

BTec 012/2024

**BOLETIM
TÉCNICO**
EDIÇÃO ESPECIAL

TRABALHOS PREMIADOS
22º COBREAP SÃO PAULO SP

Autores

Antoniél Campos

Árlon Facynek de Oliveira Carvalho

José Nilo Alves de Sousa Neto

Juliana Miura

Marcelo Ladeira

www.ibape-nacional.com.br

SUMÁRIO

INTRODUÇÃO.....	03
I. ABORDAGEM TRIDIMENSIONAL DO PLEITO EM PERÍCIAS ENVOLVENDO ANÁLISE DE PERDA DE PRODUTIVIDADE.....	04
II. ESTIMATIVA DE VALORES DE IMÓVEIS DA UNIÃO COM TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E COMPONENTES ESPACIAIS.....	40
III. PREDIÇÃO DE OCORRÊNCIAS ARQUEOLÓGICAS EM OBRAS DE INFRAESTRUTURA.....	60

INTRODUÇÃO

O IBAPE – Instituto Brasileiro de Avaliações e Perícias de Engenharia, Entidade Federativa Nacional, instituição sem fins lucrativos, congrega entidades atuantes nas áreas de Engenharia de Avaliações e Perícias de Engenharia em diversas Unidades da Federação. Dentre seus objetivos destacam-se ações visando o aprimoramento, divulgação e transmissão do conhecimento técnico.

É filiado às mais importantes entidades internacionais dedicadas ao segmento de avaliações: UPAV¹ – União Pan-americana de Associações de Avaliação e o IVSC Conselho². Internacional de Normas de Avaliação, organismos voltados para a difusão do conhecimento técnico e normalização nos âmbitos continental e global.

Os BTec - Boletins Técnicos têm por finalidade apresentar temas de alta relevância para as Avaliações e Perícias de Engenharia, sempre elaborados por autores que são referências nos assuntos.

Os boletins representam, portanto, a visão dos autores sobre o assunto, não se constituindo como um Estudo ou uma norma do IBAPE.

Nesta edição especial, o BTec 012/2024 aborda os trabalhos premiados no XXII COBREAP – Congresso Brasileiro de Engenharia de Avaliações e Perícias, correalizado com o IBAPE/SP.

¹ Unión Panamericana de Asociaciones de Valuación.

² International Valuation Standards Council.

I. ABORDAGEM TRIDIMENSIONAL DO PLEITO EM PERÍCIAS ENVOLVENDO ANÁLISE DE PERDA DE PRODUTIVIDADE

Antoniél Campos

Dentre as perícias relacionadas a pleitos de reequilíbrio econômico-financeiro de contratos, as que envolvem a análise de perda de produtividade demandam do perito o maior uso das ferramentas disponibilizadas pela Engenharia de Custos, o que requer o domínio desses conhecimentos por parte do expert.

Em perícias do gênero, é comum o perito ser instado a esclarecer não apenas as causas da perda de produtividade, mas também estimar o valor do dano financeiro dela decorrente e indicar responsabilidades. É, assim, uma perícia que requer uma abordagem tridimensional do pleito, calcada nos elementos constituintes da responsabilidade objetiva: conduta, dano e nexos causal.

A conduta das partes redonda nas causas desencadeadoras da perda de produtividade, sendo necessário, portanto, proceder ao levantamento exaustivo dessas causas, tanto quanto permitir o suporte registral disponibilizado ao especialista.

O dano financeiro decorrente da perda de produtividade é uma das parcelas do desequilíbrio econômico-financeiro do contrato, pois, havendo além da perda de produtividade a extensão do prazo, outras parcelas devem ser incluídas no reequilíbrio. Embora valores possam ser inicialmente apresentados em pleitos pela parte requerente, todas as parcelas envolvidas serão objeto de quantificação pela perícia. Para o escopo deste artigo, o objeto de estudo é o dano financeiro decorrente da perda de produtividade.

Por fim, o nexos causal, enquanto elemento que liga a conduta da parte ao dano verificado, é de obrigatória comprovação, sem a qual não há que se falar em responsabilizações.

O presente artigo abordará, pormenorizadamente, a forma de obtenção desses três elementos, de forma objetiva, apresentando novas ferramentas para suas apurações.

Serão mostradas ferramentas e técnicas de descobrimento de causa-raiz, de modo a deixar claro a ação ou inação da parte, permitindo a verificação do nexo causal entre a conduta/causa e o dano/desequilíbrio. Estabelecidas as causas, será utilizada inovadora metodologia de quantificação de sua importância, sem a subjetividade dominante nas metodologias atuais.

No tocante à apuração do dano, serão abordadas as duas metodologias mais importantes e usualmente indicadas: a *Measured Mile* (Produtividade Natural) e a Análise de Valor Agregado (AVA). Na possibilidade de uso da *Measured Mile*, será apresentada uma forma inovadora de obtenção da produtividade natural, com amparo em testes estatísticos, afastando a subjetividade na seleção de períodos impactados e não impactados. Uma vez estabelecida a produtividade natural, ou mesmo a utilização da produtividade baseline, será demonstrada mais uma inovação na quantificação do tempo ocioso por período (semanal, quinzenal, mensal) e, conseqüentemente, na extensão do dano em todo o contrato, correspondente ao impacto total da ociosidade da mão de obra e equipamentos.

Na outra ponta, será visto a forma de contemplar as parcelas do custo que sofrem a ação da perda de produtividade, notadamente a mão de obra e os equipamentos, permitindo quantificar o custo unitário do dano financeiro. Esses dois parâmetros — extensão do dano e o seu custo unitário —, permitem quantificar o dano financeiro total, que é o valor da perda de produtividade que, por sua vez, integra o desequilíbrio econômico-financeiro reclamado.

Dessa forma, as questões que se esperam ver respondidas no presente artigo são as seguintes: i.1) como identificar a(s) causa(s)-raiz deflagradora(s) da perda de produtividade? i.2) como quantificar a importância dessa(s) causa(s)? ii) como quantificar o dano financeiro decorrente da perda de produtividade? e iii) como verificar a existência de nexo de causalidade entre causa(s) e dano?

O objetivo do artigo é, portanto, responder a esses questionamentos, que são usuais em perícias de valoração de desequilíbrio econômico-financeiro de contratos, apresentando ao perito uma metodologia de cálculo do valor da perda de produtividade, a partir de parâmetros isentos de subjetividade, alicerçada na conduta da parte, no dano financeiro verificado e no nexos causal interligando esses dois elementos.

1. REFERENCIAL TEÓRICO

1.1 Definições

Conforme consta nas DIRETRIZES PARA ANÁLISE DE PRODUTIVIDADE, publicação do IBAPE Nacional, produtividade pode ser entendida, dentre várias definições, como sendo “a razão entre a quantidade de produtos ou serviços produzidos e a quantidade de recursos utilizados” (2021, p. 11). E arremata:

“produtividade refere-se a uma razão, um quociente entre duas grandezas: o produto (saídas) e os recursos (entradas).”

No mesmo sentido, ressalta Ricardo Salomão (2022, p. 212) que “Produtividade é a quantidade produzida por unidade de tempo, que também pode ser definida como a relação entre os resultados obtidos e o esforço demandado”.

Para a Prática Recomendada nº 25R-03, como estimar perda de produtividade de mão de obra em pleitos de construção, da AACE® International (2004, p. 2), a produtividade “refere-se às quantidades produzidas por hora de esforço do trabalhador”, podendo ser definida por qualquer uma das equações:

- $\text{Produtividade} = \text{Saída} \div \text{Entrada}$
- $\text{Produtividade} = \text{Unidades} \div \text{horas de trabalho}$
- $\text{Produtividade} = (\text{Produção total}) \div (\text{total de horas de trabalho})$

Definida a produtividade, a perda de produtividade pode ser compreendida como “o aumento do custo de desempenho causado por uma mudança nos recursos

previstos ou planejados, condições ou método de trabalho da Contratada.” (AACE, 2004, p. 11).

Sob a ótica das obrigações, a perda de produtividade para Eduardo Grebler:

consiste na diferença entre a produção contratada e a produção alcançada, que se reflete no cumprimento, ou não, dos prazos estabelecidos no cronograma do contrato e dos custos acordados entre as partes. (...) a improdutividade em regra se reflete no descumprimento dos prazos e custos contratados, por fato de responsabilidade de uma ou de ambas as partes. [GREBLER, 2014, p. 26;31]

O autor reforça o papel do cronograma e do acompanhamento de custos como elementos de monitoramento da perda de produtividade:

Na quase totalidade dos contratos, a perspectiva jurídica da improdutividade é analisada sobre o prisma dos prazos e custos. (...). Caracterizados os efeitos da improdutividade pelos mecanismos de aferição do andamento do empreendimento (cronograma e acompanhamento de custos), segue-se a necessidade de se determinar suas causas e seus efeitos. [GREBLER, 2014, p. 26;28]

Por sua vez, Simon Braithwaite (2017, sl. 30), destaca que o estabelecimento do direito a um ressarcimento por perda de produtividade requer três provas principais:

- Responsabilidade - direito contratual de recuperar devido a um problema
- Causação - relação de causa e efeito entre a ação / inação por uma parte e a resultante lesão ou impacto em outra
- Danos - a magnitude do impacto para a pessoa lesada

1.2 Do equilíbrio econômico-financeiro do contrato

Consoante a precisa lição de Ricardo Salomão:

Quando as partes celebram um contrato, elas estão também desenhando a equação contratual. Naquele momento, com as variáveis básicas devidamente pactuadas, entende-se que o contrato está em equilíbrio e esta noção é muito importante para a investigação pericial posterior. Em outras palavras, somente se pode falar em análise

de um desequilíbrio econômico-financeiro se as condições de equilíbrio forem conhecidas. [SALOMÃO, 2022, p. 207]

Conhecidas, então, as condições pactuadas, o autor define o que vem a ser o desequilíbrio, nos seguintes termos: “pode-se então definir um desequilíbrio econômico-financeiro como uma alteração, sem contrapartidas compensatórias, em uma ou mais das variáveis que constituem a equação contratual.” (Salomão, 2022, p. 208)

No mesmo sentido e detalhando o que seria a equação contratual, Maçahico Tisaka leciona:

Quando uma empresa vence uma licitação de uma determinada obra e assina o contrato, com as condições constantes no edital, as especificações técnicas, os quantitativos e custos unitários constantes na planilha de orçamento, as Leis sociais e encargos complementares utilizados, a composição do BDI, preço global e o cronograma físico-financeiro definido pelo prazo estipulado na proposta, significa que a manutenção de todos esses elementos representa a equação econômico-financeiro inicial do contrato, os quais se pressupõe equilibrados para todos os efeitos.

A alteração de qualquer uma dessas condições provoca o seu desequilíbrio, havendo a necessidade de buscar o seu reequilíbrio, para que qualquer uma das partes não saia prejudicada. [TISAKA, 2011, p. 89]

1.3 Da importância do cronograma físico-financeiro

Registra a Prática Recomendada nº 45R-08: Métodos de Proteção contra Pleitos Relacionados a Cronograma, da AACE® International, que:

O cronograma mais importante para a fase de execução do projeto é o cronograma de linha de base. Contudo, o segundo cronograma mais importante é o cronograma como executado (“as-built”) quando estiver relacionado à resolução bem sucedida de pleitos por atraso. (...) Quando o cronograma é desenvolvido adequadamente, além de ser rigorosamente mantido e apoiado pela documentação do projeto, ele é um elemento vital para a resolução de pleitos por atraso. [AACE, 2009, p. 11]

Roberto Sales Cardoso explica o porquê da relevância do cronograma “as built”, nos seguintes termos:

O cronograma possibilita o controle do desempenho da contratada por meio da comparação entre os valores das medições mensais propostas e as realizadas.

Percentuais mensais realizados com valores inferiores aos previstos caracterizam atraso, podendo ensejar a aplicação de multa por parte da fiscalização. [CARDOSO, 2016, p. 214]

Maçahico Tisaka adverte que o cronograma “precisa ser constantemente atualizado e a sua descontinuidade pode prejudicar os argumentos para fundamentar os pedidos” (Tisaka, 2011, p. 175).

1.4 Das metodologias utilizadas para estimar a perda de produtividade

Em relação às metodologias usualmente aceitas para estimação da perda de produtividade, a Prática Recomendada nº 25R-03 – Como estimar perda de produtividade de mão de obra em pleitos de construção, cita um conjunto delas, sendo objeto deste artigo as duas tidas como preferenciais: o estudo *Measured Mile* e a Análise de Valor Agregado (AVA).

1.4.1 O cálculo Measured Mile

A primeira delas — o cálculo *Measured Mile* —, é assim definida na já clássica definição dada por Schwartzkopf:

O método mais amplamente aceito para o cálculo de perda de produtividade da mão de obra é conhecido em toda a indústria como o cálculo “*Measured Mile*” (produtividade natural). Esse tipo de cálculo compara atividades idênticas em partes do projeto que sofreram e que não sofreram impactos para averiguar a perda de produtividade resultante do impacto de um conjunto conhecido de eventos. O cálculo *Measured Mile* é preferido porque considera apenas o efeito real do impacto alegado, eliminando disputas a respeito da validade de estimativas de custos ou de fatores que possam ter afetado a produtividade sem culpa da Contratante. [SCHWARTZKOPF, 1995, Apud AACE, 2004, p. 16]

Ante a dificuldade em se comparar atividades idênticas em períodos impactados e não impactados, a justiça norte-americana decidiu que é permitida a comparação da produtividade observada em atividades idênticas ou similares, ocorridas:

- a) em períodos não impactados versus períodos impactados; e
- b) em períodos menos impactados versus períodos impactados.

A forma de calcular essa produtividade natural é explicada no protocolo citado da seguinte forma:

Unidades físicas de trabalho concluído divididas por horas despendidas para concluir tais itens de trabalho determinam a produtividade durante o período menos afetado ou não afetado. Então, um cálculo semelhante é executado para o período do impacto. A perda de produtividade pode então ser calculada pela subtração do índice de produtividade unitária durante o período impactado, da produtividade unitária durante o período não afetado. [AACE, 2004, p. 17]

Conforme lembra Simon Braithwaite (2017, sl. 58), caso não haja período *Measured Mile* definido, a baseline pode ser adotada.

É de se ressaltar que o cálculo *Measured Mile* se reporta a identificar a diferença entre a produtividade própria da contratada, que é a produtividade independente de impacto, e aquela obtida em situação de impacto, sendo relevante alertar que, no caso de contratos, o equilíbrio se reporta à aferição com a produtividade pactuada, ou seja, a produtividade baseline, constante na proposta.

A produtividade natural pode ser inferior, igual ou superior a essa produtividade pactuada (baseline), sendo então conveniente calcular um fator que compare essas duas produtividades, a fim de se estimar, antes mesmo de analisar os impactos, qual seria o desempenho da contratada caso não houvesse impactos.

É o que diz Ricardo Salomão, ao definir os requisitos mínimos que uma metodologia de estimação de perda de produtividade deve ter: “Para quantificar o rompimento com rigor e precisão, uma metodologia precisa ser capaz de (1) explicar as causas de variações na produtividade e retrabalho e (2) avaliar o que teria acontecido em condições normais.” (2022, p. 215)

Para se estimar a relação entre a produtividade pactuada (baseline) e a produtividade natural da contratada, pode-se lançar mão do cálculo do Fator de

produtividade, assim definido pela Prática Recomendada nº 25R-03: Como estimar perda de produtividade de mão de obra em pleitos de construção (AACE, 2004, p. 3):

$$\text{Fator de Produtividade} = \frac{\text{Produtividade Real}}{\text{Baseline ou Produtividade planejada}}$$

Compreenda-se a “Produtividade Real” como sendo a produtividade natural, calculada pela *Measured Mile*.

Caso esse fator seja maior que 1, significa que, *ceteris paribus*, a contratada concluiria o projeto antes do prazo pactuado. Para um fator inferior a 1, o projeto seria concluído com atraso e, em caso de igualdade, $F_p = 1$, o projeto seria concluído no prazo pactuado.

