

Usando a análise fatorial para construir e validar um índice de inserção regional sustentável de usinas hidrelétricas

FÁTIMA R. DE G. FURTADO*

RICARDO CAVALCANTI FURTADO**

Resumo

Este artigo apresenta parte do resultado da pesquisa “Desenvolvimento de ferramenta para monitorar e avaliar a sustentabilidade econômica, social e ambiental dos municípios em áreas de influência de hidrelétricas”, no Brasil, realizada pela UFPE/ FADE em convênio com a CEMIG, no âmbito do programa de P&D da Aneel. A pesquisa teve uma amostra de 40 municípios nas áreas de influência direta de seis hidrelétricas da Cemig, localizadas, em sua maioria, no estado de Minas Gerais, sudeste brasileiro. A partir da amostra, coletou-se um conjunto de dados primários e secundários para a construção de uma matriz com 79 indicadores que cobrem quatro dimensões distintas: qualidade de vida urbana, desenvolvimento econômico, qualidade da gestão pública e qualidade ambiental. Todas as dimensões compõem o Índice de Inserção Regional Sustentável (IIRS). O propósito deste artigo é apresentar a aplicação detalhada do método de Análise Fatorial para a construção do IIRS que é uma ferramenta inédita para avaliar o desempenho das hidrelétricas e da eficácia de programas e políticas públicas locais, permitindo um monitoramento permanente das ações do setor elétrico e da gestão pública para o desenvolvimento sustentável.

Palavras-chave: Análise Fatorial; Inserção Regional Sustentável; Usinas Hidrelétricas; Desenvolvimento Sustentável.

* **FÁTIMA R. DE G. FURTADO** é Arquiteta e Urbanista. PhD em Planning Studies pela Universidade de Londres; professora do Departamento de Arquitetura e Urbanismo da Universidade Federal de Pernambuco; coordena o Laboratório de Estudos Periurbanos (LEPUR), grupo credenciado pelo CNPq desde 2004.

** **Ricardo Cavalcanti Furtado** é Engenheiro Eletricista. PhD em Políticas Energéticas e Ambientais pela Universidade de Londres; sócio-diretor da Diversa Consultoria em Sustentabilidade.

Introdução

Os impactos decorrentes da construção de grandes barragens para implantação de usinas hidrelétricas podem gerar grandes transformações e, por sua vez, acarretar conflitos de interesses nacionais, regionais, locais e setoriais. O processo de implantação de projetos hidrelétricos também causa amplos impactos ambientais, socioculturais e econômicos, além de conflitos institucionais, e acaba por interferir na dinâmica das regiões onde se instalam. Existe uma visão generalizada de que essas regiões, em geral de economias frágeis, são prejudicadas, enquanto os benefícios dos empreendimentos hidrelétricos são destinados a regiões industrializadas e de economias mais fortes. Torna-se necessário, portanto, construir instrumentos que avaliem e monitorem as regiões de inserção dessas hidrelétricas, visando minimizar os impactos negativos e potencializar os impactos positivos, o que conduzirá ao desenvolvimento regional e a uma inserção mais harmônica do empreendimento no seu local de implantação.

Para tanto, foi criado um instrumento para apoiar a tomada de decisão, na medida em que aponte quais os aspectos mais relevantes para garantir a efetividade da alocação de recursos e o monitoramento, além de permitir uma gestão mais eficiente dos recursos alocados de modo a contribuir com o desenvolvimento das áreas afetadas por hidrelétricas.

Nesse sentido, o objetivo da pesquisa, incluída no Ciclo 2013/2015 do Programa de Pesquisa e Desenvolvimento (P&D) da Agência Nacional de Energia Elétrica (ANEEL), foi elaborar uma ferramenta para monitorar e avaliar a inserção regional sustentável de empreendimentos

hidrelétricos, abrangendo a dimensão ambiental, a econômica, a qualidade de vida urbana e a qualidade da gestão pública nos municípios em suas áreas de influência. Para isso, foi construído o Índice de Inserção Regional Sustentável (IIRS), a partir de um conjunto de métodos e técnicas de diferentes campos disciplinares, que sintetiza 79 indicadores, agrupados nas quatro dimensões citadas. Os indicadores também passaram por um processo de normalização e parametrização para que seus valores fossem expressos em unidades de medidas comparáveis, utilizando percentagens, per capita e funções de densidade. Os indicadores foram normalizados em uma escala de 0 a 1. Os dados foram tratados e consolidados num banco de dados eletrônico, desenvolvido com o propósito de apoiar e dar suporte à obtenção do IIRS.

