**Journal of Modern Applied Statistical Methods**

Volume 12 | Issue 2 Article 26

11-1-2013

# On Some Properties of a Heterogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent (TANH) Transfer Functions

Christopher Godwin Udomboso
*University of Ibadan, Ibadan, Nigeria*, cg.udomboso@gmail.com

Journal of Modern Applied Statistical Methods
November 2013, Vol. 12, No. 2, 427-435.

ISSN 1538 − 9472

# *Emerging Scholars:* On Some Properties of a Heterogeneous Transfer Function Involving Symmetric Saturated Linear (SATLINS) with Hyperbolic Tangent (TANH) Transfer Functions

**Christopher Godwin Udomboso**
University of Ibadan
Ibadan, Nigeria

For transfer functions to map the input layer of the statistical neural network model to the output layer perfectly, they must lie within bounds that characterize probability distributions. The heterogeneous transfer function, SATLINS_TANH, is established as a Probability Distribution Function (*p.d.f*), and its mean and variance are shown.

*Keywords:* Statistical neural network, SATLINS, TANH, SATLINS_TANH, mean, variance

## Introduction

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

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

where $y$ is the dependent variable, $X = (x_0 \equiv 1, x_1, \ldots, x_I)$ is a vector of independent variables, $w = (\alpha, \beta, \gamma)$ is the network weight: $\alpha$ is the weight of the input unit, $\beta$ is the weight of the hidden unit, and $\gamma$ 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

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

*The author is a lecturer in the Department of Statistics. Email him at: cg.udomboso@gmail.com.*

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

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

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

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

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

where $y$ is the dependent variable, $X=\left(x_0\equiv 1,x_1,\ldots,x_I\right)$ is a vector of independent variables, $w = (\alpha, \beta, \gamma)$ is the network weight: $\alpha$ is the weight of the input unit, $\beta$ is the weight of the hidden unit, and $\gamma$ 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.

The distributional properties of the heterogeneous model arising from the convolution of SATLINS and TANH is investigated here. Let $g_1(.)$ = Symmetric Saturated Linear function (SATLINS), defined as

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

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

$$\tanh=g_2\left(.\right)=f_2\left(n\right)=\frac{e^n-e^{-n}}{e^n+e^{-n}} \tag{4}$$

## Symmetric Saturating Linear and Hyperbolic Tangent

**(i)**

Let
$$f(n) = f_1(n) \otimes f_2(n) = \int_a^b f_1(n-m)\, f_2(m)\, dm \qquad (5)$$

For $n < -1$, $f_1(n) = -1$, which implies also that $f_1(n - m) = -1$.

$$f_2(m) = \frac{e^m - e^{-m}}{e^m + e^{-m}}$$

Therefore,

$$\begin{aligned} f_1(n) \otimes f_2(n) &= \int_{-r}^{n} (-1)\left(\frac{e^m - e^{-m}}{e^m + e^{-m}}\right) dm, \qquad r < n < -1 \\ &= \log\left(e^m + e^{-m}\right)^{-1}\Big|_r^n = \log\left(\frac{e^r + e^{-r}}{e^n + e^{-n}}\right) \end{aligned} \qquad (6)$$

**(ii)**

Similarly, for $-1 \le n \le 1$, $f_1(n) = n$, which implies that $f_1(n - m) = n - m$, such that $-1 \le m \le n$.

Therefore,

$$\begin{aligned} f_1(n) \otimes f_1(n) &= \int_{-1}^{n} f_1(n-m)\, f_2(m)\, dm \\ &= \int_{-1}^{n} (n-m)\left(\frac{e^m - e^{-m}}{e^m + e^{-m}}\right) dm \end{aligned} \qquad (7)$$

Using integration by part, and noting that

$$\int uv' = uv - \int u'v$$

Let $u = n - m$. This implies that $du = -dm$.

and $v' = \dfrac{d\left[e^m + e^{-m}\right]}{e^m + e^{-m}}$. This implies that $v = \log\left(e^m + e^{-m}\right)$.

Thus,

$$f_1(n) \otimes f_2(n) = (n-m)\log\left(e^m + e^{-m}\right) + \int_{-1}^{n} \log\left(e^m + e^{-m}\right) dm \tag{8}$$

In (6), let $I = \int_{-1}^{n} \log\left(e^m + e^{-m}\right) dm$

Now, let $x = \log\left(e^m + e^{-m}\right)$, which implies that $e^x = e^m + e^{-m}$

But $x = k \in \mathbb{N}$ for $-1 \le m \le 1$. Hence $I = 0$.

