

UNIVERSIDADE FEDERAL DO PARANÁ

SHEILA REGINA ORO

ÍNDICE DE MONITORAMENTO DO COMPORTAMENTO ESTRUTURAL DOS
BLOCOS DE CONCRETO DE BARRAGENS – UMA ABORDAGEM
MULTIVARIADA

CURITIBA

2016

SHEILA REGINA ORO

ÍNDICE DE MONITORAMENTO DO COMPORTAMENTO ESTRUTURAL DOS
BLOCOS DE CONCRETO DE BARRAGENS – UMA ABORDAGEM
MULTIVARIADA

Tese apresentada ao Programa de Pós-graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, na Linha de Pesquisa em Métodos Estatísticos Aplicados à Engenharia, dos Setores de Tecnologia e de Ciências Exatas, da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Ciências.

Orientador: Prof. Dr. Anselmo Chaves Neto

CURITIBA

2016

Oro, Sheila Regina

Índice de monitoramento do comportamento estrutural dos blocos de concreto de barragens – uma abordagem multivariada / Sheila Regina Oro. – Curitiba, 2016.

139 f. : il., tabs.

Tese (doutorado) – Universidade Federal do Paraná, Setores de Tecnologia e de Ciências Exatas, Programa de Pós-Graduação em Métodos Numéricos em Engenharia.

Orientador: Anselmo Chaves Neto

Bibliografia: p. 115-19

1. Barragens - Inspeção. 2. Monitoramento ambiental. 3. Análise multivariada. I. Chaves Neto, Anselmo. II. Título.

CDD 627.82



Ministério da Educação
Universidade Federal do Paraná
Setor de Tecnologia / Setor de Ciências Exatas
Departamento de Construção Civil / Departamento de Matemática/ Departamento
de Engenharia de Produção.
Programa de Pós-Graduação em Métodos Numéricos em Engenharia –
PPGMNE/UFPR.



TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em MÉTODOS NUMÉRICOS EM ENGENHARIA da Universidade Federal do Paraná foram convocados para realizar a arguição da Tese de Doutorado de **SHEILA REGINA ORO**, intitulada: "ÍNDICE DE MONITORAMENTO DO COMPORTAMENTO ESTRUTURAL DE BLOCOS DE CONCRETO DE BARRAGENS - UMA ABORDAGEM MULTIVARIADA ", após terem inquirido a aluna e realizado a avaliação do trabalho, são de parecer pela sua aprovação.

Curitiba, 11 de Março de 2016.



Prof ANSELMO CHAVES NETO (UFPR)

(Presidente da Banca Examinadora)



Prof INÁCIO ANDRUSKI GUIMARÃES (UTFPR)

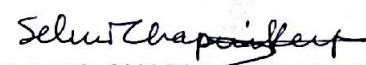


Prof JAIR MENDES MARQUES (UFPR)

Participação por vídeo conferência
Prof ROSANGELA VILLWOCK (UNIOESTE)



Prof SACHIKO ARAKI LIRA (UFPR)



Prof SELMO CHAPIRA KUPERMAN (USP)

Às minhas irmãs, Rosangela e
Karen, companheiras de todas
as horas.

AGRADECIMENTOS

Sem Deus eu nada seria e por isso eu agradeço a Sua presença em minha vida e esta graça alcançada.

Ao longo desta formação, ampliei meu conhecimento científico, vivenciei momentos de amor, comunhão, perdão, partilha e solidariedade, com pessoas especiais às quais sou grata por fazerem parte da minha história e cujos nomes cito a seguir.

Ao meu esposo Gilson, por permanecer pacientemente e amorosamente ao meu lado em todos os momentos, partilhando das alegrias e angústias.

Aos meus filhos Gabriel e Isabel e à minha enteada Júlia, por compreenderem minhas ausências, desculparem meu nervosismo e impaciência.

Aos meus pais Olevir e Diomar, por não medirem esforços para me ajudar e pelo amor incondicional.

À minha irmã Rosangela, pelas revisões gramaticais, ortográficas e pelos tantos momentos em que me ouviu e me falou.

Ao meu sobrinho Marcos, pela hospitalidade e auxílio nas traduções.

Aos demais familiares, por todo o auxílio, incentivo e compreensão.

Ao meu estimado orientador, Prof. Dr. Anselmo Chaves Neto, pelos ensinamentos, prontidão no atendimento, liberdade, generosidade e por acreditar na minha capacidade.

Aos professores do Dinter UFPR - Unioeste, em especial à Prof^a Liliana M. Gramani, pelo conhecimento transmitido e oportunidade concedida.

Ao PTI - CEASB pela oportunidade e disponibilização dos dados.

Ao engenheiro Cláudio Neumann Júnior, engenheiro de Itaipu, pela disponibilidade no atendimento presencial e à distância, por suas explicações didáticas e correções do texto.

Aos meus colegas "dinterandos" pelo companheirismo, incentivo e colaboração. Em especial: ao Geraldo, pela intervenção no momento oportuno; às amigas Tereza, Eliete, Fabiana e Suellen, por me acolherem em suas casas, me ouvirem, incentivarem e pela companhia nas horas de estudo. Ao Loreci e ao Aureo, pelas aulas de monitoria, caronas e paciência.

RESUMO

Este trabalho enfoca o monitoramento das respostas das estruturas de concreto dos blocos de contrafortes do Trecho D da barragem de Itaipu, por meio da análise das séries dos sensores de monitoramento de deslocamentos das estruturas e fundações de dois blocos da barragem e as séries temporais das condições ambientais no entorno da barragem. As técnicas de análise multivariada foram utilizadas para estudar as relações entre os deslocamentos e as condições ambientais, delimitadas pela temperatura superficial do concreto, temperatura ambiente e nível de água do reservatório. A análise de correlação canônica foi usada para avaliar a influência das variáveis ambientais nos deslocamentos das estruturas e fundações da barragem. A aplicação da análise fatorial objetivou a identificação das fontes de variabilidade dos dados e a ordenação dos sensores de acordo com a ação dos fatores. As datas das medições foram agrupadas conforme as similaridades presentes nas observações, por meio da aplicação da análise de agrupamentos. Em seguida, a análise discriminante foi usada para avaliar os grupos quanto à sua homogeneidade. Os resultados indicaram que as técnicas utilizadas permitem distinguir as respostas da barragem e identificar os efeitos das variações das condições ambientais sobre os deslocamentos das estruturas e fundações da barragem. Então, utilizando esses resultados como dados de entrada, criou-se o Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem (IMCRB), considerando a ação da temperatura ambiente e do nível de água do reservatório, com a finalidade de contribuir para o diagnóstico de mudanças no padrão de comportamento estrutural da barragem. Os valores obtidos para o IMCRB indicaram que o processo monitorado estava sob controle, era estável e que, portanto, as respostas da barragem eram previsíveis. Dados atualizados foram utilizados para a validação do método.

Palavras-chave: Monitoramento de Barragens. Deslocamentos. Condições Ambientais. Análise Multivariada. IMCRB.

ABSTRACT

This work is focused on the monitoring of the responses of the concrete structures of the buttresses blocks of Sector D of the Itaipu dam, through the analysis of series of monitoring sensors of structural displacement and foundations of two blocks of the dam and the time series of environmental conditions in the dam surroundings. Multivariate analysis techniques have been used to study the relationship between displacement and environmental conditions, defined by the surface temperature of the concrete, ambient temperature and reservoir water level. The canonical correlation analysis was used to evaluate the influence of environmental variables on the displacement of structures and dam foundations. The application of factor analysis intended to identify the sources of variability in data and the ordering of the sensors according to the action of factors. The dates of the measurements were grouped according to the similarities present in the observations, by applying cluster analysis. Thereafter, the discriminant analysis was used to evaluate the groups as to their homogeneity. The results indicated that the techniques used allow the distinguish of the dam responses and identification of the effects of variations in environmental conditions on the displacements of structures and dam foundations. Then, using these results as input data, the Joint Monitoring Index Blocks Responses (JMIBR) of dam, considering the action of temperature and water level of the reservoir was created in order to contribute to the diagnosis of changes in the dam's structural behavior pattern. The values obtained for JMIBR indicated that the monitored process was under control, stable and, thus, the dam responses were predictable. Updated data were used for method validation.

Keywords: Monitoring Dams. Displacements. Environmental Conditions. Multivariate Analysis. IMCRB.

LISTA DE FIGURAS

FIGURA 1 –	TIPOS DE BARRAGENS	20
FIGURA 2 –	ESBOÇO DE UMA BARRAGEM DE CONTRAFORTES	21
FIGURA 3 –	ESQUEMA DO SISTEMA DE UMA BARRAGEM DE CONCRETO	22
FIGURA 4 –	CORRELAÇÃO ENTRE TIPOS DE INSTRUMENTOS E DETERIORAÇÃO DE BARRAGENS DE CONCRETO	26
FIGURA 5 –	ESQUEMA DE INSTALAÇÃO DE UM EXTENSÔMETRO MÚLTIPLO DE HASTES	27
FIGURA 6 –	ESQUEMA DE INSTALAÇÃO DE PIEZÔMETRO <i>STANDPIPE</i> EM UM FURO DE SONDAGEM	29
FIGURA 7 –	ESQUEMA DE INSTALAÇÃO DE PÊNDEULOS DIRETO E INVERTIDO	30
FIGURA 8 –	ESQUEMA DE UM TERMÔMETRO PARA CONCRETO	32
FIGURA 9 –	ESQUEMA DE UMA BASE PARA MEDIDOR DE JUNTAS (ALONGÂMETRO)	33
FIGURA 10 –	INTERAÇÃO ENTRE EQUIPES NA AVALIAÇÃO DA SEGURANÇA DE BARRAGENS	36
FIGURA 11 –	LÓGICA ORIGINAL DO SOAA	38
FIGURA 12 –	ARRANJO GERAL DA USINA HIDRELÉTRICA DE ITAIPU	40
FIGURA 13 –	BLOCO CHAVE D7 DO TRECHO D DA BARRAGEM DE ITAIPU	41
FIGURA 14 –	BLOCO CHAVE D8 DO TRECHO D DA BARRAGEM DE ITAIPU	42
FIGURA 15 –	FLUXOGRAMA DO MÉTODO	77
FIGURA 16 –	EXEMPLOS DE DISPOSIÇÃO DOS VALORES DO IMCRB NA CARTA DE CONTROLE	86
FIGURA 17 –	SÉRIE TEMPORAL DO SENSOR X20	89
FIGURA 18 –	GRÁFICOS UTILIZADOS NA IDENTIFICAÇÃO DOS VALORES ATÍPICOS PARA O NÍVEL DE ÁGUA DO RESERVATÓRIO	90
FIGURA 19 –	GRÁFICO DE DISPERSÃO DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS	92
FIGURA 20 –	CORRELOGRAMA DOS 42 SENSORES	95
FIGURA 21 –	RESULTADO DO TESTE DE BARTLETT	96
FIGURA 22 –	RESULTADO DO CÁLCULO DO COEFICIENTE KMO	97
FIGURA 23 –	<i>SCREE PLOT</i> DOS AUTOVALORES DA MATRIZ DE CORRELAÇÕES AMOSTRAIS	97
FIGURA 24 –	VISTA FRONTAL (AÉREA) DO TRECHO D DA BARRAGEM DE ITAIPU	101
FIGURA 25 –	AGRUPAMENTOS PELO MÉTODO DA LIGAÇÃO MÉDIA, UTILIZANDO A DISTÂNCIA DE MAHALANOBIS	102
FIGURA 26 –	GRÁFICO DE DISPERSÃO DOS ELEMENTOS DE ACORDO COM AS FUNÇÕES DISCRIMINANTES ENTRE GRUPOS	103

FIGURA 27 –	SÉRIES TEMPORAIS DOS PRINCIPAIS FATORES QUE INFLUENCIARAM OS DESLOCAMENTOS DAS ESTRUTURAS DE CONCRETO DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU, NO PERÍODO DE JAN/90 A DEZ/13	105
FIGURA 28 –	SÉRIES DIFERENCIADAS DOS PRINCIPAIS FATORES QUE INFLUENCIARAM OS DESLOCAMENTOS DAS ESTRUTURAS DE CONCRETO DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU, NO PERÍODO DE JAN/90 A DEZ/13	106
FIGURA 29 –	CARTA DE CONTROLE PARA O ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU	107
FIGURA 30 –	CARTA DE CONTROLE PARA A SÉRIE DIFERENCIADA DO ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU	108
FIGURA 31 –	CARTA DE CONTROLE PARA O ÍNDICE IMCRB CONTENDO A SÉRIE HISTÓRICA E A PREVISÃO DOS VALORES FUTUROS	109
FIGURA 32 –	CARTA DE CONTROLE PARA O ÍNDICE IMCRBDIF CONTENDO A SÉRIE HISTÓRICA E A PREVISÃO DOS VALORES FUTUROS	110
FIGURA 33 –	VERIFICAÇÃO DA NORMALIDADE DOS RESÍDUOS DO MODELO ÁRIMA AJUSTADO À SÉRIE TEMPORAL DO IMCRB	110
FIGURA 34 –	CARTA DE CONTROLE ATUALIZADA PARA O ÍNDICE IMCRB	111
FIGURA 35 –	CARTA DE CONTROLE ATUALIZADA PARA O ÍNDICE IMCRBDIF	112

LISTA DE TABELAS

TABELA 1 –	VALORES DE CONTROLE PARA OS EXTENSÔMETROS DO TRECHO D DA BARRAGEM DE ITAIPU	43
TABELA 2 –	VALORES DE CONTROLE PARA OS PÊNDEULOS DIRETOS E INVERTIDOS DO TRECHO D DA BARRAGEM DE ITAIPU.	43
TABELA 3 –	VALORES DE CONTROLE PARA AS BASES DE ALONGÂMETRO (ABERTURAS) DO TRECHO D DA BARRAGEM DE ITAIPU	44
TABELA 4 –	VALORES DE CONTROLE ESTIMADOS PARA OS EXTENSÔMETROS DO TRECHO D DA BARRAGEM DE ITAIPU	45
TABELA 5 –	VALORES DE CONTROLE ESTIMADOS PARA OS PÊNDEULOS DO TRECHO D DA BARRAGEM DE ITAIPU, NO SENTIDO DO FLUXO	45
TABELA 6 –	NÍVEIS DE ATENÇÃO PARA OS PIEZÔMETROS DO BLOCO D8 DO TRECHO D DA BARRAGEM DE ITAIPU	47
TABELA 7 –	CORRELAÇÃO CANÔNICA ENTRE OS CONJUNTOS DES E AMB	91
TABELA 8 –	VARIÁVEIS FORTEMENTE CORRELACIONADAS	93
TABELA 9 –	PRINCIPAIS CORRELAÇÕES ENTRE O PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS E SENSORES DE CADA GRUPO..	93
TABELA 10 –	MEDIDAS DOS MAIORES AUTOVALORES, COM AS CORRESPONDENTES PERCENTAGENS DAS VARIÂNCIAS EXPLICADAS E VARIÂNCIAS ACUMULADAS..	98
TABELA 11 –	CARGAS FATORIAIS, COMUNALIDADES E VARIÂNCIAS ESPECÍFICAS DOS SENSORES	99
TABELA 12 –	CLASSIFICAÇÃO DAS DATAS DAS MEDIÇÕES EM TRÊS GRUPOS	103

LISTA DE QUADROS

QUADRO 1 –	DESCRIÇÃO DOS PRINCIPAIS SUBSISTEMAS DO SOAA ...	38
QUADRO 2 –	PRECISÃO DOS PRINCIPAIS MÉTODOS DE EXTRAÇÃO DE FATORES	58
QUADRO 3 –	FENÔMENOS MONITORADOS PELOS INSTRUMENTOS E SEUS RESPECTIVOS SENSORES	79
QUADRO 4 –	CLASSIFICAÇÃO DO PROCESSO MONITORADO E AS POSSÍVEIS CAUSAS ASSOCIADAS À OCORRÊNCIA DE ANORMALIDADES	85
QUADRO 5 –	SENSORES QUE APRESENTARAM SÉRIES INCOMPLETAS	88

LISTA DE SIGLAS

ADAS: Sistema de Aquisição de Dados Automáticos
AIC: Critério de Informação de Akaike
AMB: Condições Ambientais
APH: Análise Paralela de Horn
ARIMA: Modelo Autoregressivo Integrado Médias Móveis
ARMA: Modelo Autoregressivo Médias Móveis
CCN: Condição de Carregamento Normal
CDC: Sistema de Drenos de Concreto
CEASB: Centro de Estudos Avançados em Segurança de Barragens
CK: Critério de Kaiser
DES: Deslocamentos
FIS: Sistema de Fissuras
IMCRB: Índice de Monitoramento Conjunto das Respostas dos Blocos
KMO: Kaiser-Meyer-Olkin
MA: Médias Móveis
MD: Maquete Digital
MF: Movimentação da Fundação
MHBITA: Movimentação Horizontal dos Blocos por Influência Térmica Ambiental
MHRB: Movimentação Horizontal Relativa entre os Blocos D7 e D8
MHRBF: Movimentação Horizontal Relativa do Bloco D7 em Relação à Fundação
MMCPV: Média Mínima de Correlações Parciais de Velicer
MS: Mapa de Subpressão
PHI: Pressão Hidrostática
PTI: Parque Tecnológico Itaipu
SAA: Sistema Analítico de Auscultação
SAT: Sistema de Arquivo Técnico
SGDP: Sistema de Gestão de Dados Planimétricos
SGM: Sistema de Gestão de Medições
SOAA: Sistema de Otimização e Análise de Auscultação
SR: Sistema de Recomendações

SUMÁRIO

1	INTRODUÇÃO	15
1.1	PROBLEMA DE ESTUDO	15
1.2	OBJETIVOS	17
1.2.1	Objetivo Geral	17
1.2.2	Objetivos Específicos	17
1.3	JUSTIFICATIVA	18
1.4	ESTRUTURA DA TESE	19
2	REVISÃO DA LITERATURA	20
2.1	BARRAGENS	20
2.1.1	Segurança de Barragens	22
2.1.2	Instrumentação de Barragens	24
2.1.3	Monitoramento de Barragens de Concreto	33
2.1.4	Barragem de Itaipu	36
2.2	MÉTODOS ESTATÍSTICOS	47
2.2.1	Séries Temporais	48
2.2.2	Análise de Correlação Canônica	52
2.2.3	Análise Fatorial	54
2.2.4	Análise de Agrupamentos	60
2.2.5	Análise Discriminante	64
2.2.6	Cartas de Controle	67
2.3	MÉTODOS ESTATÍSTICOS APLICADOS NO MONITORAMENTO DE ESTRUTURAS	68
3	MATERIAL E MÉTODOS	76
3.1	SELEÇÃO DOS DADOS	78
3.2	ORGANIZAÇÃO E PROCESSAMENTO DOS DADOS	80
3.3	ANÁLISE DE CORRELAÇÃO CANÔNICA DE DESLOCAMENTOS E CONDIÇÕES AMBIENTAIS	81
3.4	ANÁLISE FATORIAL	82
3.5	MODELAGEM DO IMCRB	83

3.6	ANÁLISE DE AGRUPAMENTOS E ANÁLISE DISCRIMINANTE	86
4	RESULTADOS E DISCUSSÃO	88
4.1	PROCESSAMENTO DOS DADOS	88
4.2	INFLUÊNCIA DAS CONDIÇÕES AMBIENTAIS NOS DESLOCAMENTOS DE UMA BARRAGEM DE CONCRETO	91
4.2.1	Análise Fatorial Aplicada aos Dados da Instrumentação dos Blocos D7 e D8 do Trecho D da Barragem de Itaipu	94
4.2.2	Agrupamentos das Datas das Medições de Acordo com as Similaridades	102
4.3	ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS DA BARRAGEM	104
4.3.1	Formulação do Índice	104
4.3.2	Cartas de Controle dos Índices IMCRB e IMCRBdif	107
	CONCLUSÃO	113
	REFERÊNCIAS	115
	APÊNDICES	120
	APÊNDICE A – ALGORITMO DE GERAÇÃO DE MÉDIAS E IDENTIFICAÇÃO DE LACUNAS NO CONJUNTO DE DADOS	121
	APÊNDICE B – MEDIDAS DESCRITIVAS DAS VARIÁVEIS	123
	APÊNDICE C –TESTE DE BARTLETT E COEFICIENTE KMO NO <i>SOFTWARE R</i>	124
	APÊNDICE D – CLASSIFICAÇÃO RESULTANTE DA ANÁLISE DISCRIMINANTE DAS DATAS DE MEDIÇÃO	125
	APÊNDICE E – RESUMO DOS MODELOS ARIMA AJUSTADOS ÀS SÉRIES DO IMCRB E IMCRBdif	139

1 INTRODUÇÃO

1.1 PROBLEMA DE ESTUDO

A preocupação com a segurança estrutural de uma obra de grande porte é constante, iniciando na fase de projeto e durando por toda sua vida útil.

Na área de engenharia, o termo segurança estrutural refere-se à capacidade que a estrutura de uma obra tem de suportar todo o esforço a que está sujeita e toda espécie de ações desfavoráveis durante sua existência, sem atingir um estado limite de ruptura, mantendo assim sua estabilidade e funcionalidade para a qual foi projetada.

A NBR 8681 (2003, p. 15) estabelece que "a segurança das estruturas deve ser verificada em relação a todos os possíveis estados que são admitidos como limites para a estrutura considerada", no que diz respeito às condições analíticas e construtivas.

Os parâmetros utilizados para estabelecer as condições analíticas de segurança são oriundos de três naturezas: ações; esforços internos (solicitações, esforços solicitantes, tensões) e efeitos estruturais (deformações, deslocamentos, aberturas de fissuras). Por outro lado, para a verificação das condições construtivas de segurança, são consideradas as exigências definidas pelas normativas específicas para os tipos de materiais utilizados na construção das estruturas (NBR 8681, 2003).

No âmbito da segurança de barragens a Lei nº 12.334 (BRASIL, 2010, p. 1) define este termo como uma "condição que vise a manter a sua integridade estrutural e operacional e a preservação da vida, da saúde, da propriedade e do meio ambiente".

As estruturas de barragens de concreto são suscetíveis a uma gama de alterações provocadas pela incidência de deslocamentos, deformações, tensões, pressões, entre outros fenômenos. O monitoramento do comportamento dessas estruturas é realizado por meio da auscultação. A instrumentação utilizada nesse monitoramento pode ser composta por pêndulos, termômetros, bases de

alongâmetro, extensômetros, piezômetros, entre outros, instalados em pontos estratégicos da barragem, na época da sua construção. A tomada de decisão a respeito das condições de segurança das estruturas é baseada na análise dos numerosos gráficos das séries de dados dessa instrumentação, tendo como referência as condições analíticas e construtivas estabelecidas nos projetos das barragens.

Sendo assim, a equipe técnica designada para a avaliação das condições de segurança de uma barragem é responsável pelo estabelecimento, revisão e divulgação, sob a forma de relatórios, dos níveis de segurança em que se encontram as estruturas da barragem.

A avaliação da segurança estrutural de uma barragem pode suscitar algumas dúvidas quanto ao desempenho global da estrutura, frente à interação dela com o meio no qual está inserida. Eis algumas questões que podem surgir e que motivaram a realização deste estudo: como a estrutura da barragem de concreto responde às oscilações nas condições ambientais do seu entorno? Como as datas das medições referentes aos deslocamentos, temperaturas e nível de água do reservatório se relacionam com o comportamento dos dados? O que informam conjuntamente os dados dos diferentes instrumentos utilizados no monitoramento da barragem? As respostas a essas questões não são imediatas. Dependem, especialmente, da compreensão das respostas do sistema e da análise conjunta dos dados numéricos fornecidos pela instrumentação.

Nesse contexto, verifica-se a importância e a necessidade de desenvolvimento de um método de análise conjunta dos dados para auxiliar no diagnóstico global das condições de segurança estrutural da barragem. Este trabalho pretende contribuir com essa demanda, desenvolvendo um método para o monitoramento conjunto das respostas estruturais de uma barragem de concreto, a partir da base de dados reais da barragem de Itaipu.

Os dados oriundos das leituras manuais da instrumentação instalada nos blocos chave D7 e D8 da barragem receberam tratamento estatístico composto por um conjunto de modelos e técnicas, tais como Modelagem de Séries Temporais, Análise de Correlação Canônica, Análise Fatorial, Análise de Agrupamentos, Análise Discriminante, Intervalos de Confiança e Cartas de Controle, aplicadas convenientemente para a obtenção dos objetivos.

Em suma, neste trabalho pretende-se elaborar um método que permita avaliar globalmente as respostas de uma barragem de concreto e sirva de apoio para o diagnóstico das suas condições estruturais, utilizando técnicas de análise multivariada aplicadas às séries temporais da instrumentação, tendo como resultado um índice multivariado devidamente delimitado por intervalos de confiança e monitorado por uma carta de controle.

1.2 OBJETIVOS

1.2.1 Objetivo Geral

Desenvolver um método que permita a análise conjunta dos dados multivariados da instrumentação de uma barragem de concreto e das condições ambientais, que possa auxiliar o monitoramento estrutural e sirva de apoio para a tomada de decisão.

1.2.2 Objetivos Específicos

- Construir uma matriz de dados amostrais completa usando a modelagem de séries temporais;
- Estudar as relações entre as condições ambientais e os deslocamentos detectados pela instrumentação de uma barragem de concreto;
- Identificar os principais fatores que influenciam as respostas das estruturas e fundações de uma barragem de concreto;
- Identificar os sensores da instrumentação de maior relevância no que diz respeito à variabilidade dos dados;

- Realizar o agrupamento das datas das medições, de acordo com as similaridades.
- Elaborar um indicador multivariado das respostas das estruturas de uma barragem de concreto;
- Propor uma metodologia de controle e alerta baseado nos valores do indicador multivariado para indicar a existência de anormalidades e sua provável procedência.

1.3 JUSTIFICATIVA

A respeito da segurança estrutural da barragem de Itaipu, o parecer técnico, elaborado periodicamente pela equipe do Centro de Estudos Avançados em Segurança de Barragens - CEASB, fundamenta-se nas medições instrumentais, as quais geram mais de 2000 gráficos, que quando comparados aos valores especificados no projeto estrutural informam as condições estruturais da barragem. Até o momento as tomadas de decisões quanto ao aspecto dos dados e as análises das informações gráficas são feitas de forma empírica pelos membros da equipe de monitoramento.

A frequência de medições e a quantidade de instrumentos em uso, desde a época da construção da barragem, geram um enorme volume de informações a serem analisadas quanto à segurança da barragem. Por isso, considera-se necessária a concentração dessas informações, de forma a agilizar o processo de identificação de anomalias.

Nesse contexto, é de fundamental importância o desenvolvimento de modelos, baseados em métodos numéricos, que permitam a análise conjunta dos dados e auxiliem a identificação de anormalidades. Em especial, os métodos estatísticos podem ser muito úteis para esta tarefa, pois fornecem ferramentas para: o estudo das relações existentes nos conjuntos de dados; a identificação de fatores inerentes à variabilidade das observações; o estabelecimento de critérios de classificação de variáveis e atributos; o ajuste e a previsão de valores; o desenvolvimento de modelos multivariados; entre outros.

1.4 ESTRUTURA DA TESE

Esta tese está organizada em capítulos, sendo o primeiro a presente introdução.

O segundo traz uma revisão bibliográfica, contemplando os principais aspectos técnicos referentes ao monitoramento estrutural de barragens de concreto, descrevendo a teoria pertinente aos métodos estatísticos que nortearam o estudo, além da apresentação do estado da arte no que diz respeito à aplicação dessas técnicas no contexto do monitoramento de estruturas.

Os dados reais da instrumentação de monitoramento da barragem de Itaipu são apresentados no terceiro capítulo, juntamente com a descrição do método adotado para a modelagem e análise conjunta dos mesmos.

O quarto capítulo é dedicado à descrição e discussão dos resultados obtidos na aplicação do método, com a apresentação das ilustrações pertinentes.

As conclusões e as principais contribuições deste estudo, além das sugestões para trabalhos futuros, são apresentadas no último capítulo.

No final do texto encontram-se os Apêndices, que trazem a programação de algoritmos, a descrição de variáveis, a classificação dos atributos (datas das medições) e o resumo dos modelos ajustados às séries IMCRB e IMCRBdif.

2 REVISÃO DA LITERATURA

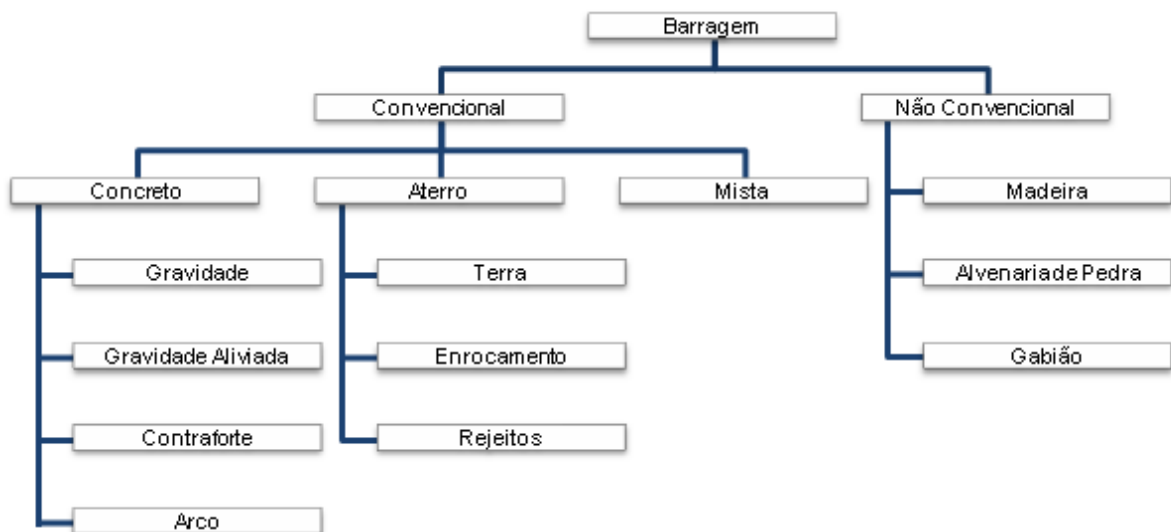
Neste capítulo são abordados os assuntos relacionados com a temática deste trabalho, bem como o referencial teórico que dá suporte à metodologia desenvolvida na execução do mesmo.

2.1 BARRAGENS

Barragem é "toda estrutura construída transversalmente a um rio ou talvegue com a finalidade de obter a elevação do seu nível d'água e/ou criar um reservatório de acumulação de água seja de regulação das vazões do rio, seja de outro fluido" (MSIB, 2002, p. 15).

Atualmente, há diversos tipos de barragens (Figura 1), caracterizadas de acordo com o projeto específico, a forma física, o objetivo e os tipos de materiais empregados na sua construção.

FIGURA 1 – TIPOS DE BARRAGENS

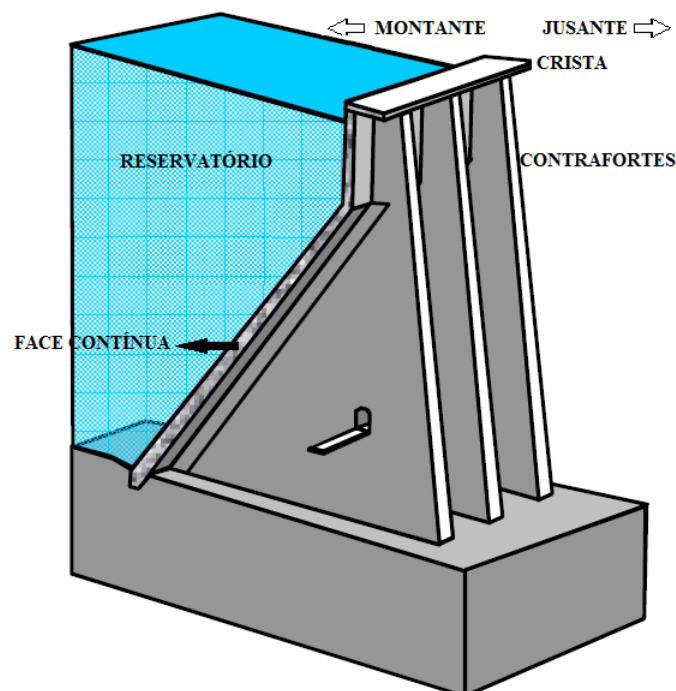


FONTE: A autora (2016).

A seguir, são apresentados alguns aspectos pertinentes à barragem de contrafortes, os demais tipos de barragens não são abordados devido ao escopo da pesquisa associada a este trabalho.

Por definição, contrafortes são pilares de sustentação. Uma barragem de concreto construída com blocos de contrafortes (Figura 2) apresenta estrutura contínua a montante, a laje, suportada a jusante por elementos descontínuos - os contrafortes.

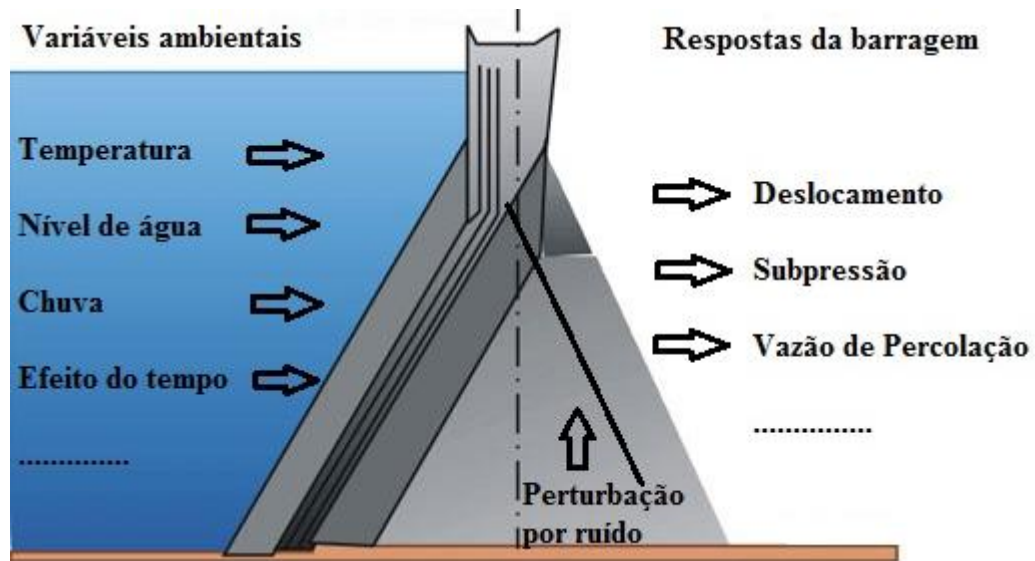
FIGURA 2 – ESBOÇO DE UMA BARRAGEM DE CONTRAFORTES



FONTE: Adaptado de EKSTROM (2009).

A relação entre uma barragem de concreto e o ambiente no qual ela está inserida pode ser representado por um sistema *input-output* (Figura 3), no qual as entradas (*input*) são as variáveis ambientais: temperatura ambiente no entorno da barragem, temperatura do concreto, temperatura da água, nível de água do reservatório, nível de água a jusante, chuva, atividade sísmica, tempo (idade da obra), entre outras; e as saídas (*output*) são as respostas da barragem, tais como: deslocamentos, subpressão, vazão de percolação e assim por diante (CHENG e ZHENG, 2013).

FIGURA 3 – ESQUEMA DO SISTEMA DE UMA BARRAGEM DE CONCRETO



FONTE: Adaptado de CHENG e ZHENG (2013).

De um modo geral, as barragens de contrafortes apresentam as seguintes características:

- Subpressão reduzida devido à pequena área da base;
- Maior compressão sobre a fundação;
- Exigem maior tratamento das fundações, como tirantes e injeção de calda de cimento;
- Maior economia de concreto, quando comparada aos outros tipos de barragem de concreto;
- Necessitam de estudos geológicos aprofundados;
- Maior emprego de armadura de aço, do que outros tipos de barragens de mesma altura.

2.1.1 Segurança de Barragens

Na legislação vigente o termo segurança de barragens é definido como uma "condição que vise a manter a sua integridade estrutural e operacional e a

preservação da vida, da saúde, da propriedade e do meio ambiente" (BRASIL, 2010, p. 1).

Art. 8º: O Plano de Segurança da Barragem deve compreender, no mínimo, as seguintes informações:

I - identificação do empreendedor;

II - dados técnicos referentes à implantação do empreendimento, inclusive, no caso de empreendimentos construídos após a promulgação desta Lei, do projeto como construído, bem como aqueles necessários para a operação e manutenção da barragem;

III - estrutura organizacional e qualificação técnica dos profissionais da equipe de segurança da barragem;

IV - manuais de procedimentos dos roteiros de inspeções de segurança e de monitoramento e relatórios de segurança da barragem;

V - regra operacional dos dispositivos de descarga da barragem;

VI - indicação da área do entorno das instalações e seus respectivos acessos, a serem resguardados de quaisquer usos ou ocupações permanentes, exceto aqueles indispensáveis à manutenção e à operação da barragem;

VII - Plano de Ação de Emergência (PAE), quando exigido;

VIII - relatórios das inspeções de segurança;

IX - revisões periódicas de segurança (BRASIL, 2010, p. 3-4).

Cabe ressaltar que a responsabilidade pela segurança da barragem, em todos os seus aspectos, é do seu proprietário. Este deve garantir que a operação da barragem e a sua manutenção sejam realizadas por pessoas treinadas e habilitadas para exercer as funções que lhes competem (MSIB, 2002).

A periodicidade da avaliação da segurança estrutural deve ser estabelecida com base na classificação de barragens, em termos de previsão quanto às consequências da ruptura. De acordo com a ANEEL (2015), essa periodicidade deve ser de 1, 2 ou 3 anos quando os níveis de risco e dano potencial associado forem considerados alto, médio ou baixo, respectivamente.

De acordo com Cruz (2006), quando a instrumentação apontar evidências de problemas que coloquem em risco a integridade da obra, é necessário colocar em prática um plano de medidas reparadoras viáveis tanto técnica como economicamente. Este plano, organizado previamente por uma equipe técnica responsável por monitorar a segurança da barragem, deve prever as prováveis situações críticas, estabelecer a priori as soluções técnicas, os recursos correspondentes, e indicar claramente os procedimentos administrativos, especialmente em situações de emergência. A esse respeito, MSIB (2002) afirma que:

Pouco adianta a leitura e análise de dados da instrumentação, à exceção de inspeções periódicas e à manutenção de arquivos com dados históricos de cada estrutura, se as medidas corretivas que se fizerem necessárias para restabelecer as condições de segurança não forem implementadas (MSIB, 2002, p. 14)

2.1.2 Instrumentação de Barragens

A respeito do aproveitamento das informações fornecidas pela instrumentação de uma barragem, deve-se considerar que:

... o valor da instrumentação não está associado apenas a obras que apresentam comportamentos não previstos, indicando a necessidade de medidas reparadoras ou acerto das hipóteses de projetos, mas também à indicação da ocorrência de condições seguras, mesmo em face das solicitações extremas atuantes ou da discordância das hipóteses de cálculo com a realidade (CRUZ, 2006, p. 609).

