ON THE APPROXIMATION OF THE CUT FUNCTIONS
BY HYPER–LOG–LOGISTIC FUNCTION

SVETOSLAV MARKOV[1], NIKOLAY KYURKCHIEV[2],
ANTON ILIEV[3], AND ASEN RAHNEV[4]

[1]Institute of Mathematics and Informatics

Bulgarian Academy of Sciences

Acad. G. Bonchev Str., Bl. 8, 1113 Sofia, BULGARIA

[2,3,4]Faculty of Mathematics and Informatics

University of Plovdiv Paisii Hilendarski

24, Tzar Asen Str., 4000 Plovdiv, BULGARIA

ABSTRACT: We study the uniform approximation of the sigmoid cut function by smooth sigmoid functions such as the Hyper-log–logistic function. The limiting case of the interval-valued step function is discussed using Hausdorff metric. Various expressions for the error estimates of the corresponding uniform and Hausdorff approximations are obtained. Numerical examples are presented using *CAS MATHEMATICA*.

1. INTRODUCTION

In this paper we discuss some computational, modelling and approximation issues related to one class of sigmoid functions.

Sigmoid functions find numerous applications in various fields related to life sciences, chemistry, physics, artificial intelligence, fuzzy set theory, insurance mathematics, debugging and test theory.

In fields such as signal processing, machine learning, artificial neural networks, sigmoid functions are also known as "activation" and "squashing" functions.

In this work we concentrate on several practically important classes of sigmoid functions. Two of them are the cut (or ramp) functions and the step function.

Cut functions are continuous but they are not smooth (differentiable) at the two endpoints of the interval where they increase.

Step functions can be viewed as limiting case of cut functions; they are not continuous but they are Hausdorff continuous (H-continuous) [1], [2].

Section 2 contains preliminary definitions and motivations.

In Section 3 we study the uniform and Hausdorff approximation of the cut functions by hyper–log–logistic function.

Curiously, the uniform distance between a cut function and the hyper–log–logistic function of best uniform approximation is an function of the scale parameter β and does not depending on the width of the underlying interval Δ, resp. on the slope k.

By contrast, it turns out that the Hausdorff distance (H-distance) depends on the slope and tends to zero when increasing the slope.

Numerical examples are presented throughout the paper using the computer algebra system *MATHEMATICA*.

2. PRELIMINARIES

Sigmoid functions. In this work we consider *sigmoid functions* of a single variable defined on the real line, that is functions s of the form $s : \mathbb{R} \longrightarrow \mathbb{R}$. Sigmoid functions can be defined as bounded monotone non-decreasing functions on $\mathbb{R}$. One usually makes use of normalized sigmoid functions defined as monotone non-decreasing functions $s(t), t \in \mathbb{R}$, such that $\lim s(t)_{t \to -\infty} = 0$ and $\lim s(t)_{t \to \infty} = 1$. In the fields of neural networks and machine learning sigmoid-like functions of many variables are used, familiar under the name *activation functions*. (In some applications the sigmoid functions are normalised so that the lower asymptote is assumed -1: $\lim s(t)_{t \to -\infty} = -1$.)

Cut (ramp) functions. Let $\Delta = [\gamma - \delta, \gamma + \delta]$ be an interval on the real line $\mathbb{R}$ with centre $\gamma \in \mathbb{R}$ and radius $\delta \in \mathbb{R}$. A cut function (on Δ) is defined as follows:

Definition 1. The cut function $c_{\gamma,\delta}$ on Δ is defined for $t \in \mathbb{R}$ by

$$
c_{\gamma,\delta}(t) = \begin{cases} 0, & \text{if} \quad t < \Delta, \\ \dfrac{t - \gamma + \delta}{2\delta}, & \text{if} \quad t \in \Delta, \\ 1, & \text{if} \quad \Delta < t. \end{cases} \tag{1}
$$

Note that the slope of function $c_{\gamma,\delta}(t)$ on the interval Δ is $1/(2\delta)$ (the slope is constant in the whole interval Δ). Two special cases are of interest for our discussion in the sequel.

