**An Acad Bras Cienc (2020) 92(3): e20200594 DOI 10.1590/0001-3756202020200594**
Anais da Academia Brasileira de Ciências | *Annals of the Brazilian Academy of Sciences*
Printed ISSN 0001-3765 I Online ISSN 1678-2690
www.scielo.br/aabc | www.fb.com/aabcjournal

ENGINEERING SCIENCES

# Trade-off between number of constraints and primary-statement robustness in entropy models: the case of the open-channel velocity field

ANTONIO VIANA DA SILVA FILHO, JOSÉ CARLOS DE ARAÚJO & ARMIN RAABE

**Abstract:** In this research, the trade-off between the number of restrictions and the robustness of the primary formulation of entropy models was evaluated. The performance of six hydrodynamic models in open channels was assessed based on 1730 Laser-Doppler anemometry data. It was investigated whether it is better to use an entropy-based model with more restrictions and a weak primary formulation or a model with fewer restrictions, but with a strong formulation. In addition, it was also investigated whether the model performance improves with the insertion of restrictions. Three of the investigated models have a weak formulation (open-channel velocity field represented by Cartesian coordinates); while the other three models have a strong formulation, according to which isovels are represented by curvilinear coordinates. The results indicated that models with two restrictions performed better than those with one restriction, since the additional restriction includes information relevant to the system. Models with three restrictions perform worse than those with two restrictions, because the information lost due to the use of a numerical solution was more substantial than the information gained by the third restriction. In conclusion, a strong primary formulation brought more information to the system than the inclusion of a third constraint.

**Key words:** Hydrodynamic model, Information theory, Laser-Doppler anemometry, Shannon entropy.

## INTRODUCTION

Models are designed to explain physical processes, to which one may assign a probability of event occurrence, considering that each system state has a level of uncertainty. Well-designed probabilistic models tend to enhance their capacity of representing reality because they consider the intrinsic uncertainties of the processes, which arise from several sources such as natural randomness, inaccuracy in data measurement, model structure imperfect parameterization, and others Gupta & Govindaraju (2019). Many water-related problems demand a probabilistic approach due to the considerable amount of uncertainty involved Mishra (2009), Cobo et al. (2017), e.g., rainfall occurrence, magnitude and intensity Mélèse et al. (2018), basin flow and sediment Shrestha et al. (2016), and hydraulics Tapoglou et al. (2019). Shannon (1948) investigated the information content and its relation to uncertainty measures while proposing uncertainty quantification, the so-called Shannon informational entropy, or simply Shannon entropy. Jaynes (1957a, b) physically formulated the informational principle of maximum entropy (PME) using Shannon entropy, which maximizes uncertainty under the given constraints and, thus, avoids

the use of unproven assumptions. As a result, the probability density function associated to a researched process can be obtained by maximizing the constrained entropy function and using the variational calculus and the Lagrange multipliers method.

Therefore, the informational principle of maximum entropy Shannon (1948), Jaynes (1957a, b), Shore & Johnson (1980) provides an adequate approach for introducing probability into complex hydrodynamic problems such as modeling the velocity field in open channels. It has been successfully applied to several fields of Hydraulic and Environmental Engineering Harmancioglu & Singh (1998), Singh (2013), (2014), Ardiclioglu et al (2005). Other researches based on PME yielded encouraging results in areas such as water resources Cheng et al. (2019), drought assessment Zuo et al. (2017), climate change Jin et al. (2016), ecology Banavar et al. (2010), assessment of river discharges Chiu et al. (2005), Alvisi et al. (2014), Farina et al. (2014), sediment yield Chiu & Hsiung (1981), de Araújo (2007), Furbish et al. (2016), pipe-network hydraulic Waldrip et al. (2016) soil moisture Al-Hamdan & Cruise (2010), river morphology Moramarco et al. (2013), and open-channel hydrodynamics Chiu (1987), (1988), (1989), (1991), Chiu et al. (2005), Barbé et al. (1991), de Araújo & Chaudhry (1998), Luo & Singh (2011), Fontana et al. (2013), Singh et al. (2013), Jiang & Chen (2016), Greco & Martino (2018), Mirauda & Russo (2019), to name but a few.

Literature presents several robust, physically-based hydrodynamic models, such as Shiono-Knight Shiono & Knight (1991), Knight (2013), MIKE-11 DHI (1992), (2017), River2D Steffler & Blackburn (2002), Beakes et al. (2014), CE-QUAL-W2 Cole & Wells (2006), and Environmental Fluid Dynamics Code Tetra Tech (2007), among others. These models use complex spatially-distributed systems of equations that encompass the principles of mass, energy, and momentum conservation; as well as the effects of turbulence. Hydrodynamic models, such as the aforementioned ones, are able to solve complex problems Beakes et al. (2014), Knight (2013), Torres-Bejarano et al. (2015), Thanh et al. (2020), but demand a large number of parameters, which are often unavailable. When not based on measured data, the parameterization process may introduce uncertainty to such an extent that simple few-parameter models yield better results than the complex ones, especially in ungauged basins. Despite the fact that entropy equations tend to demand few parameters, they often out-perform equations based on different approaches, due to the robustness of the principle Chiu (1987), (1988), (1991), de Araújo & Chaudhry (1998), Ardiclioglu et al. (2005), Fontana et al. (2013), Singh (2014), Alvisi et al. (2014), Mirauda & Russo (2019).

The entropy-equation optimization is subject to the given constraints, which represent information about the problem to be solved. Therefore, the greater the number of constraints, the more information there is about the system Jaynes (1957a). More information implies less uncertainty and more accurate models. Nevertheless, the density functions obtained by PME can be solved analytically only if a maximum of two constraints are used. If the optimization uses three or more constraints, it demands a numerical solution, which simultaneously lowers the model accuracy and increases the computational effort. This ambiguity – more than two constraints generate more information, but also weaken the computational solution – yields a non-trivial non-linear problem, which has not (to our best knowledge) been straightforward tackled in the Literature, especially by researches based on an accurate robust datasets. Therefore, the objective of this work was to assess the trade-off between number of constraints and

strength of the primary statement on the performance of hydrodynamic entropy-based models using accurate laboratory data for validation. For this purpose, the following quests were analyzed: (i) is there an improvement in model performance when two constraints are used instead of one? (ii) is there an improvement in model performance when a third constraint is introduced, considering that its solution is not analytical? (iii) is it better to use a three-constraint model with a weak primary statement, or a two-constraint model with a strong primary statement?

### Abbreviations

D : Flow depth at the channel
F(u) Probability of the longitudinal velocity being less or equal to ***u***
H : Entropy function
M : Entropy parameter.
NSE : Nash& Sutcliffe
p(u) : Probability density function
PME : Principle of maximum entropy
RMSE : Root mean square error
SRP : Steffler, Rajaratnam and Peterson (1985)
u : Longitudinal velocity
$u_{av}$ : average velocity.
$U_{max}$ : Maximum velocity in the cross section
U1y : Model with one constraint and Cartesian coordinates.
U2y : Model with two constraints and Cartesian coordinates.
U3y : Model with three constraint and Cartesian coordinates.
U1ξ : Model with one constraint and curvilinear coordinates.
U2ξ : Model with two constraints and curvilinear coordinates.
U3ξ : Model with three constraints and curvilinear coordinates.
y : The vertical distance of a any point located in the flow from the channel bed
z : The horizontal distance of a any point located in the flow from the nearest wall.
β : Boussinesq coefficient.
$\lambda_i$ : Lagrange parameters, i = 1,2,...,12.
ξ : isovel
δy, δi, βi, ε: shape parameters

## MATERIALS AND METHODS

We investigated the performance of six entropy models designed to simulate open-channel velocity fields. The simulated velocities were compared with accurately-measured laboratory data. Three of the investigated models have a weak primary statement, i.e., they assume that isovels could be well represented by Cartesian coordinates; whereas the remaining three models have a strong statement according to which isovels are better represented by curvilinear coordinates. In this work, data were extracted from the experiments made by Steffler et al. (1985): run 1 (hereafter called SRP1), run 2 (SRP2), and run 3 (SRP3). The experiments (see the main characteristics in Table I) were performed at the Thomas Blench Laboratory flume located at the University of Alberta, Canada. The velocities were accurately measured using a Laser-Doppler anemometer. The models performance was assessed with the Nash-Sutcliffe coefficient (NSE) and the root mean square error (RMSE).

