

March 2004

## Asymmetric kernel regression

M. Mackenzie  
*University of Wollongong*

K. Tieu  
*University of Wollongong, ktieu@uow.edu.au*

Follow this and additional works at: <https://ro.uow.edu.au/engpapers>



Part of the [Engineering Commons](#)

<https://ro.uow.edu.au/engpapers/29>

---

### Recommended Citation

Mackenzie, M. and Tieu, K.: Asymmetric kernel regression 2004.  
<https://ro.uow.edu.au/engpapers/29>

# Asymmetric Kernel Regression

Mark Mackenzie and A. Kiet Tieu

**Abstract**—Kernel regression is one model that has been applied to explain or design radial-basis neural networks. Practical application of the kernel regression method has shown that bias errors caused by the boundaries of the data can seriously effect the accuracy of this type of regression. This paper investigates the correction of boundary error by substituting an asymmetric kernel function for the symmetric kernel function at data points close to the boundary. The asymmetric kernel function allows a much closer approach to the boundary to be achieved without adversely effecting the noise-filtering properties of the kernel regression.

**Index Terms**—Asymmetric kernel function, beta density, gamma density, kernel regression.

## I. INTRODUCTION

MOODY and Darken [1] seem to be the first to link the radial-basis neural networks (RBNNs) with the kernel regression of statistics followed shortly after by Specht [2] and Zaknich *et al.* [3]. Kernel regression is one of the many mathematical or physical models, which is often used to explain the characteristics of RBNNs.

In view of the close relationship with kernel regression, many of the important results originally developed for kernel regression can be directly applied to the RBNN. One particularly serious problem that occurs in kernel regression is the introduction of a bias error close to the edges of the data where the symmetrical radial kernel function is truncated by the edge of the data (Fig. 1). This is referred to as the boundary error.

A number of different numerical schemes have been developed in the statistical literature to cope with the boundary error. These can be found in the work on kernel density estimation where many have been summarized by Zhang *et al.* [4]. Probably the simplest method (mentioned in the paper of Hall and Wehrly [5]) is to reduce the window width of the Kernel function in proportion to its distance from the boundary (Fig. 2). Although it is very effective in reducing the bias error, it can cause a significant increase in the random noise, which is inversely proportional to the window width of the kernel function. For this reason, it is probably not practical; but it highlights the fundamental problem with boundary bias correction, which is to correct the bias error without adversely effecting the noise-level performance.

Other boundary-error correction techniques include using specialized boundary kernels at the boundary, reflection of the data about the boundary, and transformation of the data. The boundary kernel method uses linear multiples of the kernel function in the vicinity of the boundary, which are specially

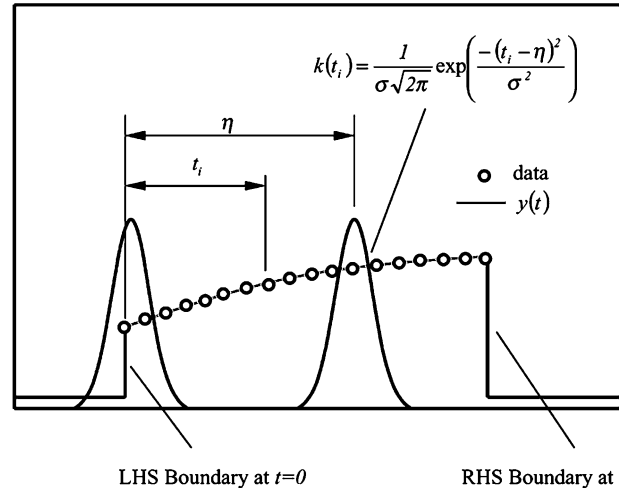


Fig. 1. Gaussian kernel regression showing truncation of the kernel at the boundary.

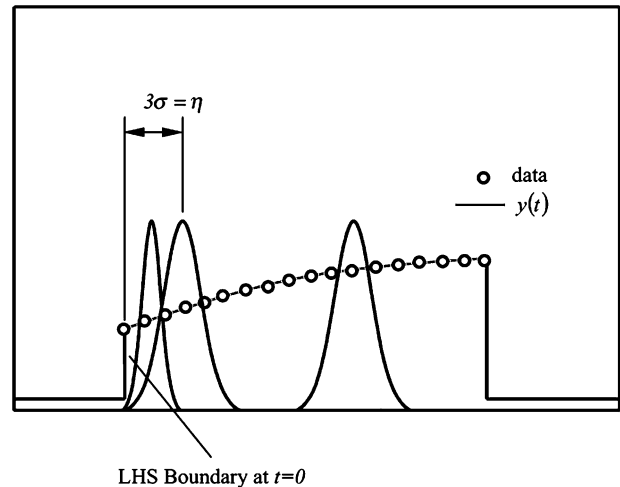


Fig. 2. Correction of boundary bias error by reduction of the variance of the Gaussian kernel close to the boundary.

tailored to reduce the bias error (Gasser and Muller [6]). The locally linear kernel regression estimator of Fan and Gijbels [7] behaves in a similar way. It has since been reported that the boundary kernel method also suffers from an increase in the variance error. The reflection method of Hall and Wehrly [5] provides reduced boundary bias error with a low variance error. Data is reflected about the boundary to prevent the kernel from being truncated by the boundary. While boundary bias error is improved, the bias error will still be larger than in the interior when the data has a nonzero first derivative at the boundary. Marron and Ruppert [8] apply a transformation to the data so that it has a first derivative equal to zero prior to applying the reflection method.

Manuscript received June 5, 2001; revised January 22, 2003.

The authors are with the Mechanical Engineering, University of Wollongong, Wollongong, N.S.W. 2522, Australia (e-mail: kiet\_tieu@uow.edu.au).

