

CORRELAÇÃO ENTRE GERAÇÃO DISTRIBUÍDA E FATORES SOCIAIS NO BRASIL UTILIZANDO CONJUNTOS FUZZY

1 INTRODUÇÃO

O crescimento da geração distribuída (GD) na última década no Brasil é notória, especialmente devido ao seu enorme potencial para energia solar fotovoltaica. Analisando o avanço da geração distribuída no Brasil, em (SILVA *et al*, 2018) destacam-se as políticas que incentivaram este crescimento a partir do ano de 2013. Uma das principais políticas foi o ProGD – Programa de Desenvolvimento da Geração Distribuída de Energia Elétrica, lançado em 2015, visando ampliar a GD com fontes renováveis em residências, instalações industriais e comerciais, escolas, universidades, hospitais, dentre outros edifícios públicos. Dentre as políticas estaduais, com destaque para os estados de Minas Gerais e Goiás, foram criadas várias formas de incentivo a instalação de sistemas de GD.

Essa expansão acarreta na geração de empregos, beneficiando a sociedade e, portanto, podendo ter efeitos significativos em indicadores sociais. Indicadores como domicílios sem água encanada ou esgoto, assim como mortalidade infantil, taxa de fecundidade, mortalidade por doenças, analfabetismos, entre outros. Apesar do caráter puramente social, podem indicar uma característica econômica importante: o nível de pobreza.

Conforme explorado em (CRESPO e GUROVITZ, 2002), a pobreza possui conceituação complexa, podendo ser estudada apenas do ponto de vista econômico ou a partir de aspectos não-econômicos. O acesso a qualidade de vida é um fator importante na pobreza, que possui caráter multidimensional, refletindo a situação da falta de acesso à cidadania, direitos básicos do cidadão e também à renda básica para sobrevivência.

Devido a pobreza possuir diferentes níveis, assim como a riqueza, não se resumindo ao binário pobre e não-pobre, a utilização da teoria dos conjuntos *fuzzy* se mostra eficaz para sua representação. Os conjuntos *fuzzy* descrevem certos fenômenos utilizando graus variando de 0 a 1 que representam incertezas. De tal forma, a pobreza não é representada entre pobre e não-pobre, mas em diferentes níveis entre os extremos. Para definir a pobreza a partir apenas de indicadores sociais, desconsiderando fatores econômicos, constrói-se um índice *fuzzy* de pobreza com base nos dados sociais manipulados em valores *fuzzy*.

Em (CARVALHO, KERSTENETZKY e DEL VECCHIO, 2000), um índice *fuzzy* de pobreza para a região metropolitana do sudeste brasileiro é obtido a partir de diferentes indicadores. Os autores em (VIEIRA *et al*, 2002) compararam métodos *fuzzy* desenvolvidos para representar a multidimensionalidade da pobreza em diferentes regiões, salientando que este tipo de estudo já está consagrado e em constante expansão. O índice *fuzzy* de pobreza descrito apresenta a flexibilidade dos indicadores utilizados como vantagem, quando comparado ao IDH – Índice de Desenvolvimento Humano. Além disso, este índice informa sobre o principal motivo de pobreza de uma região.

Em (BORGES, 2020), o autor correlaciona dados de geração distribuída de uma região com fatores socioeconômicos, utilizando a teoria de conjuntos clássica, que trata dos chamados conjuntos *crisp*. Para verificar a correlação entre dois conjuntos de variáveis utilizou-se ferramentas estatísticas. As ferramentas mais comuns são os coeficientes de correlação de Pearson, Spearman e Kendall, cada um com suas características.

Para conjuntos *fuzzy*, deve-se adotar outros métodos para calcular sua correlação, sendo baseados em métodos estatísticos ou não. Estes métodos podem retornar

informações sobre o grau de correlação (fraco, médio, forte) e o sentido de correlação (positivo ou negativo), assim como os métodos da teoria clássica dos conjuntos.

