Using maximum entropy for discharge calculation in open channel flow

V. E. M. G. Diniz[1], V. O. C. Pereira[1], L. P. Bispo[1] & P. A. Souza[2]
[1]Civil Engineering Department, Universidade Nove de Julho, Brazil
[2]Department of Hydraulic and Environmental Engineering of Polytechnic School, São Paulo University, Brazil

Abstract

This paper uses the maximum entropy model to calculate the discharge in open channels. Rivers and an artificial rectangular channel are considered in this paper. The maximum entropy model needs to gauge the flow velocity only in three points to calculate the discharge of any given open channel. The flow velocity needs to be gauged on the river thalweg. A genetic algorithm was used to determine the three parameters of the maximum entropy equation. The discharge calculation results obtained by the maximum entropy model aided by the genetic algorithm were compared to the discharge calculation results obtained by the traditional (using a rotor current meter) and acoustic methods (Acoustic Doppler Current Profiler or ADCP). It was concluded that the maximum entropy model is reliable because it yields results very close to the results obtained by the traditional and acoustic methods. It was also concluded that it makes the discharge measurement easier in any given open channel.
Keywords: maximum entropy, discharge calculation, genetic algorithm, open channel, flow velocity.

1 Introduction

It was always important the discharge calculation of any given river and nowadays it's becoming more and more important due to the occurrence of floods and drought periods. The floods and droughts disturb very much the human activities. The main objective of calculating discharges in rivers is the lifting of the relationship between stage and discharge (rating curve). Through

WIT Transactions on Ecology and The Environment, Vol 171, © 2013 WIT Press
www.witpress.com, ISSN 1743-3541 (on-line)
doi:10.2495/WRM130271

this relationship, it's generated discharge time series, essential to accomplish statistical analysis and to design structures and waterworks.

The concept of entropy was used to substantiate the connection between the deterministic and probabilistic worlds, the latter being unfamiliar to hydraulic engineers Chiu in Moraes [1].

Entropy is the cumulative probabilities function that measures the information generated and transmitted by an event, through the weighted sum by the probability of how many times an event has occurred.

According to the principle of entropy, in equilibrium state, a system tends to maximize the entropy on the entropy previously contained. By maximizing the entropy, it is estimated that the most likely event is the one that will happen. This principle can be used to model the most probable distribution of states of a system Chiu in Moraes [1].

From the concepts of entropy and information theory, Chiu in Moraes [1] rewrote the entropy equation and developed equations for open-channel flows from a conceptual form to the velocity distribution profile, shear stress distribution and sediment concentration distribution. Chiu in Moraes [1] used the method of listing the hypothesis with the highest probability of occurrence, that is, the method of maximizing the entropy functional for the development of these equations.

Chiu in Moraes [1] used the maximum entropy model and presented a relation between the entropy parameter (M) and the velocity distribution profile of any given flow in an open-channel cross section. Chiu in Moraes [1] demonstrated the usefulness of the entropy parameter (M) and concluded that "M" as a new hydraulic parameter indicates the importance and value of the information given by the location and magnitude of maximum velocity in a cross section.

Genetic algorithms are part of evolutionary algorithms and allow a population composed of many individuals to evolve under specific selection rules to a state that generally maximizes the fit of the objective function (cost function). The method was developed by John Holland (1975) during the 1960s and 70s and was popularized by one of his students, David Goldberg, who solved a complex problem involving gas transmission control through pipelines in his dissertation of 1989 (Diniz *et al*. [2]).

Genetic algorithms have some advantages like being able to make optimizations using continuous or discrete parameters, not requiring derivative information, searching simultaneously from a large number of samples of the solution space, dealing with a large number of parameters to be well suited for parallel programming, yielding a list of optimum parameters, rather than a single solution, allowing the parameters encoding so that the optimization is done with the encoded parameters, optimizing parameters in solution spaces extremely complexes, being able to avoid a local minimum and working with numerically generated data, experimental data or analytic functions (Diniz *et al*. [2]).