Além disso, de acordo com Cruz (2006), a importância da instrumentação transcende a avaliação das condições de segurança ao longo da vida útil de uma barragem, pois objetiva verificar:

- As hipóteses, os critérios e os parâmetros adotados em projeto, de modo a permitir o aprimoramento do projeto da própria obra em estudo, ou de futuras barragens, visando a condições mais econômicas e/ou mais seguras;
- A adequação de métodos construtivos;
- As condições de segurança das obras, de modo a serem adotadas medidas corretivas em tempo hábil, se necessárias.

No entanto, o mesmo autor aponta que a instrumentação apresenta limitações, especialmente pelos seguintes motivos:

- Alterações das condições locais na instalação do instrumento podem gerar valores falsos;
- Os instrumentos não apresentam os valores extremos de comportamento da estrutura, apenas os valores médios;
- Alguns medidores estão instalados em locais que não permitem a verificação cabal quanto ao funcionamento dos mesmos;

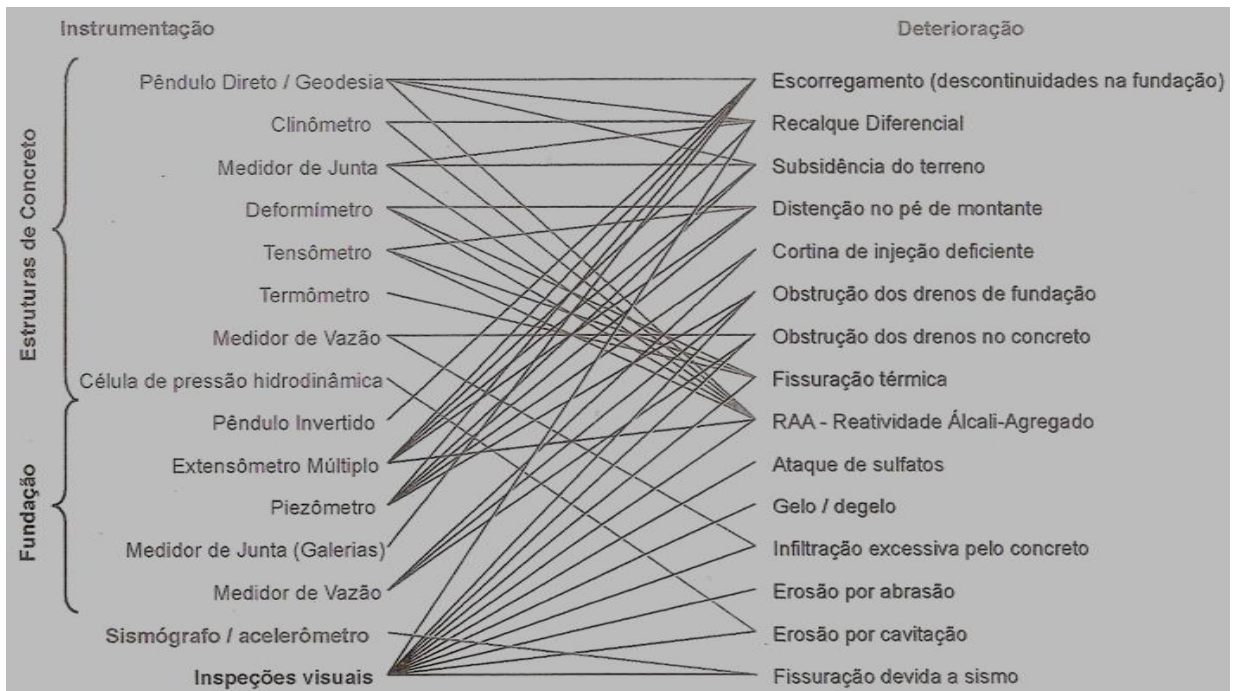
- Valores atípicos podem surgir e serem descartados por achar que o instrumento está com defeito; ou, também, aceitar valores normais de instrumentos defeituosos ou inadequados à finalidade pretendida;
- Alguns instrumentos são suscetíveis a cisalhamento se atravessam uma área submetida a deslocamentos concentrados;
- Falsa sensação de segurança decorrente da mera instalação de instrumentos, sem planejamento nem acompanhamento periódico.

Sendo assim, segundo Cruz (2006), a escolha dos instrumentos deve levar em conta as características desejáveis do mesmo, a saber:

- Confiabilidade;
- Alta durabilidade;
- Não provocar, durante ou após a instalação, alterações no valor da grandeza que pretende medir;
- Robustez;
- Alta precisão;
- Alta sensibilidade;
- Não ser influenciável por outras grandezas, que não as de interesse;
- Instalação simples;
- Não causar interferência na praça de trabalho;
- Baixo custo.

O tipo de instrumentação instalada nas estruturas e fundações de barragens depende do fenômeno que se quer monitorar. A Figura 4 traz um esquema representando os principais fenômenos monitorados e os respectivos instrumentos utilizados. Por exemplo, piezômetros e extensômetros, são instrumentos instalados na fundação da barragem e que permitem monitorar as subpressões, recalques e deslocamentos da barragem e de suas fundações; enquanto que pêndulos diretos e invertidos, deformímetros, termômetros superficiais, tensômetros múltiplos e medidores mecânicos de junta, permitem avaliar o desempenho estrutural dos blocos da barragem (DTI, 2009).

FIGURA 4 – CORRELAÇÃO ENTRE TIPOS DE INSTRUMENTOS E DETERIORAÇÃO DE BARRAGENS DE CONCRETO



FONTE: SILVEIRA (2003).

A seguir são apresentadas algumas informações referentes aos instrumentos considerados neste trabalho, sendo eles: extensômetros, piezômetros, pêndulos diretos e invertidos, bases de alongômetro e termômetros para concreto.

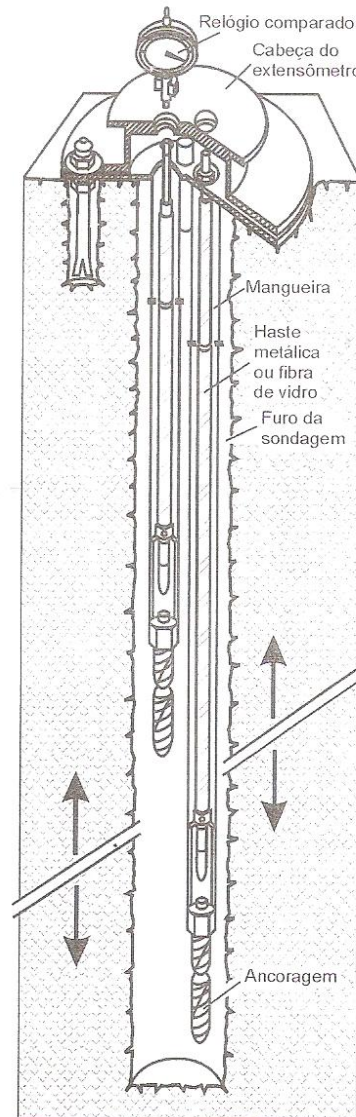
Os extensômetros múltiplos de hastes (Figura 5) possibilitam a medição de recalques e deslocamentos em diversas áreas da fundação da barragem, resultantes da ação do empuxo hidrostático, das influências térmicas ambientais entre outros condicionantes.

Estes instrumentos são instalados em furos de sondagens, na época da construção da barragem, empregam geralmente hastes de aço inox com diâmetro de 6 a 10 mm e de comprimento variável. A respeito da forma de instalação dos extensômetros, Silveira (2003) orienta o seguinte:

Para a medição dos recalques, os extensômetros múltiplos são instalados a partir de furos de sondagem verticais... Nos blocos da barragem onde se dispuser de galerias de acesso transversais ao eixo, pode-se prever a instalação de extensômetros a montante e a jusante, para a medição dos deslocamentos angulares da barragem junto à fundação. Em termos de profundidade, os extensômetros são instalados em furos correspondentes a 1/2 a 1/3 da altura da barragem, na seção instrumentada, que deve levar em consideração também as características geológicas da fundação, para

incluir as camadas de maior deformabilidade da fundação (SILVEIRA, 2003, p.224).

FIGURA 5 – ESQUEMA DE INSTALAÇÃO DE UM EXTENSÔMETRO MÚLTIPLO DE HASTES



FONTE: SILVEIRA (2003).

Para aferição dos deslocamentos das hastes dos extensômetros são utilizados relógios comparadores ou sensores de corda vibrante, de acordo com a precisão e tipo de leitura pretendida.

As principais características dos extensômetros múltiplos de hastes, de acordo com Cruz (2006), são:

- Possibilidade de medição de deslocamentos na direção do furo de sondagem onde se encontram chumbadas as hastes;

- Deslocamentos relativos cisalhantes de certa magnitude ao longo de descontinuidades no maciço rochoso, que interceptam o extensômetro e podem danificá-lo;
- Pode ser instalado facilmente em furos sub-horizontais até verticais;
- Permite avaliar a deformabilidade de partes isoladas do maciço rochoso;
- Leitura e cálculos rápidos e simples;
- Confiabilidade e durabilidade satisfatórias;
- Baixa dispersão de leituras;
- Leitura efetuada com relógio comparador com sensibilidade de centésimo de milímetro ou precisão maior;
- Dificuldade de instalação em furos inclinados para cima ou que apresentam vazão devida a artesianismo.

Outro instrumento instalado nas fundações de barragens é o piezômetro. Este tipo de instrumento permite a medição da subpressão atuante no local da sua instalação (DTI, 2009). Sobre a importância do monitoramento das subpressões na fundação de barragens de concreto, Silveira (2003) destaca que:

A observação das subpressões na fundação das barragens de concreto é de suma importância para a boa supervisão de suas condições de segurança, tendo em vista que a estabilidade dessas estruturas, em termos de escorregamento, tombamento ou flutuação, é diretamente afetada pelo nível das pressões piezométricas na interface concreto-rocha e nas descontinuidades sub-horizontais de baixa resistência existentes na fundação (SILVEIRA, 2003, p. 245).

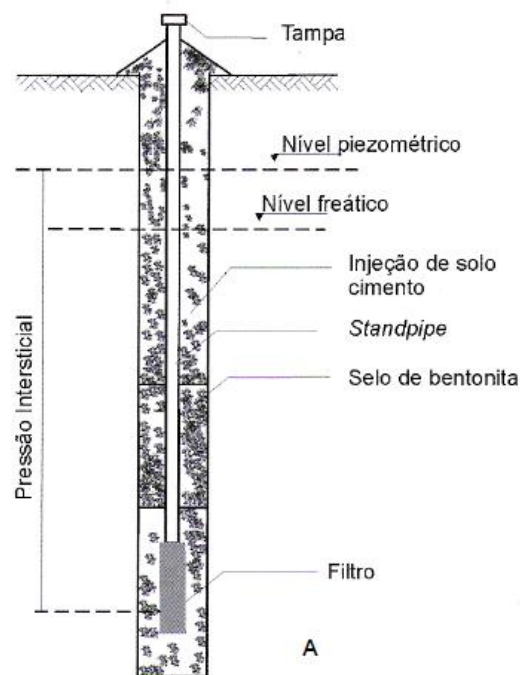
No caso específico do piezômetro do tipo *standpipe* (Figura 6), também denominado de tubo aberto, a água dos poros passa através do filtro do bulbo drenante do instrumento até atingir o equilíbrio com a poropressão na fundação. A poropressão corresponde, então, à altura da água acima do bulbo do instrumento. Em geral, usa-se a cota do ponto médio do bulbo como referência para leitura (CASTRO, 2008).

De acordo com Castro (2008), as principais características dos piezômetros *standpipe* são:

- Confiabilidade;
- Durabilidade;

- Sensibilidade;
- Possibilidade de verificação de seu desempenho por meio de ensaios de recuperação do nível d'água;
- Estimativa do coeficiente de permeabilidade do solo;
- Interferência da praça de compactação durante a construção da barragem;
- Inadequação, geralmente, para a medição de pressões neutras de período construtivo;
- Certa dificuldade de acesso aos terminais de leitura;
- Alto tempo de resposta (*time lag*), quando instalado em solos com baixa permeabilidade.

FIGURA 6 – ESQUEMA DE INSTALAÇÃO DE PIEZÔMETRO *STANDPIPE* EM UM FURO DE SONDAGEM



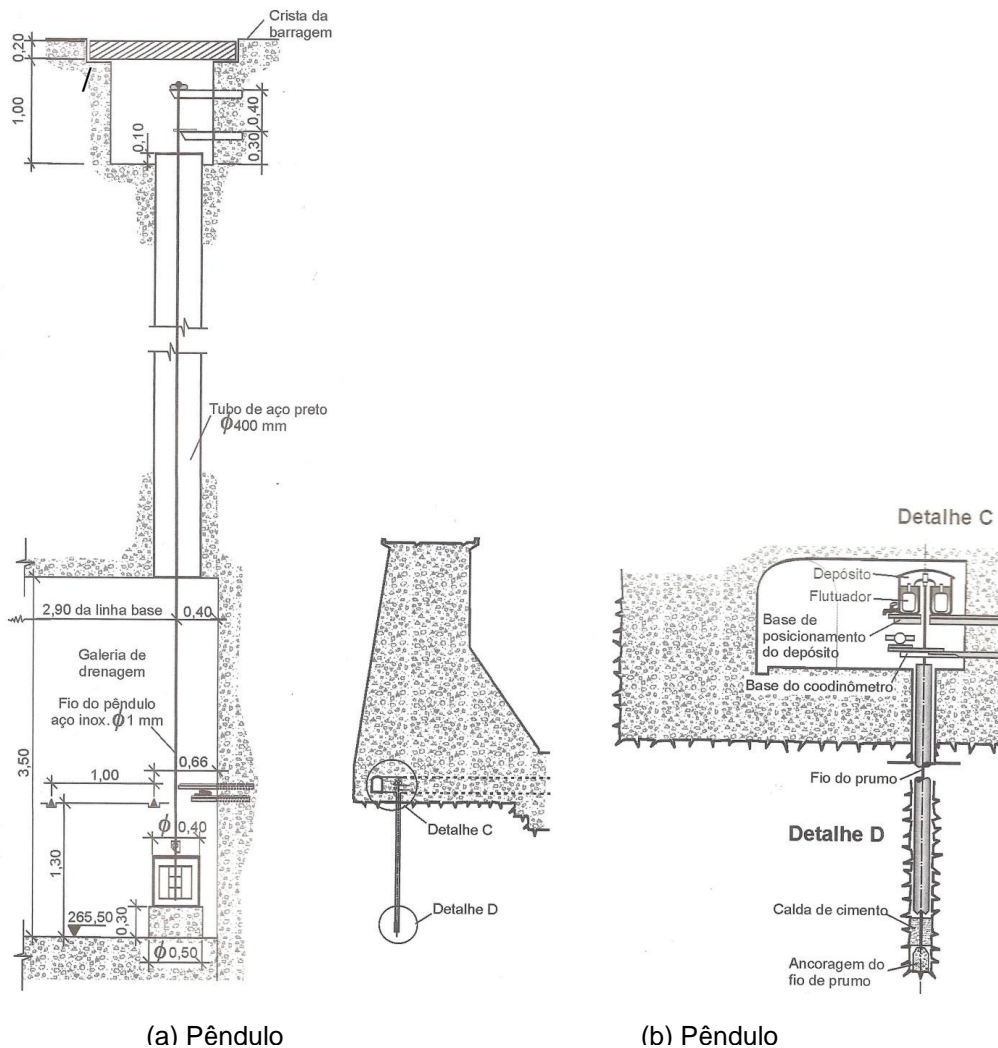
FONTE: SILVEIRA¹ (2006 *apud* CASTRO, 2008).

Os deslocamentos horizontais da barragem são medidos através da associação dos valores observados nos pêndulos diretos e invertidos (Figura 7), sendo que para a medição dos deslocamentos horizontais da crista são indicados os pêndulos diretos, enquanto que para os deslocamentos cisalhantes da base da barragem são usados os pêndulos invertidos (SILVEIRA, 2003). Tais deslocamentos

¹ SILVEIRA, J. F. A. *Instrumentação e comportamento de barragens de terra e enrocamento*. São Paulo: Oficina de Textos, 2006.

são afetados principalmente pela deflexão da estrutura de concreto, pela rotação da base da estrutura, devido à deformabilidade da fundação e pelas influências térmicas ambientais. Exceto as influências térmicas, que atuam permanentemente ao longo da vida útil da obra, esses deslocamentos evoluem de modo mais acentuado nos primeiros anos após o enchimento do reservatório da barragem, tendendo exponencialmente para um valor estabilizado com o tempo (CHI, 2002).

FIGURA 7 – ESQUEMA DE INSTALAÇÃO DE PÊNDULOS DIRETO E INVERTIDO



FONTE: Adaptado de SILVEIRA (2003).

Conforme Silveira (2003), os deslocamentos horizontais do fio do pêndulo são medidos em relação a uma das paredes da galeria, empregando-se um coordenômetro ótico ou um coordenômetro eletrônico, segundo as direções montante-jusante e margem direita-esquerda.

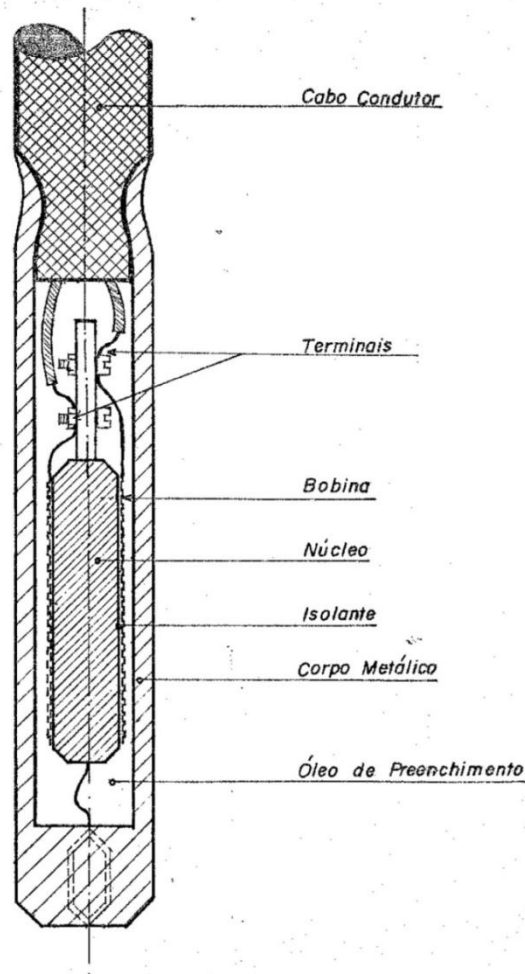
Na instalação dos pêndulos diretos uma extremidade é fixa à crista da barragem indo até o contato concreto-rocha, enquanto que os pêndulos invertidos são fixos na fundação indo até o contato concreto-rocha, sendo que a profundidade de instalação dos pêndulos invertidos está condicionada à altura da barragem e às descontinuidades sub-horizontais da fundação. Preferencialmente, ambos devem ser instalados nos mesmos blocos-chave da barragem, de tal modo que seja possível comparar os valores dos deslocamentos horizontais dos pêndulos direto/invertido com aqueles obtidos geodesicamente (MATOS, 2002).

Por meio de termômetros para concreto é possível aferir a temperatura em um determinado local das estruturas, decorrente do desenvolvimento do calor gerado pela hidratação do cimento ou devido a fontes externas geradoras de calor. A sua utilização se estende também ao longo da vida útil da estrutura, tendo como objetivo o conhecimento da distribuição da temperatura na mesma, visando caracterizar as deformações de origem térmica (SOM, 1990a).

Durante a fase de construção de uma barragem de concreto, termômetros (Figura 8) são embutidos na estrutura para medição da temperatura. Os termômetros internos, instalados em regiões centrais das massas de concreto, são úteis no monitoramento de reações exotérmicas resultantes da hidratação do cimento. Posteriormente, durante a fase de operação da barragem, as informações destes sensores também são usadas, em combinação com outros instrumentos, na avaliação do desempenho estrutural, pois, conforme Vasconcelos (1978), deslocamentos, deformações, tensões e movimentos de juntas são influenciados pela distribuição de temperaturas nos blocos que compõem a barragem.

Dado que as mudanças sazonais de temperatura representam uma das principais causas dos deslocamentos e da ocorrência de fissuras em barragens de contrafortes, o controle da temperatura durante a fase de operação da barragem apresenta a vantagem de auxiliar o monitoramento desses fenômenos (VASCONCELOS, 1978; EKSTROM, 2009).

FIGURA 8 – ESQUEMA DE UM TERMÔMETRO PARA CONCRETO



FONTE: SOM (1990a).

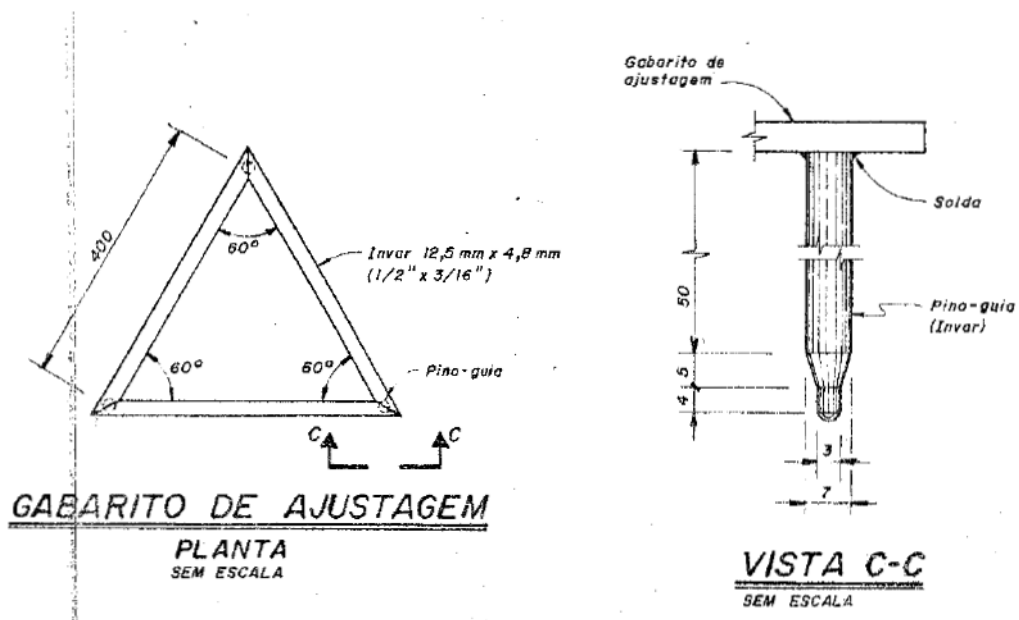
Outra vantagem, de acordo com Matos (2002), é que quando instalados superficialmente junto à face montante, os termômetros informam a temperatura da água do reservatório.

Para o acompanhamento dos deslocamentos ocorridos entre dois blocos adjacentes, separados por uma junta de contração, são utilizados alongômetros apoiados sobre bases triangulares de referência previamente fixadas na superfície do concreto (Figura 9). Os valores medidos permitem acompanhar o comportamento das estruturas durante os períodos de construção, de carregamento hidrostático e na fase de operação do reservatório (SOM, 1990b).

As bases de alongômetro possibilitam medir deslocamentos relativos horizontais e verticais entre as juntas dos blocos da barragem. A leitura é feita por um dispositivo portátil acoplado a um relógio comparador. Para verificarem-se os

deslocamentos, são feitas duas leituras em cada base de alongômetro: uma das medidas corresponde à abertura ou fechamento das juntas, ao passo que a outra medida indica o deslizamento entre os blocos. O deslizamento, por sua vez, de acordo com Neumann Jr (2015), pode ser interpretado de duas formas, conforme as bases de alongômetro estejam instaladas na parede ou no piso das galerias da estrutura. Quando no piso, o deslizamento indica um movimento relativo no sentido montante-jusante. Quando na parede, isto corresponderia a um recalque relativo entre os blocos.

FIGURA 9 – ESQUEMA DE UMA BASE PARA MEDIDOR DE JUNTAS (ALONGÂMETRO)



FONTE: SOM (1990b).

2.1.3 Monitoramento de Barragens de Concreto

Monitoramento estrutural, de acordo com Santos *et al.* (2013), pode ser definido como o desenvolvimento e aplicação de estratégias para identificar comportamentos anormais (tais como danos) em sistemas estruturais. Em estruturas de engenharia civil, os danos podem levar a ações de manutenção de alto custo e, quando ocorrem com significativa magnitude, podem resultar em consequências sociais e humanas dramáticas. Um monitoramento eficiente deve visar identificar

danos em um estágio inicial, o que está geralmente relacionado a fenômenos locais, com pequena magnitude.

A identificação de danos tem sido tema de estudos em diversas áreas, tais como, sistemas estruturais mecânicos, espaço aéreo e engenharia civil, em geral com abordagens utilizando modelos-bases ou orientadas a dados. No primeiro caso, normalmente ajusta-se um modelo numérico aos dados reais combinado com técnicas de otimização. As abordagens orientadas a dados, por outro lado, são geralmente baseadas em processamento de dados obtidos a partir do monitoramento, sem depender de modelos a priori (SANTOS *et al.*, 2013).

O interesse em técnicas que abordam o monitoramento estrutural de barragens tem crescido nas últimas décadas devido especialmente à ocorrência de acidentes, tais como rupturas, que resultaram em desastres ambientais de grandes proporções, prejuízos financeiros e, muitas vezes, altos índices de mortes. Essas ocorrências comprovam a importância do desenvolvimento de metodologias confiáveis para monitorar do comportamento das estruturas que compõem a parte física das barragens, como forma de evitar as consequências provocadas por desastres (MEDEIROS e LOPES, 2011).

A respeito do monitoramento de barragens, Cruz (2006) ressalta que esta atividade deve ser realizada por uma equipe especializada, através de inspeções visuais, medição geodésica de deslocamentos verticais e/ou horizontais, levantamentos batimétricos e acompanhamento da instrumentação.

Nessa tarefa, a auscultação é parte integrante e de extrema importância, por considerar os dados da instrumentação e as inspeções visuais. A auscultação é composta por um conjunto de formas de observação do comportamento da barragem e fundações, para controlar suas condições de segurança, comprovar a validade das hipóteses e dos métodos de cálculos utilizados no projeto e verificar a necessidade de medidas corretivas (ITAIPU BINACIONAL, 2016).

Recomenda-se a medição de grandezas associadas aos comportamentos: estrutural (tensões, deformações e deslocamentos); térmico (temperatura do concreto, dissipação do calor de hidratação, ciclos térmicos); hidráulico e hidrogeológico (vazões e pressões intersticiais); junto com fatores ambientais que influenciam o comportamento da barragem tais como temperaturas ambiente e do

reservatório, nível de água a montante e a jusante, precipitação e atividade sísmica (ELETROBRÁS, 2003; CARVALHO e ROMANEL, 2007; FIORINI, 2008).

Sendo assim, a avaliação de segurança de uma barragem envolve o monitoramento das respostas resultantes da interação das suas estruturas e fundações com o meio ambiente e pela ação de outras fontes de perturbação, tais como, material estrutural, ruído, entre outras. De acordo com Kuperman *et al.* (2005), os relatórios referentes à instrumentação e às inspeções visuais são úteis para esta tarefa porque abrangem todos os aspectos das represas, desde a sua construção até a fase de operação.

Em geral, os métodos utilizados para avaliar a segurança estrutural de barragens consistem em comparar as cargas e os fatores de segurança utilizados nos seus projetos com o comportamento de todas as suas estruturas ao longo dos anos.

Para identificar situações de normalidade ou anormalidade quanto ao comportamento da estrutura da barragem, com base na instrumentação, é necessário analisar e interpretar resultados tanto com tendência crescente quanto decrescente. Além disso, é preciso determinar os valores previstos para as grandezas de interesse, tendo como referência os critérios de cálculo adotados em projeto e, sempre que possível, os valores (ou níveis) de projeto e/ou críticos, para confrontação com os observados (CRUZ, 2006).

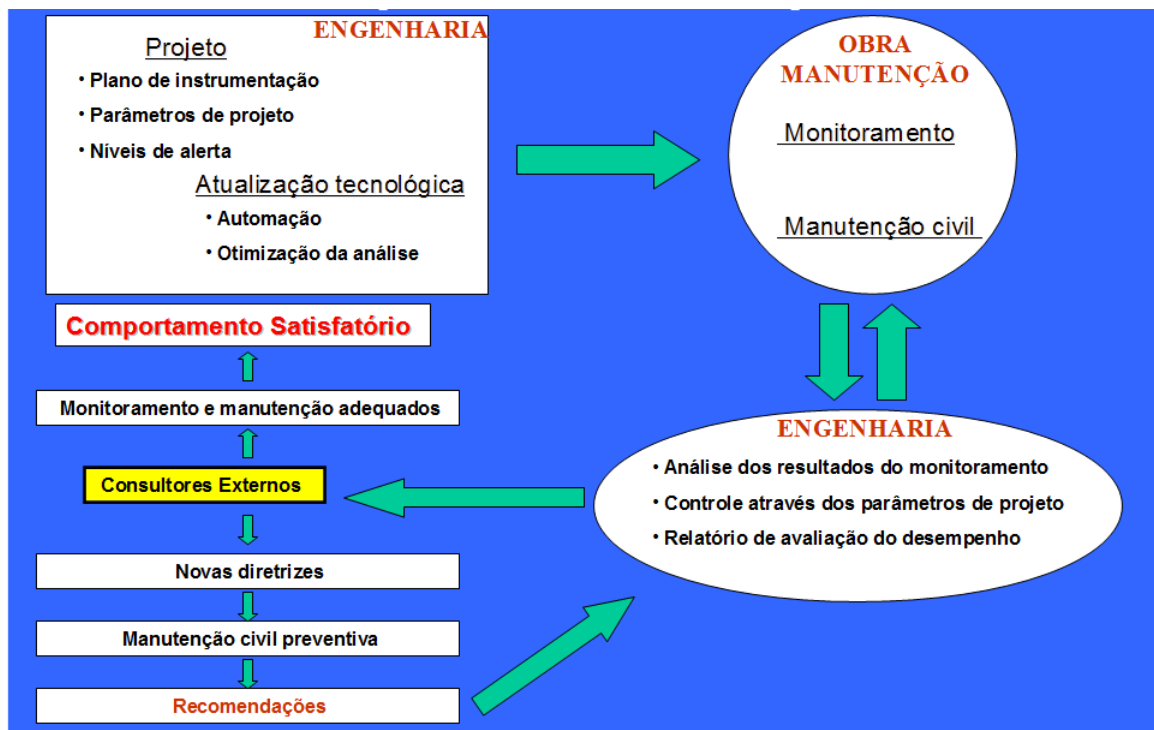
Conforme já exposto, o monitoramento de barragens deve contemplar tanto a verificação das condições de contorno de controle (temperatura, chuva, nível de água, etc), quanto a identificação da resposta estrutural (deslocamentos, rotações, percolação, etc.). Os dados obtidos, conforme De Sortis e Paoliani (2007), são úteis para a avaliação da segurança do desempenho de barragens, principalmente se as medidas atuais são comparadas com toda a série de dados registrados por meio de ferramentas de identificação estatísticas e estruturais. Sendo assim, é possível realizar uma interpretação analítica das medidas e, após a identificação dos parâmetros adequados, a verificação do comportamento normal da estrutura.

A seleção de dados durante esta atividade, segundo Farrar e Worden (2007), envolve a escolha do método, o tipo, número e localização do sensor, e o *hardware* de aquisição/armazenamento/transmissão de dados. Este processo é específico para cada aplicação. Questões econômicas desempenham um importante

papel na tomada dessas decisões. O intervalo de tempo no qual os dados devem ser recolhidos é outro ponto que deve ser considerado.

A Figura 10 ilustra como as diversas equipes que constituem o corpo técnico envolvido com a segurança de barragens interagem, para que seja possível visualizar antecipadamente todas as situações críticas mais prováveis, de modo a prover as soluções técnicas, os recursos correspondentes e estabelecer procedimentos administrativos claros, especialmente em situações de emergência (CRUZ, 2006).

FIGURA 10 – INTERAÇÃO ENTRE EQUIPES NA AVALIAÇÃO DA SEGURANÇA DE BARRAGENS



FONTE: FIORINI (2008).

2.1.4 Barragem de Itaipu

A Usina Hidrelétrica de Itaipu está localizada no rio Paraná, na divisa entre o Brasil e o Paraguai. A empresa Itaipu Binacional foi fundada em 1974 e a construção das instalações teve início no ano seguinte. As 20 unidades geradoras conferem uma capacidade máxima de 14000 MW, tornando-a a maior usina hidrelétrica em

geração de energia em operação no mundo. Em 1995 foi considerada uma das sete maravilhas do mundo moderno. Ocupa também posição de destaque quanto à produção de energia e potência instalada, estando entre as maiores represas de gravidade aliviada do mundo, além de ser referência nos estudos de concreto e na segurança de barragens.

A segurança da barragem e das suas estruturas é constantemente monitorada, estudada e avaliada. A Itaipu, em parceria com a Fundação Parque Tecnológico Itaipu (PTI) e por meio do Centro de Estudos Avançados em Segurança de Barragens (CEASB), tem dedicado especial atenção ao desenvolvimento de soluções estratégicas nesta área. No CEASB, são realizadas pesquisas nas áreas de processamento integrado de dados, realidade aumentada, modelagem 3D, simulação, geotecnia, instrumentação e realidade virtual (CEASB, 2016).

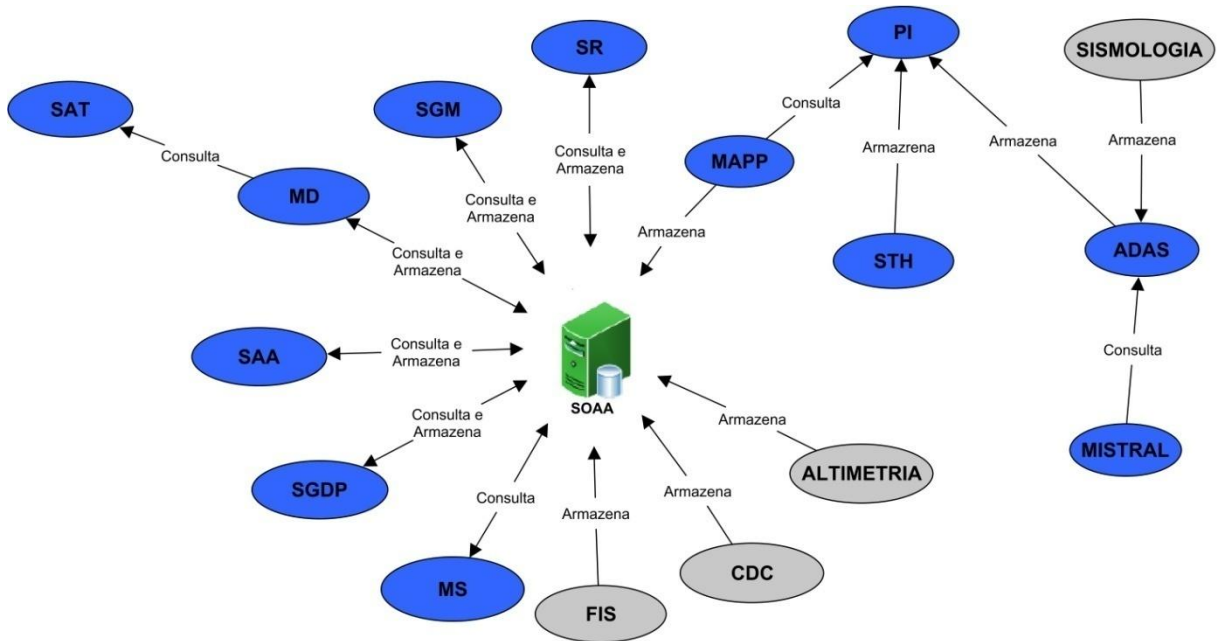
O CEASB foi responsável pelo desenvolvimento do Sistema de Cadastro Nacional de Barragens, entregue em 2014, que reúne informações sobre cerca de 1.400 barragens e é a base de dados oficial do Comitê Brasileiro de Barragens (CBDB). Trata-se de um sistema que possibilita gerenciar, de forma simples e confiável, as informações técnicas necessárias para a realização de avaliações estatísticas sobre barragens, a elaboração de projetos, a construção de novas barragens e para a pesquisa científica (CEASB, 2016).

A análise da segurança estrutural da barragem de Itaipu é feita a partir das observações da auscultação. Dentre os mais de 2300 instrumentos instalados nas estruturas e fundações da barragem encontram-se: pêndulos diretos e invertidos, bases de alongâmetro, termômetros para concreto, piezômetros, extensômetros, entre outros. Além dos dados fornecidos pela instrumentação, outros elementos são monitorados, tais como temperatura ambiente e níveis de água (montante e jusante). As medições são realizadas periodicamente, de forma manual e/ou automatizada (NEUMANN JR, 2015).

O processo de análise da auscultação das estruturas, de acordo com Coelho; Patias e Garay (2015) é realizado por meio do desenvolvimento de módulos de monitoramento, os quais fazem parte do Sistema de Otimização e Análise de Auscultação (SOAA). A partir dos subsistemas que compõem o SOAA (Figura 11) a equipe responsável pela segurança estrutural da barragem pode acessar os dados de projeto, tais como plantas e seções transversais das estruturas, fundações e

instrumentos instalados. Além disso, também é possível elaborar gráficos de leituras de instrumentos em diferentes escalas, comparando-os com gráficos dos níveis a montante e a jusante, precipitação e temperatura, favorecendo, assim, a realização de análises mais dinâmicas dos dados. O Quadro 1 apresenta a descrição dos principais subsistemas.

FIGURA 11 – LÓGICA ORIGINAL DO SOAA



FONTE: Adaptado de COELHO, PATIAS e GARAY (2015).

QUADRO 1 – DESCRIÇÃO DOS PRINCIPAIS SUBSISTEMAS DO SOAA

SUBSISTEMA	DESCRIÇÃO
Sistema de Arquivo Técnico (SAT)	Responsável pelo armazenamento e consulta de todos os documentos e projetos de ITAIPU.
Sistema de Gestão de Medições Manuais (SGM)	Responsável por introduzir e validar leituras de campo que são processadas através de fórmulas e transformadas em grandezas.
Maquete Digital (MD)	Permite de forma intuitiva navegar pela estrutura civil da barragem obter gráficos e informações da instrumentação civil.
Sistema Analítico de Auscultação (SAA)	Sistema <i>online</i> de monitoramento que contém dados de alarme dos sensores com leituras manuais e automatizadas.
Sistema de Gestão de Dados Planimétricos (SGDP)	Responsável pela importação dos dados das campanhas planimétricas, processamento e geração de gráficos de monitoramento.
Mapa de Subpressão (MS)	Permite gerar mapas de distribuição das subpressões nas feições geológicas monitoradas por piezômetros.