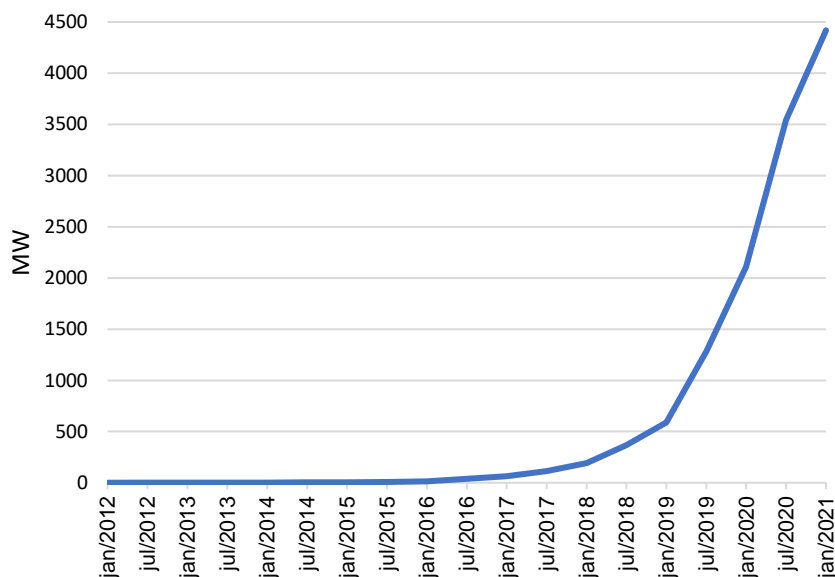
De tal forma, neste trabalho investiga-se a correlação da geração distribuída e a pobreza, obtida como um índice *fuzzy* a partir de fatores sociais, no Brasil. São realizadas análises estaduais e regionais, em dois períodos de tempo distintos, a fim de verificar a influência do crescimento da GD na correlação. Verificado se as variáveis estão fortemente ou fracamente correlacionadas, além de positivamente ou negativamente correlacionadas, apresenta-se as conclusões obtidas a partir da análise dos resultados, levando em consideração que esta correlação pode evidenciar a influência do crescimento da GD ao acesso à cidadania, a direitos básicos, a melhor renda e a qualidade de vida.

2 GERAÇÃO DISTRIBUÍDA

A geração distribuída é definida como a energia gerada próximo ou no local de consumo, podendo ser de diversas fontes de energias renováveis. Com a proximidade, investimentos em expansão dos sistemas de transmissão e distribuição são adiados, de forma benéfica. A GD diminui também impactos ambientais, carregamento das redes, perdas e diversifica a matriz energética.

As análises em (SILVA *et al*, 2018) mostraram que a implantação do ProGD e de outras políticas impulsionaram o crescimento da geração distribuída no Brasil. De 2012 até 2021, a curva de crescimento se assemelha a uma curva exponencial, conforme apresentado na Figura 1.

Figura 1 – Potência Solar Fotovoltaica Instalada (MW)



Fonte: ANEEL.

Os dados de geração distribuída apresentados na figura foram obtidos no site da ANEEL – Agência Nacional de Energia Elétrica, disponibilizados por meio da ferramenta Microsoft Power BI. Para o ano de 2021, os dados regionais de potência instalada do tipo solar fotovoltaica estão apresentados na Tabela 1. Também foram coletados dados estaduais, não apresentados devido a sua extensão.

Tabela 1 – Geração Distribuída no Brasil por Região – 2021

Região	Potência Instalada (kW)
Norte	204.305,32
Nordeste	769.199,84
Sudeste	1.633.760,3
Sul	1.052.140,93
Centro-Oeste	716.185,78

Fonte: ANEEL.

3 INDICADORES SOCIAIS PARA O ÍNDICE DE POBREZA

Os indicadores para construir o índice *fuzzy* de pobreza foram escolhidos de forma a englobar aspectos sociais, como dados de dinâmica populacional, saúde, educação, situação domiciliar e desigualdade, evitando-se assim aspectos econômicos, comumente utilizados em outras análises. A coleta foi realizada a partir dos Indicadores de Desenvolvimento Sustentável – IDS, do Instituto Brasileiro de Geografia e Estatística – IBGE, disponibilizados no Sistema IBGE de Recuperação Automática – SIDRA.

Os indicadores do IDS estão dispostos em quatro dimensões: ambiental, social, econômica e institucional. No entanto, apenas duas se encaixam na análise de pobreza proposta. A dimensão ambiental corresponde aos objetivos de preservação do meio ambiente, com um tema relevante para o presente estudo, o saneamento. A dimensão social está ligada à satisfação das necessidades humanas, qualidade de vida e justiça social e, portanto, extremamente relevante para a análise.

No total, foram coletados 17 indicadores, organizados por região e por estado. A seguir estes são apresentados, com uma breve descrição e o motivo de sua escolha.