2 Literature review

Chiu *et al.* [3] presented a new velocity distribution equation based on the probability and entropy concepts as an alternative to the existing power law and universal law equations for pipe-flow studies. The probability concept was used to formulate the velocity distribution and the maximum entropy principle was used to derive the probability law governing the velocity distribution. The new equation, with only two parameters (the entropy parameter (M) and the maximum velocity), is applicable to the entire flow field in a pipe, from laminar to turbulent and regardless whether the pipe is smooth or rough. A new equation for the friction factor (f) was also derived based on the new velocity distribution equation, the flow shear stress and the universal head loss formula. This equation relates the friction factor to the velocity distribution through the entropy parameter (M). A statistical method was also presented to analyze and quantify the uncertainty in the velocity distribution due to uncertainties in the related parameters or variables.

Zhuang *et al.* [4] developed a four-stage framework for optimal planning of regional water distribution system. Most work of the first two stages is left to the decision makers of municipal planning bureau and water utilities. Dijkstra Algorithm is used at the third stage to find the least cost for layout of transmission mains between adjacent subregions. At the fourth stage, a model for diameter (the decision variables) optimization of regional water distribution systems was proposed with the objectives of minimal investment and operation costs and maximum entropy reliability. The shuffled frog-leap algorithm (SFLA), linked to EPANET as the hydraulic engine, was adopted to solve the model. The developed framework was applied and demonstrated on planning of a real regional water distribution system in China. The demonstrated case could serve as an example to other similar planning of regional water distribution systems.

Wang and Duan [5] used entropy as an indirect index to measure the reliability of gas distribution networks. The relative entropy which is the ratio of actual entropy and maximum entropy was used as the reliability constraint to find the optimal diameters of a real gas distribution network from a reliability and economy point of view with different diameter combinations. To obtain the optimal diameters it was used an improved genetic algorithm with a new set of suitable problem-specific genetic operators and fitness functions. The use of a real gas distribution network as example proved that the total investment cost is a little larger in the new design scheme based on reliability constraint, but both the economy and reliability of the gas distribution network are overall considered, so the optimal scheme is more close to the practice.

Jamasb *et al.* [6] prepared an optimization problem using genetic algorithm and developed a computer program and linked it to EPANET to calibrate the model. The problem optimization was prepared by using the genetic algorithm tool box in MATLAB7. The authors used the pipe friction factor (Hazen William Coefficient), nodal consumptions, combinations of both and pipe diameters as variables to be determined and applied the method to a hydraulic network for different consumption scenarios. Furthermore, by studying the common operating conditions in the example hydraulic network demonstrated that, synchronized adjusting demands and roughness as decision variables and using the observations related to the fire fighting condition (with maximum demand) lead to more precise results in calibration of model and system simulations. The developed program offers a powerful approach to decrease the effects of uncertainties.

Kim [7] analyzed the oscillatory flows in pipeline systems due to excitation by valve operation by the impulse response method. The pressure head or discharge responses were derived at any point in a simple and theoretical pipeline system including the impact of leakage (single or multiple leaks) based on the impulse response method. Frequency-dependent friction was used to simulate the impact of unsteady friction for the laminar flow. Extensive development of the impulse response method was made by considering the sources of friction associated with the local and convective acceleration of velocity for turbulent flow. A genetic algorithm was integrated into the impulse response method to calibrate the location and the quantity of leakage. Transient analysis showed that the extended impulse response method provides matched results to the experimental time series and simulations of the method of characteristics with the assumption of unsteady friction.

Kim [8] developed an alternative approach, namely the address-oriented impedance matrix method (AOIMM), to the conventional methods of transient computation and parameter calibration for a heterogeneous pipeline network for a more robust calibration of a heterogeneous and multi-looped pipe network system. A genetic algorithm was incorporated in address generation of the impedance matrix for leak location as a search engine. Simultaneous calibration of wave speed, friction, leakage location and quantity is one distinct feature of the AOIMM, which means that the impact of various factors can be effectively evaluated. The potential of the proposed calibration algorithm was demonstrated when it was applied to a fictitious heterogeneous pipe network system. The computational costs of the AOIMM were substantially lower than those of numerical approaches and the address generation module allows for the adaptive calibration of leakage as well as for efficient allocation of memory storage. Transient analysis shows that the AOIMM results match those of simulations of MOC (Method of Characteristics) in several flow regimes.