(continua)

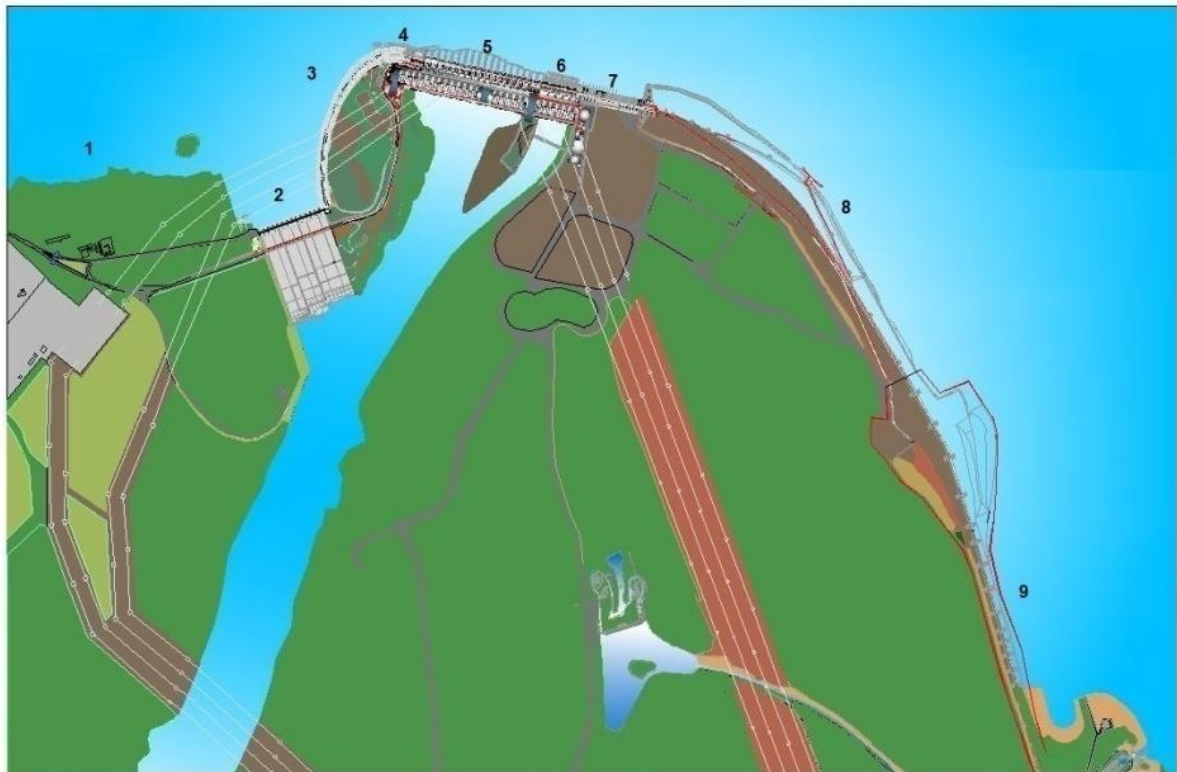
SUBSISTEMA	DESCRIÇÃO
Sistema de Recomendações (SR)	Sistema que realiza o gerenciamento das recomendações oriundas dos relatórios de análise estrutural da Barragem de Itaipu, permitindo acompanhar a situação da execução das recomendações.
PI	Plataforma do historiador de dados dos sistemas em tempo real da instrumentação de Itaipu.
Sistema de Aquisição de Dados Automáticos (ADAS)	Responsável por adquirir e armazenar diversas leituras de sensores de campo instalados na UHI.
MISTRAL	Software para análise qualitativa dos instrumentos automatizados para a análise do desempenho da barragem.
Sistema de Fissuras (FIS)	Realiza o registro e acompanhamento das fissuras. Sistema em fase de integração com o SOAA.
Sistema de Drenos de Concreto (CDC)	Responsável por introduzir e validar as leituras de campo dos drenos no sistema. Sistema em fase de integração com o SOAA.
ALTIMETRIA	Responsável por introduzir e validar as leituras de campo das campanhas altimétricas no sistema. Sistema em fase de integração com o SOAA.
SISMOLOGIA	Responsável pelo monitoramento sísmológico de ITAIPU. Sistema em fase de integração com o SOAA.

FONTE: COELHO, PATIAS e GARAY (2015).

No arranjo geral de Itaipu (Figura 12) é possível perceber a divisão da barragem em cinco classes: de enrocamento, de terra, principal, de ligação e lateral. As barragens: principal, estrutura de desvio, de ligação e laterais, são de concreto do tipo gravidade aliviada, maciça e contrafortes, respectivamente. Cada barragem está subdividida em blocos. Por exemplo, em toda a extensão da barragem há 83 blocos de contrafortes, sendo 64 localizados na margem direita (Trechos D e E), entre o Vertedouro e a Barragem Principal e outros 19 instalados na margem esquerda (Trecho I), entre a Estrutura de Desvio e a Barragem de Enrocamento (DTI, 2009).

A seguir são apresentadas apenas as informações referentes ao Trecho D, Barragem Lateral Direita, devido ao escopo deste trabalho.

FIGURA 12 – ARRANJO GERAL DA USINA HIDRELÉTRICA DE ITAIPU



Legenda:

- | | |
|---|--|
| 1 - Barragem de terra da Margem Direita | 6 - Canal de Desvio, Trecho H |
| 2 - Vertedouro, Trecho A | 7 - Barragem de Ligação Esquerda, Trecho I |
| 3 - Barragem Lateral Direita, Trecho D | 8 - Barragem de Enrocamento |
| 4 - Barragem de Ligação Direita, Trecho E | 9 - Barragem de terra da Margem Esquerda |
| 5 - Barragem Principal, Trecho F | |

FONTE: Adaptado de CEASB (2014).

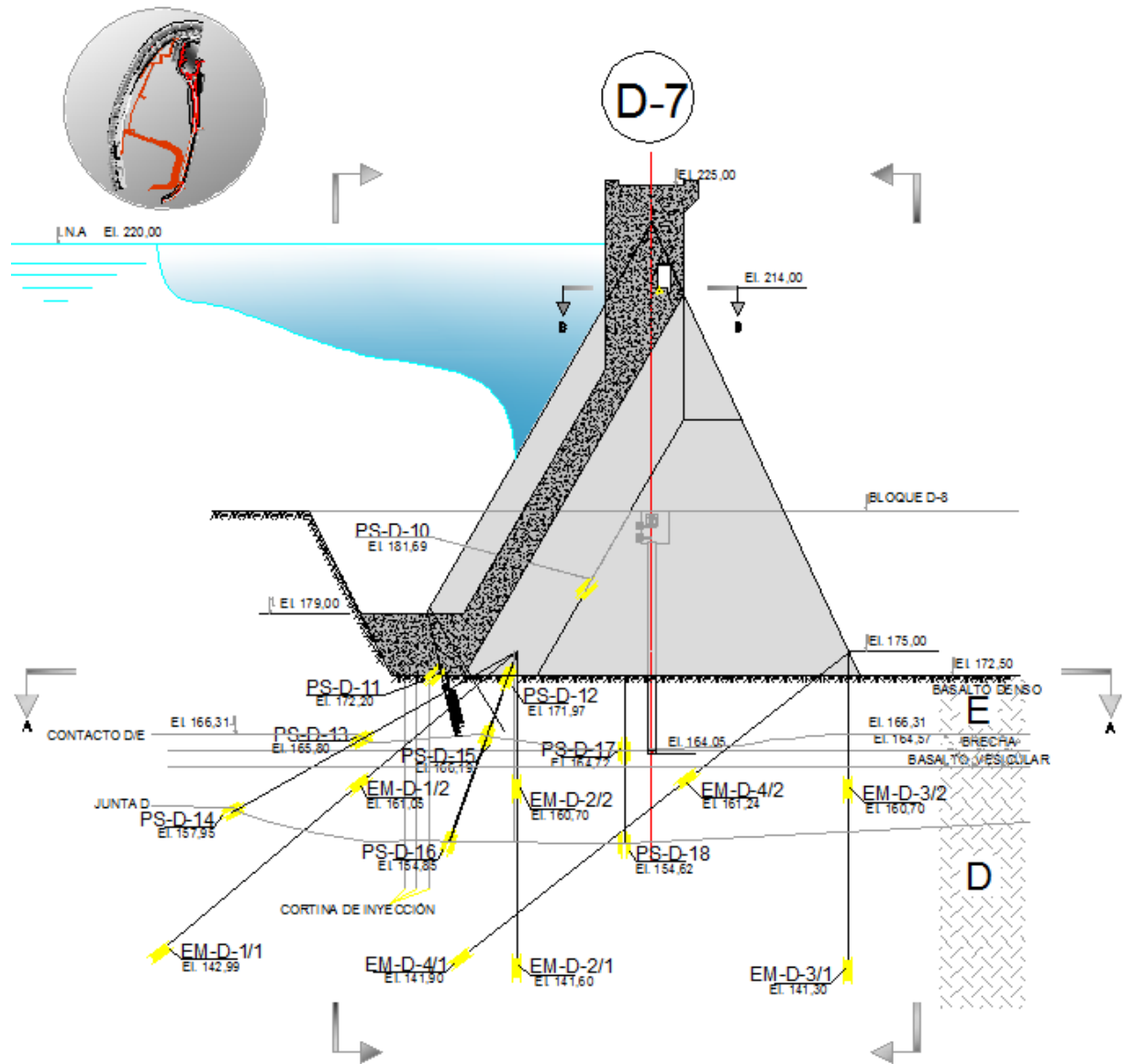
O eixo da Barragem Lateral Direita tem a forma de arco com curvatura de raio de circunferência duplo, perfazendo um comprimento de 986 m na crista. Os blocos de contraforte desse trecho possuem a mesma configuração estrutural, a saber: 17 m de largura no eixo e altura variando de 35 a 85 m; cabeça poligonal do contraforte de concreto-massa; declividade de montante e jusante iguais aquelas da Barragem Principal; cabeça de montante tal como o quadrante de um círculo, livre de tensões de tração internas; alma do contraforte com espessura crescente no sentido horizontal a partir do pescoço² até a face de jusante e no sentido vertical a partir da elevação 201,5 m até a fundação (SOM, 1984).

Do total de 58 blocos de contrafortes existentes no Trecho D, apenas seis (D7, D8, D20, D38, D54, D57) estão fortemente instrumentados. Esses são denominados blocos chave, por causa de sua representatividade perante os demais

² Pescoço do contraforte: região que vai da intersecção da parte inclinada com a vertical, até a crista.

desse setor. As Figuras 13 e 14 mostram as representações de dois desses blocos, com parte da respectiva instrumentação instalada.

FIGURA13 – BLOCO CHAVE D7 DO TRECHO D DA BARRAGEM DE ITAIPU



FONTE: CEASB (2014).

TABELA 1 – VALORES DE CONTROLE PARA OS EXTENSÔMETROS DO TRECHO D DA BARRAGEM DE ITAIPU

Bloco	Instrumento	Deslocamento Máximo Relativo ⁴ (mm)		
		Ancoragem Superior	Ancoragem Média	Ancoragem Inferior
D7	EMD1	-0,1	-	+1,8
D7	EMD2	+0,1	-	+0,3
D7	EMD3	-0,1	-	-0,2
D7	EMD4	0,0	-	+0,5
D8	EMD5	+0,1	+0,5	+0,4
D8	EMD6	-0,3	-	-0,3
D8	EMD7	-0,2	-0,4	-0,7
D8	EMD8	-0,4	-	-0,8
D8	EMD9	0,0	-0,1	0,0

FONTE: Adaptado de (SOM, 1984).

A Tabela 2 apresenta as medidas extremas previstas para os deslocamentos horizontais relativos dos pêndulos diretos e invertidos destes mesmos blocos, nas direções montante-jusante e do eixo da barragem, considerando dois casos de carregamento: carga hidrostática (CH) e variação de temperatura⁵ (ΔT_0).

TABELA 2 – VALORES DE CONTROLE PARA OS PÊNDULOS DIRETOS E INVERTIDOS DO TRECHO D DA BARRAGEM DE ITAIPU

Bloco	Instrumento	Cota (m)	Máximo Deslocamento Relativo ⁶ (mm)	
			ΔT_0	CH
D7	PID1 - Base	191,62	+1,2	+2,2
D8	PRD1 - Base Superior	215,44	+2,4	+1,4
D8	PRD1 - Ponto de Fixação	221,90	+3,2	+1,7

FONTE: Adaptado de (SOM, 1984).

Os valores máximos previstos para as aberturas das juntas entre os blocos chave D7/D8 e D8/D9, medidas pelas bases de alongâmetro, constam na Tabela 3. Para os recalques diferenciais entre blocos a estimativa do deslocamento diferencial vertical máximo é de 1 mm (SOM, 1984).

De acordo com SOM (1984), os valores mínimo e máximo previstos para a temperatura na superfície dos blocos de contrafortes da barragem de Itaipu são 5°C e 40°C, respectivamente.

⁴ O sinal positivo indica afastamento entre a cabeça e a ancoragem e o negativo significa aproximação.

⁵ A variação de temperatura corresponde à diferença entre a temperatura máxima do concreto e a sua temperatura de equilíbrio, estimada em função da temperatura média anual ambiente.

⁶ O sinal positivo indica deslocamento para jusante.

TABELA 3 – VALORES DE CONTROLE PARA AS BASES DE ALONGÂMETRO (ABERTURAS) DO TRECHO D DA BARRAGEM DE ITAIPU

Junta	Instrumento	Cota (m)	Máximo Deslocamento (mm)
D7/D8	JSD22	191,3	0,70
D7/D8	JSD23	191,3	0,70
D7/D8	JSD24	214	0,70
D8/D9	JSD25	191,3	0,70
D8/D9	JSD26	191,3	0,70
D8/D9	JSD27	214	0,70

FONTE: Adaptado de (SOM, 1984).

Os valores de controle citados anteriormente foram baseados exclusivamente nos critérios de projeto da barragem de Itaipu, e servem como simples valor de referência, já que na sua maioria encontram-se ultrapassados ou são muito conservadores, como informa a Especificação Técnica (1997):

Deve-se enfatizar que os valores de controle anteriormente determinados, o foram na fase de projeto, havendo os mesmos sido de utilidade apenas para o acompanhamento na fase do enchimento do reservatório, não sendo mais aplicáveis na fase de operação. Isto é devido ao fato de que os modelos matemáticos utilizados basearam-se apenas em análises elástico lineares, não havendo sido computadas as deformações lentas da fundação e do concreto, que ainda não se estabilizaram completamente, nem as influências térmicas ambientais (variação verão/inverno) (ESPECIFICAÇÃO TÉCNICA, 1997, p.1).

Por esse motivo esses valores foram reavaliados por meio de modelagem matemática, utilizando os dados medidos na instrumentação, após o enchimento do reservatório, considerando uma situação normal de utilização e uma relação estatística entre a temperatura ambiente e os valores medidos. Os valores máximo e mínimo da temperatura média ambiente utilizados nos cálculos foram 28,7°C e 13,9°C, respectivamente (CHI, 1999a).

Nas Tabelas 4 e 5 são apresentados os valores de controle estimados para os deslocamentos medidos pelos extensômetros múltiplos de hastes e pêndulos no sentido do fluxo (montante-jusante), respectivamente, considerando a Condição de Carregamento Normal (CCN), que corresponde às variações de temperaturas médias observadas no verão e no inverno.

TABELA 4 – VALORES DE CONTROLE ESTIMADOS PARA OS EXTENSÔMETROS DO TRECHO D DA BARRAGEM DE ITAIPU

Bloco	Instrumento	Elevação - Ancoragem (m)	Valores de Controle - CCN		
			Inverno	Média	Verão
D7	EMD1	176,2 - 143,0	-0,20	-0,26	-0,31
D7	EMD1	176,2 - 161,1	-0,10	-0,20	-0,28
D7	EMD2	176,1 - 141,6	0,14	-0,10	-0,30
D7	EMD2	176,1 - 160,7	0,10	-0,14	-0,34
D7	EMD3	176,1 - 141,3	-0,74	-0,62	-0,52
D7	EMD3	176,1 - 160,7	-0,62	-0,49	-0,38
D7	EMD4	176,0 - 141,9	-0,42	-0,72	-0,97
D7	EMD4	176,0 - 161,2	-0,44	-0,70	-0,91
D8	EMD5	190,9 - 151,8	0,75	0,69	0,64
D8	EMD5	190,9 - 169,0	0,78	0,81	0,84
D8	EMD6	191,1 - 150,4	-1,09	-0,95	-0,84
D8	EMD6	191,1 - 168,7	-0,99	-0,87	-0,77
D8	EMD7	191,0 - 151,6	-1,16	-1,43	-1,65
D8	EMD7	191,0 - 167,5	-1,06	-1,30	-1,50
D8	EMD7	191,0 - 181,3	-0,64	-0,94	-1,19
D8	EMD8	191,2 - 149,5	-0,36	-0,47	-0,56
D8	EMD8	191,2 - 167,2	-0,19	-0,34	-0,47
D8	EMD9	191,1 - 151,5	-0,17	-0,44	-0,66
D8	EMD9	191,1 - 168,4	0,27	0,49	0,67
D8	EMD9	191,1 - 181,2	3,15	0,40	-1,88

FONTE: Adaptado de CHI (1999a).

TABELA 5 – VALORES DE CONTROLE ESTIMADOS PARA OS PÊNDULOS DO TRECHO D DA BARRAGEM DE ITAIPU, NO SENTIDO DO FLUXO

Bloco	Instrumento	Cota - Referência (m)	Valores de Controle - CCN		
			Inverno	Média	Verão
D7	PID1	191,62 - 164,05	0,89	0,66	0,47
D8	PRD1	215,44 - 191,10	1,76	1,15	0,65
D8	PRD1	221,90 - 191,10	4,19	2,59	1,27

FONTE: Adaptado de CHI (1999a).

Para a reavaliação dos valores de controle de piezômetros e medidores de vazão da barragem de Itaipu também foram empregados modelos estatísticos, considerando CCN. Esses modelos, obtidos por meio das leituras cronológicas das grandezas dos tipos “causa” e “efeito”, por um determinado período de tempo, podem ser utilizados para o controle de grandezas do tipo “efeito”, tais como: deslocamentos, infiltrações, deformações, subpressões, e assim por diante. Essas grandezas são correlacionadas no intuito de expressar matematicamente a série temporal da grandeza “efeito”. No entanto, a avaliação dos valores de controle, por meio dos modelos estatísticos, nem sempre permite um julgamento completo sobre a ocorrência de todos os possíveis eventos excepcionais (CHI, 1999b).

O estabelecimento de níveis de alerta, conforme experiência com outras barragens, requer uma técnica apurada, objetivando saber o que constitui eventual problema apenas com o instrumento, o que constitui influência térmica, o que constitui eventual influência do carregamento da barragem (carga hidrostática, subpressão excessiva, transmissão de esforços de um bloco para outro, etc.), o que constitui eventual problema com o maciço rochoso de fundação, o que constitui condições limites excepcionais, etc (CHI, 2002, p. 36).

Os critérios a serem utilizados para o estabelecimento de valores de atenção e de alerta, para os instrumentos de auscultação da barragem de Itaipu, incluídos no sistema ADAS, são apresentados no relatório CHI (2002).

Neste sistema foram incluídos 205 instrumentos de auscultação, dentre os quais incluem-se os pêndulos diretos, pêndulos invertidos, medidores de vazão, parte dos piezômetros de fundação, piezômetros da interface aterro-concreto dos muros de ligação, assim como parte dos extensômetros múltiplos de hastes... Estes instrumentos foram selecionados após uma minuciosa e criteriosa análise, procurando-se selecionar aqueles instrumentos que fossem mais importantes para a pronta detecção de eventuais anomalias (CHI, 2002, p.1-2).

Tais critérios podem ser empregados tanto na avaliação das respostas dos instrumentos integrantes do sistema "ADAS", quanto para os demais instrumentos de auscultação, permitindo a supervisão do comportamento global das estruturas e fundações da barragem e auxiliando a detecção de qualquer eventual anomalia.

- Nível de Atenção A: Corresponde a um primeiro nível de controle, onde a grandeza medida ultrapassaria ligeiramente o valor pré-estabelecido para o instrumento, de modo a necessitar de uma averiguação de suas causas e implicações, porém sem uma consequência significativa para as condições de segurança da barragem;
- Nível de Atenção B: Corresponde a um segundo nível de verificação, onde o valor de controle seria também ultrapassado em maior intensidade, de modo a necessitar de uma averiguação mais rápida, porém com implicação não muito séria sobre as condições de segurança da barragem;
- Nível de Atenção C: Corresponde a um terceiro nível de controle, onde a grandeza medida ultrapassaria em muito o valor de referência, de modo a necessitar de uma averiguação urgente de suas causas e implicações imediatas. Trata-se de um valor que apesar de ser aceitável, pode implicar em danos à estrutura da barragem. Neste caso a confirmação da leitura do instrumento e a investigação das causas do problema e da necessidade de eventuais medidas reparadora deve ser imediata (CHI, 2002, p.5).

Caso os níveis de atenção sejam ultrapassados por pelo menos dois instrumentos instalados no mesmo bloco e associados entre si, imediatamente um alerta é emitido e os responsáveis pela supervisão das condições de segurança da barragem são mobilizados. A Tabela 6 apresenta os níveis de atenção para alguns desses instrumentos, calculados a partir dos valores médios medidos entre 1982 e 2001, sem considerar as variações térmicas entre verão e inverno. Esses níveis podem ser atualizados computando as influências térmicas através de um coeficiente apropriado (CHI, 2002).

TABELA 6 – NÍVEIS DE ATENÇÃO PARA OS PIEZÔMETROS DO BLOCO D8 DO TRECHO D DA BARRAGEM DE ITAIPU

Bloco	Instrumento	Níveis de atenção		
		A	B	C
D8	PD8	2,61	2,69	2,69
D8	PS19	220,00	220,00	220,00
D8	PS20	215,00	217,50	220,00
D8	PS21	190,00	205,00	220,00

FONTE: Adaptado de CHI (2002).

Os extensômetros instalados sub-horizontalmente para montante, a partir da cabeça dos blocos-chave, apresentam tendência de evolução dos recalques ou dos deslocamentos sub-horizontais da barragem de modo exponencial ao longo do tempo, com nítida influência das variações térmicas e comportamento regular (CHI, 2002).

2.2 MÉTODOS ESTATÍSTICOS

Um conjunto de n observações acumuladas sequencialmente ao longo do tempo com relação a uma determinada variável, com a característica de dependência serial, é denominado série temporal (BOX; JENKINS, 1976).

Uma coleção de n observações sobre p variáveis aleatórias distintas, tomadas de um mesmo item, compõe uma amostra multivariada, que pode ser representada por uma matriz X de ordem $n \times p$ (Equação 1), de tal forma que as linhas correspondem às observações e as colunas às variáveis.

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2j} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{ip} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{np} \end{bmatrix} \quad (1)$$

A análise multivariada fornece métodos e técnicas para a interpretação teórica dessa amostra de forma conjunta, enquanto que os modelos de previsão de séries temporais possibilitam a previsão de valores futuros, obtidas diretamente dos valores passados, sem o uso de uma teoria subjacente.

Nas subseções a seguir são apresentados os aspectos teóricos das técnicas de análise estatística utilizadas no presente estudo.

2.2.1 Séries Temporais

Em se tratando de séries temporais, de acordo com Barbão (2007), pode-se distinguir componentes sistemáticos, indicativos de movimentos regulares, e não-sistemáticos, que apontam para movimentos irregulares ou díspares.

Pertencentes à classe de componentes sistemáticos estão: a tendência de crescimento, o ciclo e a sazonalidade, sendo que estes dois últimos diferem pelos períodos de seus movimentos. Já os componentes não-sistemáticos caracterizam-se pelas irregularidades, ruídos aleatórios, que devem ser isolados e estudados a fim de melhor compreender a série e possibilitar a previsão de valores. Os vários tipos de componentes podem atuar tanto de forma independente quanto conjunta, de forma que previsões com base em séries temporais só são de fato válidas se as propriedades de seus componentes permanecem relativamente estáveis durante o tempo em que a previsão é feita (BARBÃO, 2007).

Por tendência em séries temporais entende-se um comportamento globalmente direcionado dos dados, representados graficamente por linhas continuamente ascendentes ou descendentes de suave curvatura, num determinado espaço de tempo. Já os ciclos geralmente aparecem em séries longas e

caracterizam-se por oscilações quase regulares em torno de tendências. Segundo Barbão (2007), seu estudo é importante porque permite controlar as causas e efeitos desse comportamento dos dados. Por outro lado, os componentes sazonais são ciclos curtos que também variam em torno da tendência, com periodicidade conhecida.

De acordo com Box e Jenkins (1976), o conjunto das observações feitas em intervalos regulares de tempo, como por exemplo, mensalmente, pode ser representada genericamente por $\{Z_1, Z_2, Z_3, \dots, Z_t, \dots, Z_{n-1}, Z_n\}$ ou, alternativamente, como a série $\{Z_t, t=1,2,\dots,n\}$. De tal forma que a série temporal (Equação 2) é dada em função das componentes: tendência (T_t), sazonalidade (S_t), ciclo (C_t) e ruído aleatório (a_t).

$$Z_t = f(T_t, S_t, C_t, a_t) \quad (2)$$

Box e Jenkins (1976) orientam que ao iniciar o estudo de uma série temporal as primeiras estatísticas calculadas devem ser a função de autocovariância (γ_k), a função de autocorrelação (ρ_k) e a função de autocorrelação parcial (ϕ_{kk}), usando estimativas amostrais padrão (Equações 3 e 4).

$$\hat{\gamma}_k = \frac{1}{n} \sum_{t=1}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z}) \quad (3)$$

$$\text{Onde: } \bar{z} = \frac{1}{n} \sum_{t=1}^n z_t$$

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (4)$$

$$\text{Onde: } \hat{\gamma}_0 = \frac{1}{n} \sum_{t=1}^n (z_t - \bar{z})^2$$

A solução do sistema composto pelas equações lineares de Yule-Walker (Equação 5) fornece as estimativas dos valores das autocorrelações parciais (ϕ_{ij}) para as defasagens de ordem k , representando o quanto Z_t e Z_{t+k} estão relacionadas.

$$\phi(B)Z_t = a_t \quad (6)$$

$$Z_t = \delta + \theta(B)a_t \quad (7)$$

$$Z_t = \delta + \theta(B)a_t \quad (8)$$

Onde:

Z_t é a série temporal

B é o operador de retardo

$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ é o polinômio característico autoregressivo

$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$ é o polinômio característico média móveis

a_t é o ruído aleatório

Séries temporais não estacionárias, de acordo com Box e Jenkins (1976), podem ser representadas genericamente por um Modelo Autoregressivo Integrado Médias Móveis de ordem (p, d, q) – ARIMA (p, d, q) , conforme a Equação (9), com p parâmetros na porção autoregressiva, q parâmetros na parte de médias móveis e d diferenciações. A utilidade da diferenciação é tornar a série estacionária na média. Para séries não estacionárias na variância, a aplicação de uma transformação da família Box & Cox aos dados, tal como apresentado por Johnson e Wichern (2007), supre essa demanda.

$$\phi(B)\nabla^d Z_t = \theta(B)a_t \quad (10)$$

Onde: $\nabla^d = (1 - B)^d$ é o operador de diferenciação.

A formulação geral para as séries temporais sazonais, com correlação serial entre e dentro dos períodos sazonais (s), segundo Box e Jenkins (1976), é composta pelos modelos multiplicativos (Equação 10) do tipo ARIMA $(p,d,q)\times(P,D,Q)_s$.

$$\Phi(B^S)\phi(B)\nabla_S^D \nabla^d Z_t = \Theta(B^S)\theta(B)a_t \quad (10)$$

Onde:

$$\Phi(B^S) = 1 - \Phi_1 B^S - \Phi_2 B^{2S} - \dots - \Phi_p B^{pS}$$

$$\Theta(B) = 1 - \theta_1 B^S - \theta_2 B^{2S} - \dots - \theta_Q B^{QS}$$

$$\nabla_S^D = (1 - B^S)^D$$

Após a definição do modelo para a série temporal, para gerar as previsões dos valores futuros, pode-se utilizar o procedimento incondicional proposto por Box e Jenkins (1976), enquanto que para prever os valores anteriores aos iniciais indica-se o procedimento *backforecasting*.

A escolha do modelo que melhor representa uma série temporal, dentre os vários possíveis, deve levar em conta o número de parâmetros envolvidos. A utilização de um número elevado de parâmetros na modelagem pode não trazer benefícios quanto à qualidade do ajuste aos dados. Em geral, opta-se por modelos parcimoniosos, com poucos parâmetros, por sua utilidade prática.

Existem muitos critérios descritos na literatura de séries temporais para a seleção do modelo. Morettin e Tolo (2004) indicam a escolha do modelo que minimiza o Critério de Informação de Akaike (AIC), que, para modelos ARIMA (p, d, q), é calculado (Equação 11) em função da estimativa da variância dos resíduos ($\hat{\sigma}_a^2$), do número de parâmetros do modelo (c) e do tamanho da série (n).

$$AIC = \ln(\hat{\sigma}_a^2) + c \frac{2}{n} \quad (11)$$

2.2.2 Análise de Correlação Canônica

A análise de correlação canônica é uma técnica de análise de interdependência, que permite ao pesquisador identificar e quantificar as associações lineares existentes entre dois grupos de variáveis (\underline{X} e \underline{Y}). A ideia básica é encontrar as combinações lineares das variáveis de \underline{X} e as combinações lineares das variáveis de \underline{Y} que produzem correlações elevadas entre os dois grupos (JOHNSON e WICHERN, 2007; HAIR *et al.*, 2009).

Em geral, o grupo \underline{X} é composto por p variáveis, enquanto que o outro, \underline{Y} , é formado por q variáveis, assumindo $p \leq q$. Neste caso, os vetores \underline{X} e \underline{Y} possuem matrizes de covariâncias Σ_X e Σ_Y , respectivamente, e o seu relacionamento está resumido na matriz de covariância cruzada entre esses vetores que é Σ_{XY} .

Considerando que U e V (Equações 12 e 13), doravante denominadas variáveis canônicas, são as combinações lineares dos vetores \underline{X} e \underline{Y} , respectivamente, o problema canônico consiste na obtenção dos pesos \underline{a} e \underline{b} que maximizem a correlação entre U e V (Equação 14). Os vetores de pesos \underline{a} e \underline{b} , neste caso, são soluções do sistema de equações representado na Equação 15.

$$U = \underline{a}'\underline{X} \quad (12)$$

$$V = \underline{b}'\underline{Y} \quad (13)$$

$$\text{Corr}(U, V) = \sqrt{\lambda} = \frac{\underline{a}'\underline{\Sigma}_{XY}\underline{b}}{\sqrt{\underline{a}'\underline{\Sigma}_X\underline{a}}\sqrt{\underline{b}'\underline{\Sigma}_Y\underline{b}}} \quad (14)$$

$$\begin{cases} (\underline{\Sigma}_{XY}\underline{\Sigma}_Y^{-1}\underline{\Sigma}_{YX} - \lambda\underline{\Sigma}_X)\underline{a} = 0 \\ (\underline{\Sigma}_{YX}\underline{\Sigma}_X^{-1}\underline{\Sigma}_{XY} - \lambda\underline{\Sigma}_Y)\underline{b} = 0 \end{cases} \quad (15)$$

Onde: λ é um autovalor da matriz $(\underline{\Sigma}_X^{-1}\underline{\Sigma}_{XY}\underline{\Sigma}_Y^{-1}\underline{\Sigma}_{YX})$ ou, equivalentemente, da matriz $(\underline{\Sigma}_Y^{-1}\underline{\Sigma}_{YX}\underline{\Sigma}_X^{-1}\underline{\Sigma}_{XY})$ e cada autovalor dará origem a um par de variáveis canônicas. Dessa forma, cada par de variáveis canônicas apresenta variância unitária, correlação máxima e é não correlacionado com os demais pares. Em geral, procura-se obter poucos pares de variáveis canônicas que explicam grande parte da interdependência entre os dois conjuntos de variáveis originais (observáveis).

Para interpretar as variáveis canônicas, examina-se a matriz de correlação entre as variáveis originais e as variáveis canônicas, denominadas cargas canônicas.

2.2.3 Análise Fatorial

Conforme Lattin, Carrol e Green (2011), a seleção das variáveis de um processo pode ser feita a partir de p variáveis correlacionadas e, pela aplicação da análise fatorial, obtém-se um conjunto de m variáveis latentes⁸ não correlacionadas ($1 < m < p$), denominadas fatores comuns, que possuem propriedades especiais em termos de variâncias.

Por meio da análise fatorial é possível obter a descrição de matrizes com grande quantidade de elementos, de forma exploratória (sem ter conhecimento prévio da sua estrutura), e visualizar as relações mais importantes de um extenso conjunto de variáveis entre si ou as relações entre as amostras e as variáveis.

Essa técnica é amplamente utilizada nas áreas humanas, sociais e engenharias, sendo aplicada a escalas que consistem em uma grande quantidade de itens utilizados para medir personalidade, estilos de comportamento ou atitudes, na verificação da unidimensionalidade, um pressuposto central da Teoria de Resposta ao Item. Esse pressuposto, segundo Laros (2012), implica que todos os itens de um instrumento estejam medindo um único construto. Se o instrumento está de fato medindo mais do que um fator, um escore total individual deve ser calculado para cada fator e todas as análises estatísticas subsequentes devem ser feitas independentemente para cada fator.

O processo de obtenção dos fatores é amplamente abordado na literatura atual e pode ser resumido, segundo a abordagem por componentes principais, na sequência de etapas⁹ descritas a seguir.

Etapa 1: Seleção das p variáveis, definição do tamanho da amostra n , amostragem dos dados e organização na forma matricial (Equação 1). De acordo com Hair *et al.* (2009), o tamanho da amostra deve ser de pelo menos 5 vezes o número de variáveis, sendo que a proporção mais aceitável seria de 10 para um.

Etapa 2: Estimação das médias $\hat{\mu}_j$ (Equação 16) e das variâncias $\hat{\sigma}_j^2$ (Equação 17) amostrais das p variáveis aleatórias.

⁸ Por variável latente entende-se toda aquela que não pode ser medida e/ou observada diretamente.

⁹ A notação utilizada nas etapas descritas segue aquela apresentada por Mingoti (2005) e Johnson e Wichern (2007).

$$\hat{\mu}_j = \bar{x}_j = \frac{\sum_{i=1}^n x_{ij}}{n}, \quad j = 1, 2, \dots, p \quad (16)$$

$$\hat{\sigma}_j^2 = s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2, \quad j = 1, 2, \dots, p \quad (17)$$

Etapa 3: Padronização dos dados (Equação 18). De acordo com Mingoti (2005), ao utilizar variáveis padronizadas (z_{ij}), cuja variância é unitária, não há dominância direta de nenhuma delas e supera-se o problema de trabalhar com medidas de escalas e/ou grandezas diferentes.

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, \quad i = 1, 2, \dots, n; \quad j = 1, 2, \dots, p \quad (18)$$

Etapa 4: Composição da matriz de covariâncias amostrais (Equação 19) ou, de correlações amostrais (Equação 20). Ao optar pela matriz de correlações amostrais (R), os coeficientes de ponderação dos fatores serão mais equilibrados que aqueles obtidos a partir da matriz de covariâncias (MINGOTI, 2005).

$$\hat{V}(\underline{X}) = S = \begin{bmatrix} s_1^2 & s_{12} & \dots & s_{1j} & \dots & s_{1p} \\ s_{21} & s_2^2 & \dots & s_{2j} & \dots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{j1} & s_{j2} & \dots & s_j^2 & \dots & s_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \dots & s_{pj} & \dots & s_p^2 \end{bmatrix} \quad (19)$$

Onde:

$s_{jk} = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)$ é a covariância amostral entre as variáveis x_j e x_k , com $j, k = 1, 2, \dots, p$, para $j \neq k$

s_j^2 é a variância amostral (Equação 17) da variável x_j , com $j = 1, 2, \dots, p$

$$\hat{\rho} = R = \begin{bmatrix} 1 & r_{12} & \dots & r_{1j} & \dots & r_{1p} \\ r_{21} & 1 & \dots & r_{2j} & \dots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{j1} & r_{j2} & \dots & 1 & \dots & r_{jp} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \dots & r_{pj} & \dots & 1 \end{bmatrix} \quad (20)$$

Onde: $r_{jk} = r_{kj} = \frac{s_{jk}}{\sqrt{s_{jj}}\sqrt{s_{kk}}}$ é a correlação amostral entre as variáveis x_j e x_k , com $j, k = 1, 2, \dots, p$, para $j \neq k$

Etapa 5: Cálculo dos autovalores (λ_j) da matriz de covariâncias (ou de correlações) e os correspondentes autovetores. Algebricamente, na análise fatorial, os autovalores dispostos em ordem crescente correspondem às variâncias dos fatores, permitindo a formação de auto-espacos associados aos autovetores e, conseqüentemente, os planos fatoriais, que são subespacos bidimensionais. Dessa forma, é possível reduzir o número de dimensões do espaco com o qual se está trabalhando, obtendo uma representação dos espacos das amostras e das variáveis neste espaco de dimensão menor.

Etapa 6: Verificação da viabilidade do uso do modelo fatorial, com a análise das medidas das correlações e o cálculo da medida de adequação da amostra, pelo teste de esfericidade de Bartlett (Equação 21) e pelo critério Kaiser-Meyer-Olkin (KMO), respectivamente. No teste de esfericidade de Bartlett pretende-se rejeitar a hipótese nula de que a matriz de correlações é igual à matriz identidade. O coeficiente KMO (Equação 22) varia entre 0 e 1 e quanto mais próximo da unidade melhor a adequação da amostra.

$$T = - \left[n - \frac{1}{6}(2p + 11) \right] \sum_{j=1}^p \ln \hat{\lambda}_j \quad (21)$$

$$KMO = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} q_{ij}^2} \quad (22)$$

Onde:

$T \sim \chi_v^2$ com $v = \frac{p}{2}(p - 1)$ graus de liberdade

q_{ij} é o elemento pertencente à i -ésima linha e j -ésima coluna da matriz $Q = DR^{-1}D$, com $D = (\sqrt{\text{diagonal}(R^{-1})})^{-1}$, com $i = 1, 2, \dots, n$ e $j = 1, 2, \dots, p$.

Etapa 7: Escolha do número (m) de fatores de acordo com critério adotado. Este é um momento muito importante, visto que interfere na obtenção dos resultados subsequentes e suas respectivas interpretações. De acordo com Laros(2012), os erros que podem ocorrer nessa fase são: a superextração (extração de um número superior de fatores do que o necessário) e a subextração (extração de um número

inferior de fatores do que o necessário). Ambos comprometem de forma significativa os resultados e podem levar a conclusões equivocadas.

Os métodos mais comumente utilizados para a determinação do número de fatores são os critérios: de Kaiser; *Scree* de Cattell; média mínima de correlações parciais de Velicer (MMCPV) e análise paralela de Horn (APH).

No critério de Kaiser (CK), também denominado Guttman-Kaiser, o número de fatores extraídos é igual à quantidade de autovalores maiores que 1,0. O autovalor determina a proporção da variabilidade (Equação 22 e 23) presente no conjunto de variáveis que é explicada por um fator.

$$\frac{\hat{\lambda}_j}{s_1^2 + s_2^2 + \dots + s_p^2} ; j = 1, 2, \dots, p; \text{ para análise feita a partir de } S \quad (22)$$

$$\frac{\hat{\lambda}_j}{p} ; j = 1, 2, \dots, p; \text{ para análise feita a partir de } R \quad (23)$$

O critério *Scree* de Cattell é baseado no gráfico *Scree-Plot*, no qual o eixo das abscissas representa os fatores e o eixo das ordenadas os autovalores. O número de fatores pode ser determinado de duas formas: a primeira consiste em posicionar uma linha horizontal, paralela ao eixo x, na altura em que os pontos do gráfico formam uma linha aproximadamente paralela à horizontal. O número de pontos acima dessa linha horizontal determina a quantidade de fatores. O outro modo consiste em identificar o ponto em que o gráfico apresenta uma mudança brusca de declive, formando uma espécie de cotovelo. O número de pontos antes desse cotovelo corresponde ao número de fatores.