1.4.2 Análise de Valor Agregado

A outra metodologia abordada pela Prática Recomendada nº 25R-03, e objeto deste artigo, é a Análise de Valor Agregado (AVA), cuja forma de mensuração da perda de produtividade é assim descrita:

A estimativa da Contratada ou, em alternativa, o valor em dólares dos pedidos de pagamento, valores contratuais ou preços unitários podem ser usados para determinar as horas de mão de obra, quando elas foram despendidas e, possivelmente, em quais atividades. Unidades físicas de trabalho concluídas multiplicadas pelos índices de orçamento unitário podem ser usadas para determinar horas agregadas. As horas agregadas são então comparadas às horas reais gastas para o período do impacto e a diferença entre as duas pode ser usada para calcular a perda de produtividade vivenciada. [AACE, 2004, p. 17]

Ricardo Salomão estabelece o limite máximo para a perda de produtividade, ao tempo que já adianta sobre quais parcelas é calculado esse impacto:

A quantificação da improdutividade/ociosidade requer conhecimentos específicos sobre esta área da Engenharia de Custos. A primeira coisa que deve ser calculada é o limite máximo desse impacto, que é a diferença entre os histogramas reais e

previstos, de mão de obra e de equipamentos. Nenhum cálculo de improdutividade/ociosidade poderá resultar em maior valor do que o preço destas diferenças. [SALOMÃO, 2022, p. 223]

1.5 Das parcelas do custo que sofrem a ação da ociosidade e do atraso

Para Aquiles da Costa (2016, p. 98), na hipótese de sobrevirem fatos extraordinários, impactantes, “deverá ser assegurada a necessária recomposição de todo e qualquer custo adicional”, inclusive “aqueles relativos à subutilização do pessoal e dos equipamentos mobilizados no período em que perduraram estes fatores”.

É no mesmo sentido o entendimento de Ricardo Salomão:

Especial destaque deve ser dado à identificação de ocorrências de improdutividade e ociosidade de recursos materiais e humanos (equipamentos e mão de obra), visto que esses eventos representam desembolsos realizados, por uma ou por ambas as partes, sem a contrapartida do avanço físico esperado, o que afeta diretamente a equação contratual. [SALOMÃO, 2022, p. 206]

2. METODOLOGIA

Será apresentada a metodologia proposta para a abordagem tridimensional de um pleito envolvendo perda de produtividade, contemplando: i) como quantificar física e financeiramente o dano; ii) como quantificar a importância das causas desencadeadoras do dano; e iii) como estabelecer o nexo de causalidade, como liame entre a conduta da parte que redundou nas causas e o dano financeiro decorrente.

2.1 O cálculo *Measured Mile*

Configurado o desempenho impactado de um projeto, há a necessidade de se comparar tal performance com uma outra que tenha sido pactuada entre as partes, integrante da equação contratual. Esse desempenho é representado pela produtividade baseline, ou da linha de base.

Ocorre que, uma vez apresentado um pleito de reequilíbrio econômico-financeiro, as partes tendem a discutir, de forma diametralmente opostas, a validade técnica dessa produtividade referencial: do lado da contratada, a afirmação de que não cumpriu com a produtividade pactuada em face dos impactos provocados pela contratante; do lado da contratante, a afirmação de que a prática da contratada destoava da pactuada e que, mesmo na ausência dos impactos alegados, a mesma não estaria cumprindo com o pactuado.

Para dirimir a celeuma, passou-se então a avaliar qual seria a produtividade natural da contratada num cenário em que não houvesse impactos. A partir dessa aferição, estabelecer-se-ia um novo marco temporal técnico — mas não um marco jurídico —, que permitiria avaliar a perda de produtividade advinda exclusivamente dos impactos, sem interferência da performance da contratada¹. A forma de avaliar essa produtividade natural da contratada é conhecida como Cálculo *Measured Mile*.

2.1.1 Identificação de períodos não impactados

Considere-se um projeto com custo de \$ 100, prazo de execução previsto de 8 meses, mas executado em 16, conforme representado no cronograma a seguir.

Tabela 1 – Cronograma físico-financeiro (exemplo)

Desemp.nat	0,0	1,5	3,1	4,6	6,2	7,7	9,2	10,8	12,3	13,8	15,4	16,9	18,5	20,0	21,5	23,1	24,6
mês	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VP	0	10	25	40	60	80	90	95	100								
VA	0	5	8	11	15	22	30	35	38	41	52	58	60	75	88	95	100
n1	0	0	0	1	1	1	2	2	2	3	3	3	4	4	5	7	7
n2	0	1	1	2	2	2	3	3	3	4	4	4	5	5	6	8	8
VPn1	0	0	0	10	10	10	25	25	25	40	40	40	60	60	80	95	95
VPn2	0	10	10	25	25	25	40	40	40	60	60	60	80	80	90	100	100
PA		0,50	0,80	1,07	1,33	1,80	2,33	2,67	2,87	3,05	3,60	3,90	4,00	4,75	5,80	7,00	8,00
Pplan		10,00	10,00	11,25	15,00	15,00	15,00	15,00	15,00	16,36	20,00	20,00	20,00	20,00	12,38	5,83	5,00
Pexec		5,00	3,00	3,00	4,00	7,00	8,00	5,00	3,00	3,00	11,00	6,00	2,00	15,00	13,00	7,00	5,00
Fp = Pexec/Pplan		0,50	0,30	0,27	0,27	0,47	0,53	0,33	0,20	0,18	0,55	0,30	0,10	0,75	1,05	1,20	1,00
Atraso		0,50	1,20	1,93	2,67	3,20	3,67	4,33	5,13	5,95	6,40	7,10	8,00	8,25	8,20	8,00	8,00
Impacto		0,25	0,35	0,37	0,37	0,27	0,23	0,33	0,40	0,41	0,23	0,35	0,45	0,13	-0,03	-0,10	0,00
																	4,13

Fonte: Elaborado pelo autor

¹ Do ponto de vista obrigacional, o possível atraso decorrente exclusivamente de uma performance da contratada abaixo da pactuada, deve ser abordada entre as partes de forma apartada. (Nota do autor)

Onde:

VP e VA: valores planejado e agregado (medido), respectivamente.

n1 e n2: nº dos meses que delimita o intervalo em VP que contém o VA.

VPn1 e VPn2: Valores planejados em n1 e n2, respectivamente

PA: prazo agregado, dado pela expressão:

$$PA = n1 + \frac{(VA - VPn1)}{(VPn2 - VPn1)}$$

Pplan: Produtividade planejada em “n” = $(VAn - VAn-1)/(PAn - PAn-1)$

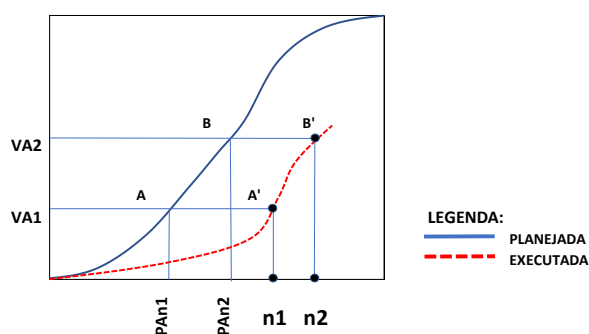
Pexec: Produtividade executada em “n” = $(VAn - VAn-1)/(n - n-1)$

Atraso: indica quanto o projeto está atrasado no mês considerado =
mês – PA

Fp: Fator de produtividade = Pexec / Pplan

Para compreensão da metodologia proposta, as curvas de desempenho previsto e realizado estão representadas na figura a seguir.

Figura 1 – Desempenho previsto x executado



Fonte: Elaborado pelo autor

O modo usual de calcular a *Measured Mile* consiste em identificar atividades iguais ou semelhantes e comparar a sua produtividade em período impactado e não impactado (ou pouco impactado). A fim de afastar a subjetividade desse processo, será admitida a premissa de que, conforme a Figura 1, e desde que não tendo havido alteração no sequenciamento, as atividades planejadas no segmento AB são as mesmas executadas em A'B', apenas variando o prazo da execução. Essa premissa

é verossímil, pois sendo a curva planejada a totalização dos VPs de todas as atividades do respectivo mês, esse conjunto de atividades será reproduzido na curva executada, apenas num prazo maior (ou mesmo menor) que o planejado. No caso de a sequência das atividades ter sido alterada, será necessário refazer a curva planejada, de modo a retratar a sequência real das atividades, a fim de que a premissa continue válida.

Dessa forma, o que se pretende doravante é encontrar a maior amostra possível de produtividades planejadas e executadas que possam ser consideradas como participantes de um mesmo universo, no caso, do universo das atividades não impactadas, pois a produtividade planejada provém desse cenário não impactado.

A estatística condiciona a comparabilidade entre grupos ao teste de igualdade das médias, no caso de amostra com distribuição normal, ou comparando os postos de posição, no caso de dados sem distribuição normal, ou não-paramétricos, e desde que, em ambas as situações, atenda-se ao nível de significância estipulado para o teste de igualdade.

Caso os dois grupos de produtividades, Pplan e Pexec, possam ser considerados partícipes de um mesmo universo, será realizado o cálculo do Fator de produtividade ($F_p = P_{exec}/P_{plan}$), considerando para cada produtividade o valor da tendência central adequada ao caso (a média aritmética, no caso de dados paramétricos, ou a mediana, para dados não-paramétricos). Para valores de F_p igual a 1, significa que $P_{plan} = P_{exec}$, e que a produtividade natural é, portanto, igual à planejada. Caso F_p seja maior que 1, significa que a produtividade natural é maior que a produtividade baseline e, por fim, caso F_p resulte menor que a unidade, significará que a produtividade natural é inferior à planejada (baseline) e que, caso o projeto não fosse impactado, o mesmo seria executado com atraso, tendo em vista a produtividade natural ser inferior à pactuada.

Para o exemplo em tela, organizou-se os dados, por comodidade, seguindo a ordem decrescente do Fp, conforme tabela a seguir.

Tabela 2 – Teste de igualdade para as produtividades

item	mês	Pplan	Pexec	Fp	Normalidade	Igualdade	
					Shapiro-Wilk	Wilcoxon	t pareado
1	15	5,83	7	1,20			
2	14	12,38	13	1,05	1,0000		0,1894
3	16	5,00	5	1,00	0,1972		0,2194
4	13	20,00	15	0,75	0,3982		0,6108
5	10	20,00	11	0,55	0,1682		0,2853
6	6	15,00	8	0,53	0,2751		0,1322
7	1	10,00	5	0,50	0,4099		0,0641
8	5	15,00	7	0,47	0,2896		0,0266
9	7	15,00	5	0,33	0,1089		0,0113
10	2	10,00	3	0,30	0,2916		0,0043
11	11	20,00	6	0,30	0,1813		0,0025
12	3	11,25	3	0,27	0,1382		0,0009
13	4	15,00	4	0,27	0,0776		0,0003
14	8	15,00	3	0,20	0,0404	0,0029	
15	9	16,36	3	0,18	0,0186	0,0018	
16	12	20,00	2	0,10	0,0255	0,0012	

Fonte: Elaborado pelo autor

Para todos os testes foi estabelecido o nível de significância de 5%. O primeiro teste aplicado é o Shapiro-Wilk, para verificação da normalidade. Caso o conjunto de dados dos dois grupos atenda ao pressuposto da normalidade, serão comparados utilizando-se o teste t pareado; caso contrário², o teste de Wilcoxon. Conforme tabela acima, apenas o subgrupo de 7 dados de cada tipo de produtividade pode ser tido como participantes de um mesmo universo³ ($p\text{-valor} > 0,05$). Estatisticamente, a um nível de significância de 5%, esses dois grupos de 7 dados cada, podem ser tidos como iguais.

A tabela a seguir ilustra a seleção final de produtividades comparáveis.

² Conforme (Souto e Souto, 2020, p. 63), a forma pareada, de “par a par”, é indicada pois se pretende comparar um mesmo grupo de atividades cujo desempenho deu-se em momentos distintos: conforme planejado e conforme executado.

³ A hipótese nula, h_0 , é de que, a um nível de significância de 5% ($\alpha = 0,05$), os grupos são iguais.

Tabela 3 – Teste de igualdade para as produtividades – amostra final

item	mês	Pplan	Pexec	Fp	Normalidade	Igualdade	
					Shapiro-Wilk	Wilcoxon	t pareado
1	15	5,83	7	1,20			
2	14	12,38	13	1,05	1,0000		0,1894
3	16	5,00	5	1,00	0,1972		0,2194
4	13	20,00	15	0,75	0,3982		0,6108
5	10	20,00	11	0,55	0,1682		0,2853
6	6	15,00	8	0,53	0,2751		0,1322
7	1	10,00	5	0,50	0,4099		0,0641
8	5	15,00	7	0,47	0,2896		0,0266
9	7	15,00	5	0,33	0,1089		0,0113
10	2	10,00	3	0,30	0,2916		0,0043
11	11	20,00	6	0,30	0,1813		0,0025
12	3	11,25	3	0,27	0,1382		0,0009
13	4	15,00	4	0,27	0,0776		0,0003
14	8	15,00	3	0,20	0,0404	0,0029	
15	9	16,36	3	0,18	0,0186	0,0018	
16	12	20,00	2	0,10	0,0255	0,0012	

Fonte: Elaborado pelo autor

Como o teste que por último selecionou os dados se trata do teste t pareado, que é um teste paramétrico, o Fator de produtividade será a razão entre a média aritmética de Pexec e a média aritmética de Pplan. Todavia, antes de se proceder a esse cálculo, faz-se necessário um saneamento dessa amostra, a fim de afastar os dados discrepantes. Optou-se pela clusterização por k-means, para 3 clusters, no qual o central será tido como o representativo e os demais, discrepantes. Seguem os cálculos do saneamento da amostra.

Figura 2 – Saneamento da amostra final (exemplo)

mês	Pplan	Pexec	Fp	Global	Item	Cluster
15	5,83333	7	1,2		1	2
14	12,381	13	1,05		2	3
16	5	5	1		3	2
13	20	15	0,75		4	3
10	20	11	0,55		5	3
6	15	8	0,53333		6	1
1	10	5	0,5		7	1
média_1	12,5	6,5	0,52	menor		
média_2	5,41667	6	1,10769	maior		
média_3	17,4603	13	0,74455	meio		

Fonte: Elaborado pelo autor

O cluster 3 é o de valor intermediário, devendo ser o selecionado para representar a amostra saneada. Vê-se que a amostra final selecionada é composta pelos meses 10, 13 e 14. De um total de 16, esses se mostram relativamente distantes

dos segmentos iniciais e finais da curva, o que é um ponto favorável. Em todo caso, estabelece-se como critério de admissão de dados, após o saneamento, aqueles contidos no intervalo de 10% a 90% do prazo total estendido. Caso não haja dados dentro desse intervalo, será considerado prejudicada a aferição da produtividade natural, ocasião em que se sugere a adoção da produtividade da linha de base (baseline), conforme pactuada.

Para o exemplo em tela, o limite seria de 1,6 meses (10% de 16) a 14,4 meses (90% de 16); como os dados finais se referem às medições dos meses 10, 13 e 14, todos os três dados serão considerados como representativos.

2.1.2 Determinação da produtividade natural

A verificação da produtividade natural é realizada a partir da comparação das médias dos dois grupos, para os meses selecionados, conforme cálculos a seguir.

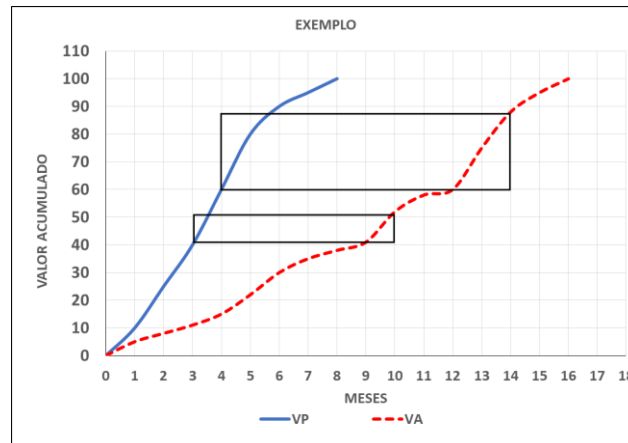
Tabela 4 – Cálculo *Measured Mile* – produtividade natural

item	mês	Pplan	Pexec	Fp	Normalidade	Igualdade	
					Shapiro-Wilk	Wilcoxon	t pareado
1	15	5,83	7	1,20			
2	14	12,38	13	1,05	1,0000		0,1894
3	16	5,00	5	1,00	0,1972		0,2194
4	13	20,00	15	0,75	0,3982		0,6108
5	10	20,00	11	0,55	0,1682		0,2853
6	6	15,00	8	0,53	0,2751		0,1322
7	1	10,00	5	0,50	0,4099		0,0641
8	5	15,00	7	0,47	0,2896		0,0266
9	7	15,00	5	0,33	0,1089		0,0113
10	2	10,00	3	0,30	0,2916		0,0043
11	11	20,00	6	0,30	0,1813		0,0025
12	3	11,25	3	0,27	0,1382		0,0009
13	4	15,00	4	0,27	0,0776		0,0003
14	8	15,00	3	0,20	0,0404	0,0029	
15	9	16,36	3	0,18	0,0186	0,0018	
16	12	20,00	2	0,10	0,0255	0,0012	
Média		17,46	13,00	0,74	=> Produtividade Natural		

Fonte: Elaborado pelo autor

Os períodos selecionados como participantes de um mesmo universo, e que por isso são considerados comparáveis, estão representados na figura a seguir.