Prasad [9] presented a method for optimum design of pumped hydraulic networks with various elements such as pipes, pumps and tanks. Prasad [9] used a steady state genetic algorithm and the objective function was to minimize the total cost, which includes capital cost and energy cost. Prasad [9] introduced a new approach for reservoir sizing, which does not require explicit consideration of some of the operational constraints, such as the limits on maximum and minimum tank levels. Prasad [9] also accomplished extended period simulation for every trial solution in the optimization process and it permitted accurate estimation of energy costs. Prasad [9] applied the proposed model to a well-known hydraulic network design example and studied two cases. Comparing the designs obtained by the proposed model to those proposed by other researchers under similar performance conditions but with different reservoir sizing methods, Prasad [9] concluded that the proposed method was found to be superior not only in terms of cost (found cheaper designs) but also in terms of hydraulic performance. The yielded results showed the advantage of using the proposed model to obtain optimum or near optimum solutions.

3 Method

It will be present a method to calculate the discharge in any given open channels (this paper focuses the rivers and an artificial rectangular channel) by gauging the flow velocity only in three points on the river thalweg. It's necessary to gauge the velocity in three points because the problem has three unknowns and three equations are necessary in order to solve it. The method uses the maximum entropy model and a genetic algorithm to determine the three parameters (unknowns) of the maximum entropy equation. The three parameters are the maximum flow velocity (U_{max}), the entropy parameter (M) and the water depth (h) where the maximum flow velocity occurs. Figure 1 depicts U_{max} and h. It's not possible to depict M. The water depth was gauged from the surface to the river bed as depicted in figure 2. It was necessary to use the genetic algorithm

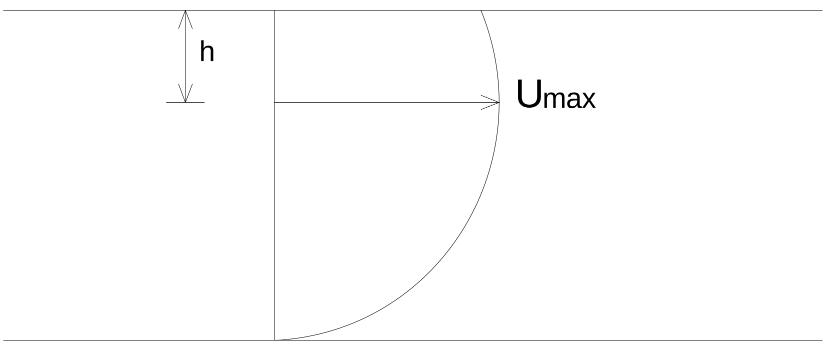

Figure 1: Water profile of any given open channel.

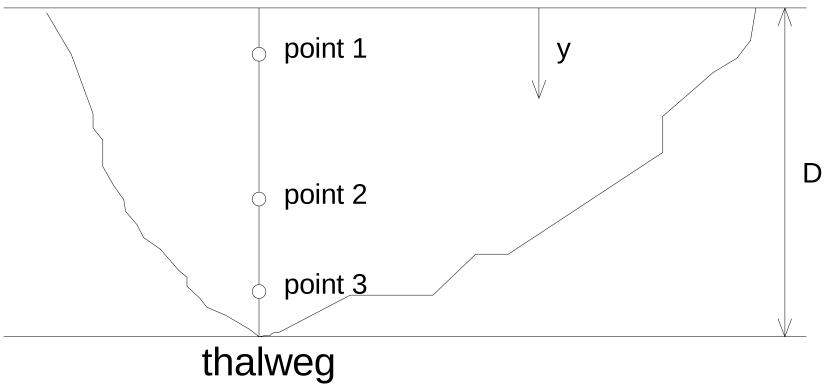

Figure 2: Cross section of any given open channel.

because experience shows it's very difficult to determine the parameters of the maximum entropy equation and it would take a lot of time to accomplish it without a computational tool. Besides being a tool to make this work easier, the genetic algorithm takes a short period of time to find the three parameters of the maximum entropy equation.

