



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

IMPLEMENTAÇÃO DE UM PROGRAMA PARA INTEGRAR MODELOS DE PREVISÃO DE CHEIAS FLUVIAIS

**RELATÓRIO FINAL DE PROJETO DE INICIAÇÃO CIENTÍFICA
(PIBIC/CNPq/INPE)**

Carmo Vinieri Filho (UNIVAP, Bolsista PIBIC/CNPq)

[E-mail: carmovinieri@yahoo.com.br](mailto:carmovinieri@yahoo.com.br)

Dr. Acioli Antonio de Olivo (LAC/CTE/INPE, Orientador)

[E-mail: acioli@lac.inpe.br](mailto:acioli@lac.inpe.br)

Junho de 2006

RESUMO

O objetivo deste trabalho é implementar um sistema de alerta contra cheias, cujos métodos foram desenvolvidos por Olivo (2004). Os modelos utilizados são de natureza “caixa-preta”, com a observação do nível fluviométrico a montante da seção de interesse onde se deseja realizar a previsão. A partir de uma série de registros de níveis fluviométricos da bacia do rio Itajaí em Santa Catarina e de técnicas como regressão múltipla por mínimos quadrados, modelos auto-regressivos e Modelos de Composição de Especialistas Locais (MCEL), implementar-se-á um sistema que sirva de alarme-resposta em tempo real, com a finalidade de fornecer informações às autoridades de Defesa Civil, uma estimativa confiável do crescimento do nível do rio, com antecedência suficiente para a tomada das decisões necessárias. Propõe-se uma implementação utilizando uma linguagem Orientada a Objetos, mais especificamente JAVA, com a finalidade de executar todos os passos supramencionados na referência. Até o presente momento os métodos implementados já executam boa parte do processo, desde o tratamento dos dados, passando por regressões polinomiais, até a determinação dos coeficientes da regressão. Os resultados já obtidos são satisfatórios, tais como aprendizagem de métodos estatísticos e matemáticos e sua implementação, bem como seu comportamento perante a massa de dados, aprendizagem sobre comportamento fluviométricos em relação ao modelo proposto, aprendizagem passo a passo dos métodos que compõem a regressão do modelo, com ferramentas para manipulação de dados, como Excel, estudos de modelos auto-regressivos, manipulação e tratamento da massa de dados e conhecimento de técnicas para transmissão de dados e armazenamento.

SUMÁRIO

RESUMO	2
SUMÁRIO.....	3
LISTA DE TABELAS	4
LISTA DE FIGURAS.....	5
1. INTRODUÇÃO.....	6
2. IMPLEMENTAÇÃO.....	9
2.1 REGRESSÃO LINEAR.....	9
2.2 AUTO-CORRELAÇÃO SERIAL DE DURBIN	12
2.3 BETA ZERO	13
2.4 REGRESSÃO DOS RESÍDUOS.....	15
2.5 O SOFTWARE PREVISÃO.....	16
3. VALIDAÇÃO.....	17
4. CONCLUSÕES E TRABALHOS FUTUROS	19
4.1 ANÁLISE DE RESULTADOS	19
4.2 TRABALHOS FUTUROS	19
5. REFERÊNCIAS BIBLIOGRÁFICAS.....	20

Lista de Tabelas

TABELA 3. 1 – COEFICIENTES RO DURANTE PROCESSO	17
TABELA 3. 2 – COEFICIENTES DE PREVISÃO EM 1º REGRESSÃO	17
TABELA 3. 3 – COEFICIENTES DE PREVISÃO FINAIS	18

Lista de Figuras

FIGURA 1. 1 – GRÁFICO DE CALIBRAÇÃO DO MODELO DE REGRESSÃO: ALCANCE 6 HORAS.....	7
FIGURA 1. 2 – GRÁFICO DE CALIBRAÇÃO DO MODELO DE REGRESSÃO CORRIGIDO PELOS RESÍDUOS: ALCANCE 6 HORAS	7
FIGURA 1. 3 – DIAGRAMA DE BLOCOS DO MODELO DE OLIVO	8
FIGURA 2. 1– AMBIENTE DE DESENVOLVIMENTO JAVA	9
FIGURA 2. 2 – MATRIZ DO MÉTODO DE GAUSS PARA CASO PROPOSTO.....	10
FIGURA 2. 3 – MATRIZ GAUSS RESULTANTE	11
FIGURA 2. 4 – FLUXOGRAMA REPRESENTANTE DO MÉTODO DA REGRESSÃO	12
FIGURA 2. 5 – DIAGRAMA EXEMPLO PARA MÉTODO MELHOR BETA 0.....	14
FIGURA 2. 6– DIAGRAMA DO MÉTODO DE REGRESSÃO DE RESÍDUOS.....	15
FIGURA 3. 1 – GRÁFICO TEMPO X NÍVEL PREVISTO.....	18
FIGURA 3. 2 – DIAGRAMA DE BLOCOS DO SISTEMA COMO UM TODO	18

1. INTRODUÇÃO

A interação entre o homem e os rios é muito antiga tão antiga quanto a sua existência, assim problemas com inundações urbanas afetam diversos municípios brasileiros causando inúmeras perdas até mesmo irreparáveis, principalmente quando atinge locais densamente habitados.

