



UNIVERSIDADE CATÓLICA DE PERNAMBUCO
PRÓ-REITORIA ACADÊMICA
COORDENAÇÃO GERAL DE PÓS-GRADUAÇÃO
MESTRADO EM DESENVOLVIMENTO DE PROCESSOS AMBIENTAIS

VICENTE NATANAEL LIMA SILVA

**MODELAGEM DE DADOS CLIMÁTICOS E
SOCIOECONÔMICOS EM MUNICÍPIOS DO ESTADO
DE PERNAMBUCO UTILIZANDO ANÁLISE DE
COMPONENTES PRINCIPAIS (ACP)**

RECIFE

2017

VICENTE NATANAEL LIMA SILVA

**MODELAGEM DE DADOS CLIMÁTICOS E
SOCIOECONÔMICOS EM MUNICÍPIOS DO ESTADO
DE PERNAMBUCO UTILIZANDO ANÁLISE DE
COMPONENTES PRINCIPAIS (ACP)**

Dissertação apresentada ao Programa de Pós-Graduação em Desenvolvimento em Processos Ambientais Universidade Católica de Pernambuco (UNICAP) como pré-requisito para obtenção do título de Mestre em **Desenvolvimento de Processos Ambientais**.

Área de Concentração: Desenvolvimento em Processos Ambientais

Linha de Pesquisa: Modelagem e Otimização de Processos.

Orientador: Profº. Drº. Valdemir Alexandre dos Santos

Recife

2017

Silva, Vicente Natanael Lima

Modelagem de Dados Climáticos e Socioeconômicos em Municípios do Estado de Pernambuco Utilizando Análise de Componentes Principais (ACP) / Vicente Natanael Lima Silva; orientador Valdemir Alexandre dos Santos, 2017, 81 p.

Dissertação (Mestrado) - Universidade Católica de Pernambuco. Pró-reitoria Acadêmica. Curso de Mestrado em Desenvolvimento de Processos Ambientais, 2017.

1. Elementos Climáticos.
2. Indicadores Socioeconômicos.
3. Análise de Componentes Principais.

Modelagem de Dados Climáticos e Socioeconômicos em Municípios do Estado de Pernambuco Utilizando Análise de Componentes Principais (ACP)

VICENTE NATANAEL LIMA SILVA

Comissão Examinadora:

Profº. Drº Valdemir Alexandre dos Santos (Orientador)
Universidade Católica de Pernambuco – UNICAP

Profº. Drº Ranyére Silva Nóbrega
Universidade Federal de Pernambuco – UFPE

Profº. Drº Francisco Madeiro Bernardino Júnior
Universidade Católica de Pernambuco – UNICAP

Defendida em ____/____/____

Coordenadora: Profª. Drª Clarissa Daisy Costa Albuquerque

AGRADECIMENTOS

Agradeço primeiramente a Deus que sempre esteve comigo em todos os momentos, por ter me dado força, sabedoria e saúde nas horas mais difíceis, por ter me permitido realizar este curso de Mestrado;

Aos meus pais, Maria Ivonete Ferreira de Lima e Carlos Roberto da Silva (in memoria), que sempre se sacrificaram para o meu melhor, pelo estímulo ao estudo, educação e amor;

Aos meus familiares e em especial a minha irmã Danielly Lima da Silva, pelo companheirismo e tolerância nos momentos difíceis e delicados que passamos juntos e ao meu primo Wagner José da Silva, pela amizade verdadeira que sempre compartilhamos.

A minha amiga Cíntia Maria do Nascimento, por tudo que vivemos, sem ela não seria possível meu ingresso no Mestrado.

Ao Prof^o Dr. Valdemir Alexandre dos Santos por ter me proporcionado uma orientação não só acadêmica mas também de ensinamentos para o resto da vida. Aos professores Ranyere Silva Nóbrega (Universidade Federal de Pernambuco) e Francisco Madeiro Bernardino Júnior (Universidade Católica de Pernambuco) por comporem a banca examinadora deste trabalho.

A Universidade Católica de Pernambuco (UNICAP), na pessoa da Coordenadora do Mestrado em Desenvolvimento de Processos Ambientais, Prof^a Dr^a. Clarissa Daisy Costa Albuquerque e a todos os meus colegas.

A Prof^a Dr^a. Aline Elesbão do Nascimento pelos ensinamentos e dedicação e pelas aulas de excepcional qualidade ministradas durante o curso.

A Coordenação de Pessoal de ensino Superior (CAPES) pela concessão da bolsa de Mestrado.

A todos que direta ou indiretamente contribuíram para a realização deste trabalho.

SUMÁRIO

LISTA DE ILUSTRAÇÕES	VIII
LISTA DE TABELAS	XI
RESUMO	X
ABSTRACT	XI
CAPÍTULO I	12
1 - INTRODUÇÃO	13
2 - OBJETIVOS.....	15
2.1 - OBJETIVO GERAL	15
2.2 - OBJETIVOS ESPECÍFICOS.....	15
3. REVISÃO DA LITERATURA	16
3.1 - RELAÇÃO ENTRE ELEMENTOS DO CLIMA E ATIVIDADES HUMANAS	16
3.1.2 – Influencia nas Atividades Econômicas em Pernambuco	18
3.1.3 - Implicações na Saúde	22
3.1.4 - Abastecimento de Água	25
3.1.5 - Eventos Extremos Climáticos.....	25
3.2 - ANÁLISE DE COMPONENTES PRINCIPAIS (ACP)	27
3.2.1 - Matrizes, autovalores e autovetores.....	32
3.2.2 - Seleção e Interpretação dos Componentes Principais.....	33
4 - PROCEDIMENTOS METODOLÓGICOS	37
4.1 - SELEÇÃO DOS MUNICÍPIOS E ANÁLISE DAS VARIÁVEIS.....	37
5. REFERÊNCIAS BIBLIOGRÁFICAS	40
CAPÍTULO II	46
RESUMO.....	47
1 - INTRODUÇÃO	48
2 - MATERIAL E MÉTODOS	49

2.1 - CARACTERIZAÇÃO DA ÁREA DE ESTUDOS	49
3.2 - ASPECTOS FISIAGRÁFICOS E SOCIOECONÔMICOS	50
3.3 - ANÁLISE DE COMPONENTES PRINCIPAIS (ACP)	51
4 - RESULTADOS E DISCUSSÃO	54
5 - CONCLUSÕES	58
6 - REFERÊNCIAS BIBLIOGRÁFICAS	59

<u>CAPÍTULO III</u>	<u>61</u>
----------------------------	------------------

CONCLUSÕES GERAIS	62
--------------------------------	-----------

<u>ANEXOS</u>	<u>64</u>
----------------------	------------------

Socioeconomic Diagnosis Based on Statistical Analysis of Meteorological Variables: A Case Study in Pernambuco/Brazil.....	65
--	-----------

Confirmação de Submissão do Periódico	78
Normas Para Submissão no Periódico Internacional	79

LISTA DE ILUSTRAÇÕES

Figura 1. Principais fatores climáticos	17
Figura 2. Principais elementos do clima	18
Figura 3. Participação dos principais produtos agrícolas em Pernambuco 2010	19
Figura 4. Composição setorial da economia de Pernambuco por região em 2011	21
Figura 5. Regime Pluviométrico de Pernambuco	26
Figura 6. Esquema da aplicação da Análise de Componentes Principais.....	28
Figura 7. Representação gráfica dos autovalores e autovetores	33
Figura 8. Explicação da proporção de variação de cada CP.....	35
Figura 9. Círculo de correlação.	36
Figura 10. Mapa da localização dos Municípios Estudados	38
Figura 11. Fluxograma referente ao processo metodológico adotado.....	39

LISTA DE TABELAS

Tabela 1. Efeitos físico-psicológicos dos elementos do clima no homem.	23
Tabela 2. Tabela de critério de Kaiser para variancia acumulada na ACP.....	35

RESUMO

No Estado de Pernambuco, assim como em toda a região do Nordeste do Brasil, é evidente a expressiva interação existente entre os elementos do clima e as atividades humanas. Inúmeros estudos científicos já demonstraram uma significativa correlação entre o comportamento climático com os aspectos sociais, econômicos, culturais, etc. Este trabalho serviu como estudo de caso da aplicação da técnica estatística multivariada de Análise de Componentes Principais (ACP) na confecção de diagnósticos socioeconômicos, onde foram utilizados os elementos do clima como independentes sobre as variáveis respostas socioeconômicas (Produto Interno Bruto e Índice de Desenvolvimento Municipal) de alguns municípios que apresentaram expressivo desenvolvimento no Estado de Pernambuco – Brasil, entre os anos de 1999 e 2013. Mesmo considerando as diferenças climáticas, socioeconômicas e a imprescindível dependência da água para o desenvolvimento econômico dos municípios estudados, a ACP demonstrou que os índices socioeconômicos dos municípios localizados no Sertão (Petrolina e Arcoverde) apresentarão maior correlação com os índices de temperaturas e Insolação, no Agreste e Zona da Mata (Garanhuns e Surubim) a evaporação e temperaturas, no Litoral (Recife) a precipitação e umidade. ACP mostrou-se também efetiva em permitir a retirada ou descarte de variáveis que apresentaram baixa variabilidade ou foram redundantes por estarem correlacionadas com as de maior importância para dois primeiros componentes principais. A compreensão do comportamento dos elementos do clima e de suas consequências sobre as atividades humanas é de fundamental importância no auxílio às políticas públicas, que visem à mitigação de efeitos adversos provocados pelas alterações ambientais.

Palavras-chaves: Modelagem de Dados, Diagnósticos Socioeconômicos, Análise de Componentes Principais.

ABSTRACT

In the State of Pernambuco, as well as throughout the Northeast region of Brazil, the expressive interaction between climate elements and human activities is evident. Numerous scientific studies have already demonstrated a significant correlation between climate behavior with social, economic, cultural, etc. This work served as a case study of the application of the multivariate statistical technique of Principal Components Analysis (PCA) in the making of socioeconomic diagnoses, where the elements of the climate were used as independent variables on the socioeconomic responses (Gross Domestic Product and Municipal Development Index) Of some municipalities that presented significant development in the State of Pernambuco - Brazil, between 1999 and 2013. Even considering the climatic, socioeconomic and essential dependence of water for the economic development of the municipalities studied, the PCA showed that the socioeconomic indexes of the municipalities located in the Sertão (Petrolina and Arcoverde) will present a higher correlation with the indices of temperature and Insulation, in the Agreste and Zona da Mata (Garanhuns and Surubim) evaporation and temperatures, in the Litoral (Recife) precipitation and humidity. The PCA was also effective in allowing the removal or disposal of variables that presented low variability or were redundant because they were correlated with those of greater importance for the first two main components. Understanding the behavior of climate elements and their consequences on human activities is of fundamental importance in helping public policies to mitigate the adverse effects of environmental change.

Keywords: Data Modeling, Socioeconomic Diagnostics, Principal Component Analysis.

CAPÍTULO

1 - INTRODUÇÃO

No Estado de Pernambuco, assim como em toda a região do Nordeste do Brasil, existe uma grande interação entre o comportamento dos elementos clima e as atividades humanas, sejam elas econômicas, sociais, culturais, etc. A vulnerabilidade dessas atividades varia de acordo com a localização geográfica, o tempo, as condições socioeconômicas e ambientais.

Sabendo que os efeitos da variabilidade, e possível mudança climática, atinge de forma diferenciada cada localidade, a partir de suas características físico-naturais e humanas, o monitoramento climático de cada região representa uma ação decisiva no enfrentamento aos efeitos adversos da mudança climática global, uma vez que considera as particularidades de cada lugar, contribuindo no âmbito do planejamento e gestão do território, na elaboração de políticas específicas adequadas a cada local (BARBOZA, et al., 2016).

De acordo com Silva (2013, p. 18), “as condições climáticas de determinada região ou localidade é constituído pelos seus fatores (massas de ar, relevo, latitude, maritimidade, etc.) e conjunto de elementos climáticos ou variáveis meteorológicas¹” (temperatura, precipitação pluviométrica, umidade do ar, etc.), que exerce influência nos aspectos socioeconômicos deste espaço. As condições climáticas materializadas sobre uma determinada porção da superfície terrestre resultam de uma destacada interação entre elementos e fatores climáticos. Estes agem contraditoriamente, às vezes, ou em ações solitárias que resultam no estabelecimento de um leque bastante amplo de tipos climáticos, nas diversas escalas geográficas (JATOBÁ, 2012).

Estudos recentes mostraram que em Pernambuco a variabilidade dos elementos climáticos vem influenciando os aspectos ambientais e socioeconômicos de diversas maneiras. Araújo et al.(2013) observaram os efeitos da mortalidade de bovinos no Agreste pernambucano em decorrência da estiagem prolongada no período 2010-2013. Visando o planejamento agropecuário e controle de doenças e pragas nas lavouras em Pernambuco, Melo, Medeiros e Souza (2015) estudaram a variação média mensal e anual da umidade relativa do ar em 11 municípios desse estado, mostrando ainda a

¹ Embora ambos os termos sejam sinônimos, o termo “*Variável Meteorológica*” é mais empregada pela meteorologia enquanto o termo “*Elementos Climáticos*” é usado pela geografia. Neste trabalho será usado o termo Elementos Climáticos, devido ao mestrando ter formação em geografia.

importância desse elemento climáticos para as atividades econômicas. Lacerda et al. (2016) evidenciaram que os cenários locais de mudanças climáticas para o futuro confirmaram as atuais tendências de aumento de temperatura do ar no sertão pernambucano, o que acarretaria uma maior vulnerabilidades não só das atividades agropecuárias mas também nos recursos hídricos dessa região.

Diante disso, por meio do processo de modelagem entre dados climáticos e socioeconômicos, buscou-se obter um cenário real de como os elementos do clima (Precipitação Pluviométrica, Temperatura do Ar, Umidade Relativa, Insolação e Evaporação) influenciam nos setores socioeconômicos (Produto Interno Bruto - PIB e o Índice de Desenvolvimento Municipal- IDM) dos municípios que melhor representassem o Estado de Pernambuco, levando em consideração também suas características geoambientais, Petrolina e Arcoverde – Sertão, Garanhuns – Agreste, Surubim – Zona da Mata e Recife – Litoral. A escolha dos municípios estudados também teve como referencial o estudo das regiões de influência das cidades de Pernambuco realizado pela Agência Estadual de Planejamento e Pesquisa De Pernambuco (CONDEPE/FIDEM), que caracteriza as áreas de concentração econômico-funcional, identificando suas influências regionais e o papel de cada cidade, em sua rede urbana.

Utilizando como ferramenta de modelagem a Análise de Componentes Principais (ACP) foi possível relacionar os elementos pesquisados, possibilitando, entre outros resultados, a redução do número de variáveis com o mínimo de perda de informações e com isso definir a importância dos elementos do clima no desenvolvimento socioeconômico dos municípios estudados. Apesar da relevância do tema, as pesquisas são limitadas, tanto devido à falta de uma base de dados climáticos sem falhas (dados inexistentes) como pela indisponibilidade de dados socioeconômicos atualizados.

Desta forma, a importância em desenvolver estudos sobre tal problemática surge, não só da necessidade de conhecer os reais efeitos os elementos climáticos sobre a sociedade, mas também como a promoção de pesquisa científica que possam ampliar o conhecimento da sociedade sobre as vulnerabilidades dos municípios e do Estado às alterações ambientais e sua necessidade de adaptação, também previsto pelo Plano Estadual de Mudanças Climáticas de Pernambuco.

2 – OBJETIVOS

2.1 - Objetivo Geral

Modelar estatisticamente dados climáticos de alguns municípios, localizados em regiões que representam estrategicamente o Estado de Pernambuco, para definir a importância dos elementos do clima no desenvolvimento socioeconômico.

2.2 - Objetivos Específicos

- I. Definir municípios de interesse para a realização de estudos socioeconômicos.
- II. Coletar dados climáticos e socioeconômicos atualizados dos municípios a serem estudados.
- III. Aplicar técnicas de análise de estatística multivariada (Análise de Componentes Principais - ACP) aos dados coletados.
- IV. Identificar as relações entre os elementos do clima e índices socioeconômicos desses municípios.
- V. Elaborar diagnósticos socioeconômicos dos municípios com base nos resultados da ACP.

3. REVISÃO DA LITERATURA

3.1 Interações Entre os Elementos do Clima e as Atividades Humanas

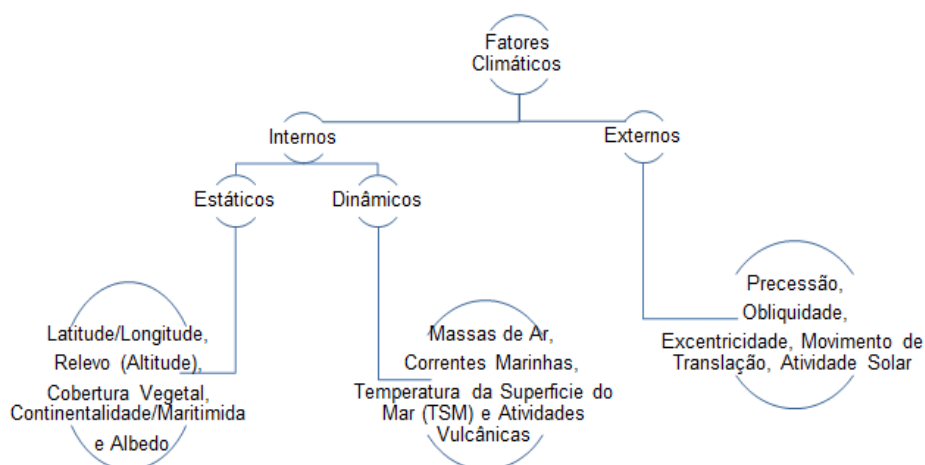
Aconteceu ao longo de toda a história da humanidade uma permanente interação entre as sociedades e as condições climáticas onde estas se desenvolviam. O entendimento dessas condições é fundamental tanto para compreensão de sua dinâmica natural quanto para o desenvolvimento dos diversos setores econômicos que dela dependem, tais como: agropecuária, saúde, economia, produção de energia, o abastecimento de água, etc., ou seja, sob várias formas o clima influencia as atividades humanas (SILVA, 2013).

O clima é um dos importantes componentes do ambiente natural. Ele afeta os processos geomorfológicos, os da formação dos solos além do crescimento e desenvolvimento das plantas, entre outros. “Os organismos, incluindo o homem, são influenciados pelo clima, as principais bases da vida para a humanidade, principalmente o ar, a água, o alimento e o abrigo, estão na dependência das condições do clima” (AYOADE, 2007).

Os fatores climáticos correspondem àquelas características geográficas diversificadoras das paisagens, são os aspectos que determinam o clima, que o condicionam e a partir dos quais as condições climáticas são originadas. Rigorosamente, uma distinção entre fator e elemento é, em muitos casos, artificial, por exemplo, a radiação pode ser um fator que promove variações das condições atmosféricas, mas considerando-a isoladamente é também um elemento climático, dependente da latitude, altitude, época do ano (VIEIRA, 2009).

Os principais fatores geográficos do clima e suas configurações estão descritas na figura 1. Os fatores climáticos podem ser de origem externas ou internas, que podem apresentar comportamentos dinâmicos ou estáticos.

Figura 1. Principais fatores climáticos



Fonte: O autor, 2017.

Em relação os elementos climáticos, Mendonça (2009, p.41) diz que eles “comunicam-se ao meio atmosférico com suas propriedades, características e peculiaridades”. Eles variam no tempo e no espaço e são influenciados pelos fatores climáticos, a este se juntam os aspectos dinâmicos do oceano e da atmosfera, como as correntes oceânicas, e as massas de ar, que atuando juntas vão qualificar os tipos de clima².

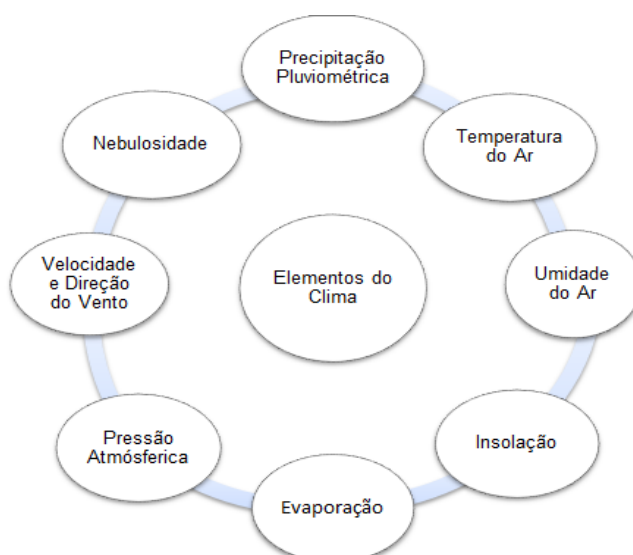
Mendonça (2009, p.41) explana ainda que

Embora seja habitual considerar os elementos à parte dos fatores climáticos, não se deve tomar com rigidez essa divisão, uma vez que os primeiros agem entre si de forma significativa, podendo, eventualmente um elemento ser ativo no controle do outro, como por exemplo, a temperatura condicionando a variação da umidade relativa e influenciando os campos barométricos. [...] e necessário examinar os princípios básicos que regem a ação dos fatores sobre os elementos e a relação dos fatores entre si.

