

# | 493 | UM MODELO PARA A PREVISÃO DE ENCHENTES EM GRANDES CIDADES: METODOLOGIA DESENVOLVIDA PARA O RIO BELÉM, EM CURITIBA

*Fabio Teodoro de Souza, Carlos Mello Garcias*

## **Resumo:**

Devido a um problema histórico de uso inadequado do solo, as regiões próximas aos rios nas grandes cidades sofrem frequentemente com as precipitações intensas que deflagram enchentes e causam impactos socioeconômicos e ambientais. As consequências das ações antrópicas incluem a degradação do solo, da fauna e da flora, o acúmulo de lixo em locais impróprios, a obstrução de obras de drenagem, a descarga de esgotos domésticos nos cursos dos rios, o aumento do coeficiente de escoamento superficial das águas pluviais causada pela impermeabilização da superfície, entre outros fatores negativos. Este artigo descreve uma metodologia para a previsão de enchentes no Rio Belém localizado na cidade de Curitiba utilizando um sistema de regras de classificação. O sistema de previsão foi desenvolvido usando um histórico de dados diários de chuva e vazão na bacia do Rio Belém durante o período de Abril de 1981 a Agosto de 2011. Este modelo de previsão poderia auxiliar na emissão de alertas da Defesa Civil e do Corpo de Bombeiros durante a ocorrência de fortes chuvas na cidade de Curitiba. O conhecimento prévio da ocorrência das enchentes, portanto, contribuiria para a gestão de ambientes urbanos no Brasil. As ferramentas computacionais descritas nesse artigo ainda poderiam também ser aplicadas em estudos de qualidade das águas dos rios urbanos ou outros estudos correlacionados em gestão urbana, tais como em transporte urbano ou infraestrutura.

**Palavras-Chave:** Regras de Classificação, Enchentes em Curitiba, Mineração de Dados.

## **1 - INTRODUÇÃO**

A expansão e urbanização das grandes cidades implicam na impermeabilização das áreas de drenagem das bacias hidrográficas (Fendrich, 1997, p.1). A remoção da cobertura vegetal e o desmatamento são ações modificadoras do ambiente natural que, em muitas circunstâncias, resultam simultaneamente na redução do tempo de concentração da bacia hidrográfica e no aumento do volume de escoamento superficial, que podem causar transbordamento dos cursos d'água (Pompêo, 2000, p. 15).

Segundo dados da UNESCO (2008, p.1), os transbordamentos causam aproximadamente metade das mortes e são responsáveis por um terço do total das perdas econômicas entre os diferentes tipos de desastres naturais que acontecem no mundo inteiro. Entre os anos de 1985 e 2005, os transbordamentos de rios causaram a morte de mais de 112 mil pessoas, atingiram mais de 354 milhões de pessoas e causaram prejuízos da ordem de 520 bilhões de euros ou 690 bilhões de dólares americanos (Bakker, 2006, p.18).

Em 2004, o Programa de Desenvolvimento das Nações Unidas estimou que em média quase 200 milhões de pessoas em mais de 90 países são expostas todo ano aos eventos catastróficos de transbordamentos (UNPD, 2004, p.3). A expectativa da vulnerabilidade aos transbordamentos é de aumentar no futuro devido às mudanças climáticas e ao sólido crescimento populacional bem como da expansão urbana (Kundzewicz, 2008, p.22).

Thiemig et al. (2011, p. 64) ressaltam que a previsão eficaz de transbordamento e a implementação de sistemas de alerta são fundamentais na gestão de risco de inundação, proporcionando tempo de preparação adicional antes de uma inundação. A previsão de eventos de transbordamento de rios nos centros urbanos, portanto, é de fundamental importância para a mobilização das instituições responsáveis pelo resgate de vítimas e pelas ações mitigadoras da catástrofe. O monitoramento de variáveis meteorológicas contribui para a construção de modelos matemáticos que permitam a previsão da ocorrência dos transbordamentos ou das enchentes nas áreas urbanas.

O trabalho de Duncan et al. (2011, p.1) descreve uma metodologia que utiliza Redes Neurais Artificiais (RNA's) para a previsão de enchentes urbanas em tempo real. Esses modelos computacionais (RNA's) foram treinados com dados de radar meteorológico, de pluviômetros, e das vazões de uma rede de bueiros em Keighley (West Yorkshire, UK).