Dentre um grande conjunto de fatores os quais causam tais desastres, um que tem responsabilidade dos nossos representantes de estado são as habitações de risco as chamadas 'ribeirinhas', essa por sua vez são a distribuição da população ao redor dos afluentes.

Uma serie de soluções são propostas ao longo dos anos, construção de diques, reservatórios, canais e outras ações fiscais, paralelamente alternativas não estruturais como adoção de sistemas de alarme-resposta, através de modelos matemáticos.

Os modelos mais utilizados são modelos de chuva-vazão, que a partir de uma série de registros de precipitação. Modelos de chuva-vazão procuram descrever uma série de etapas desde a vazão até o escoamento. No modelo de OLIVO (2004), utilizamos um modelo caixa-preta onde só nos interessa os dados de entrada e saída, isso pela dificuldade do tratamento dos modelos hidrológicos devido sua falta de linearidade, falta de inconsistência dos dados e lacunas em series pluviométricas e fluviométricas.

O objetivo do modelo é obter dados em curto prazo de 4 a 6 horas de antecedência, abandonando todos os fenômenos correlatos ao crescimento da onda como evaporação e a infiltração, admitindo que a vazão do rio pode representar todos esses fenômenos e alguns não citados.

A adoção de modelos hidrológicos conceituais embasados em fenômeno físico para a previsão de cheias fluviais são tentativas de calibração que caracteriza uma natureza empírica, são identificados vários procedimentos que sinteticamente resume-se em:

- Escolha do conjunto de eventos que represente o comportamento da bacia a ser estudada;
- Análise exploratória dos dados para a identificação das variáveis de decisão mais relevantes;

- Escolha de um modelo que permita uma estimativa tão confiável quanto possível da vazão em uma seção de interesse;
- Ajuste da previsão inicial de modo a minimizar a diferença entre a vazão prevista e a observada;
- Validação e análise de sensibilidade do modelo;

Os resultados obtidos na calibração do modelo foram excelentes como mostra os gráficos das Figuras 1.1 e 1.2:

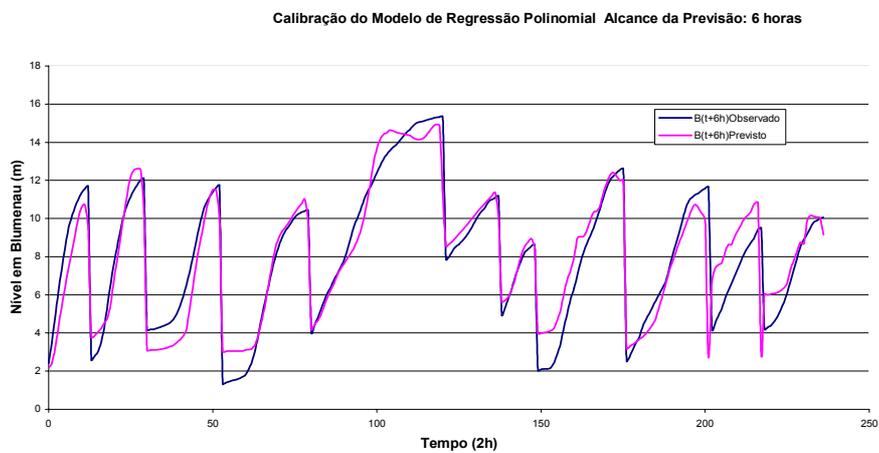


Figura 1.1 – Gráfico de Calibração do Modelo de Regressão: Alcance 6 horas

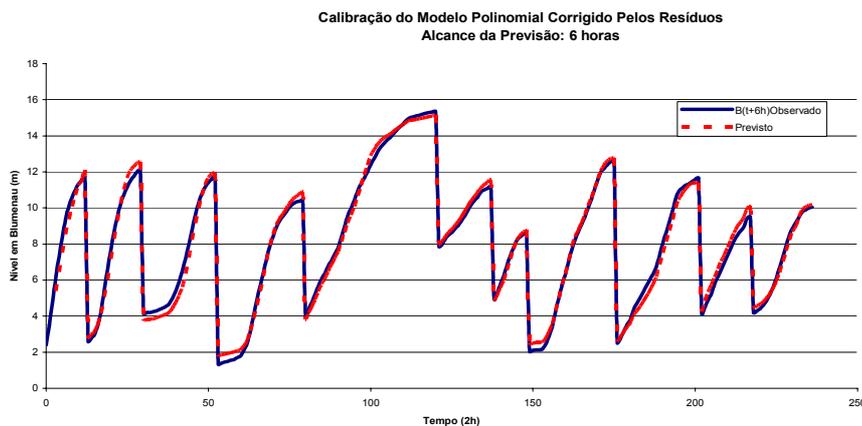


Figura 1.2 – Gráfico de Calibração do Modelo de Regressão Corrigido pelos Resíduos: Alcance 6 horas

O principal objetivo do trabalho de Olivo (2004) é mostrar que é possível construir um sistema de alarme-resposta em tempo real eficiente, com antecedência de até 6 horas, somente com dados de eventos fluviométricos, disponíveis mesmo para bacias que possuam poucas informações hidrográficas.