- A. Fecundidade:** A taxa de fecundidade total é definida como o número médio de filhos que uma mulher teria ao terminar o período reprodutivo. No Brasil, o período fértil está na faixa de 15 a 49 anos. Costuma ser menor em países desenvolvidos e maior em países em desenvolvimento. Nos países desenvolvidos, mulheres possuem maior acesso a métodos contraceptivos, planejamento familiar, educação sexual e desenvolvimento profissional. Conforme (CRESPO e GUROVITZ, 2002) e reafirmado em (CARVALHO, KERSTENETZKY e DEL VECCHIO, 2000), este indicador é comumente relacionado a pobreza em diversos trabalhos.
- B. Crescimento da população:** A taxa média geométrica de crescimento anual da população residente corresponde ao aumento da quantidade de pessoas existentes em determinada região. É fortemente ligada a taxa de fecundidade, portanto, também se relaciona à pobreza.
- C. Inverso de expectativa de vida:** Tratado como o inverso da esperança de vida ao nascer, que trata do número médio de anos de vida esperados para um recém-nascido. É relacionada ao desenvolvimento de uma região e a qualidade de vida. Como neste trabalho considera-se a pobreza, o seu inverso foi utilizado.
- D. Mortalidade Infantil:** A taxa de mortalidade infantil é definida como a relação entre o número de óbitos de crianças no primeiro ano de vida e o total de crianças nascidas vivas num período para um local. As causas da mortalidade infantil são desnutrição, doenças e, conseqüentemente, pobreza extrema. Regiões mais pobres não oferecem

acompanhamento de gestantes, assim como saneamento básico e sistemas públicos de saúde carentes.

- E. Proporção de mulheres desempregadas ou em empregos informais:** Calculado com base na proporção de mulheres de 16 anos ou mais de idade, ocupadas em trabalhos formais, é um índice que transparece desigualdade entre gêneros. De tal forma, relaciona-se com educação, oportunidades, além da forte relação com a taxa de fecundidade.
- F. Internações hospitalares por doenças transmitidas através do contato com a água:** O tratamento de água inadequado, ou até mesmo sua inexistência é uma característica de regiões menos desenvolvidas. De tal forma, este índice mensura as internações hospitalares por doenças que são transmitidas através do contato com a água. Doenças como diarreia, amebíase, leptospirose, esquistossomose, dentre outras, são causadas pelo contato com água contaminada.
- G. Internações totais por doenças relacionadas ao saneamento:** Conforme citado anteriormente, o saneamento precário ou sua inexistência são características de regiões menos desenvolvidas. De tal forma, doenças, incluindo as transmitidas pela água, podem surgir pela falta de saneamento.
- H. Analfabetos acima de 15 anos:** Indica a taxa de pessoas que não sabem ler ou escrever acima de 15 anos de uma região, relacionada à educação, e, portanto, ao desenvolvimento.
- I. Domicílios sem coleta de lixo direta ou indireta:** Representa os domicílios particulares permanentes inadequados para moradia, sendo a coleta de lixo direta ou indireta critério de adequação. Regiões menos desenvolvidas possuem coleta de lixo precária.
- J. Domicílios sem abastecimento de água da rede geral:** Neste índice, são mensurados os domicílios particulares permanentes inadequados para a moradia, sendo o critério de adequação o abastecimento de água da rede geral. Regiões pobres não possuem abastecimento de água.
- K. Domicílios sem esgotamento sanitário:** Representa os moradores urbanos em domicílios particulares permanentes sem esgotamento sanitário, que, conforme citado, é uma característica fundamental da pobreza.
- L. Índice de Gini:** Este índice mede o grau de concentração de renda em determinado grupo, apontando desigualdade. Quanto mais próximo do número 1 mais evidencia que uma pequena parcela de pessoas detém toda a riqueza. Quanto mais próximo de zero, melhor a distribuição de renda. Apesar do caráter econômico, foi considerado por representar a desigualdade, que é um fator social.
- M. Crianças faltando à escola:** Baseado na taxa de frequência escolar bruta, que expressa a quantidade de crianças que faltam à escola, seja por motivos de desinteresse, falta de materiais escolares, ou para trabalhar para ajudar a família.

- N. Crianças em séries atrasadas:** A taxa de frequência escolar líquida expressa a adequação entre série e idade, de tal forma, verificando se as crianças estão cursando a série recomendada para sua idade. De acordo com (CERQUEIRA *et al*, 2020), as taxas de abandono e atraso escolar são menores entre os jovens mais ricos.
- O. Domicílios com mais de 2 habitantes por dormitório:** Mensura os domicílios particulares permanentes inadequados para moradia, sendo o critério de adequação até dois moradores por dormitório. Este índice está ligado a conforto, e, portanto, melhor qualidade de vida.
- P. Mortalidade por homicídio:** Indica a mortalidade por homicídios em uma região. Em (CERQUEIRA *et al*, 2020), destaca-se que a segurança pública, legislação sobre armas, frequência escolar e desigualdade de renda são fatores que influem no número de homicídios.
- Q. Pessoas não usuárias da internet:** Pessoas de 10 anos ou mais que não são usuárias da internet, seja por desinteresse ou por não dispor de meios de acessá-la, característica de regiões carentes.