O artigo se propõe a descrever a aplicação do método de Análise Fatorial (AF) no processo de construção e validação do IIRS, estando dividido em duas sessões: a primeira descreve detalhadamente a aplicação do método da AF, o conjunto de técnicas e os resultados obtidos para a amostra analisada; a segunda destaca a importância dessa ferramenta para alcançar o desenvolvimento sustentável.

Análise fatorial na construção do IIRS

A técnica de Análise Fatorial (AF) foi selecionada por sua adequação para a construção de índices e sua capacidade de reduzir um grande conjunto de dados. Dessa forma, facilitou-se a construção e consolidação da matriz de indicadores. Uma vez definidos os indicadores, foram levantados dados, em fontes primárias e secundárias, e valores foram associados a cada um deles. Os dados primários foram levantados em fontes locais, a partir de

um questionário desenvolvido especificamente para essa pesquisa com perguntas tipo múltipla escolha e um espaço para respostas abertas, permitindo que o entrevistado fizesse sugestão de algum novo indicador para melhor mensurar o aspecto indicado. O objetivo do questionário foi coletar as informações pertinentes à inserção regional sustentável dos municípios afetados por empreendimentos hidrelétricos (MARTINS, 2006). Os dados secundários utilizados foram coletados em bases de dados oficiais já consolidadas¹. Os resultados desses levantamentos foram armazenados em um banco de dados por meio de um sistema de aplicativo web desenvolvido para esse fim.

Depois de consolidado o banco de dados, com todas as informações dos 40 municípios com base na matriz de indicadores, foi realizado o teste de correlação de Pearson (r), uma medida de associação bivariada que mede a direção (positiva ou negativa) e o grau de relação linear entre duas variáveis quantitativas (contínuas ou discretas). A associação das variáveis, em termos estatísticos, ocorre quando existe semelhança na distribuição das frequências ou pelo compartilhamento de variância. Especificamente para a correlação de Pearson (r), trata-se de uma medida de variância compartilhada entre duas variáveis e sua distribuição linear considera que o aumento ou diminuição de uma unidade em uma variável X gera o mesmo impacto em

uma Y (FIGUEIREDO FILHO e SILVA JÚNIOR, 2010).

A variação no coeficiente de correlação de Pearson (r) ocorre no intervalo de -1 a 1. O sinal representa a direção da relação que pode ser positiva ou negativa e o valor apresenta a força da relação. Uma correlação que apresenta valores de -1 e 1 é denominada de correlação perfeita e indica que o escore de uma variável pode ser determinado, com exatidão, ao se saber o escore de outra variável. Por outro lado, uma correlação de valor zero indica ausência de relação linear entre as variáveis. Tais valores (0, -1 e 1) raramente são observados nas ciências humanas e sociais e quando ocorrem, em geral, podem representar erros de mensuração das variáveis. De qualquer modo, é seguro afirmar que quanto mais próximo de 1 (independente do sinal) maior será a força da correlação ou o grau de dependência estatística linear entre as variáveis. E quanto mais próximo de 0 menor será a força da correlação (FIGUEIREDO FILHO *et al.*, 2010).

A interpretação dos resultados obtidos a partir da correlação pode ocorrer a partir de diferentes escalas. Na interpretação realizada, baseada em Cohen (1988), os valores entre 0,1 e 0,29 foram considerados baixos; escores entre 0,3 e 0,49 foram considerados como médios; e entre 0,50 e 1 foram considerados altos. Portanto, priorizou-se valores médios e altos de modo a garantir maior robustez na análise. Todavia, para não incorrer em problemas de correlação excessiva (multicolinearidade) entre as variáveis, os escores acima de 0,9 ($r \geq 0,9$) foram excluídos da análise de modo a preservar a adequabilidade da base de dados. Outro cuidado adicional foi a identificação e exclusão de *outliers* na amostra, de modo a evitar a

¹ Instituto Brasileiro de Geografia e Estatística (IBGE) – Banco de dados Cidades, Munic e Sidra; Programa das Nações Unidas para o Desenvolvimento (PNUD) – Atlas; Instituto Chico Mendes de Conservação da Biodiversidade (ICMBio); Ministério do Meio Ambiente (MMA); e Secretaria do Tesouro Nacional (STN) - Finanças Públicas dos Municípios – FINBRA.

multicolinearidade e assegurar que as variáveis foram medidas de forma adequada.

