. P.; LAMBERTS, R. Naturally comfortable and sustainable: Informed design guidance and performance labeling for passive commercial buildings in hot climates. **Applied Energy**, v. 174, p. 256–274, 2016. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.apenergy.2016.04.081>>.

ROMAN, N. D.; BRE, F.; FACHINOTTI, V. D.; LAMBERTS, R. Application and characterization of metamodels based on artificial neural networks for building performance simulation: A systematic review. **Energy and Buildings**, v. 217, 2020. Elsevier B.V.

RORIZ, M. Uma proposta de revisão do zoneamento bioclimático brasileiro. **Antac**, 2012. ANTAC: Associação Nacional de Tecnologia do Ambiente Construído. Disponível em: <http://www.dt.co.kr/contents.html?article_no=2012071302010531749001>.

RORIZ, M. Classificação De Climas Do Brasil - Versão 3.0., 2014. ANTAC: Associação Nacional de Tecnologia do Ambiente Construído.

DE ROSA, M.; BIANCO, V.; SCARPA, F.; TAGLIAFICO, L. A. Historical trends and current state of heating and cooling degree days in Italy. **Energy Conversion and Management**, v. 90, p. 323–335, 2015. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.enconman.2014.11.022>>.

ROSSI, M. M.; OLIVEIRA FAVRETTO, A. P.; GRASSI, C.; et al. Metamodels to assess the thermal performance of naturally ventilated, low-cost houses in Brazil. **Energy and Buildings**, v. 204, p. 23–31, 2019. Elsevier B.V.

SALAKEN, S. M.; KHOSRAVI, A.; NGUYEN, T.; NAHAVANDI, S. Seeded transfer learning for regression problems with deep learning. **Expert Systems with Applications**, v. 115, p. 565–577, 2019. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.eswa.2018.08.041>>.

SERGEI KUCHERENKO, DANIEL ALBRECHT, A. S.; KUCHERENKO, S.; ALBRECHT, D.; SALTELLI, A. Exploring multi-dimensional spaces: a Comparison of Latin Hypercube and Quasi Monte Carlo Sampling Techniques. **The 8th IMACS Seminar on Monte Carlo Methods**, , n. May, p. 1–32, 2015.

SINGARAVEL, S.; GEYER, P.; SUYKENS, J. Component-Based Machine Learning Modelling Approach For Design Stage Building Energy Prediction: Weather Conditions And Size. **Proceedings of the 15th IBPSA Conference**, p. 2617–2626, 2017. Disponível em: <<https://lirias.kuleuven.be/handle/123456789/589886>>.

SINGARAVEL, S.; SUYKENS, J.; GEYER, P. Deep-learning neural-network architectures and methods: Using component-based models in building-design energy prediction.

Advanced Engineering Informatics, v. 38, n. May, p. 81–90, 2018. Elsevier. Disponível em: <<https://doi.org/10.1016/j.aei.2018.06.004>>.

SRIVASTAVA, N.; HINTON, G.; KRIZHEVSKY, A.; SUTSKEVER, I.; SALAKHUTDINOV, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. **Journal of Machine Learning Research**, v. 15, p. 1929–1958, 2014.

UNITED STATES DEPARTMENT OF ENERGY. EnergyPlus Version 8.9.0 Documentation: Engineering Reference. , p. 1716, 2018. Disponível em: <https://energyplus.net/sites/all/modules/custom/nrel_custom/pdfs/pdfs_v8.9.0/EngineeringReference.pdf>.

VERSAGE, R. Metamodelo para Estimar a Carga Térmica de Edificações Condicionadas Artificialmente, Tese (Doutorado) — Universidade Federal de Santa Catarina, 2015.

WALSH, A.; CÓSTOLA, D.; LABAKI, L. C. Validation of the climatic zoning defined by ASHRAE standard 169-2013. **Energy Policy**, v. 135, n. January, 2019.

WANG, Q.; AUGENBROE, G.; KIM, J.; GU, L. Meta-modeling of occupancy variables and analysis of their impact on energy outcomes of office buildings. , v. 174, p. 166–180, 2016.

WEI, Y.; ZHANG, X.; SHI, Y.; et al. A review of data-driven approaches for prediction and classification of building energy consumption. **Renewable and Sustainable Energy Reviews**, v. 82, n. August 2017, p. 1027–1047, 2018. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2017.09.108>>.

WERNCKE, L.; DELAVALD, A.; SMITH, P.; et al. Methodology for ranking controllable parameters to enhance operation of a steam generator with a combined Artificial Neural Network and Design of Experiments approach. **Energy and AI**, v. 3, p. 100040, 2021. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.egyai.2020.100040>>.

WESTERMANN, P.; EVINS, R. Surrogate modelling for sustainable building design – A review. **Energy and Buildings**, v. 198, p. 170–186, 2019. Elsevier B.V. Disponível em: <<https://doi.org/10.1016/j.enbuild.2019.05.057>>.

WESTERMANN, P.; WELZEL, M.; EVINS, R. Using a deep temporal convolutional network as a building energy surrogate model that spans multiple climate zones. **Applied Energy**, v. 278, n. April, p. 115563, 2020. Elsevier. Disponível em: <<https://doi.org/10.1016/j.apenergy.2020.115563>>.

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**, v. 30, n. 1, p. 79–82, 2005.

YAN, C. W.; YAO, J. Application of ANN for the prediction of building energy consumption at different climate zones with HDD and CDD. **Proceedings of the 2010 2nd International Conference on Future Computer and Communication, ICFCC 2010**, v. 3, n. Cdd, p. 286–289, 2010. IEEE.

YILDIZ, B.; BILBAO, J. I.; SPROUL, A. B. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. **Renewable and Sustainable Energy Reviews**, v. 73, n. March 2016, p. 1104–1122, 2017. Elsevier Ltd. Disponível em: <<http://dx.doi.org/10.1016/j.rser.2017.02.023>>.

YONDO, R.; ANDRÉS, E.; VALERO, E. A review on design of experiments and surrogate models in aircraft real-time and many-query aerodynamic analyses. **Progress in Aerospace Sciences**, v. 96, n. December 2017, p. 23–61, 2018.

YONG, S. G.; KIM, J.; CHO, J.; KOO, J. Meta-models for building energy loads at an arbitrary location. **Journal of Building Engineering**, v. 25, n. December 2018, p. 100823, 2019. Elsevier Ltd. Disponível em: <<https://doi.org/10.1016/j.job.2019.100823>>.

YOSINSKI, J.; CLUNE, J.; BENGIO, Y.; LIPSON, H. How transferable are features in deep neural networks? **Advances in Neural Information Processing Systems**, v. 4, n. January, p. 3320–3328, 2014.

APÊNDICE A- Códigos desenvolvidos e bases de dados

O Apêndice A fornece o acesso aos códigos de programação desenvolvidos para realizar os processos de amostragem, de modelagem, de simulação, de tratamento, de treinamento e de validação nesta pesquisa. Os códigos estão disponíveis no repositório do autor em: <<https://github.com/mazzaferro/tese>>, como indicado na Tabela 1.A.

Tabela 1.A. Códigos e descrição.

Nome do arquivo	Descrição
whole_building_generator.py	Gerador de IDFs
1_edit_filename.py	Editor para o nome de arquivos EPW
2_filename_folder.py	Nomeador de pastas para as simulações
3_change_idf_ground_temp.py	Editor de temperaturas do solo no IDF
discrepancy.R	Cálculo da discrepância
incremental_sampling_linux.R	Metamodelagem dinâmica

Fonte: elaborada pelo autor (2020).

As bases de dados climáticas elaboradas são acessíveis através do mesmo repositório e foram nomeadas segundo a Tabela 2.A.

Tabela 2.A. Bases de dados climáticas e descrição.

Nome do arquivo	Descrição
df_annual_mean.csv	Base de dados climáticos anuais
df_monthly_mean.csv	Base de dados climáticos mensais

Fonte: elaborada pelo autor (2020).

APÊNDICE B - Diversidade climática na RNA-C

Para analisar a RNA-C, e suas eventuais limitações, é necessário compreender a elaboração da base de dados utilizada em sua metamodelagem. O Apêndice B explorou a diversidade de dados climáticos utilizados para desenvolver a RNA-C.

A base de dados da RNA-C foi composta por amostras independentes de 11340 combinações de variáveis relacionadas à edificação, simuladas em 33 cidades brasileiras, num total aproximado de 374 mil simulações. A descrição das simulações foi organizada em 24 variáveis relacionadas à edificação, e em oito variáveis climáticas.

A maior parte das 33 localidades foi selecionada de modo a representar os principais centros urbanos brasileiros, sem considerar técnicas de amostragem ou classificações climáticas. Um fator que contribuiu para aumentar a cobertura de combinações, foi a inclusão estratégica de climas extremos (dentro do contexto brasileiro) entre os 33 climas. Os climas adicionais beiraram os limites inferiores e superiores das variáveis climáticas da RNA-C, como exposto na Tabela 3.B.