4 CONJUNTOS FUZZY

A teoria dos conjuntos *fuzzy* surgiu com a intenção de dar um tratamento matemático a termos linguísticos subjetivos, como “por volta de” e “aproximadamente”. Os autores em (SOUZA e MESQUITA, 2010) descrevem estes conjuntos como modeladores de incertezas, usados para descrever fenômenos utilizando graus que representam qualidades ou verdades parciais.

A pobreza, por exemplo, é um fenômeno relacionado ao mundo sensível, tendo elementos em um conjunto com maior ou menor intensidade. Nestes conjuntos *fuzzy*, também chamados de nebulosos, é comum dizer qual elemento de um conjunto se enquadra melhor em um termo.

A função de pertinência define o grau de pertencimento de um elemento a um conjunto, assumindo valores de 0 (menos pertinente) a 1 (mais pertinente). Para a pobreza, um elemento com pertinência 1 indica que este elemento é o que melhor pertence ao termo “pobreza”, ao passo que pertinência 0 indica o que menos se enquadra no conjunto.

Apesar de estudar incertezas, a teoria dos conjuntos *fuzzy* é fundamentada e definida. De tal forma, conjuntos da teoria clássica podem ser trabalhados nos conjuntos nebulosos. Os dados de potência instalada e os indicadores sociais são dados *crisp*, isto é, dados da teoria de conjuntos clássica. Para encontrar o índice de pobreza de cada região e, posteriormente, a correlação desejada, se faz necessário transformar estes dados em índices-*fuzzy*.

A função de pertinência utilizada foi a apresentada em (CARVALHO, KERSTENETZKY e DEL VECCHIO, 2000) e mostrada a seguir, transformando os dados *crisp* de cada indicador em índices no domínio dos conjuntos *fuzzy*.

$$IF_{ri} = \frac{(R_i - Min_i)}{Max_i - Min_i} \quad (1)$$

Em que:

IF_{ri} : valor do índice-*fuzzy* para o indicador i calculado para a região r ;

R_i : valor *crisp* do indicador i para a região r ;

Min_i : valores *crisp* mínimo dentre todas as regiões para o indicador i ;

Max_i : valores *crisp* máximo dentre todas as regiões para o indicador i .

De tal forma, cada indicador possuirá um elemento totalmente pertinente, com valor 1, e um com valor 0. Na Tabela 2 são apresentados os índices-*fuzzy* para a potência instalada (PI) da Tabela 1, além dos índices para cada indicador social. Devido à extensão dos dados, apenas os dados por região e para o ano de 2021 são apresentados.

Tabela 2 – Índice *Fuzzy* por Região – 2021

Índice	Norte	Nordeste	Sudeste	Sul	Centro-Oeste
PI	0.0000	0.3952	1.0000	0.5931	0.3581
A	1.0000	0.7347	0.0204	0.0000	0.2041
B	1.0000	0.1639	0.1475	0.0000	0.8525
C	1.0000	0.8620	0.0519	0.0000	0.4450
D	1.0000	0.8902	0.1220	0.0000	0.6098
E	0.9712	1.0000	0.0755	0.0000	0.2590
F	0.3571	0.1071	0.2143	1.0000	0.0000
G	1.0000	0.8229	0.0000	0.1655	0.3771
H	0.4132	1.0000	0.0165	0.0000	0.1322
I	1.0000	0.9719	0.0000	0.1292	0.2528
J	1.0000	0.1815	0.0000	0.0040	0.1169
K	1.0000	0.8125	0.0000	0.0625	0.0625
L	0.4792	0.7083	0.5625	0.0000	1.0000
M	0.5526	1.0000	0.0000	0.6053	0.6579
N	1.0000	0.9769	0.0000	0.3815	0.3873
O	1.0000	0.4737	0.3450	0.0000	0.2047
P	0.8820	1.0000	0.0000	0.1461	0.9326
Q	0.0000	0.0088	1.0000	0.8148	0.8849

Fonte: elaborado pelos autores.

Possuindo todos os dados como conjuntos nebulosos, portanto, o próximo passo é estabelecer a pertinência de cada região ao conjunto pobreza.

5 ÍNDICE FUZZY DE POBREZA

O índice *fuzzy* de pobreza é a medida que revela em que dimensão determinada região é pobre, verificando a pertinência de cada uma ao conjunto pobreza. Para isso, se faz necessária a construção da matriz *fuzzy* pobreza, pela qual se transmite a informação de quanto cada indicador uma determinada região precisaria apresentar para pertencer ao conjunto analisado.

A matriz é construída de forma hierárquica, de tal forma, os 17 indicadores são organizados de acordo com o seu peso no conceito de pobreza. A construção da matriz, apesar de estar pautada em estudos de (CRESPO e GUROVITZ, 2002), (CARVALHO, KERSTENETZKY e DEL VECCHIO, 2000), (VIEIRA *et al*, 2002) e (CERQUEIRA *et al*, 2020), tem característica subjetiva. Portanto, a ordem dos indicadores, assim como seu peso, deve ser definida pelo pesquisador, considerando os diversos estudos.