### Models with weak primary statement

Two primary statements were assumed and models with one, two or three constraints used, respectively, in order to maximize the entropy function H (Equation 1), in which u means the longitudinal velocity; p(u) the respective probability density function; and $U_{max}$ the maximum velocity in the cross section. The six entropy models are divided into two groups: three models admit the Cartesian coordinate

**Table I. Parameters of the experiments: run 1 (SRP1), run 2 (SRP2), and run 3 (SRP3).**

| Experiment | n | n.vert | Q (m³/s) | D (m) | B (m) | B/D (-) | A (m²) | $u_{av}$ (m/s) | $U_{max}$ (m/s) | ξ (m) |
|---|---|---|---|---|---|---|---|---|---|---|
| SRP1 | 526 | 10 | 0.126 | 0.146 | 1.143 | 7.83 | 0.167 | 0.755 | 0.844 | 0.00 |
| SRP2 | 663 | 16 | 0.126 | 0.225 | 1.143 | 5.08 | 0.257 | 0.490 | 0.530 | 0.00 |
| SRP3 | 541 | 16 | 0.032 | 0.093 | 1.143 | 12.29 | 0.106 | 0.301 | 0.372 | 0.00 |
| All | 1730 | 42 | (-) | (-) | (-) | (-) | (-) | (-) | (-) | (-) |

**The symbols mean: n = number of measured points; n.vert = number of measured vertical profiles; Q = discharge; D = flow depth; B = flume width; A = wetted area; $u_{av}$ = average velocity, given by Q/A; $U_{max}$ = maximum measured velocity; ε = maximum-velocity dip. Data source : Steffler et al. (1983).**

system (weak primary statement, Equation 2), whereas the three others admit the curvilinear coordinate system, as described in (Chiu 1988) (Figure 1).

$$H(u) = \int_0^{Umax} p(u).Ln\left[p(u)\right].du \quad (1)$$

Model U1y (one constraint and Cartesian coordinates) is based on Chiu (1987), who proposes the primary statement (Equation 2), according to which F(u) is the probability of the longitudinal velocity being less or equal to ***u*** at a point located at distance ***y*** from the channel bed. In Equation 2, D is the flow depth at the channel.

$$F(u) = y / D \quad (2)$$

The first constraint is presented in Equation 3. It means that the integral of the probability density function p(u) over the whole dominium equals unity. We maximize the entropy function (H, Equation 1) subjected to one constraint (Equation 3) and further apply the result at Equation 2, yielding the velocity-distribution Equation 4, where $\lambda_1$ is the Lagrange parameter, calculated by Equation 5.

$$\int_0^{Umax} p(u)\,du = 1 \quad (3)$$

$$u(y) = e^{(1-\lambda_1)}.\left(\frac{y}{D}\right) = u_{max}.\left(\frac{y}{D}\right) \quad (4)$$

$$\lambda_1 = 1 + Ln\left(\frac{1}{U_{max}}\right) \quad (5)$$

Model U2y (two constraints and Cartesian coordinates) uses the same weak premise as model U1y (Equation 2). The entropy function (Equation 1) was maximized by Chiu (1987), using two constraints: Equations 3 and 6. The latter represents the mass conservation principle and indicates that the left-hand integral equals average velocity ($u_{av}$). As a result, the method yields the velocity-distribution Equation 7 with parameters $\lambda_2$, $\lambda_3$ and M that can be estimated using the maximum and average velocities by equations 8 and 9: Chiu (1987).

$$\int_0^{Umax} u.p(u)\,du = u_{av} \quad (6)$$

$$u(y) = \frac{1}{\lambda_3}.Ln\left(1 + \frac{\lambda_3}{e^{\lambda_2 - 1}}\frac{y}{D}\right) \quad (7)$$

$$u_{av} = \frac{(U_{max}.e^M) - \left(e^M / \lambda_3\right) + \left(1/\lambda_3\right)}{e^M - 1} \quad (8)$$

$$M = \lambda_3.Umax \quad (9)$$

Model U3y (three constraints and Cartesian coordinates), derived by Barbé et al. (1991), assumes the same premise as in the previous

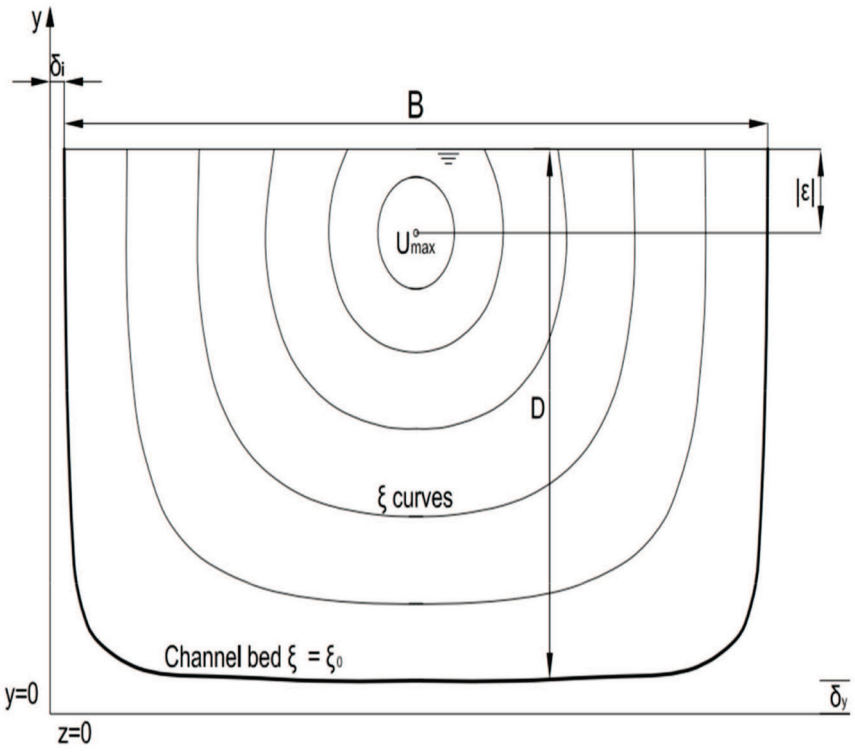


**Figure 1. Curvilinear coordinate system based on Chiu and Chiou (1986).**

models (Equation 2) and three constraints: Equations 3, 6, and 10. The third constraint represents the momentum conservation principle (Equation 10), in which β is the Boussinesq coefficient (Equation 11) and ρ is water density. The system generated by maximizing the entropy function (Equation 1) for the three constraints was solved using a numerical approach (MacLaurin series with the first two terms), which yields the approximate velocity field (Equations 13 and 14). The parameters $\lambda_4$, $\lambda_5$, and $\lambda_6$ can be obtained using the Boussinesq coefficient, maximum velocity and average flow velocity, as stated in Barbé et al. (1991).

