

# Complexidade, leis de escala urbana e perdas na distribuição de água potável: análise da rede de cidades do sul do Brasil

*Complexity, urban scaling laws and losses in the distribution of potable water: an analysis in southern Brazilian cities*

Júlio Celso Borello Vargas   
Bárbara Brzezinski Azevedo 

## Resumo

O fornecimento de água potável às populações urbanas sofre com perdas derivadas de vazamentos e uso não autorizado. Esse subproduto da urbanização é notavelmente maior em cidades grandes, o que sugere a presença de mecanismos sociotécnicos complexos cujo comportamento se torna aparentemente imprevisível à medida que aumenta a população. Com o objetivo de verificar essa hipótese e encontrar regularidades quantitativas entre o tamanho das cidades, o tamanho dos sistemas de distribuição e as perdas, analisamos a rede de municípios do sul do Brasil a partir de pressupostos e ferramentas da “Nova Ciência das Cidades”. Encontramos distribuições estatísticas hierarquizadas típicas de sistemas complexos naturais, bem como regimes de escala não lineares entre a população e indicadores selecionados: potência com expoente  $< 1$  para o tamanho da rede (extensão, nº de ligações, volume produzido) e  $> 1$  para o volume de perdas. As perdas apresentaram volume per capita com “retornos crescentes” comparáveis a produtos típicos da atividade social como diplomas universitários ou depósitos bancários. Isso sugere certo grau de previsibilidade em nível regional, oferecendo uma visão complementar do problema e uma nova contribuição às políticas de fornecimento de água nas cidades brasileiras.

**Palavras-chave:** Complexidade. Sistemas sociotécnicos complexos. Sistemas de distribuição de água. Perdas no abastecimento de água. Leis de escala urbana.

## Abstract

*The drinking water supply to urban populations suffers losses caused by leaks, theft and unauthorised use. This urbanisation by-product is remarkably greater in large cities, suggesting complex socio-technical mechanisms that apparently behave unpredictably as populations and social activity grow. To verify this hypothesis and to seek for quantitative regularities between city size, the size of the utilities and the volume of losses, we analysed a network of municipalities in southern Brazil with tools from the "New Science of Cities". We found highly hierarchical distributions, typical of natural complex systems, as well as non-linear scaling regimes between the cities' population and selected indicators: power law exponent  $< 1$  for the water distribution variables (extension, connected households and volume produced) and  $> 1$  for the volume of losses. Water losses consistently presented much higher per capita volumes as population grows, with "increasing returns" comparable to social products such as university degrees or bank deposits. This suggests a certain degree of predictability at the regional level, offering a complementary view of the problem and a new contribution to water supply policies in Brazilian cities.*

<sup>1</sup>Júlio Celso Borello Vargas

<sup>1</sup>Universidade Federal do Rio Grande do Sul  
Porto Alegre - RS - Brasil

<sup>2</sup>Bárbara Brzezinski Azevedo

<sup>2</sup>Departamento Municipal de Água e Esgoto  
Porto Alegre - RS - Brasil

Recebido em 18/07/20  
Aceito em 18/01/21

**Keywords:** Complexity. Complex socio-technical systems. Water distribution systems. Water losses, Urban scaling laws.

## Introdução

Os sistemas de distribuição de água (SDA) apresentam altos níveis de perdas, especialmente em países em desenvolvimento (TROJAN; MORAIS, 2015). De fato, os benefícios obtidos com a redução de perdas em SDA podem superar os obtidos com a ampliação do fornecimento de água tratada (AL-OMARI, 2013). Kingdom, Liemberger e Marin (2006) estimaram que reduzir 50% das perdas mundiais em SDA seria suficiente para fornecer água para mais de 100 milhões de pessoas por um ano.

No Brasil, país que conta com 12% dos recursos hídricos do planeta, o volume per capita de água tratada oferecida é 19 vezes superior ao mínimo estabelecido pela Organização das Nações Unidas, que anualmente é de 1.700 metros cúbicos por segundo por habitante. No entanto, para garantir tal consumo, os sistemas sofrem perdas que, na média nacional, alcançam 38,3% (OLIVEIRA *et al.*, 2018). Esse é um tema de alta relevância, frente aos cenários de escassez hídrica, altos custos de energia elétrica (OLIVEIRA *et al.*, 2018) e transição das políticas de saneamento no país.

Perdas em SDA são classificadas como *aparentes* ou *reais*. As perdas *reais* expressam a quantidade perdida devido aos vazamentos (RÍOS *et al.*, 2014). Globalmente, 48 bilhões de m<sup>3</sup>/ano são distribuídos e não faturados, sendo que as perdas reais representam 66% desse montante (KINGDOM; LIEMBERGER; MARIN, 2006; MAMADE *et al.*, 2015). Já as perdas *aparentes* expressam a quantidade de água não faturada devido ao consumo não autorizado e às imprecisões do sistema de medição, representando parte da água que realmente é consumida, mas não é paga.

As perdas reais são basicamente produto das características físicas locais do SDA, como sua composição material, a vazão de água, o tipo do solo e o tráfego na superfície (GOMES; MARQUES; SOUSA, 2013). Trabalhos como o de Skipworth *et al.* (1999) examinaram variáveis determinantes das perdas de água na Inglaterra, concluindo que a pressão do sistema, a idade e o tipo de rede, o comprimento da rede, a condição do solo, a carga de tráfego e a densidade das ligações de água são as mais importantes.

As perdas aparentes, por outro lado, têm maior relação com características gerais das cidades, como o crescimento populacional e a composição demográfica. González-Gómez *et al.* (2012), por exemplo, verificaram que a população atendida, a taxa de crescimento populacional e a densidade populacional influenciam nas perdas de água, enquanto Van den Berg (2015) concluiu que as perdas de água em 63 países em desenvolvimento estão associadas com a densidade populacional e o tamanho da rede de distribuição.

O uso clandestino de água é uma manifestação tipicamente social, derivada de ações individuais que buscam atender necessidades básicas, muitas vezes não supridas pelo poder público nas cidades brasileiras. Seus padrões quantitativos se comportam de maneira não linear em relação à população, variando de forma aparentemente “caótica” dentro de um sistema de cidades e tornando sua previsão, em princípio, impossível.

Esse comportamento “estranho” é, no entanto, uma característica típica de sistemas complexos auto-organizados, cujas interações locais entre os inúmeros componentes fazem emergir padrões globais que evoluem no tempo e no espaço. Essa ideia geral já é relativamente antiga, e vem sendo aplicada a inúmeras áreas, das ciências naturais à economia, da física à comunicação (BAK; TANG; WIESENFELD, 1987; BATTY, 2013; HERRERA; ABRAHAM; STOIANOV, 2015; SACCHET *et al.*, 2011).

A área do planejamento urbano vem adotando essa perspectiva para compreender os mecanismos de auto-organização presentes no processo de urbanização, anteriormente tratado como produto exclusivo do planejamento centralizado (BATTY, 2013). Chamada de “Nova Ciência das Cidades”, essa abordagem tira partido da abundância de dados e da crescente capacidade computacional, fundamentando estudos que procuram identificar padrões quantitativos e regularidades na relação entre o tamanho das cidades e das infraestruturas (BETTENCOURT, 2007; MEIRELLES *et al.*, 2018).