Nas posições iniciais, estão os indicadores que possuem maior influência na pobreza, possuindo maior peso, conforme apresentado na Tabela 3. De tal forma, nas posições finais estão os indicadores que, do conjunto, possuem menor ligação com a pobreza.

O cálculo do primeiro peso foi realizado a partir da soma dos pesos, igualados a 10^3 , portanto, o menor peso foi 0,12195 para a última posição, o dobro para a penúltima, o triplo para a antepenúltima e assim por diante.

Tabela 3 – Matriz *Fuzzy* de Pobreza

Posição	Indicador	Peso
1	D	0,8537
	J	0,8537
	K	0,8537
2	A	0,7317
	F	0,7317
	H	0,7317
	I	0,7317
3	G	0,6098
	M	0,6098
	E	0,6098
	P	0,6098
4	B	0,4878
	N	0,4878
5	C	0,3659
	L	0,3659
6	O	0,2439
7	Q	0,1220

Fonte: elaborado pelos autores.

De posse dos índices da Tabela 3, com exceção dos dados de potência instalada, e da matriz *fuzzy* de pobreza, calcula-se o índice de pobreza agregado, que corresponde ao valor máximo entre os mínimos por indicador. Conforme demonstrado em (CARVALHO, KERSTENETZKY e DEL VECCHIO, 2000), a relação *fuzzy* que descreve a função de pertinência é dada por:

$$IFP_r = \max \{ \min \{ M_i, IF_{ri} \} / i \in I \} \quad (2)$$

Em que:

IFP_r : valor do índice *fuzzy* de pobreza para a região r;

IF_{ri} : valor do índice-*fuzzy* para o indicador i calculado para a região r;

M_i : valor de referência do indicador i da matriz *fuzzy* de pobreza;

I: conjunto de todos os indicadores sociais.

Analisando a Equação (2), percebe-se que o índice *fuzzy* de pobreza de cada região é dado comparando os índices *fuzzy* de cada indicador com o valor de referência da matriz *fuzzy* de pobreza.

Para exemplificar este procedimento, considerando a região Sudeste, primeiramente, para o indicador A – taxa de fecundidade, com índice *fuzzy* de 0,0204, compara-se este valor com o de referência da matriz (0,7317). A comparação é realizada através do mínimo para garantir que a pobreza não seja superestimada, caso o índice seja maior que a referência.

O mesmo procedimento é feito para o restante dos indicadores da região, e ao fim, calcula-se o máximo entre os valores obtidos, calculando a dimensão da pobreza da região e qual dos indicadores é responsável por ela. O máximo para esta região é de 0,3658, devido ao indicador L – Índice de Gini.

Na Tabela 4 são apresentados os índices *fuzzy* de pobreza para as regiões, assim como o principal indicador responsável por esta classificação.

Tabela 4 – Índice *Fuzzy* de Pobreza por Região - 2021

Região	IFP	Indicador
Norte	0.8536	D
Nordeste	0.8536	D
Sudeste	0.3658	L
Sul	0.7317	F
Centro-Oeste	0.6098	M

Fonte: elaborado pelos autores.

6 CORRELAÇÃO ENTRE GERAÇÃO DISTRIBUÍDA E ÍNDICE *FUZZY* DE POBREZA

Na teoria clássica dos conjuntos, os coeficientes de correlação determinam qual a intensidade da relação entre duas variáveis, indicando informações sobre o grau de correlação, geralmente variando de -1 a 1, sendo os valores extremos graus fortes e valores próximos de 0 fracos. Além disso, o sinal do coeficiente evidencia o sentido de correlação, podendo ser positivo, isto é, crescem juntas, ou negativa, ao passo que uma variável cresce, outra decresce.

Na teoria *fuzzy* dos conjuntos, existem métodos semelhantes para se calcular a correlação entre dois conjuntos. Em (WATANABE e IMAIZUMI, 1999), os autores propuseram um coeficiente de correlação para a análise da relação entre o significado da palavra "alto" e das alturas individuais. Em (MITCHELL, 2004), um coeficiente de correlação para conjuntos nebulosos intuicionistas, que possuem tanto função de pertinência quanto função de não pertinência, foi desenvolvido. Como o conjunto de pobreza e de potência instalada possuem apenas função de pertinência, utilizou-se o coeficiente de correlação desenvolvido em (CHIANG e LIN, 2004) para calcular a relação entre a pobreza e a geração distribuída, conforme Equação (3):

$$C_{A,B} = \frac{\sum_{i=1}^n (\mu_A(x_i) - \bar{\mu}_A)(\mu_B(x_i) - \bar{\mu}_B)}{S_A \cdot S_B \cdot (n - 1)} \quad (3)$$

Em que:

- $C_{A,B}$: correlação entre A e B;
- $\mu_B(x_i)$: elementos (x_1, x_2, \dots, x_n) do conjunto A;
- $\mu_B(x_i)$: elementos (x_1, x_2, \dots, x_n) do conjunto B;
- $\bar{\mu}_A, \bar{\mu}_B$: média dos elementos de cada conjunto (A e B);
- S_A, S_B : desvio padrão dos elementos de cada conjunto (A e B);
- n : número de elementos de um dos conjuntos.