² Clima, num sentido restrito é geralmente definido como 'tempo meteorológico médio', ou mais precisamente, como a descrição estatística de quantidades relevantes de mudanças do tempo meteorológico num período de tempo, que vai de meses a milhões de anos. O período clássico é de 30 anos, definido pela Organização Mundial de Meteorologia (OMM). Essas quantidades são geralmente variações de superfície como temperatura, precipitação e vento. O clima num sentido mais amplo e didático é o conjunto de elementos meteorológicos que caracterizam o estado médio da atmosfera num determinado ponto sobre a superfície terrestre." (CPTEC, 2017).

Os elementos climáticos mais estudados são: Precipitação pluviométrica, Temperatura do ar, Umidade do ar, Insolação, Evaporação, Pressão atmosférica, velocidade e direção do vento, nebulosidade. A figura 2 mostra os elementos do clima e a comunicação entre eles.

Figura 2. Principais elementos do clima



Fonte: O autor, 2017

3.1.2 Influência nas Atividades Econômicas em Pernambuco

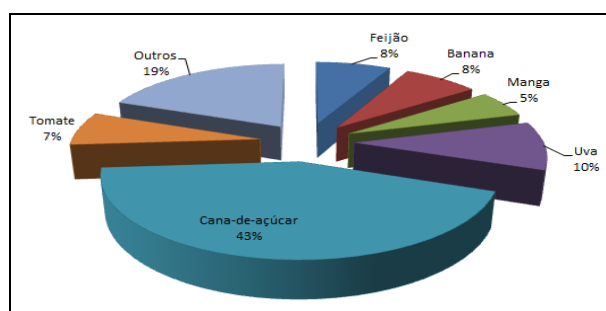
O comportamento dos elementos do clima e a sua variabilidade é um fator determinante para economia o desenvolvimento de uma região, país, estado, etc., estes desenvolvimento é mensurado através das atividades nos setores primário, secundário e terciário de acordo com os produtos produzidos, modos de produção e recursos utilizados. Estes setores econômicos podem mostrar o grau de desenvolvimento econômico de cada lugar (SILVA, 2013).

O estado de Pernambuco destaca-se por possuir atividades econômicas diversificadas, ressaltando-se as atividades ligadas aos serviços e à indústria de transformação, além da administração pública, defesa e seguridade social. A economia pernambucana, nos últimos anos, vem passando por um intenso processo de modernização e diversificação de sua base produtiva, além do intenso aporte de capital público/privado (RECIFE, 2011).

Por outro lado, diversas atividades econômicas estão sendo afetadas pela variabilidade do comportamento dos elementos do clima no estado nos últimos anos. As atividades do agronegócio, como produtividade agrícola e pecuária dependem significativamente das condições climáticas. Esta dependência é importante durante o ciclo de vida de determinadas culturas, sendo a responsável pela alternância das produções anuais destas. Os parâmetros climáticos exercem influência sobre todos os estágios da cadeia de produção agrícola, influenciando a preparação da terra, semeadura, crescimento dos cultivos, colheita armazenamento, transporte e comercialização (AYOADE, 2007).

De acordo com Brasil (2017), um dos parâmetros para mensurar o desenvolvimento econômico é o Valor Adicionado Bruto (VAB), que é a contribuição ao Produto Interno Bruto (PIB)³ pelas diversas atividades econômicas, obtida pela diferença entre o valor de produção e o consumo intermediário absorvido por essas atividades. Em Pernambuco a agropecuária responde por cerca de 5% do PIB, tratando-se da atividade mais tradicional do Estado, sujeita a frequentes oscilações de preços e de mercado, mas que se mantém como a cultura mais relevante e também dependente das condições climáticas. Na figura 3, observam-se os principais produtos agrícolas no Estado de Pernambuco. A cana-de-açúcar possui pouco mais de 43%, destacam-se ainda as culturas do feijão, da banana e do tomate com participações similares em torno de 7% da produção agrícola bruta (RECIFE, 2013).

Figura 3. Participação dos principais produtos agrícolas em Pernambuco 2010



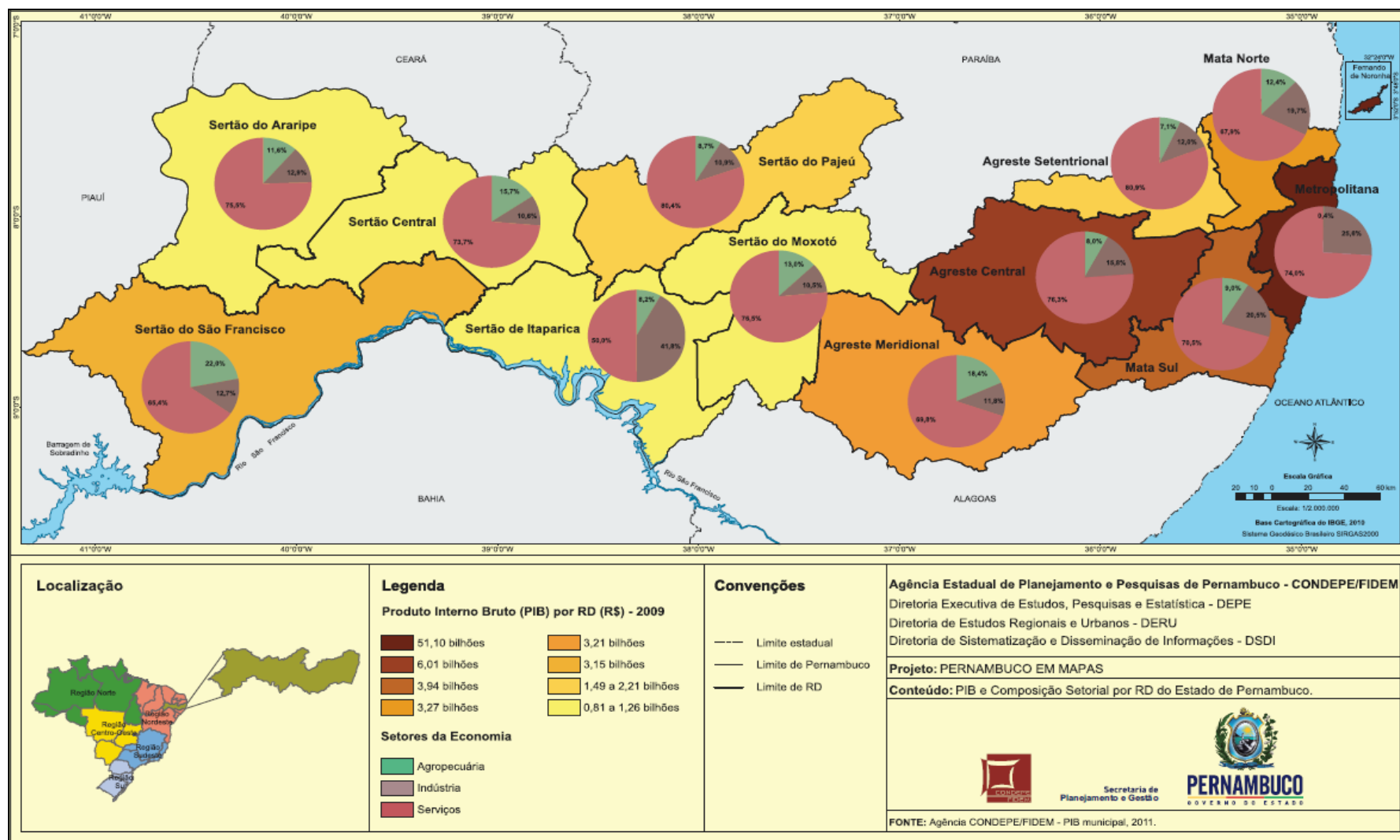
Fonte: Organizado pelo autor, com base em dados do RECIFE, 2013

³ O **Produto Interno Bruto - PIB** é um indicador de grande importância para a elaboração de políticas públicas e como fonte de informações para estudos e pesquisas. No Brasil, o cálculo do PIB é oficialmente de responsabilidade do Instituto Brasileiro de Geografia e Estatística (IBGE). Em Pernambuco, o cálculo do respectivo indicador é de competência da Agência Estadual de Planejamento e Pesquisas de Pernambuco (CONDEPE/FIDEM), órgão oficial de estatística do Estado, em parceria com o IBGE (BRASIL, 2017).

Ainda segundo Recife (2013), o setor indústria contribuía com 22,0% do VAB do Estado, destacando-se as atividades da construção e da indústria de transformação. Assim como no setor agropecuário, a produção na industrial também é afetada pelas condições de tempo e a variabilidade do clima, atingindo-o de diversas maneiras, no alimentício, Lucas (2015) analisou os efeitos decorrentes de chuvas extremas sobre o retorno das ações de algumas das principais empresas do setor alimentício brasileiro. No setor mecânico, Caldeira (2016) estudou o aproveitamento da precipitação pluviométrica para consumo não potável em uma indústria mecânica localizada no município de Belo Horizonte - MG.

No setor de bens e serviços, que participa com 73,2% do VAB de Pernambuco. A influência dos elementos do clima se dá quando esses apresentam-se de forma extrema, como em episódios de chuvas intensas na região do Litoral e zona da Mata, e dias de baixa umidade do ar no sertão do estado. Neste setor, ressalta-se o comércio e serviços de manutenção e reparação, além da administração, saúde e educação públicas e seguridade social, estas atividades contribuem, respectivamente, com 13,9% e 24,7% do VAB do estado (figura 4).

Figura 4. Composição setorial da economia de Pernambuco por região em 2011.



Fonte: CONDEPE/FIDEM, 2011

3.1.3 Implicações na Saúde

A saúde humana é fortemente influenciada pelo clima. As condições térmicas, de dispersão (ventos e poluição) e de umidade do ar exercem destacada influência sobre a manifestação de muitas doenças, epidemias e endemias humanas. Segundo Ayoade (2007), apud Silva (2013, p. 18), “a saúde humana, a energia e o conforto são mais afetados pelo clima do que por qualquer outro elemento do meio ambiente”.

O geógrafo francês Max Sorre (1880-1962) em sua obra, *“A adaptação ao meio climático e bio-social”* (1984), faz uma análise da interação entre o meio e a saúde humana enfatizando a influência daquele nesta, particularmente das condições climáticas, destacando em sua abordagem o papel dos elementos do clima na manifestação de variadas doenças, ele introduziu, ao correlacionar a ocorrência de determinadas doenças a tipos climáticos específicos, o conceito de complexo patogênico, ou “complexos patogênicos”, “cujo número e variedade são infinitos”, situando seu conhecimento como “a base de toda a geografia médica”. (SORRE, 1984).

Sobre a relação entre as condições climáticas e a saúde do homem, Sett e Ribeiro (2011, p. 42) falam que “as relações entre clima e saúde humana são complexas porque dependem da intensidade e duração da mudança de tempo (grau de contraste) e sensibilidade do receptor, sendo que esses fatores mudam continuamente de local, indivíduos e populações”.

Os efeitos meteoro-trópicos⁴ tendem a ocorrer de forma sincrônica em um grande território, sob a influência das mesmas condições de tempo e afetar significativamente a população local. Eles podem ser específicos e inespecíficos. Os efeitos meteoro-trópicos específicos do clima sobre a saúde humana ocorrem quando há uma ação direta de elementos meteorológicos em uma sequência de ações inter-relacionadas, por exemplo, chuvas torrenciais, inundações, contaminação das águas que resultam em surtos de doenças diarreicas e aumento de populações de vírus e vetores, provocando surtos de doenças transmissíveis. (SETTI; RIBEIRO, 2011).

⁴ “É o conjunto de ações complexas e diversificadas que a variabilidade do tempo tem sobre a saúde das pessoas”. (SETTI; RIBEIRO, 2011, p. 42).

Tabela 1. Efeitos físico-psicológicos dos elementos do clima no homem

Elementos Climáticos	Condições Limitantes	Manifestações Fisiológicas
Altitude (Pressão atmosférica)	Limite Máximo: 8.000m.	- Mal-das-montanhas (dor de cabeça, fadiga, alteração sensorial, depressão intelectual, indiferença, sono, descoordenação de movimentos, perda de memória). - Redução faculdades físicas e mentais. - Tristeza, apatia.
Radiação (Associada à luminosidade)	60° e 70° Latitude.	- Alta radiação/luminosidade: esgotamento nervoso, perturbações mentais, irritação, síndrome físico-psíquica "golpe de sol" (sunstroke), euforia. - Baixa radiação/luminosidade: deficiências orgânicas, raquitismo, depressão, debilidade mental.
Higrotermia	Limite Variável. Ótimo fisiológico para raça branca: 15° - 16° C/60%UR	- Diminuição da capacidade respiratória (para europeus nos trópicos). - Hiperpnéia térmica (entre negros). - Cansaço e esgotamento (entre os brancos).
Vento e Eletricidade Atmosférica	---	- Morbidez, cansaço e abatimento. - Debilidade do tonus nervoso, depressão, hipersensibilidade, irritabilidade. - Desidratação, dessecação do aparelho tegumentar. - Excitação nervosa, alucinações, delírio. - Palpitações, dispnéia, dores de cabeça, nevralgia.

Fonte: SORRE, 1984

Os efeitos específicos do clima sobre a saúde humana ocorrem quando há uma ação direta de elementos do clima em uma sequência de ações inter-relacionadas, por exemplo, chuvas torrenciais, inundações, contaminação das águas que resultam em surtos de doenças diarréicas e aumento de populações de vírus e vetores, provocando surtos de doenças transmissíveis. Além de doenças respiratórias. Em Pernambuco a incidência de doenças correlacionadas direta ou indiretamente com as condições climáticas são numerosas, tanto as que afeta o sistema respiratório como a tuberculose como as provocadas por vetores, como a dengue, zika e chikungunya, (SILVA, 2013).

Os casos de doença transmitida através da picada do mosquito *Aedes aegypti* já há bastante tempo vem afetando a vida da população do Estado de Pernambuco, entretanto nos últimos anos o aumento dos casos vem chamado cada vez mais à atenção devida não só a variabilidade climática, mas à falta de políticas públicas que minimizem a proliferação do mosquito. Em 2015, de acordo com a Secretaria de Saúde do Estado de Pernambuco, foram notificados 146.089 casos, ocorrendo 31 óbitos por causa da dengue.

Como a manifestação do clima se dá na dimensão dos episódios concretos, os eventos extremos, ou excepcionais, assumem destacada importância nas análises relacionadas à saúde do homem, uma vez que estes episódios são os que em realidade, repercutem no espaço/lugar das pessoas. Toledo Filho et. al. (2010) investigaram a relação entre o clima e o comportamento humano em Maceió-AL, mostrando que existe correlação entre o comportamento dos elementos do clima e o humano principalmente na incidência de transtornos de humor, agressão, irritabilidade e depressão.

Dessa forma, em consequência de desastres naturais também acontecem certos transtornos psicológicos que acometem àqueles expostos a uma situação ou evento estressante (de curta ou longa duração), de natureza excepcionalmente ameaçadora ou catastrófica, como no caso do Transtorno de Estresse Pós-Traumático (TEPT). Os sintomas relativos a essa síndrome correspondem a revivência do trauma, através de lembranças ou pesadelos; esquiva, através do distanciamento de pensamentos, locais e pessoas e a hiperstimulação autonômica, através da insônia, irritabilidade e hipervigilância. (FIGUEIRA E MENDLOWICZ, 2003)

Os acometimentos de transtornos psicológicos dependem da interação entre o indivíduo, ambiente e o contexto presente, não se podendo afirmar necessariamente que todas as pessoas que vivenciaram uma situação de desastre, qualquer que seja sua natureza, desenvolvam algum tipo de transtorno psicológico, nessa discussão, enquadra-se o conceito de resiliência e coping ⁵ (CARVALHO et.al., 2012).

⁵ O conceito de coping tem sido descrito como um conjunto das estratégias utilizadas pelas pessoas para adaptarem-se a circunstâncias adversas ou estressantes, essas estratégias atuam como reforços cognitivos e comportamentais que levam o indivíduo a lidar com situações de estresse, têm se constituído em objeto de estudo da psicologia social, clínica e da personalidade, encontrando-se fortemente atrelado ao estudo das diferenças individuais. (MIGUEL E BUENO, 2007).

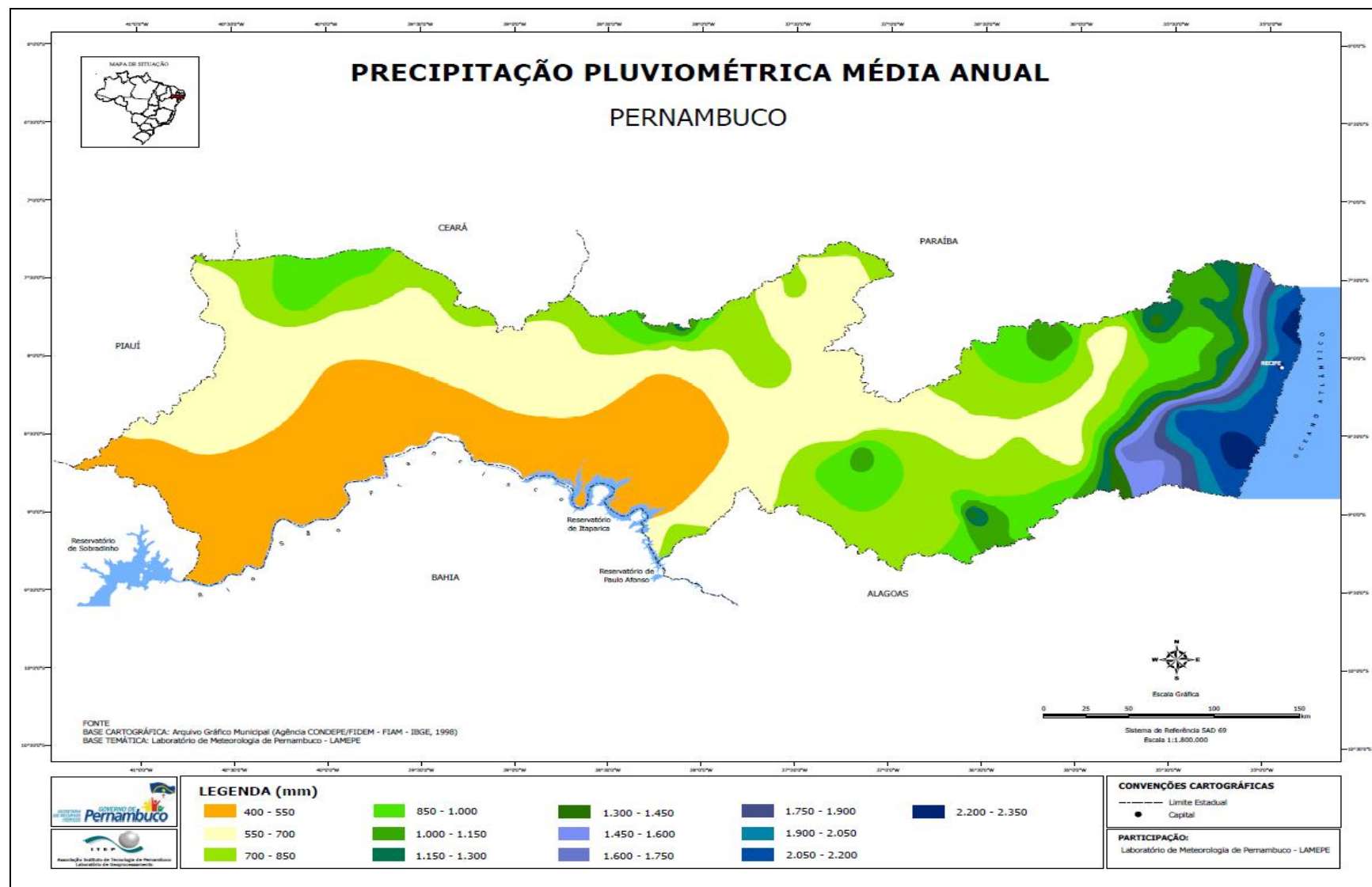
3.1.4 Abastecimento de água

A política Nacional de Recursos Hídricos estabelecida pela Lei nº 9433/97 (Lei das Águas), tem como objetivo principal assegurar à atual e às futuras gerações a necessária disponibilidade de água, em padrões de qualidade adequados aos respectivos usos, ao mesmo tempo em que busca a prevenção e a defesa contra eventos hidrológicos críticos (secas e enchentes) e o desenvolvimento sustentável, por meio da utilização racional e integrada dos recursos hídricos (BRASIL, 1997).

