



**ASSOCIAÇÃO ESPACIAL ENTRE *ICT DEVELOPMENT INDEX*
(IDI) E VARIÁVEIS SOCIOECONÔMICAS**

**SPATIAL ASSOCIATION BETWEEN *ICT DEVELOPMENT INDEX*
(IDI) AND SOCIOECONOMIC VARIABLES**

**ASOCIACIÓN ESPACIAL ENTRE EL ÍNDICE DE DESARROLLO
DE LAS TIC (IDI) Y LAS VARIABLES SOCIOECONÓMICAS**

Paulo Kuester Neto¹
Eduardo de Rezende Francisco²

DOI: 10.54751/revistafoco.v16n7-134
Recebido em: 17 de Fevereiro de 2023
Aceito em: 13 de Março de 2023



RESUMO

Este trabalho tem por objetivo analisar a associação entre índices de desenvolvimento humano (IDHM) e Coeficiente de Gini com o *ICT Development Index* (IDI) publicado pelo ITU (*International Telecommunication Union* - ONU) nas Unidades da Federação Brasileira. A hipótese é que a tecnologia tem um caráter transversal, ou seja, parte do desenvolvimento humano e social pode ser explicada pelo uso de tecnologia, já que a mesma poderia influir nos três fatores do IDH. O desenvolvimento de índices compostos tem sido objeto de análise, com vantagens ao oferecer visão holística não limitada à vertente econômica, mas também social e tecnológica. Visto que a distribuição socioeconômica não se dá de maneira uniforme no espaço geográfico, optou-se pela utilização de modelos exploratórios de Estatística Espacial. O arcabouço metodológico se vale de técnicas de *Big Data*, pela utilização e enriquecimento de dados secundários georreferenciados, e de regressão linear múltipla. Para o componente espacial utilizou-se Mapa de *clusters* espaciais e o índice *I* de Moran de auto-correlação espacial. Os resultados apresentam coeficientes de determinação (R^2) superiores a 0,85 na explicação do IDI a partir de IDHM e Gini e *I* de Moran de 0,46, que mostra influência espacial moderada no comportamento do IDI.

Palavras-chave: Big data; estatística espacial; TIC; desigualdade digital; IDH.

¹ Mestre em Tecnologias da Inteligência e Design Digital pela Pontifícia Universidade Católica de São Paulo (PUC-SP). Núcleo de Informação e Coordenação do Ponto BR (NIC.BR). Av. das Nações Unidas, 11541, CEP: 04578-000, São Paulo – SP. E-mail: paulo@nic.br

² Doutor em Administração de Empresas pela Fundação Getúlio Vargas. Fundação Getúlio Vargas - Escola de Administração de Empresas de São Paulo (FGV-EAESP). Av. Nove de Julho, 2029, Bela Vista, CEP: 01332-000, São Paulo – SP, Brasil. E-mail: eduardo.francisco@fgv.br

ABSTRACT

This work aims to analyze the association between human development indices (IDHM) and Gini Coefficient with the ICT Development Index (IDI) published by the ITU (International Telecommunication Union - UN) in the Units of the Brazilian Federation. The hypothesis is that technology has a transversal character, that is, part of human and social development can be explained by the use of technology, since it could influence the three factors of the HDI. The development of composite indices has been the object of analysis, with advantages to offer a holistic vision not limited to the economic, but also social and technological side. Given that the socioeconomic distribution does not occur uniformly in the geographic space, we opted for the use of exploratory models of Spatial Statistics. The methodological framework uses Big Data techniques for the use and enrichment of georeferenced secondary data and multiple linear regression. For the spatial component, we used the spatial clusters map and the Moran index of spatial autocorrelation. The results show coefficients of determination (R²) above 0.85 in the explanation of IDI from IDHM and Gini and I de Moran of 0.46, which shows moderate spatial influence on the behavior of IDI.

Keywords: Big data; spatial statistics; ICT; digital inequality; HDI.

RESUMEN

Este trabajo tiene como objetivo analizar la asociación entre el Índice de Desarrollo Humano (IDH) y el Coeficiente de Gini con el Índice de Desarrollo de las TIC (IDT) publicado por la UIT (Unión Internacional de Telecomunicaciones - ONU) en las Unidades de la Federación Brasileña. La hipótesis es que la tecnología tiene un carácter transversal, es decir, parte del desarrollo humano y social puede ser explicado por el uso de la tecnología, ya que podría influir en los tres factores del IDH. El desarrollo de índices compuestos ha sido objeto de análisis, con ventajas al ofrecer una visión holística no limitada al aspecto económico, sino también social y tecnológico. Dado que la distribución socioeconómica no se produce de manera uniforme en el espacio geográfico, se decidió utilizar modelos exploratorios de Estadística Espacial. El marco metodológico hace uso de técnicas de Big Data, mediante el uso y enriquecimiento de datos secundarios georreferenciados, y regresión lineal múltiple. Para el componente espacial se utilizó el Mapa de clusters espaciales y el índice I de Moran de autocorrelación espacial. Los resultados presentan coeficientes de determinación (R²) superiores a 0,85 en la explicación del IDI a partir de IDHM y Gini e I de Moran de 0,46, lo que muestra una influencia espacial moderada en el comportamiento del IDI.

Palabras clave: Big data; estadística espacial; TIC; desigualdad digital; IDH.

1. Introdução

Ondas de transformações são comuns na história humana e não são de forma alguma monopólio da tecnologia. Porém, de maneira disruptiva, a tecnologia está se tornando um catalisador de mudanças, alterando a maneira como vivemos, como lidamos com o mundo, como consumimos e lidamos com a informação, as tecnologias da informação e comunicação (TIC) estão sem dúvida se mostrando transformadoras seja no âmbito econômico, social ou

político.

O processo acelera-se e embora existam ainda grandes desafios no que concerne a pessoas digitalmente excluídas e muito por fazer em relação aos objetivos da agenda 2030ⁱ da ONU. A tecnologia por seu caráter transversal pode ter um papel inclusivo. Se trabalhada de forma adequada por agentes de políticas públicas (governo e sociedade), pode caminhar na direção de redução de desigualdade, da pobreza, assegurando uma educação inclusiva de qualidade, criando modelos sustentáveis e com infraestruturas resilientes capazes de fomentar a inovação (ONU, 2015).

Que papel teria a tecnologia nesse desenvolvimento? Ela por si só não poderia ser elencada como fator de desenvolvimento social, mas cabe aqui uma reflexão – o poder público pode valer-se de uma composição de índices tecnológico, social e econômico, para embasar políticas públicas em evidências empíricas e tentar reduzir a desigualdade social tão cara ao tecido social.

Este trabalho se insere justamente na tentativa de analisar uma possível relação entre o *ICT Development Index* (IDI) publicado pelo ITU (*International Telecommunication Union* - ONU) nas Unidades da Federação Brasileira e variáveis socioeconômicas representados pelos índices de desenvolvimento humano (IDHM) e Coeficiente de Gini (disparidade de renda). A hipótese é que a tecnologia tem um caráter transversal, ou seja, parte do desenvolvimento humano e social pode ser explicada pelo uso de tecnologia, já que a mesma poderia influir nos três fatores do IDH. Por vislumbrar pontos diferentes do desenvolvimento poderiam ser agrupados para compor um índice composto e talvez um pouco menos determinístico não focado em um simples fator mas olhando por diferentes ângulos: desenvolvimento humano (IDH), índice de Gini (disparidade de renda) e digital (*ICT Development Index* – IDI).

O cerne deste trabalho de pesquisa está na escolha de um índice de tecnologia já bem conceituado e que em sua concepção metodológica estenda a diferentes dimensões o objeto da análise, não se restringindo a dimensão do simples acesso – um índice composto, portanto.

