27th AIAA Aerodynamic Measurement Technology and Ground Testing Conference
28 June - 1 July 2010, Chicago, Illinois

AIAA 2010-4923

Model Robust Calibration: Method and Application to Electronically-Scanned Pressure Transducers

Eric L. Walker *
NASA Langley Research Center,
Hampton, VA 23681

B. Alden Starnes †
Department of Mathematics, Carson-Newman College,
Jefferson City, TN 37760

Jeffery B. Birch ‡
Department of Statistics, Virginia Polytechnic Institute and State University,
Blacksburg, VA 24061-0439

and

James E. Mays §
Department of Mathematical Sciences, Virginia Commonwealth University,
1001 West Main Street, P.O. Box 842014,
Richmond, VA 23284-2014

This article presents the application of a recently developed statistical regression method to the controlled instrument calibration problem. The statistical method of Model Robust Regression (MRR), developed by Mays, Birch, and Starnes, is shown to improve instrument calibration by reducing the reliance of the calibration on a predetermined parametric (e.g. polynomial, exponential, logarithmic) model. This is accomplished by allowing fits from the predetermined parametric model to be augmented by a certain portion of a fit to the residuals from the initial regression using a nonparametric (locally parametric) regression technique. The method is demonstrated for the absolute scale calibration of silicon-based pressure transducers.

Nomenclature

b	Bandwidth
CCMSE	Mean square error calculated using cross-check data
d	Degree of polynomial
$d.f.$	Degrees of freedom

*Research Engineer. Configuration Aerodynamics Branch. Research Directorate. NASA Langley Research Center. Lifetime Senior Member AIAA.
†Associate Professor of Mathematics. Carson-Newman College.
‡Professor of Statistics. Virginia Polytechnic Institute and State University.
§Associate Professor of Statistics. Virginia Commonwealth University.

$\mathbf{H}$	Hat matrix $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$
k	Number of regressor variables
$m.d.f.$	Degrees of freedom used by a particular model, $tr(\mathbf{H})$
MSE	Mean square error
n	Number of samples
r, R	Residual
$\dot{r}$	Predicted or estimated residuals
$\bar{R}$	Average Range
tr	Trace of a matrix
V	Voltage
V_0	Voltage at zero pressure
$\mathbf{W}$	Diagonal weight matrix
$\mathbf{X}$	Model matrix
$\mathbf{x}$	Set condition
$\mathbf{X}'$	Transpose of the model matrix
X	Regressor variable
$\mathbf{y}$	Response
$\hat{\mathbf{y}}$	Estimated/predicted response
Z	Zero shifted voltage, $Z = V - V_0$
$\boldsymbol{\beta}$	Unknown parameters
$\hat{\boldsymbol{\beta}}$	Estimate of the unknown parameters
$\hat{\sigma}$	Estimated standard deviation
σ^2	Variance
λ	MRR mixing parameter
λ^*	Asymptotically optimal MRR mixing parameter
$\hat{\lambda}^*$	Estimated Asymptotically optimal MRR mixing parameter

Sub/Superscripts

0	Quantity or vector at a specified location
LLR	Local linear regression
LPR	Local polynomial regression
MRR1	Model Robust Regression 1, see Section II.C.1
MRR2	Model Robust Regression 2, see Section II.C.2
ker	Kernel regression
NP	Nonparametric
ols	Ordinary least squares
P	Parametric

I. Introduction

It is well understood that no instrument behaves exactly according to a predefined theory. Occasionally, the theory does not adequately represent the physical behavior of a system, and there often exists an intrinsic difference in individual instruments due to variability in the manufacturing process. At times, no theory is available for an instrument's behavior, and one must rely on experience and data analysis.

This inherent deficiency in the manufacturing process is often seen with diaphragm based pressure transducers. No two transducers will behave exactly the same way. All have properties that vary from one to another, whether it is the material properties, mounting differences, or micro fractures in the diaphragm itself. However, the transducers do respond similarly in a broad sense, which allows the definition of a calibration model.

Electronically Scanned Pressure (ESP) transducers all have a nearly linear response to pressure. To account for the slight nonlinearity of the response curve, higher order terms are often included in the calibration. These higher order terms perform well for the small curvature present in the higher pressure region of the gauge; however, near zero these adjustments are inadequate and induce calibration error. This presents a problem for very low pressure measurements in both the absolute and differential regimes. Typical polynomial models do not account well for this behavior. Everhart[1] described this phenomena and proposed a more representative mathematical form, which will be discussed later. Since much effort has

been conducted in development of a parametric model for this application, this will be used to highlight the power of the current method.

For simplicity, this paper will consider the univariate calibration problem. Thus, for any given calibration, there are n pairs of measurements (x_i, y_i), $i = 1, \ldots, n$, where x_i is the set condition and y_i is the instrument response. In a typical regression model y is viewed as a function of x, $y = y(x)$; however, in the case of an instrument calibration, it is often advantageous to use an inverse calibration model, $x = x(y)$. The goal of the instrument calibration is to obtain a function or mapping of the instrument response y that will yield the condition that the instrument is measuring. Since the inverse regression model is used in practice for instrument calibration, this will be the main focus for the present article.

Another distinction that is necessary for this article is that of controlled versus random calibrations. According to Cheng and Van Ness,[2] "If the x's in the calibration stage are at fixed pre-chosen levels, the calibration is said to be controlled, but if the x's are random, the calibration is said to be random." This article will deal with controlled calibrations. The contamination or noise level dealt with here is at or below one percent of the instrument's full-scale output. Random calibrations typically deal with values that have contamination of up to ten percent of the full-scale output. Further, it is assumed that outliers are negligible, and the instrument response will incorporate any local instrument bias and inherit only the bias of the calibration standard.

In regression, the terms "parametric" and "nonparametric" signify two different approaches to the modeling process. In parametric regression, a known or specified model form is used for fitting the data. Generally, this form comes from either prior experience, or scientific theory. The problem then becomes one of determining values for the parameters of the model. In nonparametric regression, the user approaches the problem with no known model and attempts to fit a curve to the data points by joining a sequence of curve sections obtained through localized modeling. The final result is a regression curve not represented by any global form. Most often, this nonparametric approach is used when a theoretical reference curve is unavailable for a process. Semiparametric regression is a relatively new approach to the problem. This technique involves fitting the data both parametrically and nonparametrically, and then combining the results to form a curve that is based on a suitable theoretical form, yet still being able to adapt to aberrations from that form.

The purpose for this work is to demonstrate a methodology, called Model Robust Regression (MRR), for improving the use of data taken for the express purpose of controlled calibration. The advantage of MRR methodology is that it is not entirely dependent on the functional form of the parametric regression model. Given this advantage, any knowledge or understanding of the problem that can be represented in a parametric form can then be incorporated to yield an increase in calibration accuracy.

For completeness, this paper will review the pertinent regression methods, including ordinary least squares in Section II. Details of the MRR for mixing of the parametric and nonparametric models are given in Section III. The statistical measures developed for comparison of methods and overall measures of goodness of fits and predictions will be presented in Section IV. Section V gives a brief description of the pressure instrumentation used in this study. The parametric and MRR regressions are applied to the data in Sections VI and VII, respectively. Section VIII contains a discussion of the properties of the MRR procedure with cited examples. Readers interested in the mathematical foundation of the regression methods should begin in Section II; others only interested in the application and results of the current procedure should begin in Section V.

II. Regression Methodology

The following review of regression methodology is comprised of excerpts taken from Mays, Birch, and Starnes.[3] The basic problem of regression is explaining the response of a variable y by using k (with $k \geq 1$) regressor variables, $X_1, X_2, \ldots, X_k$, according to a model of the form $y = g(X_1, X_2, \ldots, X_k) + e$. The term e is a random error from the process, often assumed to have mean 0 and variance σ^2. Of interest is when the researcher has some knowledge of the form of g, but there are regions in the data that do not conform sufficiently to this underlying structure. Of particular interest is dealing with small data sets where the sample size n (the number of independent observations of $y, X_1, \ldots, X_k$) is small, say less than 50. This situation is often the case in calibration, especially when it is done in-situ.

A. Ordinary Least Squares (OLS)

The parametric approach to regression is to assume a known linear form for g, rewrite the model as $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$, and estimate the unknown parameters in $\boldsymbol{\beta}$ from the data. Here $\mathbf{X}$ is an $n \times (k+1)$ model matrix of k regressors augmented with a column of ones. Ordinary least squares (OLS) minimizes the sum of squared residuals to obtain the fitted values as

$$\hat{\mathbf{y}}_{\text{ols}} = \mathbf{X}\hat{\boldsymbol{\beta}}_{\text{ols}} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = \mathbf{H}^{(\text{ols})}\mathbf{y}, \tag{1}$$

where $\mathbf{H}^{(\text{ols})}$ is the OLS "hat" matrix (Myers[4]). The fit for y_i at location X_i may be expressed as $\hat{y}_i^{(\text{ols})} = \sum_{j=1}^{n} h_{ij}^{(\text{ols})} y_j$, a weighted sum of the n observations y_j, where $h_{i1}^{(\text{ols})}, \ldots, h_{in}^{(\text{ols})}$ are the elements of the ith row of $\mathbf{H}^{(\text{ols})}$. The weighting scheme is a direct consequence of the prescribed model. If the model is misspecified, even over subsets of the data, then inferences from OLS will be misleading.

B. Nonparametric Regression

An alternative approach is to use nonparametric regression, which is most useful in situations where the function g is considered to be unknown and the user has no knowledge of its form. Local linear regression (LLR) receives the emphasis here. The prediction of $g(\mathbf{x}_0)$ at location $\mathbf{x}_0' = (X_{10} \ldots X_{k0})$ uses weighted least squares, where the weights are a decreasing function of the distances of the respective regressor locations from $\mathbf{x}_0$. This work emphasizes the "single regressor model" where $k = 1$. That is, $\mathbf{x}_0' = (X_0)$. Extensions to cases where $k > 1$ can easily be made. Since nonparametric techniques rely totally on the data and have no specific functional form, they may fit to irregular patterns in the data and may be more variable than parametric fits. In other words, a nonparametric method can mistake noise for a signal.

Thus, a dilemma is created for the common problem that occurs when there is *partial* knowledge about the underlying model. For example, a parametric model might explain most of the data, but could be inadequate in several regions of the independent variable space, whereas a nonparametric fit would simply ignore the important information that the engineer may have about the underlying structure. Also, it is possible that even with the presence of important deviations from the underlying model, a lack of fit test would incorrectly conclude that the parametric fit is adequate (Mays, Birch, Starnes[3]).

There are several possible solutions to the above dilemma. Two recently developed model robust methods, MRR1 and MRR2, are presented in this work. MRR1 is shown in contrast to the methodology of MRR2; however, MRR2 has been selected as the regression method of choice based on analysis in Mays,[5] and Mays, Birch, Starnes.[3] Another method, partial linear regression (PLR) was developed by Speckman,[6] and is discussed by Mays[3, 5] for purposes of comparison. The goal is to obtain a simple procedure that maintains robustness to a misspecified model while adequately fitting the data (*i.e.,* to find a model that adequately represents the data, even if there is an error in the initially assumed model form). These techniques contribute to the growing list of literature dealing with semiparametric regression (described below). Such references include work by Speckman,[6] Burman and Chaudhuri,[7] Rahman, Gokhale, and Ullah,[8] Hjort and Glad,[9] Eubank and Spiegelman,[10] and Newey.[11] Work by Wooldridge,[12] Robinson,[13] and Ullah and Vinod[14] exemplify the growing use of semiparametric regression in areas such as economics.

Härdle[15] presents a thorough discussion of kernel regression and other popular nonparametric procedures. Wand and Jones[16] and Fan and Gijbels[17] also address nonparametric fitting procedures, and include detailed coverage of local least squares techniques. Kernel regression gives fitted values according to the expression $\hat{y}_i^{(\text{ker})} (= \hat{g}(X_i)) = \sum h_{ij}^{(\text{ker})} y_j$, for $j = 1, \ldots, n$. These fits may also be expressed as

$$\hat{\mathbf{y}}_{\text{ker}} = \mathbf{H}^{(\text{ker})}\mathbf{y}, \tag{2}$$

where $\mathbf{H}^{(\text{ker})} = [h_{ij}^{(\text{ker})}]$ is denoted as the kernel "hat" matrix. The current work uses the Nadaraya-Watson kernel weights

$$h_{ij}^{(\text{ker})} = K\left(\frac{X_i - X_j}{b}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{X_i - X_j}{b}\right) \tag{3}$$

(see Wand and Jones[16]), where the kernel function $K(u)$ is a decreasing function of $|u|$, and $b > 0$ is the bandwidth (smoothing parameter), which determines how rapidly the weights decrease as the distance from the location being fit increases. Other weighting schemes have been introduced and discussed by Priestley

and Chao,[18] Gasser and Müller,[19] and Chu and Marron.[20] The kernel function employed in this current work is the simplified Normal (or Gaussian) kernel given by $K(u) = e^{(-u^2)}$, where u is a real number.