Digital Object Identifier 10.1109/TNN.2004.824414

In an effort to provide boundary error correction without compromising the variance error, Zhang *et al.* [4] combine the transformation method and reflection method with the pseudodata method of Cowling and Hall [9]. Further practical application will determine the usefulness and weaknesses of this combined technique. However, the complicated nature of this technique presents difficulties in the application to neural network RBNNs.

The purpose of this paper is to introduce a new boundary error method that can easily be applied to neural networks. The intended application is to signal processing [10] and automatic control [11], where the RBNN/kernel regression has proved particularly useful. The method is based on the replacement of the symmetrical radial kernel function by an asymmetric kernel function in the region immediately adjacent to the boundary. Section III describes the proposed method. Prior to this, it is necessary to first describe the kernel regression and the associated errors, which is covered in Section II. The performance of the asymmetrical kernel is assessed in Section IV.

## II. KERNEL REGRESSION

### A. Description

Kernel regression of statistics was derived independently by Nadaraya [12] and Watson [13] in 1964 with a mathematical foundation given by Parzen's earlier work on kernel density estimation [14]. Wand and Jones [15] have given a thorough extensive review of the development of kernel density estimation since its invention.

Kernel regression is the estimation of the functional relationship  $y(t)$  between two variables  $y$  and  $t$ . Measurement produces a set of random variables  $\{t_i, y_i; i = 1 \dots N\}$  on the interval  $\{0 \leq t_i \leq T\}$ . It is assumed that

$$y_i = y(t_i) + n \quad (1)$$

where  $n$  is a random noise variable with the mean equal to zero and the variance equal to  $E[n^2]$ . The Nadaraya–Watson kernel regression estimate of  $y(t)$  at  $t = \eta$  from this random data is defined as the estimator  $\hat{y}(\eta)$  where

$$\hat{y}(\eta) = \frac{\sum_{i=1}^N y_i k(\eta - t_i)}{\sum_{i=1}^N k(\eta - t_i)}. \quad (2)$$

The function  $k(\eta - t_i)$  is the kernel function, which is traditionally chosen from a wide variety of symmetric functions [14]. In this paper, we use the Gaussian density of statistics as the kernel function, i.e.,

$$k(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-t^2}{2\sigma^2}\right). \quad (3)$$

The kernel regression averages all the data contained within the range of the kernel function centered at  $t = \eta$  (Fig. 1). The effective range of the kernel function is defined by its window width  $\sigma$ . For the Gaussian density, the window width is conveniently set to the standard deviation  $\sigma$ .

Statistical analysis of the Nadaraya–Watson kernel regression is difficult because it is defined as the ratio of two random variables [16]. In many important applications in signal processing and automatic control, a simpler form of the kernel regression may be used, which is considerably easier to analyze statistically. In these applications, the random variable  $t_i$  is the phase error, which usually has a constant pdf. With these type of random variables, the summations involving the kernel function are equivalent to a Monte Carlo integration and

$$\sum_{i=1}^N k(\eta - t_i) \rightarrow N/T. \quad (4)$$

The Nadaraya–Watson kernel regression may then be replaced by the Priestley–Chao [17] regression, which is defined as

$$\hat{y}(\eta) = \frac{T}{N} \sum_{i=1}^N y_i k(\eta - t_i). \quad (5)$$

From an engineering viewpoint, the Priestley–Chao regression can be described as the numerical computation of the Gaussian convolution filter [18], [19] of signal processing. It probably also occurs in many other scientific applications; for example, the laser Doppler anemometer [20] of fluid dynamics is very similar. Due to the simpler analytical statistical properties, in this paper, we only consider this type of kernel regression.

### B. Mean-Square Error

Of major concern in the application of the kernel regression is the mean-square error (mse). This is conveniently described using a statistical approach. The mse of the estimator  $\hat{y}(\eta)$  is defined to be

$$\begin{aligned} \text{mse}[\hat{y}(\eta)] &= E[(\hat{y}(\eta) - y(\eta))^2] \\ &= \text{var}[\hat{y}(\eta)] + \text{bias}[\hat{y}(\eta)]^2 \end{aligned} \quad (6)$$

where the first term is the variance error of  $\hat{y}(\eta)$ , which is

$$\begin{aligned} \text{var}[\hat{y}(\eta)] &= E[(\hat{y}(\eta) - E[\hat{y}(\eta)])^2] \\ &= E[\hat{y}(\eta)^2] - E[\hat{y}(\eta)]^2 \end{aligned} \quad (7)$$

and the second term is the bias error of  $\hat{y}(\eta)$

$$\text{bias}[\hat{y}(\eta)] = E[\hat{y}(\eta) - y(\eta)]. \quad (8)$$

In his 1983 work, Nadaraya [21] provides a mathematical treatment of these errors for the Nadaraya–Watson regression using Parzen's mathematical analysis of kernel density estimation [14]. Further mathematical results may also be found in Hall's 1984 paper [16]. In addition, the Taylor series approximation given by Chu and Marron [22] is useful for the practical application of the regression.

For the Priestley–Chao regression, we have

$$\text{bias}[\hat{y}(\eta)] = \int_{-\infty}^{\infty} y(t_i) k(\eta - t_i) dt_i - y(\eta) \quad (9)$$

and

$$\text{var}[\hat{y}(\eta)] = \frac{T}{N} E[n^2] \int_{-\infty}^{\infty} k^2(t_i) dt_i. \quad (10)$$

By using a Taylor series expansion for  $y(t_i)$  about  $t_i = \eta$ , the bias error is

$$\text{bias}[\hat{y}(\eta)] \cong \frac{\sigma^2}{2} y''(\eta) + \frac{\sigma^4}{8} y'''(\eta). \quad (11)$$

The odd terms of the Taylor series disappear due to the symmetric nature of the kernel function and do not contribute.