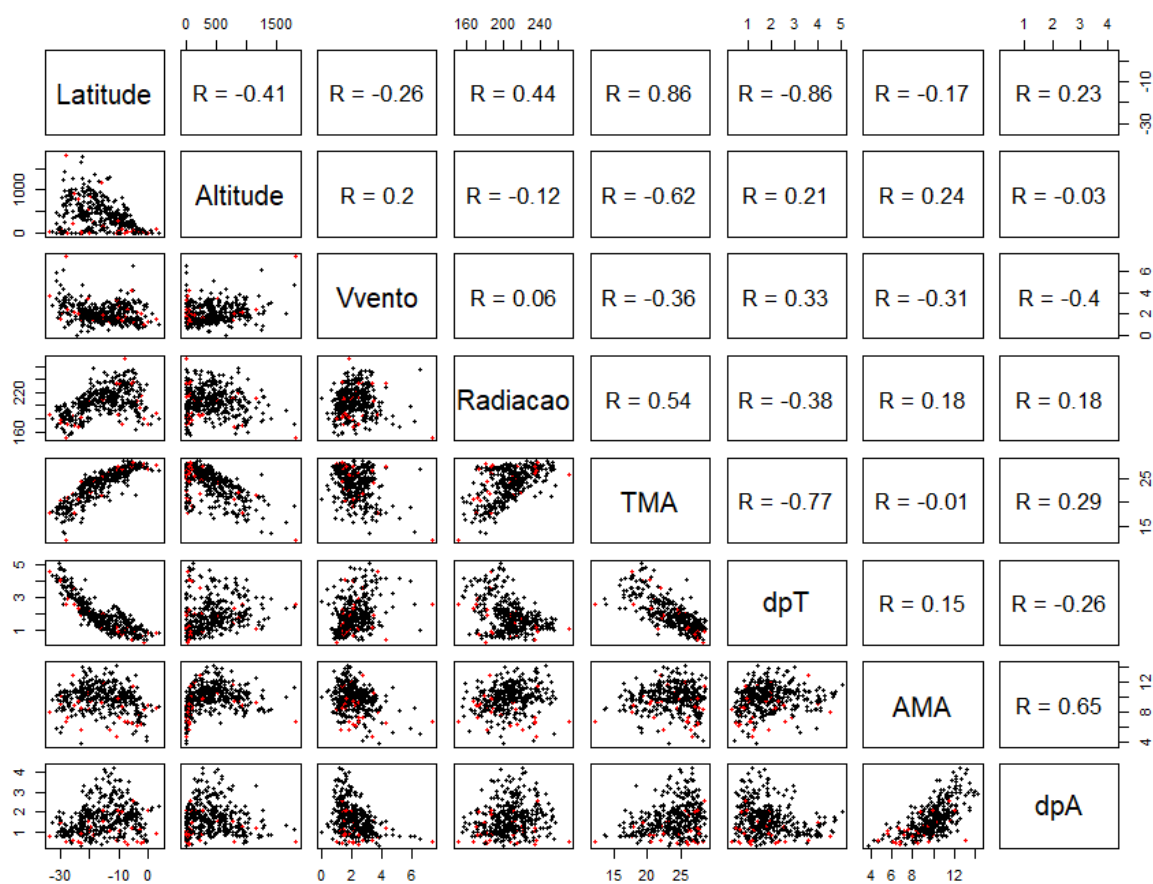
Tabela 3.B. Limites inferiores e superiores em cada variável adotada

Origem	Variável	Climas RNA-C (2016)				Arquivos climáticos INMET (2016)				Diferença entre limites	
		Mínimo		Máximo		Mínimo		Máximo		Mínimo	Máximo
		Valor	Localidade	Valor	Localidade	Valor	Localidade	Valor	Localidade		
INMET (2016)	Latitude	-33.74	Chuí	2.80	Boa Vista	-33.74	Chuí	3.81	Oiapoque	0.0%	2.7%
	Altitude	2.00	Florianópolis	1810	Urubici	0.00	Duque de Caxias	1810	Urubici	0.1%	0.0%
	Vvento	0.99	Belém	7.38	Urubici	0.00	Nova Tebas	7.38	Urubici	13.4%	0.0%
	Radiacao	151.37	Urubici	271.10	Recife	151.37	Urubici	271.10	Recife	0.0%	0.0%
RORIZ (2014)	TMA	10.88	Urubici	29.27	Teresina	10.88	Urubici	30.01	Picos	0.0%	3.9%
	dpT	0.32	Belém	4.86	Foz do Iguaçu	0.32	Belém	5.98	Quaraí	0.0%	19.8%
	AMA	2.49	Urubici	9.58	Teresina	2.49	Urubici	11.06	Itapaci	0.0%	17.3%
	dpA	0.26	Belém	2.48	Foz do Iguaçu	0.21	Arraial do Cabo	3.41	Quaraí	1.6%	29.1%

Fonte: elaborada pelo autor (2020).

A latitude, a altitude, a média anual de velocidade do vento (“Vvento”) e de radiação horizontal global (“Radiacao”) compõem as quatro primeiras variáveis climáticas. As outras quatro variáveis (“TMA, dpT, AMA e dpA”) foram calculadas a partir dos arquivos climáticos, segundo as fórmulas apresentadas por Roriz (2014) em sua proposta de zoneamento climático. Para visualizar a distribuição das 33 localidades dentre as 411 disponíveis, elaborou-se a matriz de dispersão apresentada na Figura 1.B.

Figura 1.B. Matriz de dispersão dos dados de arquivos EPW - versão INMET 2016.



Fonte: elaborada pelo autor (2020).

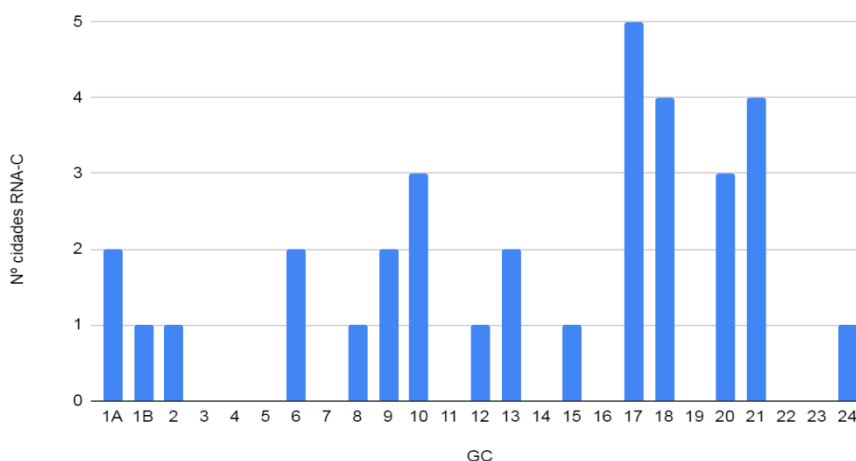
A diagonal da matriz de dispersão apresenta os nomes originais das oito variáveis climáticas que descrevem os 33 climas presentes na base de dados da RNA-C. Na parte inferior da Figura 1.B, os gráficos de dispersão fornecem uma noção do espaço amostral coberto pelos 33 climas, em relação ao total de localidades INMET. Os pontos em vermelho destacam os 33 climas que foram utilizados nas simulações que embasaram a RNA-C, enquanto os pontos em preto representam os 378 climas não utilizados. Apesar de representarem cerca de 8% dos 411 arquivos INMET, e de não terem sido escolhidos a partir

de uma amostragem estruturada, os 33 climas escolhidos estão razoavelmente distribuídos no espaço amostral.

A descrição climática por oito variáveis confere um caráter bastante exclusivo e quase categórico para as 33 combinações climáticas descritas, o que prejudicou a acurácia das predições em climas diferentes. Devido a esta limitação, a aplicação da RNA-C foi limitada a fornecer valores de carga térmica de refrigeração anual (CgTR) somente nas 33 localidades simuladas.

Inspirados na classificação proposta por Roriz (2014), 25 grupos climáticos orientam a aplicação das regulamentações de eficiência energética em edificações no Brasil (BAVARESCO et al., 2017). A mesma proposta, posteriormente, embasou a classificação de climas adotada pela INI-C. Posteriormente, parte dos climas cobertos pela RNA-C foram escolhidos como cidades representativas de grupos climáticos. A distribuição das localidades descritas na RNA-C, por grupo climático, foi apresentada na Figura 2.B.

Figura 2.B. Número de cidades descritas na RNA-C por grupo climático.



Fonte: elaborada pelo autor (2020).

A distribuição resultante evidencia uma representação maior de climas quentes (GC 17-24) em relação aos climas mais frios (GC 1-8). Idealmente, as localidades descritas na base de dados da RNA-C deveriam estar distribuídas igualmente entre os GC, porém, o desenvolvimento da RNA-C foi anterior à implementação dos GC. Aos GC que não foram representados, atribuiu-se a cidade representativa do GC com maior semelhança, de forma a garantir que a RNA-C pudesse atender todos os GC. Entretanto, a adoção de cidades representativas assume pouca divergência entre os dados climáticos das cidades de um

mesmo GC, o que pode não ser verdadeiro e configurar outra limitação para a aplicação da RNA-C.