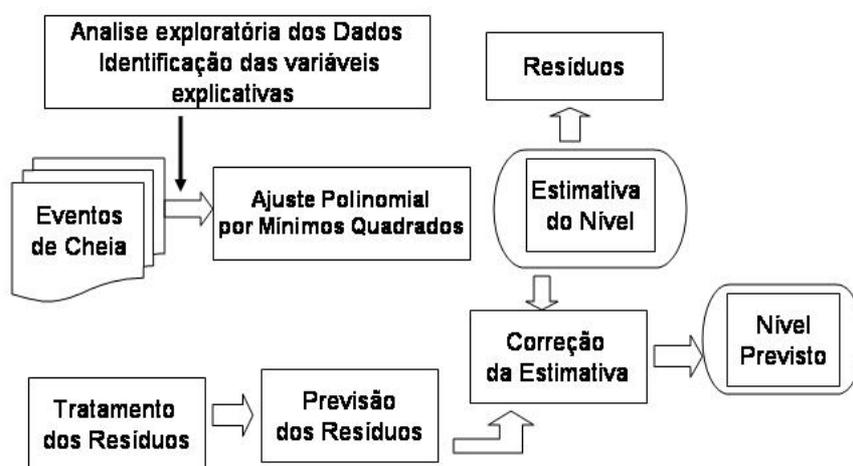


Figura 1.3 – Diagrama de blocos do Modelo de Olivo

Utilizando-se dos conhecimentos adquiridos no estudo supramencionado propõem-se uma Implementação do Modelo de Olivo.

2. IMPLEMENTAÇÃO

A Implementação proposta será uma adaptação do algoritmo do modelo de Olivo para um algoritmo baseado-se em uma linguagem de programação mais especificamente a linguagem Orientada a Objeto, JAVA.

A linguagem Orientada a Objetos vem sendo explorada por suas facilidades de uso, portabilidade, segurança e por sua fácil integração aos outros ambientes destacando-se a Internet.

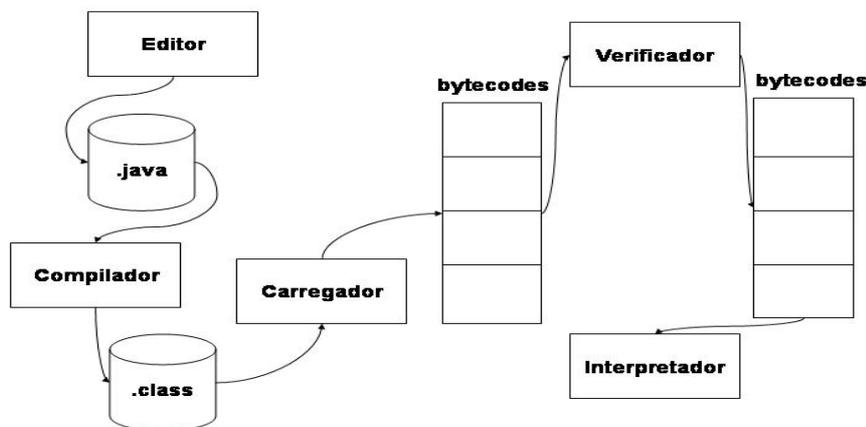


Figura 2. 1– Ambiente de Desenvolvimento Java

No diagrama acima é exemplificado como é transcodificado a linguagem de alto nível, o JAVA, para a linguagem de máquina o bytecode 0s e 1s. Para maiores informações de linguagem Java consultar o capítulo de referencias bibliográficas.

Para esta Implementação o código foi quebrado em blocos, que por sua vez executam passos do método, assim cada tópico a seguir exemplificará um bloco de operação.

2.1 Regressão Linear

O método de Regressão Linear consiste basicamente em um método de aproximação dos mínimos quadrados, é um acerto dos valores que compõem a reta. Neste caso em específico tratamos de uma Regressão Polinomial que determinara os coeficientes da curva. A expressão matemática que representa a reta é dada pela equação (1):

$$y = a_0 + a_1x + E \quad (1)$$

onde a_0 e a_1 são os coeficientes de interpretação e declive da reta, E é o resíduo ou erro.

Em uma Regressão Polinomial a equação muda conforme o grau polinomial da equação, o modelo de OLIVO, é algo parecido com a equação (2):

$$y = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (2)$$

onde a_0, a_1, a_2 são os coeficientes da regressão, e os valores de x, x^2 e x^n são as medidas das grandezas que se deseja calcular.