O critério da média mínima de correlações parciais de Velicer (MMCPV) é uma forma iterativa que consiste em recalcular a matriz de correlações parciais a cada fator extraído e obter a média do quadrado dessas correlações. A média mínima é alcançada quando a matriz de resíduos é semelhante à matriz identidade. Nesse momento, o processo para e observa-se o número de fatores extraídos.

O procedimento denominado de análise paralela de Horn (APH) consiste na geração de matrizes randômicas de correlação, seguida da análise fatorial destas matrizes e cálculo das médias dos autovalores obtidos. Para determinar o número de fatores, deve-se contar o número de vezes em que os autovalores da matriz de

correlação amostral são maiores que os respectivos valores das médias dos autovalores da matriz de correlação randômica.

O método mais popular, de acordo com a literatura pesquisada, é o critério de Kaiser (CK). No entanto, isso não significa que seja o mais apropriado em todas as situações. Laros (2012) apresenta uma revisão bibliográfica acerca da utilização e da precisão desses métodos de extração de fatores, apresentados no Quadro 2. Apesar do critério de APH ser o mais preciso, seu baixo índice de utilização decorre da falta de conhecimento a seu respeito, da não observância das sérias negligências dos outros critérios e do fato de não estar disponível nos principais pacotes estatísticos existentes no mercado.

QUADRO 2 – PRECISÃO DOS PRINCIPAIS MÉTODOS DE EXTRAÇÃO DE FATORES

Método	Precisão	Comentário
CK	22%	Tendência a superestimação
Scree	57%	Tendência a superestimação
MMCPV	84%	Tendência a subestimação
APH	92%	Tendência a superestimação

FONTE: Adaptado de LAROS (2012).

Etapa 8: Composição do modelo fatorial ortogonal (Equações 24 e 25), mediante a obtenção dos m fatores comuns e suas respectivas cargas fatoriais (Equação 26), juntamente com as medidas das comunalidades (Equação 27) e variâncias específicas (Equação 28) das p variáveis aleatórias originais padronizadas.

$$\begin{cases} \underline{Z}_1 = l_{11}\underline{F}_1 + l_{12}\underline{F}_2 + \dots + l_{1m}\underline{F}_m + \underline{\varepsilon}_1 \\ \underline{Z}_2 = l_{21}\underline{F}_1 + l_{22}\underline{F}_2 + \dots + l_{2m}\underline{F}_m + \underline{\varepsilon}_2 \\ \vdots \\ \underline{Z}_j = l_{j1}\underline{F}_1 + l_{j2}\underline{F}_2 + \dots + l_{jm}\underline{F}_m + \underline{\varepsilon}_j \\ \vdots \\ \underline{Z}_p = l_{p1}\underline{F}_1 + l_{p2}\underline{F}_2 + \dots + l_{pm}\underline{F}_m + \underline{\varepsilon}_p \end{cases} \quad (24)$$

$$\underline{Z} = L\underline{F} + \underline{\varepsilon} \quad (25)$$

Onde:

\underline{Z} é o vetor que contém os dados amostrais padronizados

L é a matriz de cargas fatoriais, composta pelos elementos l_{jk}

\underline{F} é o vetor de fatores

$\underline{\varepsilon}$ é o vetor de resíduos

$$l_{jk} = \text{cov}(Z_j, F_k) \quad (26)$$

Onde: l_{jk} é a carga fatorial (correlação) entre as variáveis padronizada Z_j e o fator F_k , com $j = 1, 2, \dots, p$ e $k = 1, 2, \dots, m$

$$h_j^2 = l_{j1}^2 + l_{j2}^2 + \dots + l_{jk}^2 + \dots + l_{jm}^2 \quad (27)$$

Onde: h_j^2 é a medida da comunalidade da j -ésima variável, com $j = 1, \dots, p$ e $k = 1, \dots, m$

$$\psi_j = s_j^2 - \sum_k^m l_{jk}^2, \quad j = 1, \dots, p \quad (28)$$

Onde: ψ_j é a medida da variância específica da j -ésima variável, com $j = 1, \dots, p$ e $k = 1, \dots, m$

A carga fatorial reflete a extensão com que cada fator comum contribui com a variância das observações em cada variável. Se o pesquisador estiver trabalhando com dados padronizados, então o parâmetro l_{jk} é o coeficiente de correlação do fator F_k com a variável Z_j . A comunalidade é a porção da variância de cada variável original proveniente dos fatores extraídos. O restante da variabilidade, que pode ser devida a fatores específicos, é medida pela variância específica.

Etapa 9: Realização da rotação *Varimax*, utilizando a matriz de transformação T , e atualização da matriz de cargas fatoriais (L^*) e dos vetores de fatores comuns (\underline{F}^*) para a composição do modelo rotacionado (Equação 29).

$$\underline{Z} = (LT)(T' \underline{F}) + \underline{\varepsilon} = L^* \underline{F}^* + \underline{\varepsilon} \quad (29)$$

O modelo fatorial possui infinitas soluções equivalentes quanto à capacidade de reproduzir a matriz de covariância amostral. Segundo Hair *et al.* (2009), isso ocorre porque a orientação da solução fatorial, isto é, a escolha dos vetores da base que descrevem o sistema de coordenadas, é arbitrária. Sendo assim, é possível escolher uma orientação da solução fatorial tal que facilite entendê-la e interpretá-la.

A interpretação dos fatores comuns usando a matriz de cargas fatoriais pode ser uma tarefa de difícil realização. Mas, é possível escolher uma orientação

diferente, rotacionando a solução fatorial de modo apropriado, para obter uma matriz de cargas fatoriais com uma estrutura mais simples de ser interpretada (HAIR *et al.*, 2009).

A rotação mais popular é a *Varimax*. Trata-se de uma transformação ortogonal dos fatores originais, que ocorre por meio da seleção dos coeficientes que maximizam a variação dos quadrados das cargas fatoriais originais, possibilitando, assim, obter para cada fator um grupo de variáveis altamente correlacionadas com um fator e com correlação desprezível ou moderada com os demais fatores.

Etapa 10: Estimação dos escores fatoriais (Equação 30). Os escores fatoriais ($\underline{\hat{f}}_i$) são os valores numéricos dos fatores para cada elemento amostral. Podem ser utilizados para classificar os elementos amostrais ou como variáveis de entrada para outras análises estatísticas.

$$\underline{\hat{f}}_i = (\hat{L}'_z \hat{L}_z)^{-1} \hat{L}'_z \underline{z}_i \quad , i = 1, 2, \dots, n \quad (30)$$

Etapa 11: Interpretação da estrutura fatorial obtida.

2.2.4 Análise de Agrupamentos

A utilização da análise de agrupamentos visa encontrar dentro de um conjunto heterogêneo de dados um pequeno número de grupos homogêneos, cuja variação dentro do grupo seja substancialmente menor do que a variabilidade total do conjunto de dados (LATTIN, CARROL e GREEN, 2011).

Para cumprir essa tarefa, é necessário que o pesquisador adote um método de comparação simultânea de observações sobre as variáveis utilizadas para caracterizar os elementos a serem agrupados.

Um método muito utilizado é o hierárquico aglomerativo, que consiste em considerar inicialmente cada observação como um grupo separado. A cada etapa do processo, os grupos juntam-se de acordo com as similaridades, formando novos agrupamentos, até que reste somente um agrupamento com o número total de observações incluído. A representação gráfica dessas etapas é conhecida como

dendrograma, caracterizada por uma estrutura hierárquica em árvore, gerada sequencialmente no processo de formação dos grupos.

De acordo com Hair *et al.* (2009), similaridade pode ser entendida como o grau de correspondência entre objetos (elementos, observações) ao longo de todas as características usadas na análise, representando uma medida da proximidade entre dois grupos. Correlações e medidas de distância são exemplos de similaridade. Estas últimas são, na realidade, uma medida de dissimilaridade, com valores maiores denotando menor similaridade. A distância é convertida em uma medida de similaridade pelo uso de uma relação inversa.

Quando as variáveis consideradas no estudo são quantitativas, os tipos de medidas de distância comumente utilizadas na análise de agrupamentos são a Euclidiana, de Mahalanobis e de Minkowsky. Não há, na literatura pesquisada, uma indicação a respeito da melhor forma de medida de distância. O pesquisador pode obter soluções distintas de agrupamentos ao usar diferentes medidas de distância ($d(\underline{x}, \underline{y})$), cabendo a ele optar por aquela que melhor represente os padrões inerentes de dados.

Dados os vetores p -dimensionais $\underline{x}' = [x_1, x_2, \dots, x_p]$ e $\underline{y}' = [y_1, y_2, \dots, y_p]$, a distância Euclidiana entre \underline{x} e \underline{y} é obtida pela Equação 30. Quando as variáveis são medidas em unidades de escala e/ou grandeza diferentes, essa formulação é geralmente aplicada aos dados padronizados. Assim, atribui-se peso igual a cada variável padronizada, de maneira que elas sejam igualmente importantes (após a padronização) na determinação da proximidade relativa dos objetos no espaço. (JOHNSON e WICHERN, 2007; LATTIN, CARROL e GREEN, 2011).

$$d(\underline{x}, \underline{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_p - y_p)^2} = \sqrt{(\underline{x} - \underline{y})' (\underline{x} - \underline{y})} \quad (30)$$

A distância estatística (métrica Mahalanobis) entre dois vetores de observações \underline{x} e \underline{y} é determinada pela Equação 31, usando a matriz de covariância amostral S .

$$d(\underline{x}, \underline{y}) = \sqrt{(\underline{x} - \underline{y})' S^{-1} (\underline{x} - \underline{y})} \quad (31)$$

Outra medida de distância é a métrica de Minkowski, cujo cálculo é definido pela Equação 32. Para $c = 1$, a medida é denominada "city-block" ou "quarteirão". Para $c = 2$, a medida é a já conhecida distância Euclidiana. Em geral, variando c muda a ponderação dada às maiores e menores distâncias. (JOHNSON e WICHERN, 2007; LATTIN, CARROL e GREEN, 2011).

$$d(\underline{x}, \underline{y}) = \left[\sum_{j=1}^p |x_j - y_j|^c \right]^{1/c} \quad (32)$$

Na abordagem aglomerativa hierárquica, além da escolha de uma fórmula de cálculo para a distância entre dois grupos, é necessário também estabelecer a forma de ligação entre os grupos para a formação de novos agrupamentos. A seguir são apresentados resumidamente os principais métodos de ligação. A teoria detalhada, discussões e exemplos de aplicação podem ser obtidos em Johnson e Wichern (2007).

Pelo método da ligação simples, em cada etapa do processo, os elementos mais próximos (Equação 33) definem a similaridade entre os grupos. Os grupos mais similares, com menores distâncias, são agrupados.

$$d_{(AB)C} = \min \{d_{AC}, d_{BC}\} \quad (33)$$

Onde:

A , B e C são grupos quaisquer;

AB é um agrupamento estabelecido anteriormente;

d_{AC} e d_{BC} são as distâncias entre os membros mais próximos dos grupos A e C e dos grupos B e C , respectivamente.

A ligação completa parte do contrário da ligação simples. As distâncias máximas são computadas (Equação 34) e os agrupamentos são formados pelos grupos mais próximos, isto é, que apresentarem as menores distâncias máximas.

$$d_{(AB)C} = \max \{d_{AC}, d_{BC}\} \quad (34)$$

Onde:

A , B e C são grupos quaisquer;

AB é um agrupamento estabelecido anteriormente;

d_{AC} e d_{BC} são as distâncias entre os membros mais distantes dos grupos A e C e dos grupos B e C , respectivamente.

O método da ligação média considera a distância entre dois grupos (Equação 35) como a média das distâncias entre todos os pares de elementos que podem ser formados com os elementos pertencentes aos grupos comparados. Tal como ocorre nos métodos anteriores, grupos mais próximos formam novos agrupamentos (MINGOTI, 2005).

$$d_{(AB)C} = \frac{\sum_{i=1}^{N_{AB}} \sum_k^{N_C} d_{ik}}{N_{AB} N_C} \quad (35)$$

Onde:

A , B e C são grupos quaisquer;

AB é um agrupamento estabelecido anteriormente;

d_{ik} é a distância entre o elemento i do grupo AB e o elemento k do grupo C ;

N_{AB} e N_C representam o número de elementos dos grupos AB e C , respectivamente.

O método de Ward, realiza os agrupamentos de menor soma de quadrados interna (Equação 36), ou seja, de variância mínima dentro do grupo.

$$SS_A = \sum_j^{N_A} (\underline{X}_j - \bar{\underline{X}})' (\underline{X}_j - \bar{\underline{X}}) \quad (36)$$

Onde:

A é um grupo qualquer;

SS_A é a soma de quadrados interna do grupo A ;

N_A é o número de elementos de A ;

\underline{X}_j é o vetor de observações do j -ésimo elemento pertencente ao grupo A ;

$\bar{\underline{X}}$ é o vetor de médias (centroide) do grupo A .

O método da ligação simples tende a produzir agrupamentos longos e encadeados, com formatos não convexos. Se os verdadeiros agrupamentos forem desta forma, isso não constituirá problema algum. Já os agrupamentos resultantes da ligação completa, são convexos na maioria das vezes. Porém, a presença de discrepâncias nos dados afeta consideravelmente os resultados deste método. Em geral, a ligação média apresenta bom desempenho no que diz respeito à minimização dos mínimos quadrados dentro dos grupos, produzindo grupos mais

homogêneos que os dois métodos anteriores. O método de Ward geralmente forma agrupamentos convexos, compactos e com aproximadamente o mesmo número de observações. Como não há conhecimento prévio dos verdadeiros agrupamentos, o pesquisador pode utilizar métodos diferentes de ligação e comparar os resultados, optando por aquele de melhor interpretação e/ou que cumpra com o objetivo de menor heterogeneidade dentro dos grupos (LATTIN, CARROL e GREEN, 2011).

Quanto ao número de agrupamentos, a escolha é muito subjetiva, dependendo dos objetivos do pesquisador no que diz respeito à simplicidade e adequação da solução.

Para testar o poder de separação entre os grupos, em geral, utiliza-se a análise discriminante.

2.2.5 Análise Discriminante

A análise discriminante é uma técnica que possibilita, a partir de um conjunto de variáveis mensuradas, estudar o perfil, realizar a diferenciação e a classificação de indivíduos (observações) em grupos previamente determinados, com um risco mínimo de erro.

A discriminação é feita com base numa regra matemática que minimiza as probabilidades de erros de classificação incorretas. A formulação desta regra, denominada função discriminante, depende das informações que se tem a respeito da população (grupos), como por exemplo, tipo de distribuição de probabilidade, média e variâncias populacionais conhecidas ou não, igualdade ou desigualdade das matrizes de covariâncias dos grupos, e assim por diante. No entanto, geralmente, essas informações a respeito da população não estão disponíveis, então, é necessário que o pesquisador realize os cálculos com base em suposições fundamentadas numa teoria subjacente, na própria experiência e que possam ser verificadas mediante testes estatísticos. Neste texto é abordado apenas o caso em que se supõe que as matrizes de covariância são iguais. A teoria a respeito dos demais casos pode ser obtida em Johnson e Wichern (2007).

O processo de construção da função discriminante inicia com a definição dos g grupos compostos por (n_1, n_2, \dots, n_g) observações cada tomadas em relação a p variáveis. O que se pretende é transformar estas observações multivariadas em univariadas, de tal forma que as populações (grupos) transformadas estejam separadas tanto quanto possível para facilitar a indicação (pertinência) de novas observações a um destes grupos (JOHNSON e WICHERN, 2007).

Para o caso em que os g grupos possuem matrizes de covariância iguais, são necessárias s funções discriminantes (y), com $s = \min(g - 1, p)$, as quais podem ser escritas como combinações lineares das p variáveis originais (x). Esta formulação (Equação 37) é conhecida como função discriminante de Fisher.

$$y_k = \hat{a}_k' x \quad (37)$$

Onde: \hat{a}_k é o vetor de coeficientes que satisfaz a condição apresentada na Equação 38, com $k \leq s$, sendo estimado pelos autovetores associados aos autovalores não-nulos da matriz $W^{-1}B$, tal que W , B e S_p são obtidas por meio das Equações 39, 40 e 41.

$$\hat{a}_k' S_p \hat{a}_k = 1 \quad (38)$$

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)' \quad (39)$$

$$B = \sum_{i=1}^g n_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \quad (40)$$

$$S_p = \frac{W}{n_1 + n_2 + \dots + n_g - g} \quad (41)$$

Onde:

W é a matriz de soma de quadrados dentro dos grupos

B é a matriz de soma de quadrados entre os grupos

S_p é a matriz de covariância amostral combinada

\bar{x}_i é o vetor de médias amostrais do i -ésimo grupo

\bar{x} é o vetor de médias amostrais geral

Cada função discriminante apresenta um poder de discriminação entre grupos diferente, o qual pode ser avaliado por meio da adoção de um ou mais critérios, a saber: proporção dos autovalores elevada (Equação 42); coeficiente de correlação canônica elevado (Equação 43) e teste qui-quadrado para o *lambda* de Wilks (Equação 44).

$$P_{\lambda_k} = \frac{\lambda_k}{\sum_{j=1}^p \lambda_j} \quad (42)$$

$$R = \sqrt{\frac{\lambda_k}{1 + \lambda_k}} \quad (43)$$

$$\chi_v^2 = \left(n_1 + n_2 + \dots + n_g - 1 - \frac{p + g}{2} \right) 1_n \Delta' \quad (44)$$

Onde:

λ_k é o k -ésimo autovetor da matriz $W^{-1}B$

$v = (p - k)(g - k - 1)$ graus de liberdade

p é o número de variáveis consideradas no estudo

$\Delta = \prod_{k=1}^p \frac{1}{1 + \lambda_k}$ é a estatística *lambda* de Wilks

Os componentes da função linear de Fisher possuem variância unitária e covariância nula. Sendo assim, é possível estabelecer uma regra para a classificação de novas observações. De acordo com esta regra (Equação 45), uma nova observação multivariada \underline{x} é alocada no grupo g_k , se o quadrado da distância entre \underline{x} e a média do grupo g_k for menor do que o quadrado da distância entre \underline{x} e a média do grupo g_i , para $i \neq k$.

$$\sum_{j=1}^r (\hat{y}_j - \bar{y}_{kj})^2 = \sum_{j=1}^r [\hat{a}'_j (\underline{x} - \bar{\underline{x}}_k)]^2 \leq \sum_{j=1}^r [\hat{a}'_j (\underline{x} - \bar{\underline{x}}_i)]^2 \quad ; \forall i \neq k ; \forall r \leq s \quad (45)$$

2.2.6 Cartas de Controle

Uma carta de controle é uma representação gráfica que apresenta a série de dados de uma variável, com uma linha central, correspondente à média, e limitada superior e inferiormente de acordo com o intervalo de confiança para a variável.

Os limites de controle são escolhidos apropriadamente de acordo com a variabilidade do processo e o nível de confiança adotado, de maneira que, se o processo for estável (sob controle estatístico), a totalidade dos pontos amostrais estará entre eles. A distribuição de uma variável que está sob controle estatístico é estável e previsível (MONTGOMERY, 2009).

Para o monitoramento de características quantitativas como, por exemplo as medidas físicas de um produto, Montgomery (2009), indica o uso das cartas de controle por variáveis, as quais permitem monitorar a média, o desvio-padrão e a amplitude das medidas. Os limites de controle para esses parâmetros podem ser estipulados por meio de intervalos de confiança baseados na adoção de porcentagens da função de distribuição da variável em questão.

Para amostras de tamanho grande o cálculo dos limites de controle independe da distribuição de probabilidade da variável, pois o Teorema Central do Limite garante que, neste caso, as médias amostrais serão normalmente distribuídas. Não há um consenso na literatura pesquisada a respeito do tamanho que a amostra deve assumir para ser considerada grande (JAMES, 2004).

Para uma amostra de tamanho grande ($n > 20$) de uma variável aleatória X , os limites superior e inferior de controle para a média dessa variável são estimados pelas Equações 46 e 47, respectivamente (CHAVES NETO, 2009).

$$LSC = \bar{x} + z_{(1-\alpha/2)} \frac{s}{\sqrt{n}} \quad (46)$$

$$LIC = \bar{x} - z_{(1-\alpha/2)} \frac{s}{\sqrt{n}} \quad (47)$$

Onde:

LSC é o limite superior de controle

LIC é o limite inferior de controle

\bar{x} é a média amostral

s é o desvio padrão amostral

$z_{\alpha/2}$ é o escore padronizado

O escore padronizado é definido com base na distribuição Normal Padrão, $Z \sim \mathcal{N}(0, 1)$, e refere-se à probabilidade $(1 - \alpha)$ da verdadeira média (populacional) pertencer ao intervalo definido. Geralmente adota-se $\alpha = 0,05$ ou $\alpha = 0,10$, o que resulta em intervalos com 95% ou 99% de confiança, respectivamente.

Uma carta de controle envia sinais estatísticos que permitem detectar a existência de uma causa especial de variabilidade ou apontam que esta pode ser atribuída ao acaso. Sendo assim, é uma ferramenta que pode ser usada para auxiliar a tomada de decisão pela equipe responsável pelo monitoramento do processo (ORO, 2010).

Na eventualidade de um ou mais pontos ficarem fora de um dos limites, evidenciará uma mudança no processo, podendo indicar que o mesmo está fora de controle, sendo necessária uma investigação e a correspondente ação corretiva para encontrar e eliminar as possíveis causas do problema. Outro indicativo de que um processo não é estável é a localização de maneira sistemática, não-aleatória ou com tendência, de uma sequência de pontos, mesmo dentro dos limites de controle (CHAVES NETO, 2009).

2.3 MÉTODOS ESTATÍSTICOS APLICADOS NO MONITORAMENTO DE ESTRUTURAS

A análise detalhada dos dados da instrumentação de auscultação de barragens, de acordo com Villwock *et al.* (2013), requer uma combinação de conhecimentos, especialmente de Engenharia, Matemática e Estatística, e experiência anterior da equipe técnica responsável pela análise destes dados. Isto pode consumir muito tempo, e muitas vezes faz com que seja impossível realizar essa tarefa de forma eficiente. Por isso, técnicas computacionais são indicadas como ferramentas auxiliares na importante tarefa de tomada de decisões.

Os algoritmos usados na modelagem estatística aplicada no monitoramento de grandes estruturas, tais como pontes e barragens, de acordo com Figueiredo *et al.* (2011), geralmente podem ser agrupados em três categorias: (i) classificação; (ii) análise de regressão; e (iii) detecção dos valores atípicos. O algoritmo a ser usado depende do modo de aprendizado ser supervisionado ou não. No aprendizado supervisionado tem-se conhecimento prévio de exemplos de dados nas condições intactas e danificadas das estruturas, enquanto que na aprendizagem não supervisionada estão disponíveis apenas os dados do estado não danificado das mesmas.

Quando aplicados em um modo de aprendizado não supervisionado, modelos estatísticos são normalmente usados para responder a questões referentes à existência e localização de danos, por meio da detecção de valores atípicos (*outlier*). Por outro lado, as aplicações de procedimentos estatísticos em um modo de aprendizado supervisionado, juntamente com modelos analíticos, tem objetivado determinar o tipo de dano, sua extensão e vida útil da estrutura remanescente.

Como os dados podem representar medidas de diferentes naturezas ou escalas, é importante a padronização dos mesmos para possibilitar o processo de identificação de danos. Em se tratando de monitoramento de estruturas, a padronização dos dados é um procedimento inerente à aquisição de dados, extração de características e modelagem estatística, que permite distinguir as alterações nas leituras dos sensores causadas por danos, daquelas causadas pela variação das condições ambientais e operacionais (FARRAR e WORDEN, 2007; FIGUEIREDO *et al.*, 2011).

Além disso, de acordo com Farrar e Worden (2007), é necessário identificar e minimizar as fontes de variabilidade no processo de aquisição de dados e do sistema monitorado. No entanto, nem todas as fontes de variabilidade podem ser eliminadas. Como por exemplo, a variação causada por diversas condições ambientais, como temperatura, umidade, carregamento e condições de contorno. Por isso, é necessário fazer as medições apropriadas de tal modo que essas fontes possam ser quantificadas estatisticamente.

A aplicação de um procedimento de limpeza de dados é indicada por Farrar e Worden (2007) para aceitar ou rejeitar o conjunto de dados no processo de seleção de características. Por exemplo, uma inspeção pode revelar que um sensor

foi instalado de forma diferente da usual e, portanto, com base no julgamento da equipe responsável pelo monitoramento das medições, este conjunto de dados pode ser desconsiderado no processo de seleção de características, uma vez que poderia influenciar a variabilidade dos dados.

Além disso, a aquisição e limpeza de dados, bem como a padronização, deve ser contínua nas atividades de monitoramento de estruturas, pois o conhecimento adquirido nos processos de seleção de características e de desenvolvimento do modelo estatístico fornecerá informações norteadoras para a realização de mudanças que podem melhorar o processo monitorado (FARRAR e WORDEN, 2007).

Em situações que possuem incertezas inerentes ao sistema adotado, segundo Farrar e Worden (2007), a análise estatística pode ser usada para classificar a alteração dos parâmetros como proveniente da modificação da condição estrutural (falhas) ou modificação das condições ambientais e/ou operacionais.

Na literatura sobre monitoramento de estruturas, de acordo com Sohn (2006), são poucos os modelos estatísticos aplicados que são voltados à verificação das mudanças observadas nos parâmetros indicadores de falha. Segundo Farrar e Worden (2007), pouca atenção tem sido dada ao desenvolvimento de modelos estatísticos para a discriminação entre as características das estruturas intactas e danificadas. O desenvolvimento desses modelos envolvem a implementação de algoritmos que operam com os recursos extraídos para quantificar o estado de dano da estrutura.

Os dados brutos, por si só nem sempre indicam a presença de danos em um estágio inicial. Por isso, é necessário o uso de métodos de seleção de características indicativas de danos precoces para auxiliar no monitoramento de estruturas. Os métodos mais utilizados para essa finalidade são modelos auto-regressivos, componentes *Wavelet*, análise de componentes principais e análise de agrupamentos (POSENATO *et al.*, 2010; SANTOS *et al.*, 2013).

A detecção de danos em grandes estruturas, quando as características estruturais indicam comportamento altamente não-linear, devido à variação das condições ambientais, pode ser feita a partir dos fatores ambientais subjacentes extraídos por meio da análise não linear de componentes principais seguida da utilização das redes neurais auto-associativas (HSU e LOH, 2010).

Os modelos estatísticos também são usados para minimizar situações de alarmes falsos de danos. Segundo Figueiredo *et al.* (2011), tais situações se dividem em duas categorias: (i) dano falso-positivo (indicação de danos quando não esteja presente) e (ii) dano falso-negativo (sem indicação de danos quando o dano está presente).

De acordo com Li, Wang e Liu (2013), a interação de grandes barragens com fatores ambientais, hidráulicos e geomecânicos, como a pressão hidrostática, temperatura do concreto e efeito do tempo é particularmente forte. Por isso modelos que levem em consideração esses fatores constituem uma ferramenta poderosa para analisar e prever a ocorrência de deslocamentos nas estruturas das barragens. Um tipo de modelagem que atende a essa demanda é a regressão linear, cuja aplicação permite, por exemplo, prever os deslocamentos horizontais dos blocos de uma barragem de contrafortes, com base na temperatura ambiente, na pressão hidrostática e na tendência dos dados (DE SORTIS e PAOLIANI, 2007).

Segundo Figueiredo *et al.* (2011), todas as fontes de variação devem ser bem caracterizadas pelos dados de simulação a fim de que os algoritmos detectem com precisão a sua influência sobre a resposta do sistema. Por exemplo, as variações sazonais de temperatura podem produzir mudanças significativas nas características de grandes estruturas, por isso devem ser consideradas nas análises dos dados dessas estruturas.

A regressão linear é um dos modelos mais utilizados para a análise das séries temporais da instrumentação de monitoramento de barragens. No entanto, de acordo com Li, Wang e Liu (2013), deve-se tomar cuidado com esta técnica, pois um falso modelo de regressão pode ser obtido quando a estacionariedade das séries não é testada, diminuindo assim a precisão da previsão. Para testar a estacionariedade das séries temporais, são utilizados os testes de Dickey-Fuller¹⁰ e de Engle-Granger¹¹. Em seguida, os autores propõem a utilização de um modelo de correção de erro para representar o equilíbrio a longo prazo e o desequilíbrio a curto

¹⁰ O teste foi sugerido por Dickey e Fuller, em 1979, no artigo intitulado "*Distribution of the estimators for autoregressive time series with a unit root*", publicado no periódico "*Journal of the American Statistical Association*", com o intuito de verificar se um modelo autorregressivo tem ou não raiz unitária.

¹¹ Teste apresentado por Engle e Granger, em 1987, no artigo "*Co-integration and error correction: Representation, estimation and testing*", na revista "*Econometrica*", vol 55, nº2, para verificar se as séries cointegram.

prazo nas relações entre variáveis, a fim de melhorar a precisão das previsões, se existirem relações de cointegração¹² entre as variáveis em análise.

A presença de multicolinearidade entre as variáveis utilizadas para ajustar os modelos de regressão linear é outro problema que deve ser tratado, uma vez que prejudica a qualidade do ajuste, a representatividade do sistema pela variável dependente e compromete as interpretações dos resultados. No caso de aplicações aos dados de barragens, a multicolinearidade pode dificultar, por exemplo, a quantificação da contribuição de carregamentos externos para a ocorrência de deformações estruturais. A regressão por componentes principais ou por mínimos quadrados parciais são formas de contornar este problema (DENG, WANG e SZOSTAK-CHRZANOWSKI, 2008; XU, YUE e DENG, 2012).

Segundo Cury e Crémona (2010), a técnica multivariada da análise de componentes principais, quando aplicada no monitoramento de grandes estruturas, é útil também para a compressão e representação dos dados originais, sem perda significativa de informação. Essa técnica também é indicada por Yu *et al.* (2010) para a extração dos componentes que influenciam a variabilidade dos dados.

As técnicas multivariadas denominadas análise de agrupamentos e análise fatorial quando aplicadas aos dados de monitoramento de barragens permitem realizar o agrupamento dos instrumentos com comportamento semelhantes, bem como realizar a classificação hierárquica das datas das medições e dos próprios instrumentos (VILLWOCK, 2009).

As comunalidades, aferidas durante a aplicação da análise fatorial, representam um fator que distingue o comportamento do instrumento, possibilitando o controle de qualidade da medida do mesmo. Esses resultados podem servir como fonte de informação para a escolha dos instrumentos cujas leituras pretende-se intensificar, mesmo quando não se tem nenhum conhecimento prévio sobre sua localização ou qualquer outra característica peculiar (VILLWOCK *et al.*, 2013).

De acordo com Villwock *et al.* (2013), a hierarquização das datas em que as medições foram realizadas, pode ser feita com base nos escores fatoriais finais, que são obtidos (Equação 48) a partir dos escores de cada fator (f_i) ponderados pelos autovalores (λ_i) da matriz de correlações.

¹² Se duas séries não-estacionárias formarem um vetor de coeficientes que gerem resíduos estacionários, diz-se que estas séries cointegram. As séries não-estacionárias são, então, ditas integradas de ordem 1, enquanto que as séries estacionárias são ditas integradas de ordem zero.

$$score_{fatorial_final} = \frac{\sum_{i=1}^m \lambda_i f_i}{\sum_{i=1}^m \lambda_i} \quad (48)$$

Outra área em que os métodos estatísticos têm sido utilizados é no estabelecimento de valores de referência para os dados de monitoramento de grandes estruturas (KUPERMAN *et al.*, 2005; SILVA, DIAS JÚNIOR e LOPES JÚNIOR, 2005). Esses valores são específicos para cada situação, pois cada estrutura apresenta um comportamento único, que reflete as condições específicas do local e do projeto. Se os valores escolhidos não são suficientemente correlacionados com a série de observações, os técnicos podem não ser alertados quanto às alterações no comportamento estrutural revelado pelos instrumentos. Neste caso, a empresa operadora estaria adiando as ações corretivas necessárias para evitar um problema maior.

Caso os valores medidos pelos mesmos instrumentos continuem a variar dentro de um intervalo, mantendo todas as outras condições semelhantes às do passado, o comportamento estrutural permanecerá dentro da normalidade. No entanto, a ocorrência de valores fora dos limites estipulados indica alteração no sistema monitorado, que pode ter sido ocasionada devido à mudança das condições operacionais, ambientais e/ou estruturais (SILVA, DIAS JÚNIOR e LOPES JÚNIOR, 2005).

A metodologia proposta por Kuperman *et al.* (2005) para estimar os valores-limite inicia com a análise das séries temporais das leituras da instrumentação. A seleção dos instrumentos é feita com base no coeficiente de correlação ($\rho \geq 0,7$). Os valores de referência (VR) que delimitam a região de aceitação para as medições de um instrumento são calculados (Equação 49) com base na média (VM) e no desvio padrão dos dados (Sd), considerando um nível de confiança de 95% e um mínimo de 120 observações disponíveis.

$$VR = VM \pm 1,96Sd \quad (49)$$

Após definidos os valores de referência para o sistema monitorado, avisos de alerta e/ou de emergência devem ser emitidos quando são identificados pontos fora da região de aceitação ou quando as leituras mostram o início de uma tendência diferente do padrão estabelecido ao longo dos anos, que é dependente

das cargas atuantes sobre a barragem e de seus efeitos. A existência de uma situação de alerta implica, de acordo com Kuperman *et al.* (2005), na verificação da instrumentação, no aumento da frequência das leituras, na inspeção detalhada da barragem e na verificação do comportamento de todas as estruturas.

A adoção de valores de referência com base nos valores históricos da instrumentação tem auxiliado os técnicos responsáveis pelo monitoramento de barragens a verificar se o valor medido pela instrumentação condiz com o esperado. Eles têm sido úteis na indicação de erros de medição, na detecção da ocorrência de anormalidade do instrumento ou da estrutura, servindo principalmente para realizar uma análise preliminar sobre o comportamento de estruturas, evidenciar variações bruscas de valores e verificar a consistência com dados anteriores.

De acordo com Silva, Dias Júnior e Lopes Júnior (2005) a definição de valores limites, para algum parâmetro indicativo que separa o estado de falha da estrutura do estado sem dano, deve ocorrer em um modo não supervisionado, o que significa que os dados do sistema, quando este apresenta falhas, não estão disponíveis no momento da definição.

A metodologia utilizada por Silva, Dias Júnior e Lopes Júnior (2005) consiste na predição dos valores das séries temporais de uma variável, com a condição de estrutura sem falha, por meio de um modelo auto-regressivo, seguida da geração de gráficos de controle para o erro residual, que é a diferença entre a predição do modelo e a medida real.

A análise fatorial também permite extrair os efeitos da variação das condições ambientais sobre as respostas de uma barragem. De acordo com Cheng e Zheng (2013), as \underline{x} respostas de uma barragem monitorada por p sensores, $\underline{x} \in R_p$, são funções (Equação 50) das \underline{v} variáveis ambientais e de outros $\underline{\eta}$ fatores específicos que a influenciam.

$$\underline{x} = f(\underline{v}) + g(\underline{\eta}) \quad (50)$$

O efeito das variáveis ambientais pode ser estimado ($\hat{\underline{x}}$) por meio de uma composição de funções (Equação 51) que projeta \underline{v} sobre um conjunto de m variáveis latentes (ξ), com $m \leq p$. Essas variáveis latentes, consideradas intermediárias, não são diretamente observáveis, mas elas podem ser calculadas a partir dos dados de monitoramento.

$$\underline{\tilde{x}} = \Lambda(\theta(\underline{v})) = \Lambda\xi \quad (51)$$

Onde: θ é uma função não linear e Λ é uma função linear.

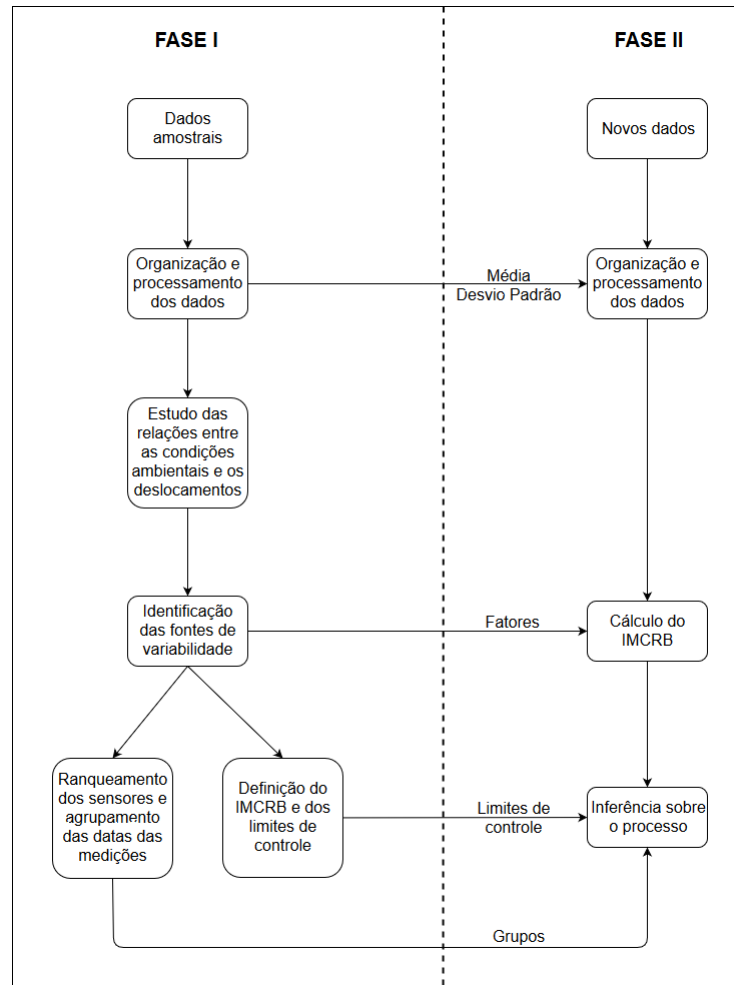
Geralmente, a influência dos fatores específicos sobre os dados de monitoramento é pequena, com a exceção das variáveis ambientais, o que implica em $g(\underline{\eta}) \ll \Lambda\xi$.

3 MATERIAL E MÉTODOS

Neste capítulo são apresentados os materiais, compostos pelos dados da instrumentação utilizados no monitoramento estrutural de dois blocos da barragem de Itaipu e o método desenvolvido para analisar os dados e auxiliar no monitoramento, baseado em métodos estatísticos.

O fluxograma (Figura 15) apresenta as duas fases e as respectivas etapas que compõem o método. Na primeira fase, foram aplicados métodos estatísticos para a modelagem e análise dos dados, objetivando principalmente a elaboração do Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem (IMCRB) e a definição dos limites de controle para o mesmo. Na segunda, o modelo do IMCRB foi validado mediante a utilização de dados atualizados, que não foram usados na Fase I.

FIGURA 15 – FLUXOGRAMA DO MÉTODO



FONTE: A autora (2016).

Estudos preliminares foram realizados em Oro, Chaves Neto e Neumann Jr (2014, 2015) e Oro *et al.* (2014, 2015) na tentativa de identificar os fatores que influenciam a variabilidade observada nos dados da instrumentação de monitoramento de barragens. Apenas o método que gerou os resultados mais promissores em termos de interpretação e aplicabilidade é apresentado neste trabalho.