Similarmente, Schellart et al. (2011, p.1) propuseram o uso de RNA's para a previsão de enchentes e testaram a acurácia dos modelos em três diferentes locais na Inglaterra: (1) em Pennine ao Norte da Inglaterra; (2) no bairro londrino Redbridge na bacia de Cranbrook, uma região muito atingida por enchentes; e (3) na área de Stockbridge em Keighley (West Yorkshire). Os dados de entrada do modelo também foram de pluviometria e do radar meteorológico.

Um modelo para a previsão de enchentes foi desenvolvido na China e considera as variáveis de declividade, uso do solo, taxas de rugosidade das redes de drenagem, taxa de infiltração, parâmetros de permeabilidade, entre outras. O modelo proposto por Fuqiang et al. (2002, p.6) integra as inúmeras variáveis em um Sistema de Informações Geográficas

(SIG's) que é acoplado a um modelo hidrológico e estima as condições de enchentes em Guangzhou.

Devido a ações antrópicas que provocam a degradação do solo, da fauna e da flora, as regiões próximas aos rios na cidade de Curitiba também sofrem frequentemente com as precipitações intensas, causando impactos socioeconômicos e ambientais. Essa pesquisa pretende desenvolver uma metodologia para auxiliar o sistema de alerta através da previsão de transbordamentos no Rio Belém, dentro do campus da PUC-PR, no bairro Prado Velho.

O método de previsão tem sido desenvolvido através da exploração de uma série cronológica de dados de enchentes (cotas e vazões) e de pluviometria, coletados e adquiridos nos órgãos estaduais. O objetivo principal da pesquisa científica é o de buscar padrões relevantes entre as variáveis explicativas e o de elaborar um modelo que possa auxiliar o sistema de alerta com um dia de antecedência.

O modelo consiste do uso das técnicas de Mineração de dados ou *Data Mining* que abrange um campo de pesquisa multidisciplinar envolvendo estatística, aprendizado de máquina, banco de dados, sistemas especialistas, técnicas de visualização de dados e computação de alto desempenho.

O presente capítulo introduziu a problemática de transbordamentos de rios em centros urbanos e descreveu alguns trabalhos relacionados à previsão de enchentes. No Capítulo 2 é descrita a metodologia utilizada para a construção de um modelo classificador baseado em regras de classificação para identificar a ocorrência de transbordamento no rio Belém. O Capítulo 3 discute os resultados obtidos, e no Capítulo 4 são elucidadas as conclusões do trabalho até o presente momento e propostas novas etapas para o projeto.

## **2 - METODOLOGIA**

Enchente ou transbordamento caracteriza-se por uma vazão relativamente grande de escoamento superficial (Villela e Matos, 1975) em que o nível da água ultrapassa o topo da cota da margem do rio ou canal. Para ser possível a previsão de enchentes, faz-se necessário conhecer a distribuição das vazões ao longo do tempo, desde o instante em que há o aumento do escoamento superficial produzido pela ocorrência de precipitações, até o fim da contribuição do escoamento superficial.

A avaliação de cenários de enchentes desta pesquisa está baseada no sistema de coleta de dados de chuva, vazões e permanência. Essas variáveis permitem a construção de um modelo que pode extrapolar condições mais críticas em função dos dados históricos.

O método utilizado na pesquisa baseia-se no uso das técnicas de regras de classificação. A classificação baseada em regras consiste basicamente em descobrir um pequeno conjunto de regras numa base de dados que forme um classificador acurado, LIU et al., (1998, p.1). Um exemplo de regra de classificação é ilustrado a seguir (SOUZA, 2004, p. 73):

Regra 20:

*SE* d\_1\_>\_65.4mm; *ENTÃO* Deslizamento de Encosta  
(25.276% 91.560% 320 293)

Interpretando-se a regra 20, o índice d\_1 refere-se ao total de chuva acumulado nas últimas 24 horas (1 dia), ou seja, se o valor de chuva acumulada no período de 24 horas ultrapassar 65,4 mm, então existe uma probabilidade de 91,56% para a ocorrência de deslizamento de encosta. O número 91,56% é também chamado confiança da regra. No banco de dados de 1266 registros, das 320 vezes em que o índice d\_1 ultrapassou 65,4mm (25,276% dos registros), 293 vezes ocorreram deslizamentos. O número 25,276% é também chamado de suporte da regra.