Special case 1. For $\gamma = 0$ we obtain a cut function on the interval $\Delta = [-\delta, \delta]$:

$$c_{0,\delta}(t) = \begin{cases} 0, & \text{if} \quad t < -\delta, \\ \dfrac{t + \delta}{2\delta}, & \text{if} \quad -\delta \leq t \leq \delta, \\ 1, & \text{if} \quad \delta < t. \end{cases} \tag{2}$$

Special case 2. For $\gamma = \delta$ we obtain the cut function on $\Delta = [0, 2\delta]$:

$$c_{\delta,\delta}(t) = \begin{cases} 0, & \text{if} \quad t < 0, \\ \dfrac{t}{2\delta}, & \text{if} \quad 0 \leq t \leq 2\delta, \\ 1, & \text{if} \quad 2\delta < t. \end{cases} \tag{3}$$

Step functions. The step function (with "jump" at $\gamma \in \mathbb{R}$) can be defined by

$$h_\gamma(t) = c_{\gamma,0}(t) = \begin{cases} 0, & \text{if} \quad t < \gamma, \\ [0,1], & \text{if} \quad t = \gamma, \\ 1, & \text{if} \quad t > \gamma, \end{cases} \tag{4}$$

which is an *interval-valued function* (or just *interval function*) [1], [2]. In the literature various point values, such as $0, 1/2$ or 1, are prescribed to the step function (4) at the point γ; we prefer the interval value $[0, 1]$. When the jump is at the origin, that is $\gamma = 0$, then the step function is known as the Heaviside step function; its "interval" formulation is:

$$h_0(t) = c_{0,0}(t) = \begin{cases} 0, & \text{if} \quad t < 0, \\ [0,1], & \text{if} \quad t = 0, \\ 1, & \text{if} \quad t > 0. \end{cases} \tag{5}$$

The step function can be perceived as a limiting case of the cut function. Namely, for $\delta \to 0$, the cut function $c_{\delta,\delta}$ tends in "Hausdorff sense" to the step function. Here "Hausdorff sense" means *Hausdorff distance*, briefly *H-distance*. The H-distance $\rho(f, g)$ between two interval functions f, g on $\Omega \subseteq \mathbb{R}$, is the distance between their completed graphs $F(f)$ and $F(g)$ considered as closed subsets of $\Omega \times \mathbb{R}$ [5], [6]. More precisely,

$$\rho(f, g) = \max\{ \sup_{A \in F(f)} \inf_{B \in F(g)} ||A - B||, \sup_{B \in F(g)} \inf_{A \in F(f)} ||A - B||\}, \tag{6}$$

wherein $||.||$ is any norm in $\mathbb{R}^2$, e. g. the maximum norm $||(t,x)|| = \max |t|, |x|$.

The logistic function was introduced by P.-F. Verhulst [7]–[9], who applied it to human population dynamics. Verhulst derived his logistic equation to describe the mechanism of the self-limiting growth of a biological population. The logistic function finds applications in an wide range of fields, including artificial neural networks, biology, ecology, population dynamics, chemistry, demography, economics, geoscience, mathematical psychology, probability, sociology, political science and bio–statistics.

Definition 2. Define the logistic (Verhulst) function v on $\mathbb{R}$ as [7]

$$v_{\gamma,k}(t) = \frac{1}{1 + e^{-4k(t-\gamma)}}. \tag{7}$$

Note that the logistic function (7) has an inflection at its "centre" $(\gamma, 1/2)$ and its slope at γ is equal to k.

In [4] we prove the following propositions

Proposition 3. *The function $v_{\gamma,k}(t)$ defined by (7) with $k = 1/(2\delta)$: i) is the logistic function of best uniform one-sided approximation to function $c_{\gamma,\delta}(t)$ in the interval $[\gamma, \infty)$ (as well as in the interval $(-\infty, \gamma]$); ii) approximates the cut function $c_{\gamma,\delta}(t)$ in uniform metric with an error*

$$\rho = \rho(c, v) = \frac{1}{1 + e^2} = 0.11920292.... \tag{8}$$

Proposition 4. *For the H-distance $h(k)$ between the function $v_{\gamma,k}(t)$ and $c_{\gamma,\delta}(t)$ the following holds for $k > 5$:*

$$\frac{1}{4k + 1} < h(k) < \frac{\ln(4k + 1)}{4k + 1}. \tag{9}$$

3. APPROXIMATION OF THE CUT FUNCTION BY HYPER–LOG–LOGISTIC FUNCTIONS

In 1968 Blumberg [10] introduced a modified Verhulst logistic equation, the so called *hyper–log–logistic equation*:

$$\frac{dN(t)}{dt} = kN^{\alpha}(1 - N)^{\gamma} \tag{10}$$

where k is the rate constant and α and γ are shape parameters.