O Índice de Desenvolvimento de TIC IDI (ITU 2017a) foi escolhido por sua maturidade e por estar sendo bem aceito porquanto tem esse rigor metodológico em constante diálogo entre uma área que está em contínua

evolução como a tecnologia e a pesquisa a qual é desenvolvida por uma rede global de pesquisadores sob os auspícios das Nações Unidas.

A despeito de não haver consenso sobre o papel das TIC seja na redução da pobreza, seja no crescimento da economia do país, pesquisadores têm debatido o tema de forma acalorada. Barrantes (2007, p.31) traz a tona um conceito de “pobreza digital”, em referência a falta de bens e serviços em TIC. Se por um lado isso pode ser visto de acordo com o modo mais usual que é a demanda ou a falta de demanda por setores menos favorecidos, por outro, esta pobreza digital poderia estar ligada potencialmente ao iletramento digital.

May, Waema e Bjåstad (2014) reforçam o conceito, sugerindo que a pobreza digital poderia ser incluída como uma sexta dimensão da pobreza, somando-se às pobrezas financeiras, ativos físicos, humanos e sociais.

O debate em torno do impacto das TIC no desenvolvimento traz o estudo divide fronteiras entre os otimistas que acreditam que a tecnologia pode de alguma forma abrir um caminho para que países em desenvolvimentos consigam dar um salto em relação a questões socioeconômicas, e os pessimistas que veem na tecnologia um possível mal que irá causar mais desemprego e aumentar o desemprego e a taxa de exclusão.

May, Waema, e Bjåstad (2014) iluminam o debate ao ilustrar um terceiro cenário intermediário (entre os otimistas e pessimistas), como que contemporizando as duas linhas, e este parece ser uma grande contribuição do estudo. Deve-se reduzir a expectativa sobre o papel das TIC no desenvolvimento dos países em desenvolvimento, porém entender seu papel no esforço para redução da pobreza.

Este percurso teórico permite refletir sobre o papel das TIC e sua relação com o desenvolvimento humano, tendo a tecnologia tomado cada vez mais parte da vida diária, sendo ela necessária para lidar com dimensões diferentes de nossas vidas, educação, saúde e bem estar, recursos governamentais, financeiros e sociais, cabe investigar a possível relação. Não se busca aqui uma relação de causa e consequência, mas uma possível intersecção que se trabalhada de forma adequada, permita a formuladores de políticas públicas, por meio de uma composição de índices já bem embasados e fundamentados, um

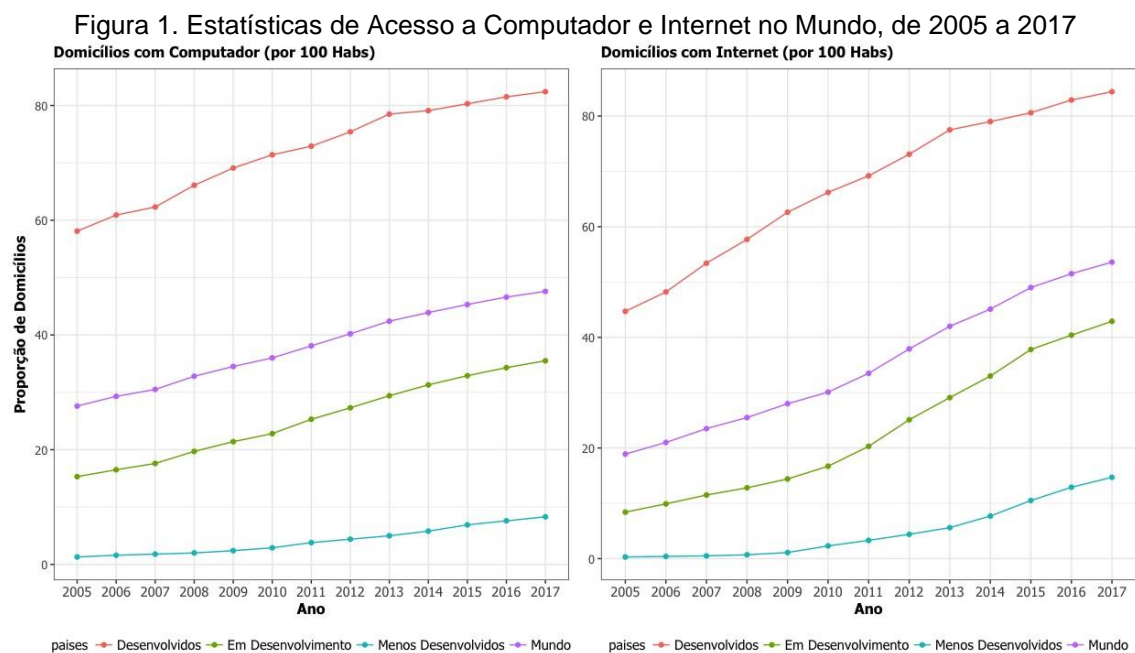
trabalho que se favoreça dessa visão de certa forma mais holística dos aspectos da desigualdade.

Tem-se como justificativa que ao usar essa análise de associação, temos como potencial verificar a possibilidade de usá-los em conjunto para prover dados multifocais que sem dúvida enriqueceria o debate sobre ações e políticas a serem implementadas frente aos desafios.

2. Referencial Teórico

2.1 Dimensões da Desigualdade Digital

A Figura 1 apresenta dados de acesso a computadores e Internet por domicílios no mundo. Enquanto em 2005 nos países desenvolvidos 58,1% dos domicílios já possuíam um computador e 44,7% tinham acesso a Internet, nos países em desenvolvimento os índices eram respectivamente de 15,3% e 8,4%, números bastantes diferentes de 2017 (*International Telecommunication Union - ITU, 2017a*).



Fonte: Elaborado pelos autores, a partir de ITU (2017a), com o uso da ferramenta R.

Em doze anos, somente a partir de 2017 os países em desenvolvimento atingiram os mesmos índices de países desenvolvidos em 2005, o que gera uma disparidade de condições, de habilidades e uso das TIC. Ainda que as

estatísticas apresentadas sejam um guia para um direcionamento mais assertivo de investimentos, visto que é possível perceber questões estruturais e geográficas no que se refere ao acesso a Internet, a exclusão digital não se dá somente nesta esfera, mas também e principalmente no seu uso. Chen e Wellman (2003) reforçam o caráter multidimensional das desigualdades no uso e acesso a Internet, e em esferas que vão desde a comunidade, municípios, estados, até nações, chegando ao estágio global. Segundo o estudo, a exclusão digital não deveria sequer ser conjugada no singular, já que falamos de diferentes exclusões se levarmos em conta seu contexto.

DiMaggio e Hargittai (2003, p.32-34) colocam ainda questões que não se limitam ao aspecto meramente estrutural, mas expandem a discussão para questões sócio-científicas. Para exemplificar as dimensões dessas exclusões, eles elencam: autonomia no uso, habilidades, suporte social, propósito e a parte tecnológica. Os autores ressaltam que a exclusão digital se dá na intersecção de variáveis socioeconômicas, tecnológicas e linguísticas, sejam essas internacionais ou intranacionais.

Não obstante seu uso seja desigual e que não haja consenso em definir um framework único para avaliar o uso das TIC, Heeks (2010) traz outro estudo no qual é possível verificar o desenvolvimento e seus impactos, o que ele chama de Desenvolvimento 2.0, modelos que podem transformar processos e estruturas de desenvolvimento. Isso permite que populações menos favorecidas, valendo-se de TICs possam desenvolver novas maneiras de transformar sua realidade.