De acordo com PERNAMBUCO (2006), “dos 185 municípios de Pernambuco 153 são abastecidas por águas provenientes exclusivamente de mananciais superficiais, reservatórios estaduais e federais”, o que os deixam mais vulneráveis as condições climáticas, principalmente a variabilidade dos totais pluviométricos. A sobrevivência da população no Semiárido, que em Pernambuco representa aproximadamente 70% do território, e o seu potencial de produção vinculam-se diretamente à disponibilidade hídrica. Dentre os elementos meteorológicos, as precipitações pluviométricas são as que apresentam alta variabilidade no tempo e no espaço. Além da irregularidade das chuvas ao longo do ano, também há uma irregularidade interanual. Há anos em que chove acima da média, enquanto que em outros as chuvas são escassas e mais concentradas, de forma que a estiagem fica mais severa. De acordo com Ab’Sáber (1999) a periodicidade em que ocorrem as secas mais severas podem ser estabelecidas em um intervalo de cerca de 9 a 12 anos, no entanto esse intervalo não é preciso. Mas, conforme ponderou Celso Furtado, “existem muitos recursos para previsão, e sabe-se mais ou menos a probabilidade da ocorrência da seca” (TAVARES; ANDRADE; PEREIRA, 1998).

A partir do regime de chuvas (figura 5) é possível traçar um cenário para a disponibilidade de água no solo do Semiárido. Ao sobrevir a estiagem, os rios secam rapidamente e os lençóis subterrâneos se aprofundam e mínguam, inviabilizando a utilização de suas águas. As temperaturas registradas no local são altas, contribuindo para a existência de um desequilíbrio entre a quantidade de água presente e a taxa de evapotranspiração. Assim, as águas acumuladas tendem a evaporar com rapidez (AB’SÁBER, 1999).

Figura 5. Regime Pluviométrico de Pernambuco



Fonte: PERNAMBUCO, 2006

3.2 Análise de Componentes Principais (ACP)

A análise de componentes principais (ACP) é uma técnica estatística multivariada de modelagem da estrutura de covariância. Ela está associada à ideia de redução de massa de dados, com menor perda possível da informação, contudo é importante ter uma visão conjunta de todas ou quase todas as técnicas da estatística multivariada para resolver a maioria dos problemas práticos. Procura-se redistribuir a variação observada nos eixos originais de forma a se obter um conjunto de eixos ortogonais não correlacionados (HONGYU, 2015).

A técnica foi inicialmente descrita pelo matemático Karl Pearson (1857-1936) em 1901. A ACP transforma linearmente um conjunto original de variáveis, inicialmente correlacionadas entre si, num conjunto substancialmente menor de variáveis não correlacionadas que contém a maior parte da informação do conjunto original (HONGYU; SANDANIELO; JUNIOR, 2015).

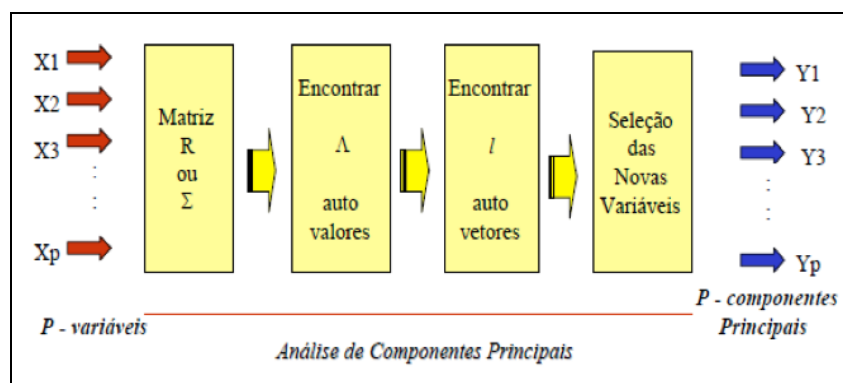
Para Lyra et. al (2010, p. 1) o objetivo principal da ACP é o de “explicar a estrutura da variância e covariância de um vetor aleatório, composto de p-variáveis aleatórias, por meio de combinações lineares das variáveis originais”. Essas combinações lineares são chamadas de componentes principais e são não correlacionadas entre si. Ainda de acordo com Hongyu, Sandanielo e Junior, (2015) “esta técnica pode ser utilizada para geração de índices e agrupamento de indivíduos”. A análise agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou seja, segundo seu comportamento dentro da população, representado pela variação do conjunto de características que define o indivíduo, ou seja, a técnica agrupa os indivíduos de uma população segundo a variação de suas características (HONGYU, 2015).

Em outras palavras, a ACP transforma um conjunto de variáveis originais em outro conjunto de variáveis de mesma dimensão denominadas de componentes principais. Os componentes principais apresentam propriedades importantes: cada componente principal é uma combinação linear de todas as variáveis originais, são independentes entre si e estimados com o propósito de reter, em ordem de estimação, o máximo de informação, em termos da variação total contida nos dados (FERREIRA, 2011).

De acordo com Tique (2015), a ACP consiste basicamente em transformar variáveis correlacionadas em grupos correlacionados, designados por componentes principais (CPs), sem, no entanto, alterar as características dos dados iniciais (originais). A primeira componente principal (1° CP) é a combinação linear normalizada das variáveis originais, que representa a máxima fração ou porcentagem da variância; a segunda componente principal (2° CP) é aquela combinação linear que não está correlacionada com o 1° CP e representa a maior parte da variância restante e assim por diante. Essas combinações são determinadas através de manipulações algébricas na matriz de covariâncias das p variáveis. Cada CP captura uma direção de variabilidade do conjunto de dados originais. As direções capturadas por cada CP são ortogonais entre si. Isso equivale a dizer que os CPs são variáveis aleatórias independentes (REGAZZI, 2001).

Para a determinação das componentes principais, é necessário calcular a matriz de variância-covariância (Σ) ou a matriz de correlação (R), encontrar os autovalores e os autovetores e, por fim, escrever as combinações lineares que serão as novas variáveis, denominadas de componentes principais (SOUZA, 2000). A figura 6 mostra o esquema da aplicação da ACP.

Figura 6 - Esquema da aplicação da Análise de Componentes Principais



Fonte: Souza, 2000

De acordo com Hongyu (2015, p. 3) A ACP tem como principais vantagens:

Retirar a multicolinearidade das variáveis, pois permite transformar um conjunto de variáveis originais inter-correlacionadas em um novo conjunto de variáveis não correlacionadas (componentes principais).

Além disso, reduz muitas variáveis a eixos que representam algumas variáveis, sendo estes eixos perpendiculares (ortogonais) explicando a variação dos dados de forma decrescente e independente.

Segundo Varella (2008), apesar das técnicas de análise multivariada terem sido desenvolvidas para resolver problemas específicos, principalmente de Biologia e Psicologia, podem ser também utilizadas para resolver outros tipos de problemas em diversas áreas do conhecimento. A análise de componentes principais é a técnica mais conhecida, contudo é importante ter uma visão conjunta de todas ou quase todas as técnicas da estatística multivariada para resolver a maioria dos problemas práticos.

As principais desvantagens da ACP, segundo Regazzi, (2000, p. 12) são

A sensibilidade a **outliers**, não recomendada quando se tem duplas ausências (muitos zeros na matriz) e dados ausentes. A ACP também não é recomendada quando se tem mais variáveis do que unidades amostrais. Ao reduzir o número de variáveis, há perda da informação de variabilidade das variáveis originais. Mas que a parte explicada seja o padrão de resposta e a outra parte o ruído, ou seja, erro de medida e redundância. A ACP nem sempre funciona (às vezes mesmo com a redução ainda continua grande). É o caso de variáveis originais pouco correlacionadas, com o caso extremo da $R = I$, os componentes principais são as próprias variáveis originais.

3.2.1 Matrizes, autovalores e autovetores.

A obtenção dos componentes principais é realizada por meio da diagonalização de matrizes simétricas positivas pré-definidas. Então, podem-se calcular os componentes principais e utilizá-los em diferentes aplicações nas mais variadas áreas científicas. Esta metodologia é em função da existência de inúmeros programas capazes de realizar cálculos matriciais para diagonalizar uma matriz simétrica positiva semi-definida. Dentre os programas comerciais podemos citar o *Unscrambler*, o *Statistica*, o *Minitab*, o *SAS*, o *SPSS*, o *Excel* (extensão Action Stat), e o *Pirouette*.

Também não é difícil encontrar, ou mesmo escrever, rotinas para ACP nos softwares *R*, *Scilab*, *Matlab* e para o programa livre *GNU Octave*. (LYRA et al. 2010).

A matriz de transformação utilizada para o cálculo da PCA consiste em uma matriz cujas linhas são os autovetores da matriz de correlação e covariância estimada dos dados. A matriz de covariância Σ é uma matriz simétrica e definida positiva, que possui informação sobre as variâncias em todos os eixos onde os dados estão distribuídos (VASCONCELOS, 2016).

Sejam as variáveis $X_1, X_2, X_3, \dots, X_p$. A matriz de dados é de ordem ' $n \times p$ ' e normalmente denominada de matriz de correlação ' R '.

$$R = \begin{bmatrix} x_{11} & x_{12} & x_{13} & \dots & x_{1p} \\ x_{21} & x_{22} & x_{23} & \dots & x_{2p} \\ x_{31} & x_{32} & x_{33} & \dots & x_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & x_{n3} & \dots & x_{np} \end{bmatrix}$$

A estrutura de interdependência entre as variáveis da matriz de dados é representada pela matriz de covariância ' Σ ' ou pela matriz de correlação ' R '.

O entendimento dessa estrutura através das variáveis $X_1, X_2, X_3, \dots, X_p$, pode ser na prática algo difícil, assim, o objetivo da ACP é transformar essa estrutura complicada, representada pelas variáveis $X_1, X_2, X_3, \dots, X_p$, em uma outra estrutura representada pelas variáveis $Y_1, Y_2, Y_3, \dots, Y_p$ não correlacionadas e com variâncias ordenadas, para que seja possível comparar os indivíduos usando apenas as variáveis Y is que apresentam maior variância. A solução é dada a partir da matriz de covariância Σ ou da matriz de correlação R (VARELLA, 2008).

A partir da matriz X de dados de ordem ' $n \times p$ ' podemos fazer uma estimativa da matriz de covariância Σ da população π que representaremos por S . A matriz S é simétrica e de ordem ' $p \times p$ '.

$$\Sigma = \begin{bmatrix} \text{Var}(x_1) & \text{Cov}(x_1 x_2) & \text{Cov}(x_1 x_3) & \dots & \text{Cov}(x_1 x_p) \\ \text{Cov}(x_2 x_1) & \text{Var}(x_2) & \text{Cov}(x_2 x_3) & \dots & \text{Cov}(x_2 x_p) \\ \text{Cov}(x_3 x_1) & \text{Cov}(x_3 x_2) & \text{Var}(x_3) & \dots & \text{Cov}(x_3 x_p) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(x_p x_1) & \text{Cov}(x_p x_2) & \text{Cov}(x_p x_3) & \dots & \text{Var}(x_p) \end{bmatrix}$$

Para a obtenção dos componentes principais de uma forma geral, seja um conjunto de p variáveis $X_1, X_2, \dots, X_p$ com médias $\mu_1, \mu_2, \dots, \mu_p$ e variância $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_p^2$, respectivamente. Estas variáveis não são independentes e, portanto, possuem covariância entre a i -ésima e k -ésima variável definida por σ_{ik} , para $i \neq k = 1, 2, \dots, p$. Então as p variáveis podem ser expressas na forma vetorial por $\mathbf{X} = [X_1, X_2, \dots, X_p]'$, com vetor de médias $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]'$ e matriz de covariância $\boldsymbol{\Sigma}$.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \dots & \sigma_{1p}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \dots & \sigma_{nn}^2 \end{bmatrix}$$

Encontram-se os pares de autovalores e autovetores $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$, em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, associados a $\boldsymbol{\Sigma}$ e então o i -ésimo componente principal é definido por (JOHNSON; WICHERN, 2007).

$$\mathbf{Z}_i = \mathbf{e}_i' \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p \quad (1)$$

A variável $\mathbf{Z}_i$, é uma variável latente, ou seja, não é mensurada a partir do experimento ou levantamento amostral (HAIR, et. al, 2005). O objetivo é determiná-la a partir das p variáveis contidas no vetor $\mathbf{X}$. A ideia é projetar os pontos coordenados originais em um plano maximizando a distância entre eles, o que equivale a maximizar a variabilidade da variável latente $\mathbf{Z}_i$. A variância de $\mathbf{Z}_i$ é dada por

$$\text{Var}(\mathbf{Z}_i) = \text{Var}(\mathbf{e}_i' \mathbf{X}) = \mathbf{e}_i' \text{Var}(\mathbf{X}) \mathbf{e}_i = \mathbf{e}_i' \boldsymbol{\Sigma} \mathbf{e}_i \quad (i = 1, \dots, p.) \quad (2)$$

Utilizando a decomposição espectral da matriz $\boldsymbol{\Sigma}$, dada por $\boldsymbol{\Sigma} = \mathbf{P} \boldsymbol{\Lambda} \mathbf{P}'$, em que $\mathbf{P}$ é a matriz composta pelos autovetores de $\boldsymbol{\Sigma}$ em suas colunas e $\boldsymbol{\Lambda}$, a matriz diagonal de autovalores de $\boldsymbol{\Sigma}$, então, tem-se que

$$\text{tr}(\boldsymbol{\Sigma}) = \text{tr}(\mathbf{P} \boldsymbol{\Lambda} \mathbf{P}') = \text{tr}(\boldsymbol{\Lambda} \mathbf{P}' \mathbf{P}) = \text{tr}(\boldsymbol{\Lambda} \mathbf{I}) = \text{tr}(\boldsymbol{\Lambda}) = \sum_{i=1}^p \lambda_i \quad (3)$$

e

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}$$

E $tr(\Sigma)$ é dada pela soma dos elementos da diagonal:

$$tr(\Sigma) = \sum_{i=1}^p \sigma_{ii} = \sum_{i=1}^p \lambda_i \quad (4)$$

Portanto, a variabilidade total contida nas variáveis originais é igual a variabilidade total contida nos componentes principais (JOHNSON; WICHERN, 2007).

A análise de componentes principais ainda tem como objetivo explicar a estrutura de variâncias e covariâncias de um conjunto de variáveis de interesse através de combinações lineares das variáveis $X_1, X_2, \dots, X_p$ que são correlacionadas entre si. Martins, Godoy e Lopes (2012, p. 34) comentam que “geometricamente, essas combinações lineares representam a seleção de um novo sistema de coordenadas obtidas através da rotação do sistema original, tendo como coordenadas os novos eixos $Y_1, Y_2, \dots, Y_p$.”

Geometricamente, os CPs podem ser representados por um sistema de eixos ortogonais com origem posicionada na média do conjunto de dados. O primeiro CP segue a direção de maior variabilidade dos dados. O segundo CP segue a segunda direção de maior variabilidade, ortogonal ao primeiro CP. Os demais CPs são posicionados no espaço p -dimensional seguindo a mesma lógica, a primeira e assim por diante (MARTINS; GODOY; LOPES, 2012).

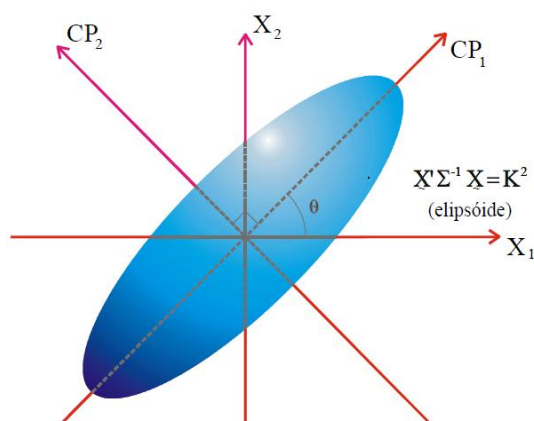
Suponha duas variáveis aleatórias X_1, X_2 , seguindo uma distribuição Normal bivariada $N_2(\mu, \Sigma)$, com vetor de médias dado por μ e matriz de covariâncias Σ . A função de densidade de $N_2(\mu, \Sigma)$ é dada pela equação da elipse centrada em μ :

$$X - \mu + \sum_{i=1}^2 (X - \mu)^T e_i e_i^T (X - \mu) = e^4 \quad (5)$$

Onde X denota o valor que contém as variáveis aleatórias X_1 e X_2 . Os eixos da elipse em (3) são dados por $\pm c (\lambda_i)^{1/2} e_i$, $i = 1, 2$ (MARTINS; GODOY; LOPES, 2012).

Os CPs $Y_1 = x^T e_1$ e $Y_2 = x^T e_2$ estão posicionados nas direções dos eixos da elipse de densidade constante, conforme ilustrado na Figura 7.

Figura 7 - Representação gráfica dos autovalores e autovetores



Fonte: Lopes, 2001

O primeiro componente corresponde ao maior eixo da elipsóide (CP_1) e o comprimento desse eixo é proporcional a $\sqrt{\lambda_1}$. O eixo de menor variância (CP_2) é perpendicular ao eixo maior. Esse eixo chama-se segundo componente principal e seu comprimento é proporcional a $\sqrt{\lambda_2}$. Assim, a análise dos componentes principais toma os eixos X_1 e X_2 e coloca-os na direção da maior variabilidade (MARTINS; GODOY; LOPES, 2012).

3.2.2 Seleção e Interpretação dos Componentes Principais

Ao realizar uma análise de componentes principais, é muito importante saber o significado de cada componente no estudo que está sendo realizado. A interpretação de uma componente principal é feita mediante o grau de importância, ou, ainda, a influência que cada variável tem sobre cada componente, sendo que esta importância é dada pela correlação entre cada variável j X e o componente i Y que estiver sendo interpretado (REGAZZI, 2001).

A contribuição de cada componente principal (Z_i) é expressa em porcentagem, e a explicação individual de cada componente pode ser calculada, por exemplo, para k -ésimo componente principal a proporção da explicação é dada por:

$$Ck = \frac{Var(Z_i)}{\sum_{i=1}^p Var(Z_i)} \cdot 100 = \frac{\lambda_i}{\sum_{i=1}^p \lambda_i} \cdot 100 = \frac{\lambda_i}{tr(S)} \cdot 100 \quad (6)$$

Pela proporção de explicação da variância total, que o modelo de k componentes principais é responsável, podemos determinar o número de componentes que deve-se reter. Em muitos casos, adota-se modelos que expliquem pelos menos 80% da variação total (JOHNSON; WICHERN, 2007). Em geral escolhe-se o componente principal de maior importância (o primeiro componente principal) como sendo aquele de maior variância ($\max \lambda_i$), que explique o máximo de variabilidade dos dados, o segundo componente de maior importância, o que apresenta a segunda maior variância e assim sucessivamente, até o componente principal de menor importância (MANLY, 1986).

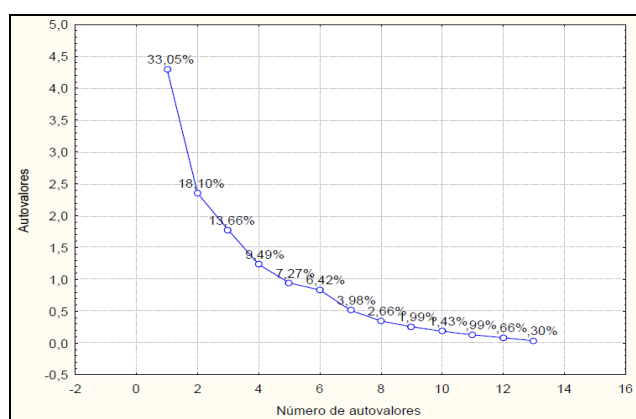
Por outro lado, os últimos componentes principais serão responsáveis por direções que não estão associadas a muita variabilidade. Em outras palavras, esses últimos componentes principais identificaram relações lineares entre as variáveis originais próximo de constante (FERREIRA, 2011).