The equation (10) is consistent with the Verhulst logistic model when $\alpha = \gamma = 1$.

A number of equations have been proposed to account for the shape of these growth curves (Turner and al. [11], Tsoularis [12]).

We will consider the following modification of the hyper–log–logistic equation (10) (see for instance [12]:

$$\frac{dN(t)}{dt} = kN^{1-\frac{1}{\beta}}(1-N)^{1+\frac{1}{\beta}} \tag{11}$$

where β is a shape parameter.

For $\beta \to \infty$ the equation (2) reduces to Verhulst equation.

The equation (11) provides a parametric interpolation formula between the predictions of the logistic equation ($\beta \to \infty$) and second order kinetics ($\beta = 1$).

Definition 5. Define the hyper-log–logistic function N on $\mathbb{R}$ as:

$$N_{\gamma,\beta,k}(t) = 1 - \frac{1}{1 + \left(1 + \frac{4k(t-\gamma)}{\beta}\right)^{\beta}}. \tag{12}$$

Note that the logistic function (12) has an inflection at its "centre" $(\gamma, 1/2)$ and its slope at γ is equal to k.

Proposition 6. *The function $N_{\gamma,\beta,k}(t)$ defined by (12) with $k = 1/(2\delta)$:*

i) is the hyper–log–logistic function of best uniform one-sided approximation to function $c_{\gamma,\delta}(t)$ in the interval $[\gamma, \infty)$ (as well as in the interval $(-\infty, \gamma]$);

ii) approximates the cut function $c_{\gamma,\delta}(t)$ in uniform metric with an error

$$\rho_1 = \rho_1(c, N) = 1 - \frac{1}{1 + \left(1 - \frac{2}{\beta}\right)^{\beta}}. \tag{13}$$

Proof. Consider functions (1) and (12) with same centres $\gamma = \delta$, that is functions $c_{\delta,\delta}$ and $N_{\delta,\beta,k}$.

In addition chose c and N to have same slopes at their coinciding centres, that is assume $k = 1/(2\delta)$, cf. Fig. 1–Fig. 2.

Then, noticing that the largest uniform distance between the cut and hyper–log–logistic functions is achieved at the endpoints of the underlying interval $[0, 2\delta]$, we have:

$$\rho_1 = N_{\delta,\beta,k}(0) - c_{\delta,\delta} = 1 - \frac{1}{1 + \left(1 - \frac{4k\delta}{\beta}\right)^{\beta}} = 1 - \frac{1}{1 + \left(1 - \frac{2}{\beta}\right)^{\beta}} = A(\beta). \tag{14}$$
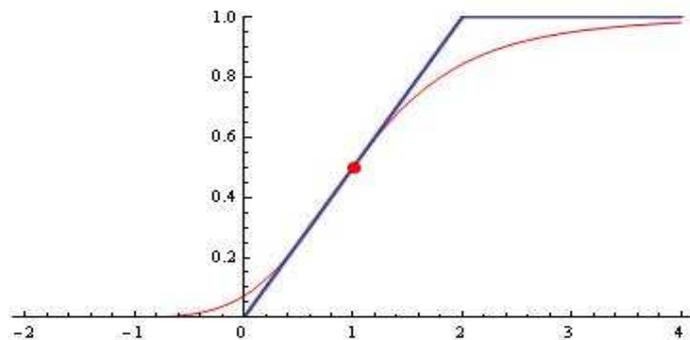
This completes the proof of the proposition.

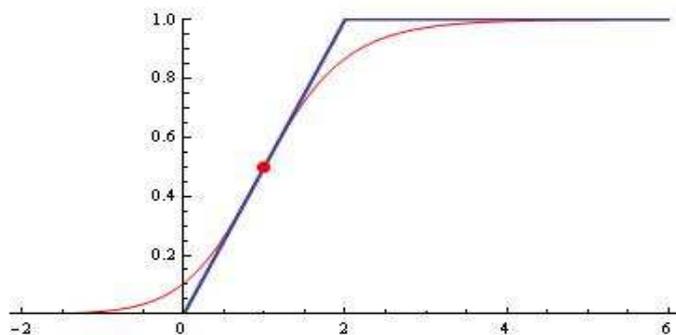Figure 1: The cut and hyper–log–logistic functions for $\gamma = \delta = 1$, $k = 1/2$, $\beta = 5$.