Para os conjuntos de potência instalada e índice de pobreza, considerando as regiões, em 2021, a correlação foi calculada como -0,8269, considerada forte, porém negativa, o que demonstra que ao passo que a variável potência instalada cresce, a variável índice de pobreza decresce.

7 ESTUDO DE CASO

O estudo foi inspirado na análise de correlação entre geração distribuída e indicadores socioeconômicos realizado em (BORGES, 2020), utilizando apenas indicadores sociais. Conforme o autor verificou, os aspectos sociais, quando analisados separadamente, possuem correlação fraca com o conjunto de potência instalada na

geração distribuída. De tal forma, através da teoria dos conjuntos *fuzzy* e o índice *fuzzy* de pobreza já apresentados, estes aspectos são incluídos num conceito mais abrangente, possibilitando uma análise mais ampla da correlação.

Primeiramente, coleta-se os dados de potência instalada fotovoltaica para os dois períodos distintos desejados na análise (2013 e 2021), para as regiões e estados. Em seguida, extrai-se dos domínios já citados os dados referentes aos aspectos sociais, para os períodos mais próximos possíveis dos dados já coletados, também para as regiões e estados.

Na íntegra, foram quatro casos analisados, considerando o período e o local. Para cada um, realiza-se os procedimentos descritos ao longo do trabalho, que descrevem o Caso II (Região, 2021). O Caso I é dado por região, porém para ano de 2013. O Caso III é realizado com dados dos estados, para o ano de 2013, ao passo que o Caso IV é realizado para o ano de 2021.

8 ANÁLISE DE CORRELAÇÃO

Para os Caso I, III e IV, realizou-se os cálculos necessários para encontrar o índice *fuzzy* de pobreza de cada um, chegando aos valores apresentados nas Tabelas 5 e 6, que também evidencia o motivo principal da pobreza. De posse destes dados, a correlação calculada para cada caso pode ser vista na Tabela 7.

Tabela 5 – Índice *Fuzzy* de Pobreza por Estado

Estado	2013		2021	
	IFP	Indicador	IFP	Indicador
RO	0.8537	J	0.8537	J
AC	0.8201	J	0.7317	A
AM	0.7317	A	0.7317	A
RR	0.7317	A	0.7059	A
PA	0.8537	J	0.7702	J
AP	0.8537	J	0.8537	D
TO	0.6098	E	0.6089	E
MA	0.8537	D	0.8537	D
PI	0.7317	H	0.8537	K
CE	0.7112	H	0.7317	H
RN	0.6791	H	0.7317	H
PB	0.7317	H	0.7317	H
PE	0.6684	H	0.7294	H
AL	0.8537	D	0.7431	D
SE	0.6845	H	0.7317	H
BA	0.6263	D	0.6176	H
MG	0.5345	M	0.3659	L
ES	0.6098	P	0.6098	P
RJ	0.4878	N	0.4878	N
SP	0.3966	M	0.4296	M
PR	0.6098	M	0.5926	M
SC	0.6098	M	0.7317	F
RS	0.6061	F	0.7317	F
MS	0.6098	M	0.6098	M
MT	0.4878	B	0.6098	M
GO	0.4814	G	0.6098	P
DF	0.4878	B	0.4878	B

Fonte: elaborado pelos autores.

Tabela 6 – Índice *Fuzzy* de Pobreza por Região - 2013

Região	IFP	Indicador
Norte	0.8537	J
Nordeste	0.8537	D
Sudeste	0.4138	F
Sul	0.7317	F
Centro-Oeste	0.5964	P

Fonte: elaborado pelos autores.