Therefore,

$$f_1(n) \otimes f_2(n) = -(n+1)\log\left(e + e^{-1}\right) \tag{9}$$

**(iii)**

Also, for $n > 1$, $f_1(n) = a = 1$. This implies that $f_1(n - m) = 1$

Therefore,

$$\begin{aligned} f_1(n) \otimes f_2(n) &= \int_{1}^{n} f_1(n-m)\, f_2(m)\, dm \\ &= \int_{1}^{n} \frac{e^m - e^{-m}}{e^m + e^{-m}} dm = log\left(\frac{e^n + e^{-n}}{e + e^{-1}}\right) \end{aligned} \tag{10}$$

The summary of the derived function is given as

$$g_1\left(\sum_{i=0}^{I}\gamma_{hi}x_i\right)g_2\left(\sum_{i=0}^{I}\gamma_{hi}x_i\right)=f(n)=\begin{cases}\log\left(\dfrac{e^r+e^{-r}}{e^n+e^{-n}}\right), & \text{for } n<-1\\(n+1)\log\left(e+e^{-1}\right)^{-1}, & \text{for -1} \le n \le 1\\\log\left(\dfrac{e^n+e^{-n}}{e+e^{-1}}\right), & \text{for } n>1\end{cases} \tag{11}$$

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

## Distributional Properties of the SATLINS_TANH SNN Model

Next it is shown that the derived transfer functions are probability density functions. By definition, the probability density function (*p.d.f*) of function *f*(*x*) of a random variable $X:\Omega\to\mathbb{R}$ is said to be a proper *p.d.f* if for $x\in(-\infty,+\infty), x\in X$ , thus,

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

From the derived transfer function in (11),

$$\int_{-\infty}^{\infty} f_1(n)\otimes f_2(n)\,dn$$
$$=\int_{-\infty}^{-1}\log\left(\frac{e^r+e^{-r}}{e^n+e^{-n}}\right)dn+\int_{-1}^{1}(n+1)\log\left(e+e^{-1}\right)^{-1}dn$$
$$+\int_{1}^{\infty}\log\left(\frac{e^n+e^{-n}}{e+e^{-1}}\right)dn$$

$$= \int_{-\infty}^{-1} \left[ \log\left(e^{r} + e^{-r}\right) - \log\left(e^{n} + e^{-n}\right) \right] dn$$

$$+ \log\left(e + e^{-1}\right)^{-1} \int_{-1}^{1} (n+1)\, dn$$

$$+ \int_{1}^{\infty} \left[ \log\left(e^{n} + e^{-n}\right) - log\left(e + e^{-1}\right) \right] dn$$

$$= \int_{-\infty}^{-1} \log\left(e^{r} + e^{-r}\right) dn + \log\left(e + e^{-1}\right)^{-1} \int_{-1}^{1} (n+1)\, dn - \int_{1}^{\infty} \log\left(e + e^{-1}\right) dn \qquad (12)$$

$$= \left[ n \log\left(e^{r} + e^{-r}\right) \right]_{-\infty}^{-1} + \left[ \log\left(e + e^{-1}\right) \left( \frac{n^{2}}{2} + n \right) \right]_{=1}^{1} - \left[ n \log\left(e + e^{-1}\right) \right]_{1}^{\infty}$$

$$= \infty + 2\log\left(e + e^{-1}\right) - \infty$$

$$= 2\log\left(e + e^{-1}\right)$$

The mean and variance of the derived transfer functions are obtained next.

For $f_1(n) \otimes f_2(n)$

$$f_1(n) \otimes f_2(n) = \begin{cases} \log\left(\dfrac{e^{r} + e^{-r}}{e^{n} + e^{-n}}\right), & \text{for } n < -1 \\ (n+1)\log\left(e + e^{-1}\right)^{-1}, & \text{for } -1 \le n \le 1 \\ \log\left(\dfrac{e^{n} + e^{-n}}{e + e^{-1}}\right), & \text{for } n > 1 \end{cases}$$

$$E(n) = \int_{-\infty}^{\infty} n\left( f_1(n) \otimes f_2(n) \right) dn$$

$$= \int_{-\infty}^{-1} n \log\left( \frac{e^{r} + e^{-r}}{e^{n} + e^{-n}} \right) dn + \int_{-1}^{1} n(n+1)\log\left(e + e^{-1}\right)^{-1} dn + \int_{1}^{\infty} n \log\left( \frac{e^{n} + e^{-n}}{e + e^{-1}} \right) dn$$

$$= \int_{-\infty}^{-1} n \log\left(e^{r} + e^{-r}\right) dn - \int_{-\infty}^{-1} n \log\left(e^{n} + e^{-n}\right) dn$$

$$+ \int_{-1}^{1} n\left(n+1\right) \log\left(e + e^{-1}\right)^{-1} dn$$

$$+ \int_{1}^{\infty} n \log\left(e^{n} + e^{-n}\right) dn - \int_{1}^{\infty} n \log\left(e + e^{-1}\right) dn$$

$$= \log\left(e^{r} + e^{-r}\right) \int_{-\infty}^{-1} n\left(dn\right) - \log\left(e + e^{-1}\right)^{-1} \int_{-1}^{1} \left(n^{2} + n\right) dn - \log\left(e + e^{-1}\right) \int_{1}^{\infty} n\left(dn\right)$$

$$= \log\left(e^{r} + e^{-r}\right) \left[\frac{n^{2}}{2}\right]_{-\infty}^{1} - \log\left(e + e^{-1}\right)^{-1} \left[\frac{n^{3}}{3} + \frac{n^{2}}{2}\right]_{-1}^{1} - \log\left(e + e^{-1}\right) \left[\frac{n^{2}}{2}\right]_{1}^{\infty}$$