Figure 2: The cut and hyper–log–logistic functions for $\gamma = \delta = 1$, $k = 1/2$, $\beta = 15$.

β	$A(\beta)$ from (14)
3	0.0357143
5	0.0721496
10	0.0969629
15	0.104658
100	0.117091
1000	0.118993
10000	0.119182
100000	0.119201
1000000	0.119203

Table 1: The function $A(\beta)$ computed by (14) for various β

Some computational examples using relation (14) for various β are presented in Table 1.

We note that the uniform distance (13) is an function of the scale parameter β and does not depending on the width of the underlying interval Δ, resp. on the slope k.

This should not surprise us. We already mentioned that the equation (10) is consistent with the Verhulst logistic model when $(\beta \to \infty)$.

Evidently

$$\lim_{\beta \to \infty} A(\beta) = 1 - \frac{1}{1 + e^{-2}} = \frac{1}{1 + e^2} = 0.11920292....$$

and we have the result from Proposition 3.

The next proposition shows that this is not the case whenever H-distance is used.

Proposition 7. *The function $N_{0,\beta,k}(t)$ with $k = 1/(2\delta)$ is the hyper–log–logistic function of best Hausdorff one-sided approximation to function $c_{0,\delta}(t)$ in the interval $[0, \infty]$ (resp. in the interval $[-\infty, 0]$).*

The function $N_{0,\beta,k}(t)$, $k = 1/(2\delta)$, approximates function $c_{0,\delta}(t)$ in H-distance with an error $h = h(c, N)$ that satisfies the relation:

$$\ln \frac{1 - h}{h} = -\beta \ln \left(1 - \frac{2}{\beta}(1 + 2kh) \right). \tag{15}$$
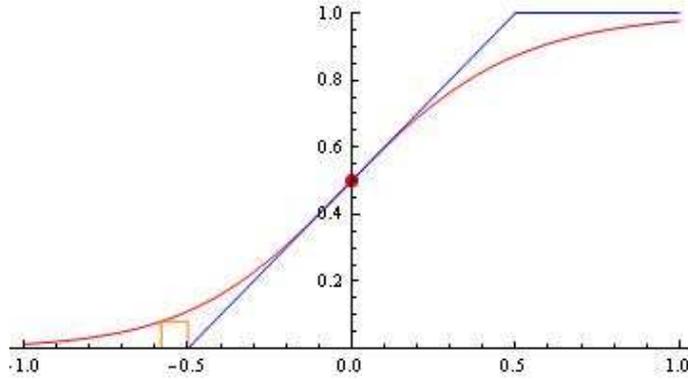
Figure 3: The cut and hyper–log–logistic functions for $k = 1$, $\delta = \frac{1}{2k}$, $\beta = 25$, H-distance $h = 0.0804196$.

Proof. Using $\delta = 1/(2k)$ we can write $\delta + h = (1 + 2hk)/(2k)$, resp.:

$$N(-\delta - h) = 1 - \frac{1}{1 + \left(1 - \frac{2}{\beta}(1 + 2kh)\right)^{\beta}}. \tag{16}$$

The H-distance h using square unit ball (with a side h) satisfies the relation

$$N(-\delta - h) = h, \tag{17}$$

which implies (15).

This completes the proof of the proposition.

Relation (15) shows that the H-distance h depends on the slope k and scale parameter β, $h = h(k, \beta)$.

The numerical results are plotted in Fig. 3 (for the case $k = 1$, $\delta = \frac{1}{2k}$, $\beta = 25$, H-distance $h = 0.0804196$) and Fig. 4 (for the case $k = 1.5$, $\delta = \frac{1}{2k}$, $\beta = 10$, H-distance $h = 0.0623606$).

The next result gives additional information on this dependence.