Figura 3 – Períodos não impactados comparáveis (exemplo)



Fonte: Elaborado pelo autor

Com efeito, percebe-se, inclusive visualmente, que os segmentos selecionados são comparáveis (inclinações semelhantes) e, portanto, pertencentes à categoria de não impactados, ou pouco impactados, uma vez que a comparação é feita com um padrão de produtividade (a de baseline) que se admite isenta de impactos.

Concluiu-se, então, que a produtividade natural da contratada foi de 74% da produtividade pactuada (de baseline), o que levaria, ceteris paribus, ao atraso do projeto independente da existência de impactos provocadores de perda de produtividade.

Com esse valor de produtividade natural, é possível agora simular o cenário do projeto com três curvas: a planejada (pactuada entre as partes), a executada (impactada) e a do desempenho natural, que seria aquela que teria ocorrido na ausência de impactos. Para o cálculo das novas datas, é necessário multiplicar as atuais por $(1/\text{Prod. Nat.})$, conforme tabela a seguir.

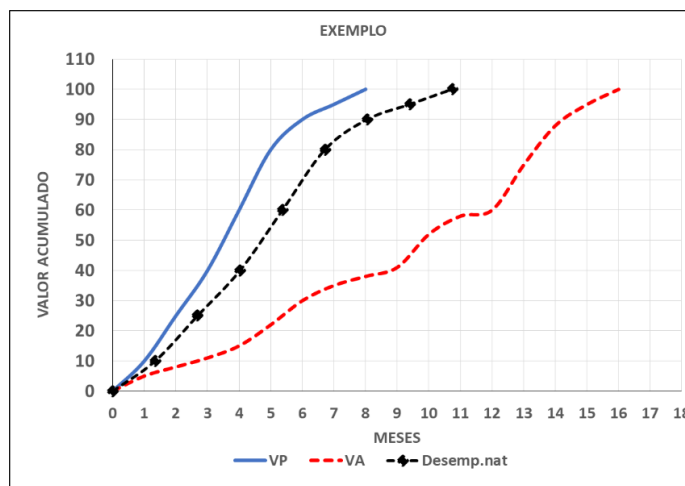
Tabela 5 – Cronograma físico-financeiro – desempenho natural (exemplo)

Desemp. Nat.	0,0	1,3	2,7	4,0	5,4	6,7	8,1	9,4	10,7	12,1	13,4	14,8	16,1	17,5	18,8	20,1	21,5
mês	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VP	0	10	25	40	60	80	90	95	100								
VA	0	5	8	11	15	22	30	35	38	41	52	58	60	75	88	95	100
n1	0	0,0	0,0	1,3	1,3	1,3	2,7	2,7	2,7	4,0	4,0	4,0	5,4	5,4	6,7	9,4	9,4
n2	0	1,3	1,3	2,7	2,7	2,7	4,0	4,0	4,0	5,4	5,4	5,4	6,7	6,7	8,1	10,7	10,7
VPn1	0	0	0	10	10	10	25	25	25	40	40	40	60	60	80	95	95
VPn2	0	10	10	25	25	25	40	40	40	60	60	60	80	80	90	100	100
PA		0,50	0,80	1,41	1,68	2,14	3,02	3,35	3,55	4,08	4,63	4,93	5,37	6,12	7,52	9,40	10,40
Atraso		0,50	1,20	1,59	2,32	2,86	2,98	3,65	4,45	4,92	5,37	6,07	6,63	6,88	6,48	5,60	5,60

Fonte: Elaborado pelo autor

n1 e n2 agora se reportam aos nº dos meses de “Desempenho natural” que delimitam o intervalo em VP que contém o VA. Com isso, o valor de PA se altera, com reflexo direto no cálculo do atraso.

O gráfico contendo os três cenários é o seguinte.

Figura 4 – Desempenho previsto x natural x executado (exemplo)


Fonte: Elaborado pelo autor

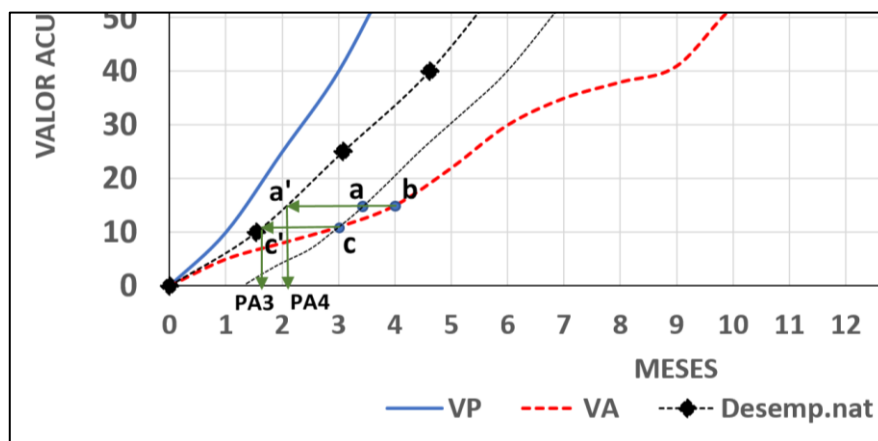
Além da curva representando o desempenho impactado (linha tracejada, em vermelho), vê-se essas duas curvas: a do desempenho previsto (traço cheio, em azul) e a do desempenho natural (traço-ponto, na cor preta), deslocada daquela em (10,7 – 8 =) 2,7 meses, que seria o atraso com que a contratada chegaria ao término do projeto, mesmo sem impactos, apenas por seu próprio desempenho.

Comparando os dois cronogramas (Tabela 1 e Tabela 5), vê-se que o atraso total caiu de 8 meses — quando a referência era a baseline —, para 5,6 meses, quando a referência passou a ser o próprio desempenho natural, baseado na produtividade natural calculada pela *Measured Mile*.

2.2 Extensão do dano

Com esse novo padrão de desempenho, serão calculados os períodos de ociosidade mês a mês, mediante à diferença de tempo entre o período impactado e o período que levaria caso a contratada pudesse ter empreendido o seu desempenho natural. Para tanto, será admitido que a contratada estava com a sua estrutura (mão de obra e equipamentos) disponível no início de cada mês, conforme a programação dos serviços a serem executados. Assim, para a mensuração mensal da perda de produtividade, a curva do desempenho natural será deslocada até atingir o mês em que se quer medir essa perda, conforme explicado na figura a seguir, onde se pretende medir a ociosidade ocorrida no mês 4.

Figura 5 – Detalhamento do cálculo da ociosidade (exemplo)



Fonte: Elaborado pelo autor

O número do mês representa o seu final. O VA do mês 4 é o indicado pela letra **b**, cujo início deu-se em 3, na letra **c**. Dessa forma, a produção do mês 4 é a diferença entre **b** e **c**. Ao deslocar a curva do desempenho natural para o início do período que se quer medir, ou seja, para o ponto **c**, e considerando que o VA em **a** = VA em **b**, vê-se que a mesma produção de **c** até **b** seria obtida de **c** para **a**, só que num período de

tempo menor. Isso significa que, caso não houvesse impacto, o VA do mês 4 seria obtido num tempo inferior a 1 mês, ou seja, nesse período de *a* até *b* a contratada ficou ociosa, com a sua estrutura mobilizada, realizando desembolsos, mas sem produzir. É esse intervalo de tempo, de *a* até *b*, que se quer então calcular. Todavia, esse período é o tempo máximo de dias no mês em que a contratada ficou ociosa. Como no início do mês (ponto *c*) esse período de ociosidade era igual a zero, o tempo médio de ociosidade no mês é a média entre 0 e *ab*.

O PA do ponto *c* (mês 3) pode ser calculado, sendo igual a PA3; da mesma forma, o PA do ponto *b* (mês 4) também pode ser calculado, sendo igual a PA4.

O atraso verificado em *c*, em relação à curva do desempenho natural, é igual a $3 - PA3$, que é igual ao segmento *c'c*; assim como o atraso verificado em *b* é igual a $4 - PA4$, que é igual ao segmento *a'b*. Como *c'c* = *a'a* — pois trata-se do quanto a curva do desempenho natural se deslocou —, a distância buscada, *ab*, é igual à diferença entre os atrasos verificados em *b* e *c*, conforme acima calculados.

Com base em todo o exposto, o impacto médio em cada período de medição pode ser dado pela expressão seguinte:

$$In = \frac{[(n - PAn) - (n-1 - PAn-1)]}{2} \quad (3)$$

Onde:

- In: impacto médio, em meses, relativo à ociosidade no mês “n”
- n: nº do mês considerado
- n-1: nº do mês anterior
- PAn: prazo agregado referente ao mês “n”
- PAn-1: prazo agregado referente ao mês anterior

Acrescentando essa parcela de impacto na tabela do exemplo, vem:

Tabela 6 – Cronograma físico-financeiro – tempo ocioso (exemplo)

Desemp. Nat.	0,0	1,3	2,7	4,0	5,4	6,7	8,1	9,4	10,7	12,1	13,4	14,8	16,1	17,5	18,8	20,1	21,5
mês	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
VP	0	10	25	40	60	80	90	95	100								
VA	0	5	8	11	15	22	30	35	38	41	52	58	60	75	88	95	100
n1	0	0,0	0,0	1,3	1,3	1,3	2,7	2,7	2,7	4,0	4,0	4,0	5,4	5,4	6,7	9,4	9,4
n2	0	1,3	1,3	2,7	2,7	2,7	4,0	4,0	4,0	5,4	5,4	5,4	6,7	6,7	8,1	10,7	10,7
VPn1	0	0	0	10	10	10	25	25	25	40	40	40	60	60	80	95	95
VPn2	0	10	10	25	25	25	40	40	40	60	60	60	80	80	90	100	100
PA		0,50	0,80	1,41	1,68	2,14	3,02	3,35	3,55	4,08	4,63	4,93	5,37	6,12	7,52	9,40	10,40
Atraso		0,50	1,20	1,59	2,32	2,86	2,98	3,65	4,45	4,92	5,37	6,07	6,63	6,88	6,48	5,60	5,60
Impacto		0,25	0,35	0,20	0,37	0,27	0,06	0,33	0,40	0,24	0,23	0,35	0,28	0,13	-0,20	-0,44	0,00
																	3,44

Fonte: Elaborado pelo autor

Os valores positivos, em vermelho, referem-se aos períodos mensais de ociosidade; os valores negativos, em verde, podem estar representando uma melhoria na performance da contratada, ou mesmo uma aceleração, a qual deve ser tratada à parte, em pleito específico.

O somatório dos períodos mensais de ociosidade, Δt , representa a extensão do dano, uma das parcelas necessárias ao cálculo do desequilíbrio decorrente da perda de produtividade. Para o exemplo acima, somando-se apenas os valores positivos, $\Delta t = 3,44$ meses de ociosidade.

2.3 Custo financeiro do dano verificado

2.3.1 Perda de produtividade usando a Análise de Valor Agregado (AVA)

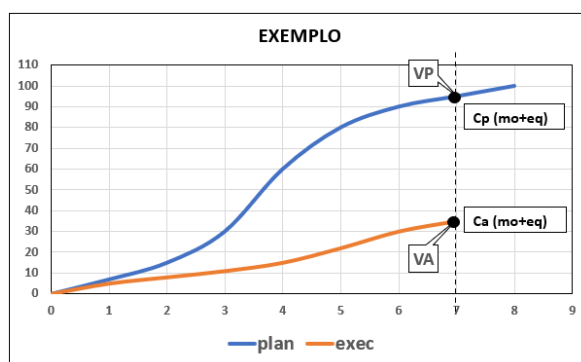
O desenvolvimento das fórmulas que permitem utilizar a AVA na quantificação da perda de produtividade, advém desse excerto da Prática Recomendada nº 25R-03 – Como estimar perda de produtividade de mão de obra em pleitos de construção:

A estimativa da Contratada ou, em alternativa, o valor em dólares dos pedidos de pagamento, valores contratuais ou preços unitários podem ser usados para determinar as horas de mão de obra, quando elas foram despendidas e, possivelmente, em quais atividades. Unidades físicas de trabalho concluídas multiplicadas pelos índices de orçamento unitário podem ser usadas para determinar horas agregadas. As horas agregadas são então comparadas às horas reais gastas para o período do impacto e a diferença entre as duas pode ser usada para calcular a perda de produtividade vivenciada. [AACE, 2004, p. 17]

Como se sabe, a AVA trabalha com três tipos de custos: o valor planejado (VP), o custo real (CR) e o valor agregado (VA), ou seja, o valor medido.

Na primeira parte da descrição acima, é ao valor planejado a que se refere; portanto, em uma determinada data e um determinado valor de VP, há um equivalente de horas de mão de obra e equipamento, que guarda proporção (aqui entra a segunda parte da descrição acima) com as horas de mão de obra e equipamento contidas no valor agregado (VA), medido na mesma data. Seja a figura a seguir.

Figura 6 – Perda de produtividade por AVA



Fonte: Elaborado pelo autor

$$\frac{VP}{Cp(mo + eq)} = \frac{VA}{Ca(mo + eq)}$$

$$Ca(mo + eq) = \frac{VA}{VP} \times Cp(mo + eq)$$

Sendo:

- Custo planejado: $CP = \frac{VP}{1+BDI}$ e $Cp(mo + eq) = \frac{VP}{1+BDI} \times \%(MO + Eq.)$
- Custo agregado: $CA = \frac{VA}{1+BDI}$ e $Ca(mo + eq) = \frac{VA}{1+BDI} \times \%(MO + Eq.)$

$$Perda\ de\ produtividade = \frac{VP}{1+BDI} \times \%(MO + Eq.) - \frac{VA}{1+BDI} \times \%(MO + Eq.)$$

$$Perda\ de\ produtividade = VP \times \left[\frac{\%(MO + Eq.)}{1+BDI} \right] - VA \times \left[\frac{\%(MO + Eq.)}{1+BDI} \right]$$

$$Perda\ de\ produtividade = (VP - VA)x \left[\frac{\%(MO + Eq.)}{(1 + BDI)} \right]$$

A expressão acima é válida se o critério de avaliação da perda de produtividade for pelo custo previsto; se o critério for pelo custo incorrido (aquele informado pelo requerente), a expressão da perda de produtividade é a seguinte:

$$Perda\ de\ produtividade = CR - VAx \left[\frac{\%(MO + Eq.)}{(1 + BDI)} \right]$$

Assim, para cada caso, o custo da perda de produtividade é dado pelas seguintes expressões:

$$VPP_{ci} = CR - VAx \left[\frac{\%(MO+Eq)}{(1+BDI)} \right] \quad (1)$$

$$VPP_{cp} = (VP - VA)x \left[\frac{\%(MO+Eq)}{(1+BDI)} \right] \quad (2)$$

Onde:

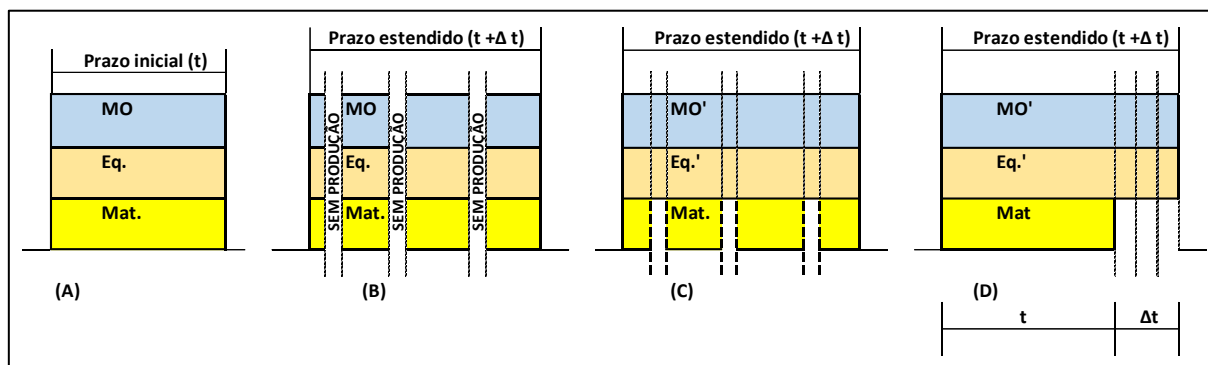
- VPPci: Valor da perda de produtividade pelo custo incorrido;
- VPPcp: Valor da perda de produtividade pelo custo previsto;
- CR: custo real (custo incorrido), isento de BDI e informado pelo requerente;
- VP: Valor planejado para a data em questão;
- VA: valor agregado (medido) para a data em questão
- %(MO+Eq): somatório dos percentuais de mão de obra e equipamentos, conforme os seus pesos na composição do custo total;
- BDI: taxa de Benefícios e Despesas Indiretas adotada, em valor decimal.