$$\int_0^{U\max} u^2.p(u).du = \beta.u_{av}^2 \quad (10)$$

$$\beta = \frac{M}{\rho.D.u_{av}^2} \quad (11)$$

$$M = \lambda_5.Umax \quad (12)$$

$$e^{\lambda_4-1}.\left\{e^{\lambda_5.u} + \lambda_6.\left[G(u)\right]\right\} = \lambda_5.\left[\frac{y}{D} + e^{\lambda_4-1}.\left(\frac{1}{\lambda_5} + \frac{2.\lambda_6}{\lambda_5^3}\right)\right] \quad (13)$$

$$G(u) = e^{\lambda_5.u}.\left(u^2 - \frac{2.u}{\lambda_5} + \frac{2}{\lambda_5^2}\right) \quad (14)$$

## Models with a strong primary statement

The strong statement admits that the longitudinal velocity is directly associated with the curvilinear, rather than with Cartesian coordinates Chiu & Chiou (1986); and that isovels can be represented by ξ coordinates (Equations 15-17), as proposed by Chiu (1986). The isovel (ξ) shape parameters (δy, ε) and variables (y, z) are defined in Figure 1. Parameter $\beta_i$ characterizes the velocity distribution of the primary flow.

$$\xi = Y.(1-Z)^{\beta_i}.e^{(\beta_i.Z-Y+1)} \quad (15)$$

$$Y = \frac{y+\delta_y}{D+\delta_y+\varepsilon} \quad (16)$$

$$Z = \frac{|z|}{(B/2) + \delta_i} \quad (17)$$

Model U1ξ (one constraint and curvilinear coordinates) is based on Chiu (1988), who proposes the strong primary premise (Equation 18), according to which F(u) is directly associated with the isovel (ξ) spatial distribution, with the key parameters $\xi_{max}$ and $\xi_0$, respectively, the maximum and minimum ξ values of the open-channel flow. The entropy function ***H*** (Equation 1) was maximized and subjected to one constraint (Equation 3), the same as in model U1y. The result was applied to Equation 18, yielding the velocity-distribution (Equation 19), where $\lambda_7$ is the Lagrange parameter, estimated by Equation 20.

$$F(u) = \frac{\xi - \xi_0}{\xi_{\max} - \xi_0} \quad (18)$$

$$u(\xi) = e^{(1-\lambda_7)} \cdot \left( \frac{\xi - \xi_0}{\xi_{\max} - \xi_0} \right) \quad (19)$$

$$\lambda_7 = 1 - Ln(U_{\max}) \quad (20)$$

Model U2ξ (two constraints and curvilinear coordinates), developed by Chiu (1988), uses curvilinear coordinates (Equation 15), Equation 18 as primary statement, and two constraints: Equations 3 and 6, the same of model U2y, resulting in Equation 21. The parameters $\lambda_8$ and $\lambda_9$ are estimated in an analogous way as in model U2y, as shown in Chiu (1988).

$$u(\xi) = \frac{1}{\lambda_9} Ln\left[ 1 + \frac{\lambda_9}{e^{\lambda_8 - 1}} \cdot \left( \frac{\xi - \xi_0}{\xi_{\max} - \xi_0} \right) \right] \quad (21)$$

In the present study, Model U3ξ (three constraints and curvilinear coordinates) used curvilinear coordinates (Equation 15), the strong statement (Equation 18) and three constraints, (Equations 3, 6 and 10 as in model U3y: Barbé et al. (1991), but substituting ratio y/D by $\left(\frac{\xi - \xi_0}{\xi_{max} - \xi_0}\right)$. The U3ξ model for the velocity field in open channels consists in solving Equations 22 and 23. The system parameters $\lambda_{10}$, $\lambda_{11}$, and $\lambda_{12}$ are estimated analogously as in model U3y, using the Boussinesq coefficient, maximum velocity and average flow velocity, as demonstrated by Barbé et al. (1991).

$$e^{\lambda_{10}-1} \cdot \left\{ e^{\lambda_{11} \cdot u} + \lambda_{12} \cdot [W(u)] \right\} = \lambda_{11} \cdot \left[ \left( \frac{\xi - \xi_0}{\xi_{\max} - \xi_0} \right) + e^{\lambda_{10}-1} \cdot \left( \frac{1}{\lambda_{11}} + \frac{2.\lambda_{12}}{\lambda_{11}^3} \right) \right] \quad (22)$$

$$W(u) = e^{\lambda_{11} \cdot u} \cdot \left( u^2 - \frac{2.u}{\lambda_{11}} + \frac{2}{\lambda_{11}^2} \right) \quad (23)$$

## RESULTS AND DISCUSSION

Table II and Figures 2 and 3 present the performance of the six entropy models applied to the three laboratory laser-Doppler experimental datasets. The weak primary statement (Equation 2) is based on the hypothesis that longitudinal velocity grows monotonically from zero at the bed to a maximum value at the water surface, i.e., the velocity dip is assumed as zero in the whole dominium, in disagreement with Literature Ardiclioglu et al. (2005), Mirauda & Russo (2019). Besides, it is also assumed that isovels are horizontal, contrasting with laboratory Steffler et al. (1985), de Araújo & Chaudhry (1998) and field Ahmadi & Maghrebi (2019) measurements. With regards to the Cartesian-coordinate entropy formulations, it can be said that model U1y performed poorly, with negative Nash-Sutcliffe coefficients (NSE) for all vertical profiles and with errors (RMSE) up to 46%. When the second constraint is added, the model (U2y) performs much better, with median NSE of +0.51 and RMSE of only 6%. The comparative results of U1y and U2y (Figures 2a and 2b, respectively) show that the addition of the second constraint definitely improves the model predictability capacity.

**Table II. Model performance of the velocity-field simulations for the 42 vertical profiles of three experiments: SRP1, SRP2, and SRP3.**