Let

$$p = -1 + \frac{1}{1 + \left(1 - \frac{2}{\beta}\right)^{\beta}}; \quad q = 1 + \frac{4k\left(1 - \frac{2}{\beta}\right)^{\beta - 1}}{\left(1 + \left(1 - \frac{2}{\beta}\right)^{\beta}\right)^2};$$

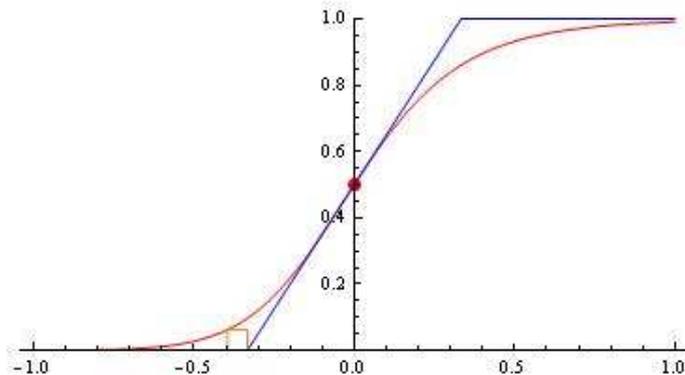$$r = -2.1q/p; \quad (p < 0; q > 0; r > 0).$$

Figure 4: The cut and hyper–log–logistic functions for $k = 1.5$, $\delta = \frac{1}{2k}$, $\beta = 10$, H-distance $h = 0.0623606$.

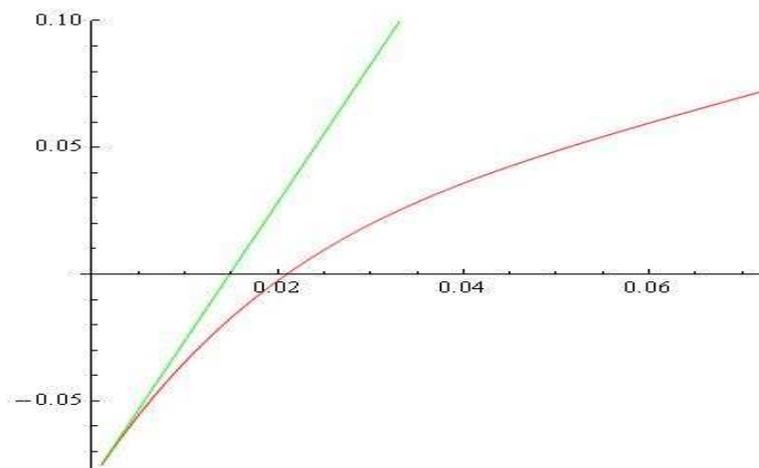Figure 5: The functions $F(d)$ and $G(d)$ for $k = 10$, $\beta = 6$.

Proposition 8. *For the H-distance $h = h(k, \beta)$ between the cut and the hyper–log– logistic functions the following holds for $r > e^{2.1}$:*

$$h_1 = \frac{1}{r} < h(k, \beta) < \frac{\ln r}{r} = h_2. \tag{18}$$

Proof. From (17) we have

$$1 - h = \frac{1}{1 + \left(1 - \frac{2}{\beta}(1 + 2kh)\right)^\beta}$$

Let us examine the function

$$F(h) = \frac{1}{1 + \left(1 - \frac{2}{\beta}(1 + 2kh)\right)^\beta} - 1 + h.$$

From $F'(h) > 0$ we conclude that function F is strictly monotone increasing.

Consider function

$$G(h) = p + qh.$$

using the Taylor expansion $G(h) - F(h) = O(h^2)$.

Hence $G(h)$ approximates $F(h)$ with $h \to 0$ as $O(h^2)$ (see, Fig.5).

In addition $G'(h) > 0$, hence function G is monotone increasing.

Further, for $r > e^{2.1}$

$$G\left(\frac{1}{r}\right) < 0, \quad G\left(\frac{\ln r}{r}\right) > 0.$$

This completes the proof of the proposition.

Some computational examples using nonlinear equation (17) and two–sided bounds (18) for various k and β are presented in Table 2.

From the above table, it can be seen that the estimates for the value of the best H–distance (see (18)) are quite precise.

For other results, see [3]-[4], [13]–[21].