Como pode ser visto, a avaliação da perda de produtividade utilizando AVA resume-se a uma proporção entre os valores planejado e medido, sem levar em conta a intensidade do impacto até a data em questão, o que justifica a utilização desta metodologia somente quando não for possível o uso do cálculo *Measured Mile* ou a adoção da produtividade baseline.

2.3.2 Perda de produtividade usando o cálculo Measured Mile

A figura a seguir ilustra as parcelas do custo total que sofrem a ação do tempo.

Figura 7 – Ociosidade da mão de obra e equipamentos



Fonte: Elaborado pelo autor

Em (A) tem-se o projeto conforme planejado: insumos e tempos conforme previstos. Em (B) a ação de micro e pequenas interrupções, em que não há produção, embora não se configure em paralisação total. Em (C) a tradução do que de fato ocorreu em (B): embora o material não tenha sido utilizado nos períodos em que não houve produção, os equipamentos e toda a mão de obra estavam mobilizadas, representando desembolsos financeiros realizados, de forma constante, de modo que o total de cada parcela foi alterado: MO => MO' e Eq. => Eq.'. Em (D), concluído o projeto, — que para este exemplo não sofreu alteração de escopo —, tem-se a mesma quantidade de material, mas uma quantidade ampliada, igual a Δt , de utilização da mão de obra e dos equipamentos. No exemplo acima, houve extensão de prazo, o que não é condição necessária para haver perda de produtividade, visto que, para essa, basta a ocorrência da ociosidade (ausência de produção), que pode ter seus efeitos temporais anulados por uma aceleração, mas não os seus efeitos financeiros, que podem ser, inclusive, agravados com essa medida.

De forma generalizada, Δt representa, então, o tempo ocioso, em que a mão de obra e equipamentos não foram utilizados na produção, desequilibrando a equação contratual.

A partir do custo total do projeto, é possível segregar os insumos de cada uma dessas parcelas, de forma que se possa obter a participação percentual individual na formação do custo total. Como não houve produção, o custo do combustível dos equipamentos deve ser deduzido e somado à parcela do material.

A expressão a seguir traduz essa composição.

Valor total do contrato (VT) = Custo total do contrato (CT) x (1+BDI) => VT = CTx(1+BDI) => CT = VT/(1+BDI)

CT = MO + Eq. + Mat.

$\alpha = MO/CT = \%MO$; $\beta = Eq./CT = \%Eq.$ e $\gamma = Mat./CT = \%Mat.$

Parcela do custo total que sofre ação do tempo: CTt = [VT/(1+BDI)]x(%MO+%Eq.)

Para um projeto com prazo de execução igual a "t" meses, o custo unitário mensal de CTt é dado por:

$$CTt(mês) = \frac{\left\{ \left[\frac{VT}{(1+BDI)} \right] x (\%MO + \%Eq.) \right\}}{t} \quad (4)$$

E, sendo Δt o total de tempo em que houve ociosidade de mão de obra e equipamento, o custo total dessa perda de produtividade é dado por:

$$VPPmm = \Delta t x \frac{\left\{ \left[\frac{VT}{(1+BDI)} \right] x (\%MO + \%Eq.) \right\}}{t} \quad (5)$$

Onde:

- VPPmm: Valor da perda de produtividade pelo cálculo *Measured Mile*;
- VT: Valor total do projeto;
- BDI: taxa de Benefícios e Despesas Indiretas adotada, em valor decimal.
- %MO e %Eq.: percentuais de mão de obra e equipamentos (descontado o combustível), conforme os seus pesos na composição do custo total;
- t : prazo inicial pactuado de execução do projeto;
- Δt : tempo de ociosidade da mão de obra e equipamentos.

2.4 CAUSAS E RESPONSABILIDADES

2.4.1 Identificação das causas

Para o projeto-exemplo considerado, a análise dos registros revelou condutas que redundaram nas causas desencadeadoras da ociosidade da mão de obra e equipamentos, o que repercutiu em perda de produtividade. Após aplicação da técnica “5 porquês”, as causas foram assim definidas:

- **PROJETOS** – os projetos fornecidos pela contratante mostraram-se deficientes, necessitando de constantes complementações, demandando tempo adicional tanto para a correção, quanto para a aprovação das alterações, impactando o prazo de execução pactuado.
- **EQUIPAMENTOS** – os equipamentos não foram mobilizados tempestivamente, bem como se observou a não execução de serviços em decorrência de falta de equipamento na patrulha, ocasionado ou por consertos ou mobilização não atempada.
- **LICENÇAS AMBIENTAIS** – observou-se atraso nas providências necessárias à obtenção de licenças ambientais, tanto as de responsabilidade da contratante, quanto as da contratada.
- **INTERFERÊNCIA DE CONCESSIONÁRIAS** – as tratativas com as concessionárias de serviços públicos, presentes no local da obra, só foram iniciadas após a emissão da ordem de início de serviço, o que impediu o início da obra por 45 dias.
- **SUPRIMENTO DE MATERIAL** – observou-se equipes paradas, à espera da chegada de material no canteiro, decorrente da aquisição tardia de insumos.

2.4.2 Estabelecimento do nexo de causalidade

A partir das causas identificadas e em cotejo com as condutas das partes, é possível elaborar o quadro a seguir, acerca da imputação da responsabilidade técnica da parte.

Quadro 1 – Estabelecimento do nexo de causalidade

CAUSA	CONDUTA	EFEITO	NEXO	RESPONSABILIDADE
Projetos	A contratante forneceu projetos deficientes	Ociosidade das equipes decorrente do tempo de espera para correção dos projetos	A conduta da contratante ensejou o efeito	Contratante
Equipamentos	A contratada não mobilizou os equipamentos no tempo, na qualidade e na quantidade previstas	Ociosidade das equipes decorrente do tempo de espera por não contar com os equipamentos previstos	A conduta da contratada ensejou o efeito	Contratada
Lic. Ambientais	Contratada e Contratante não tomaram providências tempestivas para obtenção de suas licenças ambientais	Ociosidade das equipes decorrente do tempo de espera por falta de licenças ambientais	A conduta das partes ensejou o efeito	Contratante e Contratada
Concessionárias	A contratante iniciou as tratativas para remoção da interferência de concessionárias	Ociosidade das equipes decorrente do tempo de espera para retirada das	A conduta da contratante ensejou o efeito	Contratante

CAUSA	CONDUTA	EFEITO	NEXO	RESPONSABILIDADE
	após a emissão da ordem de serviço	interferências das concessionárias		
Suprimentos	A contratada iniciava seu processo de compra em data que comprometia a chegada de material a tempo no canteiro	Ociosidade das equipes decorrente do tempo de espera para chegada de material no canteiro	A conduta da contratada ensejou o efeito	Contratada

Fonte: Elaborado pelo autor

2.4.3 Importância das causas e responsabilidades das partes

A Metodologia de Quantificação de Responsabilidades Concorrentes, objeto do Boletim Técnico BTec - 2021/011 do IBAPE Nacional, quantifica a importância das causas a partir da aplicação de um questionário onde todas as causas são confrontadas entre si, mediante a aplicação de pesos (de 1 a 9), conforme o julgamento do especialista quanto à importância relativa entre as causas. Entende-se que esse grau de subjetividade, própria da metodologia, só se justifica em situações em que as causas não possam ter o dano por ela provocado quantificado objetivamente, como na situação das causas envolvidas em acidentes de construção. No caso de pleito por perda de produtividade, a extensão do dano é mensurada pelo tempo de ociosidade, ao qual está associado a um custo financeiro, possibilitando a quantificação objetiva de cada causa envolvida. Nesse sentido, a importância das causas será mensurada objetivamente, a partir da sua participação individual na extensão total do tempo ocioso, cuja forma de cálculo já foi demonstrada.

Como forma de materializar essa importância, apresenta-se a estrutura a seguir, a partir dos elementos já identificados neste projeto-exemplo.

Tabela 7 – Importância das causas e responsabilidades das partes

mês	Ociosidade (meses)	Causas					Há Registro?
		Projetos	Equipamentos	Lic. Ambientais	Concessionárias	Suprimentos	
1	0,25	1	1	1	1	1	S
2	0,35		1	1	0,5		S
3	0,20	1					S
4	0,37					1	S
5	0,27	1					S
6	0,06					1	S
7	0,33		1				S
8	0,40	1					S
9	0,24						N
10	0,23	1					S
11	0,35	1					S
12	0,28				1		S
13	0,13			1			S
14	-0,20						N/A
15	-0,44						N/A
16	0,00						N/A
Total	3,44	1,69	0,93	0,73	0,7	0,68	4,73
	Importância	0,36	0,2	0,15	0,15	0,14	1,00
Responsabilidade das Partes	Contratante	1	0	0,5	1	0	58,2%
	Contratada	0	1	0,5	0	1	41,8%
	Total	1	1	1	1	1	100,0%

Fonte: Elaborado pelo autor

A coluna referente a ociosidade contém os valores de impacto calculados na Tabela 6.

De acordo com a pesquisa registral, são assinalados com “1” os meses de ocorrência das causas. No caso específico da Concessionária, foi registrado que os efeitos da inação da contratante perduraram pelos 45 dias iniciais dessa forma, computou-se a ocorrência no mês 1, integral, e 0,5 dias no mês 2. Após, no mês 12, essa causa voltou a ocorrer.

Observe-se que no mês 9 não foi encontrado registro que justificasse o impacto de 0,24 meses; nesse caso, admite-se que a melhor estimativa de divisão da

reponsabilidade por esta parcela mensal de impacto é a obtida sem considerar esse mês, ou seja, 58,2% de 0,24 para a contratante e 41,8% de 0,24 para a contratada.

A pontuação de cada causa é obtida pela multiplicação da ociosidade mensal por sua frequência. A totalização dessas pontuações foi igual a 4,73. Normalizando a pontuação de cada causa com esse total, obtém-se a importância percentual de cada causa na contribuição do tempo total de ociosidade.

Na parte final, a responsabilidade das partes é decorrente do nexo de causalidade configurado de acordo com o Quadro 1.

A responsabilidade das partes é, por fim, calculada mediante a soma dos percentuais das causas pelas quais a parte foi considerada responsável. Observe-se que para a causa referente ao licenciamento ambiental, houve responsabilidade concorrente, cabendo, portanto, a metade do percentual dessa causa para cada parte.

Esses valores finais podem ser rearrumados de acordo com a tabela seguinte.

Tabela 8 – Responsabilidades das partes

Causas	Importância	Ociosidade decorrente (meses)	Responsabilidade	
			Contratante	Contratada
Projetos	35,7%	1,23	35,7%	
Equipamentos	19,7%	0,68		19,7%
Lic. Ambientais	15,3%	0,53	7,7%	7,7%
Concessionárias	14,9%	0,51	14,9%	
Suprimentos	14,4%	0,49		14,4%
Total	100,0%	3,44	58,2%	41,8%

Fonte: Elaborado pelo autor

A contratante responde por 58,2% do dano decorrente da perda de produtividade, enquanto a contratada responde por 41,8%.

2.5 Avaliação do pleito de reequilíbrio decorrente da perda de produtividade

Considere-se que, para este projeto-exemplo, o BDI ofertado na proposta foi de 25%, e que a participação da mão de obra e dos equipamentos somam 50% do custo total do projeto. A partir dessas informações e acréscimos dos resultados já obtidos com a aplicação da presente metodologia, tem-se o seguinte cenário:

- Valor do projeto: \$ 100
- Prazo de execução pactuado: 8 meses
- Tempo de ociosidade: 3,44 meses
- Participação (%MO+%Eq.): 50% do custo total

2.5.1 Valor da perda de produtividade utilizando a Measured Mile

Aplicando a fórmula (5):

$$VPPmm = \Delta t \times \frac{\left\{ \left[\frac{VT}{(1+BDI)} \right] \times (\%MO + \%Eq.) \right\}}{t} \quad (5)$$

E substituindo com os valores acima, vem:

$$VPPmm = 3,44 \times \left\{ \left[\frac{100}{(1+0,25)} \right] \times 0,5 \right\} / 8 = \$ 17,20$$

O valor da perda de produtividade foi de \$ 17,20. Considerando a responsabilidade das partes na formação desse dano financeiro, a participação de cada uma é assim calculada:

- Contratante: 58,2% x \$ 17,20 = \$ 10,01
- Contratada: 41,8% x \$ 17,20 = \$ 7,19

Dessa forma, cabe de indenização à contratada a importância de \$ 10,01, como devidos ao necessário reequilíbrio econômico-financeiro do contrato, em decorrência da perda de produtividade verificada.

2.5.2 Valor da perda de produtividade utilizando a produtividade baseline e AVA

Apenas a título de objeto de pesquisa, uma vez que foi possível mensurar a perda de produtividade utilizando-se o procedimento de excelência, o cálculo *Measured Mile*, será procedida a valoração dessa perda com o uso da produtividade pactuada (baseline) e da AVA.

Inicialmente, na partir da Tabela 1 é calculado o impacto considerando a produtividade baseline. O somatório do impacto foi igual a 4,13 meses. Substituindo na fórmula (5), vem:

$$VPPBL = 4,13 \times \{[100/(1+0,25)] \times 0,5\} / 8 = \$ 20,65$$

O cálculo do valor da perda de produtividade com o uso da AVA, com base no custo previsto, é calculado por meio da fórmula (2)

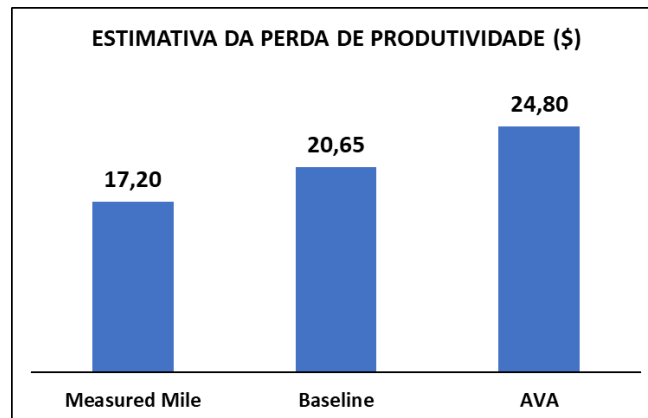
$$VPP_{cp} = (VP - VA) \times \left[\frac{\% (MO + Eq)}{(1 + BDI)} \right] \quad (2)$$

Os últimos valores disponíveis de VP e VA, e que servirão de proporcionalidade para o cálculo da perda de produtividade, são aqueles alocados no mês 8 da tabela acima, respectivamente \$ 100 e \$ 38. Substituindo os valores na expressão cima, vem:

$$VPP_{cp} = (100 - 38) \times [0,5/(1+0,25)] = \$ 24,80$$

2.5.3 Comparativo dos valores estimados conforme as metodologias

Figura 7 – Comparativo de valores da perda de produtividade



Fonte: Elaborado pelo autor

Observa-se que, para o mesmo projeto, e considerando como referencial o valor obtido com o cálculo *Measured Mile*, a valoração da perda de produtividade variou, para maior, entre 20%, se considerado o valor obtido pela produtividade baseline, e 44%, se considerado o valor obtido com o uso da AVA.

4. CONCLUSÃO

O artigo proporcionou uma abordagem completa de um pleito de reequilíbrio econômico-financeiro contratual devido à perda de produtividade, em seus três aspectos basilares e necessários à formação da responsabilidade técnica, quais sejam o dano financeiro verificado, a conduta das partes e o nexo de causalidade ligando a conduta ao dano.

Das metodologias recomendadas por protocolos internacionais para valoração dessa perda, analisou-se o cálculo *Measured Mile*, o uso da produtividade baseline (pactuada), quando a *Measured Mile* não for viável, e a Análise de Valor Agregado – AVA.