| Experiment (n.vert) | Nash-Sutcliffe coefficient (NSE) | | | | Root mean square error (RMSE) | | | |
|---|---|---|---|---|---|---|---|---|
| | Min | Average | Median | Max | Min | Average | Median | Max |
| **Model U1y: one constraint and Cartesian coordinates** | | | | | | | | |
| SRP1 (10) | -14.96 | -9.59 | -8.68 | -6.85 | 25% | 35% | 36% | 46% |
| SRP2 (16) | -9.68 | -6.43 | -6.28 | -2.84 | 17% | 25% | 26% | 31% |
| SRP3 (16) | -7.50 | -4.13 | -3.80 | -2.34 | 10% | 15% | 14% | 19% |
| All (42) | -14.96 | -6.31 | -5.99 | -2.34 | 10% | 23% | 22% | 46% |
| **Model U2y: two constraints and Cartesian coordinates** | | | | | | | | |
| SRP1 (10) | -18.37 | -3.76 | -0.17 | 0.98 | 2% | 14% | 13% | 32% |
| SRP2 (16) | -17.37 | -1.13 | 0.82 | 0.96 | 2% | 7% | 4% | 22% |
| SRP3 (16) | -24.16 | -2.57 | 0.48 | 0.97 | 1% | 7% | 5% | 18% |
| All (42) | -24.16 | -2.31 | 0.51 | 0.98 | 1% | 9% | 6% | 32% |
| **Model U3y: three constraints and Cartesian coordinates** | | | | | | | | |
| SRP1 (10) | -21.72 | -4.88 | -0.73 | 0.90 | 4% | 17% | 15% | 34% |
| SRP2 (16) | -29.12 | -3.20 | 0.05 | 0.68 | 6% | 12% | 10% | 28% |
| SRP3 (16) | -16.76 | -1.29 | 0.79 | 0.94 | 2% | 6% | 3% | 15% |
| All (42) | -29.12 | -2.87 | 0.32 | 0.94 | 4% | 11% | 9% | 34% |
| **Model U1ξ: one constraint and curvilinear coordinates** | | | | | | | | |
| SRP1 (10) | -29.14 | -15.93 | -15.80 | -6.99 | 34% | 43% | 45% | 51% |
| SRP2 (16) | -14.62 | -8.41 | -8.59 | -3.53 | 15% | 27% | 28% | 32% |
| SRP3 (16) | -13.45 | -6.43 | -6.46 | -2.79 | 11% | 17% | 18% | 21% |
| All (42) | -29.14 | -9.45 | -8.41 | -2.79 | 11% | 27% | 26% | 51% |
| **Model U2ξ: two constraints and curvilinear coordinates** | | | | | | | | |
| SRP1 (10) | -0.86 | 0.48 | 0.85 | 0.95 | 3% | 6% | 5% | 10% |
| SRP2 (16) | -0.34 | 0.68 | 0.83 | 0.97 | 2% | 4% | 4% | 6% |
| SRP3 (16) | 0.48 | 0.75 | 0.76 | 0.90 | 2% | 3% | 3% | 5% |
| All (42) | -0.86 | 0.66 | 0.77 | 0.97 | 2% | 4% | 4% | 10% |
| **Model U3ξ: three constraints and curvilinear coordinates** | | | | | | | | |
| SRP1 (10) | -1.70 | 0.14 | 0.55 | 0.76 | 6% | 9% | 9% | 12% |
| SRP2 (16) | -11.51 | -1.02 | 0.18 | 0.44 | 7% | 10% | 9% | 18% |
| SRP3 (16) | -0.01 | 0.65 | 0.79 | 0.94 | 2% | 3% | 3% | 4% |
| All (42) | -11.51 | -0.11 | 0.42 | 0.94 | 2% | 7% | 8% | 18% |

**The term "n.vert" means the number of vertical profiles. Data source: Steffler et al. (1983).**

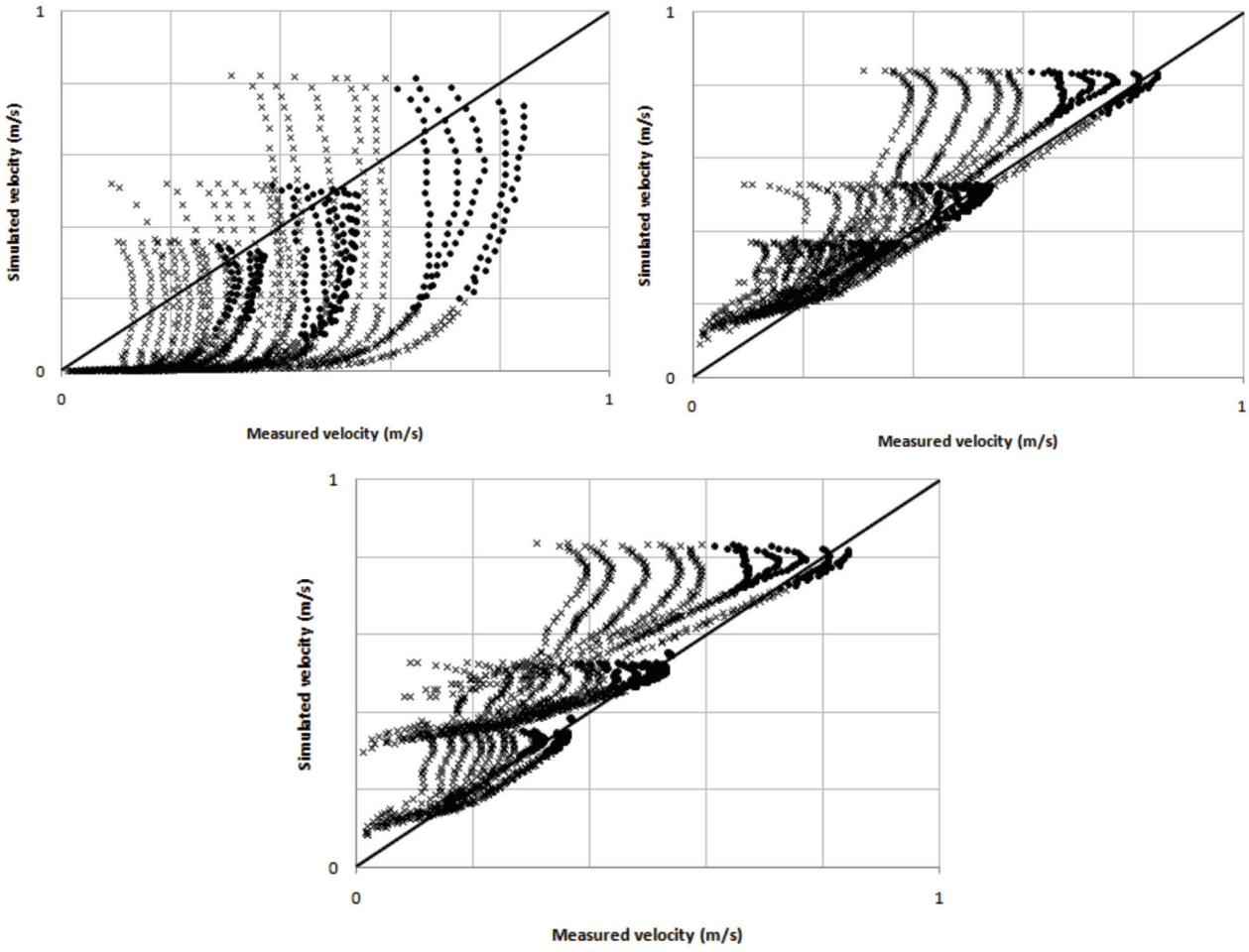


**Figure 2. Correlation between measured (experiments SRP1, SRP2, and SRP3: Steffler et al. 1985) and simulated velocities using the weak statement (Cartesian coordinates): (a) model U1y; (b) model U2y; and (c) model U3y. The black dots represent points located further than 3% of the channel width from the sidewalls and from the channel bed, whereas the *plus* (+) signs refer to the points located elsewhere (near wall and/or channel bed).**

However, although the three-constraint U3y model (Figure 2c) performs better than U1y, it is worse than U2y: the median NSE decreases to +0.32 and RMSE raises to 9%. This shows that the addition of extra information (third constraint) improves the model, i.e., U3y performs better than the one-constraint model U1y. The use of the numerical solution, as in Barbé et al. (1991), however, influences negatively this improvement and even limits its performance (U3y is worse than U2y).