Este terceiro capítulo inicia com a apresentação dos dados da instrumentação de monitoramento estrutural da barragem de Itaipu. Em seguida, são descritos os procedimentos adotados para o processamento dos dados. As seções seguintes, apresentam a forma como os métodos estatísticos: análise de correlação canônica, análise fatorial, análise de agrupamentos e análise discriminante, abordadas no capítulo 2, foram aplicados. Também é proposto o Índice de Monitoramento Conjunto das Respostas dos Blocos (IMCRB) da barragem,

considerando a ação da temperatura ambiente e do nível de água do reservatório. Os resultados numéricos obtidos são apresentados no capítulo 4.

3.1 SELEÇÃO DOS DADOS

O presente estudo utilizou uma base de dados reais do Sistema de Gestão de Medições (SGM) da barragem de Itaipu, disponibilizada pelo Centro de Estudos Avançados em Segurança de Barragens (CEASB) no formato de planilhas eletrônicas.

O conjunto de dados considerado neste trabalho é composto pelos valores aferidos manualmente na instrumentação dos blocos D7 e D8, do Trecho D, da barragem de Itaipu e pelos registros hidrometeorológicos, no período compreendido entre janeiro de 1990 e agosto de 2015.

A data inicial para a seleção dos dados, janeiro de 1990, foi sugerida pelos engenheiros que atuam no monitoramento da segurança estrutural da barragem sob a justificativa de que antes desta data a barragem ainda refletia os efeitos da construção e do enchimento do reservatório.

Os dados foram separados em dois subconjuntos: o primeiro constou das medições realizadas entre janeiro de 1990 e dezembro de 2013 e foi usado na Fase I do método; o segundo, composto pelas observações restantes, serviu para a validação do método (Fase II).

Para compor o conjunto de variáveis foram selecionados 38 sensores da instrumentação, os quais medem os possíveis deslocamentos das estruturas e fundações dos blocos D7 e D8, e outros quatro que informam as condições ambientais no entorno da barragem. O QUADRO 3 apresenta os fenômenos monitorados pelos instrumentos e os respectivos sensores, doravante também denominados variáveis observáveis.

QUADRO 3 – FENÔMENOS MONITORADOS PELOS INSTRUMENTOS E SEUS RESPECTIVOS SENSORES

FENÔMENOS	INSTRUMENTOS	SENSORES (VARIÁVEIS)	UNIDADE DE MEDIDA
Deslocamento radial	Pêndulos diretos	X1,X5	mm
Deslocamento radial	Pêndulo invertido	X3	mm
Deslocamento tangencial	Pêndulos diretos	X2,X6	mm
Deslocamento tangencial	Pêndulo invertido	X4	mm
Abertura e fechamento das juntas entre blocos	Bases de alongâmetro	X7, X9, X11, X13, X15, X17	10^{-3} mm
Deslizamento horizontal entre blocos	Bases de alongâmetro	X8, X12, X14, X18	10^{-3} mm
Recalque diferencial entre blocos	Bases de alongâmetro	X10, X16	10^{-3} mm
Temperatura na superfície do bloco (jusante)	Termômetros para concreto	X19	°C
Temperatura na superfície do bloco (montante) e da água do reservatório	Termômetros para concreto	X20	°C
Deformações do maciço rochoso	Extensômetros múltiplos de hastes	X21, ..., X40	mm
Temperatura ambiente	Termômetro	X41	°C
Nível de água do reservatório	Régua Limnimétrica	X42	m

FONTE: A autora (2016).

O conjunto de sensores usado no monitoramento dos deslocamentos considerados neste estudo foi: pêndulos diretos e invertidos, bases de alongâmetro e extensômetros múltiplos de hastes. A temperatura do concreto, a montante e a jusante, a temperatura ambiente e o nível de água do reservatório foram as variáveis consideradas como indicadoras das condições ambientais.

As análises e implementações, descritas nas seções a seguir, foram realizadas com o auxílio dos *softwares* Matlab (2015), *Statgraphics* (2010) e R (2015).

3.2 ORGANIZAÇÃO E PROCESSAMENTO DOS DADOS

Os dados utilizados tanto na primeira quanto na segunda fase do método foram codificados, tabulados e descritos, objetivando facilitar o conhecimento de suas características, possibilitar a aplicação das técnicas que compõem o método e favorecer a interpretação dos resultados.

O processo de tabulação foi realizado por meio do algoritmo "Gerar-Médias" (APÊNDICE A), implementado no Matlab, o qual também permitiu a eliminação dos dados redundantes¹³, a fixação da periodicidade mensal para as observações e a localização de lacunas no conjunto de dados. Optou-se pela periodicidade mensal porque esta é a utilizada atualmente para maioria dos sensores considerados neste estudo. No entanto, na década de 1990 as medições da instrumentação eram feitas com maior frequência e até hoje as observações das variáveis hidrometeorológicas são diárias, por isso utilizou-se a média mensal dos valores observados.

A existência de lacunas foi decorrente da ausência de medições no período (mês) e seu preenchimento foi feito por meio da modelagem e previsão de séries temporais.

A descrição dos dados foi realizada por meio dos gráficos das séries temporais, *Box-Plot* e de dispersão com o principal objetivo de identificar a presença de valores atípicos (*outlier*). Cada *outlier* detectado foi avaliado quanto à sua manutenção ou exclusão do conjunto de dados, de acordo com a justificativa de sua existência. Nos casos em que se optou pela exclusão, realizou-se a previsão de um novo valor, utilizando o mesmo procedimento do preenchimento das lacunas.

Dessa forma, foi composta a matriz de dados amostrais, com 308 linhas (médias mensais) e 42 colunas (variáveis observáveis). Optou-se pela padronização dos dados devido ao uso de instrumentos de espécies distintas, que geram valores diferentes em termos de grandeza e escala de medida. Os dados padronizados¹⁴ foram utilizados como entrada para a aplicação das técnicas e procedimentos das etapas seguintes do método.

¹³ Valores redundantes, neste estudo, foram considerados como aqueles dados que, por um motivo desconhecido, foram lançados de maneira repetida no sistema GSM, ou seja, eram compostos pelas medições realizadas nos mesmos sensores, com valores, datas e horários coincidentes.

¹⁴ Com a padronização dos dados, as variáveis originais (X_1, X_2, \dots, X_{42}) foram substituídas pelas respectivas variáveis padronizadas (Z_1, Z_2, \dots, Z_{42}).

Os dados padronizados foram separados em dois subconjuntos denominados "Dados Históricos" e "Dados Novos". Os Dados Históricos, correspondentes aos valores registrados desde o início do período considerado até dezembro de 2013, foram utilizados na Fase I, de desenvolvimento do método. Os Dados Novos, de janeiro de 2014 até agosto de 2015, foram usados na Fase II, para validar o método.

3.3 ANÁLISE DE CORRELAÇÃO CANÔNICA DE DESLOCAMENTOS E CONDIÇÕES AMBIENTAIS

A análise de correlação canônica foi utilizada para estudar a relação entre o conjunto de sensores que medem os deslocamentos (DES) das estruturas da barragem e de suas fundações e o conjunto de indicadores das condições ambientais (CAM).

Inicialmente, utilizando os dados padronizados, foram determinadas as matrizes de covariâncias entre as variáveis de cada grupo e das covariâncias cruzadas. A partir dessas matrizes foram extraídas as combinações lineares dos dois conjuntos que maximizassem as correlações entre eles.

Essas combinações, denominadas pares canônicos, foram dispostas em ordem decrescente de importância, utilizando como critério a correlação entre os dois conjuntos. Assim o primeiro par canônico formado foi aquele que apresentou maior correlação e que foi responsável pela maior porção da variância entre as variáveis. Cada par canônico formado apresentou variância unitária, correlação máxima e era não correlacionado com os demais pares.

Optou-se pela seleção do menor número de combinações que explicassem a maior parte da relação entre os dois conjuntos de variáveis.

A interpretação das variáveis canônicas presentes em cada par foi feita examinando-se as cargas canônicas, isto é, as correlações entre as variáveis originais (padronizadas) e as variáveis canônicas.

3.4 ANÁLISE FATORIAL

A análise fatorial foi utilizada para estimar a influência das condições ambientais (temperatura ambiente e nível de água do reservatório) nos deslocamentos, realizar o ranqueamento dos instrumentos segundo a sua importância no modelo fatorial e identificar os fatores que podem ser utilizados como parâmetros para a elaboração do Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem (IMCRB).

A matriz amostral foi composta pelo conjunto de Dados Históricos padronizados, de ordem 288x42. Cada coluna desta matriz guardava os dados de um sensor para o período de janeiro de 1990 até dezembro de 2013.

Inicialmente foram estimadas as matrizes de covariância e de correlações entre os sensores. Para verificar a viabilidade do uso da análise fatorial aplicou-se o teste de esfericidade de Bartlett, testando a hipótese de que a matriz de correlações era igual à matriz identidade. A matriz de correlações também serviu de base para o cálculo do coeficiente KMO, que é a outra técnica de verificação da adequação da amostra para a aplicação da análise fatorial. Tanto o teste de Bartlett quanto o coeficiente KMO foram obtidos com o auxílio do *software R* (2015), por meio dos algoritmos disponibilizados no Apêndice B.

Na sequência, os autovalores e os respectivos autovetores foram obtidos a partir da matriz de correlação. Os autovalores representam uma medida da importância do fator, enquanto que os autovetores formam um conjunto de vetores independentes (não-correlacionados entre si), que têm a propriedade de explicar a variabilidade total dos dados das medições da instrumentação.

Os autovalores foram dispostos em ordem decrescente de valor e plotados no gráfico *Scree-Plot*. A determinação da quantidade de fatores a serem extraídos levou em consideração esse gráfico, a importância dos autovalores (critério de Kaiser) e o percentual de explicação dos mesmos.

O método de extração de fatores adotado foi por componentes principais, porque este não requer que as variáveis sejam normalmente distribuídas. A correlação dos sensores (variáveis) com cada fator foi medida pelas cargas fatoriais, resultantes dessa decomposição. Para facilitar a interpretação dos fatores, utilizou-

se a transformação ortogonal denominada rotação Varimax, de tal forma que cada sensor apresentasse cargas elevadas em um fator e reduzidas nos demais.

No contexto deste estudo, a medida da comunalidade reflete o quanto da variância do sensor é devida aos fatores, representando, assim, a qualidade do ajuste do modelo fatorial aos dados e, principalmente, a importância do instrumento no contexto da análise. Usando a comunalidade como uma medida que distingue o comportamento do sensor, procedeu-se à hierarquização dos sensores, tal como descrito por Villwock (2009).

Como os fatores não podem ser diretamente medidos, seus valores, denominados escores fatoriais, foram estimados utilizando a matriz de cargas fatoriais e os dados amostrais. Esses escores formaram o conjunto de dados de entrada utilizado na modelagem do IMCRB.

3.5 MODELAGEM DO IMCRB

Os escores fatoriais serviram como dados de entrada para a modelagem do Índice de Monitoramento Conjunto das Respostas dos Blocos da Barragem (IMCRB). Por tratar-se de séries temporais, primeiro verificou-se a estacionariedade das séries dos fatores. Nos casos em que a série foi considerada não estacionária, realizou-se o número de diferenciações necessárias para torná-la estacionária, e se trabalhou com as séries diferenciadas na modelagem.

O modelo para o IMCRB foi elaborado por meio da média ponderada dos fatores, considerando como pesos os autovalores da matriz de correlações, conforme a equação (52). Nesse caso, os valores dos escores fatoriais padronizados são as entradas para a composição do índice, que é um valor adimensional.

$$IMCRB_i = \frac{\sum_{k=1}^m \lambda_k f_{ik}}{p} \quad \text{com } i = 1, 2, \dots, n \quad (52)$$

Onde:

λ_k é o k -ésimo autovalor da matriz de correlações;

f_{ik} é o i -ésimo escore do k -ésimo fator extraído;

p é o número variáveis originais (sensores);

n é o número de observações tomadas para cada variável.

Dessa forma, obteve-se uma série temporal para o IMCRB, com o mesmo tamanho do conjunto amostral, isto é, com 288 observações. Foram estimados, então, os limites de confiança (Equação 53) para essa série.

$$\begin{aligned}
 LIC_1 &= \bar{x}_{IMCRB} - 2s_{IMCRB} \\
 LSC_1 &= \bar{x}_{IMCRB} + 2s_{IMCRB} \\
 LIC_2 &= \bar{x}_{IMCRB} - 3s_{IMCRB} \\
 LSC_2 &= \bar{x}_{IMCRB} + 3s_{IMCRB}
 \end{aligned}
 \tag{53}$$

Onde \bar{x}_{IMCRB} e s_{IMCRB} são a média e o desvio padrão, respectivamente, da série histórica de valores obtidos para o Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem. Os limites LIC_1 e LSC_1 correspondem ao intervalo de confiança no qual se espera que 95% dos valores do IMCRB pertençam, enquanto que, com 99% de confiança, espera-se obter esses valores entre LIC_2 e LSC_2 . Os limites do primeiro intervalo podem ser considerados como de controle e os do segundo como de alerta para a variável monitorada.

Também foram realizadas as previsões para os valores do IMCRB para os próximos 24 meses, com base no ajuste automático do modelo ARIMA, realizado com o auxílio do pacote *forecast* do *software R*.

Por fim, os dados da série temporal desse índice, das previsões e os limites de confiança foram plotados simultaneamente num gráfico dando origem à carta de controle para a variável IMCRB.

Como a construção dos intervalos de confiança foram baseados na teoria das probabilidades, a carta de controle foi utilizada para verificar se o processo monitorado era estável, conforme os procedimentos descritos a seguir, adaptados de Cheng e Zheng (2013).

Se o IMCRB fosse superior a LSC_2 ou inferior a LIC_2 , ou, ainda, houvesse uma sequência sistemática de pontos (não aleatória), então as possíveis causas para essas anormalidades seriam: c_1 – valores extremos para as condições ambientais; c_2 – instrumentos com defeito ou em manutenção; c_3 – erros de

medição; c_4 – calibração de instrumentos; c_5 – envelhecimento da estrutura; c_6 – dano pontual na estrutura da barragem; c_7 – dano global na estrutura da barragem.

Para facilitar a leitura da carta de controle e apoiar a tomada de decisão, elaborou-se um sistema de classificação (QUADRO 4) para o estado do processo, conforme os possíveis casos de ocorrências para a disposição dos pontos no gráfico. Os casos (a) e (b) apresentados na Figura 16, adiante, ilustram um processo estável e, portanto, previsível. Por outro lado, (c) e (d) representam um processo fora de controle. A situação (c), caracterizada pela presença de um único valor extremo, fora dos limites de controle estabelecidos com 99% de confiança, pode ser justificada supostamente pelo conjunto de causas $\{c_1, c_2, c_3\}$. A ocorrência de pontos dispostos sistematicamente em ordem crescente, decrescente ou constante, ainda que dentro dos intervalos de confiança, caso (d), pode ter origem em qualquer uma das sete causas $\{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$.

QUADRO 4 - CLASSIFICAÇÃO DO PROCESSO MONITORADO E AS POSSÍVEIS CAUSAS ASSOCIADAS À OCORRÊNCIA DE ANORMALIDADES

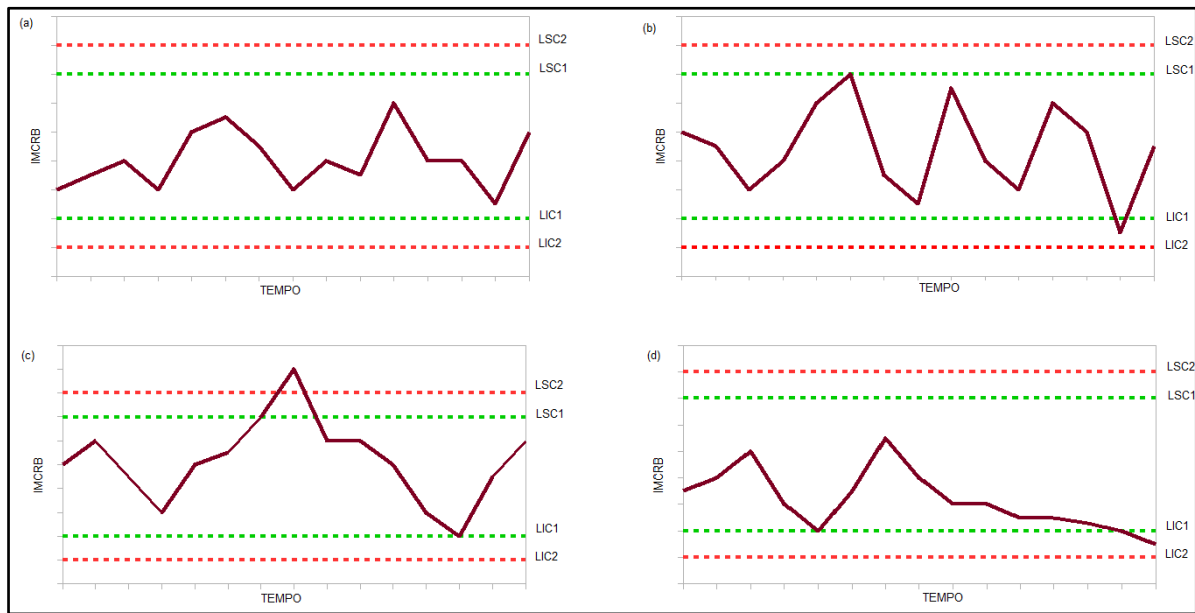
CARTA DE CONTROLE PARA O IMCRB	SITUAÇÃO DO PROCESSO MONITORADO	POSSÍVEIS CAUSAS
Caso (a)	Sob controle (estável)	–
Caso (b)	Sob controle (estável)	–
Caso (c)	Fora de controle (anormal)	$\{c_1, c_2, c_3\}$
Caso (d)	Fora de controle (anormal)	$\{c_1, c_2, c_3, c_4, c_5, c_6, c_7\}$

FONTE: A autora (2016).

As supostas causas listadas neste texto são meramente sugestivas, sendo baseadas nas principais ocorrências citadas na literatura pesquisada e nas informações recebidas durante as visitas técnicas realizadas pela pesquisadora nas dependências da barragem de Itaipu.

É importante salientar que os casos de anormalidade requerem estudos minuciosos a respeito das ações, condições analíticas e construtivas da barragem, tal como determina a legislação vigente sobre segurança de barragens, para a verificação dos parâmetros de segurança e a tomada das medidas de correção cabíveis.

FIGURA 16– EXEMPLOS DE DISPOSIÇÃO DOS VALORES DO IMCRB NA CARTA DE CONTROLE



FONTE: A autora (2016).

Após a modelagem do IMCRB e do estabelecimento dos intervalos de confiança para os seus valores, iniciou-se a Fase II do método, destinada à validação do mesmo.

O Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem, proposto neste trabalho, é uma variável que não pode ser medida diretamente e que é definida pelo modelo composto pelas variáveis latentes (fatores), originadas a partir das variáveis observáveis (sensores). Como não foi encontrada na literatura uma situação semelhante utilizou-se o conjunto de dados atualizados, compostos pelas observações registradas no período compreendido entre janeiro de 2014 e agosto de 2015, para validar os resultados obtidos.

3.6 ANÁLISE DE AGRUPAMENTOS E ANÁLISE DISCRIMINANTE

Os escores fatoriais foram submetidos à análise de agrupamentos para identificar as épocas (datas das medições) em que a instrumentação dos blocos D7 e D8 da barragem de Itaipu apresentaram desempenho semelhante. Nesta análise

utilizou-se o método de comparação hierárquico aglomerativo da média do grupo, com as distâncias calculadas pela métrica de Mahalanobis.

Os agrupamentos formados em cada etapa da aplicação da técnica foram representados graficamente por meio do dendrograma. Como o número de grupos não era conhecido antes da aplicação da análise buscou-se estabelecer uma linha de corte na altura em que fosse possível interpretar os agrupamentos formados.

Em seguida, os grupos formados foram testados mediante a análise discriminante, tendo como referência os sensores com melhor classificação na análise fatorial de acordo com as comunalidades, por considerar que estes possuíam melhor poder de distinção entre grupos.

4 RESULTADOS E DISCUSSÃO

4.1 PROCESSAMENTO DOS DADOS

Por meio da aplicação do algoritmo "Gerar_Medias", os dados dos 42 sensores foram tabulados e constatou-se a existência de 9 séries incompletas, totalizando 16 lacunas, decorrentes da inexistência de medições em algum período. O Quadro 5 apresenta os sensores e as respectivas datas em que foram detectadas as lacunas.

QUADRO 5 – SENSORES QUE APRESENTARAM SÉRIES INCOMPLETAS

DATA	SENSOR
Fev/92	X20
Jun/95	X9, X10, X11, X12
Mai/98	X27, X28
Jun/98	X27, X28
Ago/06	X19, X20
Abr/11	X40
Fev/13	X19, X20
Ago/13	X19, X20

FONTE: A autora (2016).

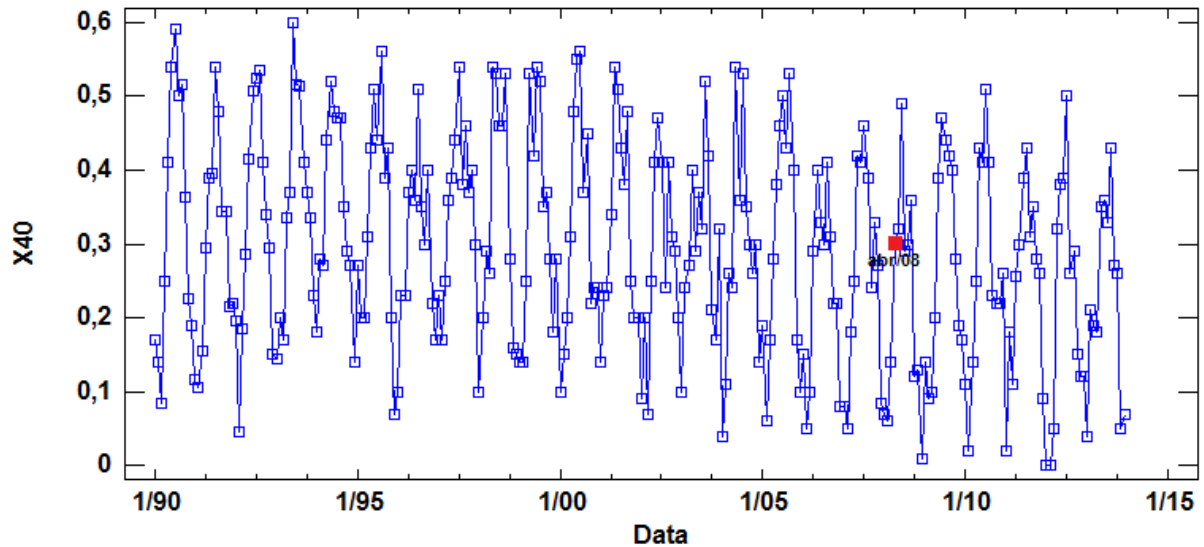
Foram criadas duas subséries para cada sensor em que detectou-se a ocorrência de lacunas: uma com as observações anteriores ao dado faltante outra com as posteriores. Os modelos ARIMA foram utilizados para a realização das previsões, com o auxílio do *software Statgraphics* (2010), adotando-se o procedimento *forecasting/backforecasting*, que consiste em modelar cada subsérie, fazer a previsão do valor e preencher a lacuna com o valor médio das duas previsões.

Por exemplo, o sensor X40 é uma haste de um extensômetro instalado na fundação do bloco D8 da barragem. Para este sensor observou-se a inexistência de leituras no mês de abril de 2008. Procedeu-se, então, ao estudo da série temporal das medições anteriores e posteriores a esta data.

Observando o gráfico da série temporal deste sensor (Figura 17), aparentemente este apresenta existência de sazonalidade (verão/inverno),

estacionariedade na média e na variância, não sendo necessária a realização de diferenciação e/ou transformação dos dados.

FIGURA 17– SÉRIE TEMPORAL DO SENSOR X40



FONTE: A autora (2016).

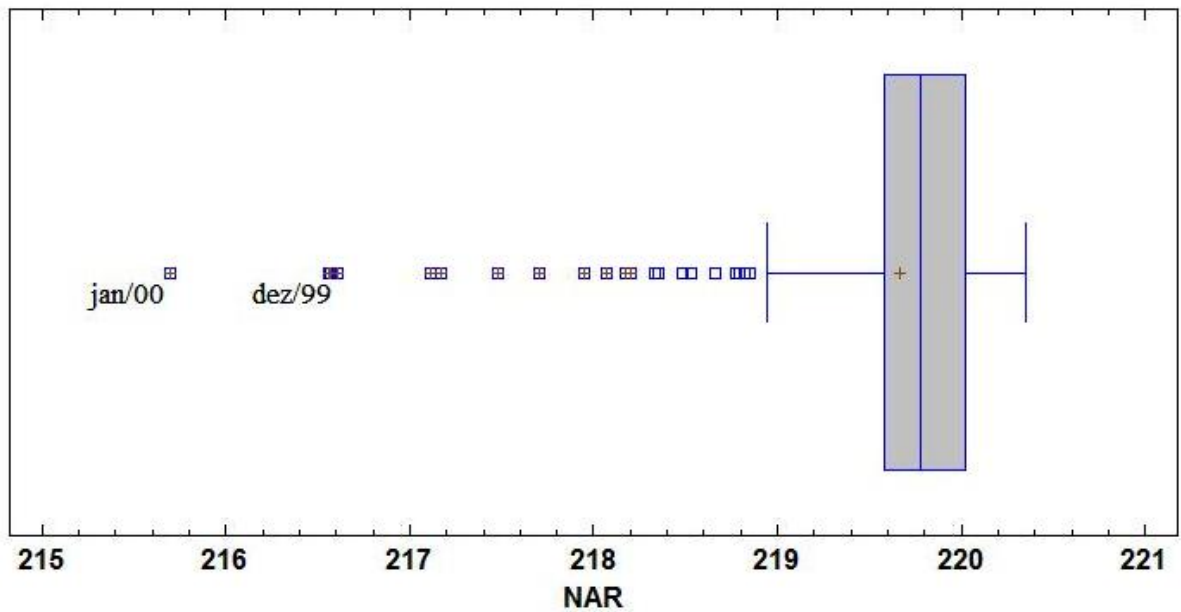
Procedeu-se, então, à modelagem de duas subséries constituídas pelas observações anteriores e posteriores a abril de 2008, com o auxílio do *software Statgraphics* (2010), utilizando o modo automático de seleção do melhor modelo, segundo o critério AIC. Para a série anterior à lacuna foi previsto um valor à frente (*forecasting*), enquanto que para a posterior foi previsto um valor um passo atrás (*backforecasting*). O valor previsto para preencher o dado faltante foi obtido por meio da média desses dois valores. De modo análogo ao descrito para o sensor X40, realizou-se o preenchimento das lacunas identificadas nos demais sensores.

Após completar as séries de dados dos sensores, foram geradas as medidas descritivas (APÊNDICE C). Observando o coeficiente de variação percebeu-se que, em geral, os dados de cada sensor apresentavam variabilidade elevada. Por este motivo e também porque havia diferença de unidade de medida e de grandeza entre os sensores, os dados foram padronizados.

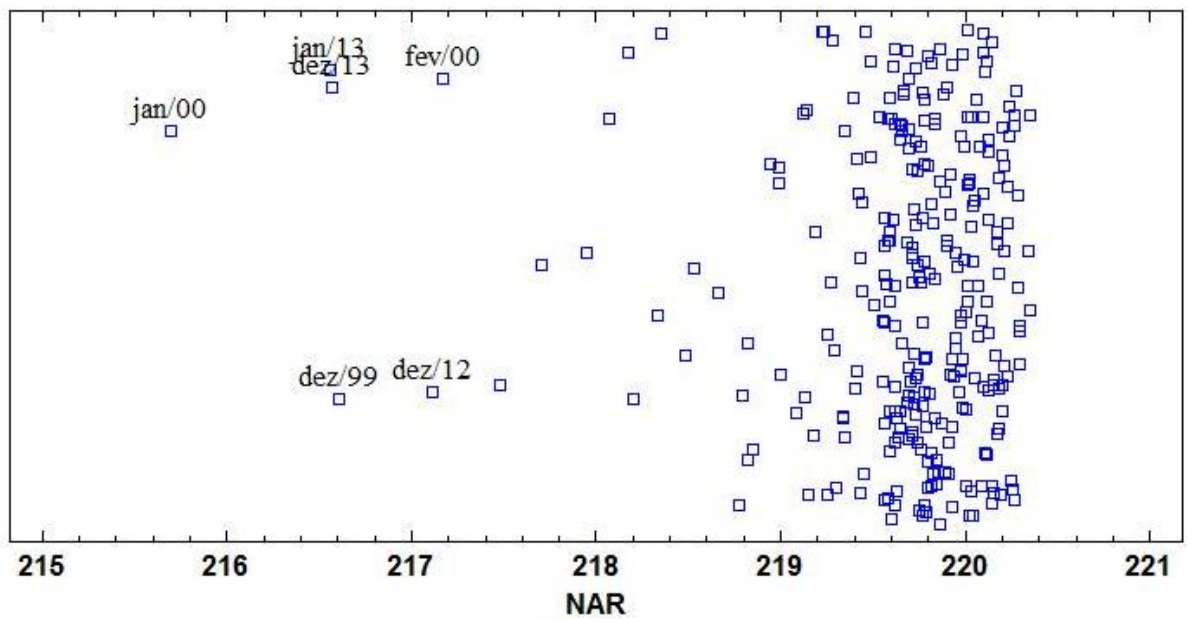
Procedeu-se à análise dos gráficos *Box-Plot* e de dispersão de cada série, em busca da ocorrência de valores atípicos. Conforme exposto na Figura 18, foram detectados vários valores atípicos nos dados do nível de água do reservatório (NAR). Segundo Neumann Jr (2015), a provável causa para essa ocorrência foi a incidência de estiagem (baixo índice pluviométrico) nos meses da estação do verão

dos anos 1999/2000 e 2012/2013. Optou-se por manter os valores tal como observados, para que fosse possível estudar a relação entre essas ocorrências e as respostas da barragem.

FIGURA 18 – GRÁFICOS UTILIZADOS NA IDENTIFICAÇÃO DOS VALORES ATÍPICOS PARA O NÍVEL DE ÁGUA DO RESERVATÓRIO



(a) Box Plot



(b) Dispersão
 FONTE: A autora (2016).

4.2 INFLUÊNCIA DAS CONDIÇÕES AMBIENTAIS NOS DESLOCAMENTOS DE UMA BARRAGEM DE CONCRETO

A análise de correlação canônica foi utilizada para estudar as relações existentes entre os grupos AMB e DES. O primeiro foi formado pelos sensores considerados como indicadores das condições ambientais (AMB), a saber: Z18, Z19, Z41 e Z42. Já o grupo DES foi composto pelos sensores que medem os deslocamentos, formado pelos demais sensores. Nesta análise trabalhou-se com os dados padronizados das variáveis (sensores).

Conforme apresentado na Tabela 7, todos os autovalores foram considerados significativos ao nível de confiança de 95% (valor-p < 0,05). Optou-se por discutir os resultados da correlação canônica de valor mais elevado.

TABELA 7 – CORRELAÇÃO CANÔNICA ENTRE OS CONJUNTOS DES E AMB

Autovalores	Correlação Canônica	Lambda de Wilks	Qui-quadrado	Graus de liberdade	Valor-p
$\lambda_1 = 0,973$	0,986	0,003	1537,08	152	0
$\lambda_2 = 0,717$	0,847	0,114	577,122	111	0
$\lambda_3 = 0,437$	0,661	0,402	242,06	72	0
$\lambda_4 = 0,286$	0,535	0,714	89,49	35	0

FONTE: A autora (2016).

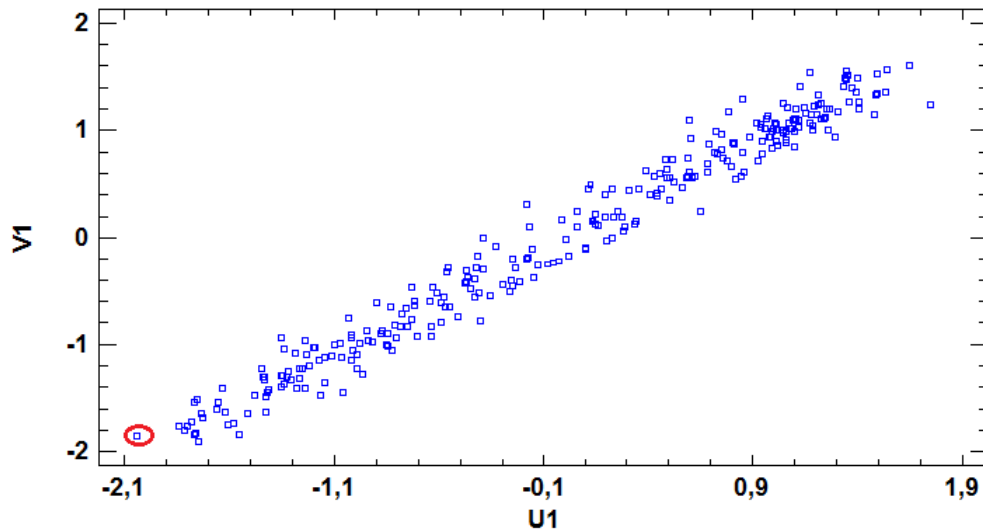
A correlação canônica entre os dois grupos foi de 0,986, indicando forte relação linear entre as variáveis que compõem o primeiro par de variáveis canônicas (U_1, V_1), conforme ilustrado na Figura 19. As Equações 54 e 55 apresentam as combinações lineares que originaram as variáveis canônicas U_1 e V_1 .

$$\begin{aligned}
 U_1 = & -0,04Z_1 - 0,09Z_2 - 0,24Z_3 + 0,02Z_4 - 0,04Z_5 - 0,05Z_6 - 0,24Z_7 \\
 & + 0,05Z_8 + 0,12Z_9 - 0,08Z_{10} - 0,04Z_{11} - 0,03Z_{12} + 0,26Z_{13} - 0,07Z_{14} \\
 & - 0,28Z_{15} + 0,11Z_{16} - 0,16Z_{17} - 0,07Z_{18} + 0,08Z_{21} - 0,07Z_{22} + 0,04Z_{23} \\
 & - 0,36Z_{24} + 0,06Z_{25} + 0,10Z_{26} + 0,02Z_{27} - 0,01Z_{28} - 0,04Z_{29} - 0,01Z_{30} \\
 & - 0,02Z_{31} - 0,02Z_{32} + 0,17Z_{33} - 0,28Z_{34} - 0,11Z_{35} + 0,04Z_{36} - 0,12Z_{37} \\
 & - 0,52Z_{38} + 0,06Z_{39} + 0,43Z_{40}
 \end{aligned} \tag{54}$$

$$V_1 = 0,20Z_{41} - 0,02Z_{42} + 0,12Z_{19} + 0,72Z_{20} \quad (55)$$

O ponto em destaque no gráfico refere-se às medições realizadas no mês de julho do ano 2000, época em que foi registrada uma das menores médias mensais para a temperatura ambiente.

FIGURA 19 – GRÁFICO DE DISPERSÃO DO PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS



FONTE: A autora (2016).

As estimativas das correlações entre cada um dos 38 sensores do grupo DES com os quatro do AMB também foram obtidas. A Tabela 8 apresenta a lista das variáveis com as maiores correlações ($\rho > |0,8|$ e valor- $p < 0,05$), com predominância dos extensômetros múltiplos de hastes (Z23, Z24, Z35, Z38, Z40) relacionados com as temperaturas (Z19, Z20 e Z41). Aparecem, também, três sensores de pêndulos (Z2, Z3 e Z5) associados às temperaturas na superfície do bloco a montante (Z20) e ambiente (Z41).

Na região onde está localizada a barragem de Itaipu, a amplitude das médias mensais das temperaturas ambiente, observadas num mesmo ano, pode chegar a 20°C. As correlações fortes dos sensores com as variáveis de temperatura confirmaram isto. As correlações negativas, com presença dominante nas primeiras colocações da lista, indicaram a presença de uma relação inversa entre as variáveis. Ou seja, nos períodos de temperatura baixa os deslocamentos foram maiores do que naqueles de temperatura elevada.

TABELA 8 – VARIÁVEIS FORTEMENTE CORRELACIONADAS

Sensores	Correlações
Z35 – Z20	-0,913
Z24 – Z20	-0,910
Z24 – Z41	-0,889
Z23 – Z20	-0,888
Z23 – Z41	-0,872
Z5 – Z41	-0,855
Z40 – Z19	-0,855
Z40 – Z41	-0,843
Z3 – Z41	-0,829
Z2 – Z20	-0,820
Z35 – Z41	-0,815
Z23 – Z19	-0,814
Z24 – Z19	-0,812
Z38 – Z19	-0,811

FONTE: A autora (2016).

Por outro lado, o nível de água do reservatório apresentou pequena correlação positiva individualmente apenas com alguns sensores. Possivelmente porque a baixa variabilidade observada nessa variável levou a forças atuantes sobre a barragem praticamente constantes. Outra justificativa pode ser a necessidade de interação do nível de água do reservatório com a temperatura para influenciar os deslocamentos.

A Tabela 9 apresenta os sensores mais correlacionados ($\rho > |0,9|$ e valor- $p < 0,05$) com o primeiro par de variáveis canônicas. Observa-se que as hastes dos extensômetros múltiplos Z23, Z24, Z35 foram os sensores mais influenciados pelas temperaturas Z20, Z41.

TABELA 9 – PRINCIPAIS CORRELAÇÕES ENTRE O PRIMEIRO PAR DE VARIÁVEIS CANÔNICAS E OS SENSORES DE CADA GRUPO

Sensor	U1	Sensor	V1
Z23	-0,924	Z20	0,989
Z24	-0,943	Z41	0,936
Z35	-0,926	-	-

FONTE: A autora (2016).