The variance error is

$$\text{var}[\hat{y}(\eta)] = \frac{TE[n^2]}{2\sigma N\sqrt{\pi}}. \quad (12)$$

### C. Bandwidth

In order to apply the regression, it is convenient to estimate the standard deviation parameter. The standard deviation can later be optimized using the well-known gradient descent algorithm or one of the techniques developed in statistics [15]. The appropriate approximate standard deviation can be determined from the frequency response of the regression. The 3-dB bandwidth of the kernel regression must be greater than the highest frequency signal that is to be modeled.

For the symmetric kernel function, it is easy to see that in the limit  $N \rightarrow \infty$ , the kernel regression approaches the convolution operator. The output  $\hat{y}(\tau)$  is the convolution of the input  $y(t)$  with the kernel function  $k(t)$ , as follows:

$$\hat{y}(\eta) \cong \lim_{N \rightarrow \infty} \frac{T}{N} \sum_{i=1}^N y_i k(\eta - t_i) = \int_{-\infty}^{\infty} y(t) k(t - \eta) dt. \quad (13)$$

In the frequency domain, convolution is equivalent to multiplication so that

$$\hat{Y}(\omega) = Y(\omega) K(\omega) \quad (14)$$

where  $\hat{Y}(\omega)$ ,  $Y(\omega)$ , and  $K(\omega)$  are the Fourier transforms of  $\hat{y}(\tau)$ ,  $y(t)$ , and  $k(t)$ , respectively.

A Gaussian kernel function has a transfer function of

$$K(\omega) = \exp\left(-\frac{\sigma^2 \omega^2}{2}\right). \quad (15)$$

The frequency response of the Gaussian kernel function to a sinusoidal input is determined by approximating  $k(\omega)$  by a Taylor series expansion, as follows:

$$\begin{aligned} K(\omega) &\approx \frac{1}{1 + \frac{\sigma^2 \omega^2}{2} + \frac{\sigma^4 \omega^4}{8} + \dots a_n \omega^n} \\ &\approx \frac{1}{(1 + \frac{\sigma^2 \omega^2}{2})(1 + \frac{\sigma^4 \omega^4}{8} + \dots)}. \end{aligned} \quad (16)$$

From this approximation, it may be seen that by ignoring the higher order terms, the bandwidth of the Gaussian kernel is

$$\text{frequency(3 dB)} \cong \frac{\sqrt{2}}{2\pi\sigma}. \quad (17)$$

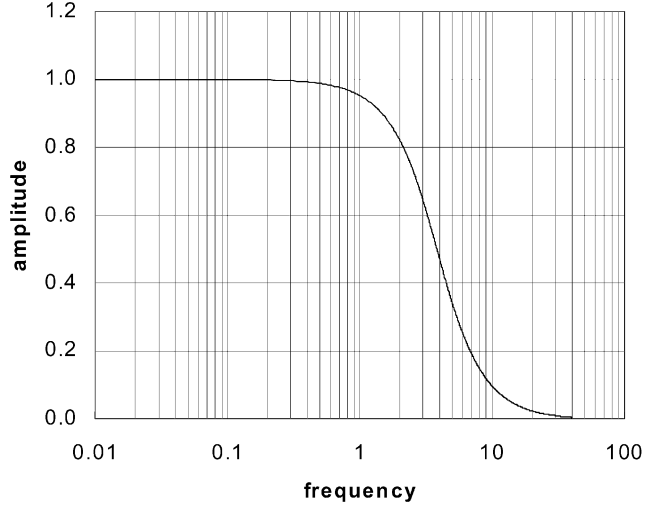


Fig. 3. Frequency response of Gaussian kernel with  $\sigma = 0.05$ .

Fig. 3 shows the frequency response for  $\sigma = 0.05$  of the Gaussian function. The actual bandwidth is about 50% lower than that predicted by (17) due to the higher order terms. From the graph of Fig. 3, the actual bandwidth is

$$\text{frequency(3 dB)} \cong \frac{\sqrt{2}}{10\sigma}. \quad (18)$$

### III. ASYMMETRICAL KERNEL REGRESSION

Although unnoticeable in Parzen's original paper [14], which describes the kernel function estimation of density, the method is defined as being applicable to data on the interval  $\{-\infty < t < \infty\}$ . In practical application of kernel regression, the available data imitates a causal function since it only occurs across a finite interval and is usually taken to be zero for  $t < 0$ . The boundary is defined as  $t = 0$ . Significant bias error occurs in the vicinity of  $t = 0$  because the kernel function odd moments are no longer zero, and these must now be included in the bias error (11). The new expression for bias error close to  $t = 0$  is

$$\begin{aligned} \text{bias}[\hat{y}(\eta)] &= \left\{ y(\eta) \int_0^{\infty} k(t - \eta) dt + y'(\eta) \int_0^{\infty} (t - \eta) k(t - \eta) dt \right. \\ &\quad \left. + \frac{1}{2} y''(\eta) \int_0^{\infty} (t - \eta)^2 k(t - \eta) dt + \dots \right\} - y(\eta). \end{aligned} \quad (19)$$

Fig. 5(a) shows the boundary error for a Gaussian kernel function  $\sigma = 0.05$  estimate of a sine wave of frequency 2.0 arbitrary units. Although  $\sigma$  is sufficiently small that a reasonably good fit has been achieved in the interior, at the boundary the bias error can be quite large.

The Gaussian function is well known in statistics as the limiting probability density of a sum of samples that each have identical probability densities defined on the interval  $\{-\infty < t < \infty\}$ . For samples, which have causal probability densities, the limiting probability density is the Gamma density [23]. The

Gamma density is naturally asymmetric to cope with the discontinuity at  $t = 0$ . In view of the causal nature of the data, we replace the symmetric Gaussian function by a Gamma density in the vicinity of  $t = 0$ .