A definição do número de componentes a serem utilizadas é feita por meio de dois critérios. O primeiro, denominado de método gráfico, representa graficamente a porcentagem de variação explicada pela componente nas ordenadas e os autovalores em ordem decrescente nas abscissas. Quando esta porcentagem diminui e a curva passa a ser praticamente paralela ao eixo das abscissas, exclui-se as componentes que restam, pois possuem pouca informação. Este critério, que considera as componentes anteriores ao ponto de inflexão da curva (HONGYU, 2015).

A figura 8 é um exemplo que mostra a seleção dos componentes principais através do método gráfico **Scree Plot**⁶, sendo que a porcentagem de variação explicada pela componente está no eixo das ordenadas, e os autovalores estão representados em ordem decrescente no eixo das abscissas.

⁶ Representa, graficamente, a porcentagem de variação explicada pela componente nas ordenadas e os autovalores, em ordem decrescente, nas abscissas, sugerido por CATTEL (1966), as comunalidades, a proporção de contribuição de cada variável factor loadings e outros valores de interesse (VICINI E, 2005).

Figura 8. Explicação da proporção de variação de cada CP



Fonte: Vicini (2005)

O segundo critério de seleção, e o critério desenvolvido por H.F. Kaiser em 1960, que consiste em incluir poucas componentes quando o número de variáveis originais é inferior a vinte e, em geral, utilizam-se aqueles componentes que conseguem sintetizar uma variância acumulada em torno de 70% (tabela 2). Além do uso na redução da dimensionalidade, a técnica de ACP pode ser utilizada como apoio à busca da variável de maior prevalência no sistema responsável, servindo-se do estudo dos coeficientes de correlação entre as componentes e às variáveis originais (CRUZ E TOPA, 2009).

Tabela 2. Tabela de critério de Kaiser para variancia acumulada na ACP

Valor	Grau da Variância Acumulada dos Dados
> 0,90	Ótimo
de 0,80 a 0,90	Boa
de 0,70 a 0,80	Razoável
de 0,60 a 0,70	Baixa
< 0,60	Inadequada

Fonte: Cruz e Topa, 2009

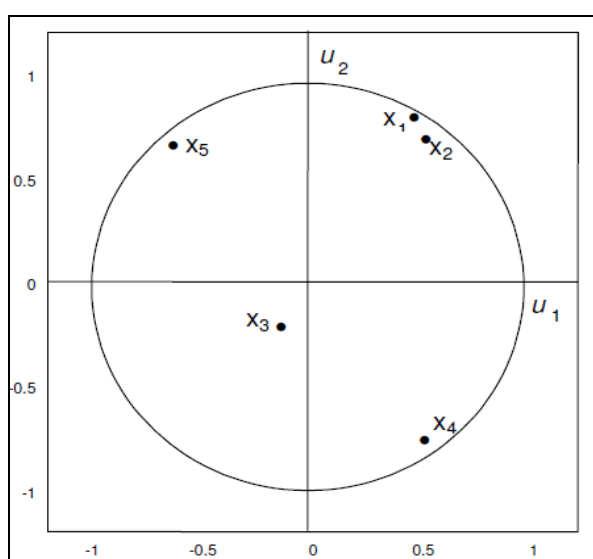
Em muitas áreas, os pesquisadores utilizam os primeiros e mais importantes componentes de uma análise para agrupar observação e itens de acordo com a representação em duas ou no máximo três dimensões. Um gráfico bidimensional pode ser obtido representando as informações das observações para duas variáveis em um diagrama, já para um gráfico **biplot** ou circulo de correlações. A ideia é de acrescentar

informações sobre várias variáveis em um único gráfico obtido pelos componentes principais (KLEFENS, 2009).

O gráfico biplot tem como objetivo a representação gráfica aproximada de uma matriz X de dados em outra $Y_{n \times p}$ de baixa dimensão, permitindo a visualização das relações e inter relações entre linhas e colunas de Y em um único gráfico (DIAS, 2008).

O biplot permite observar a posição de uma observação com relação à outra observação e a associação de cada uma das variáveis em relação a qualquer observação ou a outra variável. Assim, pode-se observar como as observações se agrupam e quais variáveis contribuem para sua posição dentro dessa representação. Deste modo, a análise das proximidades ou oposições entre variáveis é feita em termos de correlações. No exemplo mostrado na figura 9, estão representadas as projecções de 5 variáveis no plano U_1 U_2 bem como o círculo de correlação (JOHNSON e WICHERN, 1998).

Figura 9. Círculo de correlação



Fonte: KLEFENS, 2009

As variáveis x_1 , x_2 , x_4 e x_5 estão bem representadas neste plano, pois se encontram próximo da circunferência unitária: x_1 e x_2 estão fortemente correlacionadas entre si, mas são independentes das variáveis x_4 e x_5 , as quais, entre si apresentam uma correlação negativa forte. Quanto à variável x_3 , mal representada neste plano, nada se pode concluir.

4 - PROCEDIMENTOS METODOLÓGICOS

4.1 Seleção dos Municípios e Análise das Variáveis

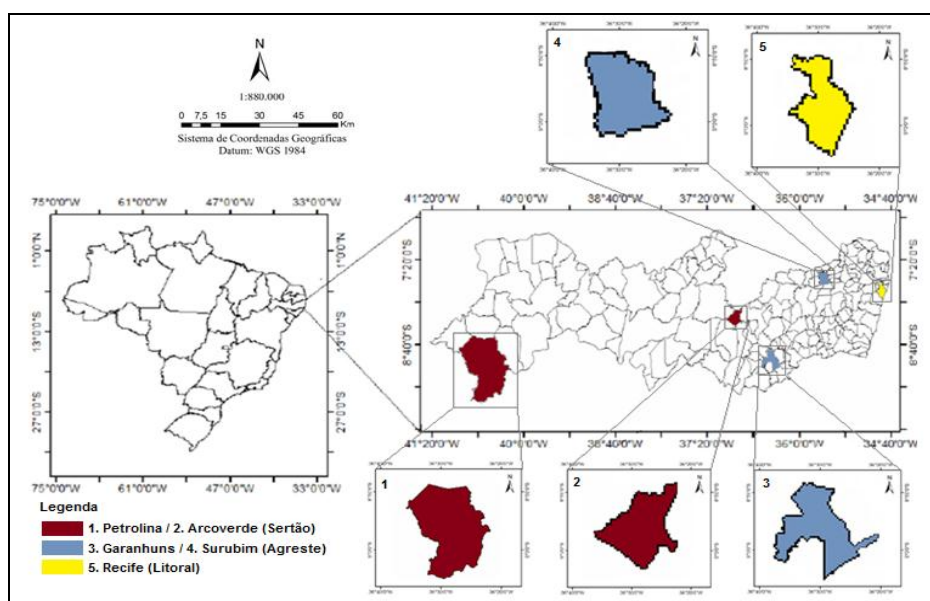
A seleção dos municípios estudados foi realizada considerando primeiramente sua localização fisiográfica (Sertão, Agreste, Zona da Mata e Litoral), a fim de escolher no mínimo um município que representasse cada região natural do Estado, a partir de suas características físico-naturais e humanas. O Estado de Pernambuco é composto por 185 municípios (incluindo a Ilha de Fernando de Noronha), e em decorrência da sua formação territorial esses municípios apresentam diversas particularidades socioeconômicas que os caracteriza tanto em razão da sua localização geográfica como das atividades econômicas desenvolvidas ao longo da sua história (ANDRADE, 2003).

Os dados climáticos foram obtidos por meio do Banco de Dados Meteorológicos para Ensino e Pesquisa do Instituto Nacional de Meteorologia (BDMEP/INMET). Foram empregados os elementos climáticos que segundo Mendonça (2009) “mais comumente caracteriza a atmosfera geográfica”, são eles: Precipitação pluviométrica (Total anual, medido em milímetros – mm), Temperatura do ar (Média anual da máxima e da mínima, medida em grau Celsius - °C), Umidade do ar (Média anual, medida através da porcentagem - %), Insolação (Total anual, medida pelo tempo de exposição ao sol – horas/minutos), Evaporação (Total anual, medido em milímetros – mm).

A partir do estudo da CONDEPE/FIDEN sobre a região de influência das cidades de Pernambuco, foram selecionados os municípios de maior expressão socioeconômica, as denominadas cidades polos⁷ de cada região fisiográfica. Todavia, algumas dessas cidades polos não possuem estações meteorológicas com dados disponíveis para a realização de pesquisas, sendo esse o motivo da escolha dos seguintes municípios: Petrolina (Sertão do São Francisco); Arcoverde (Sertão do Moxotó); Garanhuns (Agreste Meridional); Surubim (Agreste Setentrional) e Recife (Litoral), Figura 10.

⁷ **As Cidades Polos** são aquelas onde ocorre a exacerbação das características de uma cidade: aglomera a função de gestão federal e empresarial, a presença de equipamentos e de serviços, atraindo maiores fluxos de pessoas e de recursos que outras cidades numa determinada região (CONDEPE/FIDEN, 2013).

Figura 10. Mapa de localização dos municípios Estudados



Fonte: O autor, 2017

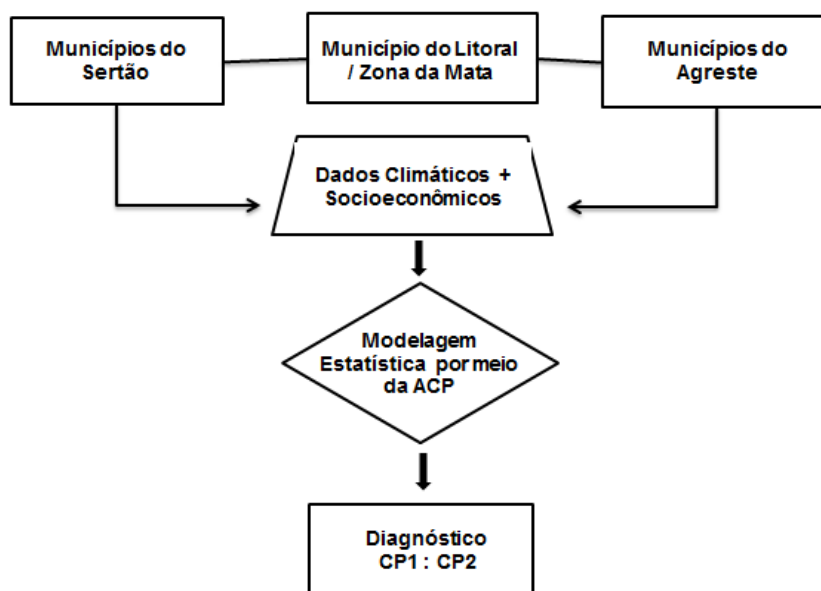
Em seguida, foram recolhidos dados de indicadores sociais e econômicos desses municípios. Como indicador econômico foi utilizado os dados do Produto Interno Bruto (PIB) que de acordo com o IBGE (2017), representa a soma (em valores monetários) de todos os bens e serviços finais produzidos numa determinada região (quer sejam países, estados ou cidades), durante um período determinado (mês, trimestre, ano etc), podendo ser ainda classificado em PIB Agropecuário, PIB Industrial e PIB dos Serviços. No Brasil, o cálculo do PIB é oficialmente de responsabilidade do Instituto Brasileiro de Geografia e Estatística (IBGE). Em Pernambuco, o cálculo do respectivo indicador é de competência da Agência Estadual de Planejamento e Pesquisas de Pernambuco (CONDEPE/FIDEM).

Como indicador social foi utilizando dados do Índice de Desenvolvimento Municipal (IDM) elaborado pelo Instituto Firjan. O IDFM é um estudo anual criado para acompanhar o desenvolvimento humano, econômico e social de todos os municípios do Brasil, com base exclusivamente em estatísticas oficiais, levando em conta três indicadores: Emprego e renda, educação e Saúde. O IDM elaborado pelo instituto Firjan é semelhante ao Índice de Desenvolvimento Humano (IDH), calculado pela ONU, entretanto, uma das diferenças entre ambos é que os dados do IDM podem ser estimados todo ano, enquanto que os do IDH só são levantados uma vez por

década, pois dependem de dados do censo demográfico realizado a cada 10 anos (FIRJAN, 2016).

O período selecionado para a análise desses dados teve como parâmetro o mesmo período de tempo de registro dos dados disponibilizados do PIB e do IFDH dos municípios estudados, ou seja, entre os anos de 1999 a 2013. Para a modelagem dessas variáveis (climáticas e socioeconômicas) e elaboração da ACPs foi utilizado o *Software Statistica*® 10.0. Na figura 11 é mostrado o fluxograma referente ao processo metodológico para obtenção dos resultados.

Figura 11. Fluxograma referente ao processo metodológico adotado



Fonte: O autor, 2017

Pretendeu-se com essa análise avaliar de que forma os elementos do clima influenciam na conjuntura socioeconômica dos municípios maior concentração econômico-funcional do Estado de Pernambuco no período de 1999 a 2013. Foram elaboradas tabelas e gráficos (*Scree plot* e *Biplot*) com a finalidade de mostrar os resultados de acordo com cada município estudado. Por se tratar de municípios com características fisiográficas pouco semelhantes, os resultados demonstrados pela ACP mostraram consonância em relação com estudos já realizados, onde determinados elementos do clima se correlacionam com as características socioeconômicas da área em estudo.

5. REFERÊNCIAS BIBLIOGRÁFICAS

AB'SÁBER, Aziz Nacib. **Sertões e sertanejos: uma geografia humana sofrida**. São Paulo: Universidade de São Paulo, 1999. 55 p.

Agência Estadual de Planejamento e Pesquisas de Pernambuco - CONDEPE/FIDEM. **Pernambuco em rede**: caracterização das regiões de influência das cidades pernambucanas / Coordenação de Ruskin Marinho de Freitas e Georgia Cavalcanti Alves de Miranda. Recife, 2013. 150 p.

Agência Estadual de Planejamento e Pesquisas de Pernambuco - CONDEPE/FIDEM. **Pernambuco em mapas** / Coordenação de Ruskin Marinho de Freitas e Kamila Soares de Arruda Santos. Recife, 2011. 159 p.

ANDRADE, Manuel Correia de Oliveira (Org.). **Atlas Escolar de Pernambuco**: Espaço Geo-Histórico e Cultural. 2. ed. João Pessoa: Grafset, 2003.

ARAÚJO, Vanessa Vasconcelos de et al. ANALISE OS EFEITOS DA MORTALIDADE DE BOVINOS NO AGRESTE PERNAMBUCANO EM DECORRÊNCIA DA ESTIAGEM PROLONGADA. In: Jornada de Ensino, Pesquisa e Extensão- XIII JEPEX., 2013, Recife. **Resumos**. Recife: UFRPE, 2013.

AYOADE, John O. **Introdução à climatologia para os trópicos**. 9. ed. Rio de Janeiro: Bertrand Brasil, 2007. 332 p

BARBOSA, Vanessa Vasconcelos et al. ANÁLISE DA VARIABILIDADE CLIMÁTICA DO MUNICÍPIO DE GARANHUNS, PERNAMBUCO – BRASIL. **Revista Brasileira de Geografia Física**, Recife, v. 2, n. 9, p.353-367, 2016.

BRASIL, Lei nº 9.433 de 8 de Janeiro de 1997. Institui a Política Nacional de Recursos Hídricos, cria o Sistema Nacional de Gerenciamento de Recursos Hídricos. Disponível

em:<<http://www.mma.gov.br/port/conama/legiabre.cfm?codlegi=370>> Acessado em 16 mar. 2017.

BRASIL. Instituto de Geografia e Estatística. Ministério do Planejamento. **Síntese de Pesquisa**. 2017. Disponível em: <<http://www.ibge.gov.br/home/estatistica/pesquisas/sintese.php>>. Acesso em: 15 mar. 2017.

CALDEIRA, Jennifer Karen Alves. **Aproveitamento de água de chuva em uma indústria mecânica**: Aproveitamento de água de chuva em uma indústria mecânica: aspectos econômicos e ambientais. 2016. 131 f. Dissertação (Mestrado) - Curso de Engenharia Ambiental, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2016.

CARVALHO, Alcione Cavalcante de et al. Atuação do Psicólogo no enfrentamento de causas naturais. **Psicologia.pt**, Jaú, v. 13, n. 3, p.1-8, 11 nov. 2012. Disponível em: <<http://www.psicologia.pt/>>. Acesso em: 15 mar. 2017.

CENTRO DE PREVISÃO E ESTUDOS CLIMÁTICOS/ INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (CPTEC/INPE). Glossário Técnico. Disponível em <www.cptec.inpe.br/glossario.shtml> Acessado em 17 mar. 2017.

CRUZ, Ivane Carneiro da; TOPA, Marcos Aurelio. **Análise multivariada como ferramenta de Gerenciamento de fornecedores visando um Relacionamento com vantagem competitiva**. 2009. 82 f. TCC (Graduação) - Curso de Estatística, Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2009.

FERREIRA, D.F. **Estatística Multivariada**. Lavras: UFLA, 2011. 675p.

FIGUEIRA, Ivan; MENDLOWICZ, Mauro. Diagnóstico do transtorno de estresse pós-traumático. **Rev. Bras. Psiquiatr.** 2003, v. 25, 12-16, p.

HAIR, Joseph. et. al. **Análise Multivariada de Dados**. 5 .ed. Porto Alegre: Bookman, 2005.

HONGYU, Kuang. **Comparação do GGE-biplot ponderado e AMMI-ponderado com outros modelos de interação genótipo × ambiente**. 2015. 155 f. Tese (Doutorado) - Curso de Estatística e Experimentação Agronômica, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2015.

HONGYU, Kuang; SANDANIELO, Vera Lúcia Martins; OLIVEIRA JUNIOR, Gilmar Jorge de. Análise de Componentes Principais: resumo teórico, aplicação e interpretação. **Engineering And Science**, Cuiabá, v. 1, n. 5, p.83-90, maio 2015.

JATOBÁ, L.O. **O Sistema Climático**. Recife: Libertas, 2012. 141 p.

JOHNSON, Richard. A; WICHERN, Dean. W. **Applied multivariate statistical analysis**. Fourth Edition New Jersey: Printice Hall, 2007.

KLEFENS, Paula Cristina de Oliveira. **O Biplot na análise fatorial multivariada**. 2009. 103 f. Dissertação (Mestrado) - Curso de Estatística, Universidade de São Paulo, Piracicaba, 2009.

LACERDA, Francinete Francis et al. Tendência do clima do semiárido frente as perspectivas das mudanças climáticas globais; o caso de Araripina, Pernambuco. **Geography Department University Of Sao Paulo**. v. 31, p.132-141, 24 jul. 2016.

LOPES, Luís Felipe Dias. **Análise de componentes principais aplicada à confiabilidade de sistemas complexos**. 2001. 121 f. Tese (Doutorado) - Curso de Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2001

LUCAS, Edimilson Costa. **Impacto de Chuvas Extremas em Ações da Indústria de Alimentos**. 4. ed. São Paulo: Fundação Getúlio Vargas, 2015. 33 p. (Textos para Discussão).

LYRA, Wellington da Silva et al. Classificação periódica: um exemplo didático para ensinar análise de componentes principais. **Química Nova**, Campina Grande, v. 33, n. 7, p.1594-1597, abr. 2010.

MANLY, Bryan. **Multivariate Statistical Methods**. New York, Chapman and Hall, 1986.159 p.

MARTINS, Sandro Luís Moresco; GODOY, Leoni Pentiado; LOPES, Luis Felipe Dias. Monitoramento do controle estatístico do processo utilizando ferramentas estatísticas. **Espacios**, Santa Maria, v. 33, n. 6, p.16-21, fev. 2012

MELO, Valneli da Silva; MEDEIROS, Raimundo Mainar de; SOUZA, Francisco de Assis Salviano. Variação média mensal e anual da umidade relativa do ar para 11 municípios no estado de Pernambuco, brasil. In: congresso técnico científico da engenharia e da agronomia, 2015, Fortaleza. **Anais...** . Fortaleza: Contecc, 2015. v. 1, p. 1 - 5.

MENDONÇA, Francisco. **Climatologia: noções básicas e climas do Brasil**. São Paulo: Oficina de Textos, 2009. 201 p.

MIGUEL, Shirly Pereira; BUENO, Maurício Haas. Habilidades sociais e estratégias de enfrentamento: um estudo correlacional. **Revista Psicologia PT**, v.4, n. 9, p. 1-8, nov. 2007. Disponível em: <<http://www.psicologia.pt/>>. Acesso em: 15 mar. 2017.

PERNAMBUCO. Secretaria de Ciência, Tecnologia e Meio Ambiente. **Atlas de Bacias Hidrográficas de Pernambuco**. Simone Rosa da Silva (Org.), Recife: Ed. A secretaria. 2006. CD-ROM.