O método dos mínimos quadrados pode ser arranjado de tal forma que obtenhamos uma série de somatórios caracterizando varias equações que variam de $i=1$ a $i=n$.

Essa série de equações, pode ser interpretada com um método iterativo, nomeado de “GAUSS JORDAN”. O método de Gauss-Jordan pode resolver n equações lineares simultaneamente, ele consiste em uma matriz que tem como componentes nada mais que os somatórios, a Figura 2.2 ilustra a matriz:

$$\left[\begin{array}{cccc|c} n & \sum X & \sum X^2 & \sum X^3 & \sum Y \\ \sum X & \sum X^2 & \sum X^3 & \sum X^4 & \sum X \cdot Y \\ \sum X^2 & \sum X^3 & \sum X^4 & \sum X^5 & \sum X^2 \cdot Y \\ \sum X^3 & \sum X^4 & \sum X^5 & \sum X^6 & \sum X^3 \cdot Y \end{array} \right]$$

Figura 2. 2 – Matriz do Método de Gauss para Caso Proposto

com essa abordagem obtêm-se os coeficientes conhecidos como betas.

O método da Regressão opera de duas maneiras diferentes, a primeira buscando os dados que estão armazenados em arquivos do tipo txt, e transformando-os em vetores com valores inteiros, assim possibilitando a manipulação de forma a obtermos os valores necessários das variáveis mencionadas acima pela matriz de Gauss. A segunda forma de abordagem do método da Regressão nada mais é que uma replica da primeira mais com alterações sutis, a mudança ocorrida nesse método é simples, ao invés de trabalharmos com as variáveis iniciais nos devemos calcular novas variáveis de estimativa, para isto foi feito o método ‘calculanovo’, que será descrito nos próximos tópicos.

Quando invocamos o método pela primeira ele cria variáveis para que possa ser efetuado os cálculos e obtermos os somatórios descritos na Figura 2.2, obtendo esses valores invocamos o método da Matriz Gaussiana, esse método monta a Matriz de forma a ficar como na Figura 2.2, executa os cálculos gaussianos, utilizando técnica de pivô e normalização das linhas da matriz, até que a Matriz fique como na Figura 2.3.

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & \beta_0 \\ 0 & 1 & 0 & 0 & \beta_1 \\ 0 & 0 & 1 & 0 & \beta_2 \\ 0 & 0 & 0 & 1 & \beta_3 \end{array} \right]$$

Figura 2.3 – Matriz Gauss Resultante

Obtendo os valores de β_0 , β_1 , β_2 e β_3 , que são os coeficientes de previsão, esses valores são guardados em um vetor para uso mais adiante.

Após obtermos os primeiros coeficientes de previsão passamos a verificar o grau de correlação entre eles, essa correlação é denominada ρ (rô), e temos um bloco que tratará especificamente dela, enquanto essa correlação estiver baixa novos coeficientes devem ser calculados invocando a Aplicação da Regressão, para isso temos que calcular novos valores de estimativa com um método parecido com a Equação (3):

$$X^* = X_i - (\rho * X_{i-1}) \quad (3)$$

Terminado esse calculo para todas as variáveis (X , X^2 , X^3 , X^4 , X^5 , X^6 e Y), é invocado novamente o método da Matriz Gaussiana, esse procedimento é executado até a correlação atingir o grau que atende ao erro e chega mais próximo a 1.

Um fluxograma é descrito na Figura 2.4 para exemplificar melhor o método da regressão:

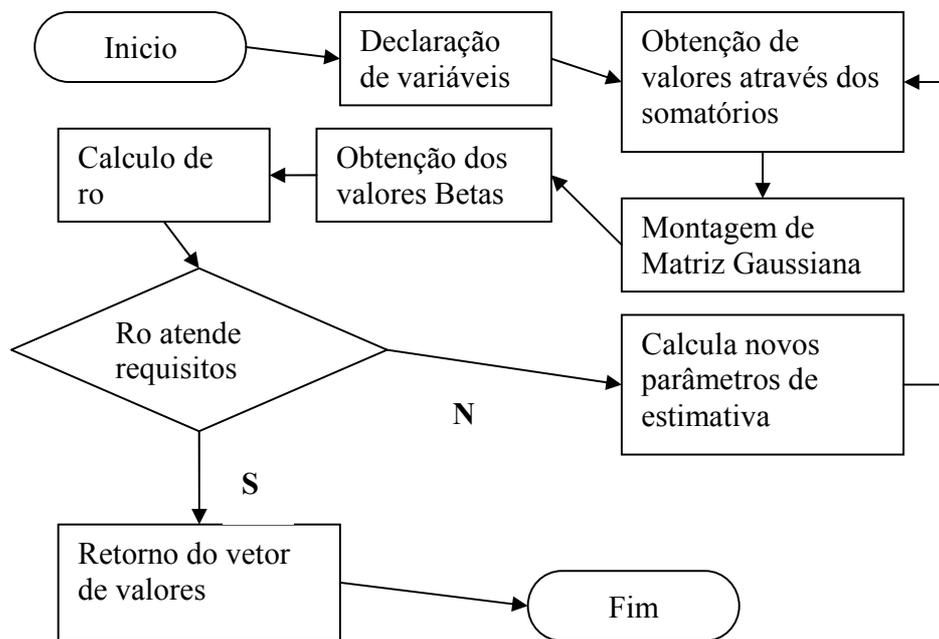


Figura 2. 4 – Fluxograma representante do método da Regressão

2.2 Auto-Correlação Serial de Durbin

Quando utilizamos um método, onde quer verificar a relação entre duas medidas quantitativas utilizamos uma análise linear, a correlação, o fator que verifica o grau de associação entre as variáveis da amostra. No método dos mínimos quadrados a teoria aponta que existe um alto grau de correlação entre as variáveis betas, para diminuirmos essa autocorrelação serial utiliza - se a Equação 2.3 para as variáveis, como supramencionado.

Após o ajuste e obtenção dos resíduos, calcula-se ρ , que é a correlação entre $E(t)$ e $E(t-1)$, aplicando a equação acima obtém-se os novos parâmetro de Regressão. Esse método é recursivo.

Para Implementar a auto correlação é necessário informar (passar parâmetros) para que o método execute os cálculos no caso do método da auto correlação passamos dois vetores contendo valores dos resíduos. Os valores dos resíduos são obtidos através de um outro método específico para obter resíduos, esse por sua vez usa os betas (coeficientes de previsão) para achar os resíduos simplesmente fazendo uma previsão e subtraindo-a do valor de Y_i que é o valor do nível observado.

Através da obtenção dos resíduos o método da autocorrelação serial ou método ‘achar ρ_0 ’, executa um calculo entre dois vetores, os vetores corresponde a resíduos e resíduos+1, ou seja um cálculo entre um vetor e sua imagem deslocada de uma posição.

Este cálculo envolve media aritmética, variância, Somatória e uma formulam que acha o valor de ρ .

$$s_{xy} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{n - 1} = \frac{\sum x_i y_i - n\bar{x}\bar{y}}{n - 1}$$

$$r = \frac{s_{xy}}{s_x s_y}$$

Equação 2. 1 – Equação para achar ρ

Sendo S_x e S_y somatórios, \bar{x} e \bar{y} medias e n o numero de elementos do vetor.

2.3 Beta Zero

Existe uma regra que diz que apenas o β_0 não se corrigi durante o processo da minimização da auto correlação serial, e isso seria um ponto muito crítico para o resultado final afetando toda a previsão porem normaliza-se o β_0 com o método ‘melhorBetazero’ ajustando assim nossos coeficientes de previsão.

O método do melhor β_0 é um método auto-ajustavel ele vai se adequando conforme a soma do quadrado dos resíduos.

A Implementação desse método é a etapa mais difícil, apesar de ser uma abordagem simples e trivial de codificar, como o ajuste é feito utilizando bom senso é difícil computacionalizar este tipo de situação.

Após obtermos os melhores coeficientes de previsão, a melhor correlação ρ , é hora de corrigir o β_0 , nesse método passamos como parâmetros um vetor com contendo os dados β (0,1,2 e 3) e ρ , assim podemos fazer previsões e compará-las aos valores observados do vetor de Y, obtém-se também os resíduos e resíduos +1 através deles é que executamos as comparações, se o somatório do quadrado dos resíduos (SQR) diminuir significa que o β_0 está se aproximando do valor que se adequa melhor na equação e verificamos fazendo previsões e verificando a proximidade entre o valor previsto e o observado, e assim vai ocorrendo até que chegue ao valor mais próximo e também quando o SQR, começar a aumentar, significa que temos o melhor β_0 .

Abaixo um Diagrama de blocos da Figura 2.5 ilustra o método β_0 :