The Gamma density [Fig. 4(a)] is defined as

$$ga(t) = \frac{t^\alpha \exp\left(-\frac{t}{\beta}\right)}{\beta^{\alpha+1} \Gamma(\alpha+1)}, \quad \alpha, \beta \geq 0 \quad (20)$$

where  $\Gamma(\alpha+1)$  is the factorial function. The parameters  $\alpha$  and  $\beta$  define the shape and location from the origin of the Gamma density. The mean value, variance, and location of the peak of the Gamma density are defined in terms of  $\alpha$  and  $\beta$ , respectively, by [23]

$$\eta = \beta(\alpha + 1) \quad (21)$$

$$\sigma = \beta\sqrt{\alpha + 1} \quad (22)$$

$$p = \alpha\beta. \quad (23)$$

Due to the asymmetry, the mean value is shifted to the right of the peak value of the Gamma density.

It has already been mentioned that the window width of the Gaussian kernel function is equal to its standard deviation, and the mean value of a Gaussian function centered at  $t = \eta$  (Fig. 1) defines the location of the estimate  $\hat{y}(\eta)$ . In the same way, we may define the window width of the Gamma kernel function to be equal to its standard deviation and the mean value  $\eta$  of the Gamma density as the location of the estimate  $\hat{y}(\eta)$ . Therefore, for a particular  $\eta$  and  $\sigma$  [defined by (21) and (22) in terms of  $\alpha$  and  $\beta$ ], the asymmetric kernel regression is

$$\hat{y}(\eta) = \frac{T}{N} \sum_{i=1}^N y_i ga(t_i, \eta, \sigma). \quad (24)$$

In order to preserve the noise performance close to the boundary, the variance  $\sigma$  of the Gamma density must be kept constant for different locations of the mean value  $\eta$ . This can be achieved by forcing  $\alpha$  and  $\beta$  to satisfy the relations

$$\beta = \frac{\sigma^2}{\eta} \quad (25)$$

and

$$\alpha = \frac{\eta^2}{\sigma^2} - 1. \quad (26)$$

Fig. 4(a) shows the Gamma density with  $\sigma = 0.1$  for various values of the mean value  $\eta$ . As  $\eta$  increases, the asymmetry of the Gamma density slowly disappears, and with  $\eta$  large enough, the Gamma density can be approximated by the Gaussian function.

As  $\eta$  decreases, the Gamma density becomes more and more asymmetric until eventually at  $\eta = \sigma$  (corresponding to  $\alpha = 0$ ), the Gamma density becomes an exponential. This is the closest distance to the boundary that is possible with the variance held constant.

The bias error of the Gamma kernel regression is

$$\begin{aligned} \text{bias}[\hat{y}(\eta)] &= \frac{\sigma^2}{2} y''(\eta) + \frac{\sigma^4}{3\eta} y'''(\eta) \\ &+ \frac{\sigma^4}{8} \left\{ 1 + 2 \left( \frac{\sigma}{\eta} \right)^2 \right\} y^{(4)}(\eta) + \dots \end{aligned} \quad (27)$$

Immediately apparent is the absence of a boundary error term. Therefore, the bias error is independent of the distance from the

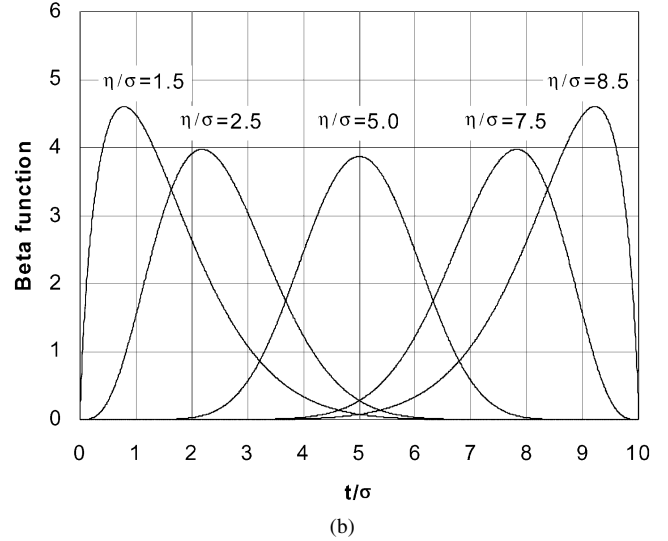
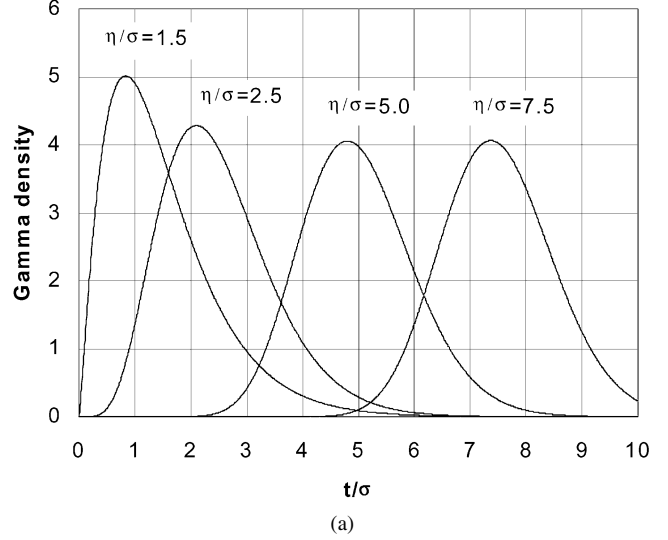


Fig. 4. (a) Gamma kernel functions at different locations from the origin with  $\sigma = 0.1$ . (b) Beta kernel functions across the interval  $\{0 \leq t/\sigma \leq 10.0\}$  with  $\sigma = 0.1$ .

boundary. Also apparent is that the odd terms of order  $\{3\}$  of the Taylor series must now be included due to the skewed nature of the asymmetric Gamma density. Consequently, at interior points, the bias error will be slightly larger than the Gaussian regression. For large  $\eta$ , the asymmetry disappears, and the bias error will approach the error of the Gaussian function.