Nahon (2006), traz a tona elementos interessantes ao discutir uma perspectiva mais holística no desenvolvimento de índices para avaliar a exclusão digital, em detrimento ao paradigma monolítico. O tema, que no início parecia centrado na questão estrutural e por conseguinte poderia ser medida por *frameworks* e ferramentas para tomada de decisão por agentes públicos, hoje se apresenta mais complexo com foco maior no indivíduo, nos diferentes modos de uso, das competências digitais, o que pode sim ser um fator de desigualdade no acesso e desenvolvimento humano.

O estudo embasa que a desigualdade existe em outros níveis: no setor,

na comunidade e em níveis individuais. Não se deve, portanto, menosprezar discrepâncias que levem em conta esses contextos.

2.2 A questão geográfica e fatores adicionais que atuam na desigualdade de renda

Como incluir as classes menos favorecidas nesse novo contexto digital? A disseminação e apropriação das TIC, segundo Spence e Smith (2010), tem sido uma questão chave da globalização. Como parte dessa transformação, temos assistido a um crescente uso dos telefones celulares ao redor do mundo, uma vez que se utilizam de redes que não necessitam de uma estrutura física de cabeamento. Dito isso, é em países ou regiões menos favorecidas que por meio deste tipo de dispositivo tem-se alterado de forma substancial a vida dos indivíduos.

O uso intenso de dispositivos móveis por parte de usuários na base da pirâmide social (*Bottom of Pyramid - BoP*), nos faz refletir, sobre alguma características interessantes. O custo de conexões banda larga em computadores domésticos em muitas regiões do mundo ainda é impeditivo, visto o padrão de renda dos menos favorecidos, mas a questão da mobilidade do espaço geográfico chama atenção como consequência dessa distorção (ITU, 2017a). Como consequência, e sabendo da não uniformidade de padrão de acesso, o espaço geográfico também se apresenta como uma dimensão na análise do estudo.

Flores (2003) em seu trabalho de pesquisa tenta verificar a relação entre desigualdade de renda e uso das TICs. Embora o estudo seja inconclusivo sobre o papel das TICs na redução da desigualdade de renda, ele aponta que existe uma combinação de fatores como raça, idade, educação, dependência da tecnologia na probabilidade de uso das TIC, embora o rendimento apareça como forte componente no modelo de regressão logística utilizado. Por fim, quando considerado o impacto da renda, fatores geográficos se mostram relevantes.

Tendo visto o arcabouço teórico no qual está suportado este trabalho, buscou-se pesquisar a possível relação das variáveis socioeconômicas com o índice de tecnologia da ITU. Seguiu-se para isso a ideia de índices compostos

desenvolvida por Nahon (2006), a contribuição das TIC para redução da pobreza por May, Waema, e Bjåstad (2014), o potencial papel transformador das TIC por Spence e Smith (2010), relação de desigualdade e uso das TIC de Flores (2003) e a importância do espaço geográfico como medida de análise.

2.3 ICT Development Index (IDI)

O Índice de Desenvolvimento de TIC IDI (ITU 2017b) é um índice dito composto que traz em sua essência onze indicadores ou medidas de referência. Tem sido usado para avaliar, monitorar e realizar comparações no uso e disseminação das TIC ao longo do tempo.

O *framework* conceitual, proposto pela ITU, leva em conta esses três estágios, suportados em um modelo de três fatores principais: infraestrutura e acesso, uso das TIC e habilidades para fazer uso das mesmas. Sendo os dois primeiros estágios representados no índice pelo acesso e intensidade do uso das TIC.

O terceiro estágio depende primordialmente de habilidades e competências para fazer uso dessas tecnologias e no IDI tem um fator fortemente ligado taxa de educação e nível de escolaridade. O *framework* pode ser visto na Figura 2, onde é possível verificar os valores de referência, pesos e medidas que compõem o índice de forma mais clara, como já discorrido anteriormente.

Figura 2. Estrutura do ICT Development Index (IDI).

Indicadores,
referências e pesos

Acesso		Valor de referência	(%)	40
1. Assinaturas de telefone fixo por 100 habitantes		60	20	
2. Assinaturas de telefones celulares por 100 habitantes		120	20	
3. Largura internacional de banda de Internet (bit /s) por usuário da Internet		976'696*	20	
4. Percentagem de agregados familiares com um computador		100	20	
5. Percentagem de residências com acesso à Internet		100	20	
Uso		Valor de referência	(%)	40
6. Percentagem de pessoas usando a Internet		100	33	
7. Assinaturas de banda larga fixa por 100 habitantes		60	33	
8. Assinaturas ativas de banda larga móvel por 100 habitantes		100	33	
Habilidades		Valor de referência	(%)	20
9. Média de anos de escolaridade		15	33	
10. Taxa de matrícula bruta secundária		100	33	
11. Taxa de matrícula bruta		100	33	

Fonte: Adaptado de ITU (2017b)

2.4 Índice de Desenvolvimento Humano (IDH)

O embasamento para o IDH (Stanton, 2007) é dado pela teoria de capacidades de Amartya Sen para melhor capturar a importância dos fins (padrão decente de vida) sobre os meios (como renda per capita). Para instrumentalizar tais capacidades, três são os pilares enfatizados: acesso à saúde, educação e bens, a partir dos quais outras podem ser alcançadas para o estado desejado de vida. Suas teorias e concepções perpassam por conceitos como a teoria da utilidade econômica para abordagem de capacidades humanas. No IDH, componentes como expectativa de vida, alfabetização, matrícula escolar e renda são combinados de forma a fornecer uma única medida de análise comparativa.

Ele é uma medida alternativa ao já bem difundido uso do Produto Interno Bruto (PIB) para medir o nível de desenvolvimento dos países, visto que o mesmo trata apenas de transações econômicas, ignorando questões como liberdade e direitos humanos, e pressupondo a distribuição uniforme de renda dentro de uma mesma sociedade analisada. Confunde-se portanto o conceito de bem estar e bem viver com o crescimento econômico, e potenciais problemas de disparidade de renda e desigualdade.

Como forma de entender e adaptar o IDH ao contexto brasileiro foi criado um índice de desenvolvimento humano municipal. Permite-se por meio do mesmo entender contextos locais da realidade brasileira facilitando a comparação e na busca por melhores parâmetros de avaliação dessas realidades. A metodologia e forma de cálculo fogem ao escopo deste trabalho de pesquisa, mas podem ser verificados em cada uma das suas três dimensões no Atlas para Desenvolvimento Humano (2013).

Para avaliar as diferenças entre o IDH e IDHM (contexto brasileiro) a Tabela 1 abaixo apresenta os valores de referências para as categorias.

Tabela 1: Diferenças do IDH e IDHM – Faixas e Valores.

	IDH Global (2014)	IDHM Brasil (2010)
Muito Baixo	-	até 0,499
Baixo	< 0,550	0,500 a 0,599
Médio	0,550 a 0,699	0,600 a 0,699
Alto	0,700 a 0,799	0,700 a 0,799
Muito Alto	0,800 a 1	0,800 a 1

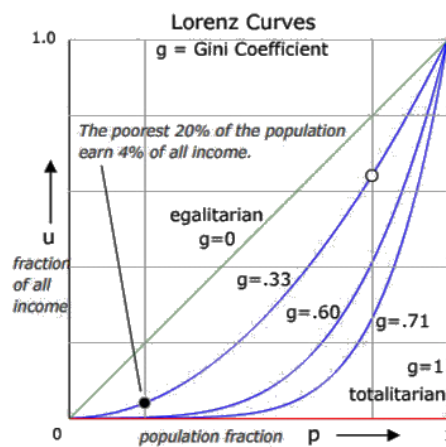
Fonte: Atlas Brasil (2013)

2.5 Coeficiente de Gini

Tendo como seu principal objetivo a mensuração de desigualdade, o coeficiente de Gini, foi desenvolvido pelo italiano Conrado Gini em publicação intitulada “*Variabilità e Mutabilità*” em 1912. De maneira mais recorrente vem sendo utilizado para cálculos de distribuição de renda em dada população, mas o mesmo método pode ser aplicado a exercícios de concentração de riqueza, de terra entre outros (IPECE, 2010).