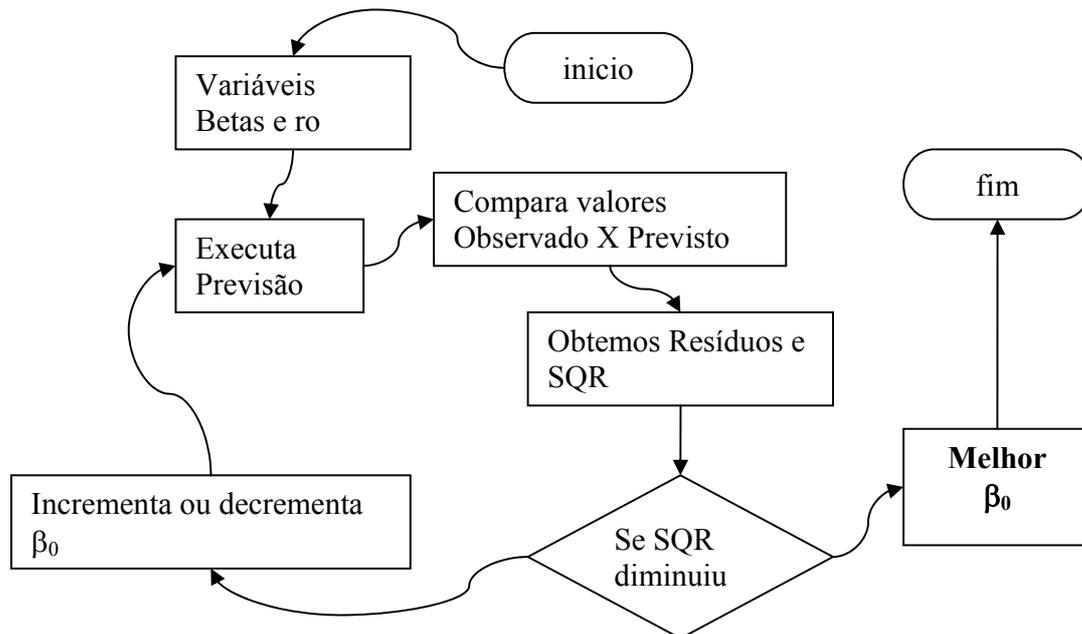


Figura 2. 5 – Diagrama exemplo para método melhor beta 0

2.4 Regressão dos Resíduos

Na calibração do modelo temos como ultima etapa a correção dos resíduos, isso é feito com uma simples Regressão reta, linear entre dois valores. A equação (4) abaixo exemplifica adequadamente:

$$R(t+\text{alcance}) = \alpha_0 + \alpha_1 * R(t) \quad (4)$$

Onde $R(t+\text{alcance})$, é a previsão do erro
 α_0 é o último resíduo e α_1 são os parâmetros da regressão.

O calculo dos resíduos é feito utilizando uma regressão linear simples entre o resíduo e o próprio resíduo deslocado de 4 posições.

Onde são calculadas as médias dos dois vetores de resíduos, os somatórios do quadrado de resíduo ou na equação (6) SQX e somatório da multiplicação entre resíduo e resíduo + 4 ou na equação (7), e tudo é calculado com as formulas:

$$\alpha_0 = \text{media } y - (b * \text{media } x) \quad (6)$$

$$\alpha_1 = (SPXY/SQX) \quad (7)$$

O processo de implementação é simples neste método, começando com declarações de variáveis, essas para armazenar valores dos cálculos de media de y e x, somatórios do produto entre x e y (SPXY) e do quadrado de x (SQX) , como em alguns outros métodos este também vai precisar de passagem de parâmetros, no caso um vetor com os últimos resíduos, desloca-se 4 posições do vetor de resíduos para obter um outro vetor de resíduos semelhante mas com shift 4, executa então os cálculos de média e somatórios e obtemos os valores de α_0 e α_1 que correspondem ao erro.

Abaixo o Diagrama na Figura 2.6 exemplifica o método:

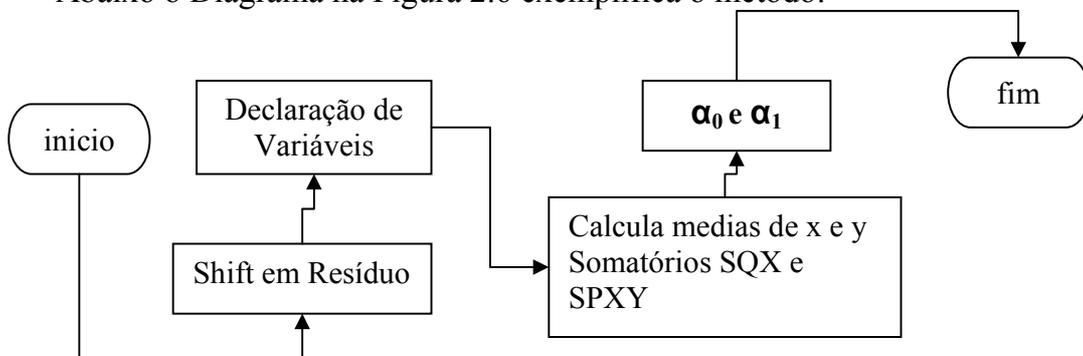


Figura 2. 6– Diagrama do método de regressão de resíduos

2.5 O Software Previsão

O Software Previsão é uma implementação dos cálculos efetuados durante o processo de obtenção das variáveis as quais utilizaremos para gerar valores da previsão, utilizando linguagem orientada a objetos 'JAVA'. O propósito de utilizar o JAVA é que essa linguagem permite diversos tipos de aplicação e tem uma vasta lista de packpages prontas, isso possibilita diversos tipos de abordagem para o Software, seja ele para fim de estudos ou comercial.

Como supramencionado os métodos se quebram em módulos como, regressão, correlação, melhorB0, achaRô, esses os mais importantes, para cada método são executados inúmeros cálculos que envolvem conceitos de matemática somados a estatística e ainda utilizamos conceitos de vetores e matrizes para manipulação e transformação de dados.