Por meio de testes estatísticos, inovou-se na detecção dos períodos não impactados, necessários ao cálculo *Measured Mile*, bem como no saneamento da

amostra, com o uso de clusterização. Tais procedimentos eliminam a subjetividade comumente envolvida nessas etapas. Com a detecção do valor da produtividade natural, foi possível, com apoio da técnica do Prazo Agregado, calcular os períodos mensais de ociosidade decorrentes exclusivamente dos impactos, os quais advêm de causas cuja responsabilidade foram também objeto de análise. O somatório das parcelas mensais de ociosidade totaliza a extensão do dano, um dos elementos da valoração da perda de produtividade.

Foram apresentadas técnicas de descobrimento de causa(s)-raiz e de realce do nexo de causalidade, como a “5 porquês”. A importância das causas deu-se de forma objetiva, a partir da frequência revelada pela investigação dos registros, sem a subjetividade inerente à Metodologia de Quantificação de Responsabilidades Concorrentes. A alocação da responsabilidade da parte pela causa deu-se com amparo na análise do nexo de causalidade. Com a importância de cada causa na formação do dano, e a responsabilidade da parte pela ocorrência da causa, foi possível quantificar a responsabilidade das partes pelo dano total verificado.

O aprofundamento da pesquisa teórica possibilitou assegurar que a perda de produtividade se materializa na ociosidade da mão de obra e dos equipamentos. Uma vez calculada a participação dessas parcelas no custo total do projeto, é possível estabelecer, a partir do prazo inicialmente pactuado, o seu custo mensal. E, uma vez conhecida a extensão do dano, o valor da perda de produtividade equivale ao produto dessas duas grandezas.

Havendo concorrência das partes na formação do dano financeiro, o mesmo deve ser repartido na proporção de suas responsabilidades, conforme assim calculadas.

REFERÊNCIAS BIBLIOGRÁFICAS

AACE (BRASIL). Prática Recomendada nº 25R-03 – Como estimar perda de produtividade de mão de obra em pleitos de construção. Rio de Janeiro: AACE, 2004. 39 p.

_____. Prática Recomendada nº 45R-08: Métodos de Proteção contra Pleitos Relacionados a Cronograma, Rio de Janeiro: AACE, 2009, 13 p.

BRAITHWAITE, Simon. Perda de produtividade e projetos interrompidos. São Paulo: AACE Brasil, 2017. 59 slides, color.

CARDOSO, Roberto Sales. A qualidade do projeto básico. In: FILIPE BONALDO ALVES (São Paulo) (org.). Pareceres de engenharia: gestão de contratos e pleitos. 1. ed. São Paulo: AACE Brasil, 2016. 244 p.

COSTA, Aquiles Augusto Diniz Martins da. Gestão contratual e o equilíbrio econômico-financeiro. In: FILIPE BONALDO ALVES (São Paulo) (org.). Pareceres de engenharia: gestão de contratos e pleitos. 1. ed. São Paulo: AACE Brasil, 2016. 244 p.

ENAP. Análise e Melhoria de Processos – Metodologia MASP: Caderno de ferramentas. 1. ed. Brasília: ENAP, 2015. 23 p.

GREBLER, Eduardo. Reflexões sobre os aspectos jurídicos da improdutividade no contrato de construção. In: MARCONDES, Fernando et al. Direito da Construção: Estudo sobre as várias áreas do Direito aplicadas ao mercado da Construção. 1. ed. São Paulo: PINI, 2014.

OCTAVIO GALVÃO NETO (São Paulo) (org.). Diretrizes para análise de produtividade em perícias de engenharia. 1. ed. São Paulo: Ibape Nacional, 2021. 51 p.

SALOMÃO, Ricardo. Desequilíbrio econômico-financeiro em contratos de obras e serviços de engenharia. In: FIGUEIREDO, Flávio Fernando de et al. Perícias de engenharia: uma visão contemporânea. 1. ed. São Paulo: Leud, 2022. 320 p.

SOUTO, Leandro Souza; SOUTO, Eduardo de Sousa. Análise de Dados Ecológicos: um guia simplificado de bioestatística para estudantes e profissionais da área ambiental, 1 ed., Brasília: Ed. do autor, 2020. 136 p.

TISAKA, Maçahico. Como Evitar Prejuízos em Obras de Construção: Construction Claim . 1. ed. São Paulo: PINI, 2011.

**Antoniél Campos**

Engenheiro Civil, especialista em avaliações e perícias de engenharia. Servidor público federal.

Vencedor das medalhas Eurico Ribeiro, Hélio de Caires, Ibá Ilha Moreira (COBREAP) e medalha Domingos de Saboya (SOBREA), todas das áreas de avaliações, perícias e meio ambiente.

Vencedor do prêmio Min. Guilherme Palmeira pelo TCU sobre soluções consensuais de conflitos.

Premiações em congressos e simpósios:

- Prêmio Ministro Guilherme Palmeira - melhor monografia na categoria sociedade civil sobre soluções consensuais de controvérsias e prevenção de conflitos no âmbito da administração pública – TCU (2023).
- Medalha Eurico Ribeiro - melhor trabalho de perícias apresentado no XXII COBREAP (2023);
- Medalha Ibá Ilha Moreira Filho – melhor trabalho na área ambiental apresentado no XXII COBREAP (2023);
- Medalha Domingos Saboya - melhor trabalho apresentado no X Simpósio da SOBREA (2022);
- Medalha Hélio de Caires - melhor trabalho de avaliações apresentado no XX COBREAP (2019);
- Menção honrosa - trabalho apresentado na área de perícias no XIX COBREAP (2017).

Acesse este trabalho premiado no XXII COBREAP através do link na biblioteca do IBAPE:

<https://ibape-nacional.com.br/biblioteca/wp-content/uploads/2023/12/abordagem-tridimensional-do-pleito-em-pericias-envolvendo-analise-de-perda-de-produtividade.pdf>

II. ESTIMATIVA DE VALORES DE IMÓVEIS DA UNIÃO COM TÉCNICAS DE APRENDIZAGEM DE MÁQUINA E COMPONENTES ESPACIAIS

***José Nilo Alves de Sousa Neto
Marcelo Ladeira (orientador PPCA UNB)
Juliana Miura***

Os bens tangíveis do tipo imobiliário possuem características intrínsecas e extrínsecas que os tornam únicos. Tal unicidade faz com que o processo de mensuração de seus valores mais prováveis e justos de mercado constitua matéria efetivamente complexa.

O Exército Brasileiro (EB) possui mais de vinte mil parcelas imobiliárias sob sua administração e necessita de ferramenta para estimativa, em massa, dos valores desses bens patrimoniais a fim de suportar processos decisórios a nível estratégico visando à arrecadação de recursos para o Tesouro Nacional ou à permuta por outros ativos que melhor atendam suas atuais demandas.

No âmbito do EB, a Diretoria de Patrimônio Imobiliário e Meio Ambiente (DPIMA) representa a instância técnica responsável por analisar e aprovar as avaliações imobiliárias dos imóveis administrados pelo EB e daqueles de interesse da Força Terrestre. Já a nível da União, a Secretaria do Patrimônio da União (SPU) faz a gestão dos ativos imobiliários. Conforme a norma ABNT NBR 14653, o Método Comparativo Direto de Dados de Mercado deve ser preferencialmente utilizado; ele propõe a inferência dos valores de mercado dos imóveis com base na oferta ou na transação de outros imóveis com determinado nível de similaridade. A valoração de imóveis, ainda que de maneira estimativa, não é uma tarefa trivial. Trabalhos realizados mostram que a localização costuma se comportar como uma variável significativa nos modelos construídos e que os imóveis ofertados ou recentemente transacionados nas adjacências daquele em análise têm influência sobre seu valor.

1. FUNDAMENTAÇÃO TEÓRICA

Segundo [Alves Dantas et al. 2010] e [Barros Antunes Campos and Almeida 2018], considerando modelo tradicional de preços hedônicos definido inicialmente por [Rosen 1974], o valor de mercado das unidades imobiliárias residenciais pode ser explicado por:

$$P = f(E, L, T, \beta) + \varepsilon \quad (\text{eq. 1})$$

Sendo o preço da habitação (P) função das suas características estruturais (E), locais (L) e da época em que foi demandado (T); f é um operador indicativo da forma funcional, β são parâmetros e ε os erros aleatórios do modelo.

Sob essa ótica, imóveis com outras vocações além da residencial também poderiam ter seus valores prováveis de mercado construídos a partir de características intrínsecas e extrínsecas.

2. TRABALHOS RELACIONADOS

Há diferentes abordagens com foco na valoração de imóveis presentes na literatura. [Alves Dantas et al. 2010] implementam modelos de abrangência geográfica municipal combinados a econometria espacial. A linearidade é considerada como um pressuposto e diversas análises são realizadas com e sem componentes espaciais. Os modelos desenvolvidos chegam a coeficientes de determinação (R^2) da ordem de 90%.

[Park e Bae 2015] aplicam e analisam algumas técnicas de aprendizagem de máquina, inclusive não lineares, no mercado de imóveis residenciais em Fairfax County, Virginia, nos Estados Unidos da América. Algoritmos C4.5, RIPPER, Naïve Bayes e AdaBoost são comparados e o RIPPER apresenta a melhor performance geral entre eles. São limitações do trabalho a abrangência geográfica restrita e a carência de interpretabilidade das variáveis independentes frente à variável explicada.

[Kiely and Bastian 2020] desenvolveram modelos regressivos de predição de valores de imóveis de diferentes tipologias em Nova York com uso de redes neurais artificiais (ANN) e conceitos de agregação espacial de variáveis. Os R^2 mais elevados obtidos são da ordem de 50%. [Dewan et al. 2019] combinaram ANN a modelos autorregressivos espaciais de caráter global (SAR) e os resultados obtidos indicaram a importância dos componentes espaciais.

[Hagenauer and Helbich 2022] propuseram um modelo de ANN combinado a regressões espaciais locais ponderadas, denominado Geographically Weighted Artificial Neural Network (GWANN). Quando aplicado ao domínio real de casas unifamiliares na Áustria, os maiores R^2 médios foram de aproximadamente 45%.

[Tchunte and Nyawa 2022] compararam o desempenho de 7 algoritmos de aprendizagem de máquina, lineares e não lineares, aplicados com e sem geocodificação a uma base de dados disponibilizada pelo Governo da França de valores de imóveis. Os modelos não lineares com geocodificação apresentaram maiores coeficientes de determinação, aproximadamente 74%. A interpretabilidade utilizando o valor de Shapley pode ser garantida por meio de definições colaborativas da Teoria dos Jogos. Conforme [Lundberg et al. 2020] e [Chakraborty et al. 2020], o valor de Shapley é a contribuição marginal média de cada valor de atributo em todas as combinações possíveis de características. Os recursos com grandes valores de Shapley absolutos são considerados importantes. Para obter a importância relativa global da variável, calcula-se a média de seus valores absolutos de Shapley em todas as ocorrências; normalmente, as medidas de tendência central são ordenadas em importância decrescente e plotadas. Fundamentado em pesquisa bibliográfica realizada, a Tabela 1 apresenta um resumo comparativo entre a solução proposta e os trabalhos da literatura.

Tabela 1 - Comparação entre trabalhos relacionados

Trabalho	$R^2 > 57\%$	Abrangência nível País	Modelagem linear	Modelagem não linear	Ajuste de hiperparâmetros	Componentes espaciais	Mais de uma vocalização de imóvel	Interpretabilidade
[Alves Dantas et al. 2010]	✓		✓			✓		✓
[Park and Bae 2015]	✓		✓	✓	✓			
[Kiely and Bastian 2020]			✓	✓	✓	✓	✓	
[Hagenauer and Helbich 2022]		✓	✓	✓		✓		
[Tchunte and Nyawa 2022]	✓	✓	✓	✓	✓	✓		
Solução proposta	✓	✓	✓	✓	✓	✓	✓	✓

Fonte: Elaborado pelos Autores

3. METODOLOGIA

Nesta seção, discorre-se acerca da extração de dados, de sua análise exploratória multivariada e da concepção inicial dos modelos associados à solução proposta.

3.1 Extração de Dados

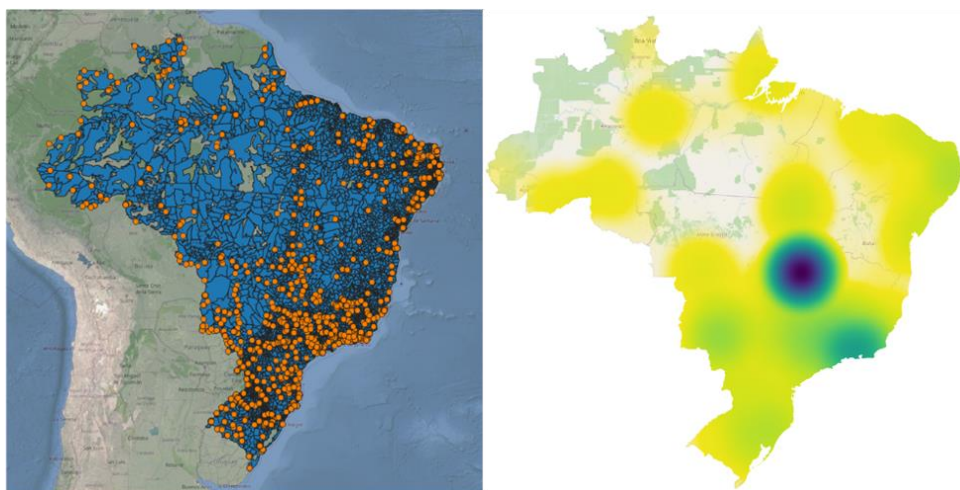
Os dados diretamente associados a cada um dos imóveis foram extraídos de bases de dados de acesso restrito do EB, com 258 imóveis anotados, e da SPU, com 4981 imóveis anotados. As ocorrências correspondem a imóveis avaliados e com laudos devidamente homologados pelas referidas instituições.

Todos os imóveis do EB foram georreferenciados em formato shapefile. Como as instâncias da SPU não continham polígonos georreferenciados, optou-se por geocodificar o campo textual Endereço. No processo de geocodificação, 925 instâncias, 17,6% do total coletado, não puderam ter suas coordenadas espaciais extraídas no Google Earth Pro. No intuito de amenizar os efeitos de tal perda, 284 ocorrências foram recuperadas manualmente, resultando em 258 instâncias do EB (5,6%) e 4340 da SPU (94,4%).

Essas bases foram enriquecidas com atributos socioeconômicos do Censo 2010 do Instituto Brasileiro de Geografia e Estatística (IBGE) tratados por área de ponderação (AP). Para associar os imóveis às AP, realizaram-se operações de união

espacial no QGIS 3.16 Hannover entre os pontos relativos aos centroides das propriedades e os polígonos das AP, extraídos de [Furtado 2020]. As 4598 instâncias resultantes e as AP encontram-se representadas na Figura 1. Nos casos em que o centroide do imóvel ficou em área descoberta pelas AP, utilizou-se a média dos atributos das AP adjacentes.

Figura 1 – AP do IBGE (em azul) sobrepostas pelos centroides de instâncias do EB e da SPU (à esquerda) e mapa densidade de imóveis que compõem a amostra urbana (à direita)



Fonte: Elaborado pelos Autores

Por fim, foram coletadas as quantidades de diversas tipologias de pontos de interesse, tais como estações de metrô, parques, shopping centers e hospitais, em um raio de 400 metros de cada propriedade, correspondente a aproximadamente cinco (5) minutos de caminhada a pé, tomando as coordenadas dos centroides dos imóveis como referência. Para tal, consumiu-se a *Application Programming Interface (API) Google Places Nearby Search*.

3.2 Tratamento de Instâncias

Foram desenvolvidas três abordagens para tratamento das instâncias coletadas, em decorrência da elevada dispersão observada das variáveis Área do Terreno, Valor Total Atualizado e Valor Unitário Atualizado, estas duas últimas atualizadas a janeiro de 2022 com uso do índice FipeZap Brasil, respeitando-se o tipo de uso do bem em questão. Segue uma síntese das abordagens abaixo:

(1) **Abordagem 1** (A₁) Remoção de outliers pelo método de Tukey baseada na variável Valor Unitário Atualizado (R\$/m²) e utilização deste atributo como variável explicada.

(2) **Abordagem 2** (A₂) Transformação das variáveis Área do Terreno com logaritmo neperiano (ln), sem remover outliers, e ln de Valor Total Atualizado como variável explicada.

(3) **Abordagem 3** (A₃) Transformação do atributo Valor Unitário Atualizado com ln, remoção da variável Área do Terreno e ln de Valor Unitário Atualizado como variável explicada.

A abordagem A₂ apresentou melhores resultados, indicando necessidade de eliminação de apenas três instâncias, todas da SPU, das 4598 provenientes do processo de geocodificação. Na modelagem mais simples, linear, as 4595 ocorrências resultantes se mostraram não influenciadas sob avaliação da distância de Cook.