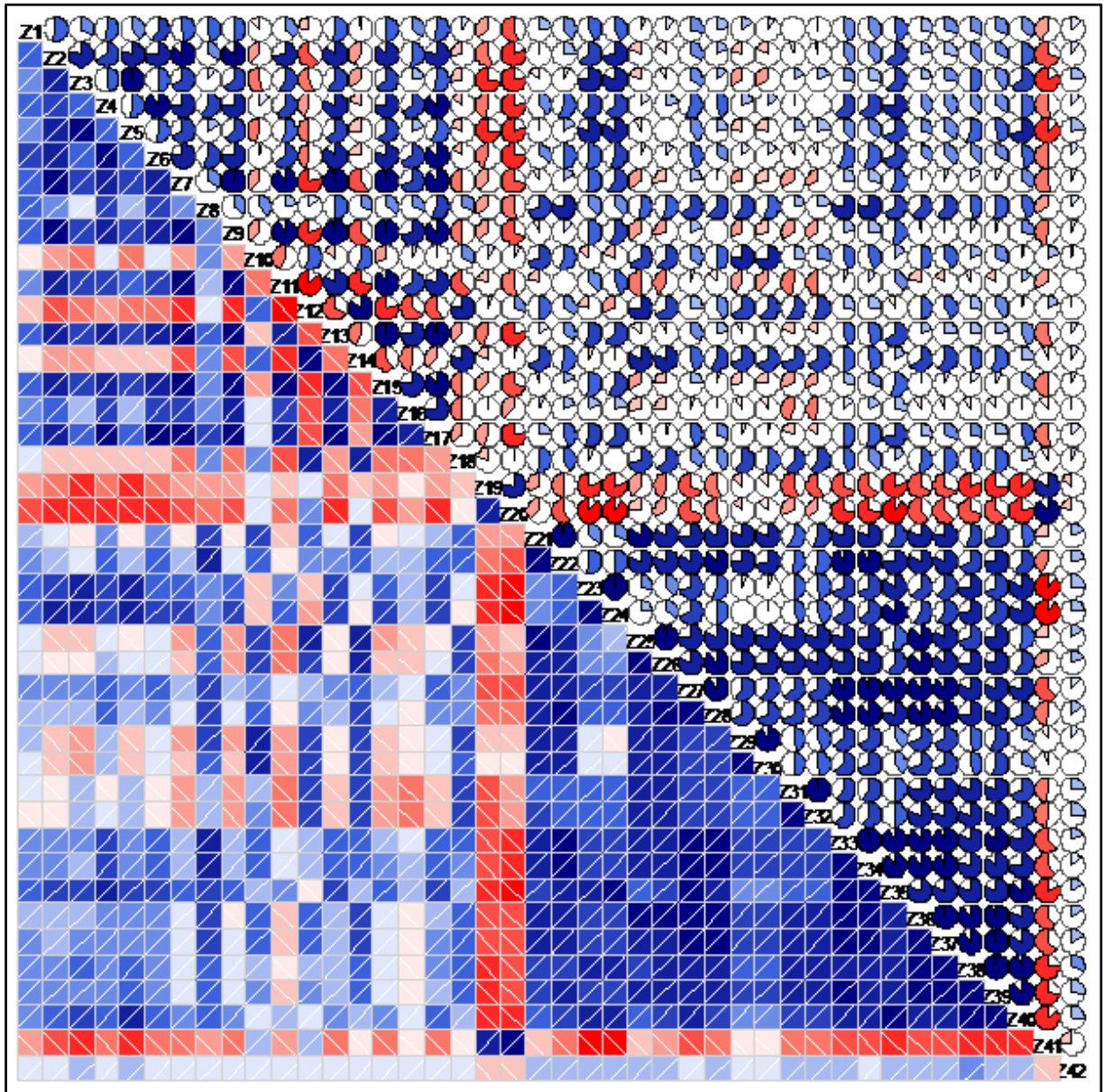
A avaliação da qualidade do potencial das variáveis canônicas foi feita com base na proporção da variância explicada pelas variáveis canônicas para cada

grupo. A variável canônica U1 explicou 38,7% da variância observada nos deslocamentos, enquanto que a proporção da variância explicada por V1 para o grupo condições ambientais foi de 65,6%. Sendo assim, os grupos “Deslocamentos” e “Condições Ambientais” foram bem representados pelo primeiro par de variáveis canônicas, uma vez que a correlação canônica entre esses grupos foi de 0,986, enquanto que os outros pares apresentaram valores menores.

4.2.1 Análise Fatorial Aplicada aos Dados da Instrumentação dos Blocos D7 e D8 do Trecho D da Barragem de Itaipu

A aplicação da análise fatorial foi feita com base na matriz de correlações dos 42 sensores considerados neste estudo. Por meio do correlograma (Figura 20) foi possível visualizar a intensidade dessas correlações. Quanto mais intensa a cor e maior o ângulo indicado, maior foi o valor dessa medida. As cores vermelha e azul foram usadas para indicar o tipo de correlação: inversa (negativa) e direta (positiva), respectivamente. A maioria das correlações foram significativas (valor- $p < 0,05$). No entanto, como a disponibilidade era de 288 observações por sensor, até mesmo valores baixos de correlação foram significativos.

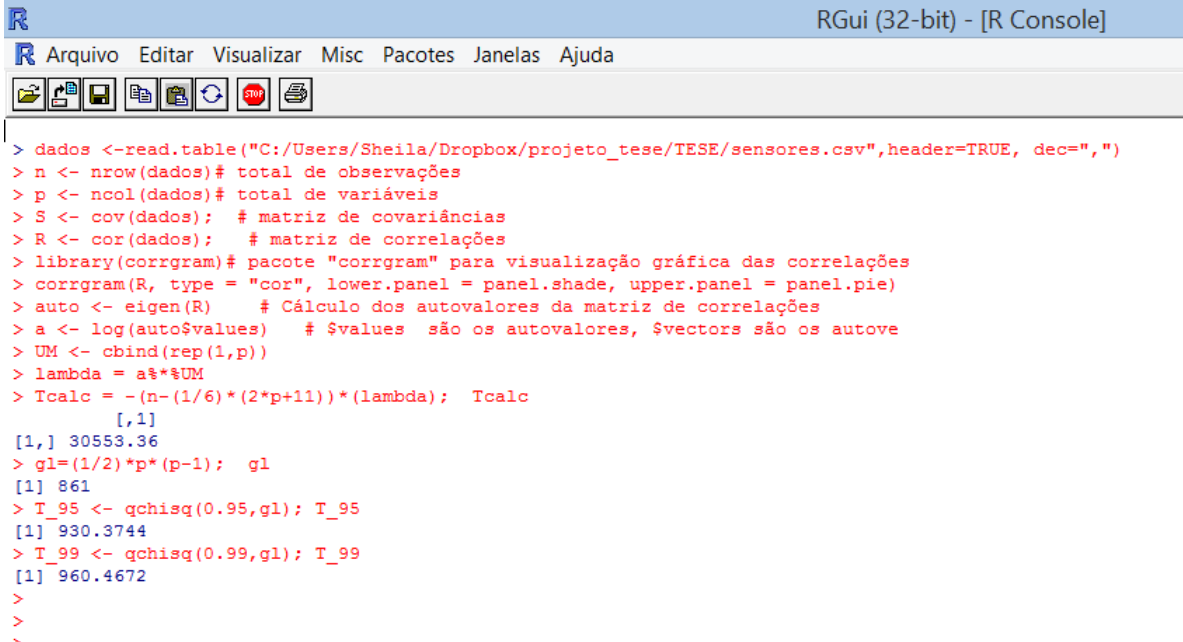
FIGURA 20– CORRELOGRAMA DOS 42 SENSORES



FONTE: A autora (2016).

Para certificar-se da significância das correlações, aplicou-se o teste de esfericidade de Bartlett, por meio de um algoritmo (Figura 21) implementado no *software R* (2015). O valor calculado para a estatística T (30553,36) foi muito superior aos valores críticos a 95% e 99% de confiança, indicando que coletivamente os sensores estavam intercorrelacionados significativamente, o que possibilitou a aplicação da análise fatorial.

FIGURA 21 – RESULTADO DO TESTE DE BARTLETT



```

RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda

> dados <- read.table("C:/Users/Sheila/Dropbox/projeto_tese/TESE/sensores.csv",header=TRUE, dec=",")
> n <- nrow(dados)# total de observações
> p <- ncol(dados)# total de variáveis
> S <- cov(dados); # matriz de covariâncias
> R <- cor(dados); # matriz de correlações
> library(corrgram)# pacote "corrgram" para visualização gráfica das correlações
> corrgram(R, type = "cor", lower.panel = panel.shade, upper.panel = panel.pie)
> auto <- eigen(R) # Cálculo dos autovalores da matriz de correlações
> a <- log(auto$values) # $values são os autovalores, $vectors são os autovetores
> UM <- cbind(rep(1,p))
> lambda = a*$UM
> Tcalc = -(n-(1/6)*(2*p+1))*(lambda); Tcalc
      [,1]
[1,] 30553.36
> g1=(1/2)*p*(p-1); g1
[1] 861
> T_95 <- qchisq(0.95,g1); T_95
[1] 930.3744
> T_99 <- qchisq(0.99,g1); T_99
[1] 960.4672
>
>
>

```

FONTE: A autora (2016).

Para obter uma medida da adequação da amostra para o modelo fatorial calculou-se o coeficiente KMO com o auxílio de uma rotina (Figura 22) implementada no *software R* (2015). O resultado obtido (0,9345) demonstrou uma excelente adequação.

Sendo assim, iniciou-se a seleção do número de fatores utilizando o critério de Kaiser. Verificou-se a possibilidade de utilizar quatro ou cinco fatores. Ao selecionar quatro fatores, observou-se que um sensor (Z1) ficou com comunalidade muito baixa (0,39). Uma comunalidade baixa indica que a variável não é importante na análise fatorial e poderia ser descartada. Porém, ao se trabalhar com cinco fatores todos os sensores obtiveram comunalidades superiores a 0,60. Como o quinto autovalor apresentou um valor muito próximo da unidade, optou-se pela extração de 5 fatores e a manutenção de todos os sensores no modelo fatorial. O gráfico do *Scree Plot* (

Figura 23) apontou claramente para 5 fatores, pois apresentou um nítido formato de cotovelo.

FIGURA 22 – RESULTADO DO CÁLCULO DO COEFICIENTE KMO

```

RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda

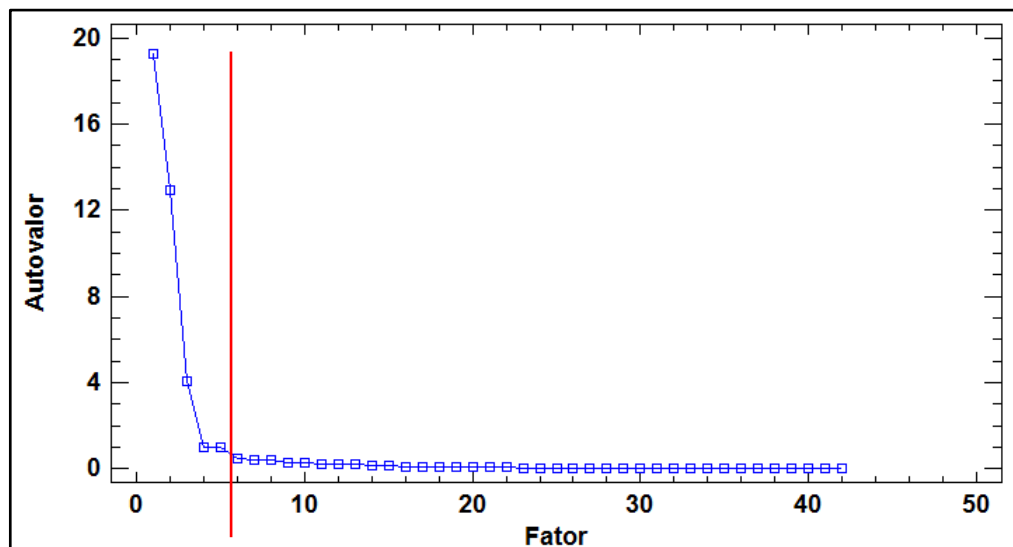
>
>
> partial.cor <- function (X)
+ {
+   RI <- solve(R)
+   D <- 1/sqrt(diag(RI))
+   Rp <- -RI * (D %o% D)
+   diag(Rp) <- 0
+   rownames(Rp) <- colnames(Rp) <- colnames(X)
+   Rp
+ }
>
> matcorp <- partial.cor(dados)# matriz das correlações parciais
>
> idiag <- seq(1, by = p + 1, length = p)
> somar2 <- sum((as.numeric(R)[-idiag])^2)
> cat("\n KMO = ",somar2 / (somar2 + sum((as.numeric(matcorp)[-idiag])^2)))

KMO = 0.9345345>
>
<

```

FONTE: A autora (2016).

FIGURA 23 – SCREE PLOT DOS AUTOVALORES DA MATRIZ DE CORRELAÇÕES AMOSTRAIS



FONTE: A autora (2016).

Sendo assim, procedeu-se à aplicação da análise fatorial por componentes principais, com o auxílio do *software Statgraphics* (2010), resultando num modelo composto por cinco fatores, responsáveis por aproximadamente 91% da variância observada do conjunto de dados. Dessa forma, obteve-se a redução na dimensão do problema, com mínima perda de informação, pois os fatores passaram a representar o conjunto de variáveis originais. A Tabela 10 apresenta os maiores

autovalores medidos, a porção da variância explicada por cada fator e o percentual da variância acumulada.

TABELA 10 – MEDIDAS DOS MAIORES AUTOVALORES, COM AS CORRESPONDENTES PERCENTAGENS DAS VARIÂNCIAS EXPLICADAS E VARIÂNCIAS ACUMULADAS

Fator	Autovalor	Variância Explicada (%)	Variância Acumulada (%)
1	19,269	45,879	45,879
2	12,948	30,828	76,707
3	4,059	9,664	86,37
4	1,013	2,413	88,783
5	0,981	2,335	91,118
6	0,507	1,208	92,326
7	0,441	1,049	93,376

FONTE: A autora (2016).

Os fatores foram nomeados de acordo com os sensores mais correlacionados com eles, isto é, conforme as maiores cargas fatoriais (Tabela 11). O primeiro fator, por apresentar correlação positiva com a maioria dos sensores de extensômetros múltiplos de hastes, foi identificado como “Movimentação da Fundação” (MF). Este foi o mais importante dentre os fatores extraídos, pois respondeu por 45,88% da variabilidade observada no conjunto de dados.

O segundo fator explicou 30,83% da variância e estava associado aos sensores das bases de alongâmetro e aos pêndulos diretos na direção margem direita-esquerda, que monitoram as aberturas e fechamentos das juntas entre blocos da barragem e os deslocamentos horizontais no sentido perpendicular ao do fluxo da água, por isso recebeu a denominação “Movimentação Horizontal Relativa entre os Blocos D7 e D8” (MHRB).

A “Movimentação Horizontal dos Blocos D7 e D8 por Influência Térmica Ambiental” (MHBITA), terceiro fator identificado, foi responsável por 9,66% da variabilidade. As variáveis dominantes neste fator foram os sensores das temperaturas (ambiente, na superfície do concreto a montante e a jusante) e dos deslocamentos horizontais da crista do bloco D8 no sentido montante-jusante (fluxo da água) medido pelos pêndulos diretos.

O quarto e o quinto fatores, denominados “Movimentação Horizontal Relativa do Bloco D7 em relação à Fundação” (MHRBF) e “Pressão Hidrostática” (PHI), estavam relacionados aos deslocamentos horizontais no sentido montante-

jusante da base do bloco D7 em relação à fundação, medidos pelo pêndulo invertido, e ao nível de água do reservatório, respectivamente. Além disso, o fator MHRBF respondeu por 2,41% da variabilidade total, enquanto que a influência de PHI foi de 2,34%.

Cada um desses últimos fatores teve associação com apenas um sensor, a saber: MHRBF com Z1 e PHI com Z42. Em especial, PHI mede o quanto as variações no nível de água do reservatório interferem nas respostas dos blocos D7 e D8, mesmo a amplitude dos valores medidos por esta variável sendo pequena, menor que 5 metros. Estas informações podem ser muito importantes para a equipe de engenheiros responsável pelo monitoramento estrutural da barragem, pois trata-se de um resultado único, sendo que não foram encontradas na literatura referências a esse respeito.

TABELA 11 – CARGAS FATORIAIS, COMUNALIDADES E VARIÂNCIAS ESPECÍFICAS DOS SENSORES

Sensor	MF	MHRB	MHBITA	MHRBF	PHI	Comunalidade	Variância Específica
Z34	0,82	0,28	0,46	0,07	-0,04	0,98	0,02
Z33	0,81	0,24	0,49	0,12	-0,02	0,98	0,02
Z35	0,55	0,46	0,67	0,03	-0,04	0,98	0,02
Z36	0,82	-0,04	0,54	0,05	0,00	0,97	0,03
Z17	0,21	0,93	0,24	0,03	0,03	0,97	0,03
Z42	0,08	0,01	0,20	0,04	0,96	0,97	0,03
Z07	-0,16	0,92	0,31	0,07	-0,05	0,97	0,03
Z09	-0,17	0,93	0,26	0,12	-0,04	0,97	0,03
Z25	0,90	-0,31	0,24	0,02	0,05	0,96	0,04
Z15	-0,13	0,93	0,28	0,05	-0,01	0,96	0,04
Z37	0,78	0,00	0,58	0,08	-0,02	0,95	0,05
Z12	0,60	-0,76	-0,10	0,12	-0,02	0,95	0,05
Z03	-0,26	0,43	0,83	-0,01	0,06	0,95	0,05
Z20	-0,28	-0,58	-0,71	-0,16	-0,02	0,95	0,05
Z26	0,88	-0,25	0,32	-0,03	0,07	0,94	0,06
Z13	0,07	0,92	0,30	0,09	-0,01	0,94	0,06
Z22	0,89	0,23	0,25	0,10	0,09	0,93	0,07
Z39	0,66	-0,05	0,68	0,12	0,08	0,93	0,07
Z38	0,55	-0,02	0,77	0,11	0,11	0,93	0,07
Z24	0,17	0,42	0,85	0,01	0,06	0,93	0,07
Z27	0,75	0,19	0,55	-0,06	-0,07	0,92	0,08
Z32	0,55	-0,45	0,63	-0,07	0,13	0,92	0,08

(continua)

Sensor	MF	MHRB	MHBITA	MHRBF	PHI	Comunalidade	Variância Específica
Z21	0,94	0,05	0,10	0,03	0,10	0,91	0,09
Z31	0,53	-0,47	0,63	-0,02	0,12	0,91	0,09
Z40	0,41	0,15	0,85	0,01	0,07	0,91	0,09
Z41	-0,15	-0,30	-0,88	-0,13	-0,05	0,91	0,09
Z05	-0,16	0,44	0,83	0,06	0,08	0,91	0,09
Z11	-0,30	0,89	0,16	0,00	-0,04	0,90	0,10
Z14	0,70	-0,59	0,10	0,19	0,02	0,89	0,11
Z23	0,23	0,36	0,84	0,03	0,08	0,89	0,11
Z28	0,84	0,06	0,40	-0,08	-0,01	0,88	0,12
Z29	0,91	-0,14	-0,09	0,14	0,03	0,87	0,13
Z10	0,82	-0,10	-0,35	-0,24	0,08	0,87	0,13
Z02	-0,09	0,73	0,54	0,20	-0,03	0,87	0,13
Z04	0,36	0,81	0,26	0,04	0,03	0,86	0,14
Z06	0,28	0,82	0,24	0,17	0,01	0,84	0,16
Z18	0,54	-0,55	0,13	0,47	-0,03	0,83	0,17
Z19	-0,26	-0,17	-0,86	-0,07	0,03	0,83	0,17
Z30	0,89	-0,12	-0,07	0,05	-0,06	0,82	0,18
Z16	0,04	0,87	-0,18	-0,09	0,09	0,81	0,19
Z01	0,17	0,47	0,21	0,71	0,08	0,80	0,20
Z08	0,71	0,42	0,09	0,32	-0,04	0,79	0,21

FONTE: A autora (2016).

A variabilidade presente nas leituras de cada sensor, advinda dos fatores identificados, foi estimada através da comunalidade, pois à medida que cresce a comunalidade, menor será a variância específica. Utilizando esta informação como uma medida da importância de cada variável para o modelo fatorial, obteve-se o ranqueamento dos sensores. Assim, os instrumentos mais importantes para os blocos D8 e D7 foram, respectivamente, as hastes de extensômetro múltiplo Z33, Z34, Z35 e as bases de alongâmetro Z7, Z9 (abertura das juntas entre blocos). Além disso, o nível de água do reservatório (Z42) foi a variável relacionada com as condições ambientais melhor classificada.

O trio de hastes com maiores comunalidades possuem em comum a característica de estarem instalados no mesmo furo de sondagem a jusante, na fundação do bloco D8, a partir da elevação 191 m, muito próximo do local em que estão instaladas as duas bases de alongâmetro que medem a abertura das juntas entre os blocos D7 e D8 (Z7 e Z9). Outra característica que pode auxiliar na

explicação desta ocorrência é que no Trecho D a região do contato concreto/rocha está exposta, como é possível visualizar na Figura 24, e, portanto, sujeita à ação das condições do meio ambiente externo. Sendo assim, essas peculiaridades podem estar influenciando de modo mais acentuado esses cinco sensores. Essas informações trazem mais conhecimento das relações entre os diferentes tipos de sensores e o meio ambiente, podendo servir, futuramente, como ponto de partida para outras investigações.

FIGURA 24 – VISTA FRONTAL DO TRECHO D DA BARRAGEM DE ITAIPU



FONTE: A autora¹⁵ (2016).

O fator, por ser uma variável latente, não pode ser medido diretamente. Sendo assim, os valores dos fatores, denominados escores fatoriais, foram estimados com base nas cargas fatoriais e nos dados dos sensores. Os escores fatoriais forneceram informações para a classificação dos períodos de aquisição de dados e para a modelagem do Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem, conforme descrito nas seções a seguir.

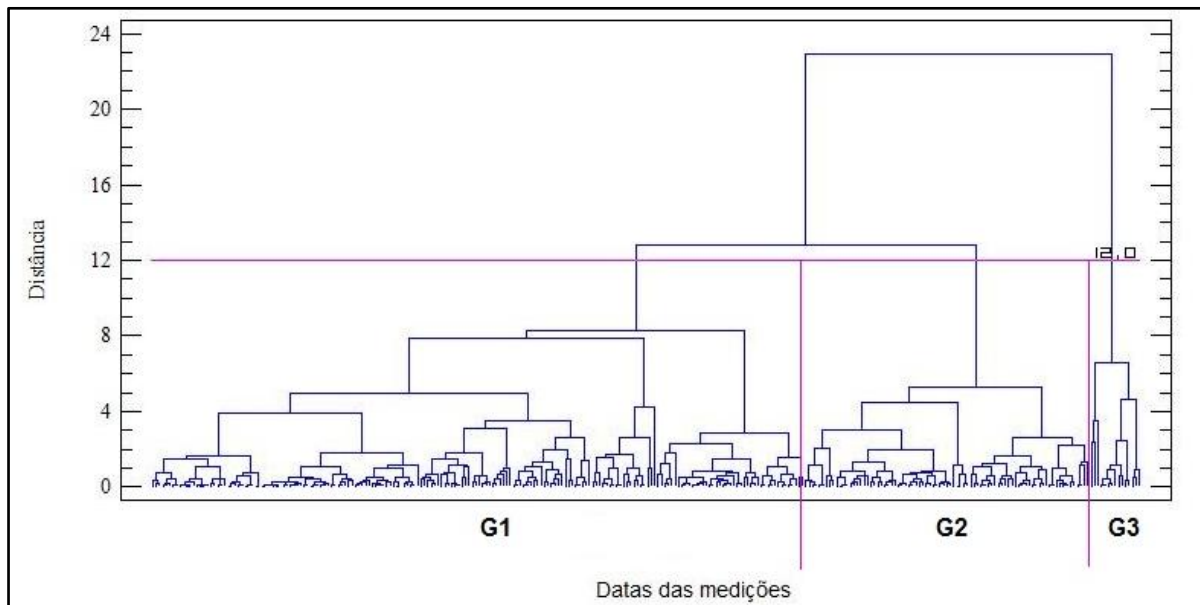
¹⁵ Foto tirada pela autora durante visita técnica realizada em 27/11/2015.

4.2.2 Agrupamentos das Datas das Medições de Acordo com as Similaridades

Obtidos os escores fatoriais, dos 5 fatores extraídos, para cada uma das 288 datas de realização das leituras foi feita a análise de agrupamentos (*cluster*) e a análise discriminante.

Utilizando o método da ligação média e a distância de Mahalanobis, foram identificados três grupos homogêneos de datas, conforme apresentado na Figura 25, agregando 190 elementos no primeiro (G1), 83 no segundo (G2) e 15 no terceiro (G3).

FIGURA 25 – AGRUPAMENTOS PELO MÉTODO DA LIGAÇÃO MÉDIA, UTILIZANDO A DISTÂNCIA DE MAHALANOBIS



FONTE: A autora (2016).

O primeiro grupo ficou constituído essencialmente dos meses de novembro a maio, período em que foram registradas as maiores temperaturas, enquanto que a maioria dos meses com menores temperaturas ficaram reunidas em G2. O terceiro grupo reuniu as datas em que foram registrados os menores níveis de água do reservatório. Cada grupo recebeu uma denominação, segundo as características responsáveis pelos agrupamentos, a saber: Temperatura Alta (G1), Temperatura Baixa (G2) e Baixo Nível de Água do Reservatório (G3).

A classificação das 288 observações foi testada por meio da análise discriminante, utilizando como variáveis discriminantes os sensores Z7, Z9, Z33, Z34, Z35 e Z42, escolhidos por apresentarem as maiores comunalidades.

Devido à grande diferença de tamanho dos grupos, considerou-se a proporcionalidade do número de observações por grupo. O resumo dos resultados é apresentado na Tabela 12. O elevado percentual de classificação correta, 94,1%, confirmou o poder discriminante dos sensores considerados na análise. A classificação de cada uma das 288 observações está disponível no Apêndice D.

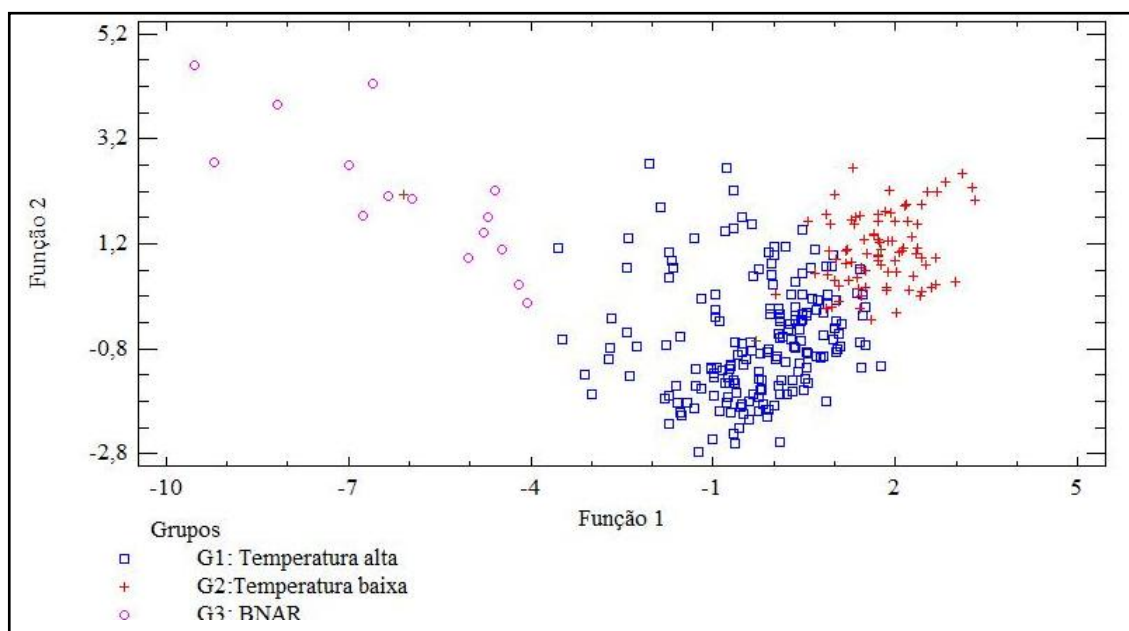
Duas funções discriminantes foram consideradas estatisticamente relevantes, ao nível de confiança de 99%, para fazer a distinção das observações pertencentes a cada grupo (Figura 26). A primeira função discrimina os grupos Temperatura Alta e Temperatura Baixa, enquanto que a segunda discrimina o grupo Baixo Nível de Água do Reservatório (BNAR) dos demais grupos.

TABELA 12 – CLASSIFICAÇÃO DAS DATAS DAS MEDIÇÕES EM TRÊS GRUPOS

Grupo Atual	Tamanho do Grupo	Proporção	Grupo Previsto			Percentual de Acerto
			1	2	3	
1	190	65,97%	178	11	1	93,68%
2	83	28,82%	5	78	0	93,98%
3	15	5,21%	0	0	15	100%
TOTAL	288	100%	183	89	16	94,10%

FONTE: A autora (2016).

FIGURA 26 – GRÁFICO DE DISPERSÃO DOS ELEMENTOS DE ACORDO COM AS FUNÇÕES DISCRIMINANTES ENTRE GRUPOS



FONTE: A autora (2016).

As funções de classificação das observações nos grupos (Equação 57) resultaram das combinações lineares dos sensores. Essas funções podem ser utilizadas para a classificação de novas datas de leituras. Para tanto, basta calcular os escores de cada novo elemento em cada grupo e, em seguida, alocá-lo naquele de maior escore.

$$\begin{aligned}
 G_1 &= -0,66 + 0,66Z_{34} - 1,33Z_{33} - 0,08Z_{35} + 0,32Z_9 + 0,13Z_{42} - 0,88Z_7 \\
 G_2 &= -3,40 - 0,45Z_{34} + 2,87Z_{33} + 0,07Z_{35} - 1,284Z_9 + 1,656Z_{42} + 2,674Z_7 \\
 G_3 &= -23,81 + 5,94Z_{34} + 1,01Z_{33} + 0,54Z_{35} + 3,01Z_9 - 10,87Z_{42} - 3,64Z_7
 \end{aligned} \quad (57)$$

4.3 ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS DA BARRAGEM

4.3.1 Formulação do Índice

Por considerar que MF, MHRB, MHBITA, MHRBF e PHI são variáveis quantitativas que estão relacionadas com a variabilidade das respostas da barragem, em relação à ação da temperatura ambiente e do nível de água do reservatório, esses fatores foram usados como parâmetros para a elaboração do Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem (IMCRB).

A Figura 27 apresenta as séries temporais das cinco variáveis (fatores) utilizadas na formulação do IMCRB. Observando os gráficos, percebeu-se que as séries apresentavam aspecto de não-estacionariedade na média. Após a realização de uma diferenciação, as séries tornaram-se estacionárias na média, conforme é possível observar nos gráficos da Figura 28. Optou-se, então, por elaborar dois índices: um com as séries originais dos fatores (IMCRB) e outro com as séries diferenciadas (IMCRBdif), e verificar ao final a adequação de ambos para a possível utilização no monitoramento da barragem.

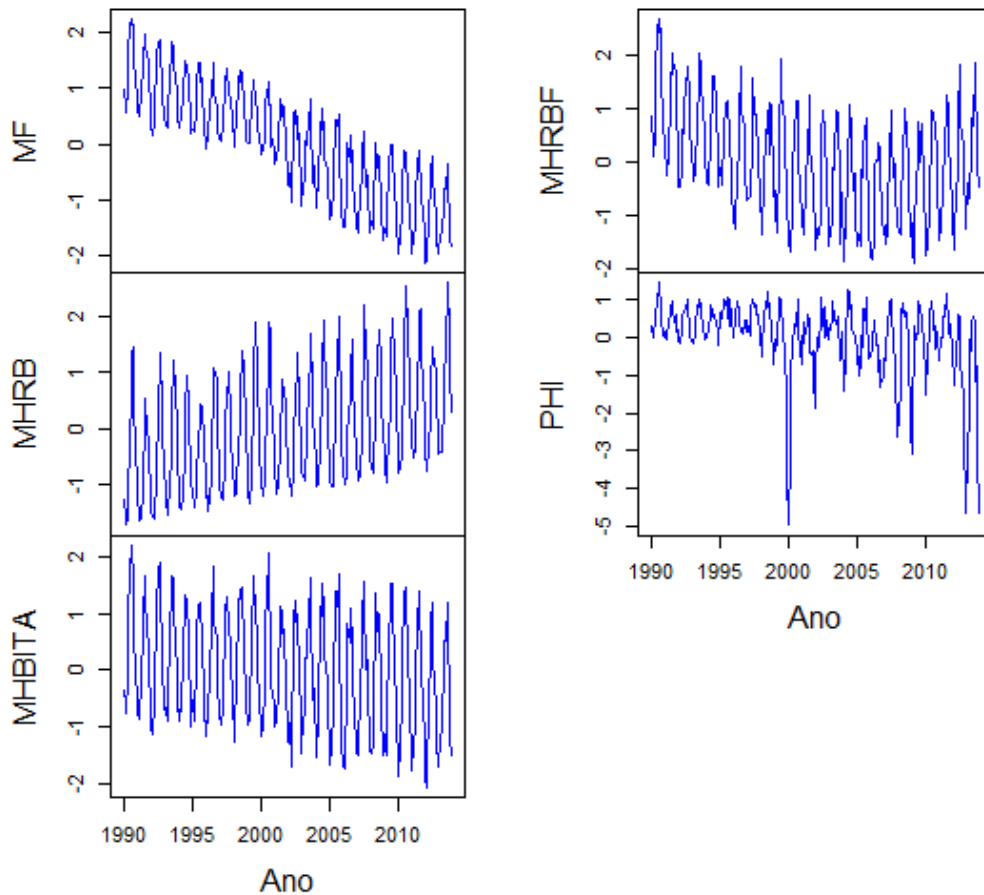
O Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem foi formulado (Equação 58) por meio da média ponderada dos escores

fatoriais, com pesos estipulados pelos autovalores da matriz de correlação amostral. De forma análoga (Equação 59), foi elaborado o índice composto pelas séries diferenciadas.

$$IMCRB \cong 0,46MF + 0,31MHRB + 0,10MHBITA + 0,02MHRBF + 0,02PHI \quad (58)$$

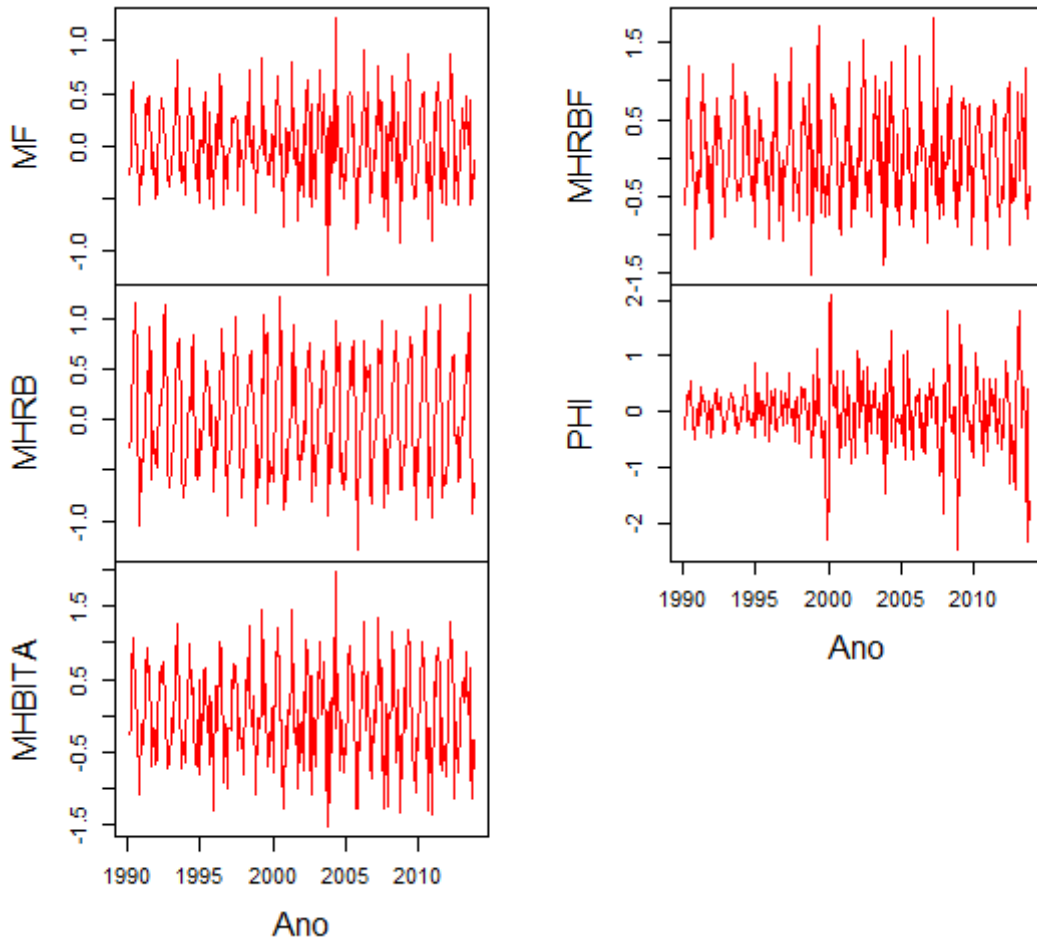
$$IMCRBdif \cong 0,46MFdif + 0,31MHRBdif + 0,10MHBITAdif + 0,02MHRBFdif + 0,02PHIdif \quad (59)$$

FIGURA 27 – SÉRIES TEMPORAIS DOS PRINCIPAIS FATORES QUE INFLUENCIARAM OS DESLOCAMENTOS DAS ESTRUTURAS DE CONCRETO DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU, NO PERÍODO DE JAN/90 A DEZ/13



FONTE: A autora (2016).

FIGURA 28 – SÉRIES DIFERENCIADAS DOS PRINCIPAIS FATORES QUE INFLUENCIARAM OS DESLOCAMENTOS DAS ESTRUTURAS DE CONCRETO DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU, NO PERÍODO DE JAN/90 A DEZ/13



FONTE: A autora (2016).

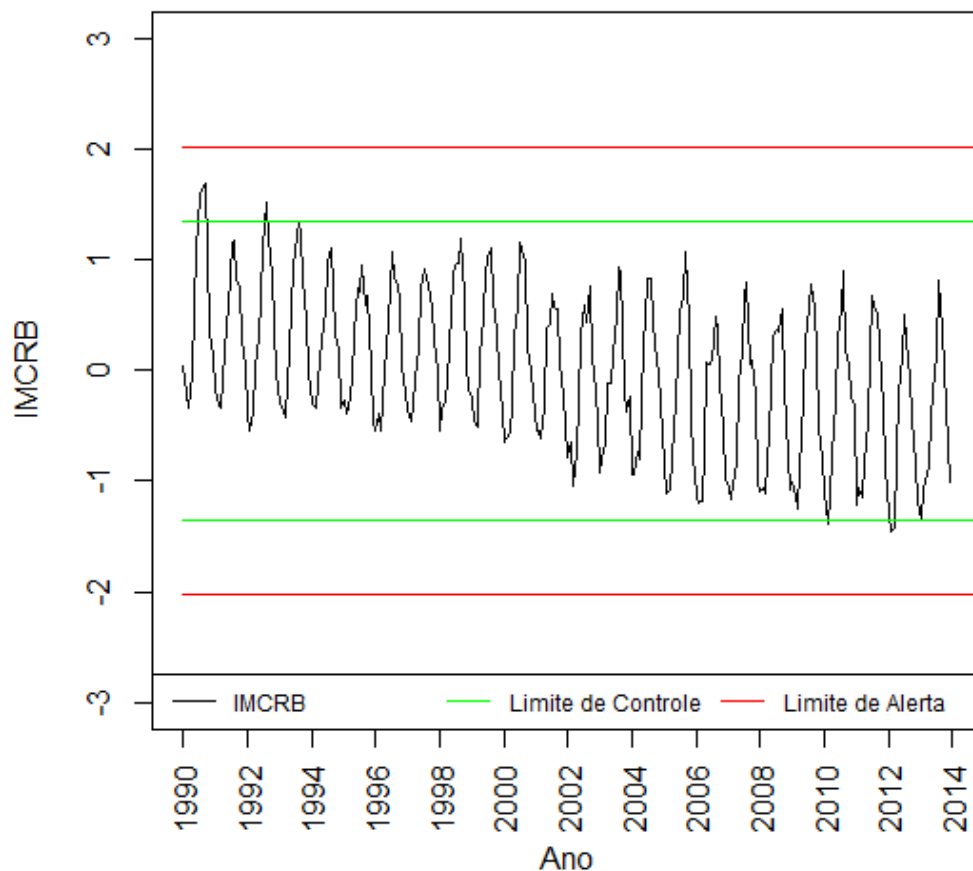
Todos os procedimentos descritos a seguir foram realizados para ambos os índices IMCRB e IMCRBdif. Por conveniência e para simplificação, foi citada apenas a forma de realização para IMCRB. No entanto, os resultados obtidos para ambos são apresentados e discutidos.