To develop the method to calculate the discharge in open channels, the authors used the following equations:

$$U(y_i) = \frac{U_{max}}{M} \ln\left[1 + (e^M - 1)\left(\frac{y_i}{D-h}\right)\exp\left(1 - \frac{y_i}{D-h}\right)\right] \quad (1)$$

$$`\epsilon = \sum_{i=1}^{N}[U_i - U(y_i)]^2 \quad (2)$$

$$\frac{U_m}{U_{max}} = \frac{e^M}{e^m - 1} - \frac{1}{M} \quad (3)$$

$$Q = U_m A \quad (4)$$

The experience shows that the three parameters range as shown hereunder:

$$3 \leq M \leq 16 \quad (5)$$

$$0.1 \leq U_{max} \leq 8 \quad (6)$$

$$0 \leq h \leq \frac{D}{3} \quad (7)$$

At first there are no constraints or penalty function. Eqns. (5), (6) and (7) aren't constraints. These equations weren't used in the genetic algorithm.

The objective function of the genetic algorithm is given by eqn (2), where (ϵ) is the error. The unknowns of the problem are U_{max}, M and h as already stated. The genetic algorithm has to search for the values of the unknowns to calculate the theoretical flow velocity ($U(y_i)$) given by eqn (1). The flow velocity calculated by eqn (1) has to minimize the error yielded by eqn (2). After finding the unknowns, the genetic algorithm uses the reached values of U_{max} and M to calculate the flow mean velocity (U_m) through eqn (3). After calculating U_m, the genetic algorithm calculates the discharge of the river or the discharge of any given open channel through eqn (4).

The internal input data, for the genetic algorithm operators, is the population size, the number of unknowns, the substring length, the total string length, the number of generations to be calculated, the crossover probability, the mutation probability, if it's going to use elitism or not and the scaling constant. The external input data (from the open channel) is the total water depth of the open channel (D) on the thalweg, the river cross section (A), the gauged flow velocities (U_i) and its respective depths (y_i) at least in three points and the number of gauges (N). The number of gauges is an input data because it's possible to gauge the flow velocity in more than three points, but three points are sufficient. Regardless the number of gauges accomplished, it's important to remember that the gauges have to be accomplished always on the river thalweg.

It's important to notice that the purpose of this paper is to show an application of the maximum entropy model aided by a genetic algorithm to calculate discharges of any given open channel. The purpose isn't in order to explain how a genetic algorithm or the maximum entropy model works. So, it won't be emphasized the technical terms referring to the genetic algorithm and to the maximum entropy model. If the reader wishes to know more about these two tools, it's necessary to look for it in other sources.

4 Results

To generate the results, it was tested some values for the internal input data and the presented values shown just ahead yielded the best results. It was used the following internal input data for the genetic algorithm: a population size of 30 individuals, the number of unknowns was 3 (U_{max}, M and h), the substring length was 5, the total string length was 15, the number of generations was 8000, the crossover probability was 95%, the mutation probability was 2%, elitism was used and it was used a scaling constant equal to 1.5. Although other values for the internal input data were tested, it doesn't take a long period of time to find out which values yield the best results and it has to be done only once, i.e. if one is going to calculate the discharge of another open channel, the same internal input data that was used for the first one can be used. It was calculated the discharge of two Brazilian rivers, Negro and Carapa, and fourteen different discharges for the rectangular channel and the internal input data remained the same for both.

To calculate the discharge of Negro River, the genetic algorithm was run ten times and calculated ten different discharges for it. The final discharge of Negro River is the result of the arithmetic average of the ten discharges calculated by the genetic algorithm. It was done to reach an accurate result, because the genetic algorithm may yield a different result every time it's run and the idea is to show that the maximum entropy model yields reliable results and that it can be used regardless other methods. It's important to say that although it was calculated ten different discharges, each discharge yielded by the genetic algorithm each time it was run was very close to each other. The same procedure was done to calculate the discharge of Carapa River and each of the fourteen discharges of the rectangular channel.