Essas regularidades de escala se mostraram estáveis em sistemas urbanos de diferentes países e culturas e, portanto, passaram a ser conhecidas como “leis de escala urbana” - *urban scaling laws* (GLERIA; MATSUSHITA; SILVA, 2004; NEWMAN, 2005; IGNAZZI, 2015; WEST, 2017; MEIRELLES *et al.*, 2018). Elas não negam as particularidades das cidades, sua identidade e atributos únicos, mas, sim, evidenciam suas notáveis similaridades, *apesar* das diferenças (STRANO; SOOD, 2016).

Este artigo analisa o sistema urbano do Sul do Brasil através de um estudo estatístico com 968 municípios da Região, em busca de regularidades que descrevam as relações entre as cidades e os sistemas de distribuição de água, incluindo as perdas. Conhecendo os padrões populacionais e o ritmo da urbanização de uma região, é possível inferir os mecanismos sociais que estão na origem das perdas na distribuição, complementando a abordagem tradicional e contribuindo com o aprimoramento do uso da água nas cidades brasileiras.

## Sistemas complexos

“Complexidade” possui diversas definições na literatura. Muitos autores enfatizam que os sistemas complexos são *abertos* e em permanente crescimento/evolução e, portanto, não convergem para equilíbrios estáticos e se mantêm em estados de estabilidade crítica (BAK; TANG; WIESENFELD, 1988). Crawford (2016), no mesmo sentido, define um sistema complexo como aquele com elementos heterogêneos que possuem um certo grau de interconexão que lhes permite adaptar-se às mudanças ao longo do tempo. Ciências como a física propõem compreender os sistemas complexos através de modelos microscópicos com poucos elementos e inter-relações baseadas em regras locais, observando sua evolução e a eventual emergência de padrões que, não raro, levam à identificação de leis gerais (BARTHELEMY, 2019a).

Portanto, este estudo trata os SDA não como meros sistemas físicos localizados, estáticos e perfeitamente controlados, mas, sim, como *sistemas sociotécnicos complexos* (SSC). SSC são sistemas baseados na interação dinâmica entre elementos materiais/tecnológicos, sociais e econômicos, sujeitos à permanente incerteza do ambiente externo ao qual estão abertos (DEY; LEE, 2017).

Considerando os SDA como sistemas sociotécnicos complexos, Azevedo e Saurin (2018) realizaram uma revisão sistemática da literatura e identificaram os principais atributos que caracterizam sua complexidade, sugerindo que eles contribuem para as perdas na distribuição:

- (a) *grande número e diversidade de elementos* (CILLIERS, 1995; CARAYON, 2006);
- (b) *interconectividade* (ELMARAGHY *et al.*, 2014; HABERFELLNER *et al.*, 2012; WECK; ROOS; MAGEE, 2013);
- (c) *interações não lineares* (PERROW, 1984; SNOWDEN *et al.*, 2007);
- (d) *círculos de retroalimentação – feedback loops* – (CILLIERS, 1995; ERDI, 2008; PERROW, 1984);
- (e) *propriedades emergentes* (ERDI, 2008);
- (f) *abertura ao ambiente* (CILLIERS, 1995); e
- (g) *capacidade adaptativa* (KURTZ; SNOWDEN, 2003).

## Cidades, hierarquia urbana e distribuições assimétricas

Os atributos “grande número”, “diversidade de elementos” e “interações não lineares” dos sistemas complexos podem ser representados estatisticamente por suas distribuições, do tipo “cauda pesada”. Elas são bastante diferentes da distribuição normal cujos valores médios representam fielmente o conjunto e sugerem relações lineares entre os elementos. Ao contrário, essas distribuições são altamente assimétricas, sendo genericamente chamadas de *powerlaws* – “lei de potência” – que incluem as conhecidas “Distribuição de Pareto” e a “Lei de Zipf” (MONTEBRUNO *et al.*, 2019).

Uma distribuição potência satisfaz (Equação 1):

$$P [X \geq x] \sim x^{-b} \quad \text{Eq. 1}$$

Onde a probabilidade da variável  $X$  ser maior ou igual a um valor  $x$  é equivalente ao valor elevado ao expoente  $b$ . Portanto (Equação 2),

$$\log P [X \geq x] = b \cdot \log x \quad \text{Eq. 2}$$

Ou seja, o logaritmo da probabilidade complementar cumulativa é linear com inclinação  $b$ .

O gráfico CCDF, portanto, deve mostrar pontos aproximadamente organizados em uma reta. Nele podem ser plotadas as linhas de ajuste, estimadas por máxima verossimilhança com o auxílio da técnica de “*bootstrapping*”, que reamostra iterativamente os dados a partir do conjunto original, recalculando os parâmetros a cada iteração e selecionando aqueles que minimizam a estatística *Kolmogorov-Smirnov* (CLAUSET; COSMA; NEWMAN, 2009).

Elas podem ser encontradas em diferentes sistemas baseados em fluxos e interações, desde os biológicos e geográficos – tamanho de estruturas pulmonares e de bacias hidrográficas, por exemplo – até os socioeconômicos como a distribuição de renda em um país, a popularidade de páginas *web* ou a população das cidades em uma região (NEWMAN, 2005). Sua “assinatura” é a mesma: um grande espaço combinatório onde os elementos se organizam e interagem através de diversas escalas, a imensa maioria deles muito pequenos e alguns poucos muito grandes. A população das cidades brasileiras, por exemplo,

varia de cerca de 800 até mais de 11.000.000 de habitantes, cobrindo seis ordens de magnitude (MEIRELLES *et al.*, 2018).

Esse tipo de distribuição implica que a posição de cada elemento na ordem por tamanho do sistema (*rank*) tenha uma relação não linear com sua frequência de aparição (*frequency*) seguindo uma função potência (SALAT; BOURDIC, 2011a). Quanto maior o elemento, menor a frequência de ocorrência. Os gráficos *rank/frequency* em escala logarítmica de uma função potência bem ajustada mostra uma linha reta, cuja inclinação reflete o expoente da função.

Mais do que curiosidade matemática, a profusão de trabalhos sobre as *powerlaws* justifica-se pela possibilidade de acessar a não linearidade de um sistema de forma simples, por meio de uma função com apenas um parâmetro. Se conhecido e bem ajustado, ele permite descrever a organização interna dos sistemas e inferir, baseado em analogias físicas e biológicas, os mecanismos que geram tal distribuição. No campo urbanístico, por exemplo, uma explicação bastante aceita é o “Processo de Yule”, que resumidamente propõe que, a um sistema, novos objetos (cidades) são ocasionalmente adicionados e, enquanto isso, elementos (habitantes) vão sendo constantemente acrescidos a elas. A probabilidade de a cidade ganhar novos habitantes é proporcional à população já existente, ou seja, existe um mecanismo aparentemente natural de “vinculação preferencial” – *preferential attachment* – que só faz aumentar a concentração e a desigualdade do sistema (BARABÁSI; ALBERT; JEONG, 1999).

## Regimes de escala nas cidades

A população das cidades e as redes de infraestrutura também costumam escalar conjuntamente com base em uma função potência. O pressuposto é que a população é um bom determinante das características das cidades, ou seja, cidades de mesma população devem apresentar semelhantes propriedades (BARTHELEMY, 2019b).

A função de escala pode ser descrita pela Equação 3:

$$Y = aX^b \quad \text{Eq. 3}$$

Onde:

$Y$  é a variável, quantidade em análise;

$X$  é a variável de tamanho da cidade (geralmente a população);

$a$  é uma constante; e

$b$  é o expoente, que em geral é positivo, pois quando uma grandeza cresce, a outra inevitavelmente também.