The variance error of the Gamma density is

$$\text{var}[\hat{y}(\eta)] \cong \frac{E[n^2]}{2\sigma N \sqrt{\pi}}. \quad (28)$$

Because the standard deviation of the Gamma density has been held fixed, this error is also independent of the distance from the boundary.

In the limit of large  $N$ , the asymmetric Gamma kernel regression approaches the integral

$$\begin{aligned} \hat{y}(\eta) &\cong \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N y_i ga(t_i, \eta, \sigma) \\ &= \int_0^\infty y(t) ga(t, \eta, \sigma) dt \end{aligned} \quad (29)$$

which is not a convolution operator. However, for large  $\eta$ , the Gamma density is approximately equal to a Gaussian function and the 3-dB bandwidth of (18) is appropriate.

The procedure described so far is suitable for a boundary on the left at  $t = 0$ . For a boundary on the right at  $t = a$ , the Gamma density can still be applied by change of variables to

$$t \rightarrow a - t. \quad (30)$$

Alternatively, the Beta density can be used. The Beta density [Fig. 4(b)] is the limiting probability density for a sum of samples each having identical probability densities on a finite interval [24]. The Beta density on the interval  $\{0 \leq t \leq 1\}$  is

$$b(t) = \begin{cases} \frac{\Gamma(\alpha+\beta+2)t^\alpha(1-t)^\beta}{\Gamma(\alpha+1)\Gamma(\beta+1)}, & \text{for } 0 < t < 1 \\ 0, & \text{for } t \geq 1 \\ 0, & \text{for } t \leq 0. \end{cases} \quad (31)$$

The variance and mean value of the Beta density, for a particular  $\alpha$  and  $\beta$ , are obtained by the solution of the simultaneous equations

$$\alpha + 1 = \frac{\eta(\eta - \eta^2 - \sigma^2)}{\sigma^2} \quad (32)$$

and

$$\beta + 1 = \frac{(1 - \eta)(\eta - \eta^2 - \sigma^2)}{\sigma^2}. \quad (33)$$

Beta kernel regression is applied in a similar way to the Gamma kernel regression. An appropriate standard deviation  $\sigma$  is chosen based on the interior data at points well away from the boundaries. To preserve the performance against noisy data, this is kept constant for each location  $\eta$  across the full range of the data. The particular Beta density to use at a location  $\eta$  is obtained from (32) and (33). Data on a different interval to  $\{0 \leq t \leq 1\}$  can be scaled so that it spans this interval.

Fig. 4(b) shows the Beta density with  $\sigma = 0.1$  at different locations across the interval  $\{0 \leq t \leq 1\}$ . It is similar in appearance to the Gamma density, although not as peaked. The variance can be held constant while  $\alpha$  and  $\beta \geq 0$ . For small  $\sigma$ , the points at which  $\alpha$  and  $\beta = 0$  may be shown to occur when  $\eta \cong \sigma$  and  $\eta \cong 1 - \sigma$ , respectively.

#### IV. APPLICATION

Simulated data consisted of a unit amplitude sine wave of frequency 0.5 (arbitrary units). The nominal variance of the kernel regression estimators was set at  $\sigma = 0.05$  which is within the frequency bandwidth of the kernel regression (18). This provides for a low bias error at interior points. All the results are shown normalized with respect to this variance.

Fig. 5(a) and (b) shows the error of a Gaussian kernel regression and the Gamma kernel regression, as a function of the distance  $t$  from the boundary, which is at  $t = 0$ . In this comparison, no noise has been added to the signal. Consequently the error shown is due to the bias error. The right-hand side (RHS) of the sine wave, which is far away from the boundary and free from boundary error, provides a useful reference. Ideally, if it were free of boundary error, the left-hand side (LHS) would be the same as the RHS. As shown in Fig. 5(b), the Gaussian regression mirrors the RHS until  $t < 2.5\sigma$ , where it begins to deviate from the RHS, indicating the boundary error begins at this point.