The bandwidth b controls the smoothness of the resulting estimate of g and in practice is crucial in obtaining a "good" kernel fit that strikes the proper balance between the variance, which may be high when b is too small, and the bias (or squared bias), which may be high when b is too large. The normalized bandwidth, b_0, is determined for the unit normalized regressor axis by subtracting the offset and dividing by the range. Mays and Birch[21,22] discuss bandwidth choice considerations in nonparametric regression in general and new considerations brought about by the model-robust procedures described below. Key issues in bandwidth choice development are the minimization of some mean squared error criterion (Härdle,[15] Rice,[23] Chiu,[24] Härdle and Marron,[25] Chu and Marron,[20] Faraway,[26] and Ruppert et al.[27]), the use of the leave-one-out criterion PRESS $= \sum(y_i - \hat{y}_{i,-i})^2$ (Allen[28]), where $\hat{y}_{i,-i}$ is the fit obtained at x_i when the data point (y_i, x_i) has been temporarily removed from the data set, and the use of penalizing functions to protect against small bandwidths (Rice,[23] Härdle,[15] and Härdle et al.[29]). Combining these methods, and adding a penalty for large bandwidths, Mays and Birch[21,22] propose a new "penalized PRESS" bandwidth selector called PRESS**. This new selector is needed due to new smoothing considerations for the model-robust methods. In general, for selecting a smoothing or mixing parameter θ for the nonparametric or model-robust methods, each having a hat matrix $\mathbf{H}$, and defining SSE_θ to be the sum of squares error when using the chosen value θ, and $\text{SSE}_{\max}$ to be the maximum sum of squares error across all θ values (e.g., when $b = 1$ for kernel regression with X values scaled from 0 to 1), PRESS** is defined as

$$\text{PRESS}^{**} = \frac{\text{PRESS}}{n - \text{tr}(\mathbf{H}) + (n-1)\frac{\text{SSE}_{\max} - \text{SSE}_\theta}{\text{SSE}_{\max}}}. \tag{4}$$

Also discussed in Mays and Birch[21,22] is the fact that PRESS** may yield one or two local minimum values (giving the appropriate bandwidth value and/or $b = 1$), and a graphical remedy for this possibility.

The above discussion deals with "global" bandwidths, but locally adaptive bandwidth selectors may prove beneficial in some instances (Altman,[30] Müller and Stadtmüller,[31] Härdle and Bowman,[32] and Faraway[26]). These more computationally advanced "improvements" to bandwidth selection may be studied in the future, but should benefit each of the procedures similarly. For a discussion on bandwidth choice in the related field of density estimation, see Park and Marron.[33]

A major problem inherent in kernel regression is predicting at the boundaries of the data, where the kernel weights become asymmetric, and bias and variance can be affected. Local polynomial regression (LPR), introduced by Cleveland[34] and discussed as a boundary bias correction by Hastie and Loader,[35] is used here to eliminate boundary bias and other problems. This technique obtains the fitted value $\hat{y}_i$ as the fitted value of a dth degree polynomial fit to the data using weighted least squares regression, where the weights w_{ij} are assigned to each observation based on an initial kernel fit to the data. The LPR fit to y_i at location X_i is obtained as

$$\begin{aligned}
\hat{\mathbf{y}}_i^{(\text{LPR})} &= \mathbf{x}_i'^{(\text{LPR})} \hat{\boldsymbol{\beta}}_i^{(\text{LPR})} \\
&= \mathbf{x}_i'^{(\text{LPR})} \left[\mathbf{X}'^{(\text{LPR})} \mathbf{W}^{\text{LPR}}(X_i) \mathbf{X}^{(\text{LPR})} \right]^{-1} \mathbf{X}'^{(\text{LPR})} \mathbf{W}^{\text{LPR}}(X_i) \mathbf{y} \\
&= \mathbf{h}_i'^{(\text{LPR})} \mathbf{y},
\end{aligned} \tag{5}$$

where $\mathbf{W}^{\text{LPR}}(X_i)$ is the diagonal weight matrix with diagonal elements consisting of the n elements of the ith row of $\mathbf{H}^{(\text{ker})}$, and $\mathbf{X}^{(\text{LPR})}$ is the model matrix for LPR determined by the degree d of the polynomial, with ith row defined as $\mathbf{x}_i'^{(\text{LPR})} = (1\ X_i\ X_i^2\ \ldots\ X_i^d)$. The n fitted values may be expressed as $\hat{\mathbf{y}}_{\text{LPR}} = \mathbf{H}^{(\text{LPR})}\mathbf{y}$, where $\mathbf{H}^{(\text{LPR})} = [\mathbf{h}_i'^{(\text{LPR})}]$. Cleveland[34] and Hastie and Loader[35] present this development with further discussion, and Stone[36,37] shows optimal convergence rates for LPR in a certain minimax sense. In choosing the order d of the local polynomial, a first order fit, yielding local linear regression (LLR), is an adequate choice for the majority of cases, and is used here as the nonparametric fitting technique. Fan[38] gives asymptotic optimality properties and advantageous small sample properties via simulations for LLR.

C. Semiparametric Regression

With partial information about the underlying model, the idea now is to combine a parametric fit that takes advantage of this knowledge with a nonparametric fit that captures any structure in the data not explained by the parametric fit.

1. Model Robust Regression 1 (MRR1)

A simple and effective procedure that addresses the shortcomings of previous semiparametric methods is Model Robust Regression 1 (MRR1), developed by Einsporn and Birch.[39] MRR1 combines the fit of a parametric model with the fit of a nonparametric model, both to the raw data, in a convex combination via a mixing parameter λ. Letting $\hat{\mathbf{y}}_{\text{ols}}$ be the OLS fitted values and $\hat{\mathbf{y}}_{\text{LLR}}$ be the local linear fitted values, the MRR1 fitted values are obtained simply as $\hat{\mathbf{y}}_{\text{MRR1}} = \lambda\hat{\mathbf{y}}_{\text{LLR}} + (1-\lambda)\hat{\mathbf{y}}_{\text{ols}}$ where $\lambda \in [0,1]$. In terms of hat matrices,

$$\begin{aligned} \hat{\mathbf{y}}_{\text{MRR1}} &= \lambda\mathbf{H}^{(\text{LLR})}\mathbf{y} + (1-\lambda)\mathbf{H}^{(\text{ols})}\mathbf{y} \\ &= \left[\lambda\mathbf{H}^{(\text{LLR})} + (1-\lambda)\mathbf{H}^{(\text{ols})}\right]\mathbf{y} \\ &= \mathbf{H}^{(\text{MRR1})}\mathbf{y}. \end{aligned} \tag{6}$$

The purpose of λ is to combine the parametric and nonparametric fits in the most efficient proportions, and λ increases from 0 to 1 as the amount of misspecification of the user's parametric model increases. The choice of λ involves considerations similar to those for bandwidth, and PRESS** may be used as the data-driven method to estimate λ. An asymptotically optimal estimator (in terms of minimizing the distance between the unknown regression function and the MRR1 fit) that behaves similar to PRESS** is presented in Section III, where convergence rates are also given for both selection criteria.

MRR1 combines two *separate* fits to the data at each X_i, selecting a value between them for the final estimate of y_i. If locations exist in the data where the two fits are either both too high or both too low, then MRR1 has no way to correct for these insufficient fits. It is important to note that MRR1 is a procedural definition. Specific use of ols for the parametric portion and LLR for the nonparametric portion of the overall fit is *not* required. Robust, weighted, etc. regression techniques can be substituted for ols and LLR as long as the substitution is of like kind (parametric or nonparametric).

2. Model Robust Regression 2 (MRR2)

The other model-robust procedure, Model Robust Regression 2 (MRR2), is motivated by the need to improve upon the shortcomings of MRR1 and other existing semiparametric methods. Similar to MRR1, two separate fits are combined to obtain the final fit. The parametric portion of MRR2 is a parametric fit, say $\hat{\mathbf{y}}_{\text{ols}}$, to the *raw data* (as in MRR1). However, the nonparametric portion of MRR2, instead of coming from a fit to the raw data, comes from a fit to the *residuals* (r) from the parametric fit. The nonparametric fit (using LLR) may be expressed as $\hat{\mathbf{r}} = \mathbf{H}_2^{(\text{LLR})}\mathbf{r}$, where $\mathbf{H}_2^{(\text{LLR})}$ is the local linear hat matrix for fitting the residuals. The final MRR2 fit is obtained by adding a *portion* of the residual fit back to the original parametric fit: $\hat{\mathbf{y}}_{\text{MRR2}} = \hat{\mathbf{y}}_{\text{ols}} + \lambda\hat{\mathbf{r}}$. As in MRR1, λ increases from 0 to 1 as model misspecification increases, and λ and the bandwidth are chosen separately. Also, λ may be chosen by PRESS** or the asymptotically optimal estimator for MRR2 discussed in Section III. The basic motivation for λ is as follows. If the parametric fit is adequate, then adding back the nonparametric fit to the residuals would increase the variability of the overall fit. A $\lambda \approx 0$ would control for this. If the parametric fit is inadequate, then the nonparametric fit should be used to improve it. The amount of misspecification of the parametric model, and thus the amount of correction needed from the residual fit, is reflected in the size of λ. MRR2 uses λ to control the variance of the final fit by allowing use of only a portion of the nonparametric fit. In terms of hat matrices,

$$\begin{aligned} \hat{\mathbf{y}}_{\text{MRR2}} &= \mathbf{H}^{(\text{ols})}\mathbf{y} + \lambda\mathbf{H}_2^{(\text{LLR})}\mathbf{r} \\ &= \left[\mathbf{H}^{(\text{ols})} + \lambda\mathbf{H}_2^{(\text{LLR})}(\mathbf{I} - \mathbf{H}^{(\text{ols})})\right]\mathbf{y} \\ &= \mathbf{H}^{(\text{MRR2})}\mathbf{y}. \end{aligned} \tag{7}$$

It is the authors' contention that MRR2 is the best overall procedure for fitting data in situations of small to moderate model misspecification. MRR2 is simpler and more intuitive than other semiparametric methods and eliminates the cases in MRR1 where both of the component fits are inaccurate in the same direction. This is accomplished by the *residual* fit, which provides flexibility to correct for these bias problems in the initial parametric fit. It is also conjectured that applying the nonparametric fit to the residuals instead of the raw data will provide fits that are less variable, because the structure left in the residuals should be less complex than that of the raw data. As with MRR1, MRR2 is a procedural definition and does not explicitly

require ols and LLR as long as a parametric method is substituted for ols and a nonparametric method for LLR.

III. Asymptotically Optimal Mixing Parameter

In this section, derivations are given for asymptotically optimal estimators of λ, the mixing parameter, that minimize the distance between the unknown regression function and the MRR1 or MRR2 fit. The term "asymptotic" means that the number of observations increases without bound, and the manifestation of these observations in the X space increases uniformly in the predictor space. This creates a fixed effects structure, as $n \to \infty$, over a uniform design. In other words, if there are an infinite number of equally spaced points the estimate would asymptotically approach the actual parameter.