A particularidade desse conjunto de regras de classificação, portanto, é de possuir o atributo classe posicionado do lado direito da regra. Esta técnica de classificação é muito eficiente e foi usada neste trabalho para predição de transbordamentos.

A metodologia desenvolvida é composta de três etapas, e a primeira é a obtenção dos dados, que consiste na aquisição do máximo de informações referentes ao domínio estudado. A segunda etapa é a preparação dos dados, onde basicamente se constrói uma matriz com os registros descritos em cada linha da matriz e as variáveis relacionadas ao problema (chuva, vazões e permanência) formam as colunas. A terceira e última etapa é a modelagem, onde os dados preparados são processados para a construção do modelo de previsão (Souza e Ebecken, 2012).

## **2.1 - Obtenção dos Dados**

Na primeira etapa da pesquisa foram obtidos 11.080 registros diários de chuva. As informações foram retiradas da Estação Pluviométrica Curitiba Prado Velho, instalada no dia 25 de Março de 1981, na bacia do Belém. Foram também coletados dados da Estação Fluviográfica Curitiba Prado Velho, cuja área de drenagem contribuinte tem 42 km<sup>2</sup> (Fendrich, 2002a). Os dados foram obtidos através da plataforma HIDROWEB (2012) da

Agência Nacional de Águas (ANA), contendo informações diárias de chuva (em milímetros) e vazão média diária em (m<sup>3</sup>/s) entre 1981 e 2011.

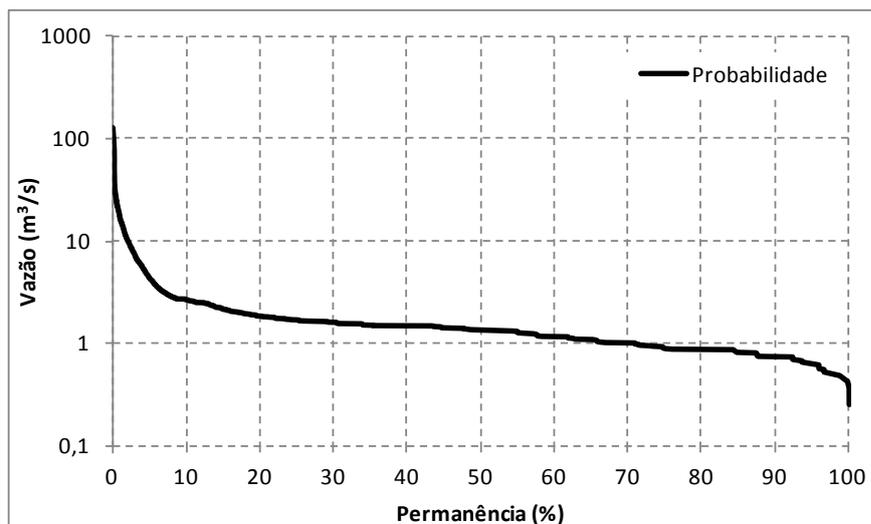
## 2.2 - Preparação dos Dados

A base de dados original é composta por dois atributos somente, ou seja, o acumulado de chuva diário (em mm) e a vazão média diária (m<sup>3</sup>/s). O objetivo do modelo computacional desenvolvido com as técnicas de regras de classificação é o de prever a vazão do próximo dia ( $t+1$ ) ou de “*amanhã*” em função dos valores de chuva e vazão do dia presente ( $t0$ ) ou de “*hoje*”. Então o modelo teria como variáveis de entrada valores de chuva e vazão ( $t0$ ) e a variável de saída seria a vazão ( $t+1$ ). Contudo, simulações preliminares usando somente essas duas variáveis de entrada do modelo não convergiram em resultados satisfatórios.

O estudo de Corani e Guariso (2005, p. 382) ressalta que os sistemas de previsão de enchentes devem considerar uma descrição do estado de saturação do solo na bacia hidrográfica, pois a taxa do escoamento superficial depende das condições de infiltração ao longo do tempo. Neste sentido, uma nova concepção do modelo foi considerada, através da inserção de novas variáveis com tempo de atraso além do dia presente ou “*hoje*” ( $t0$ ), ou seja, através da criação de novas colunas ou variáveis de dias anteriores ( $t-1$  ou “*ontem*”,  $t-2$  ou “*anteontem*”,  $t-3$  ou “*três dias atrás*” etc).