O IMCRB foi criado para concentrar as informações dos fatores que, por sua vez, representavam o conjunto de observações dos sensores utilizados neste estudo, com mínima perda de informação. Portanto, esperava-se que este índice apresentasse grande potencial de informação a respeito do comportamento dos blocos D7 e D8 da barragem e pudesse ser utilizado no seu monitoramento, como de fato ocorreu, em vista dos resultados apresentados na subseção a seguir.

4.3.2 Cartas de Controle dos Índices IMCRB e IMCRBdif

Para monitorar o índice criado, foram elaboradas as cartas de controle (Figura 29 e Figura 30) para as variáveis IMCRB e IMCRBdif, com limites de controle e de alerta estipulados com 95% e 99% de confiança, respectivamente.

FIGURA 29 – CARTA DE CONTROLE PARA O ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU

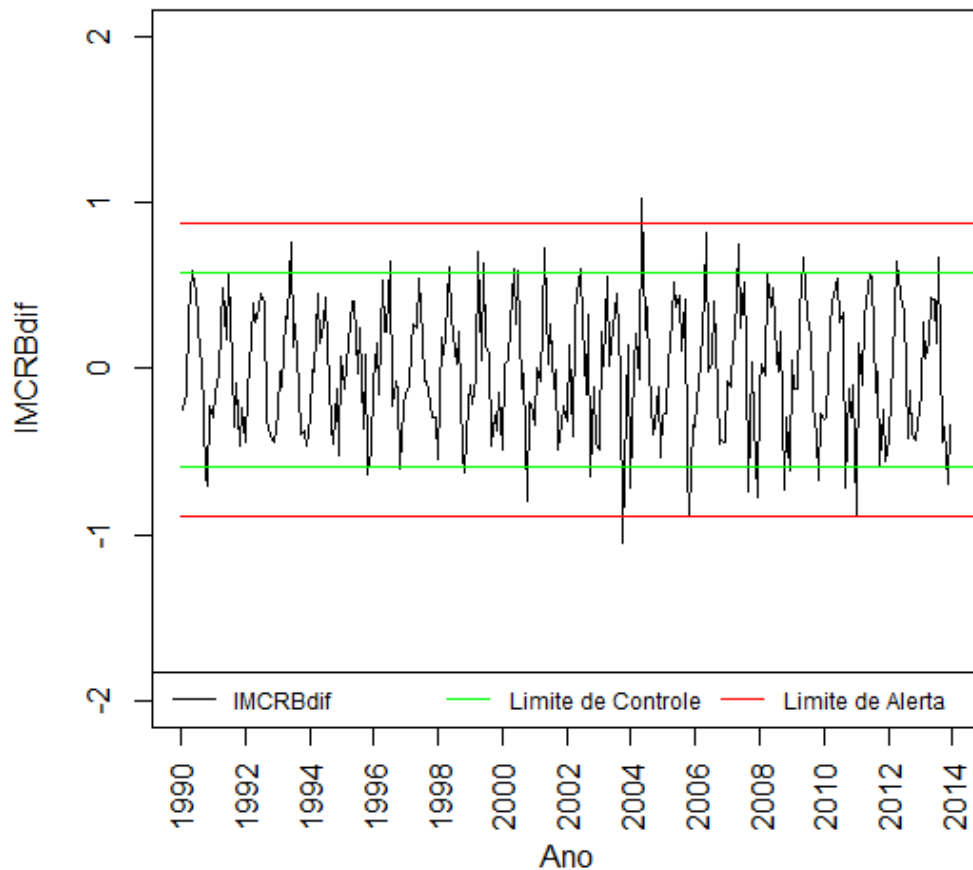


FONTE: A autora (2016).

Observando a carta de controle para o IMCRB, constatou-se que o processo monitorado apresentava-se quase que na sua totalidade dentro dos limites de controle e com tendência descendente. A provável justificativa para a presença dessa tendência é a forte influência que a Movimentação da Fundação (MF) exerce sobre o índice. Porém, outras fontes de variabilidade não consideradas neste estudo podem estar associadas a essa ocorrência, tais como, envelhecimento da estrutura, efeito da reação álcali-agregado, entre outras. Considerando que nos últimos 12

anos os valores do IMCRB estabilizaram, apresentando apenas oscilações sazonais, e conforme o sistema de classificação exposto no Quadro 5, classificou-se o processo como estável - Caso (a).

FIGURA 30 – CARTA DE CONTROLE PARA A SÉRIE DIFERENCIADA DO ÍNDICE DE MONITORAMENTO CONJUNTO DAS RESPOSTAS DOS BLOCOS D7 E D8 DA BARRAGEM DE ITAIPU



FONTE: A autora (2016).

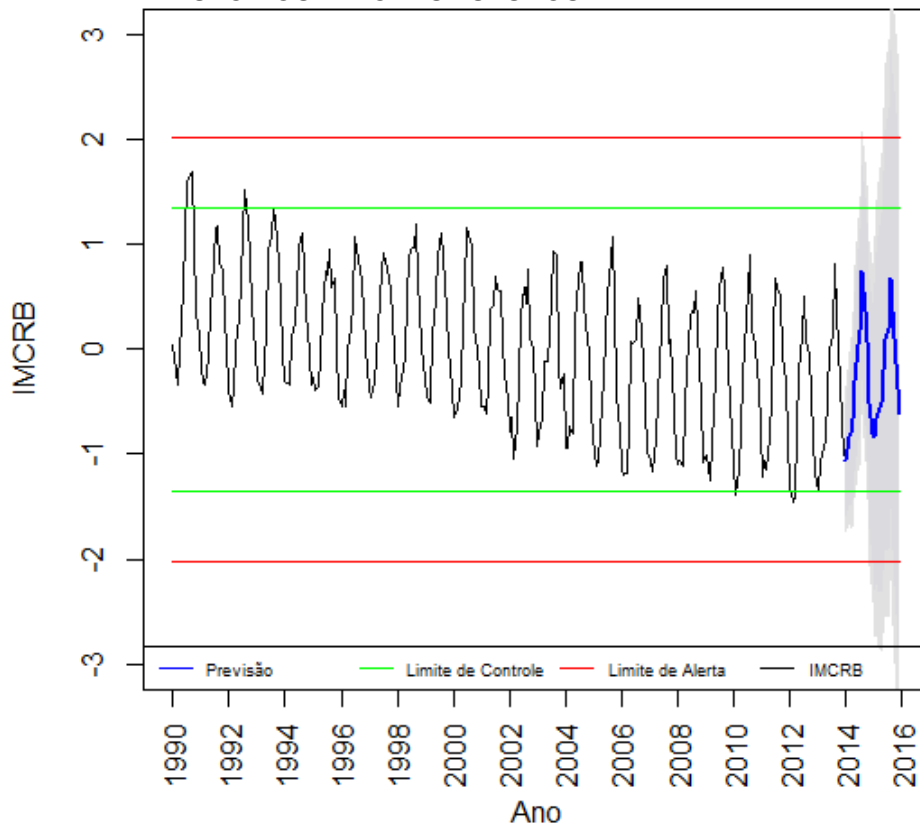
Para a série diferenciada IMCRBdif foram identificados dois pontos no ano de 2004 que excederam os limites de alerta, tratando-se de um Caso (c) - processo fora de controle. Ao comparar com os gráficos das séries diferenciadas dos fatores Movimentação da Fundação e Movimentação Horizontal dos Blocos D7 e D8 por Influência Térmica Ambiental se percebeu as mesmas ocorrências. Ao buscar informações junto à equipe técnica do CEASB, veio a informação de que em 2004 a instrumentação passou por um processo de automatização das leituras, o que poderia ter originado os valores extremos percebidos, ou seja, a provável causa para a anormalidade foi a manutenção dos instrumentos (c_2), conforme exposto na Seção 3.5 (páginas 84 e 85). Como após esse ano os valores voltaram a variar

aleatoriamente dentro do intervalo de 99% de confiança, considerou-se que o processo estava sob controle (estável) e, portanto, apresentava comportamento previsível.

Para realizar a previsão dos valores para os 24 meses subsequentes a dezembro de 2013 foi ajustado um modelo ARIMA, com o auxílio do pacote *forecast* do *software R*, utilizando a função "auto.arima", que compara diversos modelos possíveis e seleciona o de melhor ajuste segundo o critério de informação escolhido (neste caso, AIC). O Apêndice E traz um resumo das características dos modelos escolhidos, a saber: ARIMA (2,1,0)×(1,0,0)₁₂ para o IMCRB e ARIMA (2,0,0)×(1,0,0)₁₂ para o IMCRBdif.

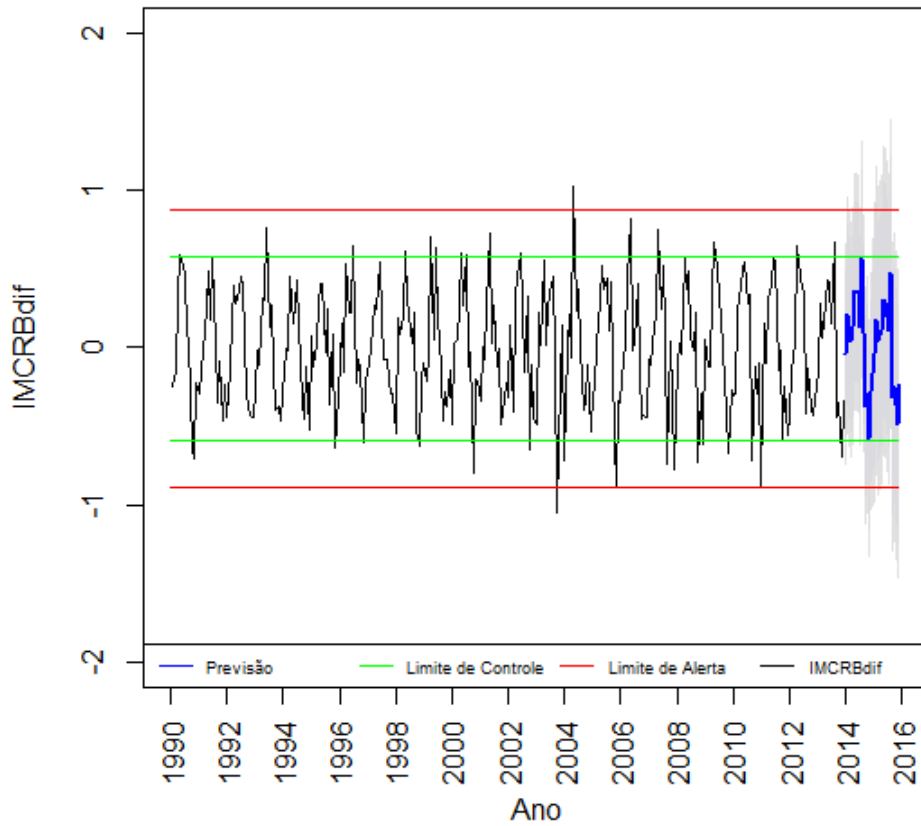
Por meio destes modelos, foi possível realizar a previsão para os 24 meses seguintes ao mês de dezembro de 2013. Os valores observados e os previstos foram plotados na mesma carta de controle (Figura 31 Figura 32). Destaca-se que em ambos os casos as previsões ficaram todas dentro dos intervalos de confiança preestabelecidos.

FIGURA 31 – CARTA DE CONTROLE PARA O ÍNDICE IMCRB CONTENDO A SÉRIE HISTÓRICA E A PREVISÃO DOS VALORES FUTUROS



FONTE: A autora (2016).

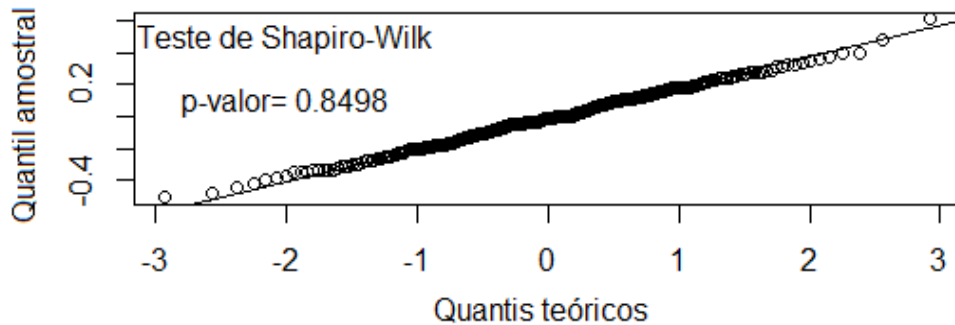
FIGURA 32 – CARTA DE CONTROLE PARA O ÍNDICE IMCRBDIF CONTENDO A SÉRIE HISTÓRICA E A PREVISÃO DOS VALORES FUTUROS



FONTE: A autora (2016).

Os resíduos dos modelos ajustados para as séries temporais do índices IMCRB e IMCRBdif foram verificados quanto à forma de distribuição, por meio do teste de Shapiro-Wilk e do gráfico Q-Q Plot (Figura 33). Em ambos os casos, os resultados indicaram que, com 95% de confiança, os resíduos eram normalmente distribuídos.

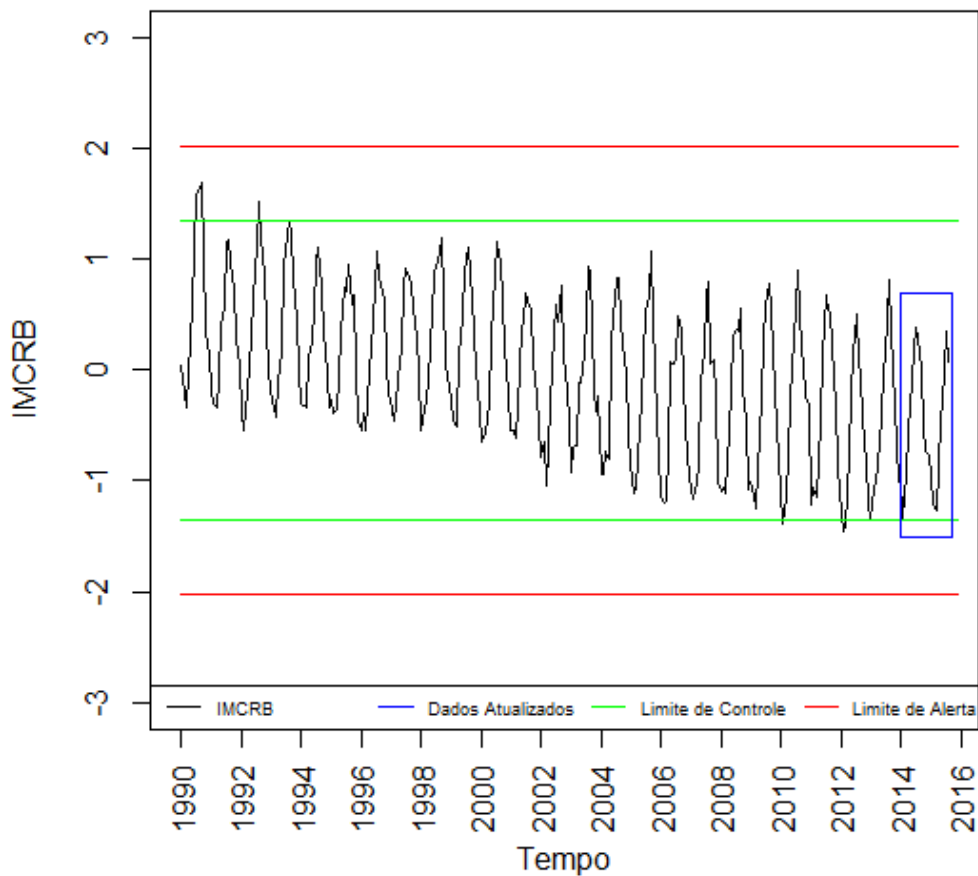
FIGURA 33 – VERIFICAÇÃO DA NORMALIDADE DOS RESÍDUOS DO MODELO ARIMA AJUSTADO À SÉRIE TEMPORAL DO IMCRB



FONTE: A AUTORA (2016).

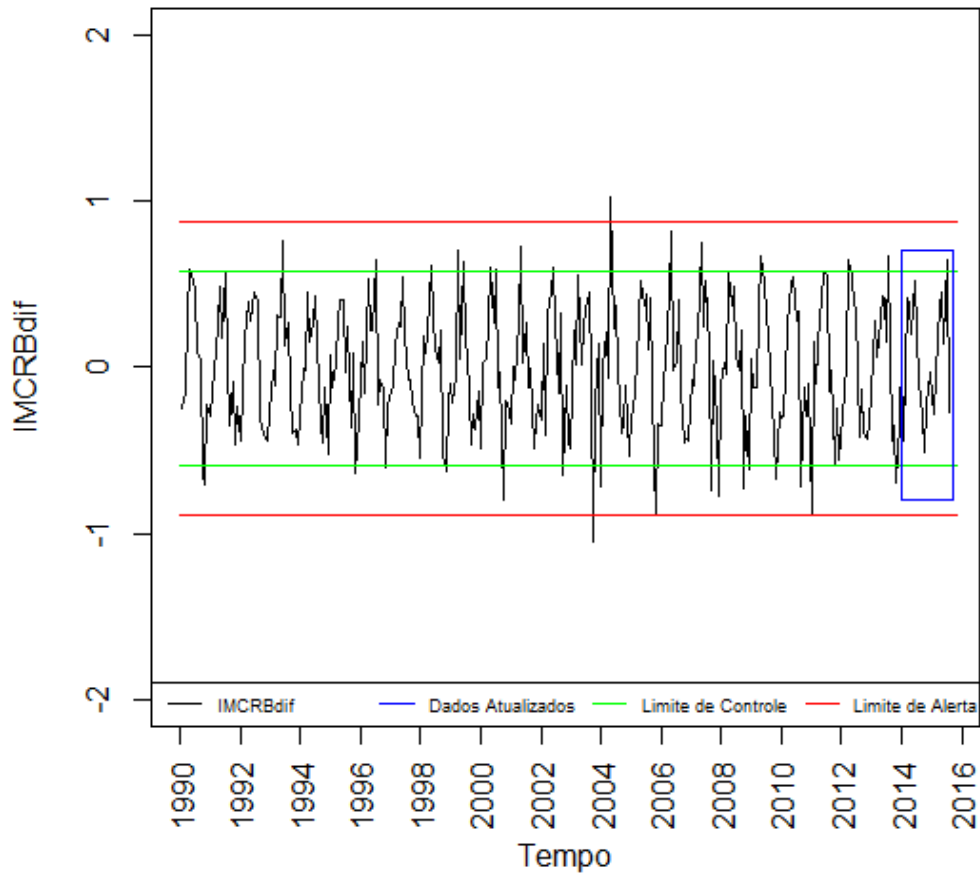
Para realizar a validação do modelo, foi utilizado o conjunto de dados referentes aos meses de janeiro de 2014 a agosto de 2015, doravante denominados dados atualizados. Esses dados já se encontravam completos e padronizados, totalizando 20 observações por sensor, visto que esta atividade foi realizada no início da aplicação do método. Utilizando os dados atualizados, foram calculados os novos valores dos escores fatoriais e, com esses, obtidos os valores atualizados para o IMCRB e IMCRBdif, os quais foram plotados nas respectivas cartas de controles (em destaque nas Figura 34 e Figura 35).

FIGURA 34 - CARTA DE CONTROLE ATUALIZADA PARA O ÍNDICE IMCRB



FONTE: A autora (2016).

FIGURA 35 - CARTA DE CONTROLE ATUALIZADA PARA O ÍNDICE IMCRBDIF



FONTE: A autora (2016).

Os valores atualizados tanto para IMCRB quanto para IMCRBdif mantiveram-se dentro dos limites de controle estabelecidos a priori e dentro do intervalo de confiança de previsão da série. Portanto, o modelo foi considerado adequado para a representação das respostas dos blocos D7 e D8 da barragem, no que diz respeito aos deslocamentos e quanto à incidência de oscilações na temperatura e no nível de água do reservatório.

CONCLUSÃO

O método proposto neste trabalho envolveu a aplicação de técnicas de análise multivariada aos dados da instrumentação utilizada no monitoramento dos deslocamentos das estruturas e fundações de uma barragem de concreto, levando em consideração a ação da temperatura e do nível de água do reservatório. O uso dessas técnicas permitiu alcançar os objetivos almejados.

Os resultados da análise de correlação canônica permitiram inferir que os deslocamentos das estruturas da barragem apresentam forte relação linear com as condições ambientais. De modo especial, verificou-se que a instrumentação registrou maiores deslocamentos nos períodos de baixas temperaturas.

A aplicação do método possibilitou verificar que o conjunto de instrumentos composto por pêndulos, bases de alongâmetro e extensômetro múltiplos de hastes pode ser usado para a previsão das respostas da barragem, no que diz respeito aos deslocamentos, segundo o critério da variabilidade nas condições ambientais.

A barragem apresentou respostas distintas dependendo da temperatura, alta ou baixa, e do baixo nível de água do reservatório, conforme resultados obtidos na análise de agrupamentos e na análise discriminante.

A maior parte da variabilidade dos dados das medições dos 42 sensores considerados neste estudo foi identificada, por meio da aplicação da análise fatorial, como advinda dos fatores: Movimentação da Fundação; Movimentação Horizontal Relativa entre os Blocos D7 e D8; Movimentação Horizontal dos Blocos D7 e D8 por Influência Térmica Ambiental; Movimentação Horizontal Relativa do Bloco D7 em relação à Fundação e Pressão Hidrostática.

O Índice de Monitoramento Conjunto das Respostas dos Blocos da barragem, elaborado com base nos escores fatoriais, mostrou-se adequado para a representação do comportamento global das estruturas dos blocos, sob a ação da temperatura ambiente e do nível de água do reservatório, por concentrar as informações, com mínima perda, e possibilitar a detecção de anormalidades.

A principal contribuição deste trabalho foi a proposição de um método para a análise conjunta dos dados dos deslocamentos das estruturas de uma barragem de concreto, das temperaturas e do nível de água do reservatório, que fornece

informações úteis para o monitoramento da barragem, permite a identificação de anormalidades nos dados e auxilia a tomada de decisão quanto ao estado do processo.

Para trabalhos futuros, sugere-se a aplicação do método com o conjunto de variáveis referentes às condições ambientais ampliado e/ou considerando outras variáveis condicionantes como, por exemplo, fluência das fundações e reação álcali-agregado. A investigação das respostas da barragem no que diz respeito à ocorrência de outros fenômenos, tais como deformações, tensões, pressões, entre outros, também é incentivada, uma vez que este trabalho considerou apenas os deslocamentos das estruturas da barragem. Adicionalmente, outros métodos estatísticos podem ser utilizados tanto de maneira exploratória, como feito neste estudo, quanto confirmatória. Por fim, sugere-se a realização de simulações para verificar a precisão do método aqui proposto quando as condições ambientais forem extremas, como por exemplo, nos eventos de nível máximo e/ou mínimo de água do reservatório.

REFERÊNCIAS

ANEEL. Agência Nacional de Energia Elétrica. **Nota Técnica nº 058/2015**. Regulamentação da Lei nº 12334 de 20 de setembro de 2010, Política Nacional de Segurança de Barragens, no âmbito do setor elétrico brasileiro. 25 jun. 2015.

BARBÃO, J. **Análise de componentes cíclicas em séries temporais uni e multivariadas via filtros HP modificados e outros métodos**. UNICAMP, Campinas, 2007.

BOX, E. L. K.; JENKINS, G. M. **Time series analysis: forecasting and control**. Holden-Day, 1976.

BRASIL. 12.334. **Lei de Segurança de Barragens**. 20 set. 2010, Sec. 1, p. 1.

CARVALHO, J. V.; ROMANEL, C. Redes neurais temporais aplicadas ao monitoramento de barragens. **Revista Eletrônica de Sistemas de Informação**, n. 10, 2007.

CASTRO, L. V. P. **Avaliação do comportamento do nível d'água em barragem de contenção de rejeito alteada a montante**. Escola Politécnica da Universidade de São Paulo, São Paulo, 2008.

CEASB. **Centro de Estudos Avançados em Segurança de Barragens**. Disponível em: <<http://www.pti.org.br/ceasb>>. Acesso em: 27 jan. 2016.

CEASB. Centro de Estudos Avançados em Segurança de Barragens. **Monitoramento da segurança estrutural da barragem de Itaipu**. Foz do Iguaçu, 2014. (Visita técnica-científica)

CHAVES NETO, A. **Técnicas estatísticas aplicadas à engenharia da qualidade**. Notas de Aula. Curitiba, 2009.

CHENG, L.; ZHENG, D. Two online dam safety monitoring models based on the process of extracting environmental effect. **Advances in Engineering Software**, v.57, p. 48–56, mar. 2013.

CHI. Central Hidrelétrica de Itaipu. **Barragem Lateral Direita: Estudos para a reavaliação dos valores de controle para a instrumentação civil das estruturas e fundações - Blocos D-54 e D-57**. Foz do Iguaçu: Itaipu, 1999a.

CHI. Central Hidrelétrica de Itaipu. **Modelo Estatístico: Estudos para a reavaliação dos valores de controle para a instrumentação civil das estruturas e fundações**. Foz do Iguaçu: Itaipu, 1999b.

CHI. Central Hidrelétrica de Itaipu. **Valores de controle para os instrumentos integrantes do “ADAS” e critérios para a geração de alarmes**. Foz do Iguaçu: Itaipu, 2002.

CNRH. **Resolução 144**. Estabelece diretrizes para implementação da Política Nacional de Segurança de Barragens, aplicação de seus instrumentos e atuação do Sistema Nacional de Informações sobre Segurança de Barragens. 10 jul. 2012a, Sec. 1.

CNRH. **Resolução 143**. Estabelece critérios gerais de classificação de barragens por categoria de risco, dano potencial associado e pelo volume do reservatório. 10 jul. 2012 b, Sec. 1.

COELHO, D. P.; PATIAS, J.; GARAY, V. R. Sistema de otimização e análise de auscultação da barragem de Itaipu. In: **XXX Seminário Nacional de Grandes Barragens**. Foz do Iguaçu, 2015.

CRUZ, P. T. **100 barragens brasileiras: casos históricos, materiais de construção, projeto**. São Paulo: Oficina dos Textos, 2006.

CURY, A.; CRÉMONA, C. Assignment of structural behaviours in long-term monitoring: Application to a strengthened railway bridge. **Structural Health Monitoring**, v. 11, n. 4, p. 422–441, 2010.

DENG, N.; WANG, J.-G.; SZOSTAK-CHRZANOWSKI, A. Dam Deformation Analysis Using the Partial Least Squares Method. In: **13th Fig. Int. Symp. on Deformation Measurements and Analysis & 4th Lag. Symp. on Geodesy for Geotechnical and Structural Engineering**. Lisbon, 2008.

DE SORTIS, A.; PAOLIANI, P. Statistical analysis and structural identification in concrete dam monitoring. **Engineering Structures**, v. 29, p. 110–120, 2007.

DTI. Diretoria técnica de Itaipu. **Usina Hidrelétrica de Itaipu: Aspectos de engenharia**. Foz do Iguaçu: Itaipu Binacional, 2009.

EKSTROM, T. A cracked concrete dam - analysis, measurements and upgrading action, Storfinnforsen dam. In: **23rd. International Congress on Large Dams**. Brasília, 2009.

ELETROBRÁS. **Critérios de projeto civil de usinas hidrelétricas**. CBDB, 2003.

ESPECIFICAÇÃO TÉCNICA. **Estudos para reavaliação dos valores limites para a instrumentação civil das estruturas da barragem de Itaipu e suas fundações**. Foz do Iguaçu: Itaipu Binacional, 1997.

FARRAR, C. R.; WORDEN, K. An introduction to structural health monitoring. **Philosophical Transactions of the Royal Society A**, v. 365, p. 303–315, 2007.

FIGUEIREDO, E.; PARK, G.; FARRAR, C. R.; WORDEN, K.; FIGUEIRAS, J. Machine learning algorithms for damage detection under operational and environmental variability. **Structural Health Monitoring**, v. 10, n. 6, p. 559 – 572, nov. 2011.

FIORINI, A. S. Interpretação de dados: análise dos resultados, interação entre áreas de Engenharia/Obra/Manutenção. In: **VI Simpósio Brasileiro sobre Pequenas e Médias Centrais Hidrelétricas**. Belo Horizonte, 2008.

HAIR, J. F.; BLACK, B.; BABIN, B.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. Porto Alegre: Bookman, 2009.

HSU, T. Y.; LOH, C. H. Damage detection accommodating nonlinear environmental effects by nonlinear principal component analysis. **Structural Control and Health Monitoring**, v. 17, n. 3, p. 338 – 354, abr. 2010.

ITAIPU BINACIONAL. **Itaipu Binacional: A maior geradora de energia limpa e renovável do planeta**. Disponível em: <<https://www.itaipu.gov.br/energia/auscultacao>>. Acesso em: 27 jan. 2016.

JAMES, B. R. **Probabilidade: um curso em nível intermediário**. 3 ed. Rio de Janeiro: IMPA, 2004.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 6 ed. Pearson, 2007.

KUPERMAN, S. C.; MORETTI, M. R.; CIFU, S.; CELESTINO, T. B.; RE, G.; ZOELLNER, K. **Criteria to establish limit values of instrumentation readings for old embankment and concrete dams**. 2005.

LAROS, J. A. **O uso da análise fatorial: algumas diretrizes para pesquisadores**. In: PASQUALI, L. (Ed.). **Análise fatorial para pesquisadores**. Brasília: LabPAM Saber e Tecnologia, 2012. p. 141–160.

LATTIN, J. M.; CARROL, J. D.; GREEN, P. E. **Análise de dados multivariados**. Cengage Learning, 2011.

LI, F.; WANG, Z. Z.; LIU, G. Towards an error correction model for dam monitoring data analysis based on cointegration theory. **Structural Safety**, v. 43, p. 12–20, 2013.

Matlab Starter Application. MathWorks, 2015.

MATOS, S. F. **Avaliação de instrumentos para auscultação de barragem de concreto. estudo de caso: deformímetros e tensômetros para concreto na barragem de Itaipu**. Dissertação, Universidade Federal do Paraná, Curitiba, 2002.

MEDEIROS, C. H.; LOPES, M. G. M. O Risco da Classificação de Barragens por Categoria de risco, com Base em Método de Ponderação de Fatores. In: **XXVIII Seminário Nacional de Grandes Barragens**. Rio de Janeiro, 2011.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Editora UFMG, 2005.

MONTGOMERY, D. C. **Introdução ao controle estatístico da qualidade**. 4. ed. Rio de Janeiro: LTC, 2009.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. Edgar Blücher, 2004.

MSIB. **Manual de Segurança e Inspeção de Barragens**. Brasília: Ministério da Integração Nacional, 2002.

NBR 8681. **Ações e segurança nas estruturas - Procedimento**. 2003.

NEUMANN JR, C. **Estrutura e funcionamento da instrumentação instalada no Trecho D da barragem de Itaipu**. Foz do Iguaçu, 2015. (Informação verbal)

ORO, S. R. **Aplicação de técnicas da engenharia da qualidade no aperfeiçoamento de um processo de produção na suinocultura**. Dissertação (Mestrado em Métodos Numéricos em Engenharia), Universidade Federal do Paraná, Curitiba, 2010.

ORO, S. R.; CHAVES NETO, A.; GARCIA, S. R. P.; MAFIOLETI, T. R. Modelagem de equações estruturais aplicada na avaliação dos fatores relacionados com a variabilidade dos dados de monitoramento de uma barragem de concreto. In: **Congresso de Métodos Numéricos em Engenharia**. Lisboa, 2015. .

ORO, S. R.; CHAVES NETO, A.; MAFIOLETI, T. R.; NEUMANN JR, C. Utilização de escores fatoriais na hierarquização dos períodos de aquisição de dados no monitoramento de uma barragem. In: **XXXV Iberian Latin American Congress on Computational Methods in Engineering**. Fortaleza, 2014. .

ORO, S. R.; CHAVES NETO, A.; NEUMANN JR, C. Fatores relacionados aos registros obtidos pela instrumentação de monitoramento de uma barragem. In: **IV Simpósio de Métodos Numéricos Computacionais da Universidade Federal do Paraná**. Curitiba, 2014. .

ORO, S. R.; CHAVES NETO, A.; NEUMANN JR, C. Joint monitoring of the displacements of the structures of a concrete dam. **International Journal of Engineering and Applied Sciences**, v. 7, n. 5, p. 1–9, dez. 2015.

POSENATO, D.; KRIPAKARAN, P.; INAUDI, D.; SMITH, I. F. . Methodologies for model-free data interpretation of civil engineering structures. **Computers and Structures**, v. 88, n. 7-8, p. 467–482, 2010.

R Core Team. 2015.

SANTOS, J. P.; CRÉMONA, C.; ORCESI, A. D.; SILVEIRA, P. Multivariate statistical analysis for early damage detection. **Engineering Structures**, v. 56, p. 273 – 285, nov. 2013.

SILVA, S.; DIAS JÚNIOR, M.; LOPES JÚNIOR, V. Detecção de falhas estruturais utilizando controle estatístico de processos. In: **4º Congresso Temático de Dinâmica, Controle e Aplicações**. Bauru, SP, 2005.

SILVEIRA, J. F. A. **Instrumentação e comportamento de fundações de barragens de concreto**. São Paulo: Oficina de Textos, 2003.

SOHN, H. Effects of environmental and operational variability on structural health monitoring. **Philosophical Transactions of the Royal Society A**, v. 365, p. 539–560, 2006.

SOM. Sistema de operação e manutenção das obras civis. **Barragem Lateral Direita e Vertedouro: Instrumentação - previsão e comportamento**. Aproveitamento Hidrelétrico de Itaipu. Foz do Iguaçu: Itaipu Binacional, 1984.

SOM. Sistema de operação e manutenção das obras civis. **Descrição de funcionamento: Termômetro para concreto**. Aproveitamento Hidrelétrico de Itaipu. Foz do Iguaçu: Itaipu Binacional, 1990a.

SOM. Sistema de operação e manutenção das obras civis. **Descrição de funcionamento: Bases para medidor de juntas (Alongâmetro)**. Aproveitamento Hidrelétrico de Itaipu. Foz do Iguaçu: Itaipu Binacional, 1990b.

Statgraphics Centurion XVI. StatPoint Technologies, 2010.

VASCONCELOS, G. R. L. **Plano de implantação e acompanhamento do instrumental de auscultação de estruturas de concreto - obra de Itaipu**. Itaipu Binacional, Divisão de Controle de Concreto, 1978.

VILLWOCK, R. **Técnicas de agrupamento e de hierarquização no contexto de KDD - aplicação a dados temporais de instrumentação geotécnica-estrutural da Usina Hidrelétrica de Itaipu**. Tese(Doutorado em Métodos Numéricos em Engenharia) -Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná, 2009.

VILLWOCK, R.; STEINER, M. T. A.; DYMINSKI, A. S.; CHAVES NETO, A. Itaipu Hydroelectric Power Plant Structural Geotechnical Instrumentation Temporal Data Under the Application of Multivariate Analysis - Grouping and Ranking Techniques. In: **Multivariate Analysis in Management, Engineering and the Sciences**. InTech, 2013. p. 81–102.

XU, C.; YUE, D.; DENG, C. Hybrid GA/SIMPLS as alternative regression model in dam deformation analysis. **Engineering Applications of Artificial Intelligence**, v. 25, p. 468–475, abr. 2012.

YU, H.; WU, Z. R.; BAO, T. F.; ZHANG, L. Multivariate analysis in dam monitoring data with PCA. **Science China, Technological Sciences**, v. 53, n. 4, p. 1088–1097, 2010.