A Equação 1 também implica que a quantidade per capita se comporta como  $Y/X \sim X^{b-1}$  e, no caso  $b=1$ , a relação se torna linear e a razão  $Y/X$  passa a ser uma constante de proporcionalidade que não depende do tamanho da cidade. Nos casos em que  $b \neq 1$  a quantidade per capita depende de  $X$  e, portanto, essa relação varia de acordo com o tamanho da cidade (BARTHELEMY, 2019b). Esse comportamento da função dependente do expoente indica o “ritmo” da urbanização e é o que se conhece por “regime de escala” (BETTENCOURT *et al.*, 2007).

Bettencourt *et al.* (2007) estudaram regimes de escala em diversos países e propuseram três classes de variáveis de acordo com  $b$ : o suprimento de água medido em “habitantes com ligação ativa” foi enquadrado como “serviço básico individualizado” (*individual basic services*, no original), escalando de forma linear com a população. Já a extensão das redes de água foi considerada variável de “infraestrutura material”, apresentando  $b < 1$  (relação sublinear). Por outro lado, indicadores de “atividade social” como número de diplomas universitários, depósitos bancários, casos de HIV, homicídios ou patentes registradas apresentaram  $b > 1$  (relação superlinear), um fenômeno sem paralelo nos sistemas biológicos (BETTENCOURT *et al.*, 2007).

Infraestruturas materiais como os SDA são resultado de planejamento, investimentos e obras centralizadas que buscam atender a demandas conhecidas e, portanto, sua implantação é mais “lenta” em relação ao crescimento da população, expressando as economias de escala. Já os produtos socioeconômicos tendem a se manifestar de forma mais “rápida”, pois emergem da miríade de interações entre os agentes que levam a ganhos crescentes com o aumento da população, sendo essa “produtividade” um vetor do crescimento urbano (WEST, 2017). Cidades maiores produzem mais per capita, atraindo desproporcionalmente mais habitantes e recursos e disparando um processo de retroalimentação da hierarquia conhecido por *rich-gets-richer* (*quem já é rico fica mais rico*, em livre tradução).

Em sistemas naturais, a prevalência da ordem hierárquica emergente expressa um melhor aproveitamento dos fluxos, levando, portanto, a uma maior eficiência (SALAT; BOURDIC, 2011b). Porém, os ganhos crescentes também se verificam para produtos negativos, que aumentam em sincronia com o aumento da riqueza na cidade, como, por exemplo, criminalidade, poluição e, possivelmente, as perdas no abastecimento de água. Bettencourt *et al.* (2007) analisaram o comportamento de uma variável análoga: as “perdas em sistemas de distribuição de energia elétrica”, cujo regime de escala com a população encontrado foi superlinear ( $b=1,11$ ), ajudando a suportar essa hipótese.

## Método

Um estudo estatístico com dados abertos do Sistema Nacional de Informações em Saneamento (SISTEMA..., 2019) foi realizado com foco no sistema urbano da Região Sul do Brasil e seus 1.188 municípios.

O SNIS contém informações preenchidas pelos prestadores de serviço, disponíveis no sítio [www.snis.gov.br](http://www.snis.gov.br). O banco de dados possui 249 campos, dos quais apenas aqueles relativos ao abastecimento de água foram utilizados, com foco nas principais medidas de tamanho do das, incluindo as ligações ativas de água, extensão da rede de distribuição de água e volume de água produzido. Elas foram enquadradas nas categorias de Bettencourt *et al.* (2007): variáveis relativas ao fornecimento de serviços individualizados e variáveis de infraestrutura material, conforme o Quadro 1.

Utilizamos a população *urbana* para contornar o viés derivado das diferentes definições de cidade, pois os limites municipais muitas vezes são diferentes do perímetro urbano legal e da área efetivamente urbanizada, à qual é majoritariamente provido o abastecimento. Os dados utilizados foram de 2017, os mais recentes disponíveis quando da consulta. Excluímos as observações incompletas e registros com valores 0 (zero) para qualquer variável, restando na amostra final 968 cidades.

Analizamos, comparamos e visualizamos as distribuições “cauda pesada” com os estimadores propostos por Gillespie (2015) por meio do código *powerLaw*, no software “R” 3.6.1 (<http://www.r-project.org/>), considerando a Equação 1. Analizamos os regimes de escala com regressões por mínimos quadrados ordinários (MQO) para estimar os expoentes entre as quantidades linearizadas através de seus logs, considerando as Equações 2 e 3. Comparamos e visualizamos os resultados também no software R, por meio de suas bibliotecas nativas.

## Resultados

A Região Sul do Brasil possui 29.754.036 habitantes e densidade populacional de 51,9 habitantes/km, com Índices de Desenvolvimento Humano (IDHM) que colocam os três Estados – Rio Grande do Sul, Santa Catarina e Paraná – entre os seis melhores do país (INSTITUTO..., 2019). Cerca de 98% da população urbana está abastecida por rede de distribuição de água – o maior índice do país – e seus operadores prestam serviços a cerca de 80% da população total. O índice de perdas na distribuição em relação ao volume de água produzido foi de 34,06%, abaixo da média brasileira, porém mais alto que nas regiões Sudeste e Centro-Oeste (SISTEMA..., 2019). A distribuição espacial e a dimensão das cidades podem ser visualizadas na Figura 1, de forma simplificada dada a escala territorial.

## Análise das distribuições

A Tabela 1 apresenta as estatísticas descritivas dos cinco indicadores.

É evidente a enorme assimetria das variáveis, com grande amplitude de valores através de muitas ordens de magnitude, desde cidades com 262 habitantes urbanos, 132 ligações ativas de água e 6.000 litros de perda por ano (Barra Bonita, SC), até metrópoles como Curitiba, que tem população acima de 1,8 milhão de habitantes, quantidade de ligações próximas a 500.000 e perdas da ordem de 70 milhões de metros cúbicos por ano.

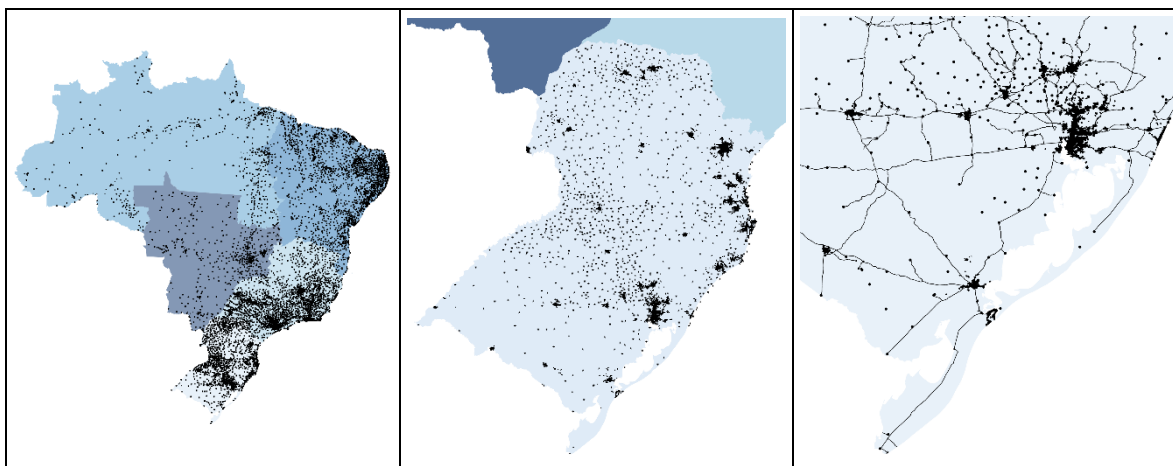
Os histogramas da coluna da esquerda da Figura 2 são similares, expressando visualmente a assimetria comum a todas as variáveis e antecipando a relação entre os indicadores. As três linhas verticais assinalam os percentis 50°, 90° e 99°. Plotamos histogramas de densidade para melhorar a visualização e evitar as inconsistências na definição dos blocos (*bins*), mas pode-se considerar que eles mostram a quantidade de ocorrências em *Y* e os tamanhos em *X*.