Em conjunto com a análise de correlação, também se empregou o teste do p valor para se obter o nível de confiabilidade da correlação. Via de regra, utiliza-se o nível de 5% de

significância. Desse modo, para um p-valor < 0,05 considerou-se a existência de correlação e para um p-valor > 0,05 a ausência de correlação.

Na Tabela 1, os resultados da correlação entre as diferentes dimensões que compõem o IIRS podem ser observadas.

Tabela 1. Matriz de correlação

		QVU	DE	QGP	QA
Correlação	QVU	1,000	0,398	0,276	0,031
	DE	0,398	1,000	0,383	0,051
	QGP	0,276	0,383	1,000	0,069
	QA	0,031	0,051	0,069	1,000
P-valor	QVU	-	0,006	0,042	0,424
	DE	0,006	-	0,007	0,378
	QGP	0,042	0,007	-	0,336
	QA	0,424	0,378	0,336	-

Como é possível observar, à exceção da dimensão qualidade ambiental (QA)², as demais dimensões apresentaram correlação positiva entre si e p-valor < 0,05, respaldando um dos pressupostos básicos para o emprego da AF que é a existência de correlação entre as variáveis.

A partir da matriz de correlação apresentada acima, pode-se verificar que as dimensões do IIRS estão

correlacionadas positivamente, satisfazendo um dos pressupostos básicos para a criação de um índice: a existência de correlação entre as variáveis.

Em seguida, a AF foi planejada com base em três etapas diferentes aplicadas na construção de cada dimensão que compõe o IIRS (FIGUEIREDO FILHO *et al.*, 2013; 2015): a) verificação da adequabilidade da base de dados com respeito ao nível de mensuração das variáveis; b) determinação da técnica de extração (componentes principais); c) tipo de rotação dos fatores (Varimax).

Definidas as etapas de planejamento da AF, foram aplicados os pressupostos metodológicos apresentados a seguir. Com base no padrão de correlação entre as variáveis, observaram-se aquelas que apresentavam coeficientes superiores ou

² A baixa correlação apresentada pela dimensão QA pode ser parcialmente explicada pela ausência de informações em algumas das variáveis independentes que a compõem, bem como pela natureza regionalizada da informação, associada ao pequeno tamanho da amostra de casos, o que implica em baixa variabilidade dos dados. Supõe-se que o aumento da quantidade de casos no futuro poderá mudar o quadro de baixa correlação da dimensão QA.

próximos a 0,3 (independente do sinal) ou padrões de correlação que apresentassem significância estatística ($p\text{-valor} < 0,05$). Com o objetivo de identificar a adequabilidade da análise fatorial ao conjunto de dados, foi realizado o teste de Kaiser-Meyer-Olkin (KMO), no qual valores superiores ao patamar de 0,5 são considerados aceitáveis. Também foi realizado o teste

de esfericidade de Bartlett (BTS), em que um $p\text{-valor}$ inferior a 0,05 é considerado aceitável.

Os valores obtidos respeitaram os parâmetros estatísticos apresentados por Figueiredo Filho *et al.* (2015) para todos os temas e dimensões do IIRS, conforme Tabela 2, indicando a existência de significância estatística no modelo proposto.

Tabela 2. Parâmetros empregados na AF para a construção das dimensões e temas do IIRS

Nível de Mensuração	Variáveis contínuas e discretas
Amostra	40
Correlação	$> 0,3$
KMO	$> 0,5$
BTS	$p\text{-valor} < 0,05$
Tipo de extração	Rotação ortogonal de tipo Varimax
Variância acumulada	$> 60\%$

Realizados os testes para verificar a adequabilidade da base de dados, empregou-se o método de extração de fatores pela Análise de Componentes Principais (ACP). A ACP procura produzir combinações lineares que apreendam o máximo possível de variância dos indicadores analisados, utilizando toda a variância observada entre as variáveis e produzindo componentes que representem a variância das variáveis observadas. Em geral, essa é a técnica mais utilizada para redução de dados. A extração de fatores foi definida com base no critério do *eigenvalue* (regra de Kaiser), que sugere a extração de fatores com valores acima de 1.