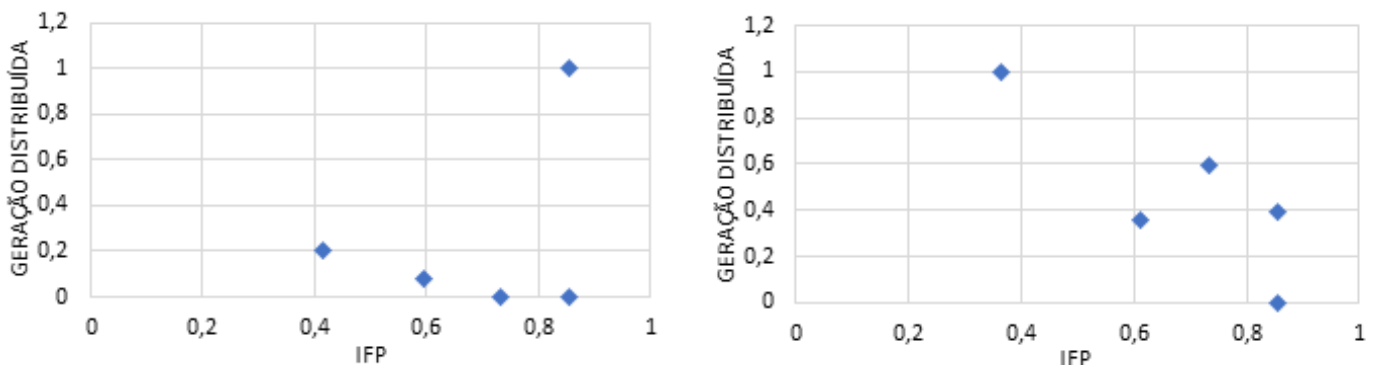
Tabela 7 – Correlação entre IFP e Geração Distribuída

	Correlação
Caso I	0.3182
Caso II	-0.8269
Caso III	-0.0935
Caso IV	-0.6163

Fonte: elaborado pelos autores.

Gráficos de dispersão são comumente utilizados para facilitar o estudo visual da correlação entre variáveis. Quanto mais próximas as marcações, mais correlacionadas são as variáveis. Na Figura 2 são apresentados visualmente a correlação entre estes dois conjuntos para a análise regional.

Figura 2 – Casos I e II: gráfico de dispersão, análise regional, 2013 e 2021, respectivamente.



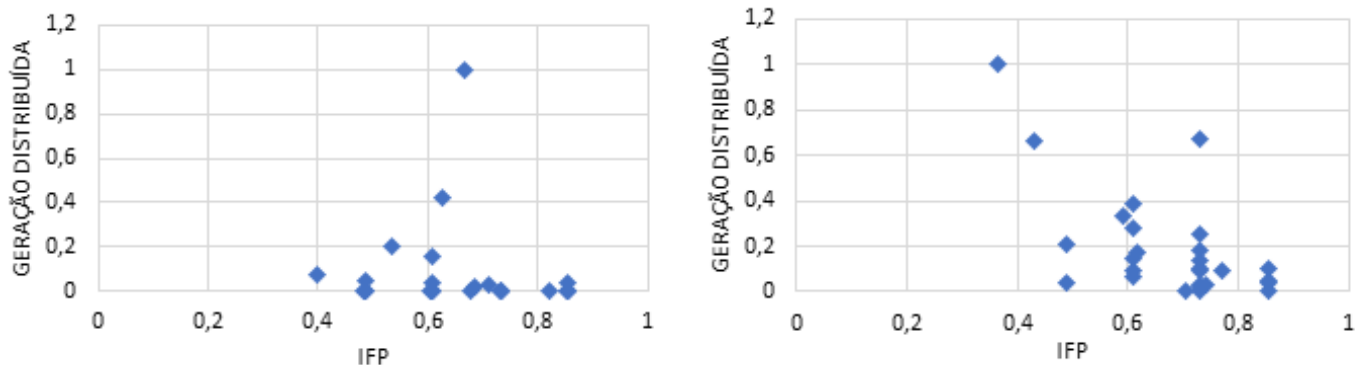
Fonte: elaborado pelos autores.

Analisando o caso regional, verifica-se que para o Caso I, as variáveis possuem correlação quase nula, e à medida que uma cresce, a outra também cresce, conforme esperado pelo coeficiente da Tabela VII, que possui correlação fraca e positiva.

Para o Caso II, verifica-se que à medida que a geração distribuída cresce, a tendência é que o índice de pobreza *fuzzy* (IFP) diminua, de tal forma, a correlação entre estes conjuntos é negativa, e conforme observado e pelo valor calculado, possuem correlação forte.

Estes resultados utilizando as regiões para as análises, devido ao baixo número de elementos no conjunto, podem estar suscetíveis a erros. Os Casos III e IV foram realizados para ter maior confiabilidade nos resultados. A Figura 3 apresenta o gráfico de dispersão para esta análise.

Figura 3 – Casos III e IV: gráfico de dispersão, análise estadual, 2013 e 2021, respectivamente.



Fonte: elaborado pelos autores.