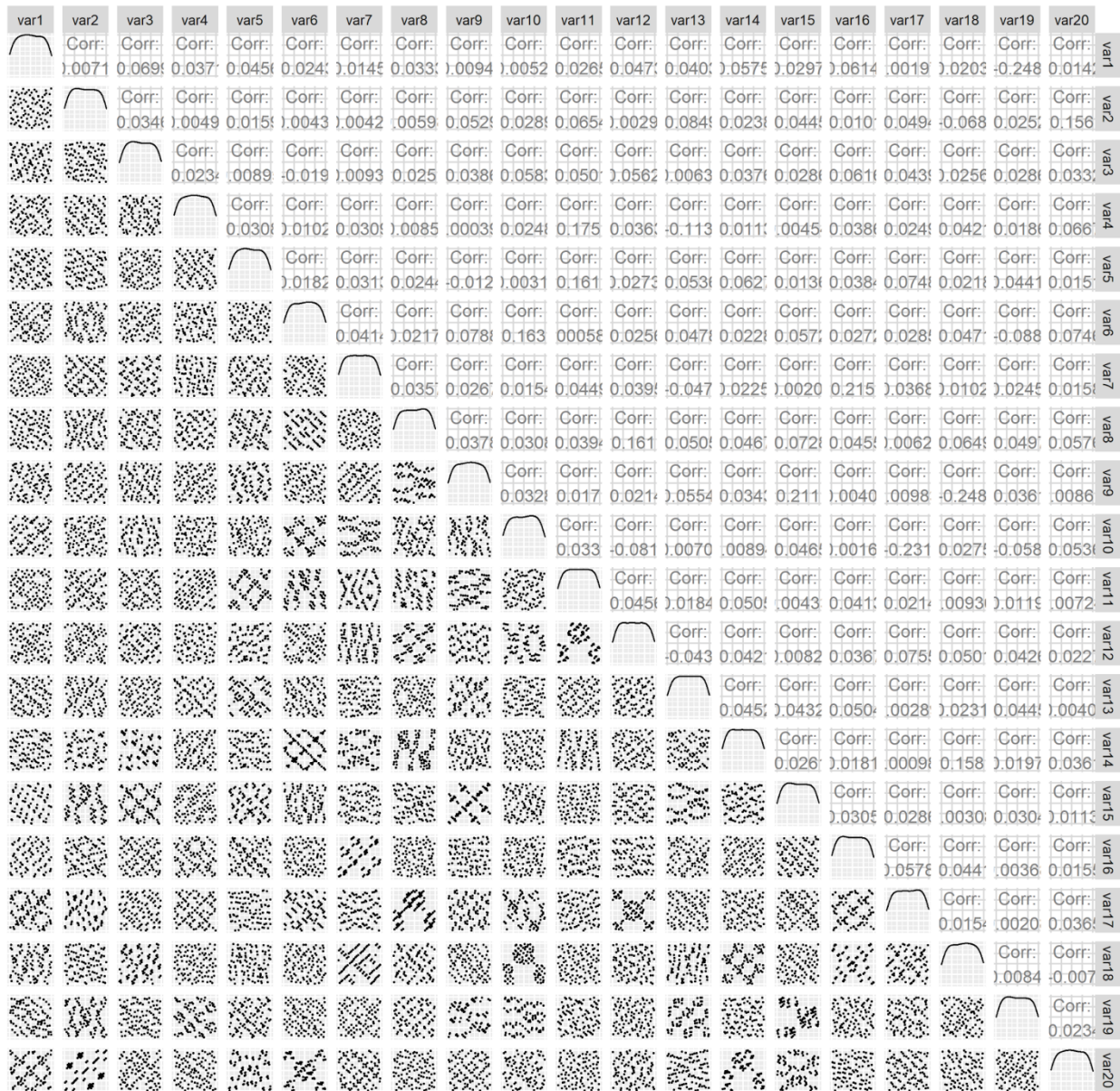
Por fim, é preferível realizar as descrições climáticas a partir de variáveis climáticas presentes nos arquivos climáticos, uma vez que os arquivos climáticos INMET têm passado por atualizações progressivas nos últimos anos. Apesar de suas limitações, a RNA-C ainda é um dos poucos metamodelos desenvolvidos a partir de simulações energéticas que consideram múltiplos climas.

APÊNDICE C - Preenchimento do espaço amostral

O Apêndice C contém as representações gráficas de amostras realizadas a partir de três sequências de baixa discrepância (Sobol, Torus e Halton) e de três tipos de amostragem por Hipercubo Latino (LHS melhorada, LHS otimizada e LHS randômica).

Figura 3.C. Amostragem de 100 pontos de acordo com a técnica de Sobol.

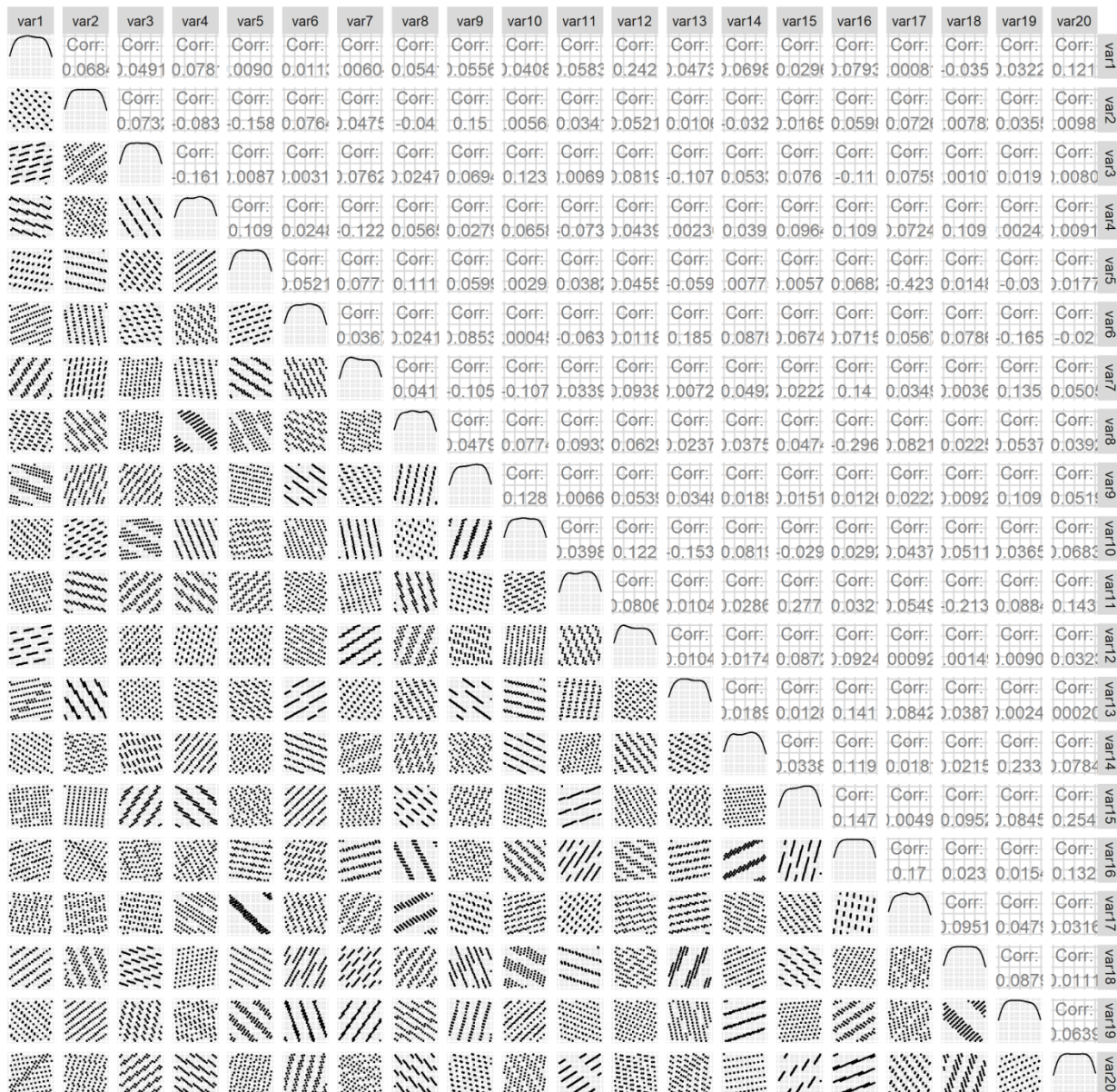
Sobol / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

Figura 4.C. Amostragem de 100 pontos de acordo com a técnica de Torus.

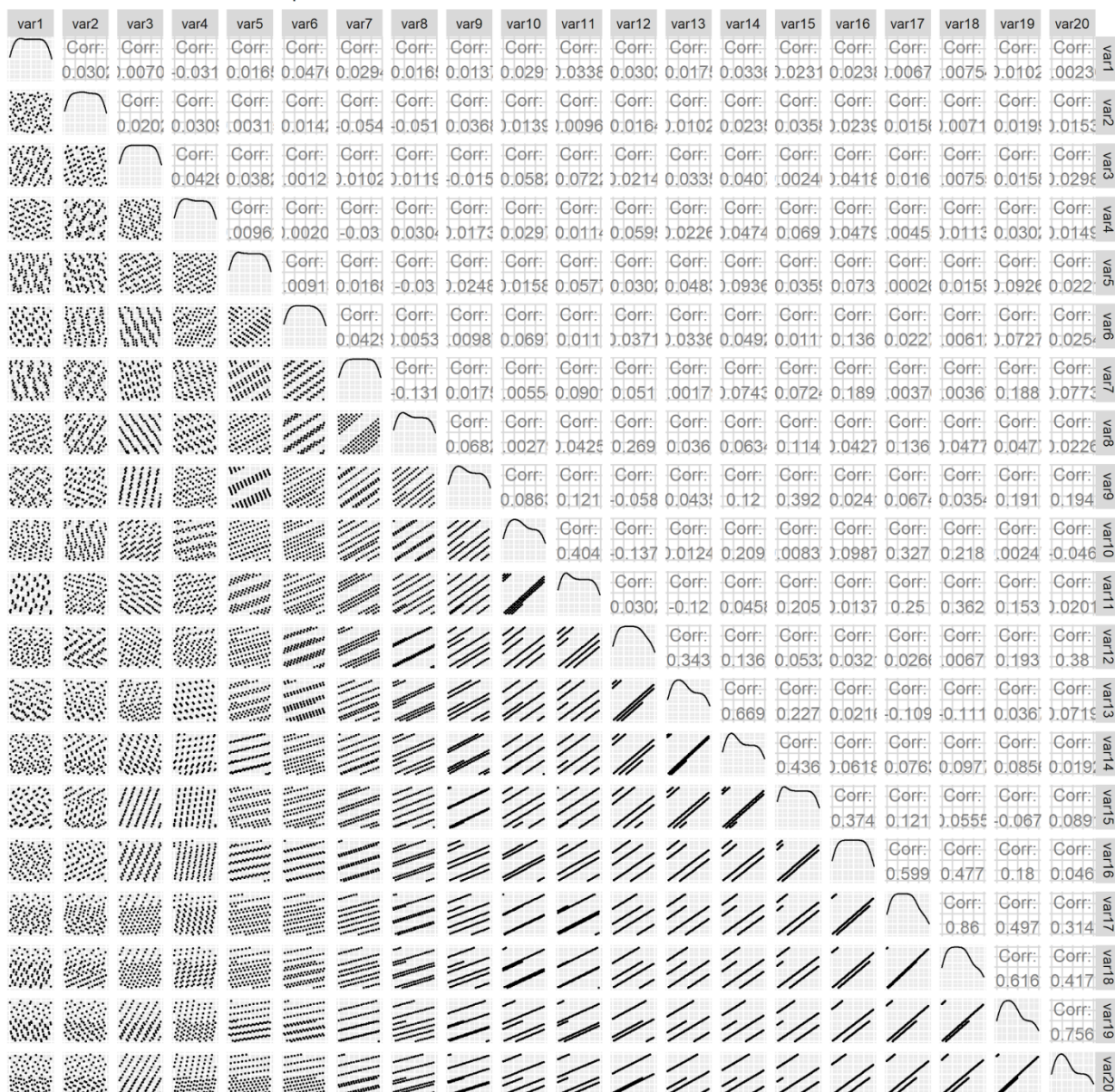
Torus / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

Figura 5.C. Amostragem de 100 pontos de acordo com a técnica de Halton.