Pode-se observar sua representação gráfica pela curva de Lorenz apresentada na Figura 3. O eixo x traz a proporção da população a ser analisada e o eixo y a proporção de riqueza acumulada; o resultado do índice Gini é um intervalo entre 0 (zero) e 1 (um), em que o número 0 corresponde a perfeita igualdade representado por uma reta de 45° e 1 uma total desigualdade onde uma única pessoa possuiria toda a renda.

Figura 3. Explicação Gráfica do Índice de Gini.



Fonte: CHESTER, 2010.

Por sua concepção, o índice de Gini pode ser utilizado para verificar questões que não se limitam ao escopo econômico e que estendem sua análise a aplicações de análise de bem-estar em uma dada população, assim como variáveis político filosóficas (IPECE, 2010).

Ainda que com algumas limitações, o índice de Gini apresenta vantagens em relação ao PIB dado que é calculado por uma análise de razão ao invés de uma variável que capturaria o perfil da maioria da população.

3. Dados e Metodologia

Um dos maiores desafios da atual sociedade da informação não é mais a ausência de dados para a realização de estudos e análises, mas a falta de dados atualizados e confiáveis. Utilizar uma metodologia que se apoie em *Big Data* é um grande potencial exatamente por tratar de bases primárias e secundárias, que permite novos olhares para os dados. Governos e agências internacionais têm-se apoiado em bases secundárias como complemento às bases oficiais (ITU, 2014).

Os novos *insights* que podem emergir dessas análises têm despertado enorme atenção de governos, pesquisadores e agências nacionais e internacionais que apoiam o desenvolvimento.

E é justo no monitoramento das TIC que o *Big Data* por meio de dados de questionários domiciliares nacionais, dados de operadores de telefonia ou agências reguladores em conjunto com bases oficiais governamentais podem iluminar a compreensão que temos de comunidades, cidades e estados, estendendo o que já hoje é feito por agências internacionais no âmbito dos países.

3.1 População do Estudo

Foram considerados para este estudo dados provenientes do PNADⁱⁱ, pesquisa nacional com amostra de 1100 municípios (IBGE, 2015), com especial atenção às variáveis demográficas e ao módulo tecnológico que avalia o acesso a Internet e TIC nos domicílios.

Devido a limitações, a amostra está condicionada a alguns níveis de desagregação como descrito no plano metodológico da PNAD, sendo possível a identificação dos níveis Brasil, Grandes Regiões, Unidades da Federação e 9 regiões metropolitanas: Belém, Recife, Salvador, Belo Horizonte, Rio de Janeiro, São Paulo, Curitiba e Porto Alegre.

Como bases auxiliares ou secundárias, e quando os dados estavam disponíveis, foram utilizados dados provenientes da Agência Nacional de Telecomunicações (ANATEL, 2015), com o intuito de compor os fatores do IDI da ITU e dados dos observatórios da Criança e Adolescente e PNE para as taxas brutas de matrículas nos ensinos médio e superior respectivamente.

Dados do Índice de Desenvolvimento Humano (IDH) e Coeficiente de Gini por unidades da federação foram obtidos a partir do site ATLAS Brasil (2013).

Todos os dados que compõem este estudo referem-se ao ano base de 2015; visa-se com isso uma padronização que permita uma análise comparável entre os índices e subíndices utilizados evitando possíveis variações de sazonalidade anual.

3.2 Fonte de Dados

- Os dados utilizados neste estudo estão abaixo descritos, de forma itemizada.
- PNAD Brasil 2015 – Desagregação em Unidades da Federação (UF);
- Sistema de Dados Abertos da ANATEL: Assinatura Telefone Fixo, Assinatura Banda Larga Fixa, Dados de Telefonia Móvel;
- Atlas Brasil – Índice de Desenvolvimento Humano (IDH);
- Atlas Brasil – Coeficiente de Gini;
- Metodologia e Cálculo IDI – ITU;
- Dados de Matrículas Brutas no ensino médio e superior – Observatório da Criança e do adolescente e Observatório do PNE;
- *Shapefiles* (mapas, objetos espaciais) por Unidades da Federação Brasil publicados pelo IBGE.

3.3 Descrição dos Métodos Analíticos

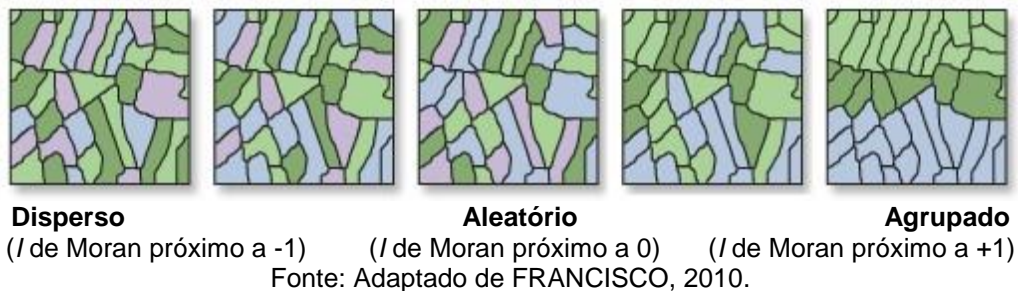
Ainda que este estudo não tenha como objetivo uma construção de modelos preditivos, não obstante em sua essência, vale-se também de técnicas de análise multivariada sobretudo para verificar o poder de relação entre as variáveis e por conseguinte técnicas estatísticas de dependência, mais especificamente a regressão linear múltipla, visando prever e/ou explicar uma variável dependente, neste caso, o IDI (ITU) calculado por UF a partir de um conjunto de variáveis independentes (IDH e Gini).

Adicionalmente à análise de regressão linear múltipla, e considerando-se a hipótese de haver uma influência do espaço, visto que tratamos de unidades

da federação, aplicou-se como passo suplementar uma regressão espacial do tipo SAR (*Spatial Auto-Regressive model*) e o cálculo do índice de autocorrelação espacial I de Moran. O SAR inclui a própria variável dependente como termo auto-regressivo (defasado espacialmente) na explicação do fenômeno, além de outras variáveis independentes (FRANCISCO, 2010; ANSELIN, 1988; GRIFFITH, 1987).

Por meio do índice I de Moran, busca-se verificar uma possível autocorrelação espacial. O resultado do indicador é dado no intervalo entre -1 e 1. Para explicar o índice, fez-se uso de uma representação visual dada pela Figura 4 que mostra a distribuição de um dado fenômeno no espaço.

Figura 4. Auto-correlação Espacial através do I de Moran.



Observa-se aqui um conceito da Estatística Espacial ao passo que deve ser razoável levar em consideração o efeito entre as regiões analisadas e, em grau até maior, suas características e estruturação no espaço geográfico.

Almeida (2012) ressalta o papel da proximidade no que tange à interação espacial dos fenômenos. Não apenas pela quantificação numérica da proximidade, mas também pelo que o autor revela como sendo uma medida de distância social, econômica e política. Ao citar dimensões de proximidade diferentes, um paralelo direto com as unidades de divisão administrativas é traçado, sendo representado pelas unidades federativas ou estados neste trabalho de pesquisa.

