CSCanada
Progress in Applied Mathematics
Vol. 5, No. 2, 2013, pp. [40–47]
DOI: 10.3968/j.pam.1925252820130502.2994

ISSN 1925-251X [Print]
ISSN 1925-2528 [Online]
www.cscanada.net
www.cscanada.org

On Some Properties of a Hetereogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent Sigmoid (TANSIG) Transfer Functions

Christopher Godwin Udomboso[a],*

[a]Department of Statistics, University of Ibadan, Ibadan, Nigeria.

* Corresponding author.
Address: Department of Statistics, University of Ibadan, Ibadan, Nigeria; E-Mail: cg.udomboso@mail.ui.edu.ng

Received: February 20, 2013/ Accepted: April 12, 2013/ Published: April 30, 2013

Abstract: Transfer functions maps the input layer of the statistical neural network model to the output layer. To do this perfectly, the function must lie within certain bounds. This is a property of probability distributions. This paper establishes the heterogeneous transfer function, SATLINS_TANSIG, as a Probability Distribution Functions (p.d.f) by showing that it is proper. It also shows the mean and variance.

Key words: Statistical neural network; SATLINS; TANSIG; SATLINS_-TANSIG; Mean; Variance

Udomboso, C. G. (2013). On Some Properties of a Hetereogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent Sigmoid (TANSIG) Transfer Functions. *Progress in Applied Mathematics, 5*(2), 40–47. Available from http://www.cscanada.net/index.php/pam/article/view/j.pam.1925252820130502.2994 DOI: 10.3968/j.pam.1925252820130502.2994

1. INTRODUCTION

Artificial Neural Networks (ANNs) are non linear mapping structures based on the function of the human brain. They imitate the training of the human brain and can process problems involving non linear and complex data even if the data are imprecise and noisy. They are powerful tools for modeling, especially when the underlying data relationship is unknown, and are ideally suited for modeling linear and non linear data. This is why their applications to statistical problems are very interesting. Such networks are called the Statistical Neural Network (SNN). This study centres on the Multi-Layer Perceptron (MLP) which happens to be the most commonly used type of ANN [1]. It has been found to be powerful in terms of model precision in the usage of homogeneous transfer functions (TFs), especially with complex or large data set. The choice of MLP is because it is the only ANN type that allows for statistical inference.

The modeling power of an SNN model lies on the transfer function that is used. There are several transfer functions that may be used in a given SNN model. Selections are made from a fixed pool of different transfer functions, and possibly using pruning techniques to drop functions that are not useful. Up till now, known literatures and researches have reported network analysis using one transfer functions (that is, homogeneous models). For example, [2–9] used the sigmoid transfer function, while [10] compared logistic and hyperbolic tangent transfer functions, [4] used the tangential transfer function (that is, family of tangents functions), [11] as well as [12] used the symmetric saturated linear transfer function.

This study endeavours to investigate an analytical derivation of a heterogeneous transfer function using the symmetric saturated linear (SATLINS) as well as the hyperbolic tangent sigmoid (TANSIG) transfer functions. It further showed that the derived transfer function is a proper probability density function (p.d.f). Hence the mean and variance were also derived.

2. THE STATISTICAL NEURAL NETWORK MODEL

Anders (1996) proposed a statistical neural network model given as

$$y = f(X, w) + u \tag{1}$$

where y is the dependent variable, $X = (x_0 \equiv 1, x_1, ..., x_I)$ is a vector of independent variables, $w = (\alpha, \beta, \gamma)$ is the network weight: 'α' is the weight of the input unit, 'β' is the weight of the hidden unit, and 'γ' is the weight of the output unit, and u_i is the stochastic term that is normally distributed (that is, $u_i \sim N(0, \sigma^2 I_n)$).

Basically, $f(X, w)$ is the artificial neural network function, expressed as

$$f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h g\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right).$$

where $g(.)$ is the transfer function.