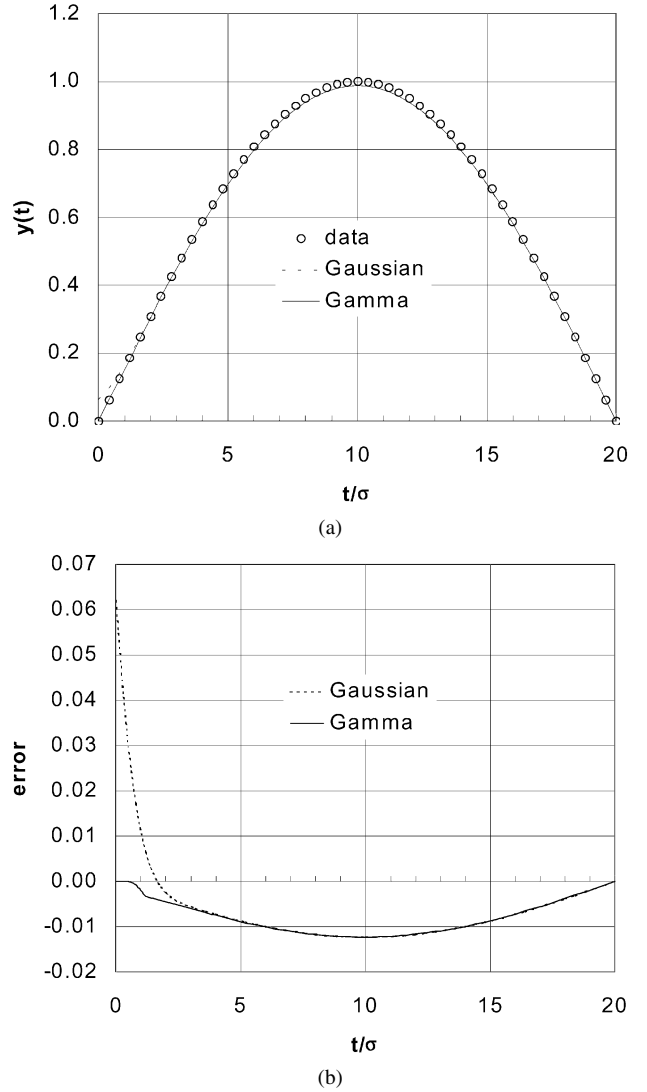


Fig. 5. (a) Gamma and Gaussian kernel regression of sine wave with boundary at  $t = 0$ . (b) Gamma kernel regression error of sine wave of Fig. 5(a).

At closer distances to the boundary, it increases rapidly. This is to be expected because the Gaussian function will begin to be truncated by the boundary for  $t < 3\sigma$ . The Gaussian bias error passes through the axes at  $t \cong 1.5\sigma$ , which is due to the positive boundary error canceling the negative interior bias error. The Gamma regression allows a much closer approach to the boundary; it mirrors the RHS until  $t = \sigma$ . In the interior away from the influence of the boundary, for  $t > 3\sigma$ , the error of the Gamma regression is almost the same as the Gaussian regression. This indicates that the third-order skew term of the Gamma bias error (27) does not significantly degrade the performance of the Gamma regression. In the interior, there is no advantage in using the Gamma function.

For distances  $t < \sigma$  the standard deviation of the Gamma regression cannot be kept constant. In this region, we fixed  $\alpha = 0$ , an exponential function, and reduced the standard deviation according to

$$\sigma = \eta = t \quad (34)$$

where  $t$  is the distance from the boundary. The reduction of variance for  $t < \sigma$  causes the discontinuity in the Gamma curve

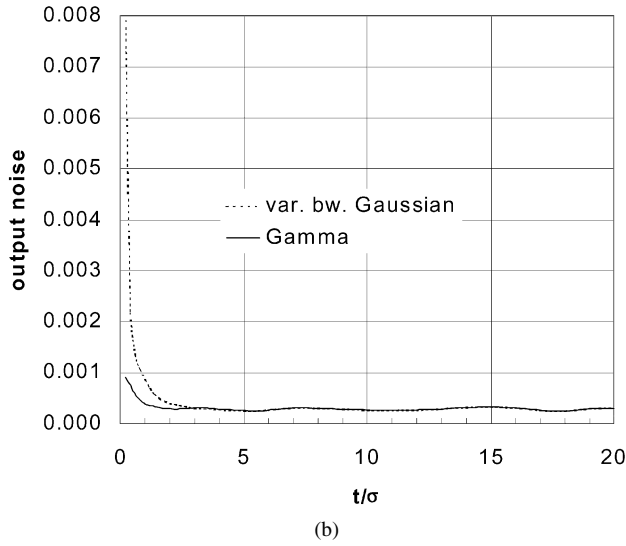
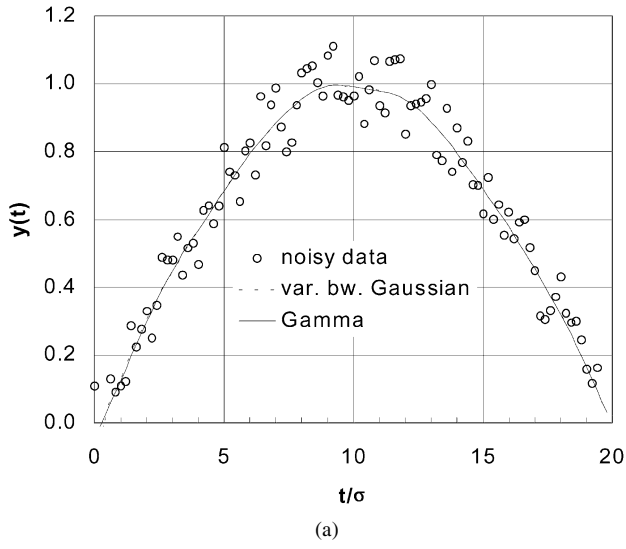


Fig. 6. (a) Comparison of Gamma regression and variable-bandwidth Gaussian regression of simulated noisy sine wave with boundary at  $t = 0$ . (b) Output noise of Gamma regression and variable-bandwidth Gaussian regression on simulated noisy sine wave (200 records).

at  $t = \sigma$ , shown in Fig. 5(b). Although this method removes the boundary error, the noise performance in this region will be degraded as the boundary is approached due to the decreasing variance.

In order to investigate the noise performance, simulated uniform density noise with amplitude in the range 0–0.25 was added to the sine wave signal [Fig. 6(a)]. The simulated noise was generated from a pseudorandom sequence. To provide a quantitative measure, the Gamma regression was compared with a boundary-error-corrected Gaussian regression using the simple method described in Hall and Wehrley [5]. For this type of boundary error correction, the standard deviation of the Gaussian function is varied according to the distance from the wall in such a way that

$$\sigma = \frac{t}{3}. \quad (35)$$

The output noise shown in Fig. 6(b) was the ensemble average from 200 records, with each record containing 100 data points.