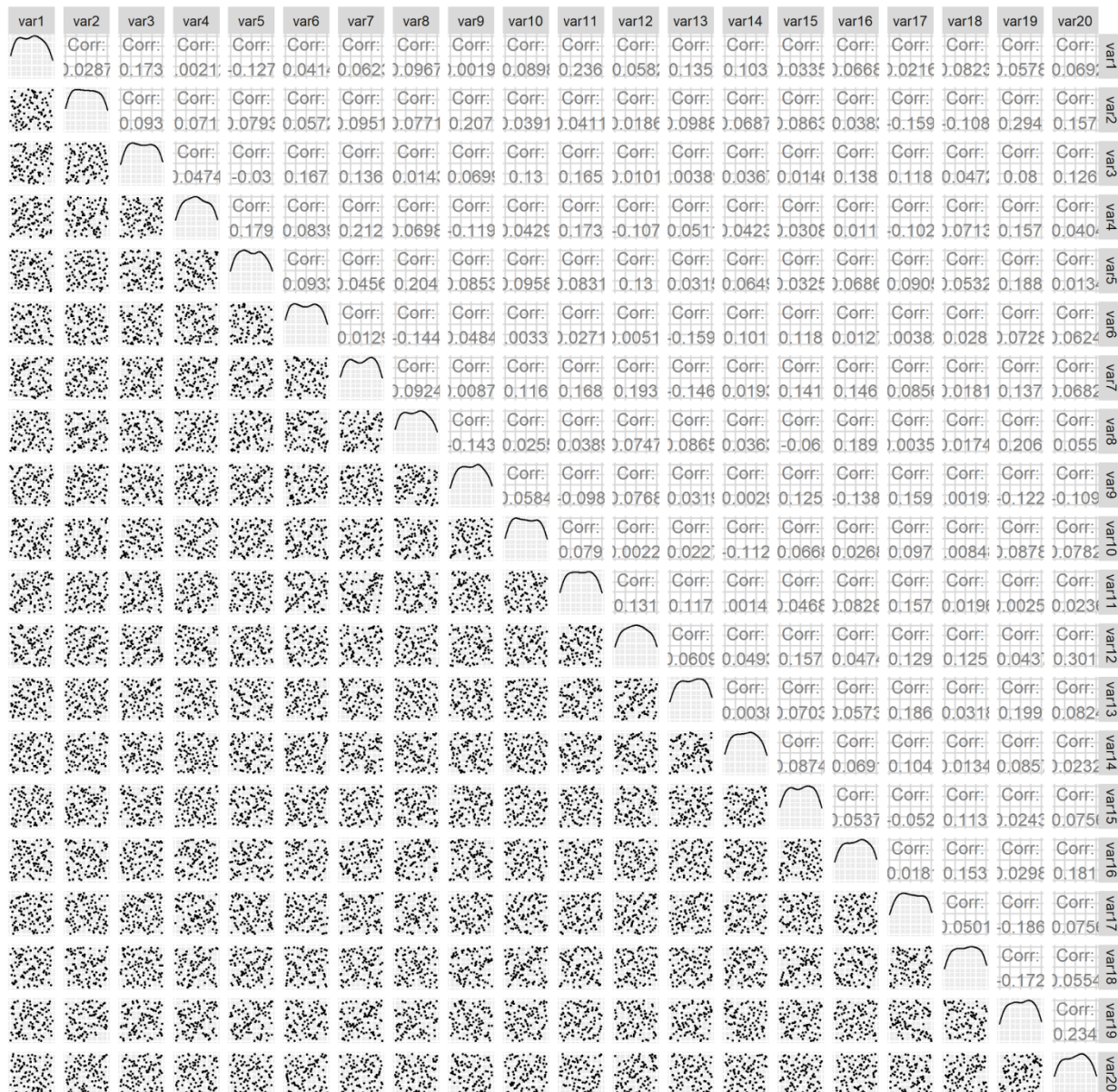
Halton / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

Figura 6.C. Amostragem de 100 pontos de acordo com a técnica de LHS melhorada.

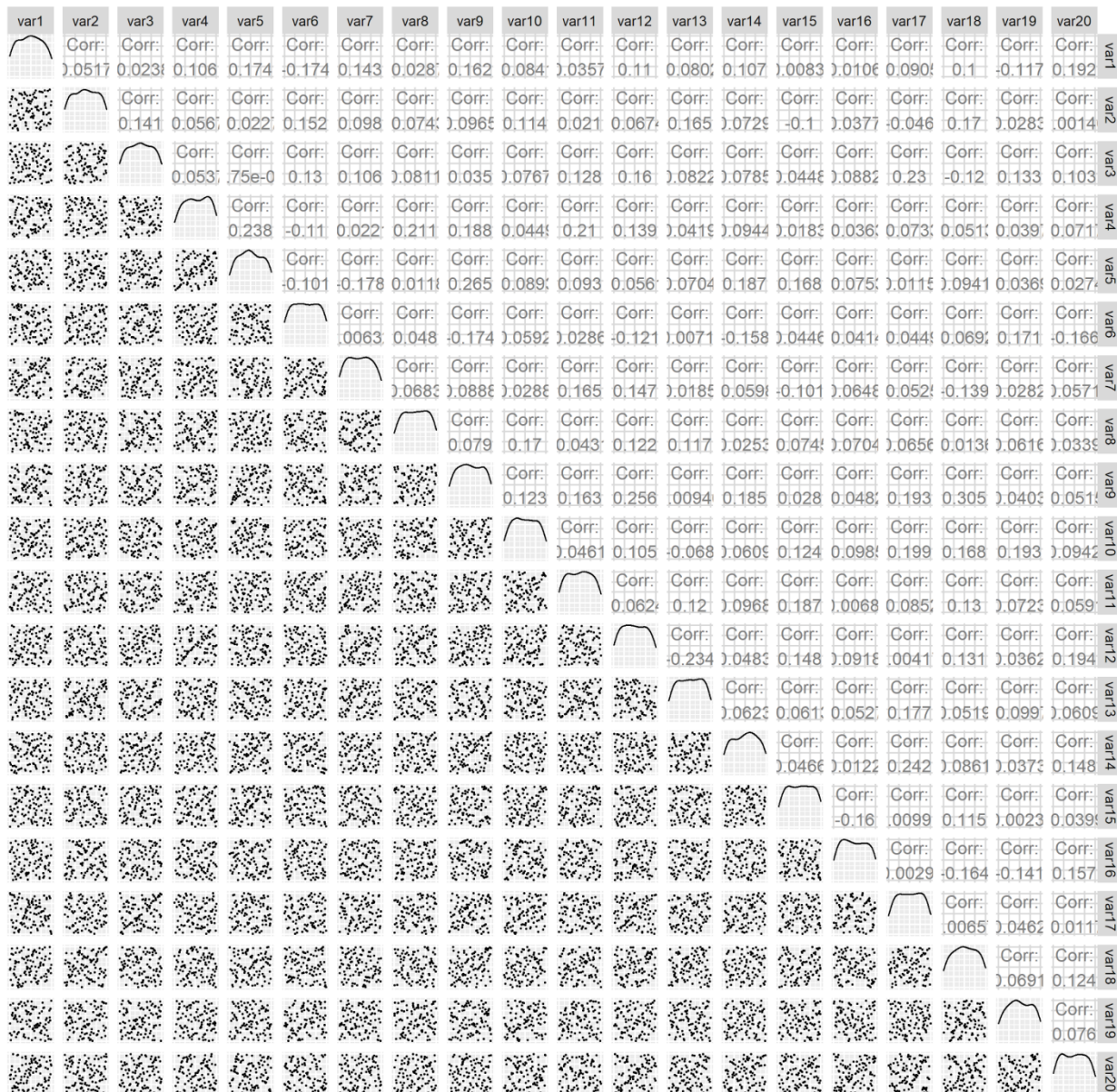
Improved LHS / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

Figura 7.C. Amostragem de 100 pontos de acordo com a técnica de LHS otimizada.

