

# **AVALIANDO A QUALIDADE DOS CURSOS DE GRADUAÇÃO BRASILEIROS: UM INDICADOR COMPOSTO BASEADO EM FOLGAS E NO “BENEFÍCIO DA DÚVIDA”**

Lívia Mariana Lopes de Souza Torres<sup>1</sup>

Felipe Alves Mendes da Silva<sup>2</sup>

Francisco S. Ramos<sup>3</sup>

## **Resumo:**

O presente trabalho tem como objetivo avaliar a qualidade dos cursos de graduação em instituições de ensino superior brasileiras. A metodologia atual utilizada pelo Ministério da Educação para obtenção do Conceito Preliminar de Curso (CPC) considera oito indicadores e utiliza pesos pré-estabelecidos para obter um valor de CPC contínuo. A definição arbitrária desses pesos é um problema significativo no método atual. Para mitigar essa limitação e proporcionar uma avaliação mais justa para os cursos, propomos um indicador composto a partir de um modelo "benefício da dúvida" baseado em folgas. Os cursos de engenharia de produção avaliados em 2019 foram utilizados como objeto de estudo para ilustrar o procedimento. Os resultados mostram a aplicabilidade do método, bem como descobertas específicas como uma maior quantidade de cursos eficientes nas regiões Sul e Sudeste e a verificação de discrepâncias entre as regiões.

Palavras chaves: Conceito Preliminar de Curso (CPC); Análise Envoltória de Dados (DEA); Benefício da Dúvida (BoD); Indicador Composto.

## **Abstract**

This work aims to evaluate the quality of undergraduate courses in Brazilian higher education institutions. The current methodology used by the Ministry of Education to obtain the Preliminary Course Concept (CPC) considers eight indicators and uses pre-established weights to obtain a continuous CPC value. The arbitrary definition of these weights is a significant problem in the current method. To mitigate this limitation and provide a fairer assessment for courses, we propose a composite indicator based on a "benefit of the doubt" model based on slack. The production engineering courses evaluated in 2019 were used as an object of study to illustrate the procedure. The results show the applicability of the method, as well as specific discoveries such as a greater number of efficient courses in the South and Southeast regions and the verification of discrepancies between regions.

Keywords: Preliminary Course Concept (CPC); Data Envelopment Analysis (DEA); Benefit of the Doubt (BoD); Composite Indicator.

---

<sup>1</sup> PPGEP-Programa de Pós-Graduação em Engenharia de Produção, UFPE

<sup>2</sup> PPGEP-Programa de Pós-Graduação em Engenharia de Produção, UFPE

<sup>3</sup> PPGEP-Programa de Pós-Graduação em Engenharia de Produção e PIMES-Programa de Pós-graduação em Economia, e LabGRC-Laboratório de Gerenciamento de Riscos, Governança e Compliance, UFPE). Os autores agradecem o apoio da CAPES (Código de Financiamento 001) e o terceiro autor agradece ao CNPq-PQ 311334/2019-9.

## 1. INTRODUÇÃO

A qualidade dos sistemas educacionais é crucial para qualquer país, e a sua avaliação é essencial para a concepção e implementação de políticas educativas. Contudo, a avaliação é um processo complexo devido aos diferentes indicadores utilizados para obter uma visão geral do desempenho de um sistema educacional (STUMBRIENE; CAMANHO; JAKAITIENE, 2020). Nesse sentido, o indicador composto (IC) surge como uma das técnicas mais populares para incorporar vários indicadores, com vastas aplicações na literatura em áreas como saúde, ambiente, economia e desenvolvimento tecnológico (ALBO; LANIR; RAFAELI, 2019). IC é uma agregação matemática de um conjunto de indicadores individuais que medem conceitos multidimensionais e geralmente não têm uma unidade de medida em comum (SHI; LAND, 2021).

O Conceito Preliminar do Curso (CPC) é um indicador de qualidade que visa avaliar os cursos de graduação brasileiros. É um parâmetro importante para que as universidades brasileiras possam obter uma visão geral do que está correto e do que precisa de ser melhorado (IKUTA; BARREYRO, 2021).

Além disso, o CPC fornece informações para a elaboração de políticas educacionais pelo governo, serve como subsídio para a supervisão e regulação do ensino superior e é utilizado pelo governo federal como critério na construção da matriz de distribuição orçamentaria (INEP, 2020a). No entanto, apesar da relevância desse indicador e de estar em curso há anos, a sua formulação continua sendo o foco de debate e crítica por parte da comunidade acadêmica, devido à falta de consenso que reside principalmente na atribuição dos pesos nos diferentes componentes que compõem o CPC (ZANELLA; OLIVEIRA, 2021). Nesse contexto, estas discussões são alimentadas pelo fato de que pequenas alterações nos pesos dos componentes podem ter um impacto significativo na mensuração do desempenho de um curso (IKUTA; BARREYRO, 2021).

Em geral, os métodos tradicionais para construir um IC consistem em atribuir pesos a cada indicador, e estas atribuições estão diretamente relacionadas com a qualidade e confiabilidade do índice calculado (BABAEE et al., 2021; STUMBRIENE; CAMANHO; JAKAITIENE, 2020). Assim, a *Data Envelopment Analysis* (DEA) tem sido usado para atribuir o melhor conjunto de pesos aos indicadores e posteriormente agregá-los para maximizar a pontuação do índice composto sem conhecimento a priori dos pesos desses indicadores (BABAEE et al., 2021).