3.4 Descrição dos Modelos Preditivos

No esforço de tentar desenvolver este trabalho de modelagem de dados, traz-se todas as variáveis utilizadas para seu cálculo do IDI por Unidade Federativa (UF) levando em conta os 3 fatores principais: acesso e

infraestrutura, uso das TIC e habilidades para fazer o uso eficaz das mesmas. Para que o cálculo pudesse seguir rigorosamente a metodologia e normalização, fez-se uso de dados da população brasileira e número de domicílios obtidos da Pesquisa Nacional por Amostra de Domicílios (PNAD) realizada anualmente entre os períodos do censo nacional brasileiro.

Para melhor acompanhamento, o desenvolvimento deste percurso de modelagem acompanha o *framework* conceitual da Figura 2.

Cada um dos onze indicadores relacionados na Tabela 2 relaciona as variáveis, dimensões ou fatores do IDI e as bases de dados (Fontes) que foram utilizadas para o cálculo do índice por UF. Todos os dados têm como referência o ano de 2015.

A partir das variáveis elencadas na primeira coluna da Tabela 2, tendo o valor bruto sendo obtido a partir das bases de dados relacionadas também na Tabela 2, inicia-se a etapa de cálculo de cada um dos onze indicadores.

O cálculo foi feito com base no relatório metodológico descrito no anexo 1 do documento *Measuring the Information Society Report 2017*, disponível no site da ITU.

Tabela 2: Variáveis, Dimensões e Fontes de Dados para Cálculo do IDI.

Variáveis	Dimensão	Fonte	Medida	Valor de Referência	Peso
A1. Assinatura de Telefone Fixo	Acesso	ANATEL	100 habitantes	60	0.20
A2. Assinatura de Telefone Celular	Acesso	ANATEL	100 habitantes	120	0.20
A3. Largura de banda internacional da Internet (bit / s) por usuário da Internet ⁱⁱⁱ	Acesso	ITU	(bit/s)	976696	0.20
A4. Domicílios com computador ou tablet	Acesso	IBGE PNAD - Módulo Tecnológico Tabela 2.3.5.1	%	100	0.20
A5. Domicílios com Acesso a Internet (A5)	Acesso	IBGE PNAD - Módulo Tecnológico Tabela 2.3.5.2	%	100	0.20
U1. Indivíduos usando a Internet	Uso	IBGE PNAD - Módulo Tecnológico Tabela 1.1.23.2	%	100	0.33
U2. Assinatura de Banda Larga Fixa	Uso	ANATEL	100 habitantes	60	0.33

U3. Assinatura de Banda Larga Móvel	Uso	ANATEL	100 habitantes	100	0.33
H1. Média de Anos de Escolaridade	Habilidade	Atlas Brasil	-	15	0.33
H2. Taxa de Matrícula Bruta no ensino médio	Habilidade	Observatório da Criança e Adolescente	-	100	0.33
H3. Taxa de Matrícula Bruta no Ensino Superior	Habilidade	Observatório do Plano Nacional de Educação	-	100	0.33

Fonte: Elaborado pelos autores, a partir de ITU (2017b).

O resultado (Normalizado*Peso) deve então ser submetido à etapa subsequente na qual serão calculados os valores para cada uma das 3 dimensões ou fatores: Acesso, Uso e Habilidades, seguindo a seguinte distribuição:

- Acesso: A1 a A5;
- Uso: U1 a U3;
- Habilidades: H1 a H3.

O resultado de cada uma das 3 dimensões foi obtido a partir da soma das variáveis de seu grupo. Seguindo o relatório metodológico deve-se então aplicar o peso das dimensões e chega-se ao valor final do IDI.

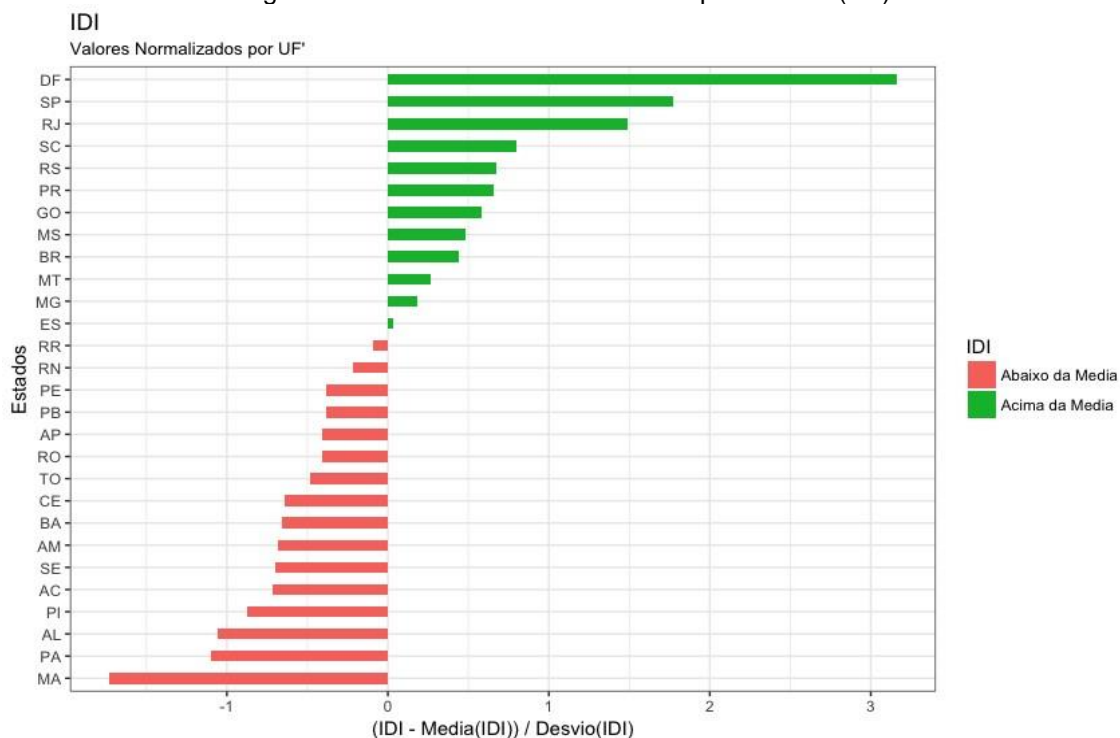
A partir do cálculo final do IDI gerou-se um *dataframe* tendo como variáveis:

- Unidades da Federação(UF);
- IDHM por UF;
- Coeficiente Gini por UF;
- IDHM Dimensão Educação IDHM_E; □ IDHM Dimensão Longevidade IDHM_L; □ IDHM Dimensão Renda IDHM_R.

4. Resultados

Para que fosse possível visualizar de forma mais clara os valores do IDI por UF, optou-se por gerar um valor normalizado: $IDI - Média(IDI) / Desvio\ Padrão(IDI)$ por UF, que pode ser visualizado na Figura 5. Adicionalmente gerou-se um diagrama de dispersão que pode ser visualizado na Figura 6.

Figura 5. Valores Normalizados do IDI pela média (UF).