The same pattern can be observed for the three curvilinear-coordinate entropy models. The one-constraint model U1ξ has negative Nash-Sutcliffe coefficients for all vertical profiles and errors ranging from 11% to 51%. According to the results, this is the worst model (Figure 3a) among the researched ones, with median NSE below –8. The inclusion of the second constraint notably improves model capacity: the median NSE is positive for all experiments (greater than +0.76) and the median error is as low as 4%. Comparison of Figures 3a and 3b also shows an improvement of the model when the second constraint is considered. The combination of two constraints and the curvilinear-coordinate

system generates the best entropy model (U2ξ) among the investigated options. As in the weak primary-statement models, the inclusion of a third constraint yields a model (U3ξ) with a performance surpassing that of the one-constraint model (U1ξ) and raising the median Nash-Sutcliffe coefficient from –8.41 to +0.42; median error decreases from 26% to 8%, at the same time. In fact, U3ξ is the second best entropy model among the six tested formulations. Comparing Figures 3b and 3c reveals, however, that the U3ξ model does not represent the velocity-field data as well as the U2ξ model: in the balance between the advantage of having more information (third constraint) and the disadvantage of using a numerical solution, the negative aspect prevails. Besides, the numerical solution of the three-constraint models generated instability during the parameterization process (when calculating the Lagrange multipliers) which augmented the computational effort. This was observed in all experiments.

In order to investigate whether it is better to use a three-constraint model with a weak primary statement, or a two-constraint model with a strong statement, the marginal improvement of model U2y was compared with models U3y and U2ξ, respectively. The results show that the third constraint has the drawback of demanding a numerical solution, which increases computer time demand and worsens result accuracy; whereas the combination of an analytical-solution system (two constraints) with a strong statement yields a high-performance model. In fact, model U2ξ has a median NSE of +0.77 against +0.32 of U3y, whereas the median error of U3y (9%) is more than twice that of U2ξ (4%). When using the Student t-test (5% significance) to compare the NSE between U2y and U3y, it shows that both models are statistically equal, i.e., the simple addition of the third constraint does not upgrade the model capability because the three-constraint model demands a numerical solution of its equations. Contrastingly, when we apply the t-test to compare U2y and U2ξ, the results indicate that they are statistically different, with clear superiority of the latter: NSE improves from +0.51 to +0.77 and the average NSE raises from –2.31 to +0.66.

Figure 4 provides a synthesis of the performance of the models, considering only the number of constraints. It is clear that models with only one constraint perform much worse than those with two or three constraints (negative NSE and high RMSE). It is also visible that, despite the similarity of the results of the U2 and U3 models, the performance of U2 models is higher. Besides, if one compares the Nash-Sutcliffe coefficient for the best-fit U2 model (U2ξ) with that for the best-fit U3 model (U3ξ) using the Student t-test with a 5% significance, it shows that the models are statistically different and that U2ξ is clearly superior.

From Figures 2, 3, and 4 it is noteworthy that all models have flaws in representing some vertical profiles. Figure 5 indicates that mal-represented profiles are those near the sidewall, which can be confirmed by Figures 2 and 3. This flaw occurs even when the best models (U2ξ, U3ξ, and U2y) are used. For the near-wall verticals, for example, model U2y exhibits an NSE coefficient as low as -24 and a corresponding error RMSE as high as 32%. It still mimics accurately the measured data for more centralized verticals with a maximum NSE of 0.98 and minimum RMSE of 1% (Table II). This fact is directly connected to sidewall proximity, as shown in de Araújo & Chaudhry (1998) and in Greco (2015). The analyzed entropy models perform well for profiles further than 3% of the channel width, however, as can be depicted from Figure 5. Nonetheless, the use of curvilinear coordinates improves model performance for

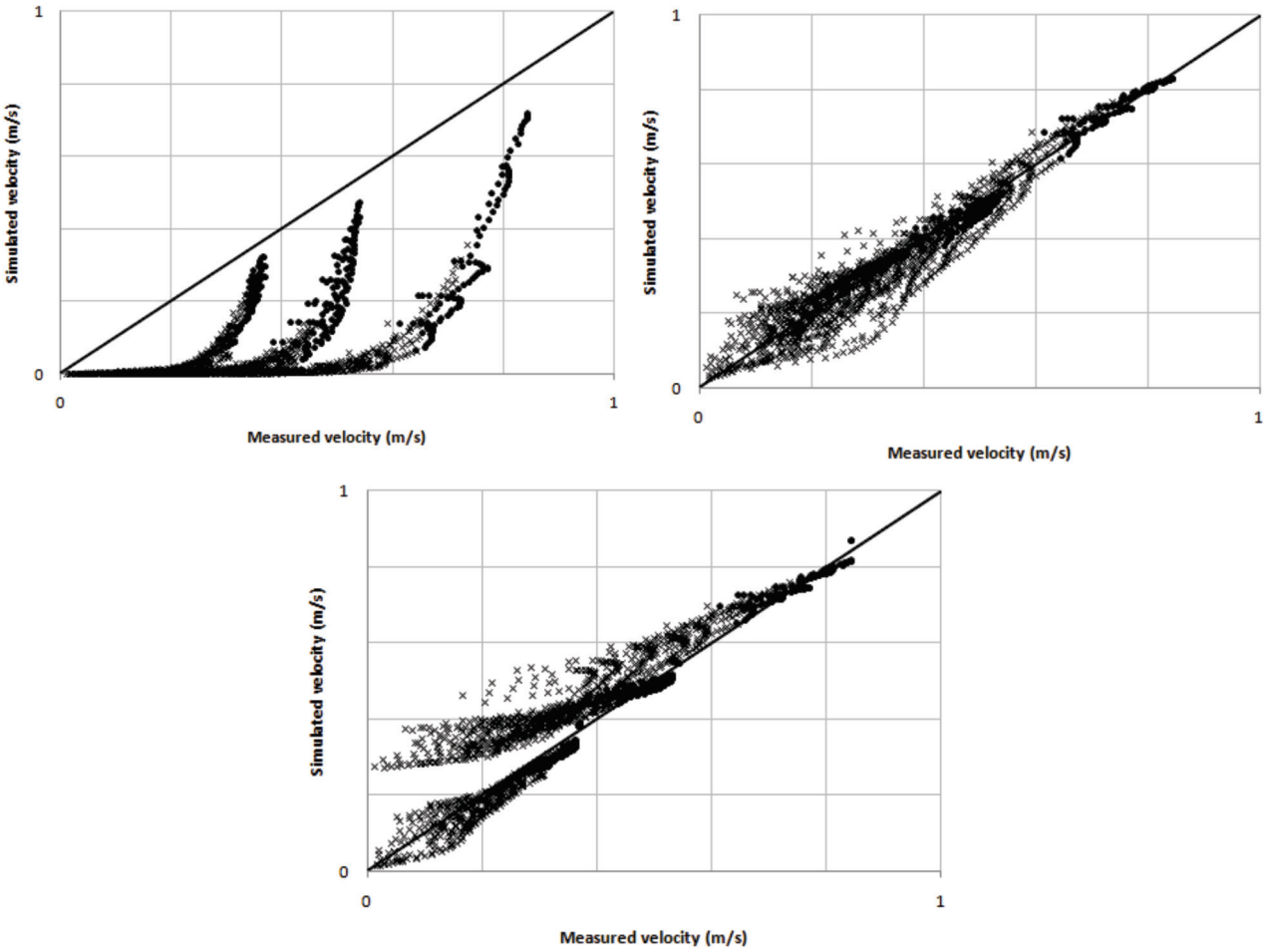


**Figure 3. Correlation between measured (experiments SRP1, SRP2, and SRP3: Steffler et al., 1985) and simulated velocities using the strong statement (curvilinear coordinates): (a) model U1ξ; (b) model U2ξ; and (c) model U3ξ. The black dots represent points located further than 3% of the channel width from the sidewalls and from the channel bed, whereas the *plus* (+) signs refer to the points located elsewhere (near wall and/or channel bed).**