Neste contexto, o *Benefit-of-the-Doubt* (BoD) consiste numa abordagem baseada em DEA para a construção de IC, resultando em pesos endógenos para indicadores agregados que variam tanto entre tais índices como entre as unidades de decisão (DMUs) avaliadas. O modelo BoD está enraizado no DEA, e é formalmente equivalente ao modelo original CCR-DEA orientado ao input,

com todos os subindicadores considerados como saídas e uma "entrada fictícia", igual a um, é considerada para todas as DMUs (CHERCHYE et al., 2007).

Neste sentido, as DMUs avaliadas podem escolher um conjunto de pesos que maximize o seu desempenho em termos do valor resultante do IC sob a restrição de que se qualquer outra DMU avaliada utilizar o mesmo conjunto de pesos, não resultará em um valor do indicador composto superior a um (KARAGIANNIS; KARAGIANNIS, 2018).

Portanto, o presente estudo avalia a qualidade dos cursos de graduação em instituições de ensino superior brasileiras através da construção de um indicador composto utilizando um modelo de "benefício da dúvida" baseado em folga. A metodologia proposta pode ser aplicada a todos os cursos de graduação brasileiros. Para ilustrar o procedimento proposto, foi selecionada uma amostra com cursos de engenharia de produção, com dados referentes a 2019, uma vez que correspondem à última avaliação destes cursos.

A primeira seção introduz brevemente o contexto de avaliação da qualidade dos cursos de graduação brasileiros. A segunda detalha aspectos do modelo atual utilizado pelo Ministério da Educação. A terceira apresenta aspectos conceituais da Análise Envoltória de Dados e o modelo utilizado para obter os indicadores compostos. A quarta elucida o método proposto, a quinta discute os resultados, e na última tem-se as considerações finais.

## **2. AVALIAÇÃO DA QUALIDADE DOS CURSOS DE GRADUAÇÃO NO BRASIL**

Os cursos de graduação no Brasil são avaliados pelo Instituto Nacional de Estudos e Pesquisa Educacional Anísio Teixeira (INEP). O INEP é uma autarquia federal, vinculada ao Ministério da Educação, e referência nacional em três grandes áreas de atuação: a) avaliações e exames educacionais; b) pesquisas estatísticas e indicadores educacionais; e c) gestão do conhecimento e estudos educacionais. Entre suas responsabilidades no Ensino Superior destaca-se o Exame Nacional do Ensino Médio (Enem), a avaliação de alunos de graduação, cursos de graduação e universidades. O Enem corresponde à prova de avaliação mais relevante para ingressar nas universidades brasileiras. A sua nota pode ser utilizada em instituições públicas, privadas e estrangeiras (as universidades portuguesas utilizam a nota como critério de admissão).

O INEP fornece anualmente um conjunto de indicadores de qualidade relativos ao ensino superior. O foco aqui é o Conceito Preliminar de Curso (CPC), um indicador de qualidade que combina, numa única medida, diferentes aspectos relacionados aos cursos de graduação. É composto por oito itens, agrupados em quatro dimensões destinadas a avaliar a qualidade dos cursos de

graduação (INEP, 2020a). A Tabela 1 detalha as dimensões e os indicadores utilizados no cálculo do CPC.

Tabela 1: Descrição das dimensões e indicadores com respectivos pesos

Dimensão	Indicadores	Peso
Desempenho dos estudantes	Nota dos concluintes no Enade (CE)	0,2
Valor agregado pelo processo formativo oferecido pelo curso	Nota do Indicador da Diferença entre os Desempenhos Observado e Esperado (IDD)	0,35
Corpo Docente	Nota de Proporção de Mestres (M)	0,075
	Nota de Proporção de Doutores (D)	0,15
	Nota de Regime de Trabalho (WRP)	0,075
Percepção Discente sobre as Condições do Processo Formativo	Nota referente à organização didático-pedagógica (DPO)	0,075
	Nota referente à infraestrutura e instalações físicas (IPI)	0,05
	Nota referente às oportunidades de ampliação da formação acadêmica e profissional (OEVTV)	0,025

Todos os indicadores relacionados à percepção dos estudantes sobre as condições do curso de graduação são obtidos por meio de questionários, preenchidos por todos os estudantes de graduação antes do exame Enade. O valor de cada indicador consiste numa média dos valores obtidos com os estudantes. Os indicadores relativos ao corpo docente representam a proporção daqueles que possuem um título de mestrado e um título de doutorado. Já a proporção do regime de trabalho indica a percentagem de professores de cursos de graduação com trabalho parcial ou em tempo integral.

Dois indicadores referem-se às duas primeiras dimensões e refletem o desempenho dos estudantes. O NC indica a nota obtida no Enade (teste que avalia os principais tópicos que um estudante deve conhecer quando está próximo de concluir o seu curso de graduação). Por outro lado, o NIDD mede o valor acrescentado pelo curso de graduação para o desenvolvimento dos estudantes graduados, considerando o seu desempenho no Enade (INEP, 2020). Na Tabela 1 apresenta-se os pesos utilizados pelo INEP para obtenção do valor do CPC. É imprescindível mencionar que no parecer técnico nº 58/2020/CGCQES/DAES, que detalha o cálculo do CPC, não há justificativa ou explicação para a escolha de tais valores.