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

Hence, the mean of derived transfer function in is

$$E\left(n\right) = \frac{2}{3} \log\left(e + e^{-1}\right)^{-1} \tag{13}$$

Similarly,

$$E\left(n^{2}\right) = \int_{-\infty}^{\infty} n^{2} \left(f_{1}\left(n\right) \otimes f_{2}\left(n\right)\right) dn$$

$$= \int_{-\infty}^{-1} n^{2} \log\left(\frac{e^{r} + e^{-r}}{e^{n} + e^{-n}}\right) dn + \int_{-1}^{1} n^{2} \left(n+1\right) \log\left(e + e^{-1}\right)^{-1} dn + \int_{1}^{\infty} n^{2} \log\left(\frac{e^{n} + e^{-n}}{e + e^{-1}}\right) dn$$

$$= \log\left(e^{r} + e^{-r}\right) \int_{-\infty}^{-1} n^{2} dn - \int_{-\infty}^{-1} n^{2} \log\left(e^{n} + e^{-n}\right) dn$$

$$+ \log\left(e + e^{-1}\right)^{-1} \int_{-1}^{1} \left(n^{3} - n^{2}\right) dn$$

$$+ \int_{1}^{\infty} n^{2} \log\left(e^{n} + e^{-n}\right) dn - \log\left(e + e^{-1}\right) \int_{1}^{\infty} n^{2} dn$$

$$= \log\left(e^r + e^{-r}\right)\left[\frac{n^3}{3}\right]_{-\infty}^{-1} + \log\left(e+e^{-1}\right)^{-1}\left[\frac{n^4}{4} - \frac{n^3}{3}\right]_{-1}^{1} + \log\left(e+e^{-1}\right)\left[\frac{n^3}{3}\right]_{1}^{\infty}$$

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

$$= \frac{2}{3}\log\left(e+e^{-1}\right)^{-1}$$

Therefore, variance of $\left(f_1(n) \otimes f_2(n)\right)$ is

$$\begin{aligned} \operatorname{var}(n) &= E\left(n^2\right) - \left[E(n)\right]^2 \\ &= \frac{2}{3}\log\left(e+e^{-1}\right)^{-1} - \left[\frac{2}{3}\log\left(e+e^{-1}\right)^{-1}\right]^2 \\ &= -\frac{2}{3}\log\left(e+e^{-1}\right) + \frac{4}{9}\left(\log\left(e+e^{-1}\right)\right)^2 \\ &= \log\left(e+e^{-1}\right)\left[\frac{4}{9}\log\left(e+e^{-1}\right) - \frac{2}{3}\right] \end{aligned} \tag{14}$$

Thus,

$$g_1\left(\sum_{i=0}^{I}\gamma_{hi}x_i\right)g_2\left(\sum_{i=0}^{I}\gamma_{hi}x_i\right) = f(n) = \begin{cases} \log\left(\dfrac{e^r + e^{-r}}{e^n + e^{-n}}\right), & \text{for } n < -1 \\ (n+1)\log\left(e+e^{-1}\right)^{-1}, & \text{for } -1 \le n \le 1 \\ \log\left(\dfrac{e^n + e^{-n}}{e + e^{-1}}\right), & \text{for } n > 1 \end{cases}$$

with mean, $E(n) = \dfrac{2}{3}\log\left(e+e^{-1}\right)^{-1}$

and variance, $\operatorname{var}(n) = \log\left(e+e^{-1}\right)\left[\dfrac{4}{9}\log\left(e+e^{-1}\right) - \dfrac{2}{3}\right]$.

## References

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

Anders, U. (1996). *Statistical Model Building for Neural Networks.* AFIR Colloqium. Nunberg, Germany.

Anderson, J. A. (2003). *An Introduction to neural networks.* Upper Saddle River, NJ: Prentice Hall.

Battiti, R. (1992). First- and second-order methods for learning: Between steepest descent and Newton's method. *Neural Computation 4*, 141–166.

Carling, A. (1992). *Introducing Neural Networks*. Ammanford, England: Sigma Press.

Foresee F. D., & Hagan, M. T. (1997). 'Gauss-Newton Approximation to Bayesian Regularization.' In *IEEE International Conference on Neural Networks (Vol. 3)*, 1930–1935. New York: IEEE.

Golden, R.M. (1996). *Mathematical Methods for Neural Network Analysis and Design*. Cambridge, MA: MIT Press.

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

Maren, A., Harston, C., & Pap, R., (1990). *Handbook of Neural Computing Applications*. San Diego, CA: Academic Press.

Nelson, M. M. & Illingworth, W. T. (1991). *A Practical Guide to Neural Nets*. Addison-Wesley Publishing Company.

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

Taylor, J. G. (1999). *Neural networks and their applications*. New York: Wiley.

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