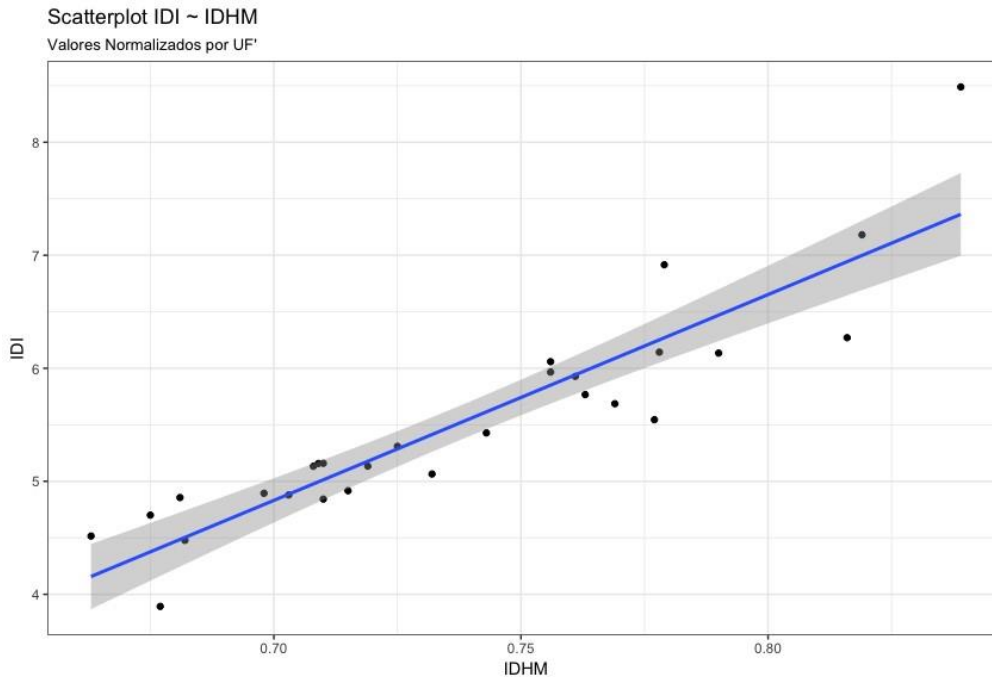


Fonte: Elaborado pelos autores, com o uso da ferramenta R.

Nota-se que parece existir uma replicação de um padrão socioeconômico. Distrito Federal e Estados da região Sudeste, Sul e Centro Oeste estão posicionados no topo da pirâmide, enquanto os estados da região Norte e Nordeste encontram-se na base.

Por meio da Figura 7 nota-se certa similaridade ou correlação numérica entre o padrão do IDI e do IDHM apesar de que as observações tenham ordens de grandeza diferentes.

Figura 6. Diagrama de dispersão entre o IDI e o IDHM por Unidade da Federação.



Fonte: Elaborado pelos autores, com o uso da ferramenta R

Esta análise corrobora o que já era possível visualizar de forma não tão clara nas bases de dados geradas. Existe uma alta correlação entre o IDI e o IDHM (0,9129) e mesmo entre as dimensões do IDHM: IDHM_E (0,8114), IDHM_L (0,7720) e IDHM_R (0,9427).

Extraí-se daqui um viés socioeconômico, uma vez que nem só a dimensão renda está ligada ao índice de tecnologia.

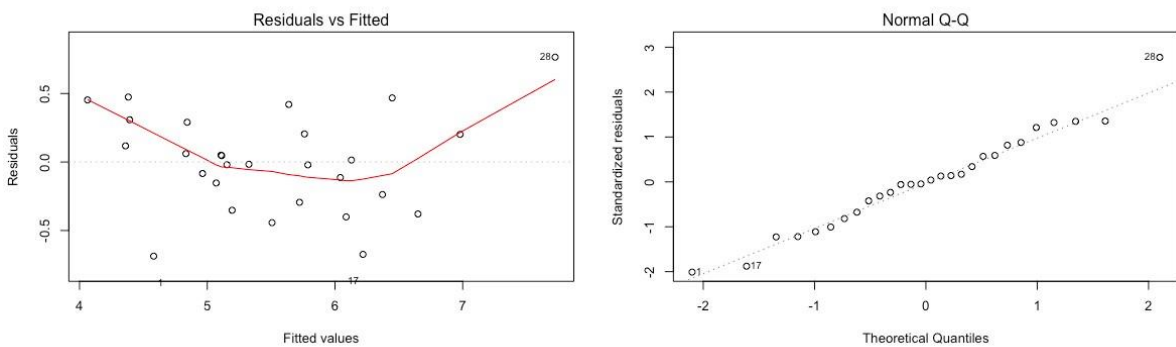
A segunda exploração vem do modelo regressivo. Duas regressões foram testadas: uma tendo apenas o índice geral do IDHM + Gini (Modelo 1), e a segunda excluindo o índice geral do IDHM e deixando as três dimensões: Educação, Renda e Longevidade (Modelo 2). Os valores do R^2 e valor-P podem ser verificados a seguir.

É possível também verificar o resultado visual da regressão por meio do gráfico Normal Q-Q que permite analisar o ajuste dos modelo, assim como o erro residual nas Figuras 7 e 8.

Figura 7. Descrição do Modelo 1: IDI a partir de IDHM e Coeficiente de Gini.

Modelo 1: IDI ~ IDHM + Gini

```
lm(formula = df_final$IDI ~ df_final$IDHM + df_final$GINI)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -10.578      1.741  -6.076 2.38e-06 ***
*** df_final$IDHM  18.757      1.537  12.204 5.00e-12 ***
*** df_final$GINI   4.555      2.284   1.995  0.0571 .
---
Residual standard error: 0.3706 on 25 degrees of freedom
Multiple R-squared:  0.8563, Adjusted R-squared:  0.8448
F-statistic: 74.49 on 2 and 25 DF, p-value: 2.939e-11
```

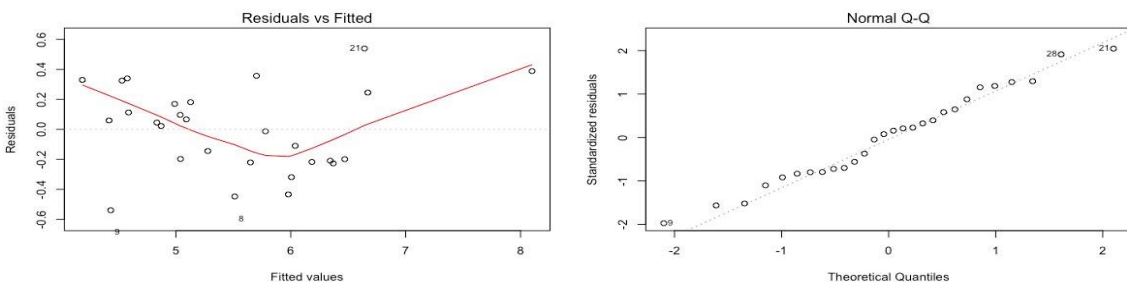


Fonte: Elaborado pelos autores, com o uso da ferramenta R

Figura 8. Descrição do Modelo 2: IDI a partir de IDHM_E, IDHM_L e IDHM_R e Coeficiente de Gini.

Modelo 2: IDI ~ IDH_E + IDH_L + IDH_R + Gini

```
lm(formula = df_final$IDI ~ df_final$GINI + df_final$IDHM_E + df_final$IDHM_L +
df_final$IDHM_R)
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -9.2569      1.7388  -5.324 2.1e-05 ***
df_final$GINI   4.3344      1.9180   2.260 0.0336 *
df_final$IDHM_E  0.9657      1.8749   0.515 0.6114
df_final$IDHM_L  1.5769      2.4695   0.639 0.5294
df_final$IDHM_R 15.1555      2.4220   6.257 2.2e-06 ***
---
Residual standard error: 0.305 on 23 degrees of freedom
Multiple R-squared:  0.9105, Adjusted R-squared:  0.8949
F-statistic: 58.48 on 4 and 23 DF, p-value: 1.015e-11
```