PERNAMBUCO. Secretária de Saúde de Pernambuco. Governo Estadual (Org.). **Casos de dengue 2015**. Disponível em: <www.saude.pe.gov.br/casosdenque2015>. Acesso em: 11 mar. 2017.

REGAZZI, A.J. **Análise multivariada**, notas de aula INF 766, Departamento de Informática da Universidade Federal de Viçosa, v.2, 2001.

SETTE, Denise Maria; RIBEIRO, Helena. Interações entre o clima, o tempo e a saúde humana. **Saúde Meio Ambiente e Sustentabilidade**, São Paulo, v. 6, n. 2, p.37-51, ago. 2011.

SILVA, Vicente Natanael Lima. **As condições climáticas do Estado de Pernambuco e seus impactos nos aspectos socioeconômicos**. 2013. 93 f. TCC (Graduação) - Curso de Licenciatura em Geografia, Ciências Geográficas, Universidade Federal de Pernambuco, Recife, 2013.

SORRE, Maximilian. **A adaptação ao meio climático e biossocial: geografia psicológica**. São Paulo: Ática; 1984.

SOUZA, Adriano Mendonça. **Monitoração e ajuste de realimentação em processos produtivos multivariados**. 2000. 198 f. Tese (Doutorado) - Curso de Engenharia de Produção, Engenharia de Produção, Universidade Federal de Santa Catarina, Florianópolis, 2000.

TAVARES, Maria da Conceição; ANDRADE, Manuel Correia de; PEREIRA, Raimundo Rodrigues. **Seca e poder: Entrevista com Celso Furtado**. São Paulo: Fundação Perseu Abramo, 1998. 100 p.

TIQUE, Alexandre Tomo. **Análise da Climatologia Dinâmica da Variabilidade da Precipitação na Região Sudeste da África**. 2015. 119 f. Dissertação (Mestrado) - Curso de Ciências Atmosféricas, Ciências Atmosféricas, Universidade de São Paulo, São Paulo, 2015.

TOLEDO FILHO, Manoel da Rocha et al. Investigação da relação entre o clima e o comportamento humano em Maceió, Alagoas. In: Congresso Brasileiro de Meteorologia, 16., 2010, Belém. **Anais...**. Belém: Sbmec, 2010. p. 1 - 5. Disponível em:

<http://www.sbmet.org.br/cbmet2010/artigos/322_70156.pdf>. Acesso em: 15 mar. 2017

VARELLA, Carlos Alberto. **Análise Multivariada Aplicada as Ciências Agrárias: Análise de Componentes Principais**. Universidade Federal Rural do Rio de Janeiro. Seropédica – RJ, 2008.

VASCONCELOS, Simone. **Introdução a Análise de Componentes Principais (ACP)**. Disponível em < <http://www2.ic.uff.br/~aconci/PCA-ACP.pdf> > Acessado em 13 mar. 2017.

VICINI, Lorena. **Análise multivariada da teoria à prática**. 2005. 215 f. Monografia (Especialização) - Curso de Estatística, Universidade Federal de Santa Maria, Santa Maria, 2005.

VIEIRA, Lima. **Meteorologia e Climatologia Agrícola**. Notas de Aula, Universidade Estadual de Maringá. 2009, p. 3- 4.

CAPÍTULO II

**ARTIGO SUBMETIDO PARA PUBLICAÇÃO NA REVISTA
ATMOSPHERIC ENVIRONMENT**

**Socioeconomic Diagnosis Based on Statistical Analysis of
Meteorological Variables: A Case Study in Pernambuco/Brazil**

DIAGNÓSTICO SOCIOECONÔMICO COM BASE EM ANÁLISE ESTATÍSTICA DE VARIÁVEIS METEOROLÓGICAS: Um Estudo de Caso em Pernambuco/Brazil

V. N. Silva¹, V. A. Santos², B. Mohand³, L. A. Sarubbo⁴

¹ Mestrando em Desenvolvimento de Processos Ambientais, Universidades Católica de Pernambuco, Rua do Príncipe n, 526 Boa Vista, CEP: 50050-900, Recife Pernambuco, Brasil

² Profº Drº da Universidade Católica de Pernambuco, Rua do Príncipe n, 526 Boa Vista, CEP: 50050-900, Recife Pernambuco, Brasil

³ Profº Drº da Universidade Federal de Pernambuco, Av. Professor Moraes do Rego, n 1235, Cidade Universitária, CEP: 50670-900, Recife, Pernambuco, Brasil

⁴ Profª Drª da Universidade Católica de Pernambuco, Rua do Príncipe n, 526 Boa Vista, CEP: 50050-900, Recife Pernambuco, Brasil

RESUMO

Na Região Nordeste do Brasil, em especial no Estado de Pernambuco, foi detectada uma correlação entre variáveis meteorológicas e de atividades humanas, sejam elas econômicas, sociais, culturais, etc. Este trabalho serviu como estudo de caso da aplicação da técnica estatística multivariada de Análise de Componentes Principais (ACP). Na confecção de um diagnóstico utilizou-se variáveis meteorológicas e respostas socioeconômicas (IDH e PIB) de alguns municípios que apresentaram expressivo desenvolvimento no Estado de Pernambuco – Brasil, entre os anos de 1999 e 2013. Mesmo considerando as diferentes características climáticas e socioeconômicas dos municípios estudados, a ACP mostrou-se efetiva e permitiu a retirada ou descarte de variáveis que apresentaram baixa variabilidade ou foram redundantes por estarem correlacionadas com as de maior importância para dois componentes principais. A compreensão do comportamento das variáveis meteorológicas e de suas consequências sobre as atividades humanas é de fundamental importância no auxílio a políticas públicas, que visem à mitigação de efeitos adversos provocados pelas alterações ambientais.

Palavras-chaves: Variáveis Meteorológicas, Indicadores Socioeconômicos, Análise de Componentes Principais.

1 – INTRODUÇÃO

Estudos recentes mostram que em diversas regiões do Planeta o comportamento das variáveis meteorológicas vêm influenciando em aspectos socioeconômicos de diversas maneiras como exemplo temos o trabalho de (Ebi e Nealon, 2016), que estudaram a correlação entre a variabilidade da precipitação pluviométrica e a temperatura com os casos de transmissão de dengue pelo mosquito *Aedes aegypti* na região tropical do planeta. O estudo de (Mehta., et al, 2013) relacionado a variabilidade climática interdecadal da precipitação pluviométrica e suas consequências na agricultura e nos recursos hídricos da bacia hidrográfica do rio Missouri, nos Estados Unidos.

O sistema climático de determinada região ou localidade são constituídas pelos fatores climáticos (massas de ar, relevo, latitude, maritimidade, etc.) e as variáveis meteorológicas (temperatura, precipitação pluviométrica, umidade do ar, etc.), que exercem influências nos aspectos socioeconômicos de determinados espaços (Kagan, 1995). As condições climáticas materializadas sobre uma determinada porção da superfície terrestre resultam de uma destacada interação entre as variáveis meteorológicas e os fatores climáticos. A atividade Solar como fator climático, por exemplo, influencia a variabilidade climática a curto e longo prazo sendo determinantes para as atividades humanas em diversas regiões do planeta como analisado por (Ineson, et. al, 2015).

Os principais indicadores dos aspectos socioeconômicos de uma região (país, estado, município, etc.), podem ser mensurados através dos valores do Produto Interno Bruto (PIB), que representa a soma, em valores monetários, de todos os bens e serviços produzidos e o Índice de Desenvolvimento Humano (IDH), no Brasil é representado anualmente pelo Índice de Desenvolvimento Municipal (IDM), é um estudo anual criado para acompanhar o desenvolvimento humano, econômico e social de todos os municípios do Brasil, com base exclusivamente em estatísticas oficiais (IBGE, 2016).

O Estado de Pernambuco é composto por 185 municípios, e em decorrência da sua formação territorial esses municípios apresentam diversas características fisiográficas e socioeconômicas que os distingue, tanto em razão da sua localização

geográfica como das atividades econômicas desenvolvidas ao longo da sua história (CONDEPE/FIDEM, 2016). Diante das atuais conjunturas socioambientais em que se encontra o Estado, Buscou-se para a realização desse estudo analisar, por meio de uma modelagem, entre os dados meteorológicos (Precipitação Pluviométrica, Temperatura do Ar, Umidade Relativa, Insolação, Evaporação) com dados socioeconômicos do PIB (Agropecuário, Industrial e Serviços) e do IDM (Emprego e Renda, Saúde e Educação) nos municípios de expressivo desenvolvimento econômicos do Estado de Pernambuco (Petrolina, Arcoverde, Garanhuns, Surubim e Recife), no período de 1999 a 2013.

Com a quantidade muito grande de variáveis envolvidas nesse processo de modelagem, o número de relações entre elas é também muito grande para ser compreendido. Assim, torna-se necessário o emprego de técnicas que possam resumir, sistematicamente, a correlação significativa existente entre as variáveis (Malekinezhad, 2014). Dentre as técnicas estatística multivariadas, associadas à Redução dos dados ou simplificação estrutural, destaca-se a Análise de Componentes Principais (ACP), que é uma técnica estatística multivariada que consiste em reduzir a dimensionalidade dos dados, de forma que a maior parte da variação do próprio conjunto de dados seja retida, proporcionando uma melhor interpretação (Hongyu, 2015).

Apesar da relevância do tema, as pesquisas ainda são muito limitadas, pela falta de emprego de ferramentas estatísticas que justifiquem o caráter científico das afirmativas contidas nos diversos estudos citados. Além disso, necessita-se que a ferramenta empregada permita a junção de diferentes variáveis possam ser analisadas simultaneamente. Isso levaria à possibilidade de se chegar a um diagnóstico, por região fisiográfica, para aplicação em várias partes do Planeta.

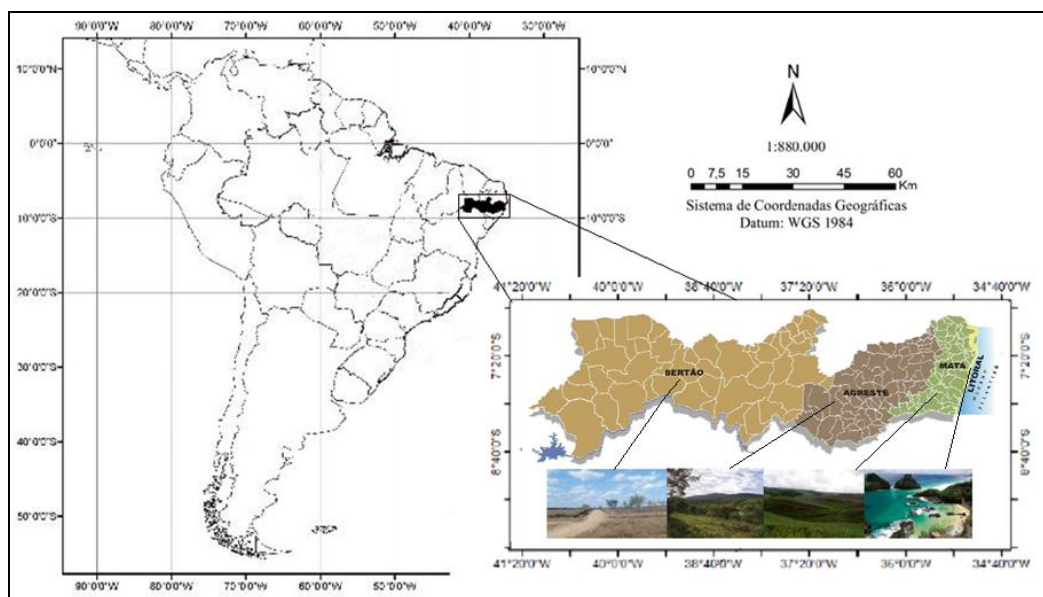
2 - Material e Métodos

2.1 - Caracterização da Área de Estudos

De acordo com dados do Instituto Brasileiro de Geografia e Estatística – IBGE, o Estado de Pernambuco localiza-se na parte oriental da região Nordeste do Brasil, entre as coordenadas geográficas 07°15'45"/ 09°28'18"S e de 034°48'35"/ 041°19'54"W,

possuindo uma área de 98.146,3 km² de extensão e sua população estimada em 2016 é de 9.410.336 de habitantes (figura 1). Limita-se ao norte, com os estados da Paraíba e do Ceará; ao sul, com os Estados de Alagoas e da Bahia; a leste, com o Oceano Atlântico; e a oeste, com o estado do Piauí. (IBGE, 2016).

Figura 1. Localização da área de estudo



Fonte: O autor, 2017

2.2 - Aspectos Fisiográficos e Socioeconômicos

O Estado de Pernambuco apresenta do litoral para o interior, uma sucessão de paisagens e de formas diferenciadas de organização do espaço, daí ter sido seu território dividido em três regiões fisiográficas: Litoral-Zona da Mata, Agreste e Sertão. Do território compreendido pelo Estado, 64% pertence ao Sertão, 25% ao Agreste e 11% a Zona da Mata/Litoral, sendo que cada uma dessas regiões apresentam tanto características climáticas distintas, quanto socioeconômicas. Assim, foram selecionados os municípios de maior expressão socioeconômica de cada região fisiográfica: Petrolina (Sertão do São Francisco); Arcoverde (Sertão do Moxotó); Garanhuns (Agreste Meridional); Surubim (Agreste Setentrional) e Recife (Litoral), (CONDEPE/FIDEM, 2016).

Pernambuco possui uma população estimada de 9.410.336 e atualmente, sua economia tem como base a agricultura, a indústria e os serviços. O setor de serviços é

predominante, seguido pela indústria naval, automobilística, química, metalúrgicos, eletroeletrônico, de minerais não-metálicos, têxtil e alimentícia. O estado assiste a uma importante mudança em seu perfil econômico com os recentes investimentos nos setores petroquímico, biotecnológico, farmacêutico, de informática, naval e automotivo, que estão dando novo impulso à economia do Estado que vem crescendo acima da média nacional (SDEC-PE, 2016).

Os municípios de Petrolina e Arcoverde, que localizam-se na região com menores índices pluviométricos do estado, o sertão, onde a economia é basicamente voltada para as atividades de comércio de bens e serviços. Em Petrolina destaca-se a agricultura irrigada, a partir do rio São Francisco. Em Arcoverde, nos períodos de estiagem prolongada à agricultura é praticamente inexistente (CONDEPE/FIDEM, 2016). Os municípios de Garanhuns e Surubim, na região agreste do estado, possuem uma economia baseada na indústria de laticínios e nas atividades de comércio de bens e serviços. Em Garanhuns são bastante desenvolvidas a atividade da bovinocultura (leite e corte) e o turismo. Surubim tem como destaques, o setor agropecuário (produção de cana-de-açúcar, leite, banana, bovinos e horticultura) e dos serviços (CONDEPE/FIDEM, 2016). A cidade do Recife é a capital, maior e principal aglomerado urbano do estado, foi o principal centro financeiro do Brasil Colônia até meados do século XVIII, quando do término do ciclo da cana-de-açúcar. Atualmente sobressaem-se as atividades da indústria de transformação, a construção civil e o comércio (IBGE, 2016).

2.3 - Análise de Componentes Principais (ACP)

A análise de componentes principais (ACP) é uma técnica estatística multivariada de modelagem da estrutura de covariância. Ela está associada à ideia de redução de massa de dados, com menor perda possível da informação, contudo é importante ter uma visão conjunta de todas ou quase todas as técnicas da estatística multivariada para resolver a maioria dos problemas práticos. Procura-se redistribuir a variação observada nos eixos originais de forma a se obter um conjunto de eixos ortogonais não correlacionados (HONGYU, 2015).

A ACP tem como principais vantagens: Retirar a multicolinearidade das variáveis, pois permite transformar um conjunto de variáveis originais intercorrelacionadas em um novo conjunto de variáveis não correlacionadas (componentes principais). Além disso, reduz muitas variáveis a eixos que representam algumas variáveis, sendo estes eixos perpendiculares (ortogonais) explicando a variação dos dados de forma decrescente e independente. (HONGYU, 2016).

Para a obtenção dos componentes principais de uma forma geral, seja um conjunto de p variáveis $X_1, X_2, \dots, X_p$ com médias $\mu_1, \mu_2, \dots, \mu_p$ e variância $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_p^2$, respectivamente. Estas variáveis não são independentes e portanto, possuem covariância entre a i -ésima e k -ésima variável definida por σ_{ik} , para $i \neq k = 1, 2, \dots, p$. Então as p variáveis podem ser expressas na forma vetorial por $\mathbf{X} = [X_1, X_2, \dots, X_p]'$, com vetor de médias $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]'$ e matriz de covariância $\boldsymbol{\Sigma}$.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1p}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \cdots & \sigma_{nn}^2 \end{bmatrix}$$

Encontram-se os pares de autovalores e autovetores $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$, em que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, associados a $\boldsymbol{\Sigma}$ e então o i -ésimo componente principal é definido por (HONGYU, 2016):

$$\mathbf{Z}_i = \mathbf{e}_i' \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p \quad (1)$$

A variável $\mathbf{Z}_i$, é uma variável latente, ou seja, não é mensurada a partir do experimento ou levantamento amostral, a ideia é projetar os pontos coordenados originais em um plano maximizando a distância entre eles, o que equivale a maximizar a variabilidade da variável latente $\mathbf{Z}_i$. A variância de $\mathbf{Z}_i$ é dada por

$$\text{Var}(\mathbf{Z}_i) = \text{Var}(\mathbf{e}_i' \mathbf{X}) = \mathbf{e}_i' \text{Var}(\mathbf{X}) \mathbf{e}_i = \mathbf{e}_i' \boldsymbol{\Sigma} \mathbf{e}_i \quad (2)$$

em que $i = 1, \dots, p$.

Utilizando a decomposição espectral da matriz Σ , dada por $\Sigma = \mathbf{P}\mathbf{\Lambda}\mathbf{P}'$, em que $\mathbf{P}$ é a matriz composta pelos autovetores de Σ em suas colunas e $\mathbf{\Lambda}$, a matriz diagonal de autovalores de Σ , então, tem-se que

$$tr(\Sigma) = tr(\mathbf{P}\mathbf{\Lambda}\mathbf{P}') = tr(\mathbf{\Lambda}\mathbf{P}'\mathbf{P}) = tr(\mathbf{\Lambda}\mathbf{I}) = tr(\mathbf{\Lambda}) = \sum_{i=1}^p \lambda_i \quad (3)$$

e

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}$$

$Etr(\Sigma)$ é dada pela soma dos elementos da diagonal:

$$tr(\Sigma) = \sum_{i=1}^p \sigma_{ii} = \sum_{i=1}^p \lambda_i \quad (4)$$

Portanto, a variável total contida nas variáveis originais é igual a variabilidade total condita nos componentes principais (JOHNSON; WICHERN, 1998). A contribuição de cada componente principal (Z_i) é expressa em porcentagem, e a explicação individual de cada componente pode ser calculada pelo critério sugerido por KAISER (1958) isto é, este critério tende a incluir poucas componentes quando o número de variáveis originais é inferior a vinte e, em geral, utiliza-se aquelas componentes que conseguem sintetizar uma variância acumulada em torno de 70%, os principais componentes que explicam a maior parte da variação no conjunto de dados (SAVEGNAGO et al., 2011).

A interpretação de uma componente principal é feita mediante o grau de importância, ou, ainda, a influência que cada variável tem sobre cada componente, sendo que esta importância é dada pela correlação entre cada variável jX e o componente iY que estiver sendo interpretado (REGAZZI, 2001).

Através de gráficos de explicação da proporção de variação de cada componente principal, **Scree Plot**, é possível identificar a porcentagem de variação explicada por cada componente principal. Já para um gráfico **biplot** ou **circulo de correlações**, A ideia é de acrescentar informações sobre várias variáveis em um único gráfico obtido pelos componentes principais (KLEFENS, 2009).

4 - Resultados e Discussão

Em todos os municípios estudados foram levadas em consideração o 1° e o 2° fator ou CP, onde foi possível demonstra quais as variáveis meteorológicas que mais influenciam nas variáveis respostas (PIB/IDM) dos municípios estudados. Utilizando o mesmo procedimento Melo, Medeiros e Souza (2015) a analisar e obter a dependência espacial da precipitação pluviométrica no Estado da Paraíba.