For any two functions of $\mathbf{x}_i'$, $\mathbf{h}_1$ and $\mathbf{h}_2$, the inner product is defined as

$$\langle \mathbf{h}_1, \mathbf{h}_2 \rangle = \frac{1}{n} \sum_{i=1}^{n} \mathbf{h}_1(\mathbf{x}_i)\mathbf{h}_2(\mathbf{x}_i), \tag{8}$$

and the norm as

$$\|\mathbf{h}_1\|^2 = \langle \mathbf{h}_1, \mathbf{h}_1 \rangle, \tag{9}$$

with

$$\|\mathbf{h}_1\| = \sqrt{\langle \mathbf{h}_1, \mathbf{h}_1 \rangle}. \tag{10}$$

For the MRR1 fit, a mixing parameter is sought to minimize

$$\left\| \hat{\mathbf{y}}^{(\mathrm{MRR1})} - \mathbf{g} \right\| = \left\| \lambda \hat{\mathbf{y}}^{\mathrm{NP}} + (1-\lambda)\hat{\mathbf{y}}^{\mathrm{P}} - \mathbf{g} \right\| \tag{11}$$

where g is the true function. As in Burman and Chaudhuri,[7] it is easy to demonstrate by simple calculus that Equation 11 attains a minimum at

$$\lambda^* = \frac{\left\langle \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}}, \mathbf{g} - \hat{\mathbf{y}}^{\mathrm{P}} \right\rangle}{\left\| \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}} \right\|^2} = \frac{\left\langle \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}}, \mathbf{g} \right\rangle - \left\langle \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}}, \hat{\mathbf{y}}^{\mathrm{P}} \right\rangle}{\left\| \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}} \right\|^2} \tag{12}$$

Since g is unknown, λ^* must be estimated; thus Mays, Birch, and Starnes[3] propose the following asymptotically optimal data driven mixing parameter:

$$\hat{\lambda}^* = \frac{\left\langle \hat{\mathbf{y}}^{\mathrm{NP}}_{-i} - \hat{\mathbf{y}}^{\mathrm{P}}_{-i}, \mathbf{y} - \hat{\mathbf{y}}^{\mathrm{P}} \right\rangle}{\left\| \hat{\mathbf{y}}^{\mathrm{NP}} - \hat{\mathbf{y}}^{\mathrm{P}} \right\|^2} \tag{13}$$

where $\hat{y}^{\mathrm{P}}_{-i}$ and $\hat{y}^{\mathrm{NP}}_{-i}$ represent the parametric and nonparametric estimates, respectively, obtained by leaving out the ith observation when fitting at x_i. These cross-validation estimates are used because the nonparametric estimate $\hat{y}^{\mathrm{NP}}_i$ at x_i may overfit the data and it is desirable to protect against putting too much weight on $\hat{y}^{\mathrm{NP}}_i$ in the mixing process.

If the parametric model is incorrect, MRR1 can achieve a consistent estimate at the asymptotic convergence rate of the nonparametric estimate. Otherwise, if the model is correct, then MRR1 achieves consistency at the parametric rate, faster than a purely nonparametric estimate. Details, including the derivation, assumptions, and proofs for the asymptotically optimal MRR1 mixing parameter are found in Burman and Chaudhuri[7] and Starnes.[40]

For obtaining the optimal estimator for MRR2, consider the parametric fit to the data, $\hat{\mathbf{y}}^{\mathrm{P}}$, and the nonparametric fit to the residuals, $\hat{\mathbf{r}}$. Similar to the MRR1 case, it can be shown that

$$\left\| \hat{\mathbf{y}}^{(\mathrm{MRR2})} - \mathbf{g} \right\| = \left\| \lambda \hat{\mathbf{r}} + \hat{\mathbf{y}}^{\mathrm{P}} - \mathbf{g} \right\| \tag{14}$$

attains a minimum at

$$\lambda^* = \frac{\left\langle \hat{\mathbf{r}}, \mathbf{g} - \hat{\mathbf{y}}^{\mathrm{P}} \right\rangle}{\left\| \hat{\mathbf{r}} \right\|^2}. \tag{15}$$

The suggested asymptotically optimal data driven mixing parameter is given by

$$\hat{\lambda}^* = \frac{\left\langle \hat{\mathbf{r}}, \mathbf{y} - \hat{\mathbf{y}}^{\mathrm{P}} \right\rangle}{\|\hat{\mathbf{r}}\|^2} \tag{16}$$

The MRR2 model possesses the same asymptotic properties as those of the MRR1 model; thus converging to the underlying model whether or not the specified model is correct. Details, including the derivation, assumptions, and proofs for the asymptotically optimal MRR2 mixing parameter are found in Starnes[40] and Mays, Birch, and Starnes.[3]

IV. Statistical Inference for Goodness of Fit (Confidence and Prediction Intervals)

Having reviewed the methods, it becomes necessary to use some metric to compare the methods. There are several metrics that can be used to infer model performance: the confidence interval, the prediction interval, and the mean square error. This section will discuss the corresponding confidence intervals (C.I.) and prediction intervals (P.I.) for the various fitting techniques and the mean square error calculation.

As developed by Mays,[5,41] the form of the confidence interval for the semiparametric methods will closely resemble that of OLS. The OLS $(1-\alpha)100\%$ C.I. for the true mean μ_{y_0} at the location x_0 for a $(p-1)$th order polynomial is given by

$$\hat{\mathbf{y}}_0^{(\mathrm{ols})} \pm t_{n-p,\frac{\alpha}{2}} \hat{\sigma} \sqrt{\mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}, \tag{17}$$

where $\hat{\mathbf{y}}_0^{(\mathrm{ols})}$ is the estimated value at the individual point $\mathbf{x}_0' = (1\, x_0\, x_0^2 \cdots x_0^{p-1})$, $t_{n-p,\frac{\alpha}{2}}$ is the $\left(\frac{1-\alpha}{2}\right)^{\mathrm{th}}$ percentile of the t-distribution with $n-p$ degrees of freedom $(d.f.)$, and $\hat{\sigma}$ is an estimate of error standard deviation. The usual estimate of σ is

$$\hat{\sigma}_{\mathrm{ols}} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{(\mathrm{ols})}\right)^2}{n-p}}. \tag{18}$$

Similarly, the OLS $(1-\alpha)100\%$ P.I. for individual measures $\hat{\mathbf{y}}_0$ at the location $\mathbf{x}_0$ is given by

$$\hat{\mathbf{y}}_0^{(\mathrm{ols})} \pm t_{n-p,\frac{\alpha}{2}} \hat{\sigma} \sqrt{1 + \mathbf{x}_0'(\mathbf{X}'\mathbf{X})^{-1}\mathbf{x}_0}. \tag{19}$$

The general form of the $(1-\alpha)100\%$ C.I. and P.I. can be expressed as

$$\hat{\mathbf{y}}_0^{(\bullet)} \pm t_{n-tr\left(\mathbf{H}^{(\bullet)}\right),\frac{\alpha}{2}} \hat{\sigma}_{(\bullet)} \sqrt{\mathbf{h}_0'^{(\bullet)}\mathbf{h}_0^{(\bullet)}}, \tag{20}$$

$$\hat{\mathbf{y}}_0^{(\bullet)} \pm t_{n-tr\left(\mathbf{H}^{(\bullet)}\right),\frac{\alpha}{2}} \hat{\sigma}_{(\bullet)} \sqrt{1 + \mathbf{h}_0'^{(\bullet)}\mathbf{h}_0^{(\bullet)}}, \tag{21}$$

respectively, where

$$\hat{\sigma}_{(\bullet)} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{(\bullet)}\right)^2}{n - tr\left(\mathbf{H}^{(\bullet)}\right)}}, \tag{22}$$

and "$\bullet$" represents the corresponding regression technique, whether it be, for example, OLS, LLR, LPR, MRR1, or MRR2. The mean square error or MSE is defined as $\hat{\sigma}^2$.

V. Calibration Experiment

A. Instrumentation

Voltages were obtained from calibrations conducted on Electronically Scanned Pressure (ESP) transducers manufactured by Pressure Systems Incorporated (PSI) using a system developed for low pressure calibrations as described by Green, Everhart, and Rhode.[42] The PSI System 8400 is a differential pressure measurement

system that uses multiplexed multi-port modules. Eight ports of a 32-port, 15 psid ESP module were randomly selected to be representative of data obtained from this measurement range.

The ESP transducers were selected to highlight the MRR2 regression model for instrument calibration because the standard ESP calibration has a documented deficiency and a proposed corrective model (Everhart[1]). In essence, the standard calibration model misspecifies the behavior of the transducer for low absolute pressures. Therefore, it is anticipated that by using the MRR2 model with the standard calibration model as the parametric form, it will asymptotically approach the corrected calibration model, provided that enough data are acquired.

B. Data Acquisition

The calibration system described by Green, Everhart, and Rhode[42] consists of a proportional integral derivative (PID) pressure control system that controls pressures on the sample side of the transducer while the reference side of the transducer is held at a hard vacuum. Calibration set pressures are measured using two Ruska transducers with 1 and 15 psid ranges accurate to 0.003% full scale and 0.005% full scale, respectively.

The calibration system controls the pressure independently from the pressure measurement system which allows pressures to be held near or within the resolution of the Ruska calibration standards. Temperature is controlled using software PID algorithms programmed into the LabVIEW programming environment. Temperature is held to within $\pm 0.1°$F. The transducer environmental pressure is controlled within an environmental control chamber which can maintain pressures from a vacuum near 0.005 psia to pressures as large as 150 psia.

Ten controlled calibrations were performed on the ESP modules. These calibration data were taken while the low pressure calibration system was demonstrated to be in statistical control. Stability and control of this calibration system is discussed by Green, Walker, and Everhart.[43] The ten calibration runs consisted of 9 to 32 points acquired from 0 to 5 psi. The data acquisition schedule details are given in Table 1. Figures 1(a) and (b) show the ten calibrations for a representative transducer on the 15 psid ESP module.

C. Establishing the Noise Level

With each replicate calibration, a zero pressure voltage V_0 was acquired. These voltage measures, which can be seen in Figure 1(b) of the 15 psid ESP data around $P = 10^{-5}$, have no correlation with the pressure settings at this level. This lack of correlation is indicative of reaching the noise floor of the instrument (*i.e.,* the instrument is no longer sensitive to changes in pressure at this level); thus the zero pressure voltage dispersion measure can be used to quantify the instrument's noise floor. Replication of measures at pressure can also yield this information; however, care must be taken in the data acquisition such that there is no error due to not repeating the set point closely enough. This can cause a local correlation effect and bias the noise measure.

The ranges from each measured port on a given module are compared for consistency. This is done to estimate an average noise floor for a given module range. Figure 2(a) shows a statistical process control (SPC) range chart for a 15 psid ESP module, and Figure 2(b) presents the same results in terms of percent of full scale of the instrument (5 volts). Calculations for Figure 2 were based on basic SPC theory outlined by Wheeler and Chambers.[44] Constants were taken from Tables A.1 and A.2 on pages 393-394. The centerline of each figure is the average range, $\bar{R}$, across each of the ports for a module. The upper and lower range limits (URL/LRL) are given by $D_4\bar{R}$ and $D_3\bar{R}$, where $D_4 = 1.777$ and $D_3 = 0.223$ for $n = 10$ (the number of calibrations).

Based on the results from Figure 2, the ports have essentially the same variation and the average variation across the ports is indicative of the behavior of transducers for this module range. The standard deviation can be estimated by $\hat{\sigma} = \bar{R}/d_2$, where $d_2 = 3.078$ for $n = 10$. The noise of the instrument will be represented by an error band of $\pm 3\hat{\sigma}$ for the absolute scale. A coverage factor of three is chosen to yield a conservative/confirmatory estimate of the intrinsic variation in the instrument. According to work presented by Wheeler,[45] a coverage factor of 3 standard deviations from the mean encompasses at least 98% of the variation of the data for all practical probability distributions, including those with high levels of skewness—far better than the 88.9% from Chebychev's inequality. The instrument noise can then be expressed as

$$\pm 3\hat{\sigma} = \pm 3\frac{\bar{R}}{d_2} = \pm\frac{3\bar{R}}{3.078} = \pm 0.975\bar{R} \approx \pm\bar{R}. \tag{23}$$

For the 15 psid ESP module, the noise band is ±1.2mV or 0.023% full scale.

VI. Calibration Model Analysis

For the purpose of discussion the 15 psid ESP module will be used to highlight the properties of MRR2. Three parametric models will be used to draw contrast to the present method: a fourth order polynomial, an "invented" model tailored for the instrument, and a simple linear fit. This section will discuss these three models and how they explain the data set. A fourth order polynomial model was chosen since that is what is used in practice. This is then compared with a model tailored for the instrument and with a simple linear fit based on the overall trend of the data seen in Figures 1(a) and (b).

The prediction interval (see Section IV) is an indication of how well the calibration model (*i.e.,* curve-fit) matches the data for a given run. As discussed in Section V.C, the instrument noise band is an indication of the instrument capability (*i.e.,* across-run reproducibility). Hence, if the prediction interval for a given fit is larger than the noise band, it is an indication of detectable lack-of-fit.