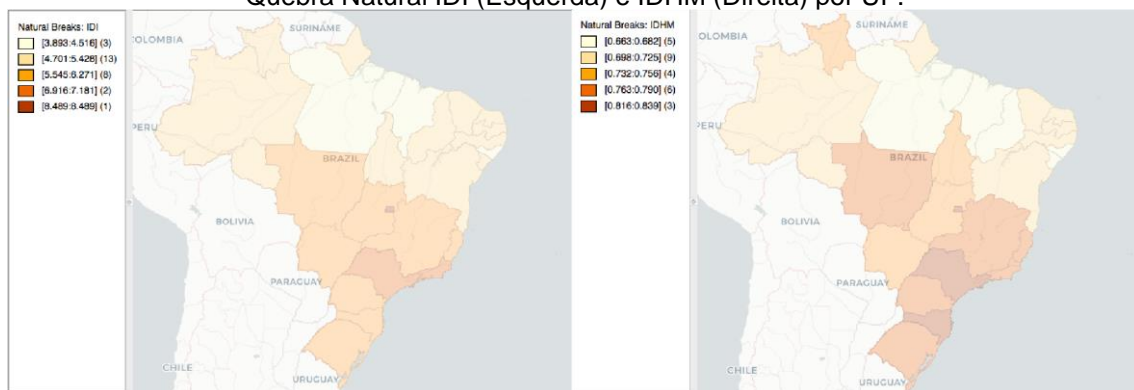
Fonte: Elaborado pelos autores, com o uso da ferramenta R

Observa-se um alto poder explicativo do modelo, tendo o R^2 e R^2 ajustado atingindo 85,63% e 84,48% no Modelo 1 e 91,05% e 89,49% no Modelo 2, respectivamente. O IDHM geral no Modelo 1 e sua dimensão de renda no Modelo 2 têm alto poder explicativo. O Gini que capta a disparidade de renda se faz presente nos dois modelos, porém com mais intensidade quando não utilizamos o índice geral do IDHM.

4.1 Análise Espacial

A fim de verificar uma possível influência do espaço no modelo, fez-se uso do software GeoDA para investigar como se daria a distribuição espacial do IDI nos Estados da Federação (UFs). O resultado pode ser visualizado na Figura 9.

Figura 9. Mapa Temático do IDI (esquerda) e do IDHM (direita) por UF de Distribuição por Quebra Natural IDI (Esquerda) e IDHM (Direita) por UF.



Nota: Classificação das variáveis baseada em Quebras Naturais (método de Jenks)
 Fonte: elaborado pelos autores, com o uso da ferramenta GeoDA.

A paleta de cores é composta por fonte de calor, ou seja, quanto mais clara a cor apresentada no mapa, menor é o valor do IDI/IDHM; o inverso se dá da mesma maneira: quanto mais forte, sendo o vermelho o expoente, mais alto o valor verificado dos índices.

Respeitando as diferenças de escala entre os índices, percebe-se de forma bem significativa a aderência do IDI ao IDHM. O padrão de distribuição por quebras naturais, embora não seja idêntico, sugere um forte relação distribuída no espaço.

A fim de seguir a investigação sobre possível relação do espaço com o modelo, optou-se por gerar os gráficos de *Cluster* e Significância LISA (*Local*

Indicators of Spatial Association), representados pela Figura 10.

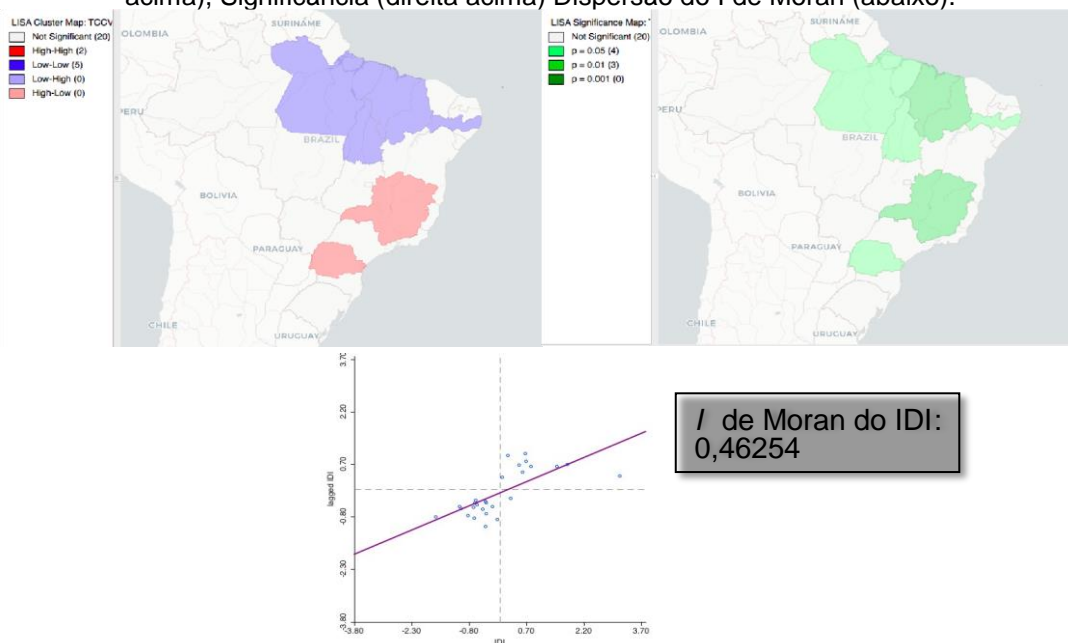
A partir da Figura 10, é possível observar que, embora na maioria dos estados não haja indícios fortes de influência do espaço no modelo, em alguns deles especialmente nos estados da região menos favorecida no Norte/Nordeste (Pará, Tocantins, Maranhão, Piauí e Pernambuco), parece se constituir um *cluster* de baixo valor do IDI.

Minas Gerais e Paraná aparecem também como dois focos de significância para o modelo espacial, incluídos nos estados que possuem valor médio/alto para o IDI, no contexto nacional.

Como última etapa da análise, calculou-se o índice de Moran para este *dataframe* espacial e chegou-se a um valor de 0.46254.

Vale destacar que a regressão espacial do tipo SAR (*Spatial Autoregressive model*) apresentou o termo espacial auto-regressivo (W_y) não significativa para a explicação do IDI. A similaridade espacial dos mapas da Figura 9 fazem-nos perceber que a variável explicativa IDHM parece capturar a influência espacial, não havendo necessidade estatística de permanência do termo espacial no modelo. Não obstante, a dependência espacial do IDI está calculada no I de Moran, não desprezível nessa relação.

Figura 10. Mapa de Clusters Espaciais: Auto-Correlação Espacial (Local Moran's I) (esquerda acima), Significância (direita acima) Dispersão do I de Moran (abaixo).



Fonte: elaborado pelos autores, com o uso da ferramenta GeoDA

5. Contribuições e Considerações Finais

Como resultado deste trabalho de pesquisa, logrou-se utilizar, calcular, normalizar e adaptar o *ICT Development Index* (IDI) para o contexto brasileiro. Por meio de pesquisas amostrais do IBGE, especialmente a Pesquisa Nacional por Amostra de Domicílios (PNAD), as bases sobre uso de Telefonia Fixa, Banda Larga (Fixa e Móvel) da Agência Nacional de Telecomunicações, e dos Observatórios da Criança e Adolescente e Observatório do PNE, foi possível obter os dados necessários para o cálculo do índice com os devidos níveis de desagregação.

Objetivou-se ainda de acordo com o arcabouço teórico desenvolvido ao longo da Introdução, verificar possíveis correlações de variáveis socioeconômicas com o índice de tecnologia da ITU. Lança-se uma semente no desenvolvimento de um índice composto, seguindo os conceitos desenvolvidos por Nahon (2006), que poderia, se devidamente fundamentado, incluir a vertente tecnológica incorporada na investigação do bem-estar social, visto que poderia valer-se da teoria de capacidades de Amartya Sen (IDH).

Ainda que não se possa pensar em relação de causa e efeito, deve-se em trabalhos futuros investigar de forma mais acentuada a contribuição das TIC para redução da pobreza e o seu potencial papel transformador. O estudo parece indicar ainda relação de desigualdade e uso das TIC e, em certa medida, alguma relação com espaço geográfico.