O CPC consiste em uma soma ponderada de oito indicadores: portanto, é um indicador composto. Devido à sua relevância para a avaliação dos cursos de graduação e das universidades em que o curso está inserido, propõe-se aqui uma alternativa para o seu cálculo. Um método baseado no “benefício da dúvida” do DEA é a formulação matemática sugerida para obtenção de um novo CPC, não requerendo pesos pré-definidos, e utilizando apenas os valores dos indicadores para realizar a avaliação.

### 3. ANÁLISE ENVOLTÓRIA DE DADOS (DEA) PARA A CONSTRUÇÃO DE INDICADORES COMPOSTOS

A Análise envoltória de Dados (DEA) é uma das técnicas mais eficazes para medir a eficiência de um conjunto de unidades tomadoras de decisão (DMUs) com múltiplas entradas e saídas (KAO, 2016). Desenvolvida por Charnes, Cooper and Rhodes (1978), essa técnica tornou-se uma das principais técnicas não paramétricas, consolidando-se nos últimos quarenta anos. Há casos específicos orientados a inputs, onde apenas os outputs são avaliados, considerando como input uma variável dummy igual a um para todas as DMUs: é denominada *Benefit-of-the-Doubt* (BoD), fundamental no contexto da construção do IC, com todos os subindicadores considerados como saídas (VAN PUYENBROECK, 2018).

Cherchye et al. (2007) foram os pioneiros a propor essa metodologia para a construção de ICs. Os autores partiram do pressuposto da credibilidade vulnerável dos ICs, afetada por possíveis divergências entre especialistas ou stakeholders, sobre o esquema de ponderação utilizado para agregar subindicadores. A metodologia é proposta para permitir um esquema de ponderação flexível para comparar DMUs. Este estudo seminal foi expandido em várias direções, modelos com restrição de peso (VAN PUYENBROECK; ROGGE, 2017), modelos robustos (DARDHA; ROGGE, 2020), modelos direcionais (FÄRE et al., 2019), modelos condicionais (FUSCO; VIDOLI; SAHOO, 2018) e modelos baseados em folgas (SBM) (MARIANO; FERRAZ; DE OLIVEIRA GOBBO, 2021).

Dada a relevância envolvida na construção de ICs, a Organização para Cooperação e Desenvolvimento Econômico (OCDE) desenvolveu um manual para a construção de ICs com aplicações em diversas áreas, como política, comunicação, economia, educação e meio ambiente (OECD, 2008). Como tal, existem mais de 150 indicadores compostos, incluindo o Índice de Desenvolvimento Humano (IDH), Qualidade de Vida, Competitividade, Desempenho Tecnológico, Desempenho do Sistema de Saúde, Desempenho Ambiental e vários índices subjetivos de bem-estar (KARAGIANNIS; KARAGIANNIS, 2020). A educação tem recebido atenção especial nesse contexto.

Rogge (2011) propôs a aplicação de uma abordagem do benefício da dúvida na avaliação de desempenho de professores. Maričić et al. (2016) desenvolveram um modelo do BoD direcional para discutir o impacto do esquema de peso no índice universitário Quacquarelli Symonds (QS). De Witte and Schiltz (2018)] aplicaram uma abordagem BoD robusta e condicional para investigar a eficácia dos distritos escolares. Szuwarzyński (2019) abordou o desempenho de pesquisa de universidades australianas com o auxílio de um modelo BoD restrito.

Com o objetivo de contribuir com a literatura sobre avaliações educacionais com modelos BoD, o presente estudo foca na qualidade dos cursos

de graduação em universidades públicas brasileiras. O Brasil é um país intercontinental, com graves problemas socioeconômicos, como desigualdade social, falta de capital humano qualificado e pobreza crescente. Assim, a superação de tais problemas passa pela consolidação de um sistema de ensino superior qualificado e comprometido com os interesses da sociedade (BERTOLIN, 2011). Portanto, avaliar e monitorar o desempenho do sistema de ensino superior nacional é uma análise do serviço público essencial para assegurar uma maior paridade entre as regiões do país.

#### 4. ESTRATÉGIA EMPÍRICA

##### 4.1 Dados

Os dados, provenientes do site do INEP, correspondem a 2019, último ano em que o Ministério da Educação avaliou os cursos de Engenharia de Produção. Também é relevante mencionar que todos os indicadores são normalizados em um intervalo de 0 a 5. Este auxilia a evitar outliers e a encontrar uma contribuição relativa mais homogênea entre as variáveis (MARIANO; FERRAZ; DE OLIVEIRA GOBBO, 2021). Em 2019 foram avaliados 654 cursos de Engenharia de Produção, 113 deles em instituições públicas. A escolha de IES públicas se justifica pelo financiamento com recursos públicos e o fato de os alunos não pagarem mensalidades. Seis cursos foram retirados por falta de informação, com a amostra final compreendendo 107 DMUs - 79 instituições federais, 27 instituições estaduais e 7 instituições municipais. A Tabela 2 apresenta as estatísticas descritivas dos indicadores. Conforme mencionado anteriormente, todos os indicadores selecionados são os usados pelo governo na análise do CPC para obter uma comparação direta. Também é relevante verificar a distribuição pelas macrorregiões do país. O Centro-Oeste (CO) brasileiro contempla oito cursos de graduação e o Norte (N) seis. O Sudeste (SE) tem a maior concentração, 48, e o Nordeste (NE) e Sul (SUL) têm 22 e 23, respectivamente.