A. Typical ESP Calibration Model

ESP transducers are usually calibrated using an intrinsic calibration method that uses a fit of a fourth order polynomial to 5 data points. In practice, this has been modified, on occasion with customized software, to include more data points. The analytical form for the standard fit is as follows:

$$P = C_0 + C_1 V + C_2 V^2 + C_3 V^3 + C_4 V^4. \tag{24}$$

Here, P is the pressure, and V is the transducer output voltage. To allow for a proper comparison of the calibration models, it is necessary to modify Equation 24 to the following form:

$$P = B_1 Z + B_2 Z^2 + B_3 Z^3 + B_4 Z^4, \tag{25}$$

where, $Z = V - V_0$, and V_0 is the voltage at zero pressure.

Figure 3 shows a typical result of regressing the data from the 15 psid ESP module using Equation 25. A brief discussion of the analysis presentation is provided here to orient the reader to the plot style that will be used for the rest of the paper. Figure 3(a), the full-scale curve fit shown on a log-log scale to highlight the low pressure behavior of the instrument, has five elements: the data used for the regression process, the curve fit result from the regression process, the confidence interval on the mean regression curve using a significance level of $\alpha = 0.05$, the prediction interval of an individual measurement using a significance level of $\alpha = 0.05$, and the cross-check data consisting of data from the 9 other calibration runs. The cross-check data will also be used to generate predicted residuals by subtracting the actual Ruska standard measure of pressure from the predicted pressure of the curve fit. This process of predicting the residual is accomplished by taking an ordered pair of measured data, (V_i, P_i), and transforming it to $(V_i, \dot{r}_i)$ by using the following equation:

$$\dot{r}_i = P_i - \hat{P}_i(V_i), \tag{26}$$

where $\dot{r}_i$ is the individual predicted residual, and $\hat{P}_i$ is the predicted response of the instrument given a particular regression model. Figure 3(b) shows the direct residuals on a log-linear scale from the curve fit as well as the predicted residuals. Confidence and prediction intervals are included in the residual plots for comparison. These intervals are taken from Figure 3(a) and applied to the axis since the estimate of the mean response has been removed. The estimate of the noise band of the instrument calculated in Section V.C is also denoted on the residual plots. Figure 3(c) shows the absolute residuals as a percent of full scale. This figure contains the same information as Figure 3(b).

Table 2 presents some of the key statistics of the calibration that will be used for discussion. This table is organized by the base parametric regression model and subcategorized by the local or nonparametric regression model. Procedures are denoted as follows: ordinary least squares (OLS), local polynomial regression (LPR), and model robust regression 2 (MRR2). Statistics included in the table are: the degrees of freedom, $m.d.f.$, used in the model calculation (the trace of the hat matrix, see Section II); the mean square error, MSE or $\hat{\sigma}^2$, statistic (calculated using Equation 22), with $\nu = n - m.d.f.$ effective degrees of freedom; the cross-check mean square error, CCMSE, statistic (calculated using Equation 22, with $\nu = 132$ for Table 2 and $\nu = 155$ for Table 3.); and the estimated asymptotically optimal mixing parameter, $\hat{\lambda}^*$. Also included

for convenience are the figure references for the corresponding regression analysis. The fourth order regression model uses 5 degree of freedom, one for each of the model coefficients in Equation 25 and one for the measurement of V_0. The coverage factor for the model dispersion is approximately 2, as is the case with all the other models in Table 2. This is indicative of having reached a point where the model degrees of freedom are not significantly impacting the predictions of the confidence and prediction intervals; thus using a 32 point calibration allows the models to be compared on an equal footing.

From Figure 3(a) it can easily be seen that there is a deviation of the data from the curve fit. This is confirmed in the residual plots, Figures 3(b) and (c). These figures show a distinct structure in the residuals that was not captured by the 4th order polynomial. Also, Figure 3(c) shows that the prediction interval for the 4th order curve fit is well above the system noise level of 0.023% full scale, indicating that error incurred due to model misspecification is significant.

B. High/Low Pressure Calibration Model

While searching for a high accuracy calibration, Everhart[1] proposed a calibration model form that better fit the calibration data. This formulation is given as:

$$P = A_1 Z^{\frac{1}{3}} + A_2 Z^{\frac{1}{2}} + A_3 Z + A_4 Z^2. \tag{27}$$

Once again the coefficients are determined using a least squares fit of the entire data set. The improvements presented by Everhart[1] show significant reduction of the curve fit residuals.

As with the 4th order polynomial, Figure 4 shows the fit and residual analysis consistent with the Everhart model. In Figure 4(a) a much different picture of the curve fit is cast than with the 4th order polynomial model, even though Table 2 shows that both models use the same number of degrees of freedom. The confidence band is much wider than the 4th order fit but the curve actually tracks with the data. On the other hand, the prediction interval is much smaller than the 4th order model. The residual plots in Figures 4(b) and (c) show little if any structure left in the residuals. These plots also show that the prediction interval is below that of the instrument noise floor; thus indicating that the Everhart calibration model will tend to vary from calibration to calibration but only within the noise of the instrument. The Everhart model will result in a much more accurate calibration than that of the 4th order polynomial.

It should be noted that the work done by Everhart to obtain this parametric form was quite extensive. Thus, one of the major points of this article is that the current method will demonstrate similar performance starting with a model containing obvious deficiencies (*e.g.,* a 4th order polynomial or a simple linear model).

C. Simple Linear Calibration Model

As previously stated, a simple linear model was chosen based on the overall trend of the data in Figure 1. For the purpose of comparison, the linear model form that is used is as follows:

$$P = D_1 Z. \tag{28}$$

Figure 5 shows the fit and residual analysis for the simple linear fit. The behavior, at least in the full scale sense, Figure 5(a), is similar to that of the 4th order fit, with the exception that the prediction interval is larger. The residual scale shows a much different story. A very distinct pattern is seen in Figure 5(b). It is also interesting to note that the prediction interval is almost twice the size of that corresponding to the 4th order fit. This is an important fact since Table 2 shows that this model only uses 2 degrees of freedom (one for the model coefficient and one for the measurement of V_0). Thus the simple linear model shows an even greater deficiency than that of the 4th order polynomial.

It has been demonstrated that the simple linear model has a much greater deficiency than either of the other parametric models presented. This will prove to be a much more severe test of the MRR2 procedure than that of the 4th order polynomial.

VII. MRR2 Calibration Analysis

This section will discuss the MRR2 fitting procedure using each of the parametric models discussed in Section VI. The nonparametric fit to the residuals from each of the parametric models is analyzed first. This is followed by an analysis of the mixed model or final form of the MRR2 procedure.

A. Local Fitting of the Residuals

Following the methodology outlined in Section II.B, three increasing order local polynomial regressions were performed on residuals from each of the three parametric fits. Using the definition of $\mathbf{x}_i^{'(LPR)}$ given below Equation 5, local linear, quadratic, and cubic models are constructed by setting $d = 1$, 2, and 3, respectively. Bandwidth choice was automated by minimization of the penalized prediction sum of squares, $PRESS^{**}$, defined in Equation 4.

Figure 6 shows the fit and residual analysis for each of the three orders of the local polynomial modeling for the residuals of the 4th order fit. Figures 7 and 8 show the corresponding analyses for the Everhart and simple linear parametric model residuals, respectively. In each of these figures, panels (a), (d), and (g) show the fit of the residuals from each respective parametric model. These panels show that no real improvement is gained by increasing the local polynomial order. An important thing to note about the curve fit panels is that the local polynomial modeling captured the structure of the residuals. Careful examination of the residual panels will yield a slight improvement in terms of the prediction interval for the local quadratic fit and slight decrease in performance for the local cubic fit. According to Fan and Gijbels,[17] the order of the local polynomial model should be odd; however, in the cases investigated by the authors, often local quadratic fits have been seen to yield a slightly better fit. Further discussion of the problem concerning the order of the local polynomial is given by Ruppert *et al.*[46] For cases involving larger gradients it may be beneficial to use a higher order local model. Table 2 shows that the number of degrees of freedom for each local modeling order is the same regardless of the base parametric model: local linear, 8.5; local quadratic, 11.5; and local cubic, 12.2. With the exception of investigating the results of the MRR2 analysis, only local linear modeling will be used for the remainder of the paper. This is done because the additional cost in terms of degrees of freedom is not worth the benefit seen in residual/prediction interval reduction for this problem. Also note that in each case the prediction intervals are well below the system noise level. This indicates that the major contribution of the model misspecification to overall variation can be removed. The next step is to combine these results with the parametric model results in such a way as to describe the behavior of the system, at least to its noise level.

B. Mixing the Models

The methodology used in the MRR2 procedure is discussed in Section II.C.2. However, the key to making MRR2 work is in the use of the asymptotically optimal mixing parameter, λ^*. This parameter was derived using a geometric argument to add back some fraction of the structure found in the residuals to the parametric fit. The asymptotically optimal mixing parameter for MRR2 is discussed in Section III. For this paper, the estimating equation, Equation 16, for $\hat{\lambda}^*$ will be used. The estimated mixing parameters for each of the base parametric models are presented in Table 2.

Figure 9 shows the regression fit and residual analysis for the MRR2 procedure using the 4th order polynomial (Equation 25) as the base parametric model and local linear (a)-(c), local quadratic (d)-(f), and local cubic (g)-(i) models. The Everhart (Equation 27) and simple linear (Equation 28) base parametric models are shown in Figures 10 and 11, respectively.

The discussion for the MRR2 figures closely follows that of the local polynomial fit of the residuals in the previous section (Section VII.A). Again, no real improvement is gained by increasing the local polynomial order. This can be seen by comparing the prediction intervals for each of the local polynomial regression orders. Another way of comparing this is by using the estimated mixing parameters given in Table 2. These statistics show very little increase with increasing local polynomial order. Recall that $\hat{\lambda}^*$ is a scaled geometric measure of the structure contained in the residuals. Thus, little, if any, additional structure is being recovered by increasing local polynomial order with the possible exception of responses with large changes in gradient.

Note from Table 2, that the number of degrees of freedom used for the MRR2 procedure is between that of the OLS and LPR procedures. This is a direct benefit of the mixing procedure discussed in Section II.C.2. Further discussion of model degrees of freedom is provided in Section VIII.A.

Each of the final MRR2 curve fits track with the data and the resulting residual plots have little, if any, structure. Also note that every one of the fits involving this procedure have prediction intervals below the system noise level. This is one of the highlights of the current method: using enough data, the solution *will* converge to the system noise level.

VIII. Discussion of MRR2 Properties

This section will discuss some of the more important properties of the MRR2 procedures by example. A discussion of the number of degrees of freedom necessary to perform this regression is included to justify the cost of the present method. Another property of MRR2 is that it will default to the parametric model when no structure in the residuals can be detected.

A. Degrees of Freedom

A discussion of the degrees of freedom used by the current procedure is important because the number of degrees of freedom is directly tied to the cost of using the model. For the purpose of this discussion, the same three parametric models will be used to regress the 15 psid ESP calibration, this time with only a 9 point calibration run. The same analysis path used in the 32 point calibration will be employed here. As with the 32 point calibration, regression fit statistics for comparison are included in Table 3. Figures 12, 13, and 14 show the parametric fit and residual analysis for the 4th order polynomial, Everhart, and simple linear models, respectively. Comparing the (a) panel of the three figures, similar results are seen when compared to Figures 3, 4, and 5. The prediction intervals are wide for both the 4th order and simple linear models with respect to that of the Everhart model. Since the models are parametric, the required number of degrees of freedom *always* remains the same. For a given model, the prediction interval is controlled by two quantities: the standard deviation of the residuals and the number of remaining degrees of freedom after subtracting those used by the model (see Equation 19 or 21). The consequence is that if a large number of degrees of freedom are remaining after the regression, the prediction (and confidence) intervals will be smaller. On the other hand, if the standard deviation of the residuals is small due to the model's explanation of the data this will also cause the prediction interval to be smaller. This small residual standard deviation explains the fact that the Everhart model prediction limits are much smaller than those of the two polynomial based models (*i.e.,* it just explains the data better). Even with a 9 point calibration the Everhart model prediction interval is within the noise for this instrument. For the two polynomial based models, only small changes are seen between the 32 point calibration (Figures 3 and 5) and the 9 point calibration(Figures 12 and 14). The differences seen here are primarily due to number of remaining degrees of freedom. As before, the residuals for the polynomial based models show a definite structure in comparison to the residuals from the Everhart model. The cross-check mean square error (CCMSE) statistics presented in Tables 2 and 3 also indicate the same trends for the parametric models.