A Implementação do sistema é feita com interfaces gráfica e está em fase de desenvolvimento, abaixo exemplo:

TELAS DO SOFTWARE

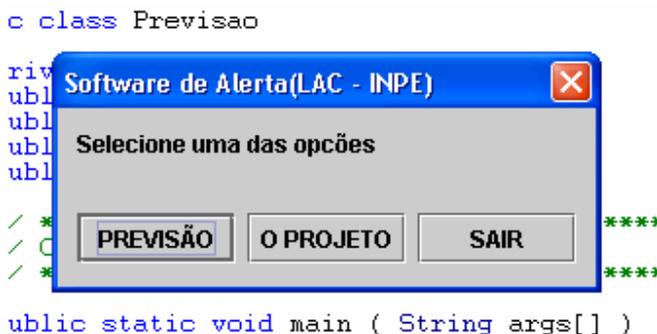


Figura 2. 7 – Tela inicial do Software



Figura 2. 8 – Segunda Tela do Software (em fase de adaptações)

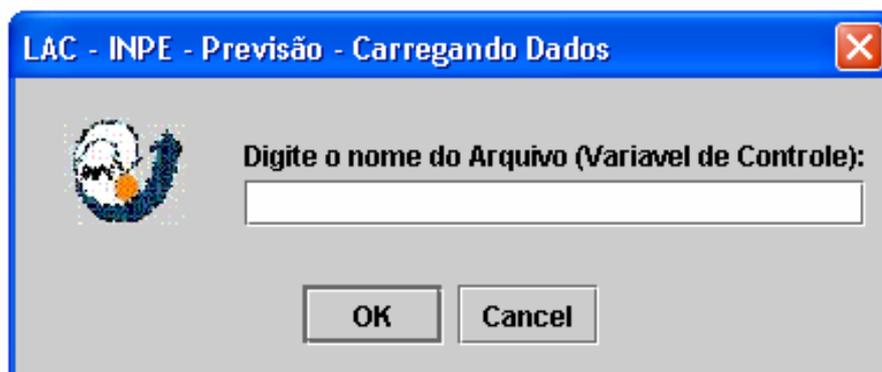


Figura 2. 9 – Campo para digitarmos o nome do Arquivo a ser estudado

3. VALIDAÇÃO

Para validarmos um sistema computacional são necessários inúmeros testes, inspeções, walkthroughs, uma serie de testes nos módulos que compõem o sistema, o proposto sistema não é muito grande portanto não é necessário fragmentá-lo em vários módulos, mas para efeito de gerencia de Software, e ainda para correção de erros verificamos e comparamos os valores de cada método para saber se a implementação está sendo feita com precisão.

Verificou-se que em determinados dados, o resultado final ($\beta_0, \beta_1, \beta_2, \beta_3$), os valores estão completamente em escala absurda o que compromete e muito o sistema, porem para efeito geral os fragmentos estão fazem cálculos muito precisos. Por exemplo o coeficiente ρ :

Método realizado	Coefficiente (ρ)
1° passagem	0.5572
2° passagem	0.884645
3° passagem	0.8585
4° passagem	0.8559
5° passagem	0.88205
6° passagem	0.9447
7° passagem	0.94577
Melhor ρ (ρ)	0.9459

Tabela 3. 1 – Coeficientes ρ durante processo

Os método para acharmos o melhor ρ ou a correlação está executando sua função com boa precisão e se aproxima muito ao valor original.

O método da matriz de Gauss também monta e calcula corretamente os valores dos betas que não ficam próximos aos originais mais tem o mesmo estilo de relação. Abaixo se apresenta uma tabela para comparação de tais:

Betas	Valores novos	Valores Existentes
β_0	1.310	0.7768
β_1	1.9622	2.3251
β_2	-0.0669	-0.13281
β_3	0.0029	0.006514

Tabela 3. 2 – Coeficientes de Previsão em 1° Regressão

Mas existe um problema por que esse método cresce como uma exponencial e seu valor final fica em escala muito altas, aproximadamente 10^7 .

Betas Finais	Valores
β_0	-2443.129
β_1	5.01038×10^7
β_2	-2.2508×10^7
β_3	2630966.578

Tabela 3. 3 – Coeficientes de Previsão finais

E esses valores geram uma seqüência de previsões com um valor de escala em 10^7 :