Tabela 2: Estatística descritiva dos indicadores

Indicador	Média	SD	Median a	Máxim o	Mínim o
Notas dos concluintes no ENADE (CE)	3,289	0,979	3,26	5	0,755
Nota do Indicador da Diferença entre Performances Observadas e Esperadas (IDD)	2,566	0,614	2,485	5	0,776
Nota de proporção de doutores (D)	3,371	1,152	3,548	5	0,394
Nota de proporção de mestres (M)	4,422	0,705	4,642	5	1,818
Nota referente à infraestrutura e instalações físicas (IPI)	2,3	0,861	2,322	4,782	0,249
Nota referente às oportunidades de ampliação da formação acadêmica e profissional (OEVT)	2,901	0,66	2,973	4,319	0,877
Nota referente à organização didático-pedagógica (DPO)	2,203	0,761	2,209	4,518	0,596
Nota de regime de trabalho (WRP)	4,871	0,569	5	5	1,225

## 4.2 Modelo SBM-BoD

Os modelos BoD fornecem uma forma de agregar múltiplos indicadores sem a pré-definição de pesos. Neste estudo, o modelo DEA BoD é baseado na hipótese SBM, e detalhado em (1). Consideremos  $n$  DMUs sob avaliação e os respectivos indicadores -  $CE$ ,  $IDD$ ,  $D$ ,  $M$ ,  $IPI$ ,  $OEVT$ ,  $DPO$ ,  $WRP$ . O modelo de programação linear apresentado em (1) corresponde a forma de envelope do SBM-BoD.  $CE_O$ ,  $IDD_O$ ,  $D_O$ ,  $M_O$ ,  $IPI_O$ ,  $OEVT_O$ ,  $DPO_O$  e  $WRP_O$  correspondem aos indicadores da DMU em observação, enquanto  $s_{CE}$ ,  $s_{IDD}$ ,  $s_D$ ,  $s_M$ ,  $s_{IPI}$ ,  $s_{OEVT}$ ,  $s_{DPO}$  and  $s_{WRP}$  correspondem as folgas das respectivas folgas, enquanto  $\lambda_k$  reflete o nível de importância do benchmark  $k$  para o curso em análise.

$$\max \frac{1}{CPC_{BOD}} = 1 + \frac{1}{8} \left( \frac{s_{CE}}{CE_O} + \frac{s_{IDD}}{IDD_O} + \frac{s_D}{D_O} + \frac{s_M}{M_O} + \frac{s_{IPI}}{IPI_O} + \frac{s_{OEVT}}{OEVT_O} + \frac{s_{DPO}}{DPO_O} + \frac{s_{WRP}}{WRP_O} \right)$$

Sujeito á

$$\sum_{k=1}^n CE_k * \lambda_k - s_{CE} = CE_O$$

$$\sum_{k=1}^n IDD_k * \lambda_k - s_{IDD} = IDD_O$$

$$\sum_{k=1}^n D_k * \lambda_k - s_D = D_O$$

$$\sum_{k=1}^n M_k * \lambda_k - s_M = M_O$$

$$\sum_{k=1}^n IPI_k * \lambda_k - s_{IPI} = IPI_O$$

$$\sum_{k=1}^n OEVT_k * \lambda_k - s_{OEVT} = OEVT_O$$

$$\sum_{k=1}^n DPO_k * \lambda_k - s_{DPO} = DPO_O$$

$$\sum_{k=1}^n WRP_k * \lambda_k - s_{WRP} = WRP_O$$

$$\sum_{k=1}^n \lambda_k = 1$$

$$\lambda_k, s_{WRP}, s_{DPO}, s_{OEVT}, s_{IPI}, s_M, s_D, s_{IDD}, s_{CE} \geq 0$$

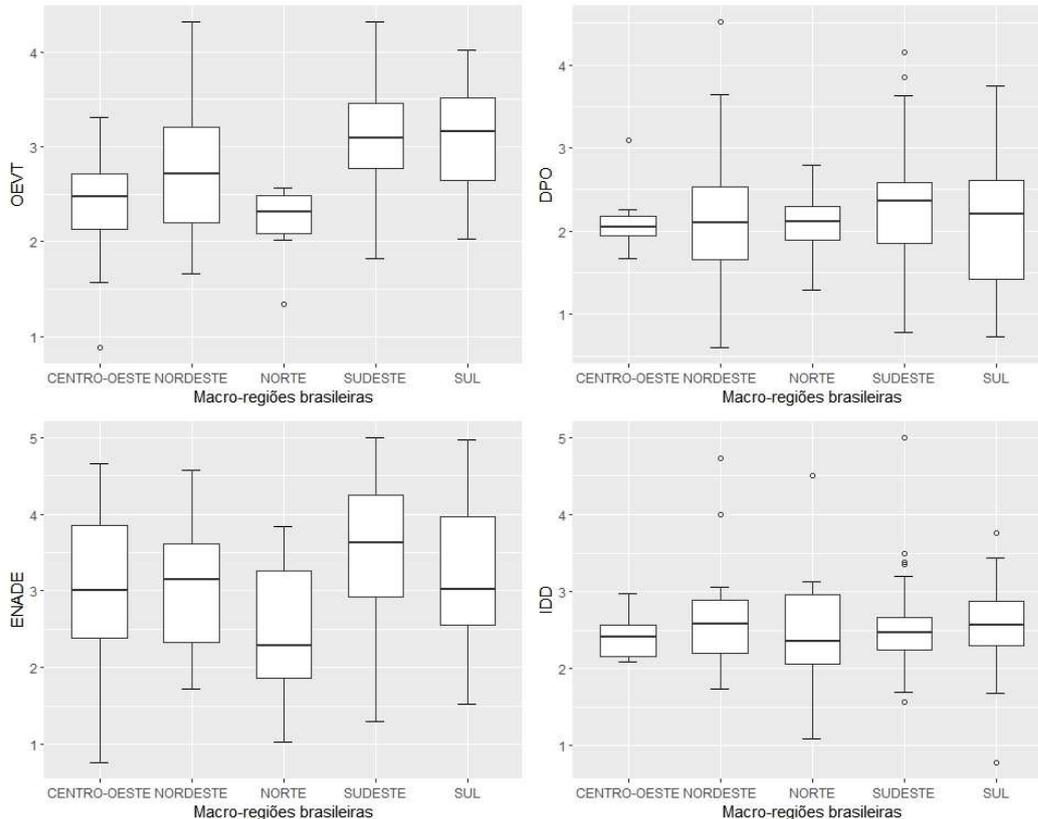
(1)