Para o Caso III, verifica-se uma dispersão no gráfico, de tal forma, há uma fraquíssima relação entre as variáveis. O mesmo não pode ser observado no Caso IV, onde verifica-se que as variáveis possuem uma relação forte, mesmo que negativa.

De forma geral, os estudos de casos mostraram que houve um aumento da correlação entre geração distribuída e IFP ao longo dos anos. Verificou-se que enquanto cresce a geração distribuída, o índice *fuzzy* de pobreza diminui. Ainda se evidenciou que ao passar dos anos, essa relação inversa entre os dois conjuntos ficou mais forte.

9 CONCLUSÕES

O crescimento da geração distribuída, principalmente de caráter fotovoltaico, acarreta na geração de empregos, beneficiando a sociedade e, possivelmente, alterando a dimensão de pobreza do local.

Conforme levantado no estudo, utilizando a teoria dos conjuntos nebulosos, a geração distribuída está fortemente ligada a pobreza no Brasil, de forma negativa, sendo que o crescimento de uma evidencia o decaimento da outra.

O estudo, além da análise de correlação proposta, apresentou, durante o cálculo do índice *fuzzy* de pobreza, as diversas causas que influenciam social e economicamente numa região, evidenciando a principal causa de pobreza para cada local.

Finalmente, o investimento na geração fotovoltaica ao longo dos anos pode ter influência nos indicadores sociais de uma localidade, devido ao observado aumento da correlação. Isto evidencia que a geração distribuída está relacionada ao acesso a direitos básicos, cidadania e melhor renda, devendo, portanto, ser levada em consideração em projetos governamentais que visam o crescimento econômico e social.

REFERÊNCIAS

BORGES, Lucas da Mata Santana. **Análise de fatores socioeconômicos em relação ao crescimento da Geração Distribuída no estado de Goiás**. 2020.

CARVALHO, Maristela; KERSTENETZKY, Célia L.; DEL VECCHIO, Renata. **Uma aplicação da teoria dos conjuntos *fuzzy* na pobreza: o caso das Regiões Metropolitanas do Sudeste brasileiro—2000**. Encontro Nacional de Economia (ANPEC), v. 35, 2007.

CERQUEIRA, Daniel Ricardo de Castro Coordenador *et al.* **Atlas da violência 2020**. 2020.

CHIANG, Ding-An; LIN, Nancy P. **Correlation of fuzzy sets**. *Fuzzy sets and systems*, v. 102, n. 2, p. 221-226, 1999.

CRESPO, Antônio Pedro Albernaz; GUROVITZ, Elaine. **A pobreza como um fenômeno multidimensional**. *RAE eletrônica*, v. 1, n. 2, p. 1-12, 2002.

MITCHELL, H. B. **A correlation coefficient for intuitionistic fuzzy sets**. *International Journal of Intelligent Systems*, v. 19, n. 5, p. 483-490, 2004.

SILVA, João Lucas de Souza *et al.* **Análise do Avanço da Geração Distribuída no Brasil**. In: VII Congresso Brasileiro de Energia Solar, Avances em Energías Renovables y Medio Ambiente. 2018. p. 98-117.

SOUZA, Osmar; MESQUITA, Marcos. **Introdução à Teoria dos Conjuntos Fuzzy**. Teoria e Aplicações de Memórias Associativas Morfológicas Nebulosas. Londrina, 2010.

VIEIRA, Cristiéle de Almeida *et. al.* **Pobreza multidimensional e método fuzzy: da teoria à aplicação**. 2002.

WATANABE, Norio; IMAIZUMI, Tadashi. **A fuzzy correlation coefficient for fuzzy random variables**. In: FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No. 99CH36315). IEEE, 1999. p. 1035-1038.

CORRELATION BETWEEN DISTRIBUTED GENERATION AND SOCIAL FACTORS IN BRAZIL USING FUZZY SETS

Abstract: *Due to the potential for photovoltaic solar energy, the growth of distributed generation has advanced in Brazil in recent decades. Government incentive policies are also reasons for this expansion. As a reflex, we can mention the increase in direct and indirect jobs, which reflects in the economy, improving aspects of society and social indicators. These indicators can be translated into an economic and social characteristic known as poverty, a multidimensional measure, having different levels. Due to this characteristic of uncertainties, poverty can be described by the theory of fuzzy sets. In this way, in this work, a fuzzy poverty index was built taking into account 17 different social indicators, which together describe development, quality of life and poverty. With this index and possessing the installed power data from the distributed generation, suitable for the nebulous set, a correlation study between the two variables was carried out. For this, four cases were considered, with regional and state data in two different periods (2013 and 2021), in order to verify the correlation. In possession of the results, an analysis was carried out, showing the existence of the correlation.*

Keywords: *correlation coefficient, distributed generation, fuzzy, poverty, social indicators.*