Para aumentar ainda a capacidade de previsão do modelo, foi inserido também um novo atributo a partir da curva de permanência de vazões. A curva de permanência é um gráfico bi-dimensional que associa cada valor de vazão (eixo das ordenadas na escala logarítmica) a uma probabilidade de ocorrência da mesma (eixo das abscissas), conforme ilustra a Figura 1.

Figura 1 - Gráfico da curva de permanência para o período de 1981 a 2011.



A curva de permanência foi construída da seguinte forma:

- 1) Organiza-se as vazões com valores variando do mínimo ao máximo;
- 2) Calcula-se a frequência de vezes que cada vazão pode ocorrer;
- 3) Calcula-se a probabilidade da vazão ocorrer.

Pode-se observar na Figura 1 que as grandes vazões têm baixa permanência ou baixa probabilidade de ocorrerem, pois dependem de eventos raros que são as chuvas muito intensas. A vazão de 10 m³/s possui permanência menor que 5%. Analogamente, percebe-se também que os baixos valores de vazão possuem alta probabilidade de ocorrência ou são mais permanentes, pois são eventos independentes de fortes chuvas. A vazão de 1 m³/s possui permanência da ordem de 70%.

Vale ressaltar que a permanência de uma vazão pode variar em função do tempo, e que o modelo, caso seja implementado, deverá atualizar a curva de permanência para cada novo registro de vazão. A permanência de uma determinada vazão também é uma variável que reflete o grau de saturação do solo e contribui para o desempenho do modelo.

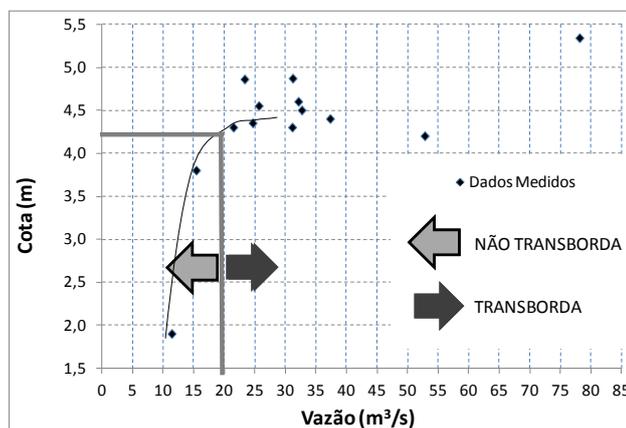
Para cada um dos três atributos (chuva, vazão e permanência) foram criadas 19 novas variáveis, considerando 20 dias anteriores ( $t_0$  ao  $t-19$ ) ao da previsão ( $t+1$ ). Também foram criadas mais três variáveis para a vazão e permanência considerando os valores mínimo, médio, e máximo para o período considerado de 20 dias, e duas variáveis para a chuva (médio e máximo), pois em uma janela de 20 dias possivelmente o valor mínimo de precipitação será igual a zero. A nova matriz possui 68 variáveis de entrada e uma variável de saída (vazão de amanhã ou  $t+1$ ).

Para a construção do modelo de classificação, todos os valores contínuos das variáveis foram transformados em valores categóricos segundo dois critérios. Nas variáveis

de entrada as classes foram subdivididas pelas métricas estatísticas da *mediana*, *1º*, e *3º quartil*. A transformação da variável de saída (vazão em  $t+1$ ) foi realizada através da consulta ao gráfico da curva chave. O gráfico da curva chave mostra as medições simultâneas de cota do nível d'água (eixo y) e vazão do rio (eixo x), realizadas para os níveis máximos observados do rio Belém entre os anos de 1981 a 1997 (Fendrich, 2002b), conforme ilustra a Figura 2.

Fendrich (2012, p. 9) relata que a partir da cota de 4,30m (aproximadamente na elevação ou altitude de 882,69m em relação ao nível do mar) o rio transborda. A vazão respectiva do “nível de transbordamento” (4,30m) observado na curva chave é de aproximadamente 20 m<sup>3</sup>/s, ou seja, todas as vazões que forem maiores que 20,0m<sup>3</sup>/s tenderão a causar enchente ou transbordamento da calha do rio.

Figura 2 - Gráfico da curva chave para verificação da vazão.