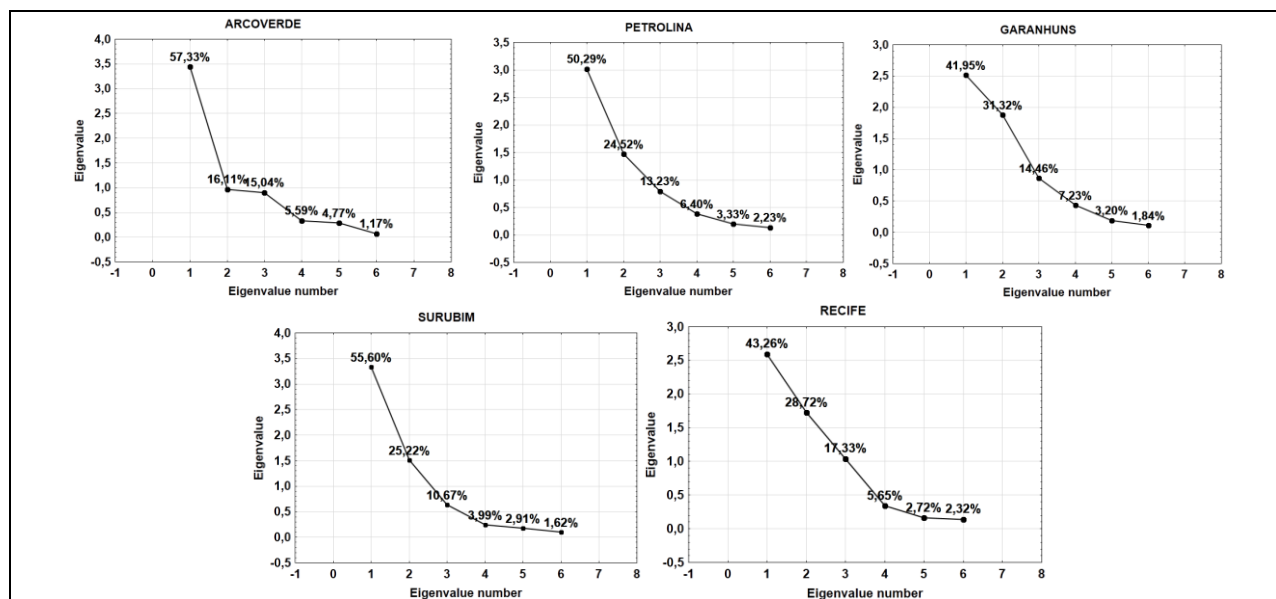
A análise das variáveis mais correlacionadas na 1° CP e na 2° CP mostrou que nos municípios de Petrolina e Arcoverde a variável temperatura máxima aparece como a mais significativa, seguida pela umidade relativa e insolação. Nos municípios de Garanhuns e Surubim, a variável insolação destaca-se por apresentar correlacionada 1° CP de ambos os municípios. Na Cidade do Recife a variável precipitação pluviométrica e a que mais contribui para a 1° CP (Tabela 1).

Tabela 1. Valores da matriz de correlação Var. Meteorológicas x PIB/IDM

Municípios	Principal Componente	Autovalores (PIB)	Proporção acumulada - PIB (%)	Autovalores (IDM)	Proporção acumulada - IDM (%)
Arcoverde	1ª CP	3,4	57,3	4,3	73,0
	2ª CP	0,96	73,4	0,84	87,1
Petrolina	1ª CP	3,0	50,2	3,1	52,8
	2ª CP	1,4	74,8	1,6	80,7
Garanhuns	1ª CP	2,5	41,9	2,4	40,5
	2ª CP	1,8	73,2	1,8	71,8
Surubim	1ª CP	3,0	50,2	3,9	65,0
	2ª CP	1,4	74,8	1,2	85,9
Recife	1ª CP	2,5	43,2	2,5	42,2
	2ª CP	1,7	71,9	1,9	75,4

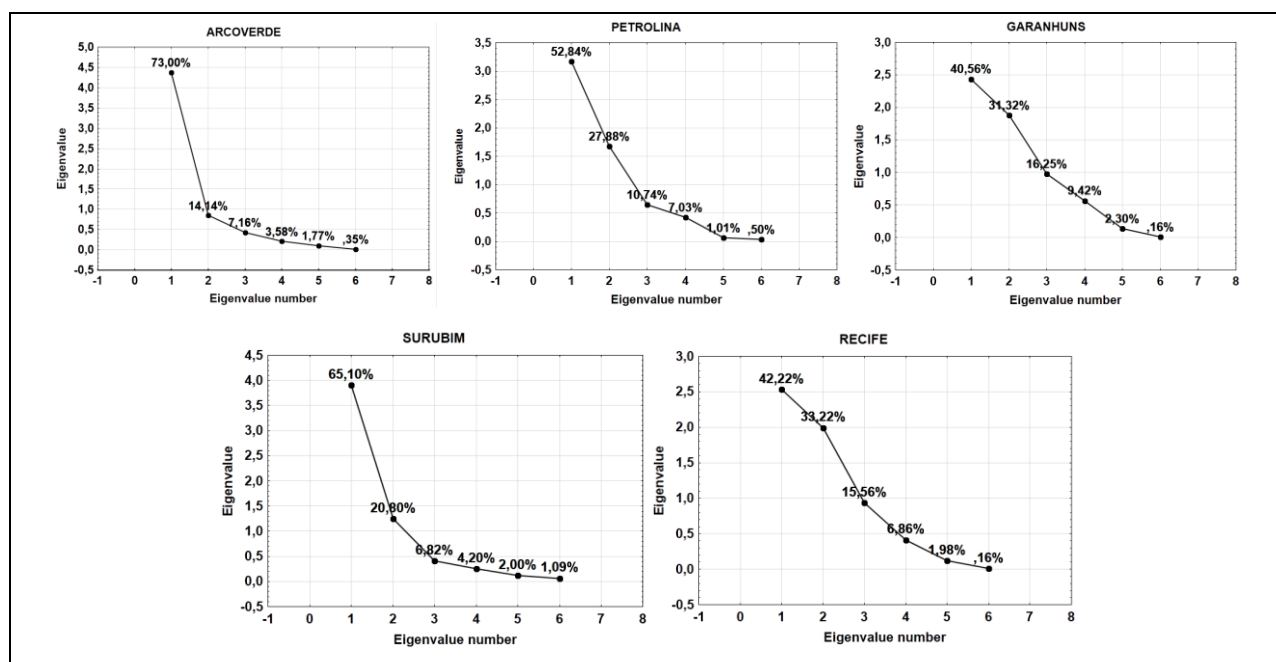
A seleção dos componentes principais através do método gráfico Scree Plot, (figura 2), mostrou uma forma geral que em todos os municípios analisados, as componentes principais e os seus valores (scores – 1° CP) explicaram a maior parcela da variância do conjunto original de dados, que pelo critério de Kaiser (1958), corresponderam valores $\geq 70\%$ da variância entre as CPs. Filho et al (2013) também usaram tal critério na construção de indicadores socioeconômicos.

Figura 2. Variáveis Meteorológicas x PIB (1999-2013)



A curva existente em cada gráfico evidência uma estabilização dos autovalores da matriz de correlação. Ou seja, os fatores com maior significância para a análise dos dados no Scree plot são aqueles com valores próximos de 1 (figura 3).

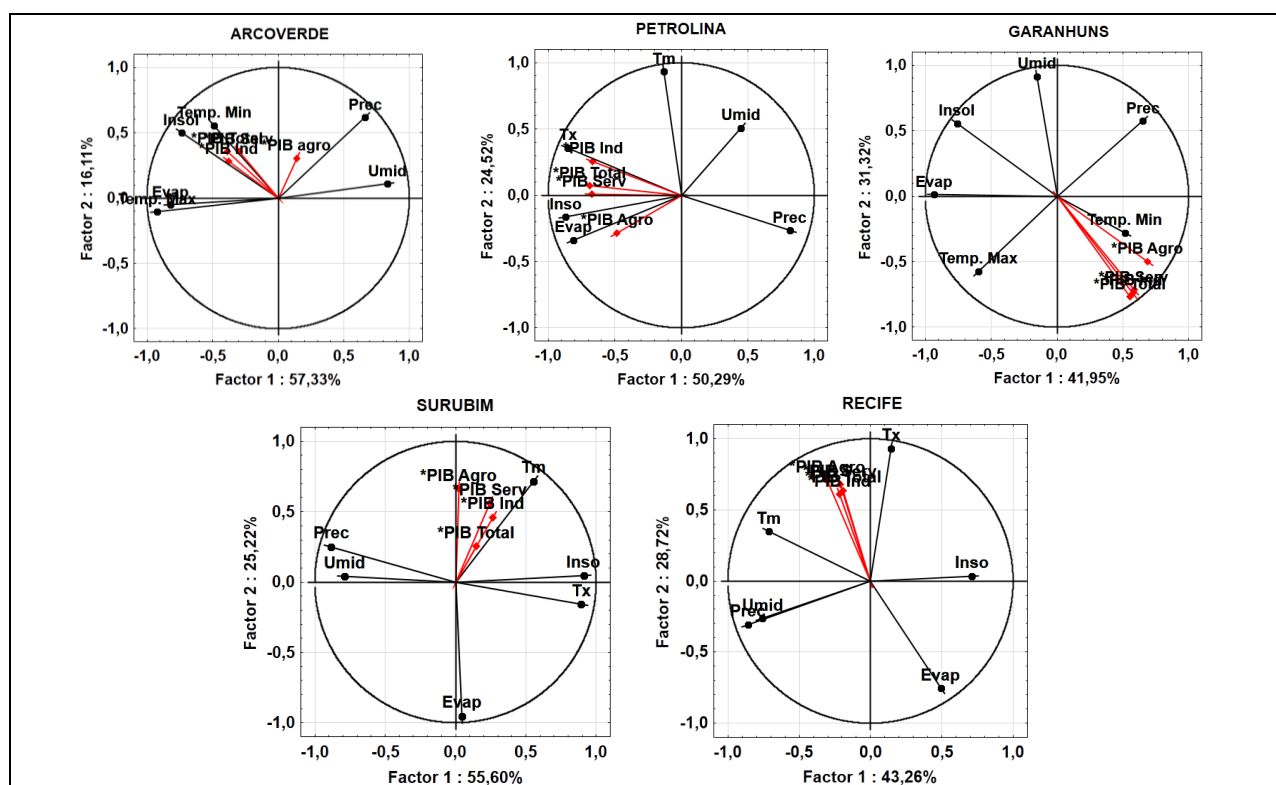
Figura 3. Variáveis Meteorológicas x IDM (2005-2013)



Nos gráficos Biplot, que representam além dos scores de cada fator, uma forma de identificar visualmente que variáveis originais estão relacionadas com o estudo, para Godoi, Hongyu e Dias (2012) a utilização do método Biplot varia de acordo como objetivo e o interesse do pesquisador, sendo muito utilizado na área do melhoramento, principalmente para o estudo de interação entre variáveis. No nosso estudo, as variáveis meteorológicas que mais se relacionam com as socioeconômicas.

Algumas variáveis estão sobrepostas umas às outras. Isso mostra que essas possuem a mesma representatividade no gráfico (figuras 4). Outro fato importante, é que algumas variáveis estão bem próximas ao círculo unitário. Isso mostra que estas possuem uma maior contribuição, em relação às variáveis que estão mais afastadas.

Figura 4. Biplot fator 1 $\times$ fator 2 sobre as variáveis meteorológicas e o PIB (1999-2013)



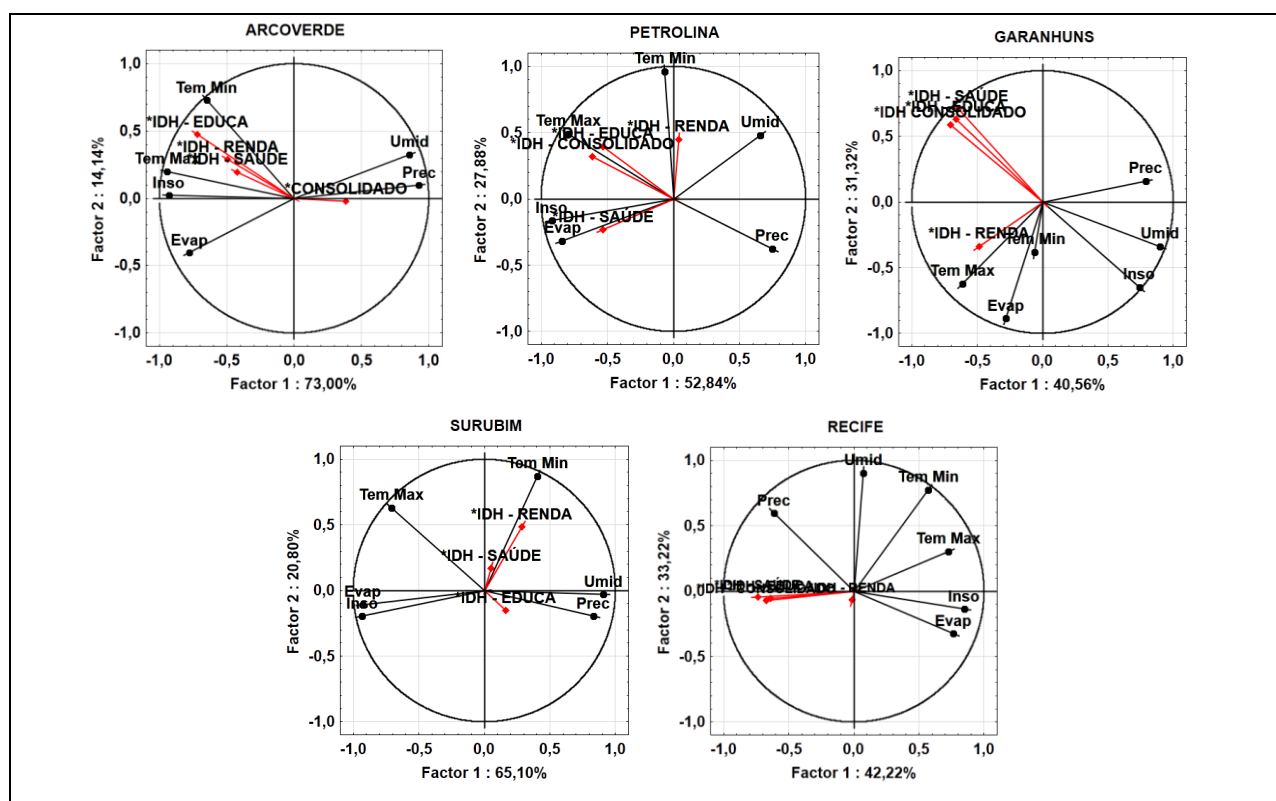
Foi diagnosticado no município de Arcoverde, foi constatada uma grande influência da variável precipitação pluviométrica e umidade sobre a variável PIB - agropecuário. No município de Petrolina, ficou constatado que a precipitação pluviométrica e a umidade não influenciam nos valores do PIB (agropecuário, industrial

e Serviços), sendo este mais correlacionado com a temperatura máxima, insolação e evaporação.

No município de Garanhuns, a variável precipitação pluviométrica e temperatura mínima exercem grande influência, no 1º fator, em todos os setores do PIB. No município de Surubim foi identificada uma influência maior, no 1º fator, da variável temperatura mínima e insolação sobre todos os setores do PIB, sendo a agropecuária com maior correlação. Na cidade do Recife, constatou-se que as variáveis meteorológicas no 1º fator pouco influenciam as variáveis socioeconômicas do PIB.

Na figura 5 são mostrados os resultados referentes a análise das variáveis meteorológicas com as do IDM.

Figura 5. Biplot CP1 × CP2 sobre as variáveis meteorológicas e o IDM (2005-2013)



Em Arcoverde, de maneira geral a precipitação pluviométrica, a umidade possui influência no IDM geral. Por outro lado, o IDM relacionado à renda, saúde e educação possui maior correlação com o comportamento da variável insolação, temperatura máxima e mínima. No município de Petrolina ficou constatado que em relação ao IDM,

a saúde sofre mais influência da insolação e da evaporação enquanto IDM renda e educação das temperaturas máxima e mínima.

Em Garanhuns as variáveis temperatura máxima e a evaporação relacionam-se com o IDM – renda, por outro lado, o IDM – geral, saúde e da educação não correlacionam-se com nenhuma variável meteorológica (1º e 2º fatores). Na cidade do Recife, constatou-se que as variáveis meteorológicas no 1º fator pouco influenciam as variáveis socioeconômicas. Por outro lado, o IDM relacionado à renda, saúde e educação possui maior correlação com o comportamento da variável precipitação pluviométrica.

3 - Conclusões

Tendo em vista os resultados obtidos, há uma relação entre o comportamento das variáveis meteorológicas e os aspectos socioeconômicos dos dados de cada município analisado. Não necessariamente a quantidade e a distribuição espacial, mais também em relação aos aspectos fisiográficas dos referidos municípios. Pode-se inferir também uma relação entre as condições de variabilidades climáticas anuais (anos secos e/ou chuvosos) com o desenvolvimento econômico e urbano dos respectivos municípios

A Análise de Componentes Principais foi eficaz para o tratamento proposto. Em todos os municípios estudados, o 1º fator de análise, em conjunto com o 2º fator, demonstraram que as variáveis meteorológicas influenciaram, em diferentes escalas, as variáveis PIB e IDM nos seus diversos setores. Um dos objetivos da ACP, neste caso, foi atingido, pois um número relativamente pequeno de componentes foi extraído (1ºCP e 2º CP) com a capacidade de explicar a maior variabilidade nos dados originais.

A ACP mostrou-se efetiva e permitiu a retirada ou descarte de variáveis que apresentaram baixa variabilidade ou foram redundantes por estarem correlacionadas com as de maior importância para dois componentes principais. Assim, um menor número de variáveis foram necessárias para explicar a variação total resultando em economia de tempo e de recursos em futuros trabalhos que utilizarão essa mesma base de dados, sem perda significativa de informação.

5 - Referências Bibliográficas

Agência Estatal de Planejamento e Pesquisa de Pernambuco-CONDEPE/FIDEM. Pernambuco aspectos naturais. Disponível em < <http://www.condepefidem.pe.gov.br>>, Acessado em 06 dez 2016.

AYOADE, J. O. Introdução à climatologia para os trópicos. 4 ed. Rio de Janeiro: Bertrand Brasil. 1996, p.262,286. Disponível em < <http://pt.slideshare.net/lcgrh/ayoadejo-introduo-climatologia-para-os-trpicos-cpia>>. Acessado em 06 dez 2016.

Ebi, K.L. Nealon. J. Dengue in a changing climate. Environmental Research. v. 151, n. 6, p. 115–123. 2016. DOI: <http://dx.doi.org/10.1016/j.envres.2016.07.026>

GODOI, Débora Robert de; HONGYU, Kuang; DIAS, Carlos Tadeu dos Santos. Os métodos Biplot de Gabriel na análise multivariada. In: REUNIÃO ANUAL DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 57., 2012, Piracicaba. **Anais...** . Piracicaba: RBRAS, 2012.

Hongyu, K. Comparação do GGE-biplot ponderado e AMMI-ponderado com outros modelos de interação genótipo × ambiente. 2015. Tese (Doutorado em Estatística e Experimentação Agronômica) - Escola Superior de Agricultura “Luiz de Queiroz”, University of São Paulo, Piracicaba, 2015. DOI. 10.11606 / T.11.2015.tde-04052015-172304.

Hongyu, et al. Principal Component Analysis: theory, interpretations and applications. E&S - Engineering and Science, v. 1, n. 5, p. 83-90. ISSN: 2358-5390 DOI: 10.18607/ES20165053.

Ineson, S. et al. Regional climate impacts of a possible future grand solar minimum. Nature Communications. Cambridge, v. 6, n. 7535, p. 1 – 8. 2015. DOI: <https://doi.org/10.1038/ncomms8535>

IBGE - Instituto Brasileiro de Geografia e estatística. Estados de Pernambuco. Disponível em < <http://www.ibge.gov.br/estadosat/perfil.php?sigla=pe>> Acessado em 06 dez 2016.

Johnson, R. A.; Wichern, D. W. Applied multivariate statistical analysis. Fourth Edition New Jersey: Printice Hall, 1998. Disponível em < <http://www1.udel.edu/oiss/pdf/617.pdf>> Acessado em 08 dez 2016.

Kagan, B.A. Ocean-atmosphere interaction and Climate Modeling. Traduce. M. A. Chazin. Cambridge: Cambridge University Press, 1995. Disponível em < <https://goo.gl/6wYLtp>> Acessado em 10 dez 2016.

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. Psychometrika, v. 23, n. 3.p. 187-200,1958. < <https://doi.org/10.1007/bf02289233>>

Klefens, P. C. O Biplot na análise fatorial multivariada. Dissertação (Mestrado em Estatística e Experimentação Agrônômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 2009. DOI.10.11606/D.11.2010.tde-25022010-085843.