The proposed convoluted form of the artificial neural network function used in this study is

$$f(X, w) = \alpha X + \sum_{h=1}^{H} \beta_h \left[g_1\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right) g_2\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right)\right]$$

and thus, the form of the statistical neural network model proposed is

$$y = \alpha X + \sum_{h=1}^{H} \beta_h \left[g_1\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right) g_2\left(\sum_{i=0}^{I} \gamma_{hi} x_i\right)\right] + u_i u_j \tag{2}$$

where y is the dependent variable, $X = (x_0 \equiv 1, x_1, ..., x_I)$ is a vector of independent variables, $w = (\alpha, \beta, \gamma)$ is the network weight: 'α' is the weight of the input unit, 'β' is the weight of the hidden unit, and 'γ' is the weight of the output unit, u_i and u_j are the stochastic terms that are normally distributed (that is, $u_i,\ u_j \sim N(0, \sigma^2 I_n)$), and $g_1(.)$ and $g_2(.)$ are the transfer functions.

In this paper, we investigate the distributional properties of the heterogeneous transfer function arising from the convolution of SATLINS and TANSIG.

Let $g_1(.)$ = Symmetric Saturated Linear function (SATLINS), defined as

$$\text{satlins} = g_1(.) = f_1(n) = \begin{cases} -1, & n < -1 \\ n, & -1 \leq n \leq 1 \\ 1, & n > 1 \end{cases} \tag{3}$$

Let $g_2(.)$ = Hyperbolic Tangent Sigmoid function (TANSIG), defined as

$$\text{tansig} = g_2(.) = f_3(n) = \frac{2}{1 - e^{-2n}} - 1 \tag{4}$$

3. SYMMETRIC SATURATING LINEAR AND HYPERBOLIC TANGENT SIGMOID MODEL (SATLINS_TANSIG)

(i). For $n < -1$, $f_1(n) = a = -1$. This implies that $f_1(n - m) = -1$.

$$f_3(m) = \frac{2}{1 - e^{-2n}} - 1 = \frac{1 + e^{-2m}}{1 - e^{-2m}} \tag{5}$$

Let

$$\begin{aligned} f(n) &= f_1(n) \otimes f_3(n) \\ &= \int_{-r}^{n} f_1(n-m) f_3(m) dm \\ &= \int_{-r}^{n} \left(1 - \frac{2}{1 - e^{-2m}}\right) dm \\ &= (n + r) - 2\int_{-r}^{n} \left(1 + e^{-2m} + e^{-4m} + ...\right) dm \\ &= \sum_{p=1}^{\infty} \frac{e^{-2pn}}{p} - \sum_{p=1}^{\infty} \frac{e^{2pr}}{p} - (n + r) \end{aligned} \tag{6}$$

(ii). For $-1 \leq n \leq 1$, $f_1(n) = n$. This implies that $f_1(n - m) = n - m$.

$$f_3(m) = \frac{2}{1 - e^{-2m}} - 1 = \frac{1 + e^{-2m}}{1 - e^{-2m}}$$

$$\begin{aligned} f_1(n) \otimes f_3(n) &= \int_{-1}^{n} f_1(n-m) f_3(m) dm, \qquad -1 \leq n \leq 1 \\ &= \int_{-1}^{n} (n - m) \left(\frac{2}{1 - e^{-2m}} - 1\right) dm \end{aligned} \tag{7}$$

$I = \int_{-1}^{n} \frac{2m}{1-e^{-2m}} dm$ decreases rapidly for any interval of m. Hence $I = 0$. Therefore, (7) becomes

$$\begin{aligned} f_1(n) \otimes f_3(n) &= 2n \int_{-1}^{n} (1-e^{-2m}) dm + \int_{-1}^{n} (n-m) dm \\ &= \left(2n^2 + 3n + \frac{1}{2}\right) - n \left(\sum_{p=1}^{n} \frac{e^{-2pn}}{p} - \sum_{p=1}^{n} \frac{e^{-2p}}{p} \right) \end{aligned} \quad (8)$$