Buscamos os parâmetros e verificamos o ajuste da lei de potência e também da distribuição *lognormal*, que possui semelhantes propriedades de assimetria, médias pequenas, alta variância e valores positivos, mas que se origina de processos matemáticos diferentes e requer dois ou três parâmetros para ser descrita (MONTEBRUNO *et al.*, 2019). A lei de potência, ao contrário, pode ser descrita por apenas um parâmetro, o expoente *b*.

Quadro 1 - Variáveis selecionadas do SNIS 2017

Categorias	Variável	Descrição	Unidade de medida
Tamanho da cidade	POP_URB	População urbana	habitantes
Serviços individualizados	AG002	Ligações ativas de água	unidades
Infraestrutura material	AG005	Extensão da rede de água	1.000 metros
	AG006	Volume de água produzido	1.000 m <sup>3</sup> /ano
	Perdas	Perdas na distribuição	1.000 m <sup>3</sup> /ano

Figura 1 - Mapa da rede urbana do Brasil e suas regiões (esquerda), da Região Sul (centro) e detalhe parcial ilustrativo do Rio Grande do Sul (direita)



Fonte: adaptado de dados espaciais de Pereira *et al.* (2019).

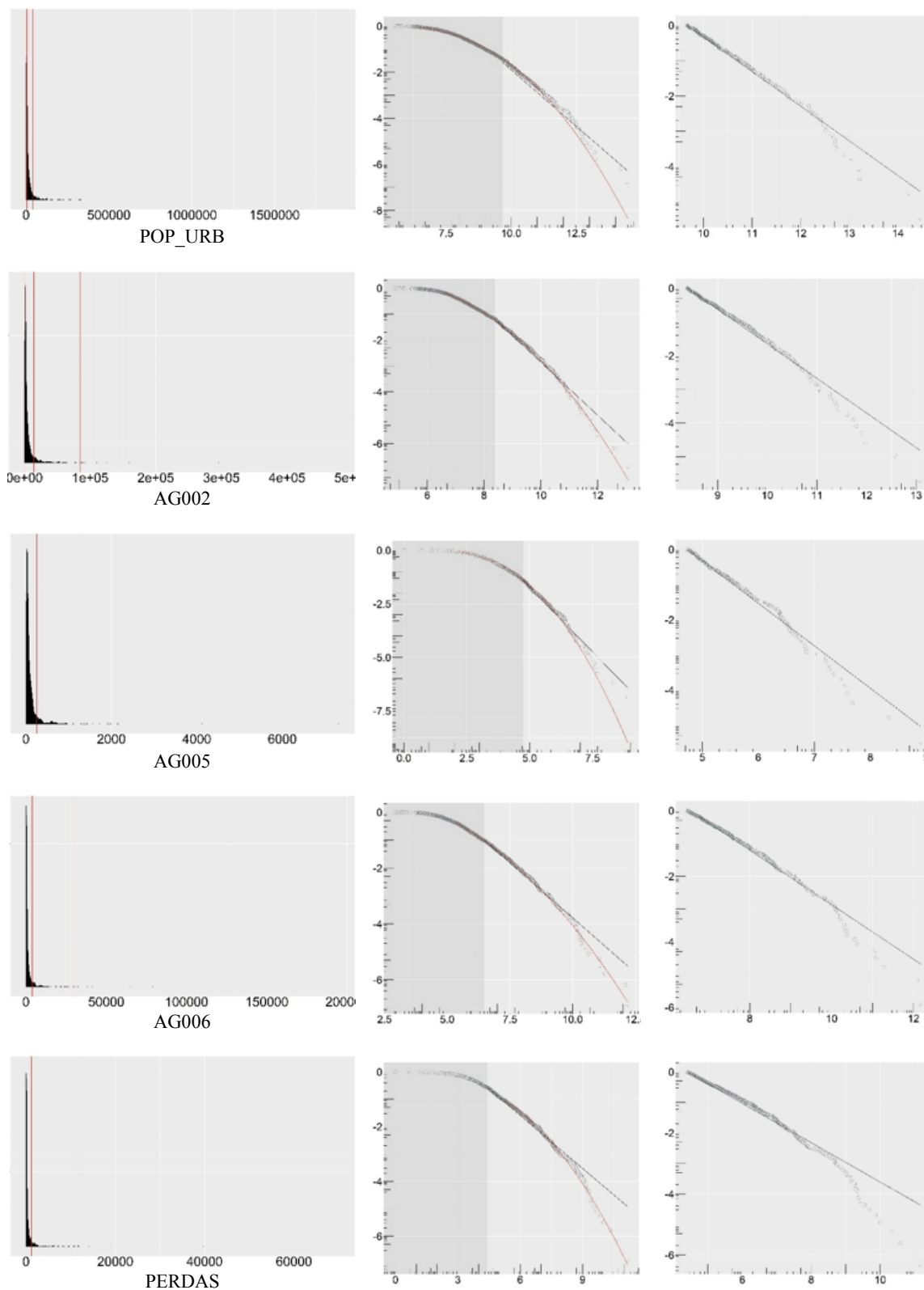
Tabela 1 - Estatísticas descritivas

Variável	N	Média	DesvP	CoefVar	Mín	Mediana	Máx	Assimetria
POP_URB	968	22.912	91.334	398,63	262	4.920	1.893.997	14,07
AG002	968	7.133	22.823	319,95	132	1.947	478.832	12,62
AG005	968	124	338	272,96	1	43,8	7.327	12,82
AG006	968	2.248	10.146	451,31	19	371	194.926	13,76
Perdas	968	766	3.307	431,86	1	103	70.080	13,03

Tabela 2 - Parâmetros das distribuições potência (PL) e lognormal (LN)

Variável	Categoria	N	PL par	PL xmin	PL ntail	LN xmin	LN ntail
POP_URB	Tamanho da cidade	968	1,983	15.870	234	644	937
AG002	Serviços básicos	968	2,020	4.399	295	776	794
AG005	Infraestrutura material	968	2,210	114	239	9	907
AG006		968	1,789	650	356	209	667
Perdas		968	1,644	82	553	163	363

Figura 2 - Esquerda: histogramas de densidade, com linhas vermelhas de percentis; centro: gráficos de probabilidade complementar acumulada CCDF (log x log) com retas de ajuste às distribuições PL em preto e LN em vermelho; e direita: ajuste à distribuição PL com valores  $>x_{min}$



Os resultados estão na Tabela 2, onde PL significa *powerlaw* e LN *lognormal*:  $x_{min}$  é o valor a partir do qual os dados se ajustam à distribuição;  $n_{tail}$  é o número de casos na porção melhor ajustada e  $PL$  par o parâmetro da distribuição potência, ou expoente  $b$ . Note-se que no gráfico CDF a cauda (*tail*) é a porção “final” do grupo de pontos, correspondente aos poucos elementos grandes com baixa probabilidade de ocorrência. A literatura reporta valores de  $b$  geralmente entre 2 e 3 para distribuições bem ajustadas (CLAUSET; COSMA; NEWMAN, 2009).