It's important to say that although the number of generations (8000) and that the genetic algorithm was run ten times to calculate each final discharge, it didn't take a long time to accomplish it. The genetic algorithm took less than 4 seconds to calculate each discharge. So, it means that it took less than 40 seconds to calculate the ten different discharges. Comparing to the period of time that the traditional and the acoustic methods take to calculate the discharge of any given open channel, it's quite reasonable to conclude that one would save a lot of time using the maximum entropy model aided by a genetic algorithm. Mainly because it's necessary to gauge the velocity flow only in three points. The velocity gauges may be accomplished using a rotor current meter, or using the ADCP or using any other device in any given water depth. It doesn't need to be accomplished at 0.2, 0.6 and 0.8 of the total depth.

The discharge results obtained for the artificial rectangular channel by the traditional method (using a rotor current meter) and by the maximum entropy model aided by the genetic algorithm are shown in table 1. The flow velocity gauges accomplished by the rotor current meter to calculate the fourteen discharges of the rectangular channel by the traditional method were accomplished in a hydraulic laboratory of Universidade de São Pàulo (USP), Brazil. It wasn't used the ADCP to calculate the discharges of the rectangular channel.

Table 1: Discharges obtained by the traditional method Mendes [10] and by the maximum entropy model for the rectangular channel.

total depth (m)	traditional method discharges (m^3/s)	maximum entropy calculated discharges (m^3/s)	error (%)
0.2565	0.00580	0.00530	9.43
0.2654	0.00778	0.00784	0.79
0.2802	0.01306	0.01308	0.19
0.3057	0.02889	0.02885	0.13
0.3351	0.06528	0.05962	9.49
0.3380	0.04361	0.04224	3.24
0.3613	0.05611	0.05336	5.16
0.3620	0.05528	0.05719	3.34
0.3840	0.05780	0.05656	2.19
0.3882	0.06028	0.05980	0.80
0.3893	0.06250	0.06497	3.80
0.3950	0.06389	0.06371	0.28
0.3955	0.06444	0.06417	0.42
0.3976	0.06583	0.06301	4.48

As it can be noticed in table 1, in a general way, the results obtained by the maximum entropy model for the rectangular channel are very close to the results obtained by the traditional method. As already commented, it was calculated fourteen discharges and only three had an error bigger than 5%. In Brazil, the permissible error when a discharge of an open channel is calculated is at most 5%.

The discharge results obtained for Negro River by the traditional method and by the maximum entropy model aided by the genetic algorithm are shown in table 2. It wasn't used the ADCP to calculate the discharge of Negro River.

As it can be noticed in table 2 just ahead, the result obtained by the maximum entropy model is very close to the result obtained by the traditional method. The error is only 2.16%.

Table 2: Discharges obtained by the traditional method Minei [11] and by the maximum entropy model for Negro River.

Traditional method discharges (m^3/s)	Maximum entropy calculated discharges (m^3/s)	Error (%)
5750.380	5877.533	2.16

The discharge results obtained for Carapa River by the traditional method, by the acoustic method and by the maximum entropy model aided by the genetic algorithm are shown in table 3.

Table 3: Discharges obtained by the traditional and acoustic methods (Melchior [12]) and by the maximum entropy model for Carapa River.

Traditional method discharges (m^3/s)	Acoustic method discharges (m^3/s)	Maximum entropy calculated discharges (m^3/s)	Error between acoustic method and maximum entropy (%)	Error between traditional method and maximum entropy (%)	Error between traditional and acoustic methods (%)
50.690	48.760	49.167	0.83	3.10	3.96

As it can be noticed in table 3, the result obtained by the maximum entropy model is very close to the results obtained by the traditional and acoustic methods. All errors are smaller than 5% and the error between the acoustic method and the maximum entropy model is only 0.83%.