Dessa forma, a variável de saída recebeu o valor categórico igual a TRANSBORDA para valores contínuos de vazão superiores a 20,0 m<sup>3</sup>/s e igual a NÃO TRANSBORDA para valores contínuos iguais ou inferiores a 20,0 m<sup>3</sup>/s. Observou-se que somente 65 registros (do total de 11.080 registros da variável vazão) possuem valores de vazão maiores que 20,0m<sup>3</sup>/s. Portanto foi construída uma matriz com 130 registros de grandes vazões, 65 registros de transbordamento e 65 outros valores que não causaram transbordamento, mas que estão muito próximos a 20,0m<sup>3</sup>/s. Todos os registros estão, portanto, na zona confusa para o classificador que deve decidir a partir das condições de chuva, vazão e permanência nos últimos 20 dias (incluindo hoje), se o rio vai ou não transbordar.

### 2.3 - DESCRIÇÃO DO MODELO

Para a construção do modelo de classificação, foi usada a técnica de regras de classificação. Para a construção do classificador foi adotado um suporte mínimo de 5% e um valor mínimo de confiança de 85%. O classificador foi construído com 22 regras, conforme descrito na Tabela 1. As regras são compostas por vários itens das variáveis na forma discreta, ou seja, representando um intervalo de valores que pode variar do mínimo a mediana, e da mediana ao máximo para a maioria das variáveis de chuva; ou do mínimo ao 1º quartil, do 1º quartil à mediana, da mediana ao 3º quartil, e do 3º quartil ao máximo para as variáveis de permanência e vazão.

Portanto, para interpretar as regras da Tabela 1 é preciso identificar a seguinte codificação: Ch=Chuva; Vaz = Vazão; Perm=Permanência, min=mínima; med=média; tn=tempo em dias anterior ao dia da previsão, sendo  $t_0$  o dia da previsão. Os valores de chuva são dados em milímetros, então  $Ch_{t10}_{<3.5}$  = chuva de 11 dias anterior ao transbordamento menor que 3,5 mm; os valores de vazão em metros cúbicos por segundo ( $m^3/s$ ) e  $1_{<Vaz_{t2}_{<1.5}} = Y$  significa valores de vazão entre 1,0 e 1,5  $m^3/s$ ; e os valores de permanência em percentual. É importante observar que todas as 22 regras identificam a classe TRANSBORDA na parte consequente da regra. Portanto, em uma situação real de chuva intensa na cidade de Curitiba, o modelo primeiramente testaria se as condições sugeridas pelas regras são encontradas e se a condição for verdadeira o modelo atribuiria a classe TRANSBORDA. Posteriormente, caso não seja satisfeita nenhuma das 22 condições ou regras geradas, o modelo atribuiria a classe NÃO\_TRANSBORDA. Este tipo de informação seria de grande relevância para as autoridades, tais como Defesa Civil e Corpo de Bombeiros para as operações de resgate ou de mitigação dos efeitos catastróficos.

Tabela 1 - Regras de Classificação

Regras de Classificação	
<p>Rule 1:           <b>SE</b>     <math>Ch_{t10}_{&lt;3.5} = Y</math>                                    <math>Vaz_{t19}_{&lt;1} = Y</math>                                    <math>0.8_{&lt;Ch_{t1}_{&lt;7}} = Y</math>                    <b>ENTÃO -&gt;</b> Class = TRANSBORDA                    (9.231% 100.000% 12 12 9.231%)</p>	<p>Rule 12:           <b>SE</b>     <math>1_{&lt;Vaz_{t2}_{&lt;1.5}} = Y</math>                                    <math>Ch_{t17}_{&gt;6} = Y</math>                    <b>ENTÃO -&gt;</b>Class=TRANSBORDA                    (6.154% 100.000% 8 8 6.154%)</p>
<p>Rule 2:           <b>SE</b>     <math>Ch_{t15}_{&lt;5.5} = Y</math>                                    <math>Vaz_{t19}_{&lt;1} = Y</math>                                    <math>0.8_{&lt;Ch_{t1}_{&lt;7}} = Y</math>                    <b>ENTÃO -&gt;</b> Class = TRANSBORDA                    (8.462% 100.000% 11 11 8.462%)</p>	<p>Rule 13:           <b>SE</b>     Perm_min_{&lt;1} = Y                                    <math>3_{&lt;Ch_{med}_{&lt;4.5}} = Y</math>                    <b>ENTÃO &gt;</b>Class = TRANSBORDA                    (6.154% 100.000% 8 8 6.154%)</p>
<p>Rule 3:           <b>SE</b>     <math>Ch_{t16}_{&lt;3.5} = Y</math>                                    <math>1_{&lt;Vaz_{t16}_{&lt;1.5}} = Y</math></p>	<p>Rule 14:           <b>SE</b>     <math>Ch_{t14}_{&lt;4} = Y</math>                                    Perm_{t13}_{&lt;25} = Y</p>