3.3 Modelagem

Os modelos urbanos propostos utilizaram os atributos constantes na Figura 2 após rigoroso processo de seleção de variáveis, sobre o qual se discorrerá na Subseção 4.4, e se valeram das transformações matemáticas da abordagem A₂. Eles foram implementados com uso dos seguintes algoritmos e bibliotecas:

- Regressão Linear Múltipla (OLS), biblioteca statsmodels versão 0.14.0 em Python;

- Regressão Espacial Spatial Two Stage Least Squares (S2SLS), conforme proposto por [Anselin 1988], pacote spreg versão 1.3.0 da biblioteca PySALem Python;
- SGDRegressor linear, biblioteca scikit-learn versão 1.2.2 em Python;
- Rede Neural Artificial ANN MLPRegressor não linear, biblioteca scikit-learn versão 1.2.2 em Python; e
- XGBRegressor não linear, biblioteca XGBoost versão 1.7.4 em Python.

Para o modelo de regressão espacial, foram criadas matrizes de pesos espaciais W fundamentadas no livro de [Anselin et al. 2014]. Testamos algumas distâncias como largura de banda δ a fim de definir os elementos diferentes de zero na matriz apenas para os casos $d_{ij} < \delta$, em que d_{ij} representa a distância euclidiana entre os elementos i e j considerando suas coordenadas geográficas e o raio de curvatura terrestre.

Para a amostra de imóveis urbanos, a distância de 1 km se mostrou um bom limite (após testes realizados entre 100 metros e 20 km), implicando boa significância fundamentada em t de Student para a matriz de vizinhança e melhora do R^2 em relação ao modelo OLS, o qual não considera a autocorrelação espacial, a influência que o valor de um imóvel exerce sobre os valores de seus vizinhos matriciais e vice-versa. Os pesos das matrizes espaciais foram calculados com função de decaimento igual ao inverso da distância entre os imóveis.

3.4 Análise e Seleção de Atributos

O critério para consideração inicial de um atributo foi baseado na sua disponibilidade sob forma estruturada ou passível de estruturação e pela sua presença nos bancos de dados do EB e da SPU.

No caso de imóveis urbanos, a identificação dos subconjuntos de preditores úteis e alinhados ao pressuposto de não multicolinearidade exigido à aplicação de modelos de regressão linear múltipla foi realizada por meio das técnicas: análise de correlações bivariadas (máximo de 80% de correlação de Pearson, conforme indicado

na NBR 14653); regressão *forward-backward stepwise* (limites p-valor de entrada e de saída de 30%, fundamentados na NBR 14653); eliminação recursiva de atributos com estimador regressivo *Random Forest*; e análise de significâncias com procedimento não paramétrico *bootstrapping* bicaudal de 10 mil subamostras aleatórias por meio da construção de modelo conceitual de equações estruturais com mínimos quadrados parciais (PLS-SEM). À exceção do PLS-SEM, implementado com uso do software *SmartPLS 4*, as demais técnicas foram desenvolvidas com as bibliotecas *pandas* e *scikit-learn* em linguagem *Python*.

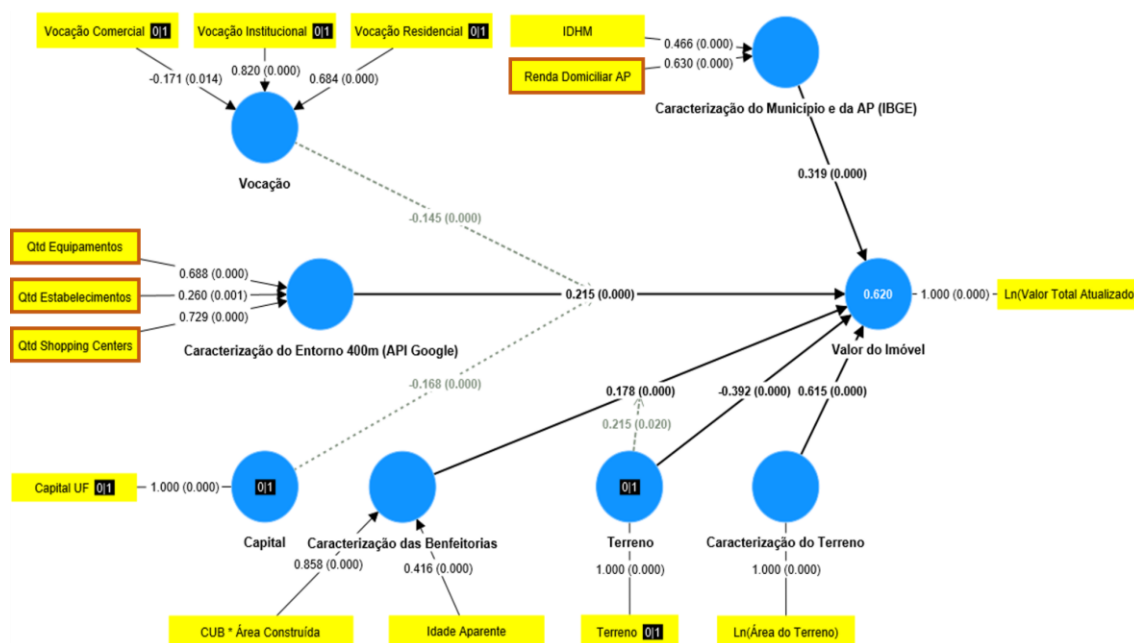
A aplicação dos algoritmos de seleção resultou nos atributos apresentados no modelo conceitual PLS-SEM representado na Figura 2, na qual constam os modelos de medida (externos), caracterizados pelo relacionamento entre os indicadores (em amarelo) e as variáveis latentes exógenas (em azul), e o modelo estrutural (interno), marcado pelo relacionamento entre as variáveis latentes exógenas, de controle, as moderadoras e a endógena. Os indicadores destacados com contorno em cor laranja são aqueles com nível de granularidade espacial intramunicipal.

Nos modelos de medida (externos), constam os pesos externos e os p-valores associados aos testes de hipótese de nulidade de regressores com base em distribuição t de Student, entre parênteses; já para os estruturais (internos), constam os coeficientes de caminho, os p-valores de seus testes de hipótese, entre parênteses, e o coeficiente de determinação para a variável latente endógena Ln (Valor Total Atualizado).

Analisando-se a relação entre o construto formativo Caracterização do Terreno e a variável latente endógena Valor do Imóvel, ilustrados na Figura 2, percebe-se que se trata da conexão mais forte do modelo estrutural. Ainda a respeito do PLS-SEM, já que os construtos se mostraram de natureza formativa, analisamos os fatores de inflação da variância (VIF) e tanto os relativos ao modelo interno quanto ao externo foram inferiores a 3,3. Apesar de nem todas as magnitudes dos coeficientes de caminho serem superiores a 0,2 em módulo, o que representa efeitos menores segundo as regras heurísticas de [Cohen 1988], todos os p-valores associados a seus testes de hipóteses e àqueles dos pesos e cargas externas do modelo externo foram inferiores a 5%.

O coeficiente de determinação do modelo conceitual PLS-SEM de 62,0% sugere poder explicativo moderado, de acordo com [Hair 2010]. Pontua-se a fragilidade dos modelos conceituais construídos em decorrência de a variável latente endógena Valor do Imóvel ter apenas um indicador a ela associado.

Figura 2 – Modelo conceitual PLS-SEM no software SmartPLS 4



Fonte: Elaborado pelos Autores

3.5 Tratamento de Variáveis

Foram aplicados testes de normalidade numéricos aos atributos explicativos selecionados. Tanto o teste de Shapiro-Wilk quanto o teste de Jarque-Bera indicaram, pelos p-valores obtidos, que há evidências de que as variáveis têm assimetria e curtose significativamente diferentes de uma distribuição normal.

Optou-se, portanto, pela aplicação de transformação do tipo *MinMaxScaler*, utilizando a biblioteca *scikit-learn* versão 1.2.2 na linguagem Python às variáveis explicativas. Na fase de pré-processamento, portanto, os atributos foram reescalados entre 0 e 1. Tal etapa evita que variáveis com maior amplitude de variação influenciem demasiadamente o modelo.

4. EXPERIMENTOS E RESULTADOS

Nesta seção, são apresentados os dados efetivamente utilizados, a configuração dos experimentos, os cenários de aplicação, os resultados dos modelos e uma análise das saídas à luz das principais métricas de avaliação adotadas, o coeficiente de determinação (R^2) e a raiz do erro quadrático médio (RMSE).

Vale ressaltar que a avaliação dos modelos de aprendizagem de máquina foi realizada com validação cruzada com 10 folds, após testes de 2 a 15 folds, programados na biblioteca scikit-learn versão 1.2.2 em Python. As médias e os desvios padrões das métricas foram calculados para o conjunto de treinamento e para o de teste. Os modelos OLS e de regressão espacial também foram avaliados quanto à probabilidade de nulidade simultânea de todos os seus regressores associados, o que representa uma hipótese grave, por meio da distribuição F de Fisher–Snedecor.

4.1 Bases de Dados e Cenários de Aplicação

Conforme etapas de pré-processamento discutidas Seção 4, a submissão de dados dos bancos do EB e da SPU, de maneira concatenada, a processos de geocodificação, enriquecimento espacial com informações de AP do IBGE (no caso de imóveis urbanos) e detecção e remoção de elementos influenciadores da amostra levou a reduções quantitativas das bases originais.

A amostra urbana efetivamente utilizada nos experimentos descritos nas subseções a seguir contemplou 4595 instâncias. O Distrito Federal concentra aproximadamente 29% dos dados.

As correlações de Pearson entre as variáveis explicativas utilizadas nos experimentos desta seção foram calculadas e não se observaram correlações parciais superiores a 80%, o que representa atendimento a um requisito normativo da NBR 14653.

4.2 Configuração dos Experimentos

A configuração dos principais hiperparâmetros foi feita cuidadosamente na modelagem de imóveis urbanos com uso de aprendizagem de máquina após testes realizados com a configuração padrão dos algoritmos e diversas outras com o *GridSearchCV* da biblioteca *scikit-learn* versão 1.2.2. A ANN foi configurada com função de ativação tangente hiperbólica e solver adam, baseado em gradiente estocástico, como otimizador de pesos. A partição entre os conjuntos de treinamento, validação e teste foi feita de forma estratificada, com base na vocação dos imóveis. Utilizaram-se termos de regularização a fim de se evitar *overfitting*.

4.3 Resultados Obtidos

Foram obtidos os coeficientes de determinação (R^2) e as raízes dos erros quadráticos médios (RMSE) especificados na Tabela 2.

Tabela 2 – Resultados das métricas de avaliação dos modelos urbanos para os dois cenários considerados

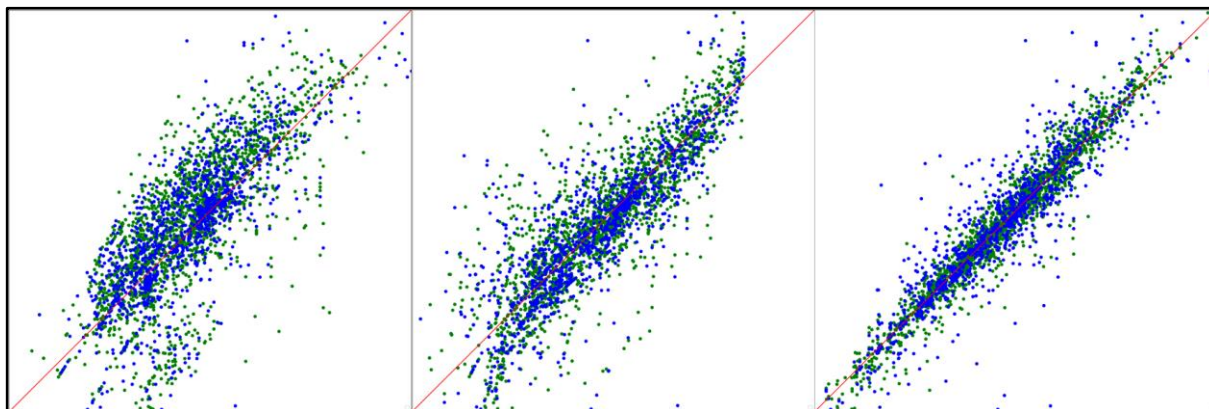
Modelo (Cenário)	R^2 médio 10 folds	dp R^2 10 folds	R^2 train	R^2 test	RMSE train	RMSE test
OLS (A)	NA	NA	61,9%	52,3%	1,28	1,73
Regressão Espacial (A)	NA	NA	63,4%	55,7%	1,26	1,39
<i>SGDRegressor</i> linear (A)	61,2%	4,7%	61,9%	61,1%	1,27	1,31
<i>SGDRegressor</i> linear (B)	60,7%	4,4%	61,7%	50,3%	1,29	1,35
ANN não linear (A)	73,8%	4,8%	74,8%	69,3%	1,04	1,17
ANN não linear (B)	71,8%	4,1%	73,8%	71,3%	1,06	1,02
<i>XGBRegressor</i> não linear (A)	86,2%	1,9%	95,7%	83,6%	0,43	0,85
<i>XGBRegressor</i> não linear (B)	86,7%	3,5%	94,8%	80,6%	0,47	0,84

NA= Não se aplica

Fonte: Elaborado pelos Autores

A maior concentração de pontos próximo às bissetrizes, destacadas em vermelho nos gráficos constantes na Figura 3, aliada às medidas de dispersão expostas na Tabela 2, são indícios de razoável assertividade das predições dos modelos não lineares. Na mesma figura, os dados de treinamento estão representados por pontos verdes e os de teste, por pontos azuis. Os eixos das ordenadas e das abscissas estão na mesma escala, na faixa de valores de 8 a 20.

Figura 3 – Gráficos de dispersão de ln de valores observados (ordenadas) versus ln de valores projetados pelos modelos (abscissas). Implementado com os algoritmos SGDRegressor (à esquerda), ANN MLPRegressor (ao centro) e XGBRegressor (à direita)



Fonte: Elaborado pelos Autores

Para o modelo de regressão espacial, analisou-se a dependência espacial, autocorrelação global, a partir do índice de Moran e do teste de Anselin-Kelejian. Ambos indicaram que as matrizes de pesos espaciais construídas com função de decaimento do inverso da distância entre os imóveis até o limite de 1 km, conforme detalhado na Subseção 4.3, são significantes.

Como os algoritmos *MLPRegressor* e *XGBRegressor* foram os únicos que atingiram consistentemente, nos dois cenários de aplicação, o objetivo específico de se obter coeficiente de determinação maior que 57%, um requisito normativo da NBR 14653, calculamos seus intervalos de confiança (IC) com 90% de grau de confiança. As semi-amplitudes inferiores e superiores dos intervalos calculados de -51,1% a +98,8%.

4.4 Interpretabilidade dos Resultados

Calcularam-se os coeficientes associados a cada variável e sua significância nas regressões lineares, com p-valores baseados na distribuição t de Student. Consequentemente, é possível se identificar as variáveis mais relevantes aos modelos

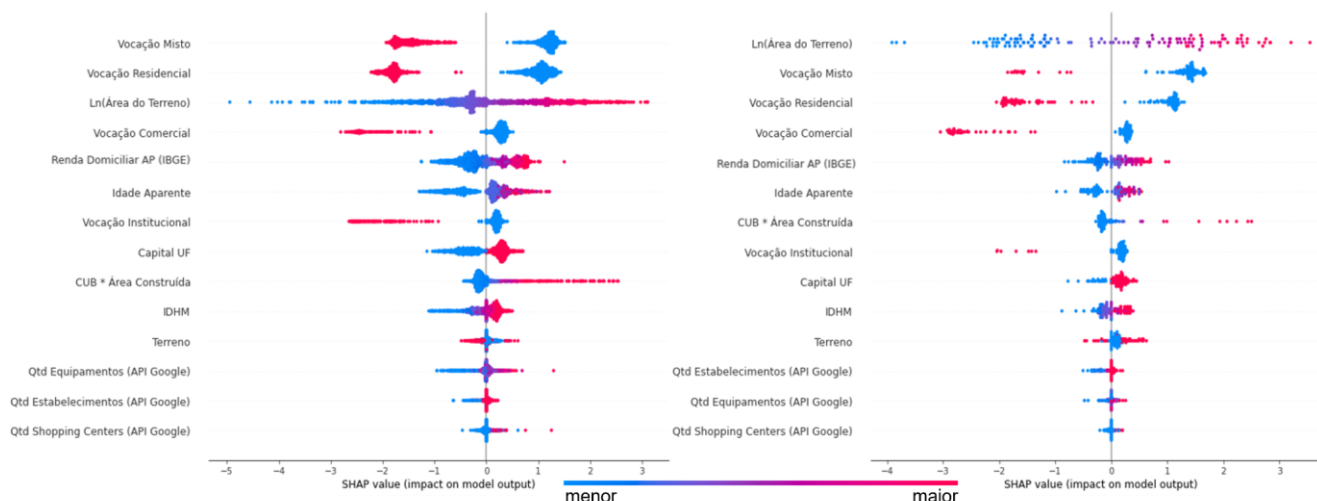
específicos mais convencionais, regressão linear múltipla (OLS) e regressão espacial; destaca-se o atributo Ln (Área do Terreno).