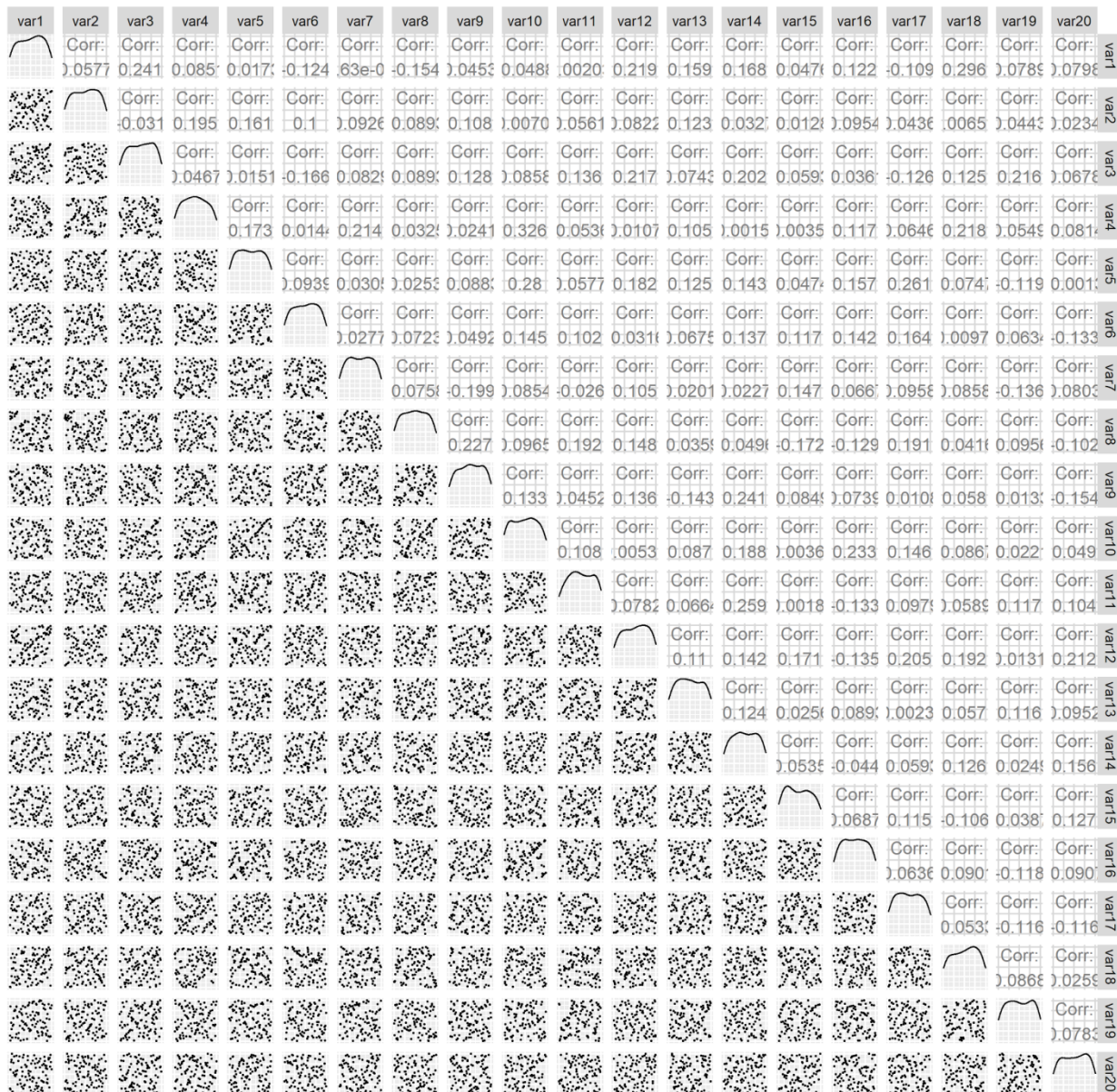
Optimum LHS / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

Figura 8.C. Amostragem de 100 pontos de acordo com a técnica de LHS randômica.

Random LHS / iteration 4 / number of points: 00100



Fonte: elaborada pelo autor (2020).

APÊNDICE D - Valores médios das médias mensais de temperatura de bulbo seco

O Apêndice D apresenta os valores médios das médias mensais de temperatura de bulbo seco (MMDBT), e o desvio padrão das médias mensais de temperatura de bulbo seco (SDMDBT), para os 411 arquivos climáticos analisados. Os valores de MMDBT foram utilizados para a amostragem dos climas ao longo da metamodelagem dinâmica.

A temperatura de bulbo seco foi priorizada, sobre a temperatura de ponto de orvalho e sobre a radiação horizontal global, por apresentar maiores correlações com a carga térmica de refrigeração das zonas térmicas das edificações de escritórios condicionados artificialmente.

A Tabela 4.D apresenta os valores médios das médias mensais de temperatura de bulbo seco, acompanhados pelos valores de desvio padrão das médias mensais de temperatura de bulbo seco, de cada arquivo climático analisado.

Tabela 4.D. Valores médios das médias mensais de temperatura de bulbo seco, e os valores de desvio padrão das médias mensais de temperatura de bulbo seco, de acordo com o clima.

ID	MMDBT	SDMDBT
BRA_AC_Feijo.AP.819240_INMET	26.08089483	0.84077554
BRA_AC_Parque.Chandless.819630_INMET	24.86928846	0.79829832
BRA_AC_Porto.Walter.819210_INMET	26.36708133	0.990891929
BRA_AC_Rio.Branco.829150_INMET	25.11143412	1.056330419
BRA_AL_Arapiraca.AP.819960_INMET	24.84862343	1.485425174
BRA_AL_Coruripe.866190_INMET	25.72631678	1.220866438
BRA_AL_Maceio.829940_INMET	25.52677662	1.210726091
BRA_AL_Palmeira.dos.Indios.829920_INMET	25.28462774	1.759673168
BRA_AL_Pao.de.Acucar.819940_INMET	27.2896184	2.222930309
BRA_AL_Sao.Luis.do.Quitunde.819970_INMET	25.42607364	1.175896453
BRA_AM_Autazes.817320_INMET	26.50106767	0.485623582
BRA_AM_Barcelos.AP.816480_INMET	26.27079421	0.642096113
BRA_AM_Coari.817700_INMET	26.24949895	0.513371527
BRA_AM_Humaita.AP.818900_INMET	26.28367066	0.612423153
BRA_AM_Itacoatiara.AP.817330_INMET	26.42436537	0.576942243
BRA_AM_Manacapuru.817290_INMET	26.69575829	0.661392568
BRA_AM_Manauas-Gomez.Intl.AP.817300_INMET	27.27682388	0.763514604
BRA_AM_Maues.817340_INMET	26.56597985	1.04836725
BRA_AM_Parintins.817030_INMET	27.48324858	1.444565539

BRA_AM_Presidente.Figueiredo.816990_INMET	26.63952349	1.053304822
BRA_AM_Urucara.AP.817020_INMET	27.45421704	1.389242005
BRA_AP_Macapa-Alcolumbre.Intl.AP.820980_INMET	25.77418893	0.969088936
BRA_AP_Oiapoque.816090_INMET	25.55771764	0.790754414
BRA_AP_Porto.Grande.816370_INMET	25.77418893	0.969088936
BRA_AP_Tartarugalzinho.816280_INMET	25.98717691	0.850716347
BRA_BA_Amargosa.AP.866750_INMET	23.18909	1.790538559
BRA_BA_Barra.AP.866340_INMET	26.7998941	1.428432363
BRA_BA_Barreiras.866520_INMET	25.11940996	1.199862032
BRA_BA_Belmonte.867440_INMET	23.79265345	1.787926723
BRA_BA_Bom.Jesus.da.Lapa.AP.866720_INMET	26.25417344	1.553520522
BRA_BA_Brumado.866960_INMET	25.17475468	2.221114056
BRA_BA_Buritirama.AP.866090_INMET	25.45548272	1.684468195
BRA_BA_Caravelas.834970_INMET	25.03127638	1.610815944
BRA_BA_Conde.866390_INMET	25.9404648	1.329504767
BRA_BA_Correntina.866710_INMET	23.82791693	1.838575357
BRA_BA_Cruz.das.Almas.866570_INMET	23.93467945	1.870183513
BRA_BA_Delfino.866100_INMET	23.8471419	1.397637489
BRA_BA_Euclides.da.Cunha.866130_INMET	24.1253337	2.402625447
BRA_BA_Feira.de.Santana.866580_INMET	24.61195653	2.170761385
BRA_BA_Guanambi.AP.866940_INMET	25.6642771	1.48905416
BRA_BA_Ibotirama.866530_INMET	26.73275546	1.642193562
BRA_BA_Ilheus.866990_INMET	23.45848452	1.529417946
BRA_BA_Ipiauí.866980_INMET	24.57222337	1.542764356
BRA_BA_Irece.AP.866350_INMET	23.95092515	1.309726089
BRA_BA_Itaberaba.832440_INMET	24.93682686	1.632190017
BRA_BA_Itapetinga.867230_INMET	24.68780317	1.94194138
BRA_BA_Itirucu.866740_INMET	21.07243424	2.013077777
BRA_BA_Jacobina.866360_INMET	24.36748245	1.586707224
BRA_BA_Lençóis.866540_INMET	24.86814767	1.685444079
BRA_BA_Macajuba.866550_INMET	24.06445762	1.785431564
BRA_BA_Marau.866770_INMET	25.67791499	1.360494639
BRA_BA_Paulo.Afonso.829860_INMET	26.33581597	2.058582906
BRA_BA_Piata.866730_INMET	19.80694058	1.568492615
BRA_BA_Porto.Seguro.867450_INMET	23.75107428	1.647996796
BRA_BA_Queimadas.866120_INMET	26.14146849	1.919106146