(iii). For $n > 1$, $f_1(n) = a = 1$. This implies that $f_1(n-m) = 1$ as

$$f_3(m) = \frac{2}{1-e^{-2m}} - 1$$

$$f_1(n) \otimes f_3(n) = \int_{1}^{n} f_1(n-m) f_3(m) dm$$

Therefore,

$$\begin{aligned} f_1(n) \otimes f_3(n) &= \int_{1}^{n} 1 \left(\frac{2}{1-e^{-2m}} - 1 \right) dm \\ &= 2 \int_{1}^{n} \left(1 + e^{-2m} + e^{-4m} + ...\right) dm - \int_{1}^{n} dm \\ &= (n-1) - \sum_{p=1}^{\infty} \frac{e^{-2pn}}{p} + \sum_{p=1}^{\infty} \frac{e^{-2p}}{p} \end{aligned} \quad (9)$$

The summary of the derived function is given as

$$g_1 \left(\sum_{i=0}^{I} \gamma_{hi} x_i \right) g_2 \left(\sum_{i=0}^{I} \gamma_{hi} x_i \right) = f(n)$$

$$= \begin{cases} \sum_{p=1}^{\infty} \frac{e^{-2pn}}{p} - \sum_{p=1}^{\infty} \frac{e^{2pr}}{p} - (n+r), & n < -1 \\ \left(2n^2 + 3n + \frac{1}{2}\right) - n \left(\sum_{p=1}^{n} \frac{e^{-2pn}}{p} - \sum_{p=1}^{n} \frac{e^{-2p}}{p} \right), & -1 \le n \le 1 \\ (n-1) - \sum_{p=1}^{\infty} \frac{e^{-2pn}}{p} + \sum_{p=1}^{\infty} \frac{e^{-2p}}{p}, & n > 1 \end{cases} \quad (10)$$

Equation (10) is the derived transfer function for the *Symmetric Saturated Linear transfer function* and the *Hyperbolic Tangent transfer function*.

4. DISTRIBUTIONAL PROPERTIES OF THE SATLINS_-TANSIG MODEL

We now show that the derived transfer function is a probability density function. By definition, the probability density function (p.d.f) of function $f(x)$ of a random variable $X : \Omega \to R$ is said to be a proper p.d.f if for $x \in (-\infty, +\infty)$, $x \in X$, we have that,

$$\int_{-\infty}^{\infty} f(x) dx = 1, \quad \forall x \in X$$

From the derived transfer function

$$\begin{aligned}
&\int_{-\infty}^{\infty} f_1(n)\otimes f_3(n)dn \\
=&\int_{-\infty}^{-1}\left[\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}-\sum_{p=1}^{\infty}\frac{e^{2pr}}{p}-(n+r)\right]dn \\
&+\int_{-1}^{1}\left[\left(2n^2+3n+\frac{1}{2}\right)-n\left(\sum_{p=1}^{n}\frac{e^{-2pn}}{p}-\sum_{p=1}^{n}\frac{e^{-2p}}{p}\right)\right]dn \\
&+\int_{1}^{\infty}\left[(n-1)-\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}+\sum_{p=1}^{\infty}\frac{e^{-2p}}{p}\right]dn \\
=&-\int_{-\infty}^{-1}\left(\sum_{p=1}^{\infty}\frac{e^{2pr}}{p}\right)dn-\int_{-\infty}^{-1}(n+r)dn \\
&+\int_{-1}^{1}\left(2n^2+3n+\frac{1}{2}\right)dn-\int_{-1}^{1}\left(\sum_{p=1}^{n}\frac{ne^{-2pn}}{p}\right)dn+\int_{-1}^{1}\left(\sum_{p=1}^{n}\frac{ne^{-2p}}{p}\right)dn \\
&+\int_{1}^{\infty}(n-1)dn+\int_{1}^{\infty}\left(\sum_{p=1}^{\infty}\frac{e^{-2p}}{p}\right)dn \\
=&\frac{7}{3}-\sum_{p=1}^{\infty}\int_{-1}^{1}\frac{ne^{-2pn}}{p}dn=\frac{7}{3}-A
\end{aligned} \tag{11}$$

where $A=\sum_{p=1}^{\infty}\left(\frac{e^{2p}-e^{-2p}}{2p^2}+\frac{1}{2p^2}\left(\frac{e^{2p}-e^{-2p}}{2p}\right)\right]$.