one, two, or three constraints, particularly in the vicinity of walls. This is most emphasized in the routine U2ξ, the best-performance model in the context of this research. In fact, Chen & Chiew (2004) experimentally observed the significant velocity gradient near the channel bed; Patel et al. (2016) showed that near-wall gradients influence turbulence and quasi-streamwise vortices in channel flow; whereas Ninto & Garcia (2006) emphasized the influence of the near-wall flow on sediment re-suspension, which was confirmed by Mohan et al. (2019). These features have been experimentally observed, among others, by Steffler et al. (1985), de Araújo & Chaudhry (1998), Chen & Chiew (2004), Ninto & Garcia (2006), Birch & Morrison (2010), Howes & Sanders (2011), and Patel et al. (2016). The curvilinear coordinates, as proposed by Chiu (1988), succeed in simultaneously representing the effect of both sidewalls and of channel bed, which explains why the 'ξ' models out-perform the Cartesian-coordinate ('y') ones.

## CONCLUSIONS

From the results, it can be concluded that entropy hydrodynamic models with two

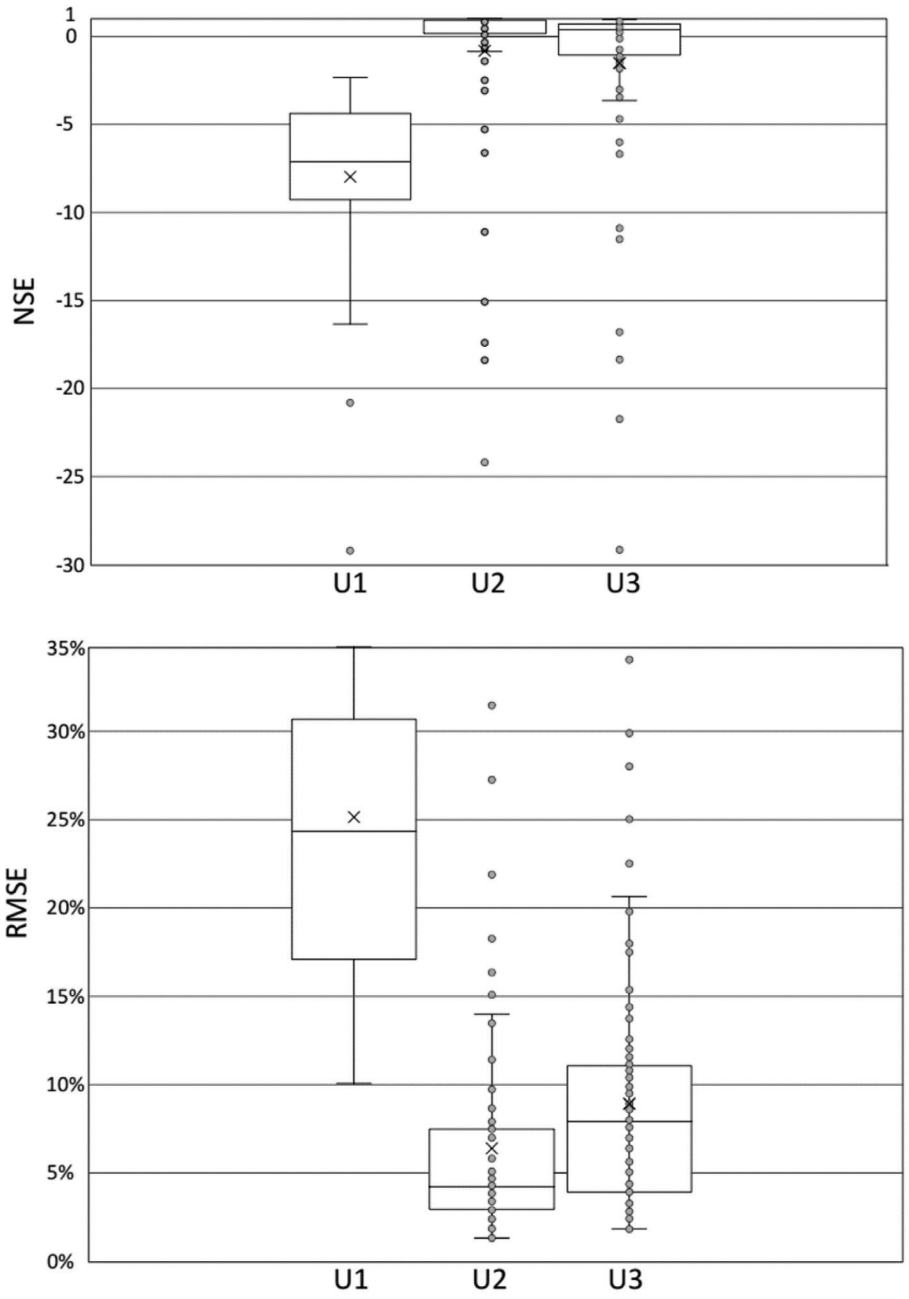


**Figure 4. Performance of the entropy models according to the number of constraints. (a) Nash-Sutcliffe coefficient (NSE). (b) Root mean square error (RMSE). U1 encompasses the joint results for both U1y and U1ξ models. The same applies to U2 and U3.**

constraints performs better than those with one constraint, and that the second constraint includes relevant information for the system. Contrastingly, models with three constraints perform worse than those with two constraints, showing that the loss of information due to the use of numerical solutions may surpass the gain of information due to the third constraint. The best-performance entropy model (that with two constraints and curvilinear coordinates – U2ξ)