APÊNDICES

APÊNDICE A – ALGORITMO DE GERAÇÃO DE MÉDIAS E IDENTIFICAÇÃO DE LACUNAS NO CONJUNTO DE DADOS

APÊNDICE B – MEDIDAS DESCRITIVAS DAS VARIÁVEIS

APÊNDICE C – TESTE DE BARTLETT E COEFICIENTE KMO NO *SOFTWARE R*

APÊNDICE D – CLASSIFICAÇÃO RESULTANTE DA ANÁLISE DISCRIMINANTE DAS DATAS DE MEDIÇÃO

APÊNDICE E – RESUMO DOS MODELOS ARIMA AJUSTADOS ÀS SÉRIES DO IMCRB E IMCRBdif

APÊNDICE A – ALGORITMO DE GERAÇÃO DE MÉDIAS E IDENTIFICAÇÃO DE LACUNAS NO CONJUNTO DE DADOS

```
function [matriz_medias] = gerar_medias(dados)
[linha,coluna]=size(dados);
%geração de uma matriz cuja primeira coluna é o instrumento, a segunda omês, a terceira o ano e a
%quarta o valor;
matriz(1:linha,1:4)=zeros;
matriz(1:linha,2:4)=dados;
j=1;
k=1;
ano_ini=matriz(1,3);
for i=k:linha
    if matriz(k,3)==ano_ini
        matriz(k,1)=j;
        k=k+1;
        while matriz(k,3)==ano_ini
            matriz(k,1)=j;
            k=k+1;
        end
        whilematriz(k,3)>ano_ini
            matriz(k,1)=j;
            k=k+1;
        end
    end
    j=j+1;
end

matriz_media=zeros(((matriz(linha,3)-matriz(1,3))+1)*12,(matriz(linha, 1)+2));
instrumento=1;
ano=matriz(1,3);
mes=1;
for i=1:((matriz(linha,3)-matriz(1,3))+1)*12
    if rem(i,12)~=0
        matriz_media(i,1)=ano;
        matriz_media(i,2)=mes;
        mes=mes+1;
    else
        matriz_media(i,1)=ano;
        matriz_media(i,2)=mes;
        mes=1;
        ano=ano+1;
    end
end

%preenchimento da terceira coluna em diante com as médias dos respectivos ano e mês do
%primeiro instrumento, e a quarta, do segundo instrumento, e assim sucessivamente para quantos
%instrumentos existirem.
mes=1;
n=0;
j=1;
soma=0;
```

```

instrumento=1;
diferenca=0;
for i=1:linha
    if matriz(i,2)==mes
        soma=soma+matriz(i,4);
        n=n+1;
    else
        matriz_media(j,instrumento+2)=soma/n;
        soma=0;
        n=0;
        diferenca=matriz(i,2)-mes;
        if diferenca>1
            for k=1:(diferenca-1)
                j=j+1;
                matriz_media(j,instrumento+2)=999999;
                mes=mes+1;
            end
            diferenca=0;
        end
        soma=matriz(i,4);
        n=1;
        if mes<12
            mes=mes+1;
        else
            mes=1;
        end
        if matriz (i,1)~=instrumento
            instrumento=instrumento+1;
            j=1;
        else
            j=j+1;
        end
    end
end
end

```

APÊNDICE B – MEDIDAS DESCRITIVAS DAS VARIÁVEIS

Variável	Unidade de Medida	Número de Observações	Média	Desvio padrão	Coefficiente de variação	Mínimo	Máximo	Amplitude	Assimetria	Curtose
X1	mm	288	0,60	0,16	27,15%	0,20	1,20	1,00	0,28	0,73
X2	mm	288	0,16	0,44	269,00%	-0,90	1,20	2,10	2,21	-2,08
X3	mm	288	2,43	1,02	42,14%	0,30	4,80	4,50	0,45	-2,22
X4	mm	288	-0,92	0,52	-56,56%	-1,60	0,60	2,20	4,66	-2,00
X5	mm	288	1,24	0,61	49,50%	0,00	2,70	2,70	0,45	-3,06
X6	mm	288	-0,60	0,43	-71,32%	-1,40	0,50	1,90	3,87	-2,13
X7	10 ⁻³ mm	288	106,97	198,95	185,98%	-340,50	606,00	946,50	1,57	-2,23
X8	10 ⁻³ mm	288	215,26	34,26	15,92%	137,00	315,00	178,00	3,92	0,19
X9	10 ⁻³ mm	288	24,10	182,65	758,00%	-290,00	509,00	799,00	2,60	-2,27
X10	10 ⁻³ mm	288	15,78	47,21	299,21%	-101,04	103,00	204,04	-2,18	-3,03
X11	10 ⁻³ mm	288	741,35	308,32	41,59%	155,00	1514,00	1359,00	0,79	-2,60
X12	10 ⁻³ mm	288	45,32	96,11	212,07%	-181,29	216,00	397,29	-1,43	-2,60
X13	10 ⁻³ mm	288	289,25	171,69	59,36%	29,00	795,00	766,00	4,62	-0,78
X14	10 ⁻³ mm	288	-401,04	45,85	-11,43%	-492,48	-281,00	211,48	1,82	-2,04
X15	10 ⁻³ mm	288	265,56	186,76	70,33%	-47,00	825,00	872,00	3,92	-0,93
X16	10 ⁻³ mm	288	-34,73	39,14	-112,72%	-117,00	91,00	208,00	5,55	1,21
X17	10 ⁻³ mm	288	470,65	540,56	114,85%	-345,00	1835,00	2180,00	3,83	-2,86
X18	10 ⁻³ mm	288	-620,93	52,09	-8,39%	-719,00	-429,00	290,00	8,24	4,27
X19	oC	288	27,92	6,32	22,62%	13,54	47,00	33,46	-0,01	-2,26
X20	oC	288	24,42	3,47	14,20%	17,28	30,89	13,62	-1,53	-4,50
X21	mm	288	-0,26	0,10	-40,44%	-0,54	-0,10	0,44	-4,26	-2,27
X22	mm	288	-0,18	0,10	-57,08%	-0,44	0,04	0,48	-2,64	-1,46
X23	mm	288	-0,09	0,14	-154,17%	-0,37	0,24	0,61	0,95	-3,15
X24	mm	288	-0,08	0,15	-199,63%	-0,36	0,26	0,62	1,31	-3,67
X25	mm	288	-0,65	0,20	-30,73%	-1,08	-0,27	0,82	-1,27	-3,10
X26	mm	288	-0,52	0,14	-26,70%	-0,85	-0,22	0,64	-2,20	-1,55
X27	mm	288	-0,72	0,12	-16,38%	-0,99	-0,44	0,55	-0,72	-1,99
X28	mm	288	-0,67	0,12	-18,68%	-1,01	-0,29	0,72	-1,96	-0,59
X29	mm	288	0,59	0,08	13,16%	0,45	0,73	0,28	-0,50	-4,40
X30	mm	288	0,69	0,09	12,56%	0,54	0,85	0,31	0,52	-4,85
X31	mm	288	-0,88	0,10	-11,32%	-1,14	-0,65	0,49	-0,95	-1,36
X32	mm	288	-0,81	0,09	-11,26%	-1,06	-0,60	0,47	-1,44	-0,59
X33	mm	288	-1,37	0,19	-13,82%	-1,79	-0,87	0,92	0,73	-1,36
X34	mm	288	-1,23	0,18	-15,00%	-1,64	-0,79	0,85	0,07	-1,73
X35	mm	288	-0,86	0,17	-19,93%	-1,23	-0,45	0,78	0,59	-3,16
X36	mm	288	-0,55	0,11	-20,18%	-0,80	-0,33	0,48	-1,85	-2,44
X37	mm	288	-0,44	0,12	-27,78%	-0,69	-0,18	0,52	-0,20	-2,35
X38	mm	288	0,34	0,15	45,20%	0,02	0,69	0,67	0,16	-2,72
X39	mm	288	0,37	0,14	38,38%	0,09	0,70	0,61	0,54	-2,65
X40	mm	288	0,30	0,14	48,57%	0,00	0,60	0,60	0,23	-3,10
X41	oC	288	22,26	3,64	16,37%	13,80	28,10	14,30	-2,28	-3,98
X42	m	288	219,66	0,64	0,29%	215,69	220,35	4,66	-20,90	41,32

APÊNDICE C –TESTE DE BARTLETT E COEFICIENTE KMO NO SOFTWARE R

```
# Carregar os dados
dados<-read.table("C:/Users/Sheila/Dropbox/projeto_tese/TESE/sensores.csv",header=TRUE,
dec=",")
n <- nrow(dados)          # total de observações
p <- ncol(dados)          # total de variáveis
S <- cov(dados);         # matriz de covariâncias
R <- cor(dados);         # matriz de correlações
print(R, digits = 2)     # visualização dos valores das correlações
library(corrgram)        # pacote "corrgram" para visualização gráfica das correlações
corrgram(R, type = "cor", lower.panel = panel.shade, upper.panel = panel.pie)

# Teste de esfericidade de Bartlett
#H0: R = I      H1: R ≠ I
# Se Tcalc>= T_95 rejeita-se H0 ao nível de significância de 5%.
# Verifica-se que as variáveis respostas são correlacionadas(não são independentes).
auto<- eigen(R);         # Cálculo dos autovalores da matriz de correlações
a<- log(auto$values);    # $values são os autovalores, $vectors são os autove
UM <- cbind(rep(1,p))
lambda = a%%UM; lambda
Tcalc = -(n-(1/6)*(2*p+11))*(lambda); Tcalc
gl=(1/2)*p*(p-1); gl
T_95 <- qchisq(0.95,gl); T_95
T_99 <- qchisq(0.99,gl); T_99

# Medidas de adequação amostral KMO e MAA
# correlações parciais
partial.cor<- function (X, ...)
{
  R <- cor(X, ...)
  RI <- solve(R)
  D <- 1/sqrt(diag(RI))
  Rp<- -RI * (D %% D)
  diag(Rp) <- 0
  rownames(Rp) <- colnames(Rp) <- colnames(X)
  Rp
}

matcorp<- partial.cor(dados) # matriz das correlações parciais

# coeficiente KMO quanto mais próximo de 1 maior a adequação
# se KMO < 0,5 é inadequado
idiag<- seq(1, by = p + 1, length = p)
somar2<- sum((as.numeric(R)[-idiag])^2)
cat("\n KMO = ",somar2 / (somar2 + sum((as.numeric(matcorp)[-idiag])^2)))

# índice de adequação amostral por variável. Valores aceitáveis > 0,5.
for (j in 1:p) {
  somar2j<- sum(R[j, -j]^2)
  cat("\n MAA", j, "=", somar2j / (somar2j + sum(matcorp[j, -j]^2)))
}
```

APÊNDICE D – CLASSIFICAÇÃO RESULTANTE DA ANÁLISE DISCRIMINANTE DAS DATAS DE MEDIÇÃO

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
jan/90	1	1	-0,103197	0,683207	0,9761	2	-3,81243	6,44532	0,0239
fev/90	1	1	0,415844	0,704612	0,9986	2	-6,18415	12,2482	0,0014
mar/90	1	1	0,72977	1,62841	0,9994	2	-6,75336	14,9383	0,0006
abr/90	1	1	0,281375	0,559268	0,9961	2	-5,25383	9,97332	0,0039
mai/90	2	*1	-0,899156	1,72467	0,729	2	-1,88846	2,04692	0,271
jun/90	2	2	0,679431	0,0198142	0,9209	1	-1,77548	6,586	0,0791
jul/90	2	2	4,66892	3,28131	0,9995	1	-2,93032	20,1362	0,0005
ago/90	2	2	4,83916	3,64613	0,9996	1	-3,07609	21,133	0,0004
set/90	2	2	4,8406	3,89241	0,9997	1	-3,19674	21,6235	0,0003
out/90	2	2	2,04956	0,303288	0,9856	1	-2,17937	10,4175	0,0144
nov/90	1	1	-1,07931	2,169	0,6194	2	-1,56612	1,48624	0,3806
dez/90	1	1	-0,658279	0,638628	0,9086	2	-2,95525	3,5762	0,0914
jan/91	1	1	-0,0313785	0,222587	0,988	2	-4,44273	7,38892	0,012
fev/91	1	1	0,253164	0,392509	0,9978	2	-5,85286	10,9482	0,0022
mar/91	1	1	0,403875	0,865714	0,999	2	-6,53231	13,0817	0,001
abr/91	1	1	0,0367295	0,097227	0,9949	2	-5,23888	8,99208	0,0051
mai/91	2	*1	-0,826226	0,877681	0,9044	2	-3,07283	3,71452	0,0956

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
jun/91	2	*1	-1,0478	2,43449	0,5808	2	-1,37383	1,43018	0,4192
jul/91	2	2	2,26665	0,488096	0,9902	1	-2,3527	11,3832	0,0098
ago/91	2	2	1,54539	0,452152	0,9788	1	-2,28899	9,77729	0,0212
set/91	2	2	-0,616438	0,603463	0,7166	1	-1,54429	4,11554	0,2834
out/91	2	2	-0,54569	0,669923	0,6799	1	-1,2989	3,83271	0,3201
nov/91	1	1	-0,758545	0,802015	0,8852	2	-2,80117	3,23091	0,1148
dez/91	1	1	-0,305056	0,152169	0,9728	2	-3,88093	5,64755	0,0272
jan/92	1	1	0,312665	0,500129	0,9982	2	-5,99606	11,4612	0,0018
fev/92	1	1	0,685736	1,77009	0,9997	2	-7,32938	16,144	0,0003
mar/92	1	1	0,353945	0,628623	0,9986	2	-6,20947	12,0991	0,0014
abr/92	1	1	-0,121044	0,068222	0,9867	2	-4,42993	7,02962	0,0133
mai/92	1	1	-0,615203	0,58253	0,9175	2	-3,02434	3,74443	0,0825
jun/92	2	2	-1,15575	0,994884	0,5411	1	-1,32071	2,98118	0,4589
jul/92	2	2	1,92124	0,364711	0,9855	1	-2,2992	10,462	0,0145
ago/92	2	2	3,74923	2,09789	0,9987	1	-2,85971	16,9721	0,0013
set/92	2	2	2,57157	0,913654	0,9941	1	-2,54716	12,8075	0,0059
out/92	2	2	1,00041	0,0940861	0,9535	1	-2,01955	7,79038	0,0465
nov/92	1	*2	-0,869433	0,760145	0,6383	1	-1,43721	3,55207	0,3617
dez/92	1	1	-0,634507	0,440859	0,9472	2	-3,52102	4,55752	0,0528
jan/93	1	1	-0,0485631	0,11555	0,9945	2	-5,24252	8,8471	0,0055
fev/93	1	1	-0,0672334	0,0533915	0,9933	2	-5,06537	8,3933	0,0067

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
mar/93	1	1	0,350595	0,685517	0,9987	2	-6,31416	12,3587	0,0013
abr/93	1	1	-0,115552	0,0121646	0,9911	2	-4,83104	7,78678	0,0089
mai/93	1	1	-0,349427	0,14876	0,9702	2	-3,83183	5,45721	0,0298
jun/93	2	2	0,726479	0,0111467	0,926	1	-1,80001	6,72049	0,074
jul/93	2	2	1,40062	0,060857	0,9689	1	-2,03984	8,59813	0,0311
ago/93	2	2	3,03884	1,20755	0,9966	1	-2,62793	14,1974	0,0034
set/93	2	2	2,49213	0,856859	0,9934	1	-2,52598	12,5495	0,0066
out/93	2	2	0,960047	0,542971	0,9585	1	-2,17884	8,4771	0,0415
nov/93	1*2		-0,619738	0,651917	0,7221	1	-1,57465	4,21812	0,2779
dez/93	1	1	-0,611991	0,421259	0,939	2	-3,34508	4,23107	0,061
jan/94	1	1	-0,0807774	0,206545	0,9857	2	-4,31363	7,01589	0,0143
fev/94	1	1	-0,0208454	0,0714772	0,9915	2	-4,77463	7,92267	0,0085
mar/94	1	1	0,26086	0,400166	0,9978	2	-5,85352	10,9726	0,0022
abr/94	1	1	-0,550326	0,324569	0,949	2	-3,47405	4,51566	0,051
mai/94	1	1	-0,715915	0,915999	0,8707	2	-2,62275	3,0733	0,1293
jun/94	2	2	-1,05814	0,949665	0,5435	1	-1,23246	2,95467	0,4565
jul/94	2	2	1,00732	0,102815	0,9541	1	-2,02685	7,82753	0,0459
ago/94	2	2	1,93668	0,5256	0,9867	1	-2,37185	10,799	0,0133
set/94	2	2	-0,0611185	0,305909	0,8398	1	-1,71792	5,27589	0,1602
out/94	1	1	-0,826267	1,31001	0,8033	2	-2,23343	2,46797	0,1967
nov/94	1	1	-0,639438	0,723346	0,901	2	-2,84757	3,48324	0,099

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
dez/94	1	1	-0,0251502	0,0415841	0,9927	2	-4,93618	8,20728	0,0073
jan/95	1	1	0,236469	2,64197	0,9831	2	-3,82983	9,11819	0,0169
fev/95	1	1	0,61521	1,774	0,9981	2	-5,67101	12,6901	0,0019
mar/95	1	1	0,416773	0,885363	0,9974	2	-5,52371	11,11	0,0026
abr/95	1	1	0,011246	0,82122	0,9828	2	-4,0359	7,25914	0,0172
mai/95	1	1	-0,64339	1,84589	0,8075	2	-2,07739	3,05752	0,1925
jun/95	2	2	0,428108	0,0580492	0,8996	1	-1,76473	6,1001	0,1004
jul/95	2	2	0,881574	0,815204	0,9162	1	-1,51063	7,25599	0,0838
ago/95	2	2	1,40868	0,317064	0,974	1	-2,21532	9,22144	0,026
set/95	2	2	-0,627722	0,734714	0,728	1	-1,61233	4,36029	0,272
out/95	2	2	1,21519	1,0487	0,9414	1	-1,56079	8,25703	0,0586
nov/95	1	1	-0,254569	1,06476	0,9504	2	-3,20665	5,31255	0,0496
dez/95	1	1	0,140967	0,930862	0,9891	2	-4,36499	8,2864	0,0109
jan/96	1	1	0,574561	1,4667	0,9982	2	-5,73123	12,4219	0,0018
fev/96	1	1	0,193351	0,740067	0,9925	2	-4,68919	8,84878	0,0075
mar/96	1	1	0,540994	1,63033	0,9975	2	-5,42858	11,9131	0,0025
abr/96	1	1	0,00636201	0,809003	0,9826	2	-4,02982	7,22499	0,0174
mai/96	1	1	-0,711115	0,837413	0,8807	2	-2,70977	3,17835	0,1193
jun/96	1	*2	-1,05136	0,935414	0,5874	1	-1,40471	3,2985	0,4126
jul/96	2	2	1,77937	2,70579	0,9893	1	-2,7458	13,4125	0,0107
ago/96	2	2	0,0422716	1,27034	0,8895	1	-2,0432	7,09765	0,1105

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
set/96	2	2	-0,0616085	0,97686	0,8679	1	-1,94456	6,39912	0,1321
out/96	2	2	0,0242867	0,21663	0,8287	1	-1,5521	5,02578	0,1713
nov/96	1	1	-0,64346	2,89134	0,7292	2	-1,63385	3,21575	0,2708
dez/96	1	1	-0,267032	1,23593	0,9437	2	-3,08526	5,21601	0,0563
jan/97	1	1	0,102286	1,41341	0,9831	2	-3,96143	7,88448	0,0169
fev/97	1	1	0,377578	1,2251	0,9955	2	-5,02314	10,3702	0,0045
mar/97	1	1	0,435548	0,879293	0,9991	2	-6,52379	13,1416	0,0009
abr/97	1	1	0,0073654	0,810622	0,9827	2	-4,03166	7,2323	0,0173
mai/97	1	1	-0,404023	1,76118	0,8944	2	-2,53999	4,37674	0,1056
jun/97	2	2	0,285301	0,527024	0,8486	1	-1,43846	5,63091	0,1514
jul/97	2	2	0,578195	0,0865238	0,9057	1	-1,68383	6,26695	0,0943
ago/97	2	2	1,58574	0,128752	0,9732	1	-2,00517	8,96695	0,0268
set/97	2	2	0,923613	0,227638	0,9317	1	-1,69026	7,11174	0,0683
out/97	2	2	-0,406471	0,407815	0,7555	1	-1,53438	4,31999	0,2445
nov/97	1	1	-0,783495	2,08904	0,7279	2	-1,76767	2,40101	0,2721
dez/97	1	1	-0,450902	0,325302	0,9524	2	-3,44709	4,6613	0,0476
jan/98	1	1	0,382844	0,827429	0,999	2	-6,49281	12,9224	0,001
fev/98	1	1	0,500871	0,920699	0,999	2	-6,39685	13,0598	0,001
mar/98	1	1	0,456315	0,78766	0,9986	2	-6,09847	12,2409	0,0014
abr/98	1	1	0,33864	1,30391	0,9943	2	-4,8161	9,95702	0,0057
mai/98	1	1	-0,946533	1,83979	0,7	2	-1,79407	1,87849	0,3

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
jun/98	2	2	1,09783	1,1462	0,9312	1	-1,50679	8,01181	0,0688
jul/98	2	2	1,89948	0,626974	0,9782	1	-1,90346	9,88923	0,0218
ago/98	2	2	2,00038	0,348352	0,9833	1	-2,07513	10,1557	0,0167
set/98	2	2	2,47121	0,580792	0,9921	1	-2,36079	11,9012	0,0079
out/98	2	2	-0,0585437	0,230587	0,8174	1	-1,55727	4,88441	0,1826
nov/98	1	1	-0,513176	0,282377	0,9532	2	-3,52693	4,65351	0,0468
dez/98	1	1	-0,267631	0,656826	0,9935	2	-5,29258	9,05037	0,0065
jan/99	1	1	0,195776	0,355613	0,9953	2	-5,15164	9,39409	0,0047
fev/99	1	1	0,431725	0,726944	0,9986	2	-6,1091	12,1522	0,0014
mar/99	1	1	0,69301	1,69036	0,9996	2	-7,2237	15,8674	0,0004
abr/99	1	1	-0,263432	0,782295	0,9567	2	-3,35987	5,3188	0,0433
mai/99	1	1	-0,348689	3,25834	0,851	2	-2,09103	5,08666	0,149
jun/99	2	2	1,78699	1,22009	0,9708	1	-1,71873	9,88789	0,0292
jul/99	2	2	2,33257	0,812439	0,9876	1	-2,0437	11,2214	0,0124
ago/99	2	2	2,34057	1,28591	0,9931	1	-2,6314	12,8862	0,0069
set/99	1*2		-0,979713	1,85481	0,682	1	-1,74265	5,03705	0,318
out/99	1	1	-1,18069	2,20434	0,7994	2	-2,56306	3,31271	0,2006
nov/99	1	1	-0,741493	4,29049	0,9908	2	-5,42089	11,9929	0,0092
dez/99	3	3	25,4998	4,77457	1	1	-1,45645	63,765	0
jan/00	3	3	44,1234	17,9225	1	1	-0,815107	112,877	0
fev/00	3	3	19,3327	0,0602347	1	1	-0,156173	44,116	0

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
mar/00	1	1	0,359215	4,55871	0,9998	2	-8,21971	20,0602	0,0002
abr/00	1	1	0,092371	0,161103	0,996	2	-5,42531	9,54009	0,004
mai/00	1	1	-1,03872	1,6192	0,8571	2	-2,8303	3,546	0,1429
jun/00	1	*2	-0,993555	1,1874	0,6406	1	-1,57152	3,99969	0,3594
jul/00	2	2	4,17901	2,78634	0,9992	1	-3,01078	18,8223	0,0008
ago/00	2	2	1,83566	0,31808	0,9788	1	-1,99871	9,64319	0,0212
set/00	2	2	2,42139	1,83821	0,9862	1	-1,84855	12,0345	0,0138
out/00	1	1	-0,637855	3,24737	0,7064	2	-1,5158	3,34689	0,2936
nov/00	1	1	-0,371676	0,54139	0,9504	2	-3,32491	4,79149	0,0496
dez/00	1	1	-0,226963	0,504437	0,9937	2	-5,28417	8,96249	0,0063
jan/01	1	1	0,399094	0,874858	0,9971	2	-5,44176	10,9002	0,0029
fev/01	1	1	0,959187	3,92165	0,9991	2	-6,00798	16,1996	0,0009
mar/01	1	1	1,08455	3,48516	0,9999	2	-8,3244	20,6467	0,0001
abr/01	1	1	0,56793	1,10646	0,9989	2	-6,27208	13,1301	0,0011
mai/01	1	1	-0,897466	1,10175	0,8498	2	-2,63075	2,91196	0,1502
jun/01	1	1	-0,947918	1,49394	0,7605	2	-2,10313	2,148	0,2395
jul/01	2	2	-0,702944	0,650692	0,6531	1	-1,33563	3,57242	0,3469
ago/01	1	1	-1,56191	4,88911	0,687	2	-2,34806	4,80505	0,313
set/01	1	1	-1,59888	5,51488	0,7072	2	-2,48091	5,62257	0,2928
out/01	1	1	-0,670136	0,515114	0,932	2	-3,28748	4,09343	0,068
nov/01	1	1	-0,509489	3,45293	0,9953	2	-5,85895	12,4955	0,0047

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
dez/01	1	*3	0,203781	7,51961	0,6321	1	-0,337926	13,681	0,3678
jan/02	1	1	0,472283	1,29873	0,9994	2	-6,97024	14,5274	0,0006
fev/02	1	1	0,623705	1,73644	0,9983	2	-5,74464	12,8168	0,0017
mar/02	1	1	1,46128	5,35691	1	2	-8,58046	23,784	0
abr/02	1	1	0,95054	2,67432	0,9998	2	-7,78983	18,4987	0,0002
mai/02	1	1	0,112175	0,158601	0,9955	2	-5,283	9,29258	0,0045
jun/02	1	1	-0,50498	1,90132	0,8554	2	-2,28281	3,80062	0,1446
jul/02	2	2	-0,314829	0,57203	0,7367	1	-1,34366	4,28606	0,2633
ago/02	1	1	-1,2047	2,19691	0,6624	2	-1,87874	1,88861	0,3376
set/02	2	2	0,602116	0,0455223	0,9113	1	-1,7274	6,36093	0,0887
out/02	1	1	-0,617172	0,684958	0,9073	2	-2,89869	3,59162	0,0927
nov/02	1	1	-0,544144	1,58602	0,8606	2	-2,36459	3,57055	0,1394
dez/02	1	1	0,268196	1,28305	0,9922	2	-4,57234	9,30776	0,0078
jan/03	1	1	0,71131	2,17379	0,9986	2	-5,8464	13,6328	0,0014
fev/03	1	1	0,685731	1,84092	0,9987	2	-5,98721	13,5304	0,0013
mar/03	1	1	0,829108	2,28142	0,9993	2	-6,4349	15,1531	0,0007
abr/03	1	1	-0,0852385	0,951897	0,9734	2	-3,68534	6,49574	0,0266
mai/03	1	1	-0,263121	0,433462	0,9666	2	-3,62748	5,50581	0,0334
jun/03	1	1	-0,505112	0,541779	0,9328	2	-3,13546	4,14611	0,0672
jul/03	2	*1	-1,27257	2,58866	0,5556	2	-1,49572	1,37858	0,4444
ago/03	2	2	0,692132	1,67618	0,9532	1	-2,32189	9,3606	0,0468

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
set/03	2	2	-0,654039	1,77772	0,7711	1	-1,8685	5,86301	0,2289
out/03	1	1	-0,929794	1,48104	0,9217	2	-3,3954	4,75588	0,0783
nov/03	1	1	-0,542036	1,22885	0,9864	2	-4,82567	8,13975	0,0136
dez/03	1	1	-0,528474	0,399555	0,941	2	-3,29716	4,28056	0,059
jan/04	1	1	0,480901	3,86068	0,9998	2	-8,20988	19,5859	0,0002
fev/04	1	1	0,744443	10,1064	0,9909	3	-3,94484	14,4071	0,0091
mar/04	1	1	0,874963	4,77739	0,9999	2	-9,02335	22,9176	0,0001
abr/04	1	1	0,923441	2,70546	0,9999	2	-7,91244	18,7209	0,0001
mai/04	1	1	-0,452083	4,41113	0,7441	2	-1,51962	4,88985	0,2559
jun/04	2	2	0,038974	1,50902	0,7709	1	-1,17456	5,59245	0,2291
jul/04	2	2	1,60187	1,20065	0,963	1	-1,65753	9,37582	0,037
ago/04	2	2	1,03554	0,0381093	0,9534	1	-1,98289	7,73134	0,0466
set/04	2	2	-0,213602	0,297327	0,7976	1	-1,58508	4,69665	0,2024
out/04	1	1	-0,827627	1,49416	0,7786	2	-2,08499	2,35251	0,2214
nov/04	1	1	-0,740031	0,941637	0,8647	2	-2,59465	2,9945	0,1353
dez/04	1	1	-0,0433781	0,405172	0,9846	2	-4,20261	7,06726	0,0154
jan/05	1	1	0,290143	0,546905	0,9964	2	-5,31986	10,1105	0,0036
fev/05	1	1	1,24298	4,03473	0,9999	2	-7,90324	20,6708	0,0001
mar/05	1	1	0,777158	6,03765	0,9999	2	-9,30848	24,5525	0
abr/05	1	1	0,71547	1,61266	0,9996	2	-6,99118	15,3696	0,0004
mai/05	1	1	-0,438915	6,41187	0,9975	2	-6,85515	17,588	0,0016

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
jun/05	1	1	-0,932953	1,22103	0,8851	2	-2,97481	3,64838	0,1149
jul/05	1	*2	-0,115095	0,255869	0,8066	1	-1,54292	4,76789	0,1934
ago/05	2	2	-0,141441	0,332548	0,8244	1	-1,68772	5,08147	0,1756
set/05	2	2	3,44031	1,7807	0,9981	1	-2,79924	15,9162	0,0019
out/05	2	2	1,92716	0,470272	0,9801	1	-1,9714	9,92375	0,0199
nov/05	1	1	-0,345577	0,327232	0,962	2	-3,57817	5,13605	0,038
dez/05	1	1	-0,088379	0,0763391	0,9881	2	-4,51027	7,26375	0,0119
jan/06	1	1	0,722259	1,68486	0,9992	2	-6,44382	14,3607	0,0008
fev/06	1	1	1,07249	3,2667	0,9997	2	-7,17203	18,0994	0,0003
mar/06	1	1	1,18241	3,92894	0,9998	2	-7,32747	19,2923	0,0002
abr/06	1	1	0,515991	1,17824	0,998	2	-5,7173	11,9885	0,002
mai/06	1	1	-1,06482	1,63916	0,769	2	-2,26735	2,38785	0,231
jun/06	1	1	-1,05223	1,8474	0,8809	2	-3,05297	4,19251	0,1191
jul/06	1	1	-1,30491	2,87433	0,7553	2	-2,43196	3,47206	0,2447
ago/06	1	1	-1,80157	14,402	0,8886	2	-4,10812	17,3587	0,0885
set/06	1	1	-1,88684	8,27803	0,5726	2	-2,17934	7,20668	0,4274
out/06	1	1	-1,33427	8,90057	0,964	2	-4,64319	13,862	0,0352
nov/06	1	1	-0,659096	3,76379	0,9923	2	-5,51275	11,8147	0,0077
dez/06	1	1	0,151792	1,53742	0,999	2	-6,73029	13,6452	0,001
jan/07	1	1	0,491087	0,987515	0,9992	2	-6,60512	13,5236	0,0008
fev/07	1	1	0,97673	3,46617	0,9999	2	-8,45799	20,6792	0,0001

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
mar/07	1	1	1,10136	3,52849	0,9999	2	-8,32049	20,7158	0,0001
abr/07	1	1	0,793095	2,00912	0,9993	2	-6,51386	14,9667	0,0007
mai/07	1	1	-0,452458	1,7932	0,8784	2	-2,43022	4,09236	0,1216
jun/07	1	*2	-0,726451	1,25923	0,5922	1	-1,09972	3,66214	0,4078
jul/07	2	2	1,10081	0,0910721	0,9486	1	-1,81484	7,57874	0,0514
ago/07	2	2	1,28243	0,0381963	0,9641	1	-2,00944	8,27831	0,0359
set/07	1	1	-1,31123	2,6954	0,5555	2	-1,53415	1,48486	0,4445
out/07	1	1	-1,42396	4,69616	0,8244	2	-2,97065	6,13318	0,1756
nov/07	1	1	-0,814858	4,85248	0,9894	2	-5,35513	12,2767	0,0106
dez/07	3	3	8,2838	2,2752	0,9998	1	-0,168193	24,2571	0,0002
jan/08	3	3	2,77275	6,44597	0,9228	1	0,291979	16,4855	0,0772
fev/08	3	3	8,78564	2,59144	0,9998	1	0,208778	24,8231	0,0002
mar/08	1	1	0,632287	5,83686	0,9999	3	-8,99786	20,0192	0,0001
abr/08	1	1	0,209592	1,52942	0,9884	2	-4,23615	8,76454	0,0116
mai/08	1	1	-0,55764	1,22708	0,8791	2	-2,54176	3,53895	0,1209
jun/08	1	*2	-0,692381	0,821781	0,6328	1	-1,23671	3,5668	0,3672
jul/08	2	2	-1,0807	0,930032	0,5483	1	-1,27468	2,97436	0,4517
ago/08	2	2	-0,381842	0,447969	0,7374	1	-1,41418	4,16901	0,2626
set/08	2	2	0,659855	0,556737	0,9402	1	-2,09443	7,72168	0,0598
out/08	1	1	-0,945151	1,33189	0,7979	2	-2,31858	2,42238	0,2021
nov/08	1	1	-0,264875	0,721651	0,9582	2	-3,39787	5,33128	0,0418

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
dez/08	3	3	5,8615	3,67031	0,9973	1	-0,0526146	20,5765	0,0027
jan/09	3	3	16,8899	0,024716	1	1	-0,237435	39,3573	0
fev/09	1	1	0,264492	5,87659	0,9996	3	-7,9388	17,2052	0,0003
mar/09	1	1	0,929375	3,15865	0,9999	2	-8,28073	19,9225	0,0001
abr/09	1	1	0,655063	1,3676	0,9993	2	-6,59763	14,2166	0,0007
mai/09	1	1	-0,407982	1,45712	0,993	2	-5,36809	9,72098	0,007
jun/09	1	1	-1,40076	3,2889	0,6783	2	-2,14656	3,12413	0,3217
jul/09	2	2	1,39279	0,0866611	0,9655	1	-1,93918	8,40698	0,0345
ago/09	2	2	0,630001	0,439676	0,9359	1	-2,0517	7,45945	0,0641
set/09	2	2	1,49242	0,11604	0,9694	1	-1,9622	8,68165	0,0306
out/09	2	2	-0,628497	0,868993	0,6435	1	-1,21915	3,70667	0,3565
nov/09	1	1	-0,340039	0,188802	0,9687	2	-3,77077	5,3939	0,0313
dez/09	1	1	0,183436	0,249409	0,9965	2	-5,4736	9,90711	0,0035
jan/10	1	1	0,776857	2,52199	0,9998	2	-7,90265	18,2246	0,0002
fev/10	1	1	1,07255	8,39343	0,9998	3	-7,54535	20,5513	0,0002
mar/10	1	1	0,967465	3,0745	0,9999	2	-8,18088	19,7148	0,0001
abr/10	1	1	0,486408	0,867877	0,9988	2	-6,27553	12,7354	0,0012
mai/10	1	1	-0,279899	0,394743	0,9661	2	-3,62892	5,43642	0,0339
jun/10	1	*2	-0,969407	1,63772	0,5051	1	-0,989725	3,33472	0,4949
jul/10	2	2	0,648349	0,121256	0,9294	1	-1,92892	6,93216	0,0706
ago/10	2	2	1,80635	0,548402	0,9848	1	-2,36345	10,5444	0,0152

(continua)

(continuação)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
set/10	1	1	-1,44916	3,3939	0,5901	2	-1,81362	2,46645	0,4099
out/10	1	*2	-0,571605	0,823279	0,6605	1	-1,23718	3,81081	0,3395
nov/10	1	1	-0,794533	1,88154	0,7447	2	-1,86519	2,36648	0,2553
dez/10	1	1	-0,695449	1,71314	0,7986	2	-2,07304	2,81196	0,2014
jan/11	1	1	0,672991	1,51303	0,9991	2	-6,3229	13,8484	0,0009
fev/11	1	1	0,977955	2,8443	0,9996	2	-6,8958	16,9354	0,0004
mar/11	1	1	0,873051	2,1873	0,9997	2	-7,14448	16,566	0,0003
abr/11	1	1	0,500735	1,04518	0,9982	2	-5,80338	11,997	0,0018
mai/11	1	1	0,0548001	1,17919	0,982	2	-3,94584	7,5241	0,018
jun/11	1	1	-0,958628	1,6188	0,7357	2	-1,98222	2,00962	0,2643
jul/11	2	2	1,29169	0,505406	0,9527	1	-1,71109	8,16734	0,0473
ago/11	2	2	0,215088	0,602249	0,8346	1	-1,40316	5,49511	0,1654
set/11	2	2	0,642096	0,187454	0,9071	1	-1,63625	6,40052	0,0929
out/11	1	1	-1,04547	2,29986	0,6025	2	-1,46116	1,47488	0,3975
nov/11	1	1	-0,300474	1,30667	0,9355	2	-2,97484	4,99904	0,0645
dez/11	1	1	0,277045	0,897179	0,9943	2	-4,87622	9,54734	0,0057
jan/12	1	1	0,832103	2,05706	0,9995	2	-6,82359	15,7121	0,0005
fev/12	1	1	1,26327	4,53885	1	2	-8,84395	23,0969	0
mar/12	1	1	1,28029	8,40076	1	3	-9,02761	23,9386	0
abr/12	1	1	0,321781	2,17672	0,9995	2	-7,31666	15,7972	0,0005
mai/12	1	1	-0,551345	0,296365	0,9574	2	-3,66345	4,86421	0,0426

(continua)

(conclusão)

Data	Grupo Atual	Grupo mais provável	Maior Pontuação	Quadrado da Distância	Probabilidade	2º Grupo mais provável	2ª Maior Pontuação	Quadrado da Distância	Probabilidade
jun/12	2	2	-0,94987	1,38262	0,5254	1	-1,05156	3,24236	0,4746
jul/12	2	2	0,987774	0,0197152	0,9497	1	-1,94965	7,55093	0,0503
ago/12	1	1	-1,34628	4,55104	0,8723	2	-3,26794	6,73799	0,1277
set/12	1	1	-1,12556	4,86034	0,9593	2	-4,28599	9,52485	0,0407
out/12	1	1	-0,802723	8,29594	0,9917	2	-6,23703	17,5082	0,0043
nov/12	3	3	8,55393	2,09752	0,9999	1	-0,38753	25,0584	0,0001
dez/12	3	3	24,5895	1,13399	1	1	-0,345945	56,0828	0
jan/13	3	3	38,196	10,0828	1	1	0,274675	91,0034	0
fev/13	3	3	21,1765	0,599101	1	1	0,199336	47,6314	0
mar/13	3	3	1,14065	8,33931	0,6601	1	0,476784	14,745	0,3399
abr/13	1	1	0,601141	1,3302	0,9987	2	-6,05847	12,9931	0,0013
mai/13	1	1	-0,348291	0,816922	0,9921	2	-5,17554	8,81505	0,0079
jun/13	2 *1		-0,94141	1,93782	0,6866	2	-1,72579	1,85022	0,3134
jul/13	2	2	-0,831451	1,76252	0,5369	1	-0,97942	3,71482	0,4631
ago/13	2	2	0,835284	0,589655	0,9522	1	-2,15583	8,22825	0,0478
set/13	1 *2		-1,92333	9,00285	0,5518	1	-2,13138	11,0753	0,4482
out/13	1	1	-1,29345	2,71617	0,7374	2	-2,32599	3,12488	0,2626
nov/13	3	3	8,955	2,24692	0,9999	1	-0,736117	26,7071	0,0001
dez/13	3	3	34,2599	7,31775	1	1	-0,741095	82,3976	0

APÊNDICE E – RESUMO DOS MODELOS ARIMA AJUSTADOS ÀS SÉRIES DO IMCRB E IMCRBdif

```
RGui (32-bit) - [R Console]
Arquivo Editar Visualizar Misc Pacotes Janelas Ajuda
[Icons]

> dados <-ts(read.table("C:/Users/Sheila/Dropbox/projeto_tese/ARTIGO_PERIODICO/escoresfatores.csv"
+ ,header=TRUE, dec=""),start=c(1990, 1), frequency=12)
> IMCRB <- 0.4588*(dados[,1])+0.3083*(dados[,2])+0.0966*(dados[,3])+0.0241*(dados[,4])+0.0233*(dados[,5])
> model.IMCRB = auto.arima(IMCRB, max.p=2, max.q=2, max.P=2, max.Q=2, max.order=5, max.d=2, max.D=1,
+ start.p=2, start.q=2, start.P=1, start.Q=1, stationary=FALSE, seasonal=TRUE, ic=c("aicc", "aic", "bic"))
> summary(model.IMCRB)
Series: IMCRB
ARIMA(2,1,0) (1,0,0) [12]

Coefficients:
      ar1      ar2      sar1
-0.393  -0.1904  0.8421
s.e.    0.067   0.0645  0.0363

sigma^2 estimated as 0.07403: log likelihood=-41.17
AIC=90.34  AICc=90.48  BIC=104.98

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 0.0007358259 0.2716127 0.2123171 11.2052 119.4627 1.010245
              ACF1
Training set -0.0310246
>
> IMCRBdif <- 0.4588*(dadosdiff[,1])+0.3083*(dadosdiff[,2])+0.0966*(dadosdiff[,3])+0.0241*(dadosdiff[,4])+0.0233*(dadosdiff[,5])
> model.IMCRBdif = auto.arima(IMCRBdif, max.p=2, max.q=2, max.P=2, max.Q=2, max.order=5, max.d=2, max.D=1,
+ start.p=2, start.q=2, start.P=1, start.Q=1, stationary=FALSE, seasonal=TRUE, ic=c("aicc", "aic", "bic"))
> summary(model.IMCRBdif)
Series: IMCRBdif
ARIMA(2,0,0) (1,0,0) [12] with zero mean

Coefficients:
      ar1      ar2      sar1
-0.393  -0.1904  0.8421
s.e.    0.067   0.0645  0.0363

sigma^2 estimated as 0.07403: log likelihood=-41.17
AIC=90.34  AICc=90.48  BIC=104.98

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 0.0007382229 0.2720855 0.2130567 24.62983 168.3937 0.8527876
              ACF1
Training set -0.03102868
\ |
```