Malekinezhad, H. Regionalization Approach for Modeling of Monthly Evaporation Based on Cluster Analysis. Natural Resources and Conservation, v. 2, n. 2, p. 25-32, 2014. DOI: 10.13189/nrc.2014.020202

Regazzi, A. J. Análise Multivariada. Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas. Departamento de Informática, 2001. 166p. Di <https://goo.gl/AeyhJM>> Acessado em 10 dez 2016.

Savegnago, R.P. et.al. Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population, Poultry Science, 90, p.2174-2188. 2011. DOI: [10.3382 / ps.2011-01474](https://doi.org/10.3382/ps.2011-01474).

CAPÍTULO III

CONCLUSÕES GERAIS

A técnica estatística multivariada Análise de Componentes Principais (ACP) mostrou-se eficaz em diagnosticar quais as variáveis meteorológicas que mais influenciam nos setores socioeconômicos dos municípios analisados. Um dos objetivos da ACP, redução no número de variáveis sem perda significativa da informação, foi atingido, pois foram utilizadas os somente 2 variáveis (1°CP e 2° CP) de um total de 12 variáveis originais.

ACP demonstrou que existe correlação significativas entre o comportamento das variáveis meteorológicas e os aspectos socioeconômicos nos municípios analisados. Os Valores da Matriz de correlação entre todas as variáveis analisadas, contidas na 1° CP e na 2° CP, obtiveram proporção acumulada acima de 70% em todos os casos estudados.

Em relação ao diagnostico socioeconômico dos municípios, pôde-se concluir que:

- No município de Arcoverde, foi constatada uma maior influência das variáveis Temp. máxima e umidade sobre a variável PIB - agropecuário. Em relação ao IDM, este foi mais correlacionado com variáveis Precipitação e Temp. Mínima.
- No município de Petrolina, foi diagnosticado que a Insolação e a Temp. máxima e a mais que influencia nos valores do PIB (agropecuário, industrial e Serviços). Em relação ao IDM, a saúde sofre mais influência da Insolação e da Evaporação enquanto IDM renda e educação das Temp. máxima e mínima.
- No município de Garanhuns, a variável Precipitação e Temp. mínima exercem grande influência em todos os setores do PIB. Em relação ao IDM, as variáveis Temp. máxima e a Evaporação relacionam-se com o IDM - Renda, por outro lado, o IDM - Geral, Saúde e da Educação não demonstraram correlação significativa com nenhuma variável meteorológica.
- No município de Surubim, foi diagnosticada que a variável Temp. mínima e insolação exerceram sobre todos os setores do PIB sendo a agropecuária com

maior correlação. A variável temperatura mínima e a umidade influenciaram os valores do IDM – renda, saúde.

- Na cidade do Recife, constatou-se que a variável meteorológica Temp. mínima e máxima demonstraram serem as mais influentes sobre os indicadores do PIB. Por outro lado, o IDM relacionado à renda, saúde e educação possui maior correlação com o comportamento da variável precipitação pluviométrica.

Por fim, a modelagem de dados climáticos e socioeconômicos utilizando ACP mostrou-se efetiva e permitiu uma análise coerente da realidade de cada município estudado, podendo ser utilizada para futuras pesquisas.

ANEXOS

SOCIOECONOMIC DIAGNOSIS BASED ON STATISTICAL ANALYSIS OF METEOROLOGICAL VARIABLES: A CASE STUDY IN PERNAMBUCO/BRAZIL

V. N. Silva¹, V. A. Santos², B. Mohand³, L. A. Sarubbo⁴

¹ Post-Graduate Program in Development of Processes, Catholic University of Pernambuco, Recife, Pernambuco, Brazil

² Centre of Science and Technology, Catholic University of Pernambuco, Recife, Pernambuco, Brazil

³ Centre of Technology and Geoscience, Federal University of Pernambuco, Recife, Pernambuco Brazil

⁴ Centre of Science and Technology, Catholic University of Pernambuco, Recife, Pernambuco, Brazil

ABSTRACT

In the Northeast Region of Brazil, especially in the State of Pernambuco, detected a correlation between weather variables and human activities, be they economic, social, cultural, etc. The aim of paper was analyze the application of the statistical technique of multivariate Principal Component Analysis (PCA) on the study of case in state of Pernambuco. In confection a diagnosis using weather variables and socioeconomic responses (HDI and GDP) of some municipalities that presented a significant development in the State of Pernambuco - Brazil, between the years of 1999 and 2013. Even seeing the different climate characteristics and socioeconomic characteristics of the cities studied. The PCA proved to be effective and allowed the removal or disposal of variables that had low variability, or were redundant because they are correlated with those of greatest importance for two main components To understand the behavior of the meteorological variables and their aftereffect on human activities is of fundamental importance in aid to public politics, aimed at mitigation of adverse effects caused by environmental changes.

Keywords: Weather Variables; Socioeconomic Indicators; Principal component analysis.

1. Introduction

Recent studies demonstrate that in various regions of the planet, the behavior of the meteorological variables are influencing on socioeconomic aspects in various ways as an example Ebi e Nealon (2016), We was studied the correlation between the variability of rainfall and temperature with the cases of transmission of dengue by the mosquito *Aedes aegypti* in the tropical region of the planet. The study of Mehta., *et al* (2013), related to climate variability of the interdecadal rainfall and its consequences in agriculture and water resources of the basin of the Missouri River, in the United States.

The climate system of a determine region or locality are constituted by climatic factors (air masses, relief, latitude, maritimidade,) and the meteorological variables (temperature, rainfall, moisture from the air, etc.), who exercise influence on socioeconomic aspects of certain spaces (Kagan, 1995). The climatic conditions on material a certain portion of the earth's surface are the result of an outstanding interaction between the meteorological variables and the climatic factors. Solar activity as a factor, for example, influences the climate variability in the short and long term being decisive for human activities in various regions of the planet as analyzed by (Ineson, *et. al*, 2015).

The main indicators of socioeconomic aspects of a region (country, state, county), can be measured through the values of Product Internal Brute (PIB), which represents the somatory of monetary values, of all goods and services produced and the Human Development Index (IDH), in Brazil is represented every year by the Index of Municipal Development (IDM) is a annual study created to accompany the human development, economic and social development of all Brazilian municipalities, based exclusively on official statistics (IBGE, 2016).

The State of Pernambuco is composed of 185 municipalities, and as a result of their training planning these municipalities feature several characteristics physiographic and socioeconomic characteristics which distinguishes them, both because of its geographical location as the economic activities developed along the story (CONDEPE/FIDEM, 2016). In the face of current social and environmental contexts in which is the condition, we sought to conduct this study to analyze, by means of a modeling, between the data meteorological (precipitation, air temperature, relative humidity, heat, Evaporatio) with socioeconomic data of PIB (agricultural, industrial and services) and the IDM (Employment and Income, Health and Education) in the municipalities of significant economic development of the State of Pernambuco (Petrolina, Arcoverde , Garanhuns, Surubim and Recife), between the period from 1999 to 2013.

With a very large quantity of variables involved in the modeling process, the number of relations between them is also too big to be understood. Thus, it becomes necessary to the use of techniques that can be summed, systematically, the significant correlation exists between the variables (Malekinezhad, 2014). Among the techniques multivariate statistical, associated to the reduction of data or structural simplification, what stands out is the Principal Component Analysis (PCA). It is a statistical technique of multivariate analysis that consists in reducing the dimensionality of the data, so that the most variation in the own set of data to be retained, providing a better interpretation (Hongyu, 2015).

Despite the relevance of the topic, the research are still very limited, by the lack of employment of statistical tools to justify the scientific character of the statements contained in the various studies cited. Besides that, Requires that the tool employed to the junction of different variables can be analyzed simultaneously. This would lead to the possibility of arriving at a diagnosis, by physiographic region, for application in various parts of the planet.

2. Materials and methods

2.1 Study area

According to data from the Brazilian Institute of Geography and Statistics - IBGE, the Pernambuco State is located in the Eastern part of the Northeastern Brazil, between the geographic coordinates $07^{\circ}15'45''$ / $09^{\circ}28'18''$ S e de $034^{\circ}48'35''$ / $041^{\circ}19'54''$ W, possessing an area of $98,146.3 \text{ km}^2$ of extension and its estimated populational in 2016 is 9,410,336 in habitants (figure 1). The state of Pernambuco it bordres in the north, with the states of Paraíba and Ceará; to the south with The States of Alagoas and Bahia; to the east, with the Atlantic Ocean; and to the west, with the state of Piauí (IBGE, 2016).

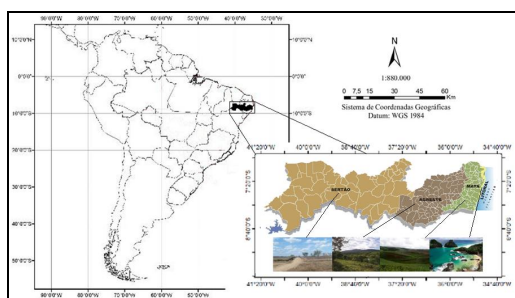


Figure 1. Brazil, highlight the state of Pernambuco, different natural regions

2.2 Physiographic aspects and socioeconomic factors

The State of Pernambuco displays from the coast to the interior, a succession of landscapes and differentiated forms of organization of space, there have been its territory divided into three physiographic regions: Coast – Zone Atlantic, Agreste and Semiarid Region. The territory of the State, 64% belongs to the Sertão, 25% of the Agreste and 11% to the Coast – Zone forest, being that each of these regions has both climatic characteristics distinct, and socioeconomic characteristics. Thus, we selected the municipalities of greater socioeconomic expression of each region Physiographic: Petrolina (Sertão of San Francisco); Arcoverde (Sertão of Moxotó); Garanhuns (Agreste Southern); Surubim (Agreste Northern) and Recife (coast), (CONDEPE/FIDEM, 2016).

Pernambuco has an estimated population of 9,410,336 inhabitant and currently, its economy is based on agriculture, industry and services. The services sector is dominant, followed by the naval industry, automotive, chemical, metallurgy, electronics, minerals non-metals, textile and food industries. The state is witnessing a significant change in its economic profile with the recent investments in the petrochemical, biotechnology, pharmaceutical, information technology, marine and automotive, we are giving new impetus to the economy of the state which has been growing above the national average (SDEC-PE, 2016).

The municipalities of Petrolina and Arcoverde, which are located in the region with the lowest rainfall in the state, the sertão region, where the economy and basically geared toward the activities of trade in goods and services. In Petrolina is the irrigated agriculture, from the São Francisco River. In Arcoverde, in period of drought extended to agriculture and almost non-existent (CONDEPE/FIDEM, 2016). The municipalities of Garanhuns and Surubim, the region Agreste of status, have an economy based on the dairy industry and the activities of trade in goods and services. In Garanhuns are quite developed the activity of cattle (milk and leathern) and tourism. Surubim has as highlights, the agricultural sector (sugar cane production, milk, banana, cattle and horticulture) and services (CONDEPE/FIDEM, 2016). The city of Recife is the capital, largest and main urban area of the state, was the main financial center of colonial Brazil until the mid 18th century, when the end of the sugar cane cycle. Currently stands to the activities of the manufacturing industry, construction and trade (IBGE, 2016).

2.3 Principal Component Analysis (PCA)

Principal component analysis (PCA) is a statistical technique of multivariate modeling of the structure of covariance. It is associated with the idea of reduction of mass of data, with less loss of information; however it is important to have a joint vision of all or almost all of the techniques of multivariate statistics to solve most of the practical problems. It seeks to redistribute the variation observed in the original shafts in order to obtain a set of orthogonal axes are not correlated (HONGYU, 2015).

The PCA has as main advantages: Remove the multicollinearity of the variables, since it allows transforming a set of original variables Inter-related in a new set of variables not correlated (principal components). In addition, reduces many variables to axes that represent some variables, these being perpendicular axes (orthogonal) explaining the variation of the data in order of decreasing and independent (HONGYU, 2016).

To obtain the key components of a general form, is a set of p variables $X_1, X_2, \dots, X_p$ with Medium $\mu_1, \mu_2, \dots, \mu_p$ and variance $\sigma_1^2, \sigma_2^2, \sigma_3^2, \dots, \sigma_p^2$, respectively. These variables are not independent and therefore, have covariance between the i -th and k -th variable is defined by σ_{ik} to $i \neq k = 1, 2, \dots, p$. Then the p variables can be expressed in the form of vector $\mathbf{X} = [X_1, X_2, \dots, X_p]'$, with vector of medium $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]'$ and the covariance matrix $\boldsymbol{\Sigma}$.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11}^2 & \cdots & \sigma_{1p}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{n1}^2 & \cdots & \sigma_{nn}^2 \end{bmatrix}$$

Are the pairs of eigenvalues and eigenvectors $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2), \dots, (\lambda_p, \mathbf{e}_p)$, in which $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$, Associated with $\boldsymbol{\Sigma}$ and then the i -th principal component is defined by (HONGYU, 2016):

$$\mathbf{Z}_i = \mathbf{e}_i' \mathbf{X} = e_{i1}X_1 + e_{i2}X_2 + \dots + e_{ip}X_p$$

The variable $\mathbf{Z}_i$ it is a latent variable, i.e., it is measured from the experiment or sample survey, The idea is to design the original coordinated in a plan by maximizing the distance between them, which is equivalent to maximize the variability of the latent variable $\mathbf{Z}_i$. The variance is $\mathbf{Z}_i$ given by:

$$\text{Var}(\mathbf{Z}_i) = \text{Var}(\mathbf{e}'_i \mathbf{X}) = \mathbf{e}'_i \text{Var}(\mathbf{X}) \mathbf{e}_i = \mathbf{e}'_i \boldsymbol{\Sigma} \mathbf{e}_i$$

Using the spectral decomposition of the matrix $\boldsymbol{\Sigma}$, Given by $\boldsymbol{\Sigma} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}'$, where $\mathbf{P}$ is the matrix composed by eigenvectors of $\boldsymbol{\Sigma}$ in its columns and $\boldsymbol{\Lambda}$, the matrix diagonal of eigenvalues of $\boldsymbol{\Sigma}$, then, is that...

$$\text{tr}(\boldsymbol{\Sigma}) = \text{tr}(\mathbf{P}\mathbf{A}\mathbf{P}') = \text{tr}(\mathbf{A}\mathbf{P}'\mathbf{P}) = \text{tr}(\mathbf{A}\mathbf{I}) = \text{tr}(\mathbf{A}) = \sum_{i=1}^p \lambda_i$$

e

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_k \end{bmatrix}$$

$\text{Etr}(\boldsymbol{\Sigma})$ it is given by the sum of the elements of the diagonal:

$$\text{tr}(\boldsymbol{\Sigma}) = \sum_{i=1}^p \sigma_{ii} = \sum_{i=1}^p \lambda_i$$

Therefore, the variable price contained in the original variaves is equal to total variability condita in the main components (JOHNSON; WICHERN, 1998). The contribution of each main component ($\mathbf{Z}_i$) is expressed in percentage, and the explanation of each individual component can be calculated by the criterion suggested by (KAISER (1958) this is, this criterion tends to include a few components when the number of original variables is less than twenty and, in general, it uses those components that are able to synthesize a cumulative variance of around 70%, the principal components that explain most of the variation in the data set (SAVEGNAGO *et al.*, 2011).

The interpretation of a principal component is made by the degree of importance, or, even, the influence that each variable has on each component, and that this importance is given by the correlation between each variable jX And the component iY which is being interpreted (REGAZZI, 2001).

Through charts to explain the proportion of variation of each main component, **scree plot**, and can identify the percentage of variation explained by each principal component. Already for a graphic **biplot** or circle of correlations, the idea is to add information about multiple variables in a single graph obtained by major components (KLEFENS, 2009).

3. Results And Discussion

In all the municipalities studied, the 1st and 2nd factor or CP were taken into account, where it was possible to show which meteorological variables most influenced the response variables (Gross Domestic Product and Municipal Development Index) of the municipalities studied. Using the same procedure Melo, Medeiros and Souza (2015) to analyze and obtain the spatial dependence of rainfall in the State of Paraíba.

The analysis of the most correlated variables in the 1st CP and 2nd CP showed that in the municipalities of Petrolina and Arcoverde the maximum temperature variable appears as the most significant, followed by relative humidity and insolation. In the municipalities of Garanhuns and Surubim, the insolation variable stands out because it presents correlated 1 ° CP of both municipalities. In the city of Recife, rainfall is the variable that contributes most to the 1st CP (Table 1).

Municipalities	Principal Component	Eigenvalues (PIB)	Accumulated proportion - PIB (%)	Eigenvalues (IDM)	Accumulated proportion - IDM (%)
Arcoverde	1ª CP	3,4	57,3	4,3	73,0
	2ª CP	0,96	73,4	0,84	87,1
Petrolina	1ª CP	3,0	50,2	3,1	52,8
	2ª CP	1,4	74,8	1,6	80,7
Garanhuns	1ª CP	2,5	41,9	2,4	40,5
	2ª CP	1,8	73,2	1,8	71,8
Surubim	1ª CP	3,0	50,2	3,9	65,0
	2ª CP	1,4	74,8	1,2	85,9
Recife	1ª CP	2,5	43,2	2,5	42,2
	2ª CP	1,7	71,9	1,9	75,4

Table 1. Values of the matrix of correlation Weather Variables x GDP/MDI

A selection of the main components using the Scree Plot graphical method (figure 2), shown in a general form in all municipalities analyzed, as main components and their systems (score - 1 ° CP) explained a larger portion of the variance of the original set of Data, which according to Kaiser's criterion (1958) corresponded values $\geq 70\%$ of the variance between CPs. Filho et al (2013) also used this criterion in the construction of socioeconomic indicators.

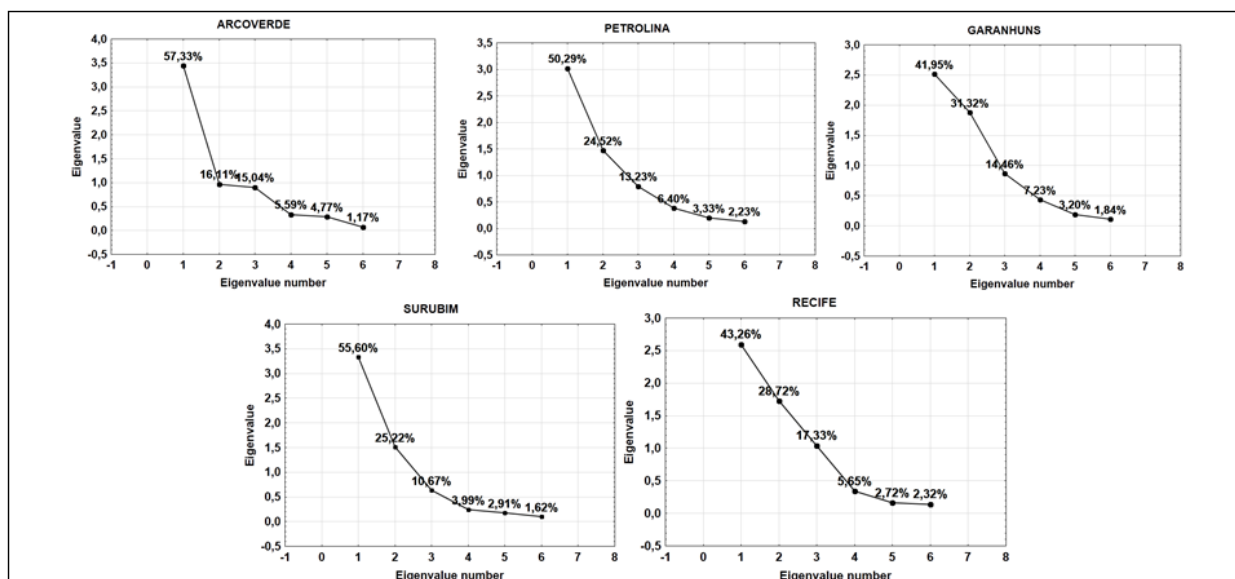


Figure 2. Weather Variables x PIB (1999-2013)

The curve in each graph shows a stabilization of the eigenvalues of the correlation matrix. That is, the factors with the greatest significance for the analysis of the data in the Scree plot are those with values close to 1 (figure 2)

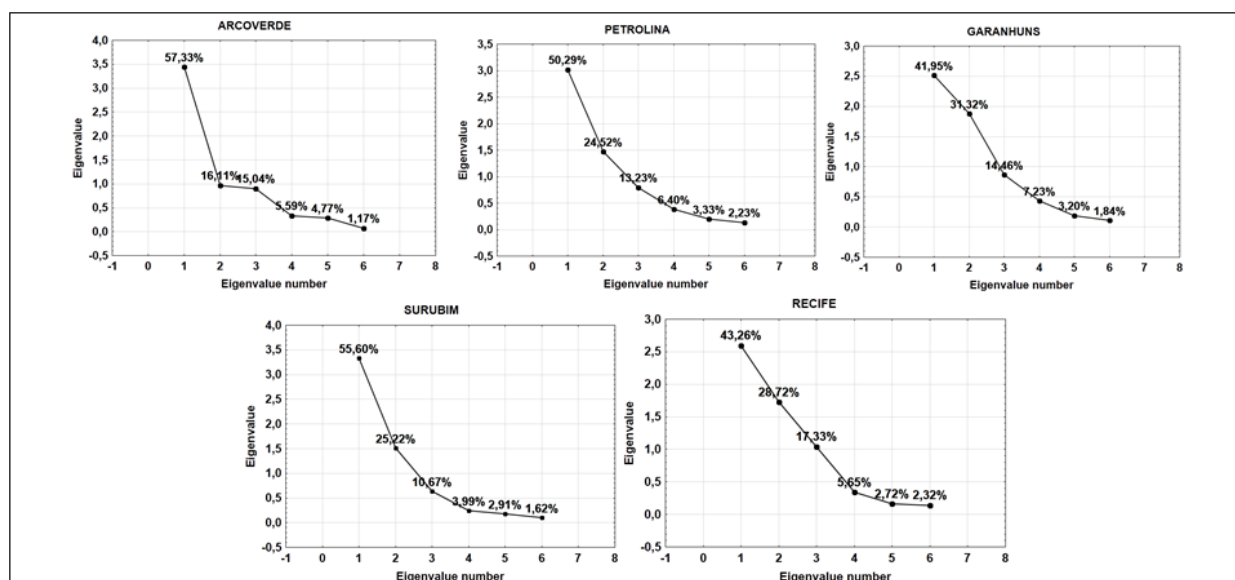


Figure 3. Weather Variables x IDM (2005-2013)

In Biplot charts, which represent, besides the scores of each factor, a way of visually identifying which original variables are related to the study, for Godoi, Hongyu and Dias (2012) the use of the Biplot method varies according to the objective and the interest of the Researcher, being very used in the area of improvement, mainly for the study of interaction between variables. In our study, the meteorological variables that are most related to socioeconomic variables.