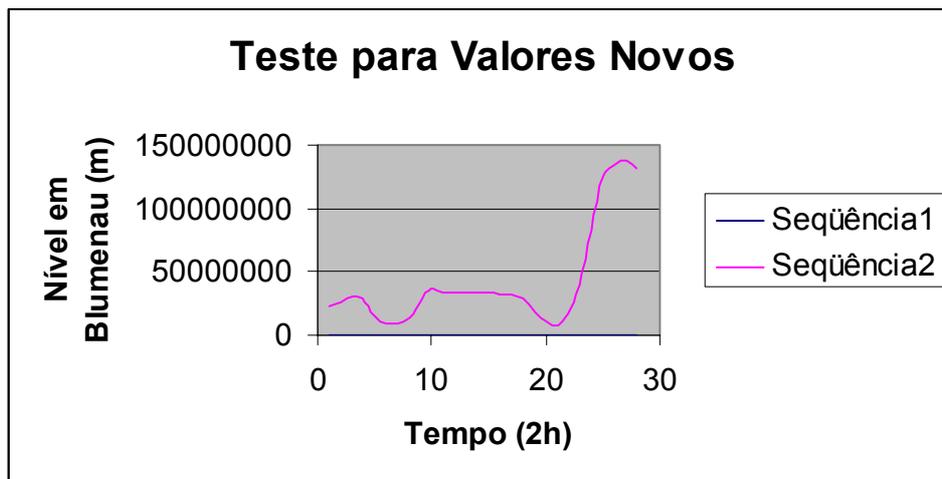


Figura 3. 1 – Gráfico Tempo X Nível Previsto

Os valores como visto no gráfico sobem como uma exponencial e os valores na escala observada de 10^7 .

O diagrama da Figura 3.2 abaixo exemplifica o processo como um todo:

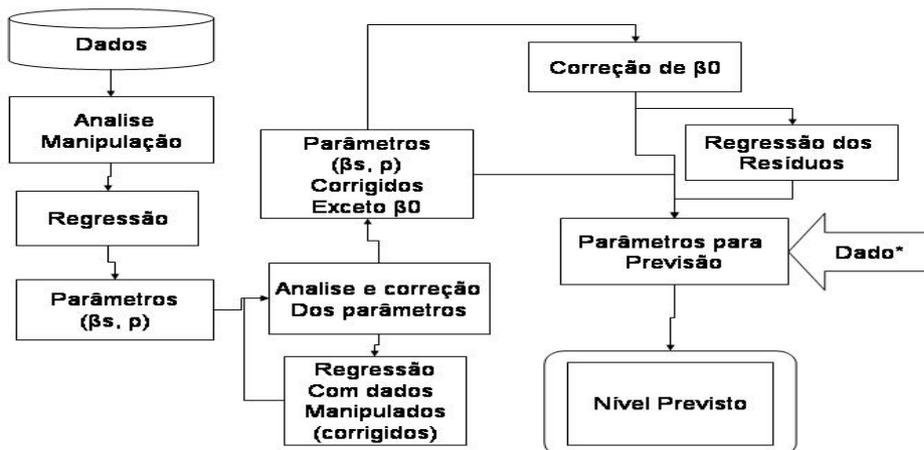


Figura 3. 2 – Diagrama de blocos do Sistema como um todo

4. CONCLUSÕES E TRABALHOS FUTUROS

A partir de uma série de eventos de cheias, com registros fluviométricos, pode-se obter, através de regressões lineares combinadas com a análise de resíduos, um sistema de alerta para monitorar o crescimento do nível de um rio, com a antecedência e a confiabilidade necessária e suficiente para servir de auxílio à defesa civil em caso de eventos extremos.

A implementação do modelo objetiva automatizar os procedimentos descritos em Olivo para a obtenção de equações que descrevam o crescimento do rio, assim podendo fazer previsões de nível com 4 horas de antecedência, apenas com a leitura do ponto de controle no afluente do rio.

4.1 Análise de Resultados

Como supramencionado os resultados obtidos até o presente momento estão em escala adequada para que este trabalho tem um caráter de um projeto completo representando uma parcela significativa de o que pode vir a ser um produto, que beneficiará a sociedade, caso venha a ser concretizado.

Os valores de dos resultados quebrados em blocos estão correspondendo aos valores existentes na calibração do modelo, como mostra a tabela 3.2 supramencionada, isso nos faz concluir que os módulos estão fazendo os cálculos necessários obtendo um bom resultado. Porém quando se integrou todo os módulos do sistema os resultados fugiram da realidade ficando muito fora da escala, como ilustra a tabela 3.3.

De modo geral o trabalho tem resultados satisfatórios, apesar de não estarem rigorosamente corretos, pois o sistema ainda precisa de ajustes para melhora dos resultados finais.

4.2 Trabalhos Futuros

Algumas etapas ainda estão previstas a serem executadas, além da correção dos coeficientes de previsão estima-se integrar ao sistema os seguintes tópicos:

Estudar e Implementar Especialistas Locais

Estudar e Implementar Redes Neurais

Teste para validação

5. REFERÊNCIAS BIBLIOGRÁFICAS

[1] OLIVO, A.A. (2004). Modelos matemáticos para a previsão de cheias fluviais. 2004. 154 f. Tese de Doutorado. Instituto Tecnológico de Aeronáutica (ITA). São José dos Campos, SP.

[2] Chapra, C.; Canale, R.P. (1985). Numerical methods for engineers with personal computer applications. McGraw-Hill.

[3] Jandi, P. Jr. (1999). Introdução ao Java. 1999. Apostila p. cm. Universidade São Francisco.