Cada ponto no gráfico da Figura 4 representa um valor de Shapley para atributos individuais e instâncias específicas. As variáveis mais importantes à formação do valor dos imóveis no modelo *ANN MLPRegressor* estão posicionadas na porção superior da imagem.

Verifica-se que a relação das variáveis explicativas com a variável dependente se mostra coerente com a realidade observada no domínio em estudo para os modelos de aprendizagem *SGDRegressor* e *ANN MLPRegressor*. Para o *XGBRegressor*, atributos como Terreno e Qtde Shopping Centers (API Google) apresentam relação direta e inversa com o valor do imóvel, respectivamente, o que não faz sentido sob a ótica do senso comum. Tentou-se aumentar a força dos termos de regularização L1 e L2, mas o referido comportamento não se alterou.

O que talvez possa explicar a mencionada inversão do *XGBRegressor* quanto à variável Terreno é a assimetria negativa acentuada da variável $CUB * \text{Área Construída}$ em relação ao ponto de equilíbrio do valor de Shapley. Ou seja, o simples fato de um imóvel não possuir benfeitorias construtivas já é tão penalizado pelo valor nulo de $CUB * \text{Área Construída}$, que Terreno entra com efeito compensatório. Tal comentário carece de análise de causalidade cuidadosa e detalhada.

Figura 4 – Valores de Shapley calculados para o algoritmo ANN MLPRegressor no Cenário A (à esquerda) e no Cenário B (à direita)



Fonte: Elaborado pelos Autores

4.5 Discussão dos Resultados

O PLS-SEM se mostrou um critério relevante e visualmente acessível no sentido de selecionar variáveis explicativas urbanas, contribuindo para a interpretabilidade da metodologia. O atributo Ln(Área do Terreno) se mostra como a variável quantitativa mais importante na formação de valor dos imóveis urbanos, em modelos lineares e não lineares. A partir da inclusão de variáveis com granularidade espacial nível AP (IBGE) e relativas ao entorno de 400 metros (API Google) dos imóveis urbanos, obteve-se incremento de poder de diferenciação entre ocorrências dentro de um mesmo município. Os referidos atributos possibilitaram a obtenção de R2 mais elevados, RMSE menores e intervalos de confiança de menor amplitude para um mesmo grau de risco.

A despeito de o algoritmo *XGBRegressor* ter performance melhor em todos os cenários urbanos, as *ANN MLPRegressor* desenvolvidas parecem mais adequadas à valoração dos imóveis administrados pelo EB. Os pontos levantados nas Subseções 5.3 e 5.4 indicam que as redes neurais treinadas têm uma performance relativamente equilibrada e que a análise de sua interpretabilidade à luz do valor de Shapley se mostra mais aderente à realidade; a valorização e a depreciação da variável dependente a partir da variação de cada atributo explicativo ocorrem de forma mais

alinhada ao senso comum. Com a análise do p-valor associado à matriz de vizinhança W e com a autocorrelação espacial dos valores dos imóveis urbanos atestada pelo i de Moran, fica ainda mais evidente a importância da localização na formação das valorações imobiliárias. Em decorrência da importância de fatores locacionais, os modelos destinados a imóveis urbanos tendem a performar de maneira superior nas regiões de maior densidade de bens anotados, ilustradas na Figura 1.

Para os modelos urbanos, as estimativas de valor são mais robustas, com menor grau de incerteza, para locais como Brasília e Rio de Janeiro. As projeções para ativos situados em lugares ermos, no interior da região Norte, por exemplo, distantes de outros imóveis, pela sua própria natureza, envolvem grau maior de incerteza. Imóveis com valor histórico-cultural associado, tais como fortalezas históricas, podem ter componente intangível e valor de bens integrados, como pinturas e monumentos, acrescidos às estimativas deste trabalho.

Os *scripts* relativos aos experimentos encontram-se parcialmente disponíveis no seguinte repositório GitHub: <https://github.com/joseniloneto/imoveisml>. Os modelos treinados e as bases de dados são de acesso restrito.

5. CONCLUSÕES E RECOMENDAÇÕES

Este trabalho é um subproduto de uma dissertação de mestrado apresentada no Programa de Pós-Graduação em Computação Aplicada (PPCA) da Universidade de Brasília (UnB) e foi apresentado em versão estendida no COBREAP 2023. Ele explorou o problema de se estimar, em níveis iniciais, valores de imóveis urbanos da União, em especial daqueles administrados pelo EB, de forma a permitir que as autoridades decisórias tenham acesso a tal informação antes de solicitar formalmente laudos de avaliação imobiliária.

Os modelos implementados com os algoritmos *ANN MLPRegressor* e *XGBRegressor* tiveram coeficiente de determinação consistentemente superior a 57%, em todos os cenários desenhados. Ressalta-se, entretanto, que se trata de estimativas com grau de incerteza associado, mais acentuado na região Norte e

menos acentuado em zonas de maior densidade de dados anotados, tais como Brasília, Rio de Janeiro e São Paulo.

A partir dos experimentos realizados e dos resultados alcançados nesta pesquisa, enxerga-se como promissor realizar novos experimentos incluindo informações do Censo IBGE 2022, quando disponíveis. O valor venal referencial à cobrança de Imposto Predial e Territorial Urbano (IPTU) extraído das plantas de valores genéricos dos municípios talvez constitua um atributo significativo e, consequentemente, ajude a aperfeiçoar os modelos; ressalta-se que, muitas vezes, ele não se encontra disponível de maneira estruturada e que os municípios não o calculam de uma única forma, o que exigiria intensa etapa de tratamento sobre ele.

Complementarmente, fotografias dos imóveis poderiam ser analisadas e ter padrões próprios reconhecidos por meio de redes convolucionais, com potencial de aprimoramento dos modelos preditivos construídos neste trabalho, possivelmente com a agregação de uma variável indicativa do estado de conservação dos imóveis.

Abordagens sob a ótica de demanda mercadológica por imóveis também seria um ponto interessante. Tentou-se incluir a variável Google Trends com granularidade espacial estadual na etapa de seleção de atributos, mas ela não se mostrou significativa. O desemprego local pode exercer influência sobre este aspecto.

REFERÊNCIAS BIBLIOGRÁFICAS

ALVES DE SOUSA NETO, J. N.; LADEIRA, M. Value Estimation of Properties Administered by the Brazilian Army Using Machine Learning and Spatial Components. **Anais do X Symposium on Knowledge Discovery, Mining and Learning (KDMiLe)**, Campinas, p. 17-24, 2022.

ANSELIN, L. Spatial Econometrics: Methods and Models. **Springer Dordrecht**, ISBN 978-94-015-7799-1, 1988.

ANSELIN, L.; REY, S. Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL. **Geoda Press LLC**, ISBN 978-0-986-34210-3, 2014.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **14653-1**: Avaliação de bens — Procedimentos gerais — Apresentação. Rio de Janeiro: ABNT, 2019.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **14653-2**: Avaliação de bens — Imóveis urbanos — Apresentação. Rio de Janeiro: ABNT, 2011.

CAMPOS, R. B. A.; ALMEIDA, E. Decomposição espacial nos preços residenciais no município de São Paulo. *Estudos Econômicos*, v. 48(1), p. 5-38, 2018.

COHEN, J. Statistical Power Analysis for the Behavioral Sciences. **Routledge**, ISBN 978-0-203-77158-7, 1988.

CHAKRABORTY, D.; ELZARKA, H. A novel construction cost prediction model using hybrid natural and light gradient boosting. **Advanced Engineering Informatics**, v. 46, 2020.

DANTAS, R. A.; MAGALHÃES, A. M.; VERGOLINO, J. R. O. Um Modelo Espacial de Demanda Habitacional para a Cidade do Recife. **Estudos Econômicos**, v. 40(4), p. 891-916, 2010.

DEWAN, P.; GANTI, R.; SRIVATSA, M.; STEIN, S. NN-SAR: A neural network approach for spatial autoregression. **2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)**, Kyoto, p. 783-789, 2019.

FURTADO, B. A. T DISET 78 - Gerando Famílias Artificiais Intraurbanas: Censo 2010. **Ipea**, 2020.

HAGENAUER, J.; HELBICH, M. *A geographically weighted artificial neural network*. **International Journal of Geographical Information Science**, v. 36(2), p. 215-235, 2022.

HAIR, J.; BLACK, W.; BABIN, B.; ANDERSON, R. *Multivariate Data Analysis: A Global Perspective*. **Pearson Education**, ISBN 978-0-135-15309-3, 2010.

HAIR, J.; HULT, G. T. M.; RINGLE, C.; SARSTEDT, M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*. **Sage Publishing**, ISBN 978-1-5443-9640-8, 2022.

KIELY, T.; BASTIAN, N. *The spatially conscious machine learning model*. **Statistical Analysis and Data Mining: The ASA Data Science Journal**, v. 13(1), p. 31-49, 2020.

LUNDBERG, S. M.; ERION, G.; CHEN, H.; DEGRAVE, A.; PRUTKIN, J. M.; NAIR, B.; KATZ, R.; HIMMELFARB, N. B.; LEE, S. I. *From local explanations to global understanding with explainable AI for trees*. **Nature Machine Intelligence**, v. 2, p. 56-67, 2020.

PARK, B.; BAE, J. *Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data*. **Expert Systems with Applications**, v. 42(6), p. 2928-2934, 2015.

ROSEN, S. *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*. **Journal Of Political Economy**, v. 82(1), p. 34-55, 1974.

TCHUENTE, D.; NYAWA, S. *Real estate price estimation in French cities using geocoding and machine learning*. **Annals of Operations Research**, v. 308(1-2), p. 571-608, 2022.

**José Nilo Alves de Sousa Neto**

Possui graduação em Engenharia de Fortificação e Construção pelo Instituto Militar de Engenharia (IME, 2015), durante a qual participou de programa de intercâmbio em engenharia com ênfase em mecânica dos fluidos na École des Ponts ParisTech (ENPC, 2014/2015), especialização em planejamento e controle financeiro pela Fundação Getúlio Vargas (FGV, 2016), MBA em Business Analytics e Big Data pela FGV (2020) e Mestrado em Computação Aplicada pela Universidade de Brasília (UnB, 2023) em linha de pesquisa de ciência de dados. Atualmente, é engenheiro na Diretoria de Patrimônio Imobiliário e Meio Ambiente (DPIMA) do Exército Brasileiro, capacitando novos avaliadores de imóveis e desenvolvendo modelos estatísticos

**Marcelo Ladeira**

Possui graduação em Engenharia Mecânica pela Universidade de Brasília (1976), especialização em Curso Avançado de Operação de Sistema Hidrotérmico pela Universidade Federal do Rio de Janeiro (1981), mestrado em Análise de Sistemas e Aplicações pelo Instituto Nacional de Pesquisas Espaciais (1981), doutorado em Computação pela Universidade Federal do Rio Grande do Sul (2000) e pós-doutorado pela Universidade de Lisboa (2004). Atualmente, é professor no Departamento de Ciência da Computação da Universidade de Brasília e Membro de Comitê Consultivo do CNPq. Tem experiência na área de Ciência da Computação, com ênfase em Metodologia e Técnicas da Computação, atuando principalmente nos seguintes temas: mineração de dados ou textos, inteligência artificial, inteligência computacional, raciocínio probabilístico, diagramas de influências, e redes bayesianas.

**Juliana Minura**

Possui graduação em Engenharia Civil pelo Instituto de Ensino Superior Planalto (IESPLAN, 2013), pós-graduação em avaliações e perícias de engenharia pelo Instituto de Pós-graduação e Graduação (IPOG, 2017) e créditos, como aluna especial de mestrado, na Arquitetura da UnB, cursando as disciplinas de Estudos Especiais em Teoria, História e Crítica 1 e Conservação do Patrimônio Moderno. Atualmente, é avaliadora de imóveis da Coordenadoria de Vistoria e Avaliação Imobiliária (CVAI) da Secretaria de Estado de Administração do Mato Grosso do Sul (SAD/MS). Recentemente, integrou equipe responsável pela publicação do Manual para Elaboração de Relatório de Valor Específico Patrimonial da CVAI/SAD.

Acesse este trabalho premiado no XXII COBREAP através do link na biblioteca do IBAPE:

<https://ibape-nacional.com.br/biblioteca/wp-content/uploads/2023/12/estimativa-de-valores-de-imoveis-da-uniao-com-tecnicas-de-aprendizagem-de-maquina-e-componentes-espaciai.pdf>

III. PREDIÇÃO DE OCORRÊNCIAS ARQUEOLÓGICAS EM OBRAS DE INFRAESTRUTURA

**Antoniél Campos
Árlon Facynek de Oliveira Carvalho**

O artigo apresenta um modelo matemático capaz de prever ocorrências de sítios arqueológicos em empreendimentos de infraestrutura

As ocorrências arqueológicas durante a execução de obras de infraestrutura tem sido uma realidade no país. A providência imediata a ser tomada, nos casos de achados fortuitos desses artefatos, é o imediato comunicado ao órgão responsável pelo patrimônio histórico e artístico nacional, no caso, ao IPHAN – Instituto do Patrimônio Histórico e Artístico Nacional. Como consequência, a área no entorno da ocorrência deve ser preservada, até que estudos de prospecção possam ser realizados por profissionais especializados, com posterior pronunciamento e deliberação da autoridade.

Nesse sentido, é objetivo do presente artigo apresentar um modelo matemático de predição de registros arqueológicos, capaz de identificar as áreas mais propensas a essas ocorrências, como medida primeira a ser tomada ainda na fase de definição da localização do empreendimento.

1. REFERENCIAL TEÓRICO

Algumas premissas são adotadas na formulação de modelos matemáticos voltados à predição de sítios arqueológicos. Segundo Tsuchiya (2002, p. 17), essas premissas podem ser elencadas da seguinte forma:

- a) Primeira: a suposição de que a escolha do local dos acampamentos dos povos pré-históricos foi influenciada por elementos ambientais, naturais e físicos;
- b) Segunda: essas variáveis ambientais sobreviveram e podem ser representadas por dados atuais. Estes dados podem estar em mapas, monografias ou podem ser coletados no campo;

- c) Terceira: considera-se que as correlações entre locais arqueológicos e as características do ambiente físico/natural, observado por arqueólogos, são fatos que indicam as decisões dos povoadamentos pretéritos.

Essas três condições formam a base das formulações voltadas às predições de locais com interesse arqueológico, mediante a elaboração de um modelo que congregue as variáveis ambientais, naturais e físicas que determinaram a escolha do local e a permanência, ainda que residual, dessas variáveis no tempo presente, validando as correlações consideradas pelos especialistas.

Para a construção de modelos preditivos, há duas abordagens clássicas: a indutiva e a dedutiva, cujas características principais estão descritas no quadro a seguir.

Quadro 1 – Abordagens usuais em MPA

INDUTIVA	DEDUTIVA
Orientada apenas por dados empíricos tratados de forma estatística para obtenção de padrões	A modelagem é orientada pelas experiências prévias do arqueólogo com áreas preferenciais de localização de sítios arqueológicos.
Parte de dados empíricos para verificação das correlações entre sítios arqueológicos e variáveis ambientais e culturais, através da aplicação de testes estatísticos.	Formulação de hipóteses sobre a localização dos sítios, como por exemplo, proximidade de água ou elevação.
Uso de técnicas inferenciais, como a regressão logística	Para cada variável considerada relevante, atribui-se diferentes pesos através de uma equação matemática.
Possibilidade de extrapolar os resultados para uma área maior	Não há a criação de novos conhecimentos e sim, a utilização de um conhecimento prévio.

Fonte: adaptado de Merencio (2020) e Deblasis e Kozlowski (2019)

As variáveis independentes são utilizadas nos modelos preditivos, cuja variável dependente é uma resposta entre 0 e 1, indicativa da probabilidade de se encontrar sítios arqueológicos no local em estudo. Esse modelo pode ser obtido por meio da abordagem dedutiva, que é aquela onde são empregados “pesos” às variáveis, ou pela abordagem indutiva, que utiliza técnicas estatísticas, como a regressão logística.