Ao solucionar o modelo (1), o inverso da função objetivo corresponde ao indicador composto e esse valor permite ranquear as DMUs. Esse modelo deve ser estimado  $n$  vezes. A escolha de um modelo SBM visa obter uma avaliação não radial, ou seja, não é necessário considerar um aumento proporcional nas saídas para melhorias no escore de eficiência. Dessa forma, a modelagem permite uma avaliação mais realista da situação.

## 5. RESULTADOS

Primeiramente, são apresentados os indicadores e feita uma breve discussão sobre sua dispersão pelas macrorregiões brasileiras. Em seguida, são detalhados os resultados obtidos pela aplicação do modelo DEA BoD SBM. A Figura 1 ilustra os boxplots criados para análise dos oito indicadores. É fundamental notar que alguma variabilidade era esperada uma vez que as macrorregiões brasileiras são muito distintas. Os valores dos indicadores

confirmaram esse fato, mas, em média, todas as regiões obtiveram valores semelhantes para o indicador Nota de regime de trabalho. Isso pode ser explicado pelo fato de que, nas instituições públicas, a maioria dos professores contratados deve trabalhar em tempo integral para a instituição, mas esse padrão não se aplica às instituições privadas no Brasil. Em relação à qualificação do corpo docente, o percentual de mestres contratados é superior ao de doutores e, embora haja variabilidade entre as regiões, o Norte apresenta pior desempenho em ambos os casos.



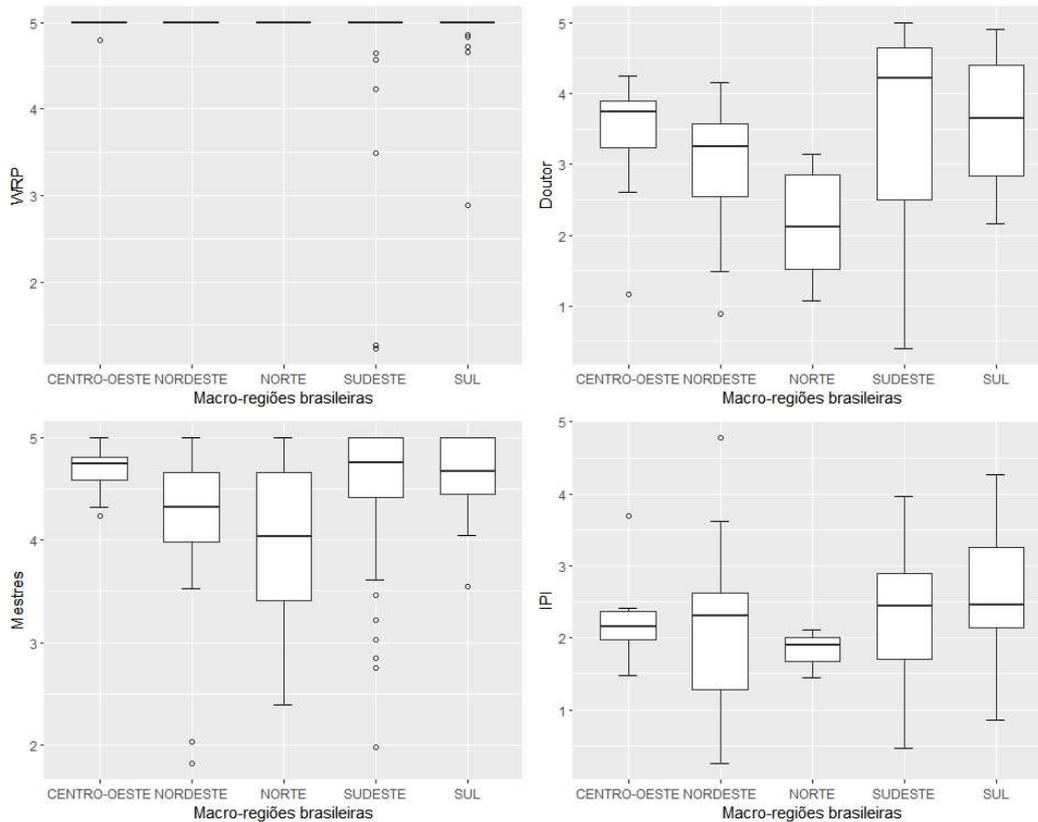


Figura 1: Dispersão dos indicadores de macrorregiões brasileiras

Também é interessante que, em média, todas as regiões tenham desempenho semelhante no fator IDD, mas há uma dispersão significativa no desempenho dos alunos na avaliação do ENADE. Ao observar a percepção dos alunos sobre seus cursos, há maior variabilidade nas Notas referentes às oportunidades de ampliação da formação acadêmica e profissional do que nos outros dois fatores.