BRA_BA_Remanso.829790_INMET	26.81721476	0.964476859
BRA_BA_Salvador.866780_INMET	25.86890686	1.21896405
BRA_BA_Santa.Rita.de.Cassia.866330_INMET	26.06336214	1.333250918
BRA_BA_Senhor.do.Bonfim.830880_INMET	23.87449949	2.171302039
BRA_BA_Serrinha.866370_INMET	24.16192855	2.293105794
BRA_BA_Uaua.819920_INMET	25.41940076	1.785616103
BRA_BA_Una.867240_INMET	24.16082717	1.383892516
BRA_BA_Valenca.866760_INMET	24.21265372	1.432753628
BRA_BA_Vitoria.da.Conquista-Figueiredo.AP.866970_INMET	20.60583949	1.708766392
BRA_CE_Acarau.817550_INMET	26.6138272	0.900910512
BRA_CE_Barbalha.819110_INMET	25.59007518	1.325287448
BRA_CE_Campos.Sales.827770_INMET	24.86276556	1.74093074
BRA_CE_Crateus.818300_INMET	27.10720409	1.844430411
BRA_CE_Fortaleza-Pinto.Martins.Intl.AP.817580_INMET	26.65419115	0.757810955
BRA_CE_Guaramiranga.824870_INMET	20.69029021	0.50931744
BRA_CE_Iguatu.818730_INMET	27.02626467	1.422101436
BRA_CE_Itapipoca.817560_INMET	26.89055553	0.999634736
BRA_CE_Jaguaribe.818330_INMET	27.77716419	1.545326895
BRA_CE_Jaguaruana.824930_INMET	27.25360682	1.053212867
BRA_CE_Morada.Nova.825940_INMET	27.1366324	0.980650853
BRA_CE_Quixeramobim.825860_INMET	26.43609442	1.482521919
BRA_CE_Taua.818720_INMET	26.1918926	1.993931868
BRA_DF_Brasilia.867150_INMET	21.09254603	0.978373131
BRA_DF_Planaltina.867160_INMET	20.96112791	1.214962243
BRA_ES_Alegre.868280_INMET	23.96260198	2.258736145
BRA_ES_Alfredo.Chaves.868290_INMET	23.59051569	1.847576331
BRA_ES_Linhares.868050_INMET	24.28288306	2.005446737
BRA_ES_Nova.Venecia.867850_INMET	24.26917299	1.864971567
BRA_ES_Presidente.Kennedy.868530_INMET	23.9841544	2.134200027
BRA_ES_Santa.Teresa.868040_INMET	18.59410266	1.869949216
BRA_ES_Sao.Mateus.867860_INMET	23.86923062	2.176453236
BRA_ES_Vitoria.868300_INMET	24.40370864	2.072168303
BRA_GO_Alto.Paraiso.de.Goias.866910_INMET	20.59222956	1.216696057
BRA_GO_Aragarcas.833740_INMET	25.79429646	1.338271845
BRA_GO_Caiaponia.867300_INMET	23.25559852	1.206149268
BRA_GO_Catalao.867770_INMET	22.50863895	1.399964937

BRA_GO_Cristalina.867370_INMET	20.73175153	1.343226814
BRA_GO_Goianesia.867130_INMET	24.53854124	1.137050437
BRA_GO_Goiania.834230_INMET	23.23781885	1.667544817
BRA_GO_Goias.867120_INMET	24.97969393	1.042898514
BRA_GO_Itapaci.866890_INMET	23.90512582	1.621904857
BRA_GO_Itumbiara.867740_INMET	23.92417264	1.608627523
BRA_GO_Jatai.867520_INMET	22.41964096	1.744671874
BRA_GO_Luziania.867360_INMET	21.98860063	1.092307701
BRA_GO_Mineiros.867510_INMET	22.37798963	1.763943098
BRA_GO_Monte.Alegre.de.Goias.866700_INMET	24.96517964	1.196829464
BRA_GO_Morrinhos.867550_INMET	22.38687353	1.579195475
BRA_GO_Niquelandia.866900_INMET	24.31633475	1.166353778
BRA_GO_Parauna.867320_INMET	24.00106225	1.509354563
BRA_GO_Pires.do.Rio.867560_INMET	22.94882526	1.490217564
BRA_GO_Posse.866920_INMET	24.44711472	0.860205793
BRA_GO_Rio.Verde.867530_INMET	22.80245426	1.205126872
BRA_GO_Sao.Simao.867730_INMET	23.99158245	1.524019071
BRA_MA_Alto.Parnaiba.829700_INMET	25.74547512	0.950394981
BRA_MA_Bacabal.817900_INMET	27.27241791	1.201197386
BRA_MA_Balsas.819030_INMET	25.9465593	1.017286844
BRA_MA_Barra.do.Corda.818250_INMET	26.28092531	1.013292672
BRA_MA_Buriticupu.817880_INMET	26.19186148	1.071637657
BRA_MA_Carolina.819010_INMET	26.57599489	1.123436334
BRA_MA_Caxias.817920_INMET	26.3747563	1.136878223
BRA_MA_Chapadinha.817490_INMET	26.88832024	1.235510444
BRA_MA_Colinas.818660_INMET	26.0519953	0.909665241
BRA_MA_Estreito.818630_INMET	25.84397577	1.016320787
BRA_MA_Farol.Preguicas.817170_INMET	27.49267545	0.64800029
BRA_MA_Farol.Santana.822770_INMET	27.35370906	0.609460684
BRA_MA_Grajau.818230_INMET	26.12383998	0.918684517
BRA_MA_Imperatriz.818220_INMET	26.88542907	0.895423644
BRA_MA_Sao.Luis.817150_INMET	26.7484104	0.577079152
BRA_MA_Turiacu.821980_INMET	26.52040267	0.941276085
BRA_MG_Aguas.Vermelhas.867220_INMET	21.40349974	1.938610272
BRA_MG_Aimores.868030_INMET	24.68292107	1.981584728
BRA_MG_Almenara.867430_INMET	25.23039768	2.078563718

BRA_MG_Araxa.867960_INMET	21.41422824	1.549984889
BRA_MG_Belo.Horizonte-Pampulha.AP.868000_INMET	21.76060697	1.68976784
BRA_MG_Buritis.867180_INMET	22.29005838	1.403164784
BRA_MG_Caldas.868460_INMET	17.88428334	2.921934576
BRA_MG_Capelinha.867610_INMET	20.09374136	1.801967848
BRA_MG_Carangola.868230_INMET	22.04700106	2.69165513
BRA_MG_Caratinga.835920_INMET	21.78435135	2.123738996
BRA_MG_Chapada.Gaucha.867190_INMET	22.70861111	1.420154163
BRA_MG_Conceicao.das.Alagoas.867940_INMET	22.78059436	2.022900357
BRA_MG_Curvelo.867800_INMET	22.85638281	2.02946437
BRA_MG_Diamantina.AP.867810_INMET	18.63233378	1.688343857
BRA_MG_Dores.do.Indaia.867970_INMET	21.99670011	1.631076691
BRA_MG_Espinosa.866950_INMET	25.029106	1.561411826
BRA_MG_Florestal.867980_INMET	20.21183964	2.793908878
BRA_MG_Formiga.868200_INMET	20.93572341	1.865211803
BRA_MG_Governador.Valadares.835430_INMET	24.050185	2.044477843
BRA_MG_Guarda-Mor.867570_INMET	21.14098508	1.540988578
BRA_MG_Ibirite.868210_INMET	19.96476647	1.576843129
BRA_MG_Itaobim.867420_INMET	25.22964475	1.9139682
BRA_MG_Ituiutaba.867750_INMET	23.26941311	1.821594033
BRA_MG_Joao.Pinheiro.AP.867580_INMET	22.45869978	1.497290786
BRA_MG_Juiz.de.Fora.836920_INMET	18.60805892	1.61381422
BRA_MG_Mantena.867840_INMET	24.27364999	2.055188647
BRA_MG_Maria.da.Fe.830150_INMET	16.26890844	2.455801744
BRA_MG_Mocambinho.867200_INMET	24.84659245	1.663625702
BRA_MG_Montalvania.866930_INMET	23.7382745	1.885868214
BRA_MG_Monte.Verde.AP.868700_INMET	15.28670331	2.241655608
BRA_MG_Montes.Claros.867400_INMET	23.67826925	1.761858174
BRA_MG_Muriae.868520_INMET	23.26413666	2.236229389
BRA_MG_Ouro.Branco.868230_INMET	19.68777546	1.815113776
BRA_MG_Passa.Quatro.837370_INMET	18.30443391	2.30132834
BRA_MG_Passos.868190_INMET	20.68253864	1.976575488
BRA_MG_Patrocínio.867780_INMET	20.33206213	1.784927027
BRA_MG_Pirapora.AP.867590_INMET	23.73818039	1.809345819
BRA_MG_Rio.Pardo.de.Minas.867210_INMET	21.59040339	1.823345138
BRA_MG_Sacramento.867950_INMET	21.80697495	1.754417735