Figures 15, 16, and 17 show the nonparametric residual fit and the resulting residual analysis for the 4th order polynomial, Everhart, and simple linear models, respectively. In this particular case, the residuals from the 4th order fit are more complex than what is represented by the 9 calibration points. This can be seen in Figures 12(b) and 15(a) with the cross-check data. Having the residual structure under-represented by the data causes the nonparametric model to fail to adequately recover this structure. As can be seen from the two residual plots (b) and (c) in Figure 15, the prediction interval is much greater than the instrument noise level. Note that the spikes seen in the prediction (and confidence) intervals are caused by the gaps in the data being too large for the calculated global bandwidth to give a good estimate in these regions. This could be addressed by using a locally defined bandwidth (see references in Section II.B), but is beyond the scope of this article. Little structure is seen in the residuals from the Everhart model in Figure 16. This is primarily due the performance of the parametric model that was seen in Figure 13. The simple linear fit does have a complex structure remaining in the residuals, but it is well represented by the 9 calibration points. This can be verified in Figure 17(a) using the cross-check data. Here the prediction interval for the remaining residuals from the local fit are at the instrument noise level. As before, the number of degrees of freedom for the local polynomial regression are the same no matter what the base parametric model is. Note that in Tables 2 and 3, the degrees of freedom required are not identical for the 9 point and 32 point calibrations. This is due to a combination of the bandwidth and the number of points that are being fit (*i.e.,* the number of terms required in a local polynomial regression to fit the data). The more data points there are and the more narrow the bandwidth, the more degrees of freedom will be required.

The MRR2 augmentation for each of the three parametric models is shown in Figures 18, 19, and 20. Results of this analysis follow directly from the preceding discussion of the local fit. Again, the 9 point calibration failed to adequately represent the 4th order polynomial model misspecification completely; however, some improvement is made by the MRR2 procedure which can be seen from comparing the parametric fit analysis in Figure 12 to the MRR2 fit in Figure 18. The CCMSE for the MRR2 augmentation of the 4th

order fit also indicates some improvement; however, the overall improvement is not as great at that seen with the 32 point calibration or either of the simple linear fits as indicated in Tables 2 and 3. Both the augmented Everhart and simple linear models have prediction intervals at or below the instrument noise level.

From this analysis, it is apparent that a trade-off exists for using MRR2. If an adequate parametric model is established, fewer data points will be necessary to define the parameters for that model. However, if misspecification is present in the model, more data must be acquired in the areas of misspecification to ensure that the structure of the parametric model residuals is adequately represented. In effect, the MRR2 procedure takes the specified parametric model and augments it with a geometrically scaled portion of locally fit data model residuals that contain a nonrandom structure. In addition, the MRR2 procedure can only augment or correct the model to the level of the noise in data.

B. Default to the Parametric Form

Another property of the current method is its tendency to default to the parametric form if the parametric residuals do not have a detectable structure. An example of this property is seen with the Everhart model. This model has a parametric description that fits the data to within the noise level consistently. Evidence for this can be seen by comparing Figures 4 and 10 for the 32 point calibration and Figures 13 and 19 for the 9 point calibration. In both cases the local linear nonparametric augmentation of the Everhart model yielded only marginal improvement. Evidence for this is seen in Tables 2 and 3. The mixing parameters for the Everhart models for both the 32 and 9 point calibrations are significantly lower than those of the polynomial models. Also, the number of degrees of freedom used for the MRR2 procedure using the Everhart base model are approximately the same as the Everhart model itself. This is not the case for either of the two polynomial models. The true advantage here is that the MRR2 procedure returned the parametric form at no additional cost. Since there was no real structure in the residuals from the Everhart model, essentially none of the local modeling was included in the final result.

IX. Final Remarks

This paper has demonstrated the robustness to regression/calibration model selection of the MRR2 method. Using a geometrically scaled portion of the residuals from the parametric model, essentially a data model, the initially chosen model is augmented and will converge to within the noise level of the instrument as the data points both increase in number and better characterize the parametric residual structure. This paper has also shown that at least for a mild regression in terms of gradients and misspecification–the number and coverage demands of the MRR2 procedure are not excessive. One other beneficial property is that the MRR2 procedure will not augment parametric models that truly fit the data. In this respect, the MRR2 procedure can also be used as a tool to detect model misspecification and drive toward a more correct parametric form if that is what is desired. This is advantageous only if a parametric form of the calibration model is required or desired.

Since instrument calibration is typically a low noise environment in comparison with the random calibration problem, the current method is vastly superior to a basic parametric model calibration. This is due to instrument behavior being defined by the data model itself. If the output of an instrument can be accurately referenced to a standard and enough data taken at a fine enough spacing, this would essentially allow a lookup table calibration of the instrument. Lack of resources often make this approach unfeasible, and either parametric or nonparametric techniques are employed to reduce the amount of data. By using the current method, the advantages of both the parametric and nonparametric methods are gained. Any basic understanding of the instrument behavior can be incorporated through the parametric form, and any misspecification or deficiency of the parametric form can be augmented by mixing in a portion of a nonparametric modeling technique.

References

[1]Everhart, J., "Calibration Improvements to Electronically-Scanned Pressure Systems and Preliminary Statistical Assessment," AIAA Paper 96-2217, 1996.

[2]Cheng, C.-L. and VanNess, J., "Robust Calibration," *Technometrics*, Vol. 39, No. 4, November 1997, pp. 401–411.

[3]Mays, J., Birch, J., and Starnes, B., "Model Robust Regression–Combining Parametric, Nonparametric, and Semiparametric Methods," *Journal of Nonparametric Statistics*, Vol. 13, 2001, pp. 245–277.

[4]Myers, R., *Classical and Modern Regression with Applications*, PWS-KENT, Boston, 2nd ed., 1990.

[5]Mays, J., *Model Robust Regression–Combining Parametric, Non-parametric, and Semiparametric Methods*, Ph.D. thesis, Virginia Polytechnic Institute and State University, September 1995.

[6]Speckman, P., "Kernel Smoothing in Partial Linear Models," *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 50, No. 3, 1988, pp. 413–436.

[7]Burman, P. and Chaudhuri, P., "A Hybrid Approach to Parametric and Nonparametric Regression," Technical Report No. 243, Division of Statistics, University of California-Davis, 1992.

[8]Rahman, M., Gokhale, D., and Ullah, A., "A Note on Combining Parametric and Nonparametric Regression," *Communications in Statistics–Simulation and Computation*, Vol. 26, No. 2, 1997, pp. 519–529.

[9]Hjort, N. and Glad, I., "Nonparametric Density Estimation With a Parametric Start," *The Annals of Statistics*, Vol. 23, No. 3, June 1995, pp. 882–904.

[10]Eubank, R. and Spiegelman, C., "Testing the Goodness of Fit of a Linear Model Via Nonparametric Regression Techniques," *Journal of the American Statistical Association*, Vol. 85, No. 410, June 1990, pp. 387–392.

[11]Newey, W., "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, Vol. 79, No. 1, July 1997, pp. 147–168.

[12]Wooldridge, J., "A Test for Functional Form Against Nonparametric Alternatives," *Econometric Theory*, Vol. 8, No. 4, December 1992, pp. 452–475.

[13]Robinson, P., "Root-N-Consistent Semiparametric Regression," *Econometrica*, Vol. 56, No. 4, July 1988, pp. 931–954.

[14]Ullah, A. and Vinod, H., "General Nonparametric Regression Estimation and Testing in Econometrics," *Handbook of Statistics*, edited by G. Maddala, C. Rao, and H. D. Vinod, Vol. 11, Elsevier Science Publishers, New York, 1993.

[15]Härdle, W., *Applied Nonparametric Regression*, Cambridge University Press, New York, 1990.

[16]Wand, M. and Jones, M., *Kernel Smoothing*, Chapman and Hall, London, 1995.

[17]Fan, J. and Gijbels, I., *Local Polynomial Modelling and Its Applications*, Chapman and Hall, London, 1996.

[18]Priestley, M. and Chao, M., "Non-Parametric Function Fitting," *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 34, No. 3, 1972, pp. 385–392.

[19]Gasser, T. and Müller, H., "Kernel estimation of Regression Functions," *Smoothing Techniques for Curve Estimation*, edited by Gasser and Rosenblatt, Springer-Verlag, Heidelberg, 1979.

[20]Chu, C.-K. and Marron, J., "Choosing a Kernel Regression Estimator (with discussion)," *Statistical Science*, Vol. 6, No. 4, November 1991, pp. 404–436.

[21]Mays, J. and Birch, J., "Smoothing Considerations in Nonparametric and Semiparametric Regression," Technical Report Number 98-2, Dept. of Statistics, Virginia Polytechnic Institute and State University, 1998.

[22]Mays, J. and Birch, J., "Smoothing for Small Samples with Model Misspecification: Nonparametric and Semiparametric Concerns," *Journal of Applied Statistics*, Vol. 29, 2002, pp. 1023–1045.

[23]Rice, J., "Bandwidth Choice for Nonparametric Regression," *Annals of Statistics*, Vol. 12, No. 4, December 1984, pp. 1215–1230.

[24]Chiu, S.-T., "Why Bandwidth Selectors Tend to Choose Smaller Bandwidths, and a Remedy," *Biometrika*, Vol. 77, No. 1, March 1990, pp. 222–226.

[25]Härdle, W. and Marron, J., "Optimal Bandwidth Selection in Nonparametric Regression Function Estimation," *Annals of Statistics*, Vol. 13, No. 4, December 1985, pp. 1465–1481.

[26]Faraway, J., "Bootstrap Selection of Bandwidth and Confidence Bands for Nonparametric Regression," *Journal of Statistical Computation and Simulation*, Vol. 37, No. 1-2, 1990, pp. 37–44.

[27]Ruppert, D., Sheather, S., and Wand, M., "An Effective Bandwidth Selector for Local Least Squares Regression," *Journal of the American Statistical Association*, Vol. 90, No. 432, 1995, pp. 1257–1270.

[28]Allen, D., "The Relationship Between Variable Selection and Data Augmentation and a Method for Prediction," *Technometrics*, Vol. 16, 1974, pp. 125–127.

[29]Härdle, W., Hall, P., and Marron, J., "How Far are Automatically Chosen Regression Smoothing Parameters from Their Optimum? (with discussion)," *Journal of the American Statistical Association*, Vol. 83, No. 401, March 1988, pp. 86–101.

[30]Altman, N., "An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression," *American Statistician*, Vol. 46, No. 3, August 1992, pp. 175–185.

[31]Müller, H. and Stadtmüller, U., "Variable Bandwidth Kernel Estimators of Regression Curves," *Annals of Statistics*, Vol. 15, No. 1, March 1987, pp. 182–201.

[32]Härdle, W. and Bowman, A., "Bootstrapping in Nonparametric Regression: Local Adaptive Smoothing and Confidence Bands," *Journal of the American Statistical Association*, Vol. 83, No. 401, March 1988, pp. 102–110.

[33]Park, B. and Marron, J., "Comparison of Data-Driven Bandwidth Selectors," *Journal of the American Statistical Association*, Vol. 85, No. 409, March 1990, pp. 66–72.

[34]Cleveland, W., "Robust Locally Weighted Regression and Smoothing Scatter Plots," *Journal of the American Statistical Association*, Vol. 74, No. 368, December 1979, pp. 829–836.

[35]Hastie, T. and Loader, C., "Local Regression: Automatic Kernel Carpentry (with discussion)," *Statistical Science*, Vol. 8, No. 2, May 1993, pp. 120–143.

[36]Stone, C., "Optimal Rates of Convergence for Nonparametric Estimators," *Annals of Statistics*, Vol. 8, No. 6, November 1980, pp. 1348–1360.

[37]Stone, C., "Optimal Global Rates of Convergence for Nonparametric Regression," *Annals of Statistics*, Vol. 10, No. 4, December 1982, pp. 1040–1053.

[38]Fan, J., "Design-adaptive Nonparametric Regression," *Journal of the American Statistical Association*, Vol. 87, No. 420, December 1992, pp. 998–1004.