A Tabela 3 apresenta o ranking dos cursos de acordo com o indicador composto obtido com o modelo BoD. O desempenho médio dos cursos de graduação em Engenharia de Produção corresponde a 0,7431. Inicialmente, podemos observar que 16 cursos obtiveram as notas máximas, e um deles os três últimos colocados obtiveram notas abaixo de 0,4. O melhor desempenho das 16 DMUs pode ser explicado pela pontuação máxima em pelo menos dois indicadores. Além disso, 4 deles alcançaram desempenho máximo em três dimensões. Também é importante notar que apenas dois cursos eficientes não estão localizados nas regiões Sudeste e Sul. Por outro lado, o curso com menor desempenho situa-se na região Nordeste, e a baixa eficiência é consequência do pior desempenho no IPI, o quarto pior desempenho no DPO e o sétimo pior desempenho no CE.

Tabela 3- Ranking obtido com o SBM-BoD

Rank	Universidade	Escore	Rank	Universidade	Escore	Rank	Universidade	Escore
1	UNESP-Itapeva	1	37	UERJ-Rio de Janeiro	0,8163	73	UFPB-João Pessoa	0,6681
1	UNESP-Itapeva	1	38	UTFPR-Londrina	0,8156	74	FEMASS-Macaé	0,6663
1	UFF-Niterói	1	39	UFF-Petrópolis	0,8145	75	IFMG-Congonhas	0,6592
1	UFMG-Belo Horizonte	1	40	UFF-Volta Redonda	0,8034	76	UEM-Goioerê	0,6592
1	UFRGS-Porto Alegre	1	41	UFV-Rio Paranaíba	0,7980	77	UFERSA-Angicos	0,6464
1	UFC-Russas	1	42	UFF-Volta Redonda	0,7959	78	IFSP-São Paulo	0,6461
1	UFSC-Florianópolis	1	43	UDESC-São Bento do Sul	0,7954	79	UEM-Maringá	0,6439
1	UFSC-Florianópolis	1	44	UFPE-Caruaru	0,7937	80	UFGD-Dourados	0,6438
1	UFRJ-Rio de Janeiro	1	45	UNB-Brasília	0,7879	81	UFBA-Salvador	0,6412
1	UFRJ-Macaé	1	46	UFU-Ituiutaba	0,7835	82	UFCG-Campina Grande	0,6270
1	UTFPR-Apucarana	1	47	UFAL-Penedo	0,7817	83	UEA-Manaus	0,6241
1	CEFET/RJ-Nova Iguaçu	1	48	UFMS-Campo Grande	0,7793	84	UFAM-Itacoatiara	0,6210
1	CEFET/RJ-Itaguaí	1	49	UNESPAR-Campo Mourão	0,7779	85	FESURV-Rio Verde	0,6133
1	UENF-Campos dos Goytacazes	1	50	UFTM-Uberaba	0,7766	86	USCS-São Caetano do Sul	0,6014
1	IFCE-Quixadá	1	51	IFES-Cariacica	0,7666	87	UNIR-Cacoal	0,5868
1	UNESP-Guaratinguetá	1	52	UNIPAMPA-Bagé	0,7664	88	UEM-Maringá	0,5823
17	UFSM-Santa Maria	0,9382	53	UFS-São Cristóvão	0,7644	89	UFES-Vitória	0,5823
18	UFOP-Ouro Preto	0,9303	54	UFCA-Juazeiro do Norte	0,7626	90	UFPR-Curitiba	0,5740
19	UFABC-São Bernardo do Campo	0,9249	55	UFVJM-Teófilo Otoni	0,7573	91	UFMS-Três Lagoas	0,5722
20	UNIFEI-Itajubá	0,9079	56	FURB-Blumenau	0,7570	92	IFMG-Governador Valadares	0,5509
21	UFOP-João Monlevade	0,8974	57	UNIRIO-Rio de Janeiro	0,7549	93	UEM-Maringá	0,5150
22	UFF-Rio das Ostras	0,8881	58	UFG-Catalão	0,7479	94	UFES-São Mateus	0,5082
23	UFV-Viçosa	0,8710	59	UFPI-Teresina	0,7476	95	UFOB-Luís Eduardo Magalhães	0,5038
24	CEFET/RJ-Rio de Janeiro	0,8668	60	UFSCAR-Sorocaba	0,7312	96	UERJ-Resende	0,5011
25	UTFPR-Medianeira	0,8626	61	UFPE-Recife	0,7252	97	UFAL-Delmiro Gouveia	0,4958
26	CEFET/MG-Belo Horizonte	0,8613	62	UNIFEI-Itabira	0,7222	98	UEZO-Rio de Janeiro	0,4753
27	UNICAMP-Limeira	0,8526	63	UNIVASF-Juazeiro	0,7204	99	UFPEL-Pelotas	0,4668
28	UFG-Aparecida de Goiânia	0,8515	64	UEMA-São Luís	0,7162	100	UEM-Maringá	0,4629
29	UNESP-Bauru	0,8500	65	IFMG-BambuÍ	0,7104	101	UNEMAT-Barra do Bugres	0,4620
30	UFSJ-São João del Rei	0,8421	66	UDESC-Joinville	0,7045	102	UNITAU-Taubaté	0,4614
31	UNICAMP-Limeira	0,8364	67	UFPB-João Pessoa	0,7037	103	UEMG-Passos	0,4540
32	UFSC-Florianópolis	0,8361	68	UFAM-Manaus	0,7007	104	UESC-Ilhéus	0,4104
33	UFPEL-Pelotas	0,8354	69	UFSCAR-São Carlos	0,6892	105	UEAP-Macapá	0,3960
34	UTFPR-Ponta Grossa	0,8348	70	UFCG-Sumé	0,6850	106	UEMG-Divinópolis	0,2925
35	UFRN-Natal	0,8231	71	UFPA-Abaetetuba	0,6811	107	URCA-Juazeiro do Norte	0,2616
36	UFPR-Jandaia do Sul	0,8192	72	UFERSA-Mossoró	0,6757			