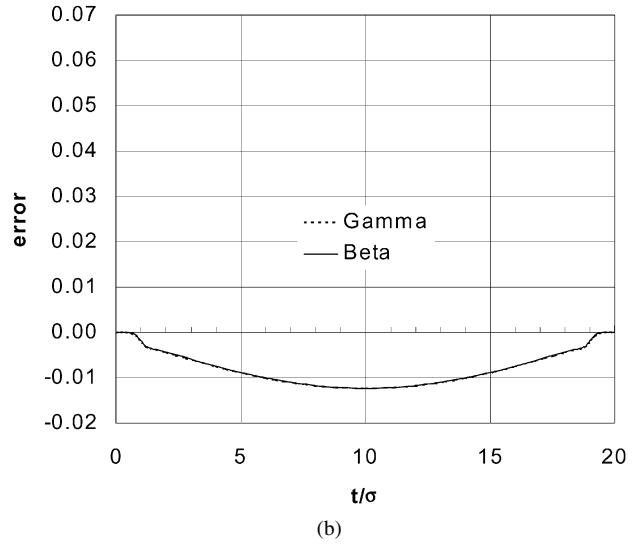
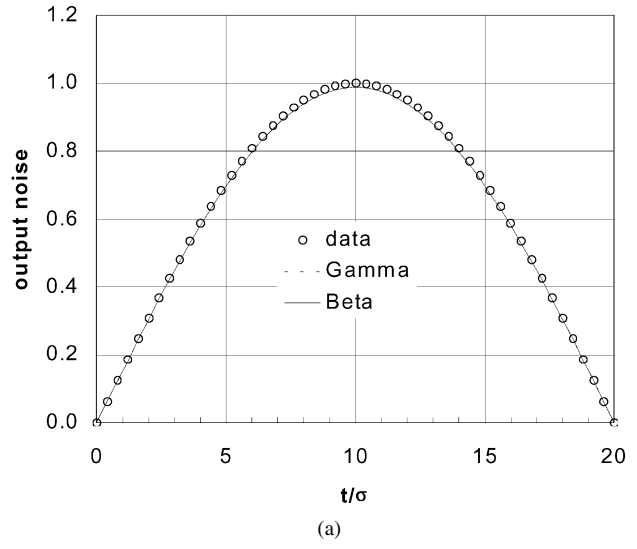


Fig. 7. (a) Beta and Gamma kernel regression boundary error correction of finite duration sine wave. (b) Beta and Gamma kernel regression error of finite duration sine wave of Fig. 7(a).

A single record is shown in Fig. 6(a). The output noise is the square of the difference between the regression operating on the noisy signal and the regression operating on the signal without noise

$$\text{output noise}(\eta) = E[(\hat{y}(\eta; t_i, y_i + n_i) - \hat{y}(\eta; t_i, y_i))^2]. \quad (36)$$

As shown in Fig. 6(b), the output noise of the Gamma regression remains flat as the boundary is approached until  $t \cong \sigma$ . Beyond this point, it increases inversely as the distance from the boundary decreases. The increase in this region is due to the gradual reduction of the standard deviation, which from (34) is proportional to the distance from the boundary. This result is in agreement with that predicted theoretically (28). The boundary-corrected Gaussian regression behaves in a similar way except that the noise begins to increase at a greater distance from the boundary at  $t \cong 3\sigma$ , in accordance with (35). This leads to a much higher output noise close to the boundary than the Gamma regression.

The Beta kernel regression was investigated by adding another boundary on the right at  $t = 1.0(t/\sigma = 20)$ . Figs. 7(a)

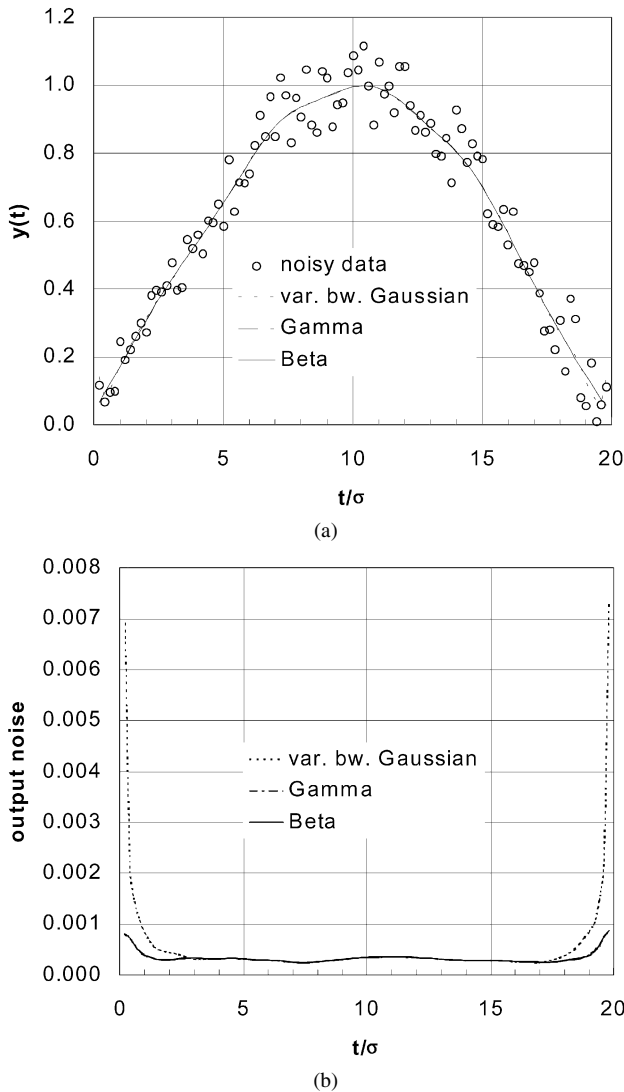


Fig. 8. (a) Comparison of Beta regression, Gamma regression, and variable-bandwidth Gaussian regression of simulated noisy finite sine wave. (b) Output noise of Beta regression, Gamma regression, and variable-bandwidth Gaussian regression of simulated noisy finite sine wave (200 records).

and (b) and 8(a) and (b) show the results. Its performance was indistinguishable from the Gamma density where, for the Gamma density, the change of variables given by (30) was used to correct the RHS boundary. At very large  $\sigma$  with respect to the interval  $\{0 \leq t \leq 1.0\}$ , of the order  $\sigma \approx 0.2$ , neither performed very well due to the poor frequency response.