Hence, for suitable values of p such that $A=\frac{4}{3}$, we conclude that $f_1(n)\otimes f_3(n)$ is a probability density function.

We next obtain the mean and variance of the derived transfer function.

$$\begin{aligned}
E(n)=&\int_{-\infty}^{-1}n\left[\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}-\sum_{p=1}^{\infty}\frac{e^{2pr}}{p}-(n-r)\right]dn \\
&+\int_{-1}^{1}n\left[\left(2n^2-3n+\frac{1}{2}\right)-\sum_{p=1}^{\infty}n\frac{e^{-2pn}}{p}+\sum_{p=1}^{\infty}n\frac{e^{2p}}{p}\right]dn \\
&+\int_{1}^{\infty}n\left[(n-1)-\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}+\sum_{p=1}^{\infty}\frac{e^{2p}}{p}\right]dn \\
=&-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-\infty}^{-1}ne^{-2pr}dn-\int_{-\infty}^{-1}(n^2-nr)dn+\int_{-1}^{1}\left(2n^3-3n^2+\frac{n}{2}\right)dn \\
&-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^2e^{-2pn}dn+\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^2e^{-2p}dn
\end{aligned}$$

$$\begin{aligned}&+\int_1^{\infty}(n^2-n)dn+\sum_{p=1}^{\infty}\frac{1}{p}\int_1^{\infty}ne^{-2p}dn\\&=-2+\frac{2}{3}\sum_{p=1}^{\infty}\frac{e^{2p}}{p}-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^2e^{-2pn}dn\end{aligned}\tag{12}$$

Let $I=\int udv=\int_{-1}^{1}n^2e^{-2pn}dn$, where $\int udv=uv-\int vdu$. Also, let $u=n^2$ and $v=\int e^{-2pn}dn$.

Then,

$$\begin{aligned}I=&\left[-n^2\frac{e^{-2pn}}{2p}\right]_{-1}^{1}+\frac{1}{p}\int_{-1}^{1}ne^{-2pn}dn\\=&\frac{1}{2p}\left(e^{2p}-e^{-2p}\right)-\frac{1}{2p^2}\left(e^{2p}+e^{-2p}\right)+\frac{1}{4p^3}\left(e^{2p}-e^{-2p}\right)\end{aligned}$$

Therefore, mean is given as

$$E(n)=-2+\sum_{p=1}^{\infty}\frac{1}{p}\left[\frac{2e^{2p}}{3}+\frac{1}{2p}\left(e^{-2p}-e^{2p}\right)+\frac{1}{2p^2}\left(e^{-2p}+e^{2p}\right)+\frac{1}{4p^3}\left(e^{-2p}-e^{2p}\right)\right]\tag{13}$$