Como 14 dos 16 cursos eficientes situavam-se nas regiões Sudeste e Sul, investigamos os resultados do indicador composto agregado por macrorregião para analisar o desempenho dos cursos em relação à sua posição geográfica no país. A Figura 2 ilustra os resultados obtidos com o modelo DEA por macrorregiões. É possível verificar que, em média, a região Sudeste obteve o

melhor desempenho, seguida pelas regiões Sul, Nordeste e Centro-Oeste, enquanto o Norte obteve o pior desempenho. O teste estatístico não paramétrico de Kruskal-Wallis foi a ferramenta selecionada para verificar se as medianas das eficiências das regiões brasileiras são iguais. O teste retornou p-valor de 0,009802 e qui-quadrado de 13,323 considerando quatro graus de liberdade e nível de significância de 0,05. Esse valor indica que pelo menos uma das regiões difere significativamente das demais e pode ser entendido como evidência de uma diferença significativa de desempenho em todo o país.

Após a comparação dos resultados das macrorregiões, foi realizada uma análise de correlação entre o ranking fornecido pelo modelo BoD e o fornecido pelo valor contínuo do CPC. Os coeficientes de correlação Rho de Spearman e Tau de Kendall foram calculados com nível de significância de 0,05. Os valores significativos de 0,8737 e 0,7053 foram obtidos para os coeficientes de Spearman e Kendall ( $p < 0,0001$ ). Ambos os valores indicam uma forte correlação entre os dois rankings. No entanto, o uso do modelo BoD não requer uma predefinição de pesos, enquanto os valores de CPC são obtidos com os pesos detalhados na seção 2.

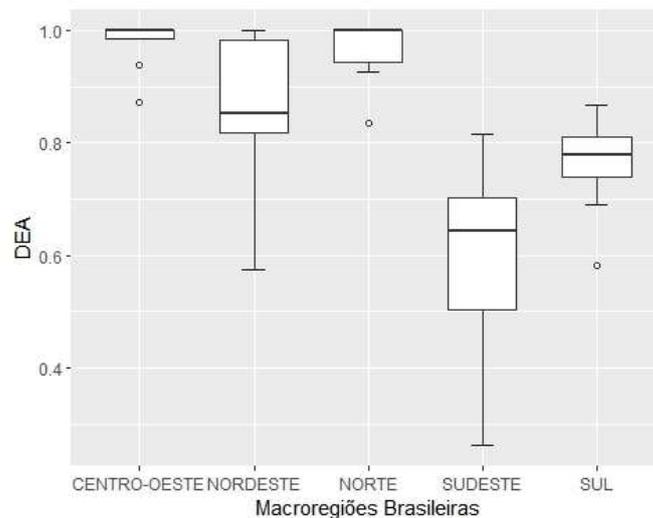


Figura 2: Dispersão do indicador composto entre as macrorregiões brasileiras

## 6. CONCLUSÕES

Este artigo delineou um novo procedimento para avaliar a qualidade dos cursos de graduação brasileiros com o auxílio de um modelo BoD baseado em folgas. A intenção é apresentar uma alternativa válida para agregar os oito indicadores considerados na avaliação dos cursos de graduação sem uma predefinição de pesos. A escolha de uma abordagem baseada em folgas em detrimento de outros modelos DEA é baseada na possibilidade de uma avaliação não radial dos indicadores e, como não há opiniões prévias de especialistas, a abordagem SBM é capaz de fornecer uma avaliação mais equilibrada como evidenciado pelos resultados de Mariano et al. (2021).

A identificação das melhores unidades é útil para desenvolver políticas específicas para melhorar seu desempenho com base nas melhores práticas. É fundamental mencionar que o presente estudo investigou cursos públicos de graduação em Engenharia de Produção para ilustrar a aplicabilidade do método, mas a proposição não se limita a esse contexto e pode ser utilizada para avaliar os demais cursos.

Como direções futuras para expandir os estudos, é necessário discutir os indicadores utilizados na avaliação e comparar os resultados do SBM com alternativas de BoD, como modelos direcionais ou condicionais. O BoD condicional pode apresentar uma alternativa interessante porque permite a consideração de aspectos como a localização geográfica, uma vez que os testes estatísticos indicaram discrepâncias entre as macrorregiões brasileiras.

## REFERÊNCIAS

ALBO, Y.; LANIR, J.; RAFAELI, S. A Conceptual Framework for Visualizing Composite Indicators. **Social Indicators Research**, v. 141, n. 1, p. 1–30, 15 jan. 2019.