[39]Einsporn, R. and Birch, J., "Model Robust Regression: Using Nonparametric Regression to Improve Parametric Regression Analyses," Technical Report Number 93-5, Dept. of Statistics, Virginia Polytechnic Institute and State University, 1993.

[40]Starnes, B., *Asymptotic Results for Model Robust Regression*, Ph.D. thesis, Virginia Polytechnic Institute and State University, 1999.

[41]Mays, J., "Capturing Nonconformity Points in Regression," *InterStat*, Vol. 3, August 2001, pp. 1–25, published online at http://interstat.stat.vt.edu/interstat/ARTICLES /2001/abstracts/G01003.html-ssi.

[42]Green, D., Everhart, J., and Rhode, M., "Development and Characterization of a Low-Pressure Calibration System for Hypersonic Wind Tunnels," AIAA 2004-0600, January 2004, Presented at *AIAA 42nd Aerospace Sciences Meeting and Exhibit, Reno, NV*.

[43]Green, D., Walker, E., and Everhart, J., "Statistical Characterization of Environmental Error Sources Affecting Electronically Scanned Pressure Transducers at Low Pressure," AIAA 2006-1386, January 2006, Presented at *44th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV*.

[44]Wheeler, D. and Chambers, D., *Understanding Statistical Process Control*, SPC Press, Knoxville, 2nd ed., 1992.

[45]Wheeler, D., *Normality and the Process Behavior Chart*, SPC Press, Knoxville, 2000.

[46]Ruppert, D., Wand, M., Holst, U., and Hössjer, O., "Local Polynomial Variance-Function Estimation," *Technometrics*, Vol. 39, No. 3, 1997, pp. 262–273.

Table 1. Data Acquisition Schedule

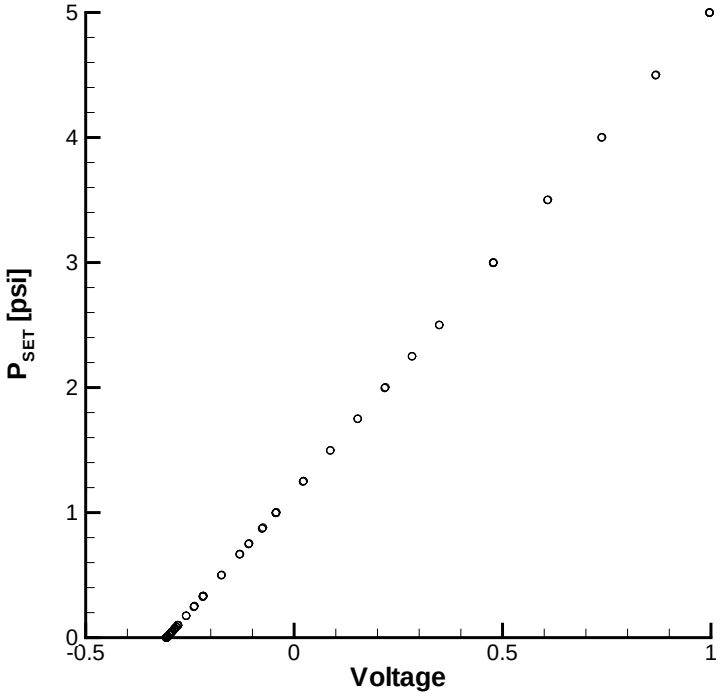
Replicates	5	1	3	1
Points Per Run	9	10	32	22
Pressure Schedule (psi)	0.0000	0.0000	0.0000	0.0000
	0.0010	0.0010	0.0010	0.0010
	0.0100	0.0100	0.0030	0.0030
	0.0330	0.0330	0.0060	0.0060
	0.1000	0.1000	0.0075	0.0075
	0.3300	0.2500	0.0100	0.0100
	1.0000	0.3300	0.0250	0.0250
	3.0000	1.0000	0.0330	0.0330
	5.0000	3.0000	0.0500	0.0500
		5.0000	0.0670	0.0670
			0.0750	0.0750
			0.0870	0.0870
			0.1000	0.1000
			0.1750	0.1750
			0.2500	0.2500
			0.3300	0.3300
			0.5000	0.5000
			0.6700	0.6700
			0.7500	0.7500
			0.8700	0.8700
			1.0000	1.0000
			1.2500	1.2500
			1.5000	
			1.7500	
			2.0000	
			2.2500	
			2.5000	
			3.0000	
			3.5000	
			4.0000	
			4.5000	
			5.0000	

Table 2. 32 Point Calibration Statistics [15 psid]

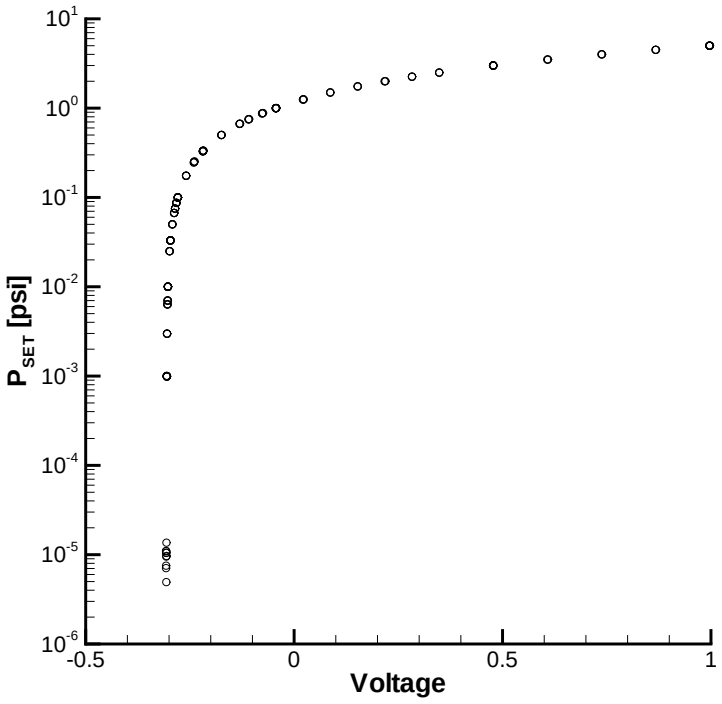
Model	$m.d.f.$	MSE($_{10^{-5}}$)	CCMSE($_{10^{-5}}$)	$\hat{\lambda}^*$	**Figure**
4th Order Polynomial Regression Model					
OLS	5.0	1.68	1.35		3
Local Linear					
LPR	7.1				6 (a)-(c)
MRR2	7.3	0.189	0.301	0.893	9 (a)-(c)
Local Quadratic					
LPR	9.6				6 (d)-(f)
MRR2	9.3	0.130	0.263	0.936	9 (d)-(f)
Local Cubic					
LPR	10.4				6 (g)-(i)
MRR2	10.0	0.135	0.261	0.934	9 (g)-(i)
Everhart Regression Model					
OLS	5.0	0.105	0.253		4
Local Linear					
LPR	7.1				7 (a)-(c)
MRR2	5.2	0.105	0.253	0.074	10 (a)-(c)
Local Quadratic					
LPR	9.6				7 (d)-(f)
MRR2	5.5	0.105	0.253	0.115	10 (d)-(f)
Local Cubic					
LPR	10.4				7 (g)-(i)
MRR2	5.7	0.105	0.253	0.126	10 (g)-(i)
Simple Linear Regression Model					
OLS	2.0	5.21	5.05		5
Local Linear					
LPR	7.1				8 (a)-(c)
MRR2	7.0	0.138	0.285	0.964	11 (a)-(c)
Local Quadratic					
LPR	9.6				8 (d)-(f)
MRR2	9.4	0.127	0.284	0.980	11 (d)-(f)
Local Cubic					
LPR	10.4				8 (g)-(i)
MRR2	10.3	0.128	0.283	0.984	11 (g)-(i)

Table 3. 9 Point Calibration Statistics [15 psid]

Model	$m.d.f.$	MSE($_{10^{-5}}$)	CCMSE($_{10^{-5}}$)	$\hat{\lambda}^*$	**Figure**
4th Order Polynomial Regression Model					
OLS	5.0	2.02	6.28		12
Local Linear					
LPR	5.3				15
MRR2	5.6	0.623	5.69	0.770	18
Everhart Regression Model					
OLS	5.0	0.0744	0.206		13
Local Linear					
LPR	5.3				16
MRR2	5.0	0.0746	0.206	0.019	19
Simple Linear Regression Model					
OLS	2.0	6.05	5.17		14
Local Linear					
LPR	5.3				17
MRR2	5.2	0.126	0.246	0.978	20

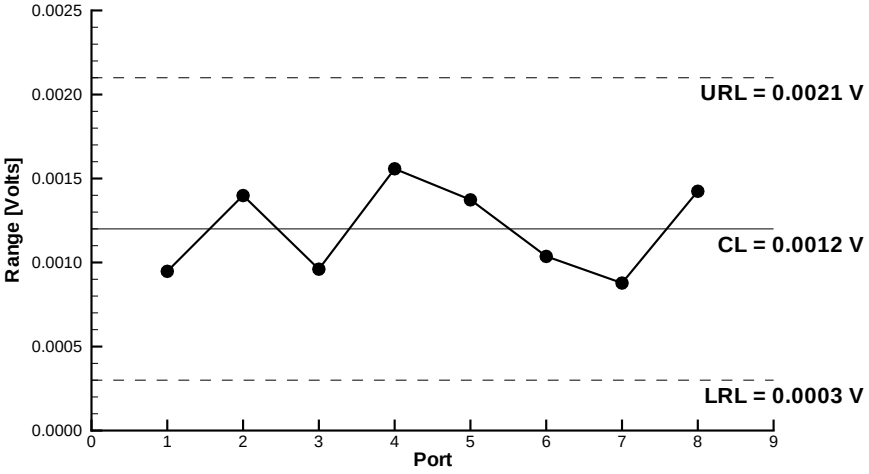


(a) Linear Scale (15 psid)

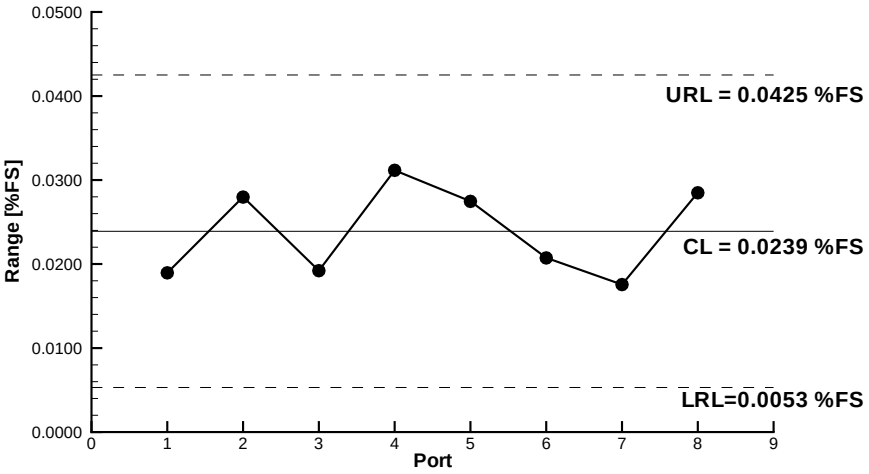


(b) Semilog Scale (15 psid)

Figure 1. ESP Calibration Data

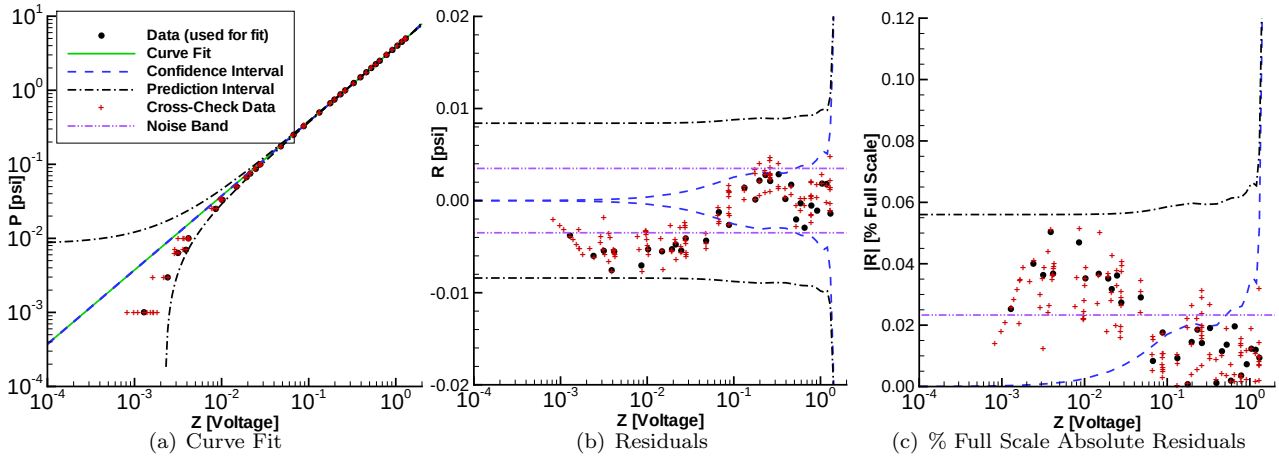


(a) Voltage (15 psid)