<p style="text-align: center;">Ch_t8_&gt;_2_5 = Y</p> <p><b>ENTÃO</b>-&gt; Class = TRANSBORDA (8.462% 100.000% 11 11 8.462%)</p>	<p style="text-align: center;">Ch_t2_&gt;_3_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.154% 100.000% 8 8 6.154%)</p>
<p>Rule 4:       SE       35_&lt;_Perm_t1_&lt;_70 = Y</p> <p style="text-align: center;">Ch_t8_&gt;_2_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class = TRANSBORDA (7.692% 100.000% 10 10 7.692%)</p>	<p>Rule 15:     SE       Ch_t12_&lt;_7_5 = Y</p> <p style="text-align: center;">45_&lt;_Perm_t7_&lt;_70 = Y</p> <p style="text-align: center;">Ch_t13_&gt;_3_5 = Y</p> <p><b>ENTÃO</b>-&gt; Class = TRANSBORDA (6.154% 100.000% 8 8 6.154%)</p>
<p>Rule 5:       SE       50_&lt;_Perm_t14_&lt;_80 = Y</p> <p style="text-align: center;">45_&lt;_Perm_t7_&lt;_70 = Y</p> <p><b>ENTÃO</b>       Class =TRANSBORDA (6.923% 100.000% 9 9 6.923%)</p>	<p>Rule 16:     SE       Ch_t9_&gt;_3_5 = Y</p> <p style="text-align: center;">1_5_&lt;_Vaz_t10_&lt;_2 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (5.385% 100.000% 7 7 5.385%)</p>
<p>Rule 6:       SE       Vaz_t14_&lt;_1 = Y</p> <p style="text-align: center;">1_&lt;_Vaz_t11_&lt;_1_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class = TRANSBORDA (6.923% 100.000% 9 9 6.923%)</p>	<p>Rule 17:     SE       1_5_&lt;_Vaz_t4_&lt;_2 = Y</p> <p style="text-align: center;">1_&lt;_Vaz_t15_&lt;_1_5 = Y</p> <p><b>ENTÃO</b>-&gt;Class = TRANSBORDA (5.385% 100.000% 7 7 5.385%)</p>
<p>Rule 7:       SE       Perm_t13_&lt;_25 = Y</p> <p style="text-align: center;">Ch_t12_&gt;_7_5 = Y</p> <p style="text-align: center;">Ch_t10_&gt;_3_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class = TRANSBORDA (6.923% 100.000% 9 9 6.923%)</p>	<p>Rule 18:     SE       Ch_t3_&lt;_4 = Y</p> <p style="text-align: center;">1_5_&lt;_Vaz_t9_&lt;_2 = Y</p> <p style="text-align: center;">Ch_t11_&gt;_5 = Y</p> <p><b>ENTÃO</b> -&gt;Class= TRANSBORDA (5.385% 100.000% 7 7 5.385%)</p>
<p>Rule 8:       SE       Ch_t7_&lt;_2_5 = Y</p> <p style="text-align: center;">Perm_t13_&lt;_25 = Y</p> <p style="text-align: center;">Ch_t10_&gt;_3_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.923% 100.000% 9 9 6.923%)</p>	<p>Rule 19:     SE       Ch_t15_&lt;_5_5 = Y</p> <p style="text-align: center;">1_&lt;_Vaz_t10_&lt;_1_5 = Y</p> <p style="text-align: center;">Ch_t5_&gt;_4_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (8.462% 90.909% 11 10 7.692%)</p>
<p>Rule 9:       SE       Ch_t7_&lt;_2_5 = Y</p> <p style="text-align: center;">Vaz_t19_&lt;_1 = Y</p> <p style="text-align: center;">Ch_t12_&gt;_7_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.923% 100.000% 9 9 6.923%)</p>	<p>Rule 20:     SE       Ch_t17_&gt;_6 = Y</p> <p style="text-align: center;">20_&lt;_Ch_max_&lt;_35 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.923% 88.889% 9 8 6.154%)</p>
<p>Rule 10:     SE       Ch_max_&gt;_45 = Y</p> <p style="text-align: center;">Ch_t8_&gt;_2_5 = Y</p> <p><b>ENTÃO</b>-&gt; Class=TRANSBORDA (6.154% 100.000% 8 8 6.154%)</p>	<p>Rule 21:     SE       1_5_&lt;_Vaz_t17_&lt;_2 = Y</p> <p style="text-align: center;">25_&lt;_Perm_t10_&lt;_45 = Y</p> <p><b>ENTÃO</b> -&gt;Class =TRANSBORDA (6.923% 88.889% 9 8 6.154%)</p>
<p>Rule 11:     SE       1_5_&lt;_Vaz_t6_&lt;_2 = Y</p> <p style="text-align: center;">Ch_t10_&gt;_3_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.154% 100.000% 8 8 6.154%)</p>	<p>Rule 22:     SE       Perm_t3_&gt;_70 = Y</p> <p style="text-align: center;">Ch_t20_&gt;_2_5 = Y</p> <p><b>ENTÃO</b> -&gt; Class=TRANSBORDA (6.154% 87.500% 8 7 5.385%)</p>