A ideia inicial de investigar a relação do IDHM e Índice de Gini com o IDI foi explorada como um possível caminho de relacionar dados socioeconômicos com uso de TIC, ainda que o Gini não tenha se mostrado tão aderente ao modelo quanto o IDHM.

A tecnologia parece ter um caráter transversal de fato, mas cabe uma melhor investigação histórica para verificar como o modelo se comportaria ao longo do tempo. Acompanhar o IDI por UF seria interessante, visto que por valer-se de bases já bem conceituadas, seria desejável entender sua variação e impacto em outras esferas sociais e econômicas, inclusive com sazonalidades, de curto e médio prazo.

REFERÊNCIAS

ADERA, E., WAEMA, T., & MAY, J. D. (2014). ICT pathways to poverty reduction: Empirical evidence from East and Southern Africa. (IDRC, Ed.). Recuperado em 05 Janeiro, 2018, de <https://www.idrc.ca/en/book/ict-pathways-poverty-reduction-empirical-evidence-east-and-southern-africa>

ALMEIDA, E. (2012). *Econometria espacial*. Campinas, SP: Alínea.

ANATEL. (2017). *Relatório Anual*. Recuperado em 05 de Dezembro, 2017, de Agência Nacional de Telecomunicações: <http://www.anatel.gov.br/dados>

ANSELIN, Luc. *Spatial Econometrics: Methods and Models*. Dordrecht, Holanda: Kluwer Academic, 1988.

ATLAS BRASIL. (2013). *Atlas do Desenvolvimento Humano*. Recuperado em 27 de Novembro, 2017, de <http://atlasbrasil.org.br/2013>

BARRANTES, R. (2007). Analysis of ICT demand: what is digital poverty and how to measure it? *Digital Poverty: Latin American and Caribbean Perspectives*, , 29-53. Recuperado em 10 de Janeiro, 2018, de http://dirsi.net/sites/default/files/dirsi_07_DP02_en.pdf

BARZILAI-NAHON, K. (2006). Gaps and bits: Conceptualizing measurements for digital divide/s. *The information society* , 22 (5), 269-278. Recuperado em 15 de Janeiro, 2018, de <https://doi.org/10.1080/01972240600903953>

CHEN, W., & WELLMAN, B. (2003). Charting and bridging digital divides. *I-Ways, Digest of electronic commerce policy and regulation* , 26 (4), 155-161. Recuperado em 16 de Janeiro, 2018, de https://www.finextra.com/finextra-downloads/featuredocs/international_digital_divide.pdf

CHESTER, M. (2010). The Visualization of Wealth Distribution and the Gini Index. Recuperado em 18 de Janeiro 2018, de <http://chesters.org/marvin/GINI/index.html>

DIMAGGIO, P., HARGITTAI, E., CELESTE, C., & SHAFER, S. (2004). Digital inequality: From unequal access to differentiated use. *Social Inequality* , 355-400. Recuperado em 17 de Dezembro, 2017, de <http://webuse.org/p/c05>

FLORES, C. P. (2003). Measuring the relationship between ICT use and income inequality in Chile. *University of Texas Inequality Project , Working Paper*, 26. Recuperado em 18 de Janeiro, 2018, de <https://pdfs.semanticscholar.org/7a88/32a22fb910243a4b7aad396538421a5ea35f.pdf>

FRANCISCO, E. R. *Indicadores de renda baseados em consumo de energia elétrica: Abordagens domiciliar e regional na perspectiva da estatística espacial*. 2010. 381 f. Tese (Doutorado em Administração de Empresas) - Escola de Administração de Empresas de São Paulo, Fundação Getulio Vargas, São

Paulo, 2010.

GRIFFITH, D. A. *Spatial Autocorrelation – a primer*. Washington, DC: Association of American Geographers, 1987.

HEEKS, R. (2010). Development 2.0: Transformative ICT-enabled development models and impacts. *Development Informatics Short Paper*, 11. Recuperado em 08 de Janeiro, 2018, de http://hummedia.manchester.ac.uk/institutes/gdi/publications/workingpapers/di/di_sp11.pdf

IBGE. (2015). *Pesquisa Nacional por Amostra de Domicílios PNAD*. Recuperado em 16 de Dezembro, 2017, de Instituto Brasileiro de Geografia e Estatística: ftp://ftp.ibge.gov.br/Trabalho_e_Rendimento/Pesquisa_Nacional_por_Amostra_de_Domicilios_continua/Nota_Tecnica/Nota_Tecnica_Diferencas_Metodologicas_das_pesquisas_PNAD_PME_e_PNAD_Continua.pdf

IPECE. (2010). *Entendendo o índice de Gini*. Instituto de Pesquisa e estratégia do Ceará, Fortaleza.

Recuperado em 04 de Fevereiro, 2018, de

http://www.ipece.ce.gov.br/publicacoes/Entendendo_Indice_GINI.pdf

ITU. (2017a). *Estatísticas de Uso das TIC*. Recuperado em 28 de Novembro, 2017, de International Telecommunication Union: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>

ITU. (2017b). *The ICT Development Index (IDI): conceptual framework and methodology*. Recuperado em 10 de Outubro, 2017, de [https://www.itu.int/en/ITU-](https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2017/methodology.aspx)

[D/Statistics/Pages/publications/mis2017/methodology.aspx](https://www.itu.int/en/ITU-D/Statistics/Pages/publications/mis2017/methodology.aspx) ITU. (2014). *The role of big data for ICT monitoring and for development*. Recuperado em 15 de Janeiro, 2018, de https://www.itu.int/en/ITU-D/Statistics/Documents/bigdata/MIS2014_Chapter5.pdf

MAY, J. W. (2014). Introduction: The ICT/poverty nexus in Africa. ICT pathways to poverty reduction: Empirical evidence from East and Southern Africa. 1-32. Recuperado em 13 de Dezembro de 2017, de <https://www.idrc.ca/en/book/ict-pathways-poverty-reduction-empirical-evidence-east-and-southern-africa>

OBSERVATÓRIO DA CRIANÇA E ADOLESCENTE. (2017). (FUNDAÇÃO ABRINQ). Recuperado em 02 Fevereiro, 2018, de Observatório da Criança e Adolescente: <http://observatoriocrianca.org.br>

ONU. (2015). *Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável*.

(Nações Unidas no Brasil) Recuperado em 20
de Janeiro, 2018, de
<http://nacoesunidas.org/pos2015/agenda2030>

PNE. (2017). *Todos pela Educação*. Recuperado em 10 de Dezembro, 2018, de
Observatório do PNE: <http://www.observatoriodopne.org.br>

SPENCE, R., & SMITH, M. (2010). ICT, development, and poverty reduction:
Five emerging stories. *Information Technologies & International Development* ,
6 (SE), 11. Recuperado em 23 de Janeiro, 2018, de :
<http://itidjournal.org/itid/article/view/616/256>

ⁱ 17 objetivos de desenvolvimento e 169 metas que buscam atingir o que não foi possível do legado dos objetivos do milênio. Combinados se traduzem por meio de três dimensões do desenvolvimento sustentável: econômica, social e ambiental.

ⁱⁱ Pesquisa Nacional por Amostra de Domicílios realizada anualmente entre o período do censo que tem periodicidade de 10 anos. Tem por objetivo principal verificar características demográficas e socioeconômicas da população.

ⁱⁱⁱ Dada a concepção estrutural da Internet e especialmente pela distribuição regional assimétrica no contexto brasileiro, esse indicador foi obtido a partir do valor informado pela ITU, não sendo possível seu cálculo por UF. O valor foi padronizado para todo o dataframe.