Selecionados os fatores, foram determinados aqueles que melhor

representavam o padrão de correlação entre os indicadores. Com a finalidade de simplificar e clarificar a estrutura dos dados, empregou-se o método da rotação ortogonal de tipo *Varimax*, que procura minimizar o número de variáveis que apresentam altas cargas em cada fator. Cabe destacar também que todas as variáveis empregadas eram do tipo discreta ou contínua, evitando-se o uso, tanto quanto possível, da inclusão de variáveis dicotômicas, que distorcem as tendências centrais. Outro fator importante foi a manutenção de algumas variáveis por razão teórica, pois o pequeno número da amostra (quarenta casos) limitou a obtenção de melhores resultados nas estimativas encontradas. Todavia, a pertinência teórica como parâmetro para manter

algumas variáveis não invalidou os resultados obtidos.

A literatura sugere uma quantidade mínima de 50 casos por variável para se conseguir estimativas razoáveis (HAIR *et al.*, 2006). Todavia, em razão do número de municípios com áreas afetadas pelos reservatórios das seis hidrelétricas avaliadas, foram realizadas análises em apenas 40 municípios referentes às usinas hidrelétricas (UHEs) Emborcação, Irapé, Nova Ponte, Queimado, Rosal e Volta Grande.

Em relação ao pressuposto da medição adequada das variáveis, deparou-se com limitações inerentes à pesquisa social,

com variáveis mal medidas e ausência de informações. Apesar das limitações inerentes à natureza da pesquisa, buscou-se sempre a produção de estimativas confiáveis, esperando-se que, no futuro, a inclusão de mais casos venha proporcionar ajustes ao modelo e, por conseguinte, o seu aperfeiçoamento.

Após o emprego da AF, consolidou-se a versão final da matriz de indicadores, disponível em Furtado e Furtado (2016). Em seguida, foi possível consolidar as dimensões qualidade de vida urbana (QVU), desenvolvimento econômico (DE), qualidade da gestão pública (QGP) e QA, que compõem o IIRS como uma média aritmética.

$$IIRS = \frac{(QVU + DE + QGP + QA)}{4}$$

Para cada dimensão do IIRS, foram verificadas as estatísticas descritivas apresentadas na Tabela 3.

Tabela 3. Estatística descritiva das dimensões do IIRS

Dimensão	N	Mínimo	Máximo	Média	Desvio Padrão	Variância
QVU	40	0,4415	0,8638	0,6815	0,0800	0,006
DE	40	0,2252	0,4535	0,3125	0,0529	0,003
QGP	40	0,2778	0,6531	0,4779	0,0939	0,009
QA	40	0,1756	0,3700	0,2919	0,0491	0,002
Casos válidos	40					

Para cada dimensão, os indicadores foram devidamente normalizados³. Destaca-se, apenas, a baixa variância encontrada nas dimensões, que pode ser explicada em função do baixo número de casos analisados. O baixo quantitativo de casos também reverberou nas estimativas de correlação.

Na Tabela 4, a seguir, apresentam-se os resultados dos testes de KMO e Bartlett, que indicam a adequabilidade da AF aos dados da matriz de indicadores proposta, bem como a existência de significância estatística.

³ À exceção da dimensão QA, na qual alguns indicadores não apresentaram informações.

Tabela 4. Testes de KMO e Bartlett.

Medida de adequação da amostra de Kaiser-Meyer-Olkin	0,632
Teste de esfericidade de Bartlett	0,039

Os resultados obtidos sugerem que a construção do IIRS cumpre os requisitos de uma AF. Apesar da forte limitação quantitativa da amostra, abaixo do mínimo recomendado pela literatura, foi possível verificar sua adequabilidade para a construção e validação da matriz dos indicadores do IIRS.