Os gráficos da coluna central da Figura 2 apresentam curvatura bastante acentuada no início, indicando efetivamente uma distribuição PL apenas na porção acima de  $PL x_{min}$  – destacada em cinza claro. Na coluna da direita da Figura 2 vemos o mesmo tipo de gráfico, plotados apenas com os valores acima do  $x_{min}$ , com as retas de ajuste PL.

A comparação dos modelos através do teste de Vuong – razão de verossimilhanças – mostrou que, com exceção de AG006, a distribuição das variáveis se ajusta melhor à distribuição lognormal. Porém, o teste requer que os  $x_{min}$  sejam iguais em ambos os modelos, o que é contraintuitivo. De qualquer maneira, os percentis 50º, 90º e 99º de uma lognormal teórica gerada pelos parâmetros extraídos da análise da variável POP\_URB são 7.016, 26.740 e 79.595 habitantes, contra 4.917, 40.552 e 305.466 habitantes da distribuição original, respectivamente.

A análise, portanto, não é conclusiva. Na prática, a melhor descrição é de um comportamento *powerlaw tail*, ou seja, a partir de um limiar a distribuição passa a obedecer à função potência (NEWMAN, 2005).

## Modelando as relações de escala

As estatísticas das regressões entre cada variável (Y) e a população urbana (X) estão na Tabela 3, e os gráficos de dispersão com retas de ajuste na Figura 3.

Os expoentes – coeficientes dos X – ficaram, com exceção das perdas, consistentemente abaixo de um. Isso é coerente com um regime sublinear, mesmo que as diferenças sejam pequenas em relação à escala linear. Também não há diferença clara entre os expoentes das categorias *serviços individualizados* e *infraestrutura*, mesmo que os primeiros sejam fornecidos como unidades (ligações) e, portanto, sugiram a relação “um-para-um” com a população.

As regressões MQO separadas entre as variáveis e a porção dos valores da população urbana acima e abaixo de  $x_{min}$  retornaram expoentes semelhantes, com piores  $r^2$ .

No entanto, a qualidade de ajuste das regressões ( $r^2$ ) com AG005 e AG006 (medidos “a granel”) é sensivelmente inferior à de AG002. Essa diferença, aliada ao sinal dos interceptos – positivo para serviços individualizados, negativo para infraestrutura – sinaliza para a distinção entre as categorias.

As perdas não seguem o mesmo regime da infraestrutura, o que pode corroborar a hipótese sobre sua natureza distinta. Ela apresentou expoente pouco maior que um, sugerindo seu escalamento com a população em regime semelhante ao encontrado na literatura para indicadores socioeconômicos.

## Estimando as perdas

Talvez o aspecto mais prático dos regimes de escala seja a possibilidade de antecipar padrões locais a partir do comportamento do sistema regional. Dois tipos de estimativas podem ser calculados:

- aplicando a equação de regressão estimada no item Modelando as relações de escala com o parâmetro  $b$  e, a seguir, retornando ao valor absoluto das perdas através da função *exp*; e
- utilizando uma cidade do sistema como referência para a determinação de um expoente “efetivo” (BARTHELEMY, 2019b).

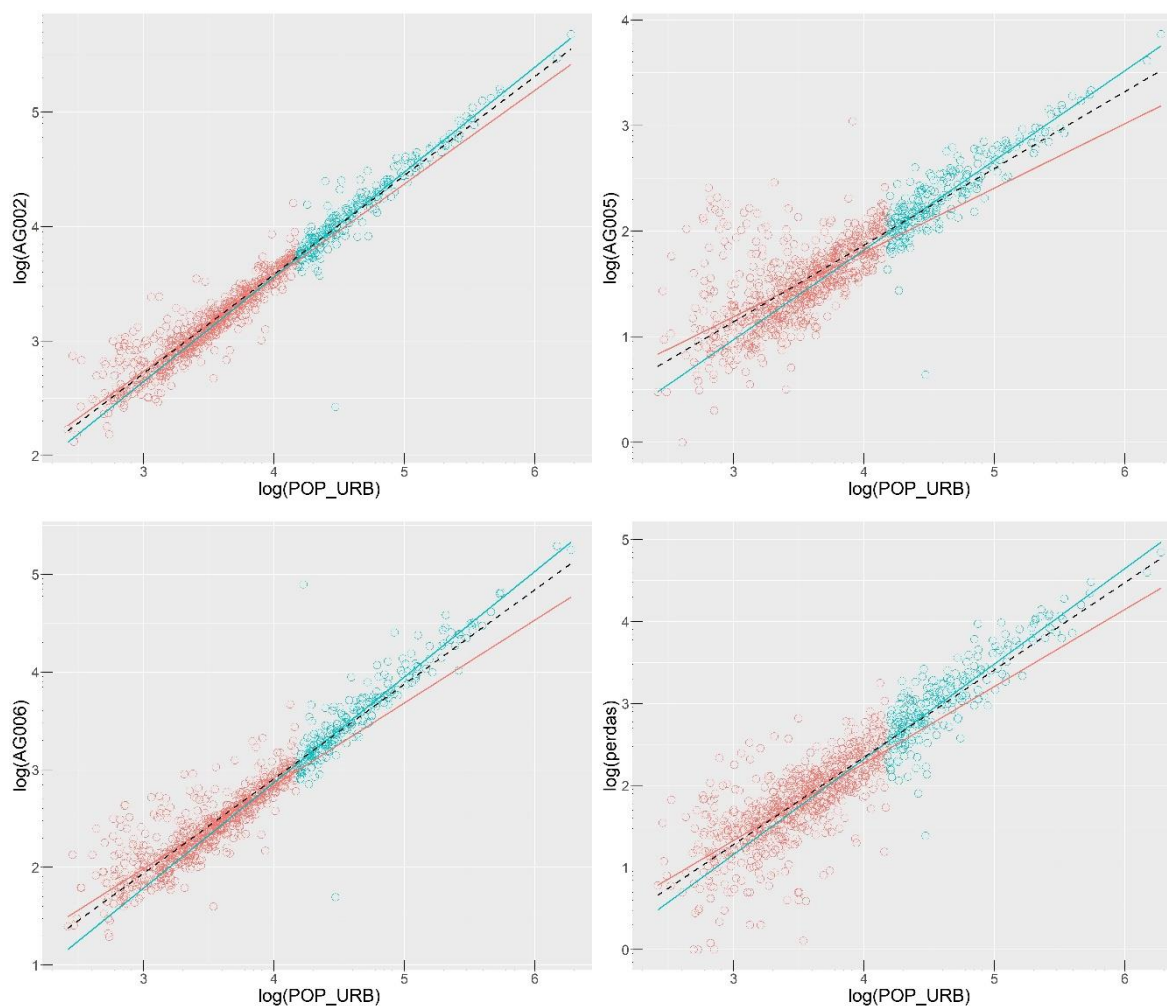
Tabela 3 - Estatísticas dos modelos de regressão linear entre as medidas dos SDA ( $\log Y$ ) e a população urbana ( $\log X$ )

Y	Intercepto	X coeficiente (b)	Erro padrão Xcoef	Adj $r^2$ reg.
AG002	0,292	0,863	0,006	0,943*
AG005	-2,167	0,734	0,014	0,729*
AG006	-2,224	0,968	0,010	0,902*
Perdas	-4,461	1,065	0,016	0,806*

Nota: \*indica p-valor < 0,01.