5 Conclusions

It's concluded that the maximum entropy model is a useful tool to calculate discharges in any given open channel because it yielded results very close to the ones yielded by the traditional and acoustic methods. It's also concluded that it's easier to calculate the discharge of open channels using the maximum entropy model when comparing it to the traditional and acoustic methods because it needs to gauge the flow velocity only in three points on the thalweg to calculate the discharge. It's also concluded the genetic algorithm makes easier the calculation of the parameters of the maximum entropy equation because as already commented, the experience shows it's very difficult to determine these parameters without a computational tool. Consequently, the genetic algorithm, when coupled to the maximum entropy model, makes even faster the discharge calculation of open channels. It's also important to observe that the maximum entropy model embraces all amounts of discharges as shown in the three tables. The discharge may range from some few liters per second to millions of liters per second.

6 Recommendations

It's recommended to use test the maximum entropy model aided by the genetic algorithm in pipe flow to asses if it yields reliable results at it did for open channel flows.

References

[1] Moraes, A.G., Entropia máxima na modelação do fator de atrito (f) de escoamento forçado, 2010, DoctorateThesis, Escola Politécnica, Departamento de Engenharia Hidráulica e Sanitária, Universidade de São Paulo, São Paulo, Brazil.

[2] Diniz, V.E.M.G., Souza, P.A. and Luvizotto, E., *Hydraulic Networks: Optimizing Operations Through a Genetic Algorithm*, VDM Verlag Dr. Müller:Saarbrücken, pp. 20–21, 2010.

[3] Chiu, C.-L., Lin, G.-F. and Lu, J.-M., Application of probability and entropy concepts in pipe-flow study. *Journal of Hydraulic Engineering*, **119(6)**, pp. 742–756, 1993.

[4] Zhuang, B., Zhao, X. and Gao, B., Optimal planning of regional water distribution systems: a case study. *Proceedings of the World Environmental and Water Resources Congress 2010: Challenges of Change*, ed. R. N. Palmer, Providence, USA, pp. 4293–4302, 2010.

[5] Wang, X. and Duan, C., Study on parameter optimization of gas network based on reliability. *Proceedings of the International Conference on Pipelines and Trenchless Technology*, eds. M. Najafi and B. Ma, Shanghai, China, pp. 695–702, 2009.

[6] Jamasb, M., Tabesh, M. and Rahimi, M., Calibration of EPANET using genetic algorithm. *Proceedings of the 10th Annual Water Distribution Systems Analysis Conference*, eds. J. E. Van Zyl, A. A. Ilemobade and H. E. Jacobs, Kruger National Park, South Africa, pp. 695–702, 2008.

[7] Kim, S.H., Extensive Development of Leak Detection Algorithm by Impulse Response Method, *Journal of Hydraulic Engineering*, **131(3)**, pp. 201–208, 2005.

[8] Kim, S.H., Address-Oriented Impedance Matrix Method for Generic Calibration of Heterogeneous Pipe Network Systems. *Journal of Hydraulic Engineering*, **134(1)**, pp. 66–75, 2008.

[9] Prasad, T.D. Design of Pumped Water Distribution Networks with Storage. *Journal of Water Resources Planning and Management*, **136(1)**, pp. 129–132, Jan. 2010.

[10] Mendes, J.B., Soleira Elíptica-circular, 2008, DoctorateThesis, Escola Politécnica, Departamento de Engenharia Hidráulica e Sanitária, Universidade de São Paulo, São Paulo, Brazil.

[11] Minei, N., Um Método Expedito para a Medição de Vazão em Rios e Canais Abertos, 1999, DoctorateThesis, Escola Politécnica, Departamento de Engenharia Hidráulica e Sanitária, Universidade de São Paulo, São Paulo, Brazil.

[12] Melchior, C., Comparativo de resultados de medição de vazão pelos métodos: convencional e acústico, 2006, Course Completion Work, União Dinâmica de Faculdades Cataratas, Civil Engineering, Foz do Iguaçu, Brazil.