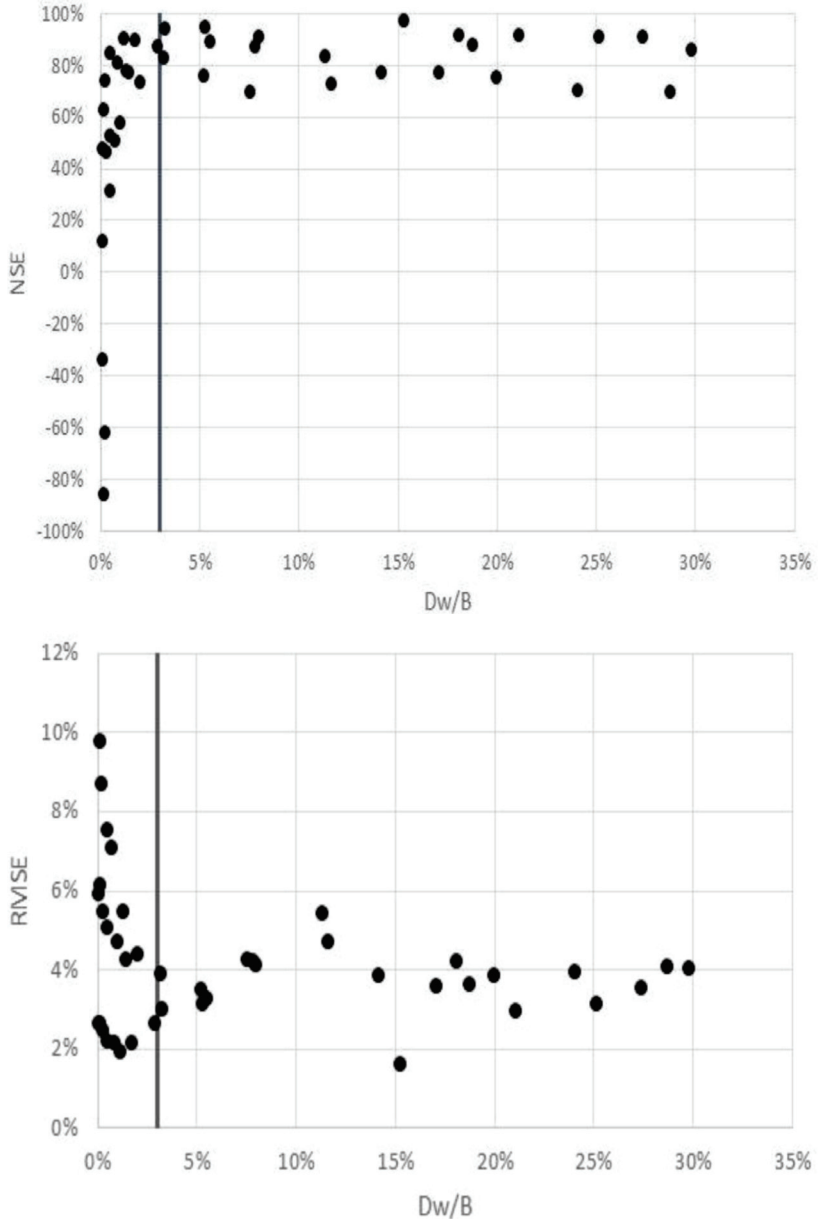


**Figure 5. Performance of the entropy models as a function of the distance to the vertical wall. Each point refers to a measured vertical profile. The variable $D_w$ is the horizontal distance to the nearest wall and B is channel width. Vertical grey lines indicate the threshold of the model applicability ($D_w/B > 0.03$). (a) Nash-Sutcliffe coefficient (NSE). (b) Root mean square error (RMSE).**

was able to mimic well the accurately-measured laboratory data for vertical profiles further than 3% of channel width. For vertical profiles closer than 3% of the width, the studied models do not perform well because the specific information concerning the prevailing processes is neither provided in the primary statement, nor in the constraints. In the present case study, the replacement of a weak primary-statement (use of Cartesian coordinates, Equation 2) by a strong one (use of curvilinear coordinates, Equation 18) brings more information to the system than the inclusion of a third constraint.

## Acknowledgments

This research was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, PRINT Grant 88881.311770/2018-01), and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico(CNPq, Grant 407999/2016-7). The authors also acknowledge the Federal Rural University of Pernambuco UFRPE for the support to the first author.

## REFERENCES


AHMADI A & MAGHREBI MF. 2019. A robust approach for rating curves estimation in open channels using isovel contours. Int J River Basin Manag 11: 1814-2060.

AL-HAMDAN OZ & CRUISE JF. 2010. Soil moisture profile development from surface observations by principle of maximum entropy. J Hydrol Eng 15: 327-338.

ALVISI S, BARBETTA S, FRANCHINI M, MELONE F & MORAMARCO T. 2014. Comparing grey formulations of the velocity-area method and entropy method for discharge estimation with uncertainty. J Hydroinform 16(4): 797-811.

ARDICLIOGLU M, DE ARAÚJO JC & SENTURK AI. 2005. Applicability of velocity distribution equations in rough-bed open-channel flow. Lhbl 4: 73-79.

BANAVAR JR, MARITAN A & VOLKOV I. 2010. Applications of the principle of maximum entropy from physics to ecology. J Phys Conf Ser 22(6): 13.

BARBÉ DE, CRUISE JF & SINGH VP. 1991. Solution of three-constraint entropy-based velocity distribution. J Hydraul Eng 117(10): 1389-1396.

BEAKES MP, MOORE JW, RETFORD N, BROWN R, MERZ JE & SOGARD SM. 2014. Evaluating statistical approaches to quantifying juvenile Chinook salmon habitat in a regulated California river. River Res Applic 30: 180-191.

BIRCH DM & MORRISON JF. 2010. Similarity of the streamwise velocity component in very rough-wall channel flow. J Fluid Mech 668: 174-201.

CHEN X & CHIEW Y-M. 2004. Velocity distribution of turbulent open-channel flow with bed suction. J Hydraul Eng 130(2): 140-149.

CHENG K, WEI S, FU Q, PEI W & LI T. 2019. Adaptive management of water resources based on an advanced entropy method to quantify agent information. J Hydroinform 21(3): 381-396.

CHIU CL. 1987. Entropy and probability concepts in hydraulics. J Hydraul Eng 113(5): 583-599.

CHIU CL. 1988. Entropy and 2-D velocity distribution in open channels. J Hydraul Eng 114(7): 738-756.

CHIU CL. 1989. Velocity distribution in open channels flow. J Hydraul Eng 115(5): 576-594.

CHIU CL. 1991. Application of entropy concept in open-channel flow study. J Hydraul Eng 117(5): 615-628.

CHIU CL & CHIOU JD.1986. Structure of 3-D flow in rectangular open channels. J Hydraul Eng 112(11): 1050-1067.

CHIU CL & HSIUNG DE. 1981. Secondary flow, shear stress and sediment transport. J Hydraul Eng 107(7): 879-898.

CHIU CL, HSU SM & TUNG NC. 2005. Efficient methods of discharge measurements in rivers and streams based on the probability concept. Hydrol Process 19(20): 3935-3946.

COBO JLM, MENDIZÁBAL R, MIQUEL A, BERNA C & ESCRIVÁ A. 2017. Use of the principles of maximum entropy and maximum relative entropy for the determination of uncertain parameter distributions in engineering applications. Entropy 19(9): 486-522.

COLE TM & WELLS SA. 2006. CE-QUAL-W2: a two-dimensional, laterally averaged, hydrodynamic and water quality model, version 3.5. Instruction Report EL-06-1. US Army Engineering and Research Development Center, Vicksburg, USA.

DE ARAÚJO JC & CHAUDHRY FH. 1998. Experimental evaluation of a 2D entropy model in open channels. J Hydraul Eng 124(10): 1064–1068.

DE ARAÚJO JC. 2007. Entropy-based equation to assess hillslope sediment production. Earth Surf Proc Land 32: 2005-2018.

DHI. 1992. MIKE 11 User Manual. Danish Hydraulic Institute, Horsholm.

DHI. 2017. MIKE 11: A Modelling System for Rivers and Channels. User Guide. Danish Hydraulic Institute, Horsholm, 510 p.

FARINA G, ALVISI S, FRANCHINI M & MORAMARCO T. 2014. Three methods for estimating the entropy parameter M based on a decreasing number of velocity measurements in a river cross-section. Entropy 16: 2512-2529.

FONTANA N, MARINI G & PAOLA F. 2013. Experimental assessment of a 2-D entropy-based model for velocity distribution in open channel flow. Entropy 15(3): 988-998.

FURBISH DJ, HAFF PK, SCHMEECKLE MW, SCHUMER R & FATHEL SL. 2016. Probability distributions of bed load particle velocities, accelerations, hop distances and travel times informed by Jaynes´s principle of maximum entropy. J Geophys Res Earth Surf 121: 1373-90.

GRECO M. Effect of bed roughness on 1-D entropy velocity distribution in open channel flow. 2015. Hydrol Res 46(1): 1-10.