Figura 3 - Gráficos de dispersão das regressões lineares entre  $\log(\text{POP\_URB})$  e  $\log(\text{demais variáveis})$ : AG002 - Ligações ativas de água; AG005 - Extensão da rede de água, AG006 - Volume de água produzido e perdas. A reta de ajuste pontilhada corresponde à regressão com a população total, enquanto as linhas sólidas correspondem aos ajustes separados para as porções acima (azul) e abaixo (vermelho) do  $x_{min}$  populacional (15.870 habitantes)



No primeiro caso, utilizando  $b=1,07$ , calculamos o erro das estimativas através da razão  $Y_{previsto}/Y_{real}$  (=1 quando a estimativa é perfeita), chegando a 40% dos valores entre 0,5 e 1,5 ( $\pm 50\%$  de erro).

A segunda abordagem utiliza a Equação 4:

$$Y_i = Y_r * (P_i/P_r)^{b_{ef}} \quad \text{Eq. 4}$$

Onde:

$Y_i$  são as perdas na cidade  $i$ ;

$Y_r$  as perdas na cidade  $r$  (referência);

$P_r$  e  $P_i$  as populações; e

$b_{ef}$  o expoente de escala efetivo de  $Y$  em função de  $P$ .

Para aplicar essa fórmula, é preciso antes calcular os expoentes locais ( $b_{loc}$ ) para cada par de cidades aplicando a Equação 5:

$$b_{loc} = \log(Y_i/Y_r)/\log(P_i/P_r) \quad \text{Eq. 5}$$

Se a relação de escala fosse perfeita, apenas um  $b_{loc}$  seria estimado para todos os pares de cidades, igual àquele obtido pelo ajuste direto (1,07). No entanto, há flutuação nos vetores de expoentes, exigindo a adoção

daquele com menor variância (BARTHELEMY, 2019b). As 6 maiores cidades tiveram resultados próximos, e a mais baixa variância foi a de Curitiba, com média de  $1,096 \pm 0,10$ .

Com esse expoente efetivo, estimamos as perdas em todas as cidades com base na Equação 4 e comparamos novamente os dados previstos e os reais, chegando a 64% dos valores entre 0,5 e 1,5.

O gráfico da esquerda da Figura 4 mostra a média do expoente local para cada razão populacional da amostra completa, enquanto o da direita mostra a mesma medida apenas para as 234 cidades acima do  $x_{min}$  populacional.

O valor de  $b_{loc}$  varia fortemente quando a razão é pequena, impactado pelas diferenças entre as perdas em cidades de população semelhante, especialmente as menores. Até  $Pr/Pi < 10$ ,  $b_{loc} = 0,936 \pm 1,85$ . A partir de  $Pr/Pi > 10$  há diminuição da variação e estabilização em valores pouco acima de 1: entre 10 e 100 vezes de razão populacional,  $b_{loc} = 1,199 \pm 0,72$  e acima de 100 vezes,  $b_{loc} = 1,109 \pm 0,11$ .

Finalizando, realizamos uma inspeção no comportamento das perdas de água per capita – tanto o volume por habitante quanto o volume por ligação ativa em busca de reforço de nossa hipótese. Se não houvesse relação dessas medidas com o tamanho da cidade, as distribuições de ambas deveriam se aproximar de uma curva normal, com média representando um padrão comum. Mas a assimetria causada pelas cidades com perdas per capita muito altas lança a média para a direita e desvia a distribuição da normalidade. Os valores a partir do 75º percentil (segunda linha vermelha nos gráficos da Figura 5) correspondem a 243 cidades, incluindo o grupo das 234 com populações acima do  $x_{min}$ , em ambas as distribuições.

Mesmo que também ocorram valores grandes em cidades pequenas, as perdas por habitante se estabilizam em patamares mais altos no grupo das 234 cidades da cauda da população, com média de  $0,032 \pm 0,001$  m<sup>3</sup>/hab/ano, contra  $0,028 \pm 0,001$  e um mínimo de 0,001 contra 0,0004.

Figura 4 - Esquerda: gráfico de dispersão entre a média do vetor de expoentes locais  $b_{loc}$  (eixo Y) de cada cidade em função da sua razão  $Pr/Pi$  (eixo X); e direita: mesma plotagem com X restrito

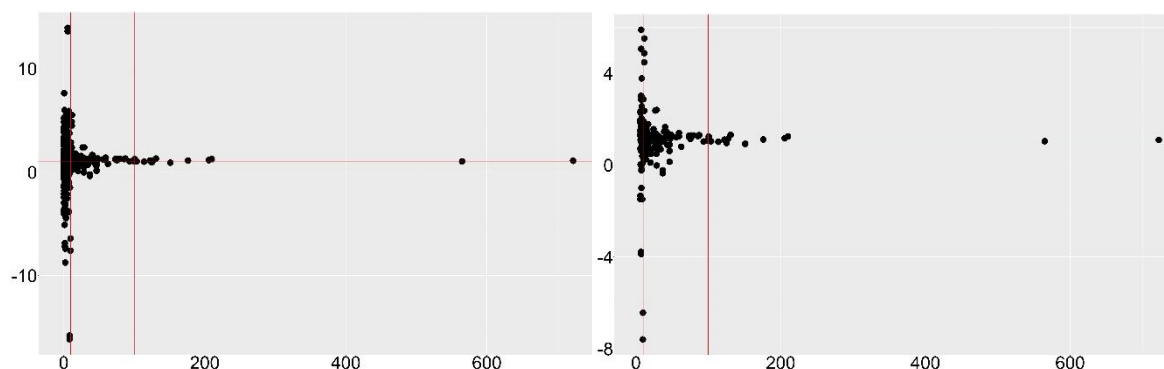
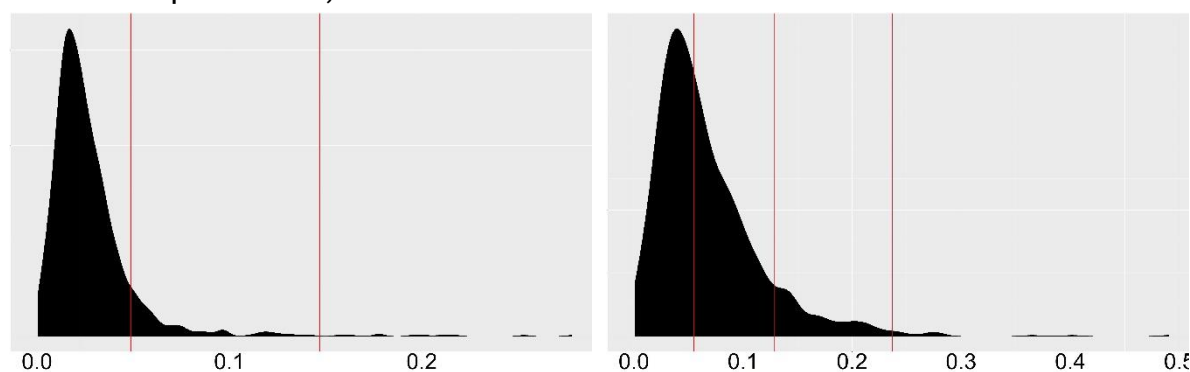


Figura 5 - Histogramas (densidade) de perdas por habitante (esquerda) e perdas por ligação (direita), com linhas de percentis 50o, 75o e 90o



## Discussão

Os resultados não trouxeram parâmetros conclusivos para as distribuições nem para os regimes de escala, mas apresentaram indícios quantitativos favoráveis à hipótese de não linearidade dos sistemas e, portanto, à teoria de que as perdas são produto das interações entre os componentes do sistema sociotécnico, especialmente a partir de um limiar mínimo de população. Eles provêm interessantes *insights* que conectam os mecanismos sociotécnicos que, em tese, lhes dão origem aos processos matemáticos por trás de cada distribuição. Enquanto uma lognormal pode ser gerada por simples multiplicações em sequência do tipo  $C \cdot X$  ou  $C/X$ , onde  $C$  é um fator de crescimento aleatório interno independente e igualmente distribuído para cada cidade (MONTEBRUNO *et al.*, 2019), a distribuição potência não. Porém, se considerada a restrição do  $X$  mínimo, a iteração das multiplicações eventualmente leva a distribuição a uma *powerlaw* (MITZENMACHER, 2004).