BABAEI, S. et al. A new approach for index construction: The case of the road user behavior index. **Computers & Industrial Engineering**, v. 152, 2021.

BERTOLIN, J. Uma proposta de indicadores de desempenho para a educação superior brasileira. **Estudos em Avaliação Educacional**, v. 22, n. 50, p. 471–490, 2011.

CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European Journal of Operational Research**, v. 2, n. 6, p. 429–444, 1978.

CHERCHYE, L. et al. **An introduction to “benefit of the doubt” composite indicators**. **Social Indicators Research**, maio 2007.

DARDHA, E.; ROGGE, N. How’s Life in Your Region? Measuring Regional Material Living Conditions, Quality of Life and Subjective Well-Being in OECD Countries Using a Robust, Conditional Benefit-of-the-Doubt Model. **Social Indicators Research**, v. 151, n. 3, p. 1015–1073, 1 out. 2020.

DE WITTE, K.; SCHILTZ, F. Measuring and explaining organizational effectiveness of school districts: Evidence from a robust and conditional Benefit-of-the-Doubt approach. **European Journal of Operational Research**, v. 267, n. 3, p. 1172–1181, 2018.

FÄRE, R. et al. A benefit-of-the-doubt model with reverse indicators. **European Journal of Operational Research**, v. 278, n. 2, p. 394–400, 16 out. 2019.

FUSCO, E.; VIDOLI, F.; SAHOO, B. K. Spatial heterogeneity in composite indicator: A methodological proposal. **Omega (United Kingdom)**, v. 77, p. 1–14, 1 jun. 2018.

IKUTA, C.; BARREYRO, G. Análise da qualidade dos cursos do Programa Universidade para Todos (PROUNI). **Revista da FAEEBA: Educação e Contemporaneidade**, v. 30, n. 61, p. 344–363, 2021.

INEP. (2020a). Nota Técnica nº 58/2020/CGCQES/DAES. // <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/indicadores-de-qualidade-da-educacao-superior/outros-documentos>. Acessado: 200-05-01.

INEP. (2020). CPC 2019 de 91,6% dos cursos está entre as faixas 3 e 5. // <https://www.gov.br/inep/pt-br/assuntos/noticias/indicadores-de-qualidade-da-educacao-superior/cpc-2019-de-91-6-dos-cursos-esta-entre-as-faixas-3-a-5>. Acessado: 2022-05-06.

KAO, C. Efficiency decomposition and aggregation in network data envelopment analysis. **European Journal of Operational Research**, v. 255, n. 3, p. 778–786, 1 dez. 2016.

KARAGIANNIS, R.; KARAGIANNIS, G. Intra- and inter-group composite indicators using the BoD model. **Socio-Economic Planning Sciences**, v. 61, p. 44–51, 1 mar. 2018.

KARAGIANNIS, R.; KARAGIANNIS, G. Constructing composite indicators with Shannon entropy: The case of Human Development Index. **Socio-Economic Planning Sciences**, v. 70, p. 1–14, 2020.

MARIANO, E. B.; FERRAZ, D.; DE OLIVEIRA GOBBO, S. C. The Human Development Index with Multiple Data Envelopment Analysis Approaches: A Comparative Evaluation Using Social Network Analysis. **Social Indicators Research**, v. 157, n. 2, p. 443–500, 1 set. 2021.

MARIČIĆ, M. et al. Multivariate approach to imposing additional constraints on the Benefit-of-the-Doubt model: The case of QS World University Rankings by Subject. **Croatian Review of Economic, Business and Social Statistics**, v. 2, n. 1, p. 1–14, 1 set. 2016.

OCDE. (2008). Centro Comum de Investigação-Comissão Europeia. Manual de construção de indicadores compostos: Metodologia e guia do usuário. Paris: Editora OCDE.

ROGGE, N. Granting teachers the “benefit of the doubt” in performance evaluations. **International Journal of Educational Management**, v. 25, n. 6, p. 590–614, 2011.

SHI, C.; LAND, K. C. The Data Envelopment Analysis and Equal Weights/Minimax Methods of Composite Social Indicator Construction: a Methodological Study of Data Sensitivity and Robustness. **Applied Research in Quality of Life**, v. 16, n. 4, p. 1689–1716, 1 ago. 2021.

STUMBRIENE, D.; CAMANHO, A. S.; JAKAITIENE, A. The performance of education systems in the light of Europe 2020 strategy. **Annals of Operations Research**, v. 288, n. 2, p. 577–608, 1 maio 2020.

SZUWARZYŃSKI, A. Benefit of the doubt approach to assessing the research performance of Australian universities. **Higher Education Quarterly**, v. 73, n. 2, p. 235–250, 2019.

VAN PUYENBROECK, T. On the Output Orientation of the Benefit-of-the-Doubt-Model. **Social Indicators Research**, v. 139, n. 2, p. 415–431, 1 set. 2018.

VAN PUYENBROECK, T.; ROGGE, N. Geometric mean quantity index numbers with Benefit-of-the-Doubt weights. **European Journal of Operational Research**, v. 256, n. 3, p. 1004–1014, 1 fev. 2017.

ZANELLA, A.; OLIVEIRA, R. M. E S. DE. Avaliação de desempenho na educação superior: uma abordagem utilizando a Análise Envoltória de Dados. **Ciência e Natura**, v. 43, p. e81, 25 ago. 2021.