4. CONCLUSIONS

In this paper we study the uniform and Hausdorff approximation of the cut functions by hyper–log–logistic function. We demonstrate that the best uniform approximation between a cut function and the respective logistic function is an function of the scale parameter β and does not depending on the slope k. On the other side we show

k	β	h_1	h from (17)	h_2
10	6	0.00704963	0.020959	0.0349294
12	6	0.00605996	0.0187194	0.0309425
14	8	0.00603839	0.019563	0.0308539
8	10	0.0102551	0.0295792	0.046968
6	12	0.0132971	0.036255	0.0574461
8	20	0.0116352	0.0338921	0.05182
4	30	0.0197223	0.05013	0.0774298
4	60	0.0204527	0.0520565	0.0795537

Table 2: The h–distance computed by (17) and two–sided bounds computed by (18)

that the Hausdorff distance (H-distance) depends on the slope k and tends to zero whenever $k \to \infty$.

For basic results on H-continuous functions and their application to problems in abstract areas such as Real Analysis, Approximation Theory, Set-valued Analysis and Fuzzy Sets and Systems we recommend [1], [3], [26], [27].

Logistic functions are also used in artificial neural networks [22]–[31]. Any neural net element computes a linear combination of its input signals, and uses a logistic function to produce the result; often called "activation" function.

Constructive approximation by superposition of sigmoidal functions and the relation with neural networks and radial basis functions approximations is discussed in [27].

A recurrent neurodynamic model of the neuron with a broad class of activation functions, including sigmoidal, stepwise, bounded linear and other ones is proposed in [33].

The explored feature family can find application in the field of debugging and test theory [36]–[37].

We hope that the results will be of interest to specialists working in the field of constructive approximation [22]–[35].

ACKNOWLEDGMENTS

REFERENCES

[1] R. Anguelov, S. Markov, Hausdorff Continuous Interval Functions and Approximations, In: *SCAN 2014 Proceedings, LNCS*, ed. by J.W.von Gudenberg, Springer, Berlin (2015).

[2] J. H. van der Walt, The Linear Space of Hausdorff Continuous Interval Functions, *Biomath*, **2** (2013), 1311261.

[3] J. Dombi, Z. Gera, The approximation of piecewise linear membership functions and Likasiewicz operators, *Fuzzy Sets and Systems*, **154** (2005), 275–286.

[4] A. Iliev, N. Kyurkchiev, S. Markov, On the Approximation of the Cut and Step Functions by Logistic and Gompertz Functions, *Biomath*, **4** No. 2 (2015).

[5] F. Hausdorff, *Set theory (2 ed.)*, Chelsea Publ., New York (1962) [1957], ISBN 978-0821838358, Republished by AMS-Chelsea (2005).

[6] B. Sendov, *Hausdorff Approximations*, Kluwer, Boston (1990).

[7] P.-F. Verhulst, Notice sur la loi que la population poursuit dans son accroissement, *Correspondance mathematique et physique*, **10** (1838), 113–121.

[8] P.-F. Verhulst, Recherches mathematiques sur la loi d'accroissement de la population (Mathematical Researches into the Law of Population Growth Increase), *Nouveaux Memoires de l'Academie Royale des Sciences et Belles-Lettres de Bruxelles*, **18** (1845), 1–42.

[9] P.-F. Verhulst, Deuxieme memoire sur la loi d'accroissement de la population, *Memoires de l'Academie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, **20** (1847), 1–32.

[10] A. A. Blumberg, Logistic growth rate functions, *J. Theoret. Biol.*, **21** (1968), 42–44.

[11] M. Turner, E. Bradley, K. Kirk, K. Pruitt, *A theory of growth, Math. Biosci.*, **29** (1976), 367–373.

[12] A. Tsoularis, Analysis of logistic growth models, *Les. Lett. Inf. Math. Sci.*, **2** (2001), 23–36.

[13] N. Kyurkchiev, S. Markov, Sigmoidal Functions: Some Computational and Modelling Aspects, *Biomath Communications*, **1** No. 2 (2014).

[14] N. Kyurkchiev, S. Markov, *Sigmoid functions: Some Approximation and Modelling Aspects. Some Moduli in Programming Environment MATHEMATICA*, LAP LAMBERT Acad. Publ. (2015).

[15] N. Kyurkchiev, S. Markov, On the Hausdorff distance between the Heaviside step function and Verhulst logistic function, *J. Math. Chem.*, **54** No. 1 (2016), 109–119.

[16] N. Kyurkchiev, A new transmuted cumulative distribution function based on the Verhulst logistic function with application in population dynamics, *Biomath Communications*, **4** No. 1 (2017).