Conclusões

O método apresentado neste artigo se mostrou adequado para a criação do IIRS. Porém, vale salientar que a construção do índice foi um processo mais complexo e contou com a incorporação de outros aportes metodológicos, além da AF. Para maiores informações, consultar o título “Inserção Regional Sustentável de Usinas Hidrelétricas” (FURTADO e FURTADO, 2016).

Os resultados obtidos sugerem, apesar das várias restrições, a adequabilidade da base de dados para a formulação do IIRS que é uma ferramenta inédita e pode contribuir para o monitoramento e avaliação da inserção regional sustentável de empreendimentos hidrelétricos, em suas diferentes dimensões.

Os resultados obtidos indicam que IIRS é um modelo coerente para interpretação e comparação de áreas afetadas por empreendimentos hidrelétricos. O objetivo dessa ferramenta é facilitar o acompanhamento e a avaliação do progresso alcançado pelos municípios das áreas de influência de UHEs, em

razão das ações socioeconômicas e ambientais implantadas e do dinamismo econômico resultante dos investimentos e encargos tributários pagos pelos concessionários. Se bem direcionados, acredita-se que os recursos gerados pelas UHEs podem alavancar o desenvolvimento na região. Portanto, esse instrumento poderá subsidiar tomadas de decisões e alocação de recursos para otimizar a inserção regional sustentável, compreendida como a potencialização dos benefícios socioeconômicos provenientes da implantação de usinas hidrelétricas, e promover melhorias na qualidade de vida da população local.

Ainda é escassa a produção de ferramentas e estudos que objetivam o monitoramento e o controle da sustentabilidade desses grandes projetos, principalmente os que superam os limites dos aspectos socioeconômicos, incluindo também os aspectos ambientais e especificidades da região de alcance do empreendimento. Destaca-se que a ferramenta apresentada pode ser aplicada em futuros trabalhos. Nesse sentido, a transparência dos procedimentos, apresentados para a construção do IIRS a partir do emprego da AF abordados neste trabalho, pretende facilitar a difusão e replicação desta ferramenta.

Referências

COHEN, J. **Statistical power analysis for the behavioral sciences**. Hillsdale Lawrence: Earlbaum, 1988. 567 p.

FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. da. Visão além do alcance: uma introdução à análise fatorial. **Opinião Pública**, Campinas, v. 16, n. 1, p.160-185, jun. 2010.

FIGUEIREDO FILHO, D. B. *et al.*. Análise Fatorial Garantida ou o Seu Dinheiro de Volta: Uma Introdução à Redução de Dados, **Revista Eletrônica de Ciência Política** [s.l.], v. 5, n. 2, p.185-211, 19 maio 2015.

_____, D. B. *et al.*. ANÁLISE DE COMPONENTES PRINCIPAIS PARA CONSTRUÇÃO DE INDICADORES SOCIAIS. **Revista Brasileira de Biometria**, São Paulo, v. 31, n. 1, p.61-78, dez. 2013.

_____, D. B. *et al.*. ANÁLISE FATORIAL GARANTIDA OU O SEU DINHEIRO DE VOLTA: UMA INTRODUÇÃO À REDUÇÃO DE DADOS. **Revista Eletrônica de Ciência Política**, [s.l.], v. 5, n. 2, p.185-211, 19 maio 2015.

Universidade Federal do Paraná. DOI: 10.5380/recp.v5i2.40368.

FURTADO, F. R. G; FURTADO, R. C. **Inserção Regional Sustentável de Usinas Hidrelétricas**. UFPE/ FADE/ P&D ANEEL/ CEMIG, 1. ed., Belo Horizonte – MG: Rona Editora, 2016.

HAIR, J. F. et al. **Multivariate Data Analysis**. 6. ed. Upper Saddle River (NJ): Pearson Prentice Hall, 2006.

MARTINS, Gilberto de Andrade. **Estudo de Caso: uma estratégia de pesquisa**. São Paulo: Atlas, 2006.

AGRADECIMENTOS

À Cemig Geração e Transmissão, à Universidade Federal de Pernambuco (UFPE), à Fundação de Apoio ao Desenvolvimento da UFPE (FADE) e à Agência Nacional de Energia Elétrica (ANEEL), nossos sinceros agradecimentos pelo apoio financeiro, logístico e científico.