GRECO M & MARTINO G. 2018. 1-D versus 2-D entropy velocity law for water discharge assessment in a rough ditch. Entropy 20: 638.

GUPTA A & GOVINDARAJU RS. 2019. Propagation of structural uncertainty in watershed hydrologic models. J Hydrol 575: 66-81.

HARMANCIOGLU NB & SINGH VP. 1998. Entropy in environmental and water resources. Encyclopedia of Earth Science. Springer, Dordrecht. Boston, p. 225-241.

HOWES DJ & SANDERS BF. 2011. Velocity contour weighting method. I: algorithm development and laboratory testing. J Hydraul Eng 137(11): 1359-1367.

JAYNES ET. 1957a. Information theory and statistical mechanics, I. Phys Rev 106: 620-630.

JAYNES ET. 1957b. Information theory and statistical mechanics, II. Phys Rev 108: 171-190.

JIANG Y, LI B & CHEN J. 2016. Analysis of the velocity distribution in partially-filled circular pipe employing the principle of maximum entropy. PLoS One 11 (3): 1-17.

JIN H, HE W, LIU Q, WANG J & FENG G. 2016. The applicability of research on moving cut data-approximate entropy on abrupt climate change detection. Theor Appl Climatol 124: 475-486.

KNIGHT D. 2013. Hydraulic problems in flooding: from data to theory and from theory to practice. In: Rowiński P (Ed), Experimental and Computational Solutions of Hydraulic Problems. GeoPlanet: Earth and Planetary Sciences. Springer, Berlin, Heidelberg, 425 p.

LUO H & SINGH VP. 2011. Entropy theory for two-dimensional velocity distribution. J Hydrol Eng 16(4): 303-315.

MÉLÈSE V, BLANCHET J & MOLINIÉ G. 2018. Uncertainty estimation of intensity-duration-frequency relationships: a regional analysis. J Hydrol 558: 579-591.

MIRAUDA D & RUSSO MG. 2019. Information entropy theory applied to the dip-phenomenon analysis in open channel flows. Entropy 21(6): 554.

MISHRA S. 2009. Uncertainty and sensitivity analysis techniques for hydrologic modelling. J Hydroinform 11(3-4): 282-296.

MOHAN S, KUMBHAKAR M, GHOSHAL K & KUMAR J. 2019. Semianalytical solution for simultaneous distribution of fluid velocity and sediment concentration in open-channel flow. J Eng Mech 145(11).

MORAMARCO T, CORATO G, MELONE F & SINGH VP. 2013. An entropy-based method for determining the flow depth distribution in natural channels. J Hydrol 497: 176-188.

NINTO Y & GARCIA MH. 1996. Experiments on particle—turbulence interactions in the near–wall region of an open channel flow: implications for sediment transport. J Fluid Mechanics 326: 285-319.

PATEL A, BOERSMA BJ & PECNIK R. 2016. The influence of near-wall density and viscosity gradients on turbulence in channel flows. 2016. J Fluid Mech 809: 793-820.

SHANNON CE. 1948. Mathematical theory of communications, I and II. Bell Syst 27: 379-443.

SHORE JE & JOHNSON RW. 1980. Properties of cross-entropy minimization. IEEET Inform Theory 17: 472-482.

SHIONO K & KNIGHT DW. 1991. Turbulent open channel flows with variable depth across the channel. J Fluid Mech 222: 617-646.

SHRESTHA B, COCHRANE TA, CARUSO BS, ARIAS ME & PIMAN T. 2016 Uncertainty in flow and sediment projections due to future climate scenarios for the 3S Rivers in the Mekong Basin. J Hydrol 540: 1088-1104.

SINGH VP. 2013. Entropy theory and its application in environmental and water engineering, ASCE Press., New York, 662 p.

SINGH VP. 2014. Entropy theory in hydraulic engineering: an introduction. ASCE Press. New York, 656 p.

SINGH VP, MARINI G & FONTANA N. 2013. Derivation of 2D power-law velocity distribution using entropy theory. Entropy 15: 1221-1231.

STEFFLER PM, RAJARATNAM N & PETERSON AW. 1985. LDA measurements in open channel. J Hydraul Eng 111(1): 119-130.

STEFFLER P & BLACKBURN J. 2002. River2D: Two-dimensional depth averaged model of river hydrodynamics and fish habitat. Introduction to depth averaged modeling and user's manual. Technical Report. Edmonton, Canada. University of Alberta, 119 p. http://www.river2d.ualberta.ca/download.htm.

TETRA TECH. 2007. The Environmental Fluid Dynamics Code User Manual USEPA Version 1.01.

THANH VQ, ROELVINK D, VAN DER WEGEN M, REYNS J, KERNKAMP H, VAN VINH G & LINH VTP. 2020. Flooding in the Mekong Delta: the impact of dyke systems on downstream hydrodynamics. Hydrol Earth Syst Sci 24: 189-212.

TAPOGLOU E, VAROUCHAKIS EA, TRICHAKIS IC & KARATZAS GP. 2019. Hydraulic head uncertainty estimations of a complex artificial intelligence model using multiple methodologies. J Hydroinform 22(1): 205-218

TORRES-BEJARANO F, PADILLA J, RODRÍGUEZ-CUEVAS C & CANTERO R. 2015. Hydrodynamics modelling utilizing the EFDC Explorer model for the sustainable management of Canal del Dique-Guajaro hydrosystem, Colombia. WIT Transactions on The Built Environment, 168. doi:10.2495/SD150371.

WALDRIP SH, NIVEN RK, ABEL M & SCHLEGEL M. 2016. Maximum entropy analysis of hydraulic pipe flow networks. J Hydraul Eng 142(9): 04016028.

ZUO D, HOU W & HU J. 2017. An entropy-based investigation into bivariate drought analysis in China. Entropy 9(9): 1-17.

**How to cite**

SILVA FILHO AV, DE ARAÚJO JC & RAABE A. 2020. Trade-off between number of constraints and primary-statement robustness in entropy models: the case of the open-channel velocity field. An Acad Bras Cienc 92: e20200594. DOI 10.1590/0001-3765202020200594.

*Manuscript received on April 21, 2020;*
*accepted for publication on July 24, 2020*

**ANTONIO VIANA DA SILVA FILHO**[1]
https://orcid.org/0000-0002-7926-3609

**JOSÉ CARLOS DE ARAÚJO**[2]
https://orcid.org/0000-0002-2649-151X

**ARMIN RAABE**[3]
https://orcid.org/0000-0002-8819-3294

[1]Federal Rural University of Pernambuco (UFRPE), Av. Gregório Ferraz Nogueira, Caixa Postal 063, 56909-535 Serra Talhada, PE, Brazil

[2]Federal University of Ceará, Department of Agricultural Engineering, Campus do Pici, Bloco 804, Pici, 60455-760 Fortaleza, CE, Brazil

[3]Leipzig University, Leipzig Institute of Meteorology, Stephanstr. 3, 04103, Leipzig, Saxonia, Germany

Correspondence to: **Antonio Viana da Silva Filho**
*E-mail: viana.filho.ufrpe@hotmail.com*

## Author contributions

Antonio Viana da Silva Filho codified the models; developed the model with three restrictions and curvilinear coordinates (U3ξ); parameterized and simulated the models; co-wrote; and formatted the text. José Carlos de Araújo designed the research project; reviewed the models, their parameterization and their results; and co-wrote the text. Armin Raabe reviewed the models and their results; and co-wrote the text.