BRA_MG_Salinas.867410_INMET	23.28435249	2.047942787
BRA_MG_Sao.Joao.del.Rei.836880_INMET	19.37222561	2.306859686
BRA_MG_Sao.Romao.867390_INMET	24.03195831	2.234475604
BRA_MG_Serra.dos.Aimores.867630_INMET	23.26256718	1.894490199
BRA_MG_Teofilo.Otoni.867620_INMET	22.39144351	1.841187112
BRA_MG_Timoteo.868010_INMET	22.84667333	1.621235972
BRA_MG_Tres.Marias.867790_INMET	21.97885315	1.531666111
BRA_MG_Uberlandia.867760_INMET	22.95100492	1.559525877
BRA_MG_Unai.AP.834280_INMET	23.44272697	1.587598278
BRA_MG_Varginha.868480_INMET	19.56367365	1.915651041
BRA_MG_Vicosa.836420_INMET	20.00593214	2.278928184
BRA_MS_Amambai.868940_INMET	21.94225812	3.130437921
BRA_MS_Campo.Grande.Intl.AP.868100_INMET	23.44796331	1.811418157
BRA_MS_Chapadao.do.Sul.867720_INMET	22.39628659	1.458387623
BRA_MS_Corumba.Intl.AP.835520_INMET	25.92213334	2.773213397
BRA_MS_Coxim.867700_INMET	24.72361586	2.089963702
BRA_MS_Dourados.836590_INMET	22.68565679	2.483406642
BRA_MS_Ivinhema.868600_INMET	23.69792005	2.563539786
BRA_MS_Juti.868590_INMET	22.90305732	3.318019866
BRA_MS_Miranda.868070_INMET	24.63205464	2.847287915
BRA_MS_Nhumirim.867680_INMET	24.89145727	2.499033364
BRA_MS_Ponta.Pora.868570_INMET	21.31366317	2.775239768
BRA_MS_Porto.Murtinho.AP.868330_INMET	25.46757243	3.985205509
BRA_MS_Rio.Brilhante.868360_INMET	23.2920588	2.615428058
BRA_MS_Sete.Quedas.868950_INMET	22.3447546	3.224586143
BRA_MS_Sidrolandia.868090_INMET	23.35354247	2.546733301
BRA_MS_Tres.Lagoas.AP.868130_INMET	24.35751941	2.353036315
BRA_MT_Agua.Boa.866860_INMET	24.94741469	0.928137481
BRA_MT_Alto.Taquari.867500_INMET	22.33504091	1.17584
BRA_MT_Apiacas.819760_INMET	25.64930518	0.783142535
BRA_MT_Campo.Novo.do.Parecis.866620_INMET	24.08198983	0.65237981
BRA_MT_Campo.Verde.867070_INMET	23.03333947	1.361404215
BRA_MT_Carlinda.819780_INMET	25.49017998	0.689769686
BRA_MT_Comodoro.866610_INMET	22.43624355	1.390349291
BRA_MT_Confresa.866060_INMET	26.12932966	0.80266868
BRA_MT_Cotriguacu.829270_INMET	24.62139081	0.755137525

BRA_MT_Cuiaba.867050_INMET	26.26789413	1.700065197
BRA_MT_Gaucha.do.Norte.866660_INMET	25.21685167	0.733224196
BRA_MT_Guaranta.do.Norte.819790_INMET	25.03847292	0.536950969
BRA_MT_Guiratinga.867290_INMET	24.69505227	1.505269315
BRA_MT_Itiquira.867480_INMET	24.03101817	1.402903157
BRA_MT_Juara.866250_INMET	25.88712995	0.943352256
BRA_MT_Juina.866240_INMET	24.84284405	0.692742432
BRA_MT_Nova.Maringa.866630_INMET	26.33523143	1.172519831
BRA_MT_Novo.Mundo.866430_INMET	24.76572365	1.29318435
BRA_MT_Paranatinga.866840_INMET	24.45516009	1.366780896
BRA_MT_Pontes.de.Lacerda.867020_INMET	25.33802958	1.504554881
BRA_MT_Porto.Estrela.867040_INMET	25.39924769	1.909046754
BRA_MT_Querencia.866470_INMET	25.02658909	0.689942577
BRA_MT_Rondonopolis.867280_INMET	25.49157445	1.111061585
BRA_MT_Salto.do.Ceu.867030_INMET	24.56653805	1.629699309
BRA_MT_Santo.Antonio.do.Leste.866850_INMET	23.68030703	1.003500562
BRA_MT_Sao.Felix.do.Araguaia.866280_INMET	26.22461198	0.688262311
BRA_MT_Sao.Jose.do.Rio.Claro.866640_INMET	24.78220555	1.022535305
BRA_MT_Sinop.866260_INMET	25.07103225	1.204539554
BRA_MT_Sorriso.866450_INMET	25.36158789	0.557876512
BRA_MT_Tangara.da.Serra.866820_INMET	24.62591918	0.99928268
BRA_MT_Vila.Bela.da.Santissima.Trindade.867010_INMET	24.81285226	1.774570273
BRA_PA_Belem.816800_INMET	27.15899146	0.370122073
BRA_PA_Braganca.816850_INMET	26.46625592	1.040895265
BRA_PA_Cameta.817100_INMET	27.21038909	0.834467476
BRA_PA_Castanhal.816820_INMET	26.20104129	0.574819853
BRA_PA_Conceicao.do.Araguaia.819400_INMET	26.76184118	1.076898087
BRA_PA_Itaituba.817780_INMET	26.85757622	0.736076906
BRA_PA_Maraba.825620_INMET	26.49434929	0.469681821
BRA_PA_Novo.Repartimento.817840_INMET	26.31459293	0.771103896
BRA_PA_Pacaja.817420_INMET	26.09527991	0.627738251
BRA_PA_Paragominas.817450_INMET	26.36612954	1.011333268
BRA_PA_Placas.817370_INMET	26.26274426	0.808513641
BRA_PA_Rondon.do.Para.817860_INMET	26.33300393	0.995534148
BRA_PA_Salinopolis.816600_INMET	27.1521376	0.935552423
BRA_PA_Santana.do.Araguaia.819820_INMET	26.32694858	0.708527274

BRA_PA_Serra.dos.Carajas.818600_INMET	23.65655669	1.243627706
BRA_PA_Soure.821410_INMET	28.07383267	0.430253825
BRA_PA_Tome.Acu.817110_INMET	25.99251672	0.771085164
BRA_PA_Tucurui.817430_INMET	27.35888366	0.493982294
BRA_PB_Areia.818770_INMET	22.47395377	1.072180545
BRA_PB_Cabaceiras.819150_INMET	25.37948189	1.480379294
BRA_PB_Camaratuba.818780_INMET	24.85519537	1.017062553
BRA_PB_Campina.Grande.819160_INMET	23.62588038	1.091942981
BRA_PB_Joao.Pessoa.819180_INMET	26.026897	1.174637742
BRA_PB_Monteiro.819140_INMET	24.24218881	1.661564817
BRA_PB_Patos.819130_INMET	27.24028871	1.450388345
BRA_PB_Sao.Goncalo.817740_INMET	26.83304956	1.229998192
BRA_PE_Arcoverde.819530_INMET	23.59191591	1.822926031
BRA_PE_Cabrobo.828860_INMET	26.78157293	1.849527245
BRA_PE_Caruaru.828950_INMET	23.34198967	1.312160552
BRA_PE_Floresta.819770_INMET	26.60490279	1.892969672
BRA_PE_Garanhuns.819550_INMET	21.05441399	1.835525996
BRA_PE_Ibimirim.819540_INMET	25.60417717	2.101493742
BRA_PE_Palmares.819570_INMET	25.05627584	1.303973771
BRA_PE_Petrolina.819910_INMET	26.91135014	1.557381833
BRA_PE_Recife.819580_INMET	25.76279045	1.174391055
BRA_PE_Serra.Talhada.819120_INMET	25.27170038	1.901203645
BRA_PE_Surubim.819170_INMET	24.22194316	1.379011854
BRA_PI_Alvorada.do.Gurgueia.818460_INMET	26.29735228	1.512979378
BRA_PI_Bom.Jesus.do.Piaui.829750_INMET	26.39083181	1.679875173
BRA_PI_Caracol.819880_INMET	24.71437921	1.923770425
BRA_PI_Castelo.do.Piaui.818290_INMET	27.33869362	1.60940293
BRA_PI_Esperantina.817510_INMET	26.97543429	1.24253759
BRA_PI_Floriano.818680_INMET	27.5727182	1.317772781
BRA_PI_Gilbues.819860_INMET	26.49912818	1.380249726
BRA_PI_Oeiras.818690_INMET	26.79274562	1.666261201
BRA_PI_Parnaiba.817520_INMET	27.34955664	1.076530013
BRA_PI_Paulistana.828820_INMET	27.5901627	1.407472521
BRA_PI_Picos.827910_INMET	28.22573242	1.646632439
BRA_PI_Piripiri.817940_INMET	27.90965678	0.860027472
BRA_PI_Sao.Joao.do.Piaui.818480_INMET	28.24562721	1.144474375

BRA_PI_Sao.Pedro.do.Piaui.818280_INMET	26.24051019	1.084346404
BRA_PI_Sao.Raimundo.Nonato.819890_INMET	27.0458871	1.061191494
BRA_PI_Teresina.818320_INMET	27.68650748	0.979280984
BRA_PI_Urucui.819050_INMET	25.80740818	1.340344309
BRA_PI_Valenca.do.Piaui.818700_INMET	27.07710448	1.451129286
BRA_PR_Castro.869210_INMET	16.89397599	2.61766029
BRA_PR_Cidade.Gaucha.868980_INMET	22.82977233	2.791842945
BRA_PR_Clevelandia.869390_INMET	17.58247	2.78424825
BRA_PR_Curitiba.838420_INMET	17.39727945	2.247120643
BRA_PR_Diamante.do.Norte.868610_INMET	22.80462816	2.639409774
BRA_PR_Dois.Vizinhos.869270_INMET	20.0276982	3.080743654
BRA_PR_Foz.do.Iguacu-Cataratas.Intl.AP.869250_INMET	21.56694188	4.356835261
BRA_PR_General.Carneiro.869410_INMET	15.86095515	3.056446944
BRA_PR_Ibaiti.869020_INMET	20.22863381	2.312893322
BRA_PR_Icaraima.868970_INMET	22.44312796	2.738346137
BRA_PR_Ilha.do.Mel.869350_INMET	21.0803948	2.568176194
BRA_PR_Inacio.Martins.869290_INMET	16.3156809	2.672972717
BRA_PR_Ivai.869300_INMET	19.1063427	3.08187384
BRA_PR_Joaquim.Tavora.869030_INMET	20.5251035	2.690436227
BRA_PR_Londrina.837680_INMET	21.84908509	2.6816385
BRA_PR_Marechal.Candido.Rondon.869160_INMET	21.40980359	3.008714086
BRA_PR_Maringa.868990_INMET	23.04541248	2.394671707
BRA_PR_Nova.Fatima.869010_INMET	20.79175286	1.927888107
BRA_PR_Nova.Tebas.869190_INMET	20.21106623	2.497162572
BRA_PR_Paranapoema.868620_INMET	22.8810842	2.671240684
BRA_PR_Planalto.869260_INMET	21.33988893	3.479603928
BRA_RJ_Arraial.do.Cabo.868920_INMET	22.94219123	1.31370846
BRA_RJ_Cambuci.868540_INMET	23.31410656	2.283402593
BRA_RJ_Campos.dos.Goytacazes.868550_INMET	24.0316209	2.20228769
BRA_RJ_Duque.de.Caxias-Xerem.868770_INMET	23.06087382	2.190698738
BRA_RJ_Farol.de.Sao.Tome.868900_INMET	23.8032568	1.730942571
BRA_RJ_Macae.868910_INMET	23.58723798	2.114708343
BRA_RJ_Niteroi.868810_INMET	21.96033666	1.827150853
BRA_RJ_Paraty.869130_INMET	22.03288354	2.186500332
BRA_RJ_Petropolis-Pico.do.Couto.868760_INMET	13.61347668	1.711906111
BRA_RJ_Resende.837380_INMET	20.71339246	2.057449794