Este parâmetro  $x_{min}$  da população sugere que a sub-rede de cidades acima de 15.870 habitantes da região se organiza quantitativamente de forma análoga a um sistema complexo, tal como os muitos outros naturais e sociotécnicos discutidos na literatura nos anos recentes.

Os valores encontrados para o expoente de escala entre o volume de perdas e a população dos municípios apresentaram grande flutuação e pequena diferença em relação àquele da relação trivial ( $b=1$ ), não permitindo afirmar categoricamente que ela obedece ao regime superlinear. Este e outros achados, no entanto, sugerem que o comportamento das perdas não é igual ao das variáveis de infraestrutura da qual derivam – extensão da rede, número de ligações, volume produzido – mas sim ao de produtos socioeconômicos, conforme apontado na literatura (MEIRELLES *et al.*, 2018).

## Conclusões

Este artigo relata estudo com 968 municípios do Sul do Brasil e seus sistemas de distribuição de água a partir do paradigma da “complexidade” e da “Nova Ciência das Cidades” com o objetivo de verificar a existência de regularidades quantitativas não lineares entre a população, os SDAs e os volumes de perdas. Encontramos um sistema altamente hierárquico, no qual os tamanhos de cidades e redes se distribuem estatisticamente com “cauda pesada”, com concentração de recursos no topo e não na média. A partir de um limiar de população, essa distribuição se aproxima à dos sistemas complexos naturais, biológicos e sociais.

Encontramos expoentes para as funções potência que descrevem a relação dos indicadores de infraestrutura e a população das cidades coerentes com a literatura. O regime sublinear para as quantidades da rede representa sua origem em processos *top-down* nos quais o crescimento da cidade não requer expansão equivalente do sistema de água, garantindo o abastecimento com economias de escala.

O regime superlinear para as perdas indicaria a existência de retornos crescentes – negativos – típicos da produtividade social que emerge de processos *bottom-up*. Quanto maior a cidade, muito mais interações e mais produção per capita do que se poderia esperar em um sistema trivial. O uso anormal ou furto de água parece seguir um padrão mais alto nas cidades maiores, gerando efeitos também sobre as perdas reais, pois o consumo da água furtada faz aumentar a pressão na rede de água projetada para um consumo menor, levando a mais vazamentos, conforme já apontado por Azevedo e Saurin (2018).

Cidades grandes atraem desproporcionalmente mais habitantes, retroalimentando a concentração e a hierarquia não linear do sistema. Nelas, as redes materiais se expandiram de forma incremental ao longo do tempo, anexando ramais e ligações regulares em ritmo relativamente lento e controlado. Porém, essa expansão torna também mais provável o gradativo surgimento de ligações clandestinas, “gatos” e outros usos não autorizados, manifestações eminentemente sociais que tendem a crescer em ritmo acelerado especialmente em áreas de alta densidade e/ou informalidade.

Esses geradores de perdas seriam, portanto, dificilmente elimináveis por completo em sociedades com alta desigualdade como a brasileira.

Nesse sentido, a revisão das políticas de saneamento no Brasil em andamento deve ser observada com atenção, pois os operadores privados tendem a privilegiar as cidades grandes, onde é possível lucrar mais com a operação. A tendência é, portanto, de aumento da concentração do sistema, pois os investimentos provavelmente serão alocados prioritariamente nas cidades que, contraditoriamente, já possuíam a capacidade de garantir o abastecimento aos seus habitantes (MEIRELLES *et al.*, 2018). Mesmo com maiores níveis de perdas per capita e sua aparente imprevisibilidade, pode valer mais a pena operar nas cidades maiores, o que vai demandar cuidado do poder público com as questões de equidade.

Este estudo não fornece modelos precisos na escala micro nem estimativas exatas das perdas reais, as quais podem ser desenvolvidas a partir de informações detalhadas dos operadores e do rico banco do SNIS, contemplando dados relativos a investimento, faturamento e medição de consumo, entre outros. Mas ele revela as regularidades macroscópicas de um sistema de cidades e suas redes, associando infraestrutura material a fenômenos sociais e sugerindo explicações para os mecanismos sociotécnicos que regem a evolução das cidades no tempo. Essa visão pode complementar as abordagens tradicionais, visando aprimorar o uso da água nas cidades brasileiras.

Sua principal limitação é a natureza eminentemente fenomenológica. As teorias são bastante convincentes quando propõem que as distribuições hierárquicas e os regimes de escala são expressões quantitativas da “complexidade”, articulando ao longo do tempo o planejamento centralizado com as emergências auto-organizadas. Mas elas são incapazes de explicitar os mecanismos causais que relacionam as interações microscópicas locais e o comportamento emergente global (BARTHELEMY, 2019a).

Faz-se necessária, portanto, a realização de estudos em outras regiões e culturas, bem como a análise de séries temporais para investigar a variação longitudinal dos elementos do sistema, fundamentais para efetivamente compreender a dinâmica populacional e de evolução das redes materiais. Também se deve buscar outras técnicas quantitativas para avançar nesse entendimento e, assim, colaborar mais concretamente com o problema das perdas em sistemas de abastecimento de água.

## Referências

- AL-OMARI, A. A Methodology for the Breakdown of NRW into Real and Administrative Losses. **Water Resources Management**, v. 1, p. 1913–1930, 2013.
- AZEVEDO, B. B.; SAURIN, T. A. Losses in water distribution systems: a complexity theory perspective. **Water Resources Management**, v. 32, p. 2919–2936, 2018.
- BAK, P.; TANG, C.; WIESENFELD, K. Self-organized criticality. **Physical Review A**, v. 38, p. 364–374, 1987.
- BARABÁSI, A.; ALBERT, R.; JEONG, H. Mean-field theory for scale-free random networks. **Physica A**, v. 272, n. 3–4, p. 173–187, 1999.
- BARTHELEMY, M. The statistical physics of cities. **Nature Review Physics**, v. 1, p. 406–415, 2019a.
- BARTHELEMY, M. Tomography of scaling. **Journal of the Royal Society Interface**, v. 16, n. 160, 2019b.
- BATTY, M. **The New Science of Cities**. Cambridge: The MIT Press, 2013.
- BETTENCOURT, L. *et al.* Growth, innovation, scaling, and the pace of life in cities. **Proceedings of the National Academy of Sciences of the United States of America**, v. 104, n. 17, p. 7301–7306, 2007.
- CARAYON, P. Human factors of complex sociotechnical systems. **Applied Ergonomics**, v. 37, p. 525–535, 2006.
- CILLIERS, P. **Approaching complexity: complexity and postmodernism understanding complex systems**. London; New York: Routledge, 1995.
- CLAUSET, A.; COSMA, R.; NEWMAN, M. E. J. Power-law distributions in empirical data. **SIAM Review**, v. 51, n. 4, p. 661–703, 2009.
- CRAWFORD, R. What can complexity theory tell us about urban planning? **New Zealand Productivity Commission Research Note 2016/2**, Abril, 2016.
- DEY, S.; LEE, S. REASSURE: Requirements elicitation for adaptive socio-technical systems using repertory grid R. **Information and Software Technology**, v. 87, p. 160–179, 2017.
- ELMARAGHY, H. *et al.* Technical paper A model for assessing the layout structural complexity of manufacturing systems. **Journal of Manufacturing Systems**, v. 33, n. 1, p. 51–64, 2014.
- ERDI, P. **Complexity Explained**. Berlin: Springer, 2008.
- GILLESPIE, C. S. Fitting heavy tailed distributions: the powerLaw package. **Journal of Statistical Software**, v. 64, n. 2, 2015.
- GLERIA, I.; MATSUSHITA, R.; SILVA, S. Sistemas complexos, criticalidade e leis de potência. **Revista Brasileira de Ensino de Física**, v. 26, n. 2, p. 99–108, 2004.