[17] S. Markov, A. Iliev, A. Rahnev, N. Kyurkchiev, A note on the Log-logistic and transmuted Log-logistic models. Some applications, *Dynamic Systems and Applications*, **27** No. 3 (2018), 593–607.

[18] N. Kyurkchiev, S. Markov, A family of recurrence generated sigmoidal functions based on the log-logistic function. Some approximation aspects, *Biomath Communications*, **5** No. 1 (2018), (18 pp).

[19] A. Iliev, N. Kyurkchiev, S. Markov, Approximation of the Cut Function by Stannard and Richard Sigmoid Functions, *International Journal of Pure and Applied Mathematics*, **109** No. 1 (2016), 119–128.

[20] N. Kyurkchiev, S. Markov, On the Approximation of the Generalized Cut Function of Degree $p+1$ By Smooth Sigmoid Functions, *Serdica Journal of Computing*, **9** No. 1 (2015).

[21] R. Anguelov, N. Kyurkchiev, S. Markov, A note on the Blumberg's hyper–log–logistic model, *BIOMATH*, (2018). (to appear)

[22] Z. Chen, F. Cao, The Approximation Operators with Sigmoidal Functions, *Computers & Mathematics with Applications*, **58** (2009), 758–765.

[23] Z. Chen, F. Cao, The Construction and Approximation of a Class of Neural Networks Operators with Ramp Functions, *Journal of Computational Analysis and Applications*, **14** (2012), 101–112.

[24] Z. Chen, F. Cao, J. Hu, Approximation by Network Operators with Logistic Activation Functions, *Applied Mathematics and Computation*, **256** (2015), 565–571.

[25] D. Costarelli, *Sigmoidal Functions Approximation and Applications*, Dottorato di Ricera in Matematica, Departiment di Matematica e Fisica, Sezione di Matematica, Roma, UNIVERSITA DEGLI STUDI (2014).

[26] D. Costarelli, R. Spigler, Approximation Results for Neural Network Operators Activated by Sigmoidal Functions, *Neural Networks*, **44** (2013), 101–106.

[27] D. Costarelli, R. Spigler, Constructive Approximation by Superposition of Sigmoidal Functions, *Anal. Theory Appl.*, **29** (2013), 169–196.

[28] D. Costarelli, G. Vinti, Pointwise and uniform approximation by multivariate neural network operators of the max-product type, *Neural Networks*, (2016), doi:10.1016/j.neunet.2016.06.002.

[29] D. Costarelli, R. Spigler, Solving numerically nonlinear systems of balance laws by multivariate sigmoidal functions approximation, *Computational and Applied Mathematics*, (2016), doi:10.1007/s40314-016-0334-8.

[30] D. Costarelli, G. Vinti, Convergence for a family of neural network operators in Orlicz spaces, *Mathematische Nachrichten*, (2016), doi:10.1002/mana.20160006.

[31] N. Guliyev, V. Ismailov, A single hidden layer feedforward network with only one neuron in the hidden layer san approximate any univariate function, *Neural Computation*, **28** (2016), 1289–1304.

[32] I. A. Basheer, M. Hajmeer, Artificial neural networks: fundamentals, computing, design, and application, *Journal of Microbiological Methods*, **43** (2000), 3–31.

[33] Y. Prostov, Y. Tiumentsev, Recurrent neurodynamic model of neuron with variable activation characteristic, *Biologically Inspired Cognitive Architectures*, (2018), (8 pp.)

[34] Y. Chalco-Cano, H. Roman-Flores, F. Gomida, A new type of approximation for fuzzy intervals, *Fuzzy Sets and Systems*, **159** No. 11 (2008), 1376–1383.

[35] Z. Quan, Z. Zhang, The construction and approximation of the neural network with two weights, *J. Appl. Math.*, (2014), doi:10.1155/2014/892653.

[36] N. Pavlov, A. Iliev, A. Rahnev, N. Kyurkchiev, *Some software reliability models: Approximation and modeling aspects*, LAP LAMBERT Academic Publishing, (2018).

[37] N. Pavlov, A. Iliev, A. Rahnev and N. Kyurkchiev, *Nonstandard Models in Debugging Theory (Part 2)*, LAP LAMBERT Academic Publishing, (2018).