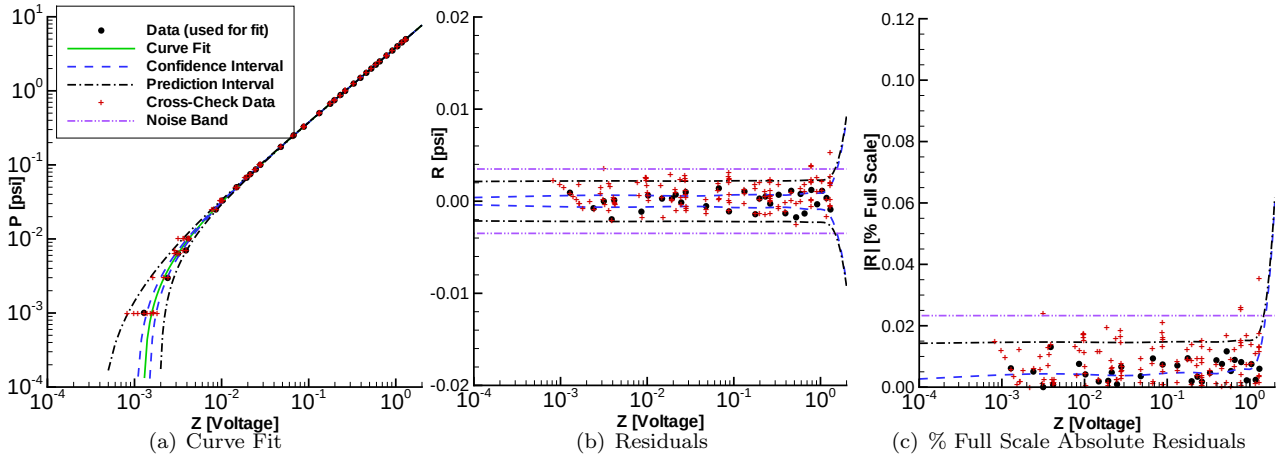
(b) % Full Scale (15 psid)

Figure 2. Zero Voltage Range Analysis



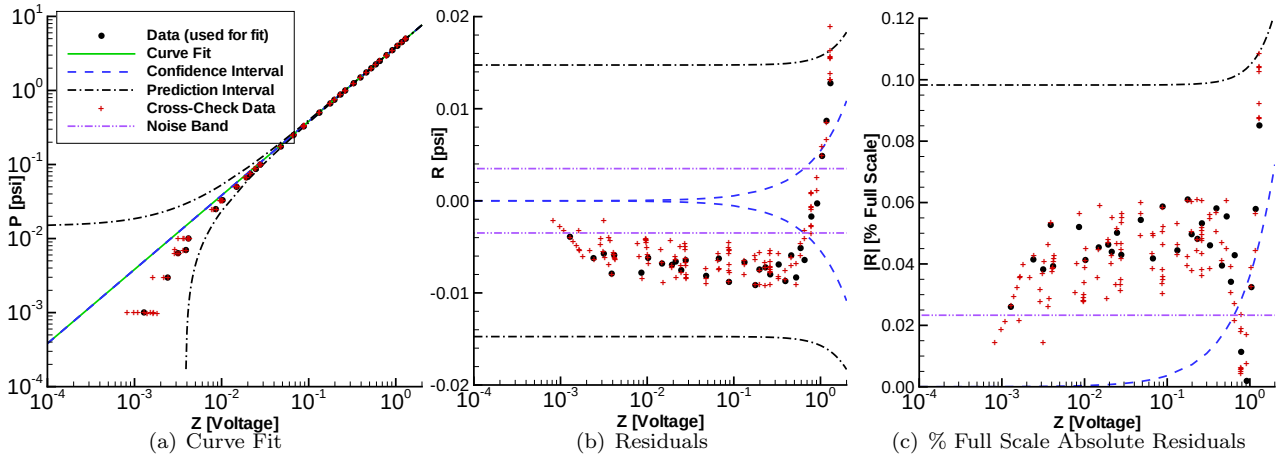
(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 3. 4th Order Polynomial Curve Fit [15 psid ESP/32 point Calibration]



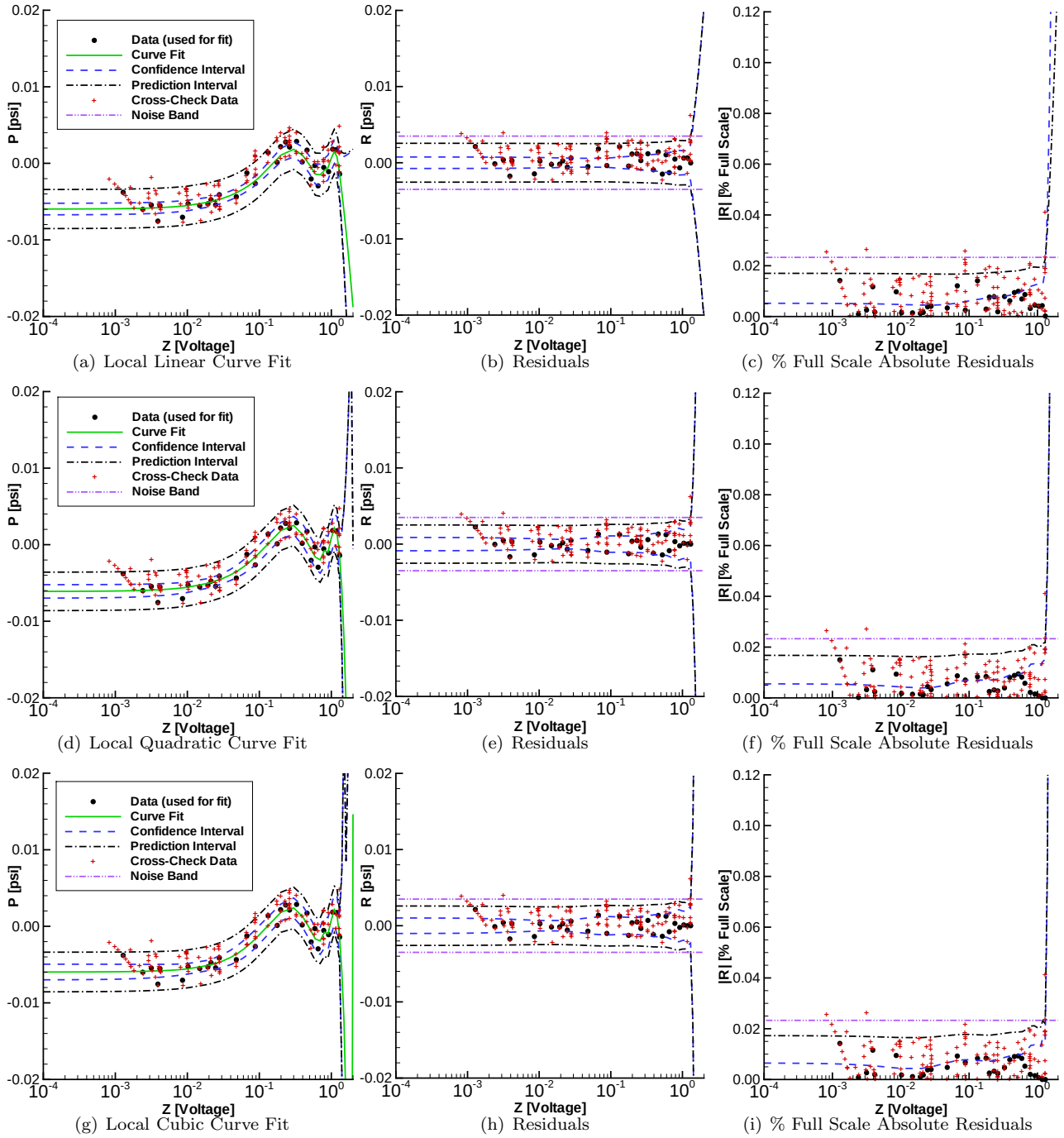
(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 4. Everhart Model Curve Fit [15 psid ESP/32 point Calibration]



(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 5. Simple Linear Fit [15 psid ESP/32 point Calibration]

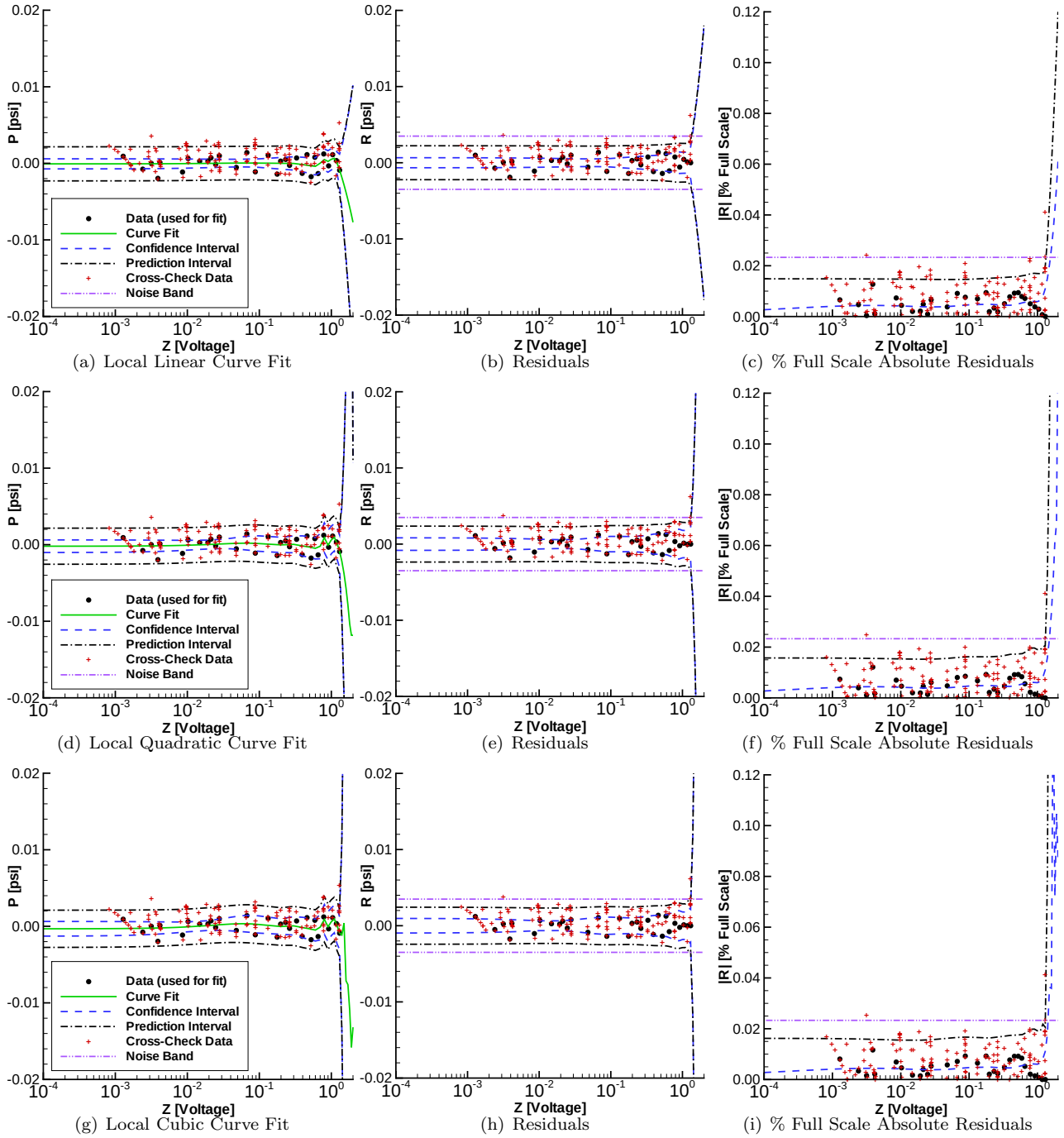


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 6. Nonparametric Curve Fit of 4th Order Polynomial Fit Residuals [15 psid ESP/32 point Calibration]