Also,

$$\begin{aligned}E(n^2)=&\int_{-\infty}^{-1}n^2\left[\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}-\sum_{p=1}^{\infty}\frac{e^{2pr}}{p}-(n-r)\right]dn\\&+\int_{-1}^{1}n^2\left[\left(2n^2-3n+\frac{1}{2}\right)-\sum_{p=1}^{\infty}n\frac{e^{-2pn}}{p}+\sum_{p=1}^{\infty}n\frac{e^{2p}}{p}\right]dn\\&+\int_1^{\infty}n^2\left[(n-1)-\sum_{p=1}^{\infty}\frac{e^{-2pn}}{p}+\sum_{p=1}^{\infty}\frac{e^{2p}}{p}\right]dn\\=&\sum_{p=1}^{\infty}\frac{1}{p}\int_{-\infty}^{-1}n^2e^{-2pn}dn-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-\infty}^{-1}n^2e^{-2pr}dn-\int_{-\infty}^{-1}n^2(n-r)dn\\&+\int_{-1}^{1}n^2\left(2n^2-3n+\frac{1}{2}\right)dn-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^3e^{-2pn}dn+\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^3e^{-2p}dn\\&+\int_1^{\infty}n^2(n-1)dn-\sum_{p=1}^{i}nfty\frac{1}{p}\int_1^{\infty}n^2e^{-2pn}dn+\sum_{p=1}^{\infty}\frac{1}{p}\int_1^{\infty}n^2e^{-2p}dn\\=&\frac{23}{15}-\sum_{p=1}^{\infty}\frac{1}{p}\int_{-1}^{1}n^3e^{-2pn}dn\end{aligned}$$

Let $J=\int udv=\int_{-1}^{1}n^3e^{-2pn}dn$, where $\int udv=uv-\int vdu$. Also, let $u=n^3$

and $v = \int e^{-2pn} dn$. Then,

$$J = \left[-n^3 \frac{e^{-2pn}}{2p} \right]_{-1}^{1} + \frac{3}{2p} \int_{-1}^{1} n^2 e^{-2pn} dn$$

$$= -\frac{1}{2p} \left\{ \left(e^{-2p} + e^{2p}\right) - 3 \left[\frac{1}{2p} \left(e^{2p} - e^{-2p}\right) + \frac{1}{2p^2} \left(e^{2p} + e^{-2p}\right) - \frac{1}{4p^3} \left(e^{2p} - e^{-2p}\right) \right] \right\}$$

Thus,

$$E(n^2)$$

$$= \frac{23}{15} + \frac{1}{2p} \left\{ \left(e^{-2p} + e^{2p}\right) - 3 \left[\frac{1}{2p} \left(e^{2p} - e^{-2p}\right) + \frac{1}{2p^2} \left(e^{2p} + e^{-2p}\right) - \frac{1}{4p^3} \left(e^{2p} - e^{-2p}\right) \right] \right\}$$

Hence,

$$\text{var}(n) = E(n^2) - [E(n)]^2$$

$$= \frac{23}{15} + \frac{1}{2p} \left\{ \left(e^{-2p} + e^{2p}\right) - 3 \left[\frac{1}{2p} \left(e^{2p} - e^{-2p}\right) + \frac{1}{2p^2} \left(e^{2p} + e^{-2p}\right) - \frac{1}{4p^3} \left(e^{2p} - e^{-2p}\right) \right] \right\}$$

$$- \left\{ -2 + \sum_{p=1}^{\infty} \frac{1}{p} \left[\frac{2e^{2p}}{3} + \frac{1}{2p} \left(e^{-2p} - e^{2p}\right) + \frac{1}{2p^2} \left(e^{-2p} + e^{2p}\right) + \frac{1}{4p^3} \left(e^{-2p} - e^{2p}\right) \right] \right\}^2$$

5. CONCLUSION

This study derived a heterogeneous transfer function involving the symmetric saturated linear and hyperbolic tangent sigmoid transfer functions. It went further to show that the derived transfer function is a proper probability distribution function (p.d.f), having mean and variance.

REFERENCES

[1] Resop, J. P. (2006). *A comparison of artificial neural networks and statistical regression with biological resources applications.* (M.S. Thesis). University of Maryland, College Park, USA.

[2] Adepoju, G. A., Ogunjuyigbe, S. O. A., & Alawode, K. O. (2007). Application of neural network to load forecasting in Nigerian electrical power system. *The Pacific Journal of Science and Technology, 8*(1), 68-72.

[3] Adewole, A. P., Akinwale, A. T., & Akintomide, A. B. (2011). Artificial neural network model for forecasting foreign exchange rate. *World of Computer Science and Information Technology Journal, 1*(3), 110-118.