### 3 -RESULTADOS E DISCUSSÃO

O resultado da validação de um modelo de classificação é usualmente apresentado na forma de uma matriz de confusão, conforme ilustra a Tabela 2. Os componentes da diagonal principal da matriz de confusão fornecem o número de exemplos corretamente classificados para cada classe correspondente (TRANSBORDA / NÃO TRANSBORDA). Por exemplo, para a classe TRANSBORDA, foram corretamente classificados 63 exemplos, e para a classe NÃO TRANSBORDA foram corretamente classificados 61 exemplos. Os valores da outra diagonal apresenta o número de registros erroneamente classificados pelo modelo.

Tabela 2 - Matriz de confusão do classificador na validação do modelo.

	TRANSBORDA	NÃO TRANSBORDA
TRANSBORDA	63	2
NÃO TRANSBORDA	4	61

A taxa de classificação correta (TCC) desse modelo é a soma dos valores classificados corretamente ( $63+61=124$ ) divididos pelo número total de estimativas ( $63+61+4+2=130$ ). Portanto, o modelo possui uma TCC igual a 95,4%, que pode ser considerada alta devida à complexidade do estudo de vazões de enchentes.

É importante ressaltar que a acurácia do modelo ainda poderia ser melhorada pela inserção de novos dados (variáveis com as características do uso do solo da bacia hidrográfica do Rio Belém) e/ou por artifícios matemáticos na construção dos modelos, tais como as técnicas de *Bagging* (Han e Kamber, 2001), *Ponderação* (Duda, 2001), e *Boosting* (Ting e Zheng, 1998).

#### 4 - CONCLUSÃO E TRABALHOS FUTUROS

A previsão de vazões de enchentes é uma tarefa complexa que envolve um grande número de taxonomias (climatológicas, atmosféricas, uso e ocupação do solo, topografia, densidade da rede de drenagem etc). O modelo de previsão proposto neste trabalho foi construído com a utilização de somente dois atributos monitorados (chuva e vazão) e apresentou resultado satisfatório com uma TCC igual a 95,4%.

Contudo, percebe-se que as regras geradas usaram um grande número de variáveis para a construção do classificador, o que indica um alto grau de complexidade do fenômeno estudado. A inserção de mais atributos deveria ser considerada na continuação deste trabalho, tais como as variáveis relacionadas ao uso e ocupação do solo, mapas

litológicos, pedológicos etc. O trabalho de Fendrich (2012, p. 7), por exemplo, descreve a evolução da impermeabilização dos solos da bacia hidrográfica do Rio Belém.

Os resultados deste trabalho potencializam o uso dessas ferramentas para a previsão de enchentes no rio Belém e em outras bacias hidrográficas na cidade de Curitiba.

Deve-se ressaltar também que todo o conhecimento adquirido na metodologia proposta foi obtido de dados monitorados; resta ainda a consideração do conhecimento dos especialistas que em muito pode acrescentar em benefício da qualidade das predições.