- GOMES, R.; MARQUES, A.; SOUSA, J. District metered areas design under different decision makers' options: cost analysis. **Water Resources Management**, v. 27, p. 4527–4543, 2013.
- GONZÁLEZ-GÓMEZ, F. *et al.* Explanatory factors of urban water leakage rates in Southern Spain. **Utilities Policy**, v. 22, p. 22–30, set. 2012.
- HABERFELLNER, R. *et al.* **Systems Engineering**. 12. Auflage. Zürich: Orell Füssli Verlag, 2012.
- HERRERA, M.; ABRAHAM, E.; STOIANOV, I. Graph-theoretic surrogate measures for analysing the resilience of water distribution networks. **Procedia Engineering**, v. 119, n. 1, p. 1241–1248, 2015.
- IGNAZZI, A. C. **Coevolution in the Brazilian system of cities**. PhD thesis, Universidade de Paris, Paris, 2015.
- INSTITUTO DE PESQUISA ECONÔMICA E APLICADA. **Ipea data**. Disponível em: <https://www.ipea.gov.br/portal/>. Acesso em: 13 jul. 2019.
- KINGDOM, B.; LIEMBERGER, R.; MARIN, P. **The challenge of reducing non-revenue water (NRW) in developing countries - how the private sector can help** : a look at performance-based service contracting. 2006. Disponível em: <http://documents.worldbank.org/curated/en/2006/12/7531078/challenge-reducing-non-revenue-water-nrw-developing-countries-private-sector-can-help-look-performance-based-service-contracting>. Acesso em: 14 mar. 2019.
- KURTZ, C. F.; SNOWDEN, D. J. The new dynamics of strategy: sense-making in a complex and. **IBM Systems Journal**, v. 42, n. 3, p. 462–483, 2003.
- MAMADE, A. *et al.* Energy auditing as a tool for outlining major inefficiencies: results from a real water supply system. **Procedia Engineering**, v. 119, n. 1, p. 1098–1108, 2015.
- MEIRELLES, J. *et al.* Evolution of urban scaling: evidence from Brazil. **PLoS ONE**, v. 13, n. 10, p. 1–15, 2018.
- MITZENMACHER, M. A brief history of generative models for power law and lognormal distributions. **Internet Mathematics**, v. 1, n. 2, p. 226–251, 2004.
- MONTEBRUNO, P. *et al.* A tale of two tails: do Power Law and Lognormal models fit firm-size distributions in the mid-Victorian era? **Physica A: Statistical Mechanics and its Applications**, v. 523, p. 858–875, mar. 2019.
- NEWMAN, M. E. J. Power laws, Pareto distributions and Zipf's law. **Contemporary Physics**, v. 46, n. 5, p. 323–351, 2005.
- OLIVEIRA, G. *et al.* **Desafios para disponibilidade hídrica e avanço da eficiência do saneamento**. Trata Brasil, Go associadoss, 2018.
- PEREIRA, R. H. M. *et al.* **geobr**: Loads Shapefiles of Official Spatial Data Sets of Brazil. R package. R Found. Stat. Comput., Vienna. 2019.
- PERROW, C. **Normal accidents**: living with high-risk technologies. Princeton: Princeton University Press, 1984.
- RÍOS, J. C. *et al.* Methodology for the identification of apparent losses in water distribution networks. **Procedia Engineering**, v. 70, p. 238–247, 2014.
- SACCHET, O. F. *et al.* Chaotic behavior of consumer in the digital age: made an approach based on the theory of chaos. **Revista de Administração da UNIMEP**, v. 9, n. 3, p. 87–111, 2011.
- SALAT, S.; BOURDIC, L. Power laws for energy efficient and resilient cities. **Procedia Engineering**, v. 21, p. 1193–1198, 2011.
- SISTEMA NACIONAL DE INFORMAÇÕES EM SANEAMENTO. **Diagnóstico do SNIS**. Disponível em: <http://www.snis.gov.br>. Acesso em: 13 jul. 2019.
- SKIPWORTH, P. J. *et al.* The effect of regional factors on leakage levels and the role of performance indicators. **Water and Environment Journal**, p. 184–188, 1999.
- SNOWDEN, D. J. *et al.* A leader's framework for decision making. **Harvard Business Review**, p. 69–76, 2007.

STRANO, E.; SOOD, V. Rich and poor cities in Europe. An urban scaling approach to mapping the European economic transition. **PLoS ONE**, v. 11, n. 8, p. 1–8, 2016.

TROJAN, F.; MORAIS, D. Maintenance management decision model for reduction of losses in water distribution networks. **Water Resources Management**, p. 3459–3479, 2015.

VAN DEN BERG, C. Drivers of non-revenue water: a cross-national analysis. **Utilities Policy**, v. 36, p. 71–78, 2015.

WECK, O. L. de; ROOS, D.; MAGEE, C. L. **Engineering systems**: meeting human needs in a complex technological world. London: PHI Learning Private Limited, 2013.

WEST, G. **Scale**: the universal laws of growth, innovation, sustainability, and the pace of life in organisms, cities, economies, and companies. New York: Penguin Press, 2017.

**Júlio Celso Borello Vargas**

Departamento de Urbanismo | Universidade Federal do Rio Grande do Sul | Rua Sarmento Leite, 320, sala 521, Farroupilha | Porto Alegre - RS - Brasil | CEP 90050-170 | Tel.: (51) 3308-3123 | E-mail: julio.celso@ufrgs.br

**Bárbara Brzezinski Azevedo**

Departamento Municipal de Água e Esgoto | Rua 24 de Outubro, 200, Moinhos de Vento | Porto Alegre - RS - Brasil | CEP 90510-000 | E-mail: babibrzezinski@gmail.com

***Ambiente Construído***

Revista da Associação Nacional de Tecnologia do Ambiente Construído

Av. Osvaldo Aranha, 99 - 3º andar, Centro

Porto Alegre - RS - Brasil

CEP 90035-190

Telefone: +55 (51) 3308-4084

[www.seer.ufrgs.br/ambienteconstruido](http://www.seer.ufrgs.br/ambienteconstruido)

[www.scielo.br/ac](http://www.scielo.br/ac)

E-mail: [ambienteconstruido@ufrgs.br](mailto:ambienteconstruido@ufrgs.br)



This is an open-access article distributed under the terms of the Creative Commons Attribution License.