[4] Adeyiga, J. A., Ezike, J. O. J., Omotosho, A., & Amakulor, W. (2011). A neural network based model for detecting irregularities in e-Banking transactions. *African Journal of Computer and ICTs, 4*(2), 7-14.

[5] Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*, 716-723.

[6] Akinwale, A. T., Arogundade, O. T., & Adekoya, A. F. (2009). Translated Nigeria stock market prices using artificial neural network for effective prediction. *Journal of Theoretical and Applied Information Technology*, 36-43.

[7] Anders, U. (1996). *Statistical model building for neural networks.* Nunberg, Germany: AFIR Colloqium.

[8] Anderson, J. A. (2003). *An introduction to neural networks.* Prentice Hall.

[9] Ashigwuike, C. E. (2012). Estimation of solar power generation in some Nigerian cities using artificial neural network. *Journal of Chemical, Biological and Physical Sciences, 2*(2), 929-936.

[10] Battiti, R. (1992). First- and second-order methods for learning: Between steepest descent and Newtons method. *Neural Computation, 4*, 141-166.

[11] Carling, A. (1992). *Introducing neural networks.* England: Sigma Press.

[12] Falode, A. O., & Udomboso, C. G. (2012). Predictive modeling of gas production, utilization and flaring in Nigeria using TSRM and TSNN: A comparative approach. *International Journal of Environmental Engineering.* (Accepted)

[13] Foresee, F. D., & Hagan, M. T. (1997). Gauss-Newton approximation to Bayesian regularization. In *IEEE International Conference on Neural Networks, 3*, (pp. 1930-1935). New York: IEEE.

[14] Gan, C., Limsombunchai, V., Clemes, M., & Weng, A. (2005). Consumer choice prediction – artificial neural networks versus logistic models. *Journal of Social Sciences, 1*(4), 211-219.

[15] Golden, R. M. (1996). *Mathematical methods for neural network analysis and design.* Cambridge: MIT Press.

[16] Ibeh, G. F., Agbo, G. A., Rabia, S., & Chikwenze, A. R. (2012). Comparison of empirical and artificial neural network models for the correlation of monthly average global solar radiation with sunshine hours in Minna, Niger State, Nigeria. *International Journal of Physical Sciences, 7*(8), 1162-1165.

[17] Lawrence, J. (1994). *Introduction to neural networks: design, theory, and applications.* Nevada City, CA: California Scientific Software Press.

[18] Maren, A., Harston, C., & Pap, R. (1990). *Handbook of neural computing applications.* San Diego, CA: Academic Press.

[19] Nelson, M. M., & Illingworth, W. T. (1991). *A practical guide to neural nets.* USA: Addison-Wesley Publishing Company, Inc.

[20] Omole, O., Falode, O.A., & Deng, A. D. (2009). Prediction of Nigerian crude oil viscosity using artificial neural network. *Petroleum & Coal, 51*(3), 181-188.

[21] Smith, M. (1993). *Neural networks for statistical modeling.* New York: Van Nostrand Reinhold.

[22] Taylor, J. G. (1999). *Neural networks and their applications.* Wiley.

[23] Tayfur G. (2002). Artificial neural networks for sheet sediment transport. *Hydrol. Sci. J., 47*(6), 879-892.

[24] Toprak, F., & Cigizoglu, H. K. (2008). Predicting longitudinal Dispersion coefficient in natural streams by artificial neural networks. *Hydrol. Processes, 22*(20), 4106-4129.

[25] Udomboso, C. G., & Amahia, G. N. (2011). Comparative analysis of rainfall prediction using statistical neural network and classical linear regression model. *Journal of Modern Mathematics and Statistics, 5*(3), 66-70.

[26] Warner, B., & Misra, M. (1996). Understanding neural networks as statistical tools. *The American Statistician, 50*(4), 284-293.