Some variables are overlapping each other. This shows that these have the same representativity in the graph (figures 4). Another important fact is that some variables are very close to the unit circle. This shows that these have a greater contribution, in relation to the variables that are further away.

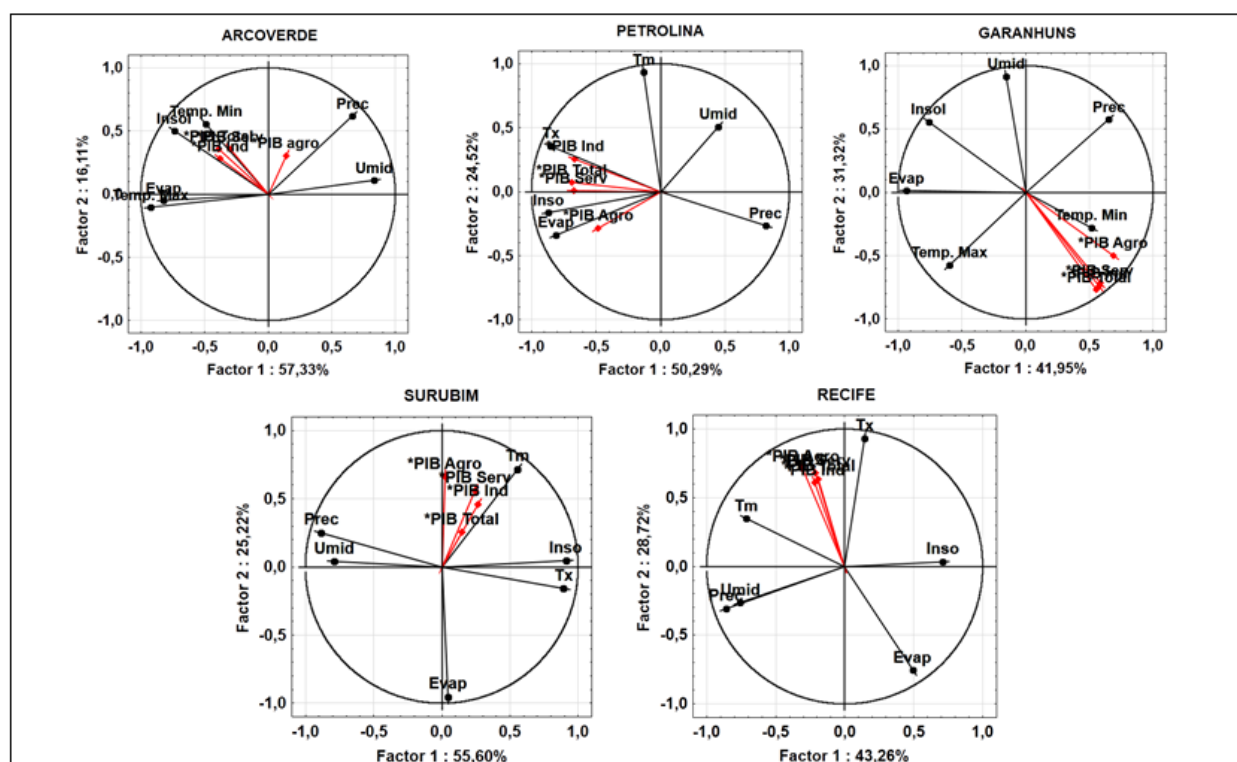


Figure 4. Biplot PC1 × PC 2 about the Weather Variables and PIB (1999-2013)

It was diagnosed in the municipality of Arcoverde, a great influence of the rainfall and humidity variable on the PIB - agriculture and livestock variable was observed. In the municipality of Petrolina, it was observed that rainfall and humidity do not influence GDP (agricultural, industrial and services), which is more correlated with maximum temperature, insolation and evaporation.

In the municipality of Garanhuns, the variable pluviometric precipitation and minimum temperature exert great influence, in the 1st factor, in all sectors of the GDP. In the municipality of Surubim, a greater influence was identified in the 1st factor of the variable minimum temperature and isolation on all sectors of GDP, being the agriculture with the highest correlation. In the city of Recife, it was verified that the meteorological variables in the 1st factor have little influence on the socioeconomic variables of the GDP. Figure 5 shows the results concerning the analysis of the meteorological variables with those of the IDM.

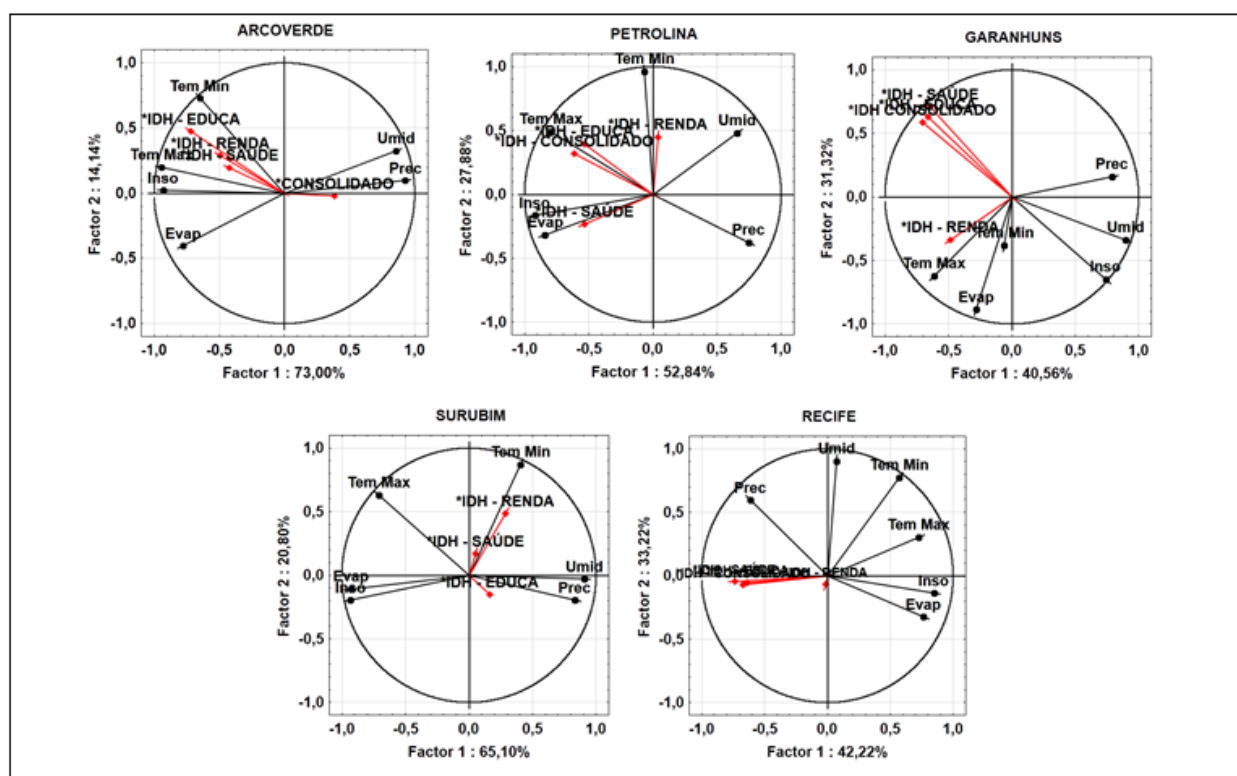


Figure 5. Biplot PC1 x PC 2 about the Weather Variables and IDM (2005-2013)

In Arcoverde, in general rainfall, the humidity has influence in the general IDM. On the other hand, the WDI related to income, health and education has a higher correlation with the behavior of the insolation variable, maximum and minimum temperature. In the municipality of Petrolina, it was observed that in relation to the IDM, health suffers more influence of the insolation and the evaporation as IDM income and education of the maximum and minimum temperatures.

In Garanhuns the variables maximum temperature and evaporation are related to the income-IDM. On the other hand, the general MDI, health and education do not correlate with

any meteorological variable (1st and 2nd factors). In the city of Recife, it was verified that the meteorological variables in the 1st factor have little influence on the socioeconomic variables. On the other hand, the IDM related to income, health and education has a higher correlation with the behavior of the rainfall variable.

4. Conclusions

Considering the results obtained, there is a relation between the behavior of the meteorological variables and the socioeconomic aspects of the data of each municipality analyzed. Not necessarily the quantity and the spatial distribution, but also in relation to the physiographic aspects of said municipalities. One can also infer a relationship between the conditions of annual climatic variability (dry and / or rainy years) with the economic and urban development of the respective municipalities.

Principal Component Analysis was effective for the proposed treatment. In all municipalities studied, the 1st factor of analysis, together with the 2nd factor, showed that the meteorological variables influenced, at different scales, the variables GDP and MDI in their different sectors. One of the objectives of the PCA in this case was reached, because a relatively small number of components were extracted (1 ° CP and 2 ° CP) with the capacity to explain the greater variability in the original data.

The ACP was effective and allowed the removal or disposal of variables that presented low variability or were redundant because they were correlated with those of greater importance for two main components. Thus, a smaller number of variables were necessary to explain the total variation resulting in saving of time and resources in future works that will use this same database, without significant loss of information.

Acknowledgements

We gratefully acknowledge the National Council for Scientific and Technological Development (CNPq) and the Federal Agency for the Support and Evaluation of Graduate Education (CAPES).

References

State Agency of Planning and Research of Pernambuco - CONDEPE/FIDEM. Pernambuco natural aspects Available in < <http://www.condepefidem.pe.gov.br>>, (accessed 08.12.16).

Ayoade, J. O. Introduction to climatology for the tropics. 4 ed. Rio de Janeiro: Bertrand Brasil. 1996, p.262,286. Available in < <http://pt.slideshare.net/lcgrh/ayoadejo-introduo-climatologia-para-os-trpicos-cpia>>. (accessed 06.12.16).

Ebi, K.L. Nealon. J. Dengue in a changing climate. Environmental Research. v. 151, n. 6, p. 115–123. 2016. DOI: <http://dx.doi.org/10.1016/j.envres.2016.07.026>

Hongyu, K. Comparison of the GGE-biplot Weighted and AMMI- With other models of interaction genotype $\times$ environmen. 2015. Thesis (Doctorate in Statistics and experimentation Agronomy) - The Escola Superior de Agricultura “Luiz de Queiroz”, University of São Paulo, Piracicaba, 2015. DOI. 10.11606 / T.11.2015.tde-04052015-172304.

Hongyu, et al. Principal Component Analysis: theory, interpretations and applications. E&S - Engineering and Science, v. 1, n. 5, p. 83-90. ISSN: 2358-5390 DOI: 10.18607/ES20165053.

Ineson, S. et al. Regional climate impacts of a possible future grand solar minimum. Nature Communications. Cambridge, v. 6, n. 7535, p. 1 – 8. 2015. DOI: <https://doi.org/10.1038/ncomms8535>

The Brazilian Institute of Geography and Statistics – IBGE . States of Pernambuco. Available in < <http://www.ibge.gov.br/estadosat/perfil.php?sigla=pe>> Acessado em (accessed 12.12.16).

Johnson, R. A.; Wichern, D. W. Applied multivariate statistical analysis. Fourth Edition New Jersey: Printice Hall, 1998. Disponivel em < <http://www1.udel.edu/oiss/pdf/617.pdf> > (accessed 11.12.16).

Kagan, B.A. Ocean-atmosphere interaction and Climate Modeling. Traduce. M. A. Chazin. Cambridge: Cambridge University Press, 1995. Available in < <https://goo.gl/6wYLtp> > (accessed 10.12.16).

Kaiser, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, v. 23, n. 3, p. 187-200, 1958. < <https://doi.org/10.1007/bf02289233> >

Klefens, P. C. The Biplot in factor analysis multivariate analysis. Dissertation (Master in Statistics and experimentation Agronomic) - The Escola Superior de Agricultura “Luiz de Queiroz”, University of São Paulo, Piracicaba, 2009. DOI.10.11606/D.11.2010.tde-25022010-085843.

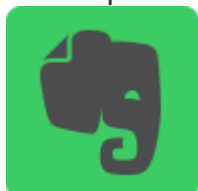
Malekinezhad, H. Regionalization Approach for Modeling of Monthly Evaporation Based on Cluster Analysis. *Natural Resources and Conservation*, v. 2, n. 2, p. 25-32, 2014. DOI: 10.13189/nrc.2014.020202

Regazzi, A. J Multivariate analysis. University of de Viçosa, Center for Exact and Technological Sciences. Department of Informatic , 2001. 166p. Available in <https://goo.gl/AeyhjM> > (accessed 10.12.16).

Submission Confirmation

AE

Atmospheric Environment <eesserver@eesmail.elsevier.com>



Responder|

Hoje, 02:35

Você;

vicentenael@gmail.com

Re: SOCIOECONOMIC DIAGNOSIS BASED ON STATISTICAL ANALYSIS OF METEOROLOGICAL VARIABLES: A CASE STUDY IN PERNAMBUCO/BRAZIL by VICENTE NATANAEL SILVA; VALDEMIR ALEXANDRE, PHD; Mohand Benachour, PHD; Wanderson Santos, M.Sc; Leonie Sarubbo, PHD Research Paper

Dear Dr SILVA,

Your submission entitled "SOCIOECONOMIC DIAGNOSIS BASED ON STATISTICAL ANALYSIS OF METEOROLOGICAL VARIABLES: A CASE STUDY IN PERNAMBUCO/BRAZIL" has been received by Atmospheric Environment

You may check on the progress of your paper by logging on to the Elsevier Editorial System as an author. The URL is <http://ees.elsevier.com/atmenv/>.

Your username is: vicentenael@hotmail.com

If you need to retrieve password details, please go to: http://ees.elsevier.com/atmenv/automail_query.asp

Your manuscript will be given a reference number once an Editor has been assigned.

Thank you for submitting your work to this journal.

Kind regards,

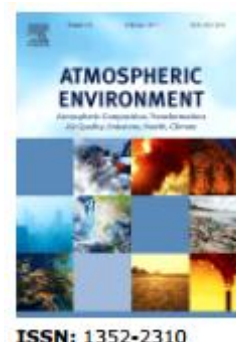
Elsevier Editorial System
Atmospheric Environment

For further assistance, please visit our customer support site at <http://help.elsevier.com/app/answers/list/p/7923>. Here you can search for solutions on a range of topics, find answers to frequently asked questions and learn more about EES via interactive tutorials. You will also find our 24/7 support contact details should you need any further assistance from one of our customer support representatives.

Normas Para Submissão no Periódico Internacional



ATMOSPHERIC ENVIRONMENT



GUIDE FOR AUTHORS

Your Paper Your Way We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article.

BEFORE YOU BEGIN

Ethics in publishing Please see our information pages on Ethics in publishing and Ethical guidelines for journal publication. Policy and ethics The work described in your article must have been carried out in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki) for experiments involving humans <http://www.wma.net/en/30publications/10policies/b3/index.html> EU Directive 2010/63/EU for animal experiments http://ec.europa.eu/environment/chemicals/lab_animals/legislation_en.htm Uniform Requirements for manuscripts submitted to Biomedical journals <http://www.icmje.org>. This must be stated at an appropriate point in the article

Declaration of interest - All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

Article structure - Subdivision - numbered sections.

Divide your article into clearly defined and numbered sections. Subsections should be numbered 1.1 (then 1.1.1, 1.1.2, ...), 1.2, etc. (the abstract is not included in section numbering). Use this numbering also for internal cross-referencing: do not just refer to 'the text'. Any subsection may be given a brief heading. Each heading should appear on its own separate line.

Introduction - State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Material and methods - Provide sufficient detail to allow the work to be reproduced. Methods already published should be indicated by a reference: only relevant modifications should be described.

Theory/calculation - A Theory section should extend, not repeat, the background to the article already dealt with in the Introduction and lay the foundation for further work. In contrast, a Calculation section represents a practical development from a theoretical basis.

Results- Results should be clear and concise.

Discussion - This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Conclusions- The main conclusions of the study may be presented in a short Conclusions section, which may stand alone or form a subsection of a Discussion or Results and Discussion section.

Appendices - If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

Essential title page information

Title. Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible. Locations should be contained in the keywords rather than in the title. Titles should not be sentences, questions or contain redundant, obscure or vernacular terms.

Author names and affiliations. Where the family name may be ambiguous (e.g., a double name), please indicate this clearly. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower-case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name, and, if available, the e-mail address of each author.

Corresponding author. Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. Ensure that telephone and fax numbers (with country and area code) are provided in addition to the e-mail address and the complete postal address. Contact details must be kept up to date by the corresponding author

Present/permanent address. If an author has moved since the work described in the article was done, or was visiting at the time, a "Present address" (or "Permanent address") may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Abstract - A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

Graphical abstract - Although a graphical abstract is optional, its use is encouraged as it draws more attention to the online article. The graphical abstract should summarize the contents of the

article in a concise, pictorial form designed to capture the attention of a wide readership. Graphical abstracts should be submitted as a separate file in the online submission system. Image size: Please provide an image with a minimum of 531 × 1328 pixels (h × w) or proportionally more. The image should be readable at a size of 5 × 13 cm using a regular screen resolution of 96 dpi. Preferred file types: TIFF, EPS, PDF or MS Office files. You can view Example Graphical Abstracts on our information site. Authors can make use of Elsevier's Illustration and Enhancement service to ensure the best presentation of their images and in accordance with all technical requirements: Illustration Service.

Highlights Highlights are mandatory for this journal. They consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate editable file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). You can view example Highlights on our information site.

Keywords Immediately after the abstract, provide a maximum of 6 keywords, using American spelling and avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.

References Citation in text - Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list they should follow the standard reference style of the journal and should include a substitution of the publication date with either 'Unpublished results' or 'Personal communication'. Citation of a reference as 'in press' implies that the item has been accepted for publication.

Reference links Increased discoverability of research and high quality peer review are ensured by online links to the sources cited. In order to allow us to create links to abstracting and indexing services, such as Scopus, CrossRef and PubMed, please ensure that data provided in the references are correct. Please note that incorrect surnames, journal/book titles, publication year and pagination may prevent link creation. When copying references, please be careful as they may already contain errors. Use of the DOI is encouraged.

A DOI can be used to cite and link to electronic articles where an article is in-press and full citation details are not yet known, but the article is available online. A DOI is guaranteed never to change, so you can use it as a permanent link to any electronic article. An example of a citation using DOI for an article not yet in an issue is: VanDecar J.C., Russo R.M., James D.E., Ambek W.B., Franke M. (2003).