## V. CONCLUSION

Symmetric kernel regression is subject to significant bias error caused by the boundaries of the experimental data. For one-dimensional data, there are two boundaries. A Gamma kernel function was proposed to correct the error at both boundaries. Theoretical arguments supported by numerical simulation indicate that the Gamma kernel is capable of correcting the boundary error, without sacrificing the noise performance at distances from the boundary as close as  $\eta = \sigma$ . A Beta kernel function was also investigated. Numerical simulations indicate that its performance is similar to the Gamma kernel regression.

## REFERENCES

- [1] J. Moody and C. J. Darken, "Fast learning in networks of locally-tuned processing units," *Neural Comput.*, vol. 1, pp. 281–294, 1989.
- [2] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Networks*, vol. 2, pp. 568–576, Nov. 1991.
- [3] A. Zaknich, C. de Silva, and Y. Attikouzel, "The probabilistic neural network for nonlinear time series analysis," in *Proc. IEEE Int. Joint Conf. Neural Networks (IJCNN)*, Singapore, Nov. 17–21, 1991, pp. 1530–1535.
- [4] S. Zhang, R. J. Karunamuni, and M. C. Jones, "An improved estimator of the density function at the boundary," *J. Amer. Statistical Assn.*, vol. 94, pp. 1231–1241, 1999.
- [5] P. Hall and T. E. Wehrly, "A geometrical method for removing edge effects from kernel-type nonparametric regression estimators," *J. Amer. Statistical Assn.*, vol. 86, pp. 665–672, 1991.
- [6] T. Gasser and H.-G. Müller, "Kernel estimation of regression functions," in *Smoothing Techniques for Curve Estimation*, T. Gasser and M. Rosenblatt, Eds. Heidelberg, Germany: Springer-Verlag, 1979, pp. 23–68.
- [7] J. Fan and I. Gijbels, "Variable bandwidth and local linear regression smoothers," *Annals Statistics*, vol. 20, pp. 2008–2036, 1992.
- [8] J. S. Marron and D. Ruppert, "Transformations to reduce boundary bias in kernel density estimation," *J. Royal Statistical Soc., Ser. B*, vol. 56, pp. 653–671, 1994.
- [9] A. Cowling and P. Hall, "On pseudo data methods for removing boundary effects in kernel density estimation," *J. Royal Statistical Soc., Ser. B*, vol. 58, pp. 551–563, 1996.
- [10] A. Zaknich, "Introduction to the modified probabilistic neural network for general signal processing applications," *IEEE Trans. Signal Processing*, vol. 46, pp. 1980–1990, July 1998.
- [11] D. Sbarbaro, J. P. Segovia, S. Alcozer, and J. Gonzales, "Applications of radial basis network technology to process control," *IEEE Trans. Contr. Syst. Technol.*, vol. 8, pp. 14–22, Jan. 2000.
- [12] E. A. Nadaraya, "On estimating regression," *Theory Probab. Appl.*, vol. 10, pp. 186–190, 1964.
- [13] G. S. Watson, "Smooth regression analysis," *Sankhya Ser. A*, vol. 26, pp. 101–116, 1964.
- [14] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statistics*, vol. 33, pp. 1065–1076.
- [15] M. P. Wand and M. C. Jones, *Kernel Smoothing*. London, U.K.: Chapman & Hall, 1995.
- [16] P. Hall, "Integrated square error properties of kernel estimators of regression functions," *Annals Statistics*, vol. 12, pp. 241–260, 1984.
- [17] M. B. Priestley and M. T. Chao, "Non-parametric function fitting," *J. Royal Statistical Soc., Ser. B*, vol. 34, pp. 385–392, 1972.
- [18] M. Dishal, "Gaussian-response filter design," *Elec. Commun.*, vol. 36, pp. 3–26, 1959.
- [19] H. J. Blinichoff and A. I. Zverev, *Filtering in the Time and Frequency Domains*. New York: Wiley, 1976, ch. 3.
- [20] F. L. Crosswy and J. O. Hornkohl, "Signal conditioning electronics for a laser vector velocimeter," *Rev. Sci. Instrum.*, vol. 44, pp. 1324–1332, 1973.
- [21] E. A. Nadaraya, *Nonparametric Estimation of Probability Densities and Regression Curves*. Dordrecht, The Netherlands: Kluwer Academic, 1989, ch. 4.
- [22] C. K. Chu and J. S. Marron, "Choosing a kernel regression estimator," *Statistical Sci.*, vol. 6, pp. 404–436, 1992.
- [23] A. Papoulis, *The Fourier integral and its applications*. New York: McGraw-Hill, 1962, pp. 234–237.
- [24] —, *The Fourier integral and its applications*. New York: McGraw-Hill, 1962, pp. 238–239.

**Mark Mackenzie** received the B.E. Honors degree in mechanical engineering from the University of Adelaide, Adelaide, Australia, in 1988 and the Master's degree in mechanical engineering from the University of Wollongong, Wollongong, Australia, in 1995. He is currently working toward the Ph.D. degree at the same university.

**A. Kiet Tieu** received the B.E. Honors degree in mechanical engineering and the Ph.D. degree from the University of Western Australia, Crawley, in 1970 and 1975, respectively.

He worked for B.H.P. Steel for eight years and has been with the University of Wollongong, Wollongong, Australia, for 20 years, where he is now a Professor and Head of School of Mechanical, Materials and Mechatronics Engineering.