O conhecimento prévio da ocorrência das enchentes, portanto, contribuiria para a gestão de ambientes urbanos no Brasil. As ferramentas computacionais descritas nesse artigo não se limitam ao estudo de enchentes, e poderiam também ser aplicadas em estudos de qualidade das águas dos rios urbanos e sua relação espaço temporal com as fontes de poluição e a saúde da população, entre outros estudos correlacionados em transporte urbano ou infraestrutura.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

BAKKER, M. H. N., 2006. Transboundary river floods: vulnerability of continents, international river basins and countries. Corvallis, OR: Oregon State University.

CORANI, G., GUARISO, G., 2005. Coupling Fuzzy Modeling and Neural Networks for River Flood Prediction. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and reviews, vol. 35, no. 3, August.

DUDA, R. O., HART, P. E., STORK, D. G., 2001, Pattern Classification, Wiley Interscience, Second Edition.

DUNCAN, A. P. et al., 2011. Urban flood prediction in real-time from weather radar and rainfall data using artificial neural networks. Weather Radar and Hydrology. (Proceedings of a symposium held in Exeter, UK, April) (IAHS Publ. 351, 2012), 568-573.

FENDRICH, R., 2012. Relação de “Permeabilidade Artificial” dos solos “Rp” de uma bacia hidrográfica urbana: detenção distribuída das águas pluviais. Revista Técnica do IEP. Ano II, no. 4, Jan/Jun, PP. 5-14.

FENDRICH, R., 2002a. Diagnóstico dos recursos hídricos da Bacia Hidrográfica do Rio Belém.

FENDRICH, R., 1997. Bacia Hidrográfica Urbana do Rio Belém. INTEC/PUC-PR, Curitiba.

FUQIANG, T., Heping, H., Jianzhong, G., 2002. An Urban Flood Simulation Model based on Geographic Information System. International Commission on Irrigation and Drainage. Eighteenth Congress. Montreal. Q. 51. P. 3.05.

HAN, J., KAMBER, H., 2001, Data Mining - Concepts and Techniques - Chapters 6, 7 and 8.

HIDROWEB (2012). Plataforma de Dados Hidrológicos. Agência Nacional de Águas (ANA). Disponível em <http://hidroweb.ana.gov.br/>

KUNDZEWICZ, Z.W., 2008. Flood risk and vulnerability in the changing climate. Land Reclamation, 39, 21-31.

LIU, B., HSU, W., CHEN, S., MA, Y., 1998. Integrating Classification and Association Rule Mining, KDD-98, August, New York.

POMPÊO, C. A. 2000. Drenagem Urbana Sustentável. RBRH - Revista Brasileira de Recursos Hídricos, Volume 5, n.1 Jan/Mar, 15-23.

SCHELLART, A., 2011. Urban pluvial flood modelling with real time rainfall information - UK case studies. 12<sup>nd</sup> International Conference on Urban Drainage, Porto Alegre/Brazil, 10-15, September 1.

SOUZA, FT., EBECKEN, N., 2012. A Data Based Model to Predict Landslide Induced by Rainfall in Rio de Janeiro City. Geotechnical and Geological Engineering. 30, 85-94.

SOUZA, F. T. (2004) Predição de Escorregamentos das Encostas do Município do Rio de Janeiro através de Técnicas de Mineração de Dados. Anexo I. Tese de Doutorado da COPPE/UFRJ.

TING, K. M., ZHENG, Z., 1998, "Boosting Trees for Cost-Sensitive Classifications", Proceedings of the Tenth European Conference on Machine Learning, LNAI-1398, Berlin: Springer-Verlag, pp. 190-195.

THIEMIG, V., Roo, A., Gadain, H., 2011. Current status on flood forecasting and early warning in Africa, International Journal of River Basin Management, 9:1, 63-78.

UNESCO - United Nations Educational, Scientific and Cultural Organization, 2008. Floods. Available from: [http://portal.unesco.org/science/en/ev.php-URL\\_ID=6082&URL\\_DO=DO\\_PRINTPAGE&URL\\_SECTION=201.html](http://portal.unesco.org/science/en/ev.php-URL_ID=6082&URL_DO=DO_PRINTPAGE&URL_SECTION=201.html) [Accessed 10May 2010].

UNPD - United Nations Development Programme, 2004. A global report: reducing disaster risk, a challenge for development. New York: United Nations Development Programme.

VILLELA, S. M. e MATOS, A., 1975. Hidrologia Aplicada. Editora McGraw-Hill, São Paulo.