BRA_RJ_Rio.de.Janeiro-Vila.Militar.868790_INMET	23.16633065	2.222456339
BRA_RJ_Teresopolis.868880_INMET	17.85431196	2.073659594
BRA_RJ_Valenca.868750_INMET	22.16706051	2.514714645
BRA_RN_Apodi.818350_INMET	26.76046816	1.21201165
BRA_RN_Caico.818750_INMET	27.78271015	1.615058227
BRA_RN_Macau.818360_INMET	26.90564047	0.554347289
BRA_RN_Mossoro.818340_INMET	26.72082429	1.109739523
BRA_RN_Natal.818390_INMET	26.90564047	0.554347289
BRA_RN_Touros.818380_INMET	26.85039787	1.057015573
BRA_RO_Ariquemes.819700_INMET	25.62758361	0.713551763
BRA_RO_Cacoal.866220_INMET	26.14048283	0.942138233
BRA_RO_Porto.Velho.819320_INMET	25.92840907	0.672713253
BRA_RO_Vilhena.866420_INMET	24.20993245	0.733212676
BRA_RR_Boa.Vista.816150_INMET	27.01856834	0.970666226
BRA_RS_Alegrete.869750_INMET	19.17727492	4.686581576
BRA_RS_Bage_fixed.839800_INMET	17.65109628	4.173672781
BRA_RS_Bento.Goncalves.869790_INMET	17.18990546	3.310554438
BRA_RS_Cacapava.do.Sul.869860_INMET	17.60460347	3.756298713
BRA_RS_Camaqua.869890_INMET	17.93891222	3.90482847
BRA_RS_Canela.AP.869800_INMET	15.66568559	3.56994845
BRA_RS_Cangucu.869930_INMET	16.27179147	3.854322918
BRA_RS_Chui.869980_INMET	17.27632091	4.468953866
BRA_RS_Cruz.Alta.839120_INMET	18.58727839	3.851815028
BRA_RS_Erechim.869540_INMET	17.46264668	3.136556317
BRA_RS_Frederico.Westphalen.869510_INMET	19.26707843	3.60822491
BRA_RS_Jaguarao.869960_INMET	17.88751528	4.270455951
BRA_RS_Lagoa.Vermelha.869650_INMET	16.629779	3.104256967
BRA_RS_Mostardas.839700_INMET	18.89834984	3.838501368
BRA_RS_Palmeira.das.Missoes.869530_INMET	18.38842758	3.29934189
BRA_RS_Passo.Fundo.869630_INMET	17.51132275	3.309511103
BRA_RS_Porto.Alegre.869880_INMET	20.04291949	3.810534632
BRA_RS_Quarai.869820_INMET	19.07792955	5.044183925
BRA_RS_Rio.Pardo.869780_INMET	18.81128667	3.963648302
BRA_RS_Santa.Maria.839360_INMET	19.07624048	3.911663127
BRA_RS_Santana.do.Livramento.839530_INMET	17.60955648	4.521338656
BRA_RS_Santiago.869760_INMET	18.65081869	4.48923058

BRA_RS_Santo.Augusto.869520_INMET	19.08010371	3.539954441
BRA_RS_Sao.Borja.869600_INMET	20.68275695	4.603242105
BRA_RS_Sao.Gabriel.869840_INMET	19.0956402	4.628063487
BRA_RS_Sao.Jose.dos.Ausentes.869670_INMET	13.8211545	1.83165707
BRA_RS_Sao.Luiz.Gonzaga.869610_INMET	20.28058553	4.126780262
BRA_RS_Soledade.869640_INMET	17.42431009	3.919573896
BRA_RS_Torres.869810_INMET	18.86882966	3.089099271
BRA_RS_Tramandai.869900_INMET	18.98846121	3.205374307
BRA_RS_Uruguaiana.869730_INMET	19.58355625	4.740622957
BRA_RS_Vacaria.869660_INMET	15.66718969	3.672344776
BRA_SC_Ararangua.869710_INMET	19.39507515	3.81090898
BRA_SC_Cacador.869430_INMET	16.48003251	2.764024361
BRA_SC_Curitibaños.AP.869560_INMET	16.10632688	2.684626163
BRA_SC_Dionisio.Cerqueira.869360_INMET	19.22092675	3.560021786
BRA_SC_Florianopolis.838970_INMET	20.93189169	2.819079243
BRA_SC_Indaial.838720_INMET	19.9738747	3.119003027
BRA_SC_Itapoa.869470_INMET	19.96212518	2.942430927
BRA_SC_Ituporanga.869570_INMET	18.17916325	4.009940363
BRA_SC_Joacaba-Santa.Terezinha.Muni.AP.869550_INMET	17.6300407	2.781180884
BRA_SC_Major.Vieira.869420_INMET	16.99042053	3.660326384
BRA_SC_Novo.Horizonte.869380_INMET	18.12361754	3.43255242
BRA_SC_Rio.Negrinho.869450_INMET	16.41788304	2.835062963
BRA_SC_Santa.Marta.839250_INMET	19.4413063	2.684812985
BRA_SC_Sao.Joaquim.869690_INMET	13.65118741	2.367597846
BRA_SC_Sao.Miguel.do.Oeste.AP.869370_INMET	19.66874149	3.569112828
BRA_SC_Urubici.869680_INMET	10.84095913	1.752170255
BRA_SC_Urussanga.869700_INMET	19.67112306	3.855000986
BRA_SC_Xanxere.869400_INMET	18.05274426	2.784003687
BRA_SE_Aracaju.866160_INMET	26.53916515	1.132403943
BRA_SE_Brejo.Grande.866180_INMET	26.31556444	1.070394228
BRA_SE_Carira.866150_INMET	24.18963656	2.150606864
BRA_SE_Itabaianinha.866380_INMET	24.60835496	1.793854912
BRA_SE_Poco.Verde.866140_INMET	24.02143247	2.2045585
BRA_SP_Ariranha.868410_INMET	22.19822042	2.248405744
BRA_SP_Avare.869040_INMET	19.60557446	2.16534931
BRA_SP_Bauru.837220_INMET	20.85211003	2.016860295

BRA_SP_Campinas.837210_INMET	22.0336186	2.067387416
BRA_SP_Campos.do.Jordao.868720_INMET	15.32566151	2.196871141
BRA_SP_Casa.Branca.868440_INMET	21.39232687	1.750952651
BRA_SP_Franca.868180_INMET	20.9682965	1.356928453
BRA_SP_Ibitinga.868430_INMET	21.92536728	2.137623439
BRA_SP_Iguape.869230_INMET	21.1214182	2.884816678
BRA_SP_Itapeva.837740_INMET	19.12223524	2.137393029
BRA_SP_Itapira.868690_INMET	21.23719086	2.237143549
BRA_SP_Ituverava.868170_INMET	22.28719289	2.284298322
BRA_SP_Jales.868140_INMET	23.9117611	1.714282618
BRA_SP_Jose.Bonifacio.868390_INMET	22.92934785	2.28991465
BRA_SP_Lins.868400_INMET	22.83910109	2.119751177
BRA_SP_Ourinhos.868660_INMET	20.933569	2.710915533
BRA_SP_Piracicaba.868680_INMET	20.9344358	2.254036734
BRA_SP_Presidente.Prudente.837160_INMET	23.78326621	2.171965524
BRA_SP_Rancharia.868640_INMET	21.39354511	2.981142831
BRA_SP_Sao.Carlos.868450_INMET	20.29945367	1.973525167
BRA_SP_Sao.Luis.do.Paraitinga.869120_INMET	19.72950367	3.093354293
BRA_SP_Sao.Paulo.837810_INMET	19.58220953	1.943040743
BRA_SP_Sorocaba.838510_INMET	20.45430742	2.706458605
BRA_SP_Taubate.869110_INMET	20.72776282	2.895183805
BRA_SP_Valparaiso.868380_INMET	23.3156653	2.175566865
BRA_SP_Votuporanga.868150_INMET	23.73697285	1.751402792
BRA_TO_Araguaina.819000_INMET	25.128328	0.498601408
BRA_TO_Araguatins.818210_INMET	27.08307169	1.026639311
BRA_TO_Dianopolis.866320_INMET	24.20881259	1.19610511
BRA_TO_Formoso.do.Araguaia.866290_INMET	26.62110036	1.452324585
BRA_TO_Gurupi.866300_INMET	25.85143556	1.096748574
BRA_TO_Palmas.866070_INMET	26.81260937	1.128799564
BRA_TO_Parana.866500_INMET	25.92952048	1.031468477
BRA_TO_Pedro.Afonso.819410_INMET	26.38471502	0.876434582
BRA_TO_Peixe.866490_INMET	25.91926075	0.957435868

Fonte: elaborada pelo autor (2020).