Para Gonzales (2018, p. 15), A regressão logística é uma técnica estatística que tem como objetivo produzir, a partir de um conjunto de observações, um modelo que permita a predição de valores tomados por uma variável categórica, frequentemente binária, em função de uma ou mais variáveis independentes contínuas e/ou binárias.

Autores como Deblasis e Kozlowski, consideram que a variável resposta pode advir de uma incerteza:

O conjunto de dados geográficos necessários para a elaboração dos modelos é composto por dois elementos: os sítios conhecidos e um mesmo número de locais aleatórios onde não há presença de sítios arqueológicos, pois a regressão logística opera com valores binários de certeza ou incerteza da presença de sítios. [DEBLASIS e KOZLOWSKI, 2019, p. 81]

Mas para Tsuchiya (2002, p. 60), a presença ou ausência de ocorrências arqueológicas não é uma incerteza, mas sim uma condição necessária para que as variáveis independentes possam explicar a ausência ou a presença dessas ocorrências: “na regressão logística utiliza-se variáveis independentes, que explicam a variável dependente, a presença/ausência de ocorrências arqueológicas.”

Em abordagem puramente teórica quanto à técnica estatística em comento — a regressão logística —, Gujarati (2019, p.195) apresenta uma alternativa ao uso dessa técnica em situações em que a variável resposta é dicotômica, ou seja, assume apenas dois resultados possíveis: “sim” ou “não”, como é o caso do presente estudo. A estratégia consiste em agrupar os dados conforme a frequência das variáveis utilizadas. Dessa forma, tem-se que a variável resposta tem por fim estimar, conforme a combinação dessas variáveis, o total de ocorrências 3 arqueológicas. Sabendo-se, a priori, o total de ocorrências, a frequência relativa de cada grupo responderia pela probabilidade de ocorrências para aquela determinada combinação de variáveis. Estabelecida assim essa nova amostra, pode-se usar então os Mínimos Quadrados Ordinários – MQO, para obtenção do modelo de regressão.

2. METODOLOGIA

Das duas abordagens analisadas, a indutiva e a dedutiva, é incontestável que a segunda traz maior subjetividade na aplicação. Ocorre que, para a abordagem indutiva, e por extensão, o uso da regressão logística, há que se ter uma amostra representativa composta por respostas “sim” e “não”, o que significa, na prática, que não basta apenas a caracterização do local em que se encontra o sítio arqueológico, mas igualmente se faz necessário que tenha havido a prospecção com respostas “não”, ou seja, sondagens fracassadas de pesquisa por artefatos para, daí então, caracterizar o ambiente com as variáveis apropriadas.

Tal é a dificuldade de se utilizar o que seria a técnica de excelência para o fim de detectar sítios arqueológicos, que é a regressão logística.

No outro pórtico, tem-se, como já citado, a alternativa de utilizar a abordagem dedutiva, com a limitação imposta pela subjetividade do arqueólogo, por mais experimentado que seja.

Assim, entre esses extremos, optou-se por adotar uma estratégia que se configura como intermediária entre as duas abordagens clássicas, que é a trazida por Gujarati (2019, p. 195), ao aproveitar o manancial de informações contidas na amostra composta por respostas “sim”, ou seja, o próprio universo das ocorrências de sítios arqueológicos da região de interesse, tirando da amostra as probabilidades de ocorrências arqueológicas, aqui denominado de risco arqueológico (Ra), permitindo estabelecer uma métrica para o potencial arqueológico, do tipo “Alto”, “Médio” ou “Baixo”, conforme o risco arqueológico calculado para o local de interesse.

2.1 Região de interesse

Foram estudadas as ocorrências de sítios arqueológicos no Estado do Rio Grande do Norte, conforme mapeamento constante na figura a seguir.

Figura 1 – Mapeamento de sítios arqueológicos do RN



Fonte: IPHAN, 2023.

Foram selecionados 85 sítios (Δ) que continham a localização geográfica, sendo que um deles, o de nº 36, era uma ocorrência marítima, tendo sido descartado. A amostra final é composta, portanto, por 84 dados.

2.2 Variáveis utilizadas

Foram provisoriamente selecionadas as seguintes variáveis independentes:

- **Distância da água:** distância euclidiana entre o sítio arqueológico e a formação aquática mais próxima;
- **Classes de solo:** Chernossolo, Neossolos, Planossolos, Latossolos, Cambissolos, Argissolos e Luvisolos
- **Declividade:** inclinação do terreno no sentido sítio/fonte aquática.

A identificação dos sítios arqueológicos e as coordenadas geográficas foram fornecidas pelo IPHAN (2023); a distância da água ao sítio (m), a elevação do sítio (m) e a elevação da massa aquática (m) mais próxima ao sítio foram obtidos diretamente no Google Earth. A cota da água é a diferença entre a elevação do sítio e a elevação da água. Dividindo-se a cota pela distância da água ao sítio, tem-se a declividade predominante. O tipo de solo vem de Santos (2011).

2.3 Agrupamento de variáveis

Quadro 2 – Segregação de variáveis

Dist.Água (A1)		Declividade (D1)		Classe de solo			
< 1.500 m		< 0,090		Neossolos (S1)		Planossolos (S2)	
Sim	Não	Sim	Não	Sim	Não	Sim	Não
1	0	1	0	1	0	1	0

Fonte: Elaborado pelo autor

Combinando-se as possibilidades do Quadro 6 com a amostra de sítios arqueológicos, obtém-se a seguinte amostra agrupada:

Tabela 1 – Amostra agrupada

#	A1	D1	S1	S2	Quant.	Frequência (Ra)
1	1	1	1	0	14	0,1667
2	1	1	0	1	28	0,3333
3	1	1	0	0	18	0,2143
4	1	0	1	0	5	0,0595
5	1	0	0	1	0	0,0000
6	1	0	0	0	0	0,0000
7	0	1	1	0	12	0,1429
8	0	1	0	1	3	0,0357
9	0	1	0	0	4	0,0476
10	0	0	1	0	0	0,0000
11	0	0	0	1	0	0,0000
12	0	0	0	0	0	0,0000
Total:					84	

Fonte: Elaborado pelo autor

A frequência representa o risco arqueológico (Ra) de se encontrar um sítio, de acordo com a combinação de variáveis presentes no local; ou seja, é a variável resposta buscada.

2.4 Modelo de predição de sítios arqueológicos

Foi utilizada a técnica de Regressão Linear, uma vez que se buscava estudar o comportamento de uma variável dependente — no caso, o risco de ocorrência arqueológica, Ra — em relação a outras variáveis, chamadas independentes, que são responsáveis pela variabilidade observada na probabilidade da ocorrência.

Para a elaboração dos cálculos, utilizou-se o software SisDea Home, v. 1.60.

O modelo de regressão de melhor ajuste, ou seja, o que melhor representa a probabilidade de ocorrência de sítio arqueológico, é dado pela equação:

$$Ra = -0,01871785714 + 0,05640952381 \times A1 + 0,1119761905 \times D1 + 0,0268 \times S1 - 0,02551071429 \times S2$$

Onde:

Ra: Probabilidade de encontrar um sítio arqueológico no local de interesse;

A1: Distância da massa aquática ao sítio arqueológico, em metros, assumindo o valor “1”, caso a distância seja < 1.500 m, ou valor “0”, caso a distância seja ≥ 1.500 m;

D1: Declividade predominante no local de interesse, adimensional, assumindo o valor “1”, para uma declividade $< 0,090$, ou valor “0”, caso a declividade seja $\geq 0,090$;

S1: Classe de solo. Assume o valor “1”, caso o solo seja da classe Neossolos, ou valor “0”, caso não seja;

S2: Classe de solo. Assume o valor “1”, caso o solo seja da classe Planossolos, ou valor “0”, caso não seja.

2.5 Potencial arqueológico retornado pelo modelo preditivo

Aplicou-se o modelo à amostra agrupada (Tabela 3), segregando-se os valores obtidos em três grupos: “Alto”, “Médio” e “Baixo”, conforme o potencial de ocorrência de sítio arqueológico. O ponto de mudança entre “Alto” e “Médio”, obtido por meio da clusterização por k-means, foi de 0,124; portanto, para um risco arqueológico (Ra) maior que 0,124 (12,4%), considera-se alto o potencial de ocorrência. Já o ponto de

mudança ente “Médio” e “Baixo”, também obtido por meio da clusterização, foi de 0,068; portanto, para um risco arqueológico (Ra) menor que 0,068 (6,8%), considera-se baixo o potencial de ocorrência. Para valores intermediários a esses pontos de mudança, considera-se que o risco é médio para a ocorrência de sítio arqueológico. Esses horizontes constam na tabela a seguir.

Tabela 2 – Potencial arqueológico

Potencial arqueológico	
Risco (Ra)	Potencial
$< 0,068$	Baixo
$0,068 \leq Ra \leq 0,124$	Médio
$> 0,124$	Alto

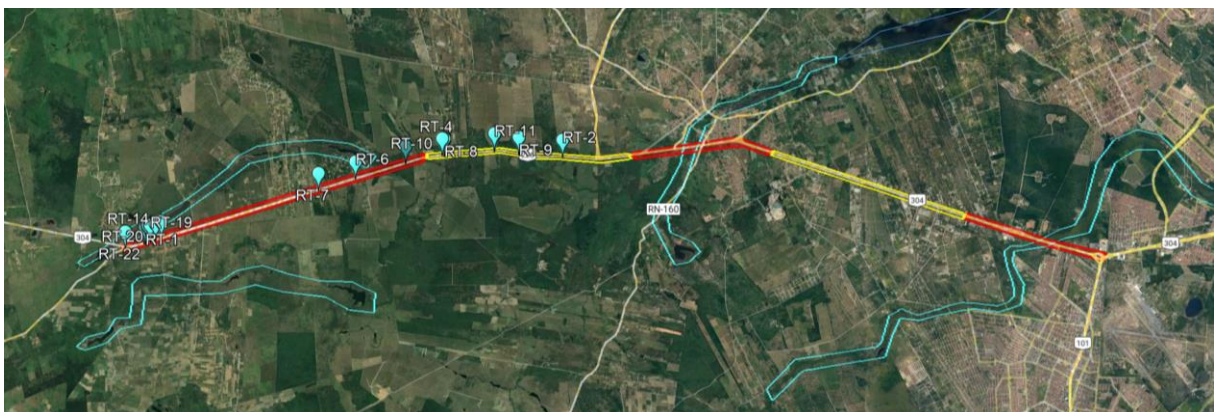
Fonte: Elaborado pelo autor

3 ESTUDO DE CASO

Obra: Duplicação de uma rodovia, com 27 km.

Foi aplicado o modelo (Eq. 1) em toda a extensão da obra rodoviária de duplicação, tendo sido obtida a seguinte configuração de potencial arqueológico.

Figura 2 – Potencial arqueológico de uma duplicação de rodovia (2022)



Fonte: Elaborado pelo autor, com imagens de satélite do Google Earth

Os segmentos considerados como de “Alto” potencial arqueológico estão destacados em vermelho, ao passo que os segmentos de “Médio” potencial estão indicados em amarelo. Não foi detectado segmento de “Baixo” potencial arqueológico.

Uma vez que a obra de duplicação da rodovia encontra-se ainda em execução, especial atenção deve ser dada aos segmentos de “Alto” potencial arqueológico ainda não iniciados.

Essa antecipação da ação de monitoramento arqueológico só se faz possível com o amparo de um modelo preditivo como o ora utilizado.

REFERÊNCIAS BIBLIOGRÁFICAS

DEBLASIS, Paulo e KOZLOWSKI, Henrique de Sena. **Perspectivas de modelagem preditiva arqueológica: estudo de caso na Bacia do Rio Tubarão – Santa Catarina.** In: Especiaria - Caderno de Ciências Humanas. p. 72-94, mar. 2019.

GONZALEZ, Leandro de Azevedo. **Regressão Logística e suas Aplicações.** 2018. 106 f. Monografia (Graduação) - Curso de Ciência da Computação, Centro de Ciências Exatas e Tecnológicas, Universidade Federal do Maranhão, São Luís, 2018.

GUJARATI, Damodar N. **ECONOMETRIA: princípios, teoria e aplicações práticas.** 1 ed. São Paulo: Saraiva Educação, 2019. 522 p.

HUAYANAY, Alex de La Cruz. **Modelos De Regressão Para Resposta Binária Na Presença De Dados Desbalanceados.** 2019. 91 f. Dissertação (Mestrado) - Curso de Estatística, Centro de Ciências Exatas e Tecnologia, UFSCar-Usp, São Carlos, 2019.

IPHAN. Instituto do Patrimônio Histórico e Artístico Nacional. <
<https://www.gov.br/iphan/pt-br/patrimonio-cultural/patrimonio-arqueologico/cadastro-de-sitios-arqueologicos> > (acesso em 12/04/2023).

KIPNIS, Renato. **O uso de modelos preditivos para diagnosticar recursos arqueológicos em áreas a serem afetadas por empreendimentos de impacto ambiental.** IN: CALDARELLI, Solange Bezerra. Atas do Simpósio sobre Política Nacional do Meio Ambiente e Patrimônio Cultural. Goiânia: Instituto Goiano de Pré-História e Antropologia, 1997, p. 34–51.

MERENCIO, Fabiana Terhaag. **A modelagem preditiva na Arqueologia: um olhar crítico a partir do estudo de caso no território Xetá.** Cadernos do Lepaarq, v. XVII, n.33, p. 25-45, Jan-Jun. 2020.

PAULA, Eduardo Henrique Andrade de. **INSTITUTIONAL VOIDS: uma análise taxonômica e fenomenológica de sua influência na paralisação de obras públicas no brasil.** 2022. 727 f. Tese (Doutorado) - Curso de Administração, Ciências Empresariais, Universidade Fumec, Belo Horizonte, 2022.

PEREZ, Glauco Constantino; AFONSO, Marisa Coutinho e MOTA, Lúcio Tadeu. **Métodos de análise espacial para sítios arqueológicos**: um modelo preditivo para o estado de São Paulo. In: Cadernos do Lepaarq, v. XV, n.30., p. 98-120, Jul-Dez. 2018.

SANTOS, Humberto Gonçalves dos. **O novo mapa de solos do Brasil**: legenda atualizada. Rio de Janeiro: Embrapa Solos, 2011. 67 p. Modo de acesso: < <http://www.cnps.embrapa.br/solosbr/publicacao.html> >. (acesso em 12/04/2023).

TSUCHIYA, Ítalo. **Regressão Logística Aplicada Na Análise Espacial De Dados Arqueológicos**. 2002. 103 f. Dissertação (Mestrado) - Curso de Ciências Cartográficas, Cartografia, Universidade Estadual Paulista, Presidente Prudente, 2002.

**Antoniél Campos**

Engenheiro Civil, especialista em avaliações e perícias de engenharia. Servidor público federal.

Vencedor das medalhas Eurico Ribeiro, Hélio de Caires, Ibá Ilha Moreira (COBREAP) e medalha Domingos de Saboya (SOBREA), todas das áreas de avaliações, perícias e meio ambiente.

Vencedor do prêmio Min. Guilherme Palmeira pelo TCU sobre soluções consensuais de conflitos.

Premiações em congressos e simpósios:

- Prêmio Ministro Guilherme Palmeira - melhor monografia na categoria sociedade civil sobre soluções consensuais de controvérsias e prevenção de conflitos no âmbito da administração pública – TCU (2023).
- Medalha Eurico Ribeiro - melhor trabalho de perícias apresentado no XXII COBREAP (2023);
- Medalha Ibá Ilha Moreira Filho – melhor trabalho na área ambiental apresentado no XXII COBREAP (2023);
- Medalha Domingos Saboya - melhor trabalho apresentado no X Simpósio da SOBREA (2022);
- Medalha Hélio de Caires - melhor trabalho de avaliações apresentado no XX COBREAP (2019);
- Menção honrosa - trabalho apresentado na área de perícias no XIX COBREAP (2017).

**Árlon Facynek de Oliveira Carvalho**

Bacharel e Mestre em Arqueologia pela Universidade Federal do Piauí, com experiência em pesquisa e gestão do patrimônio arqueológico.

Cargo atual: Arqueólogo na empresa Zago Engenharia.

Acesse este trabalho premiado no XXII COBREAP através do link na biblioteca do IBAPE:

<https://ibape-nacional.com.br/biblioteca/wp-content/uploads/2023/12/predicao-de-ocorrencias-arqueologicas-em-obras-de-infraestrutura.pdf>