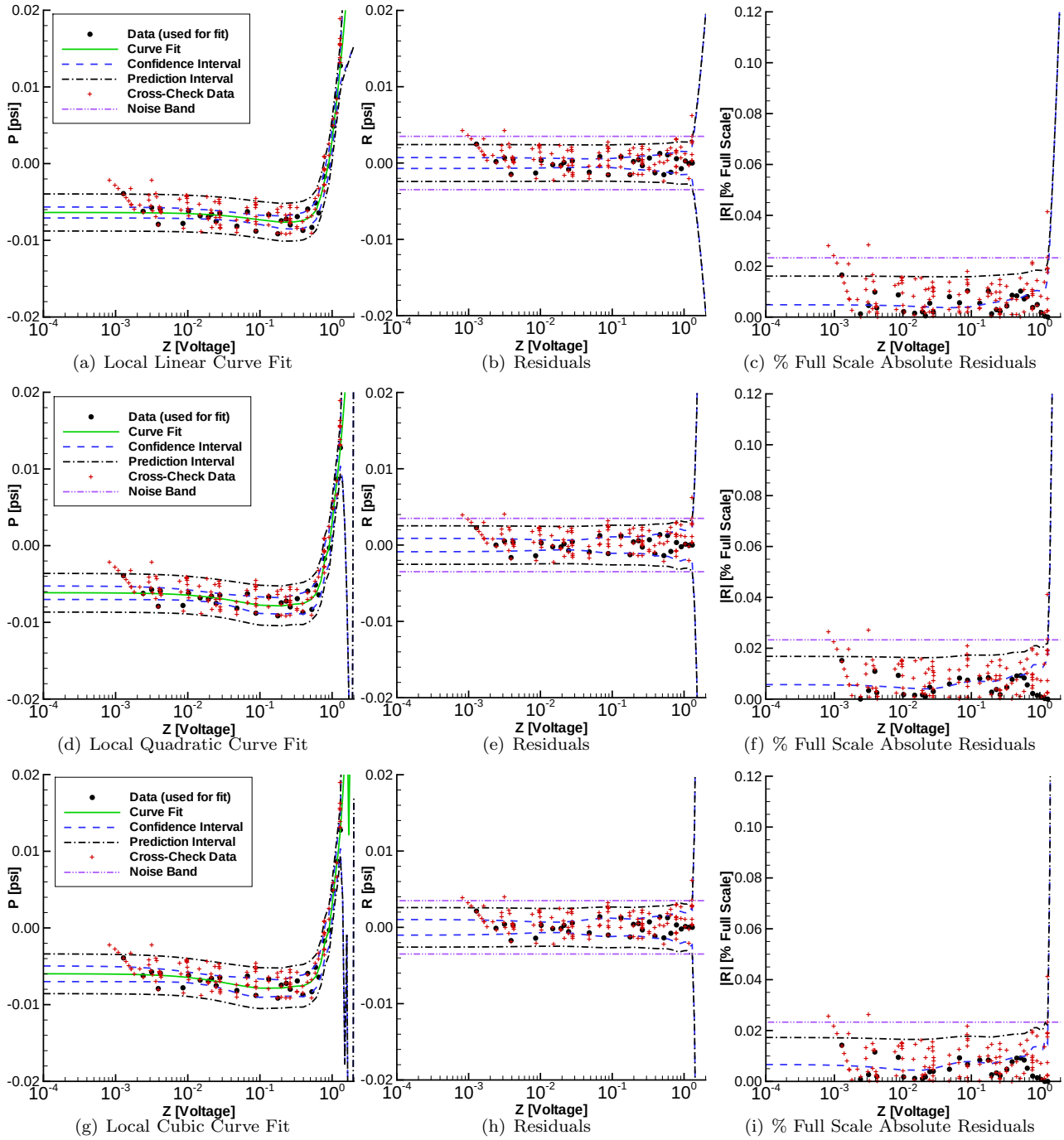


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 7. Nonparametric Curve Fit of Everhart Model Residuals [15 psid ESP/32 point Calibration]

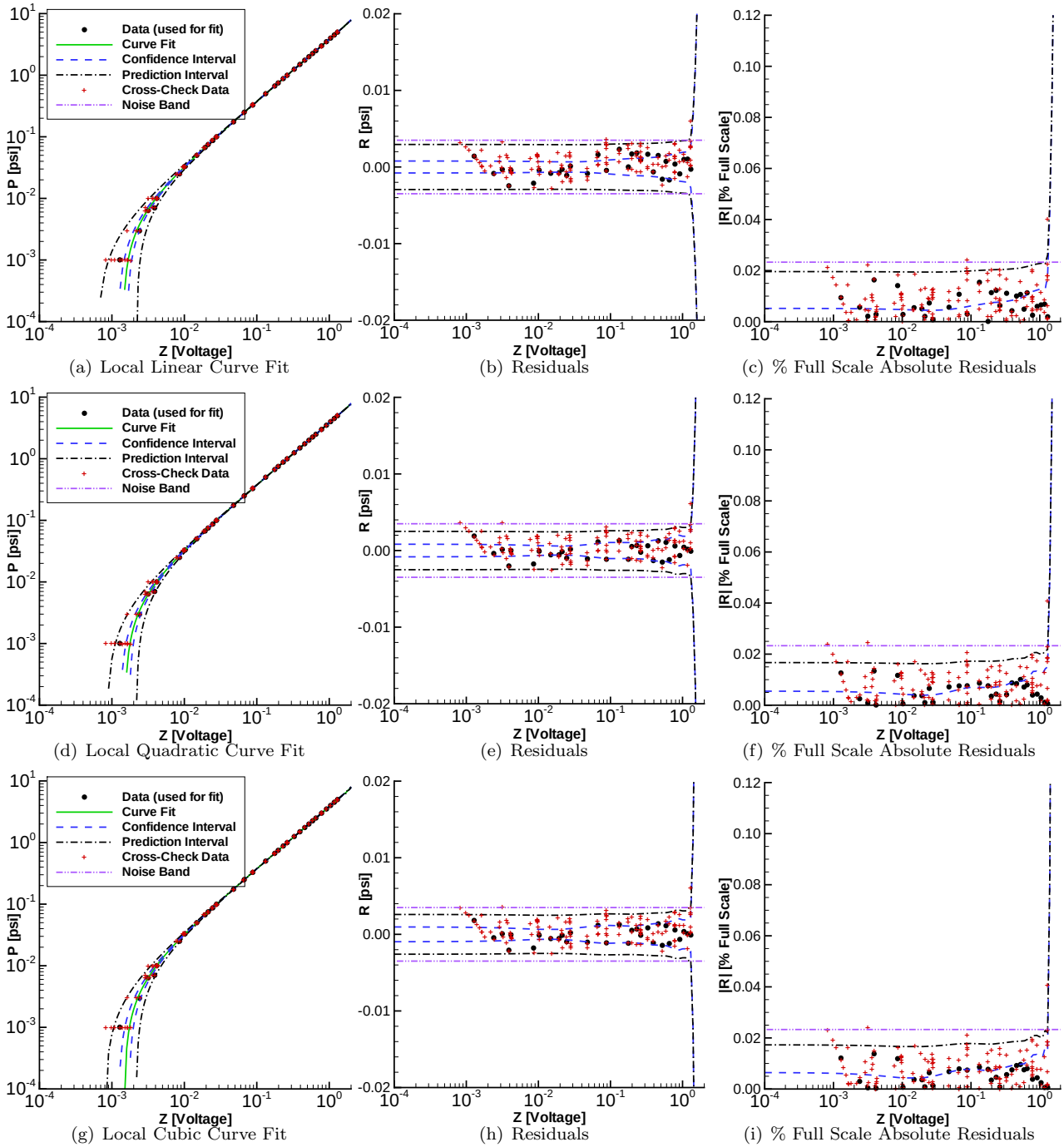


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 8. Nonparametric Curve Fit of Simple Linear Fit Residuals [15 psid ESP/32 point Calibration]

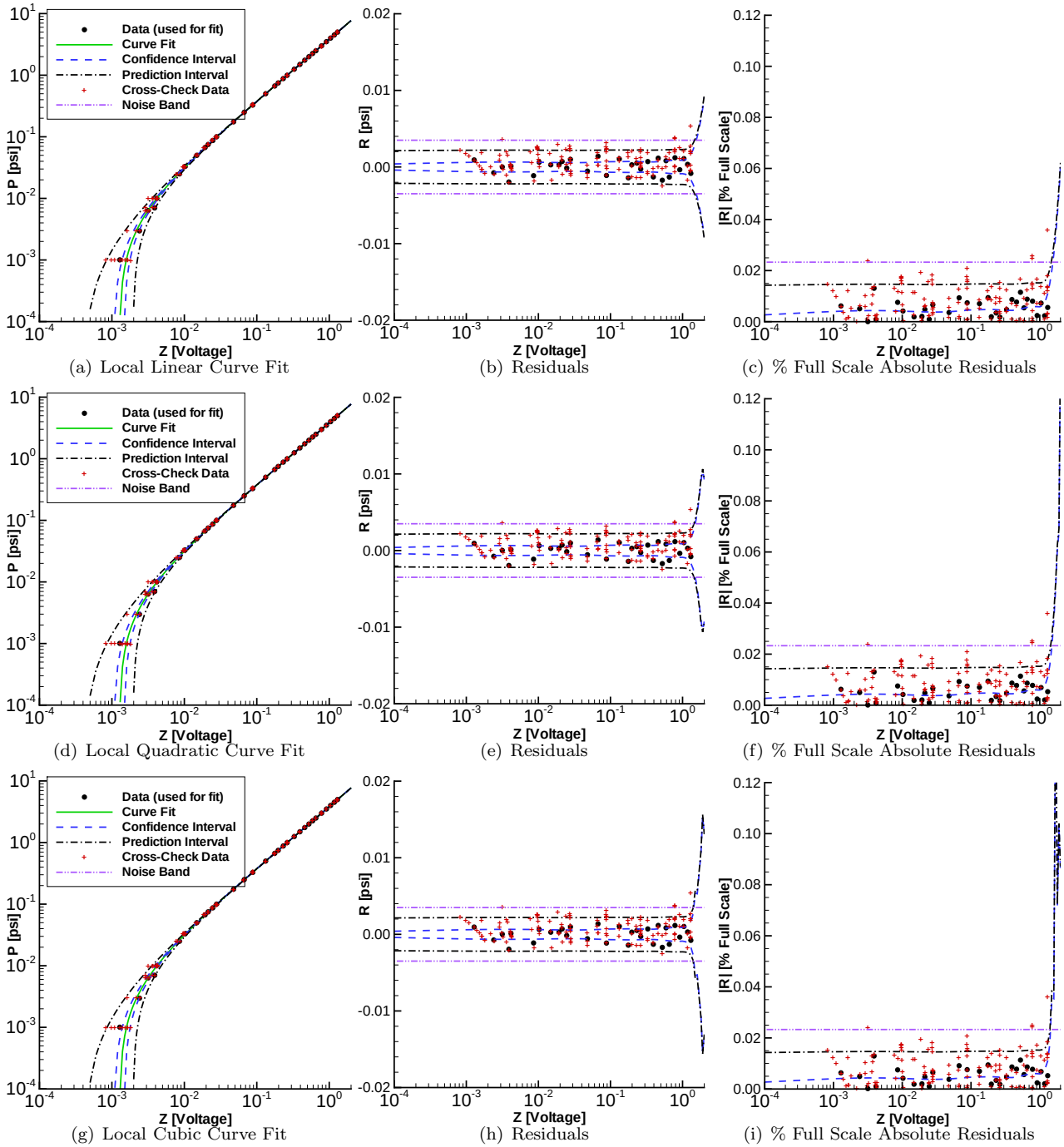


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 9. MRR2 Agumentation of the 4th Order Polynomial Fit [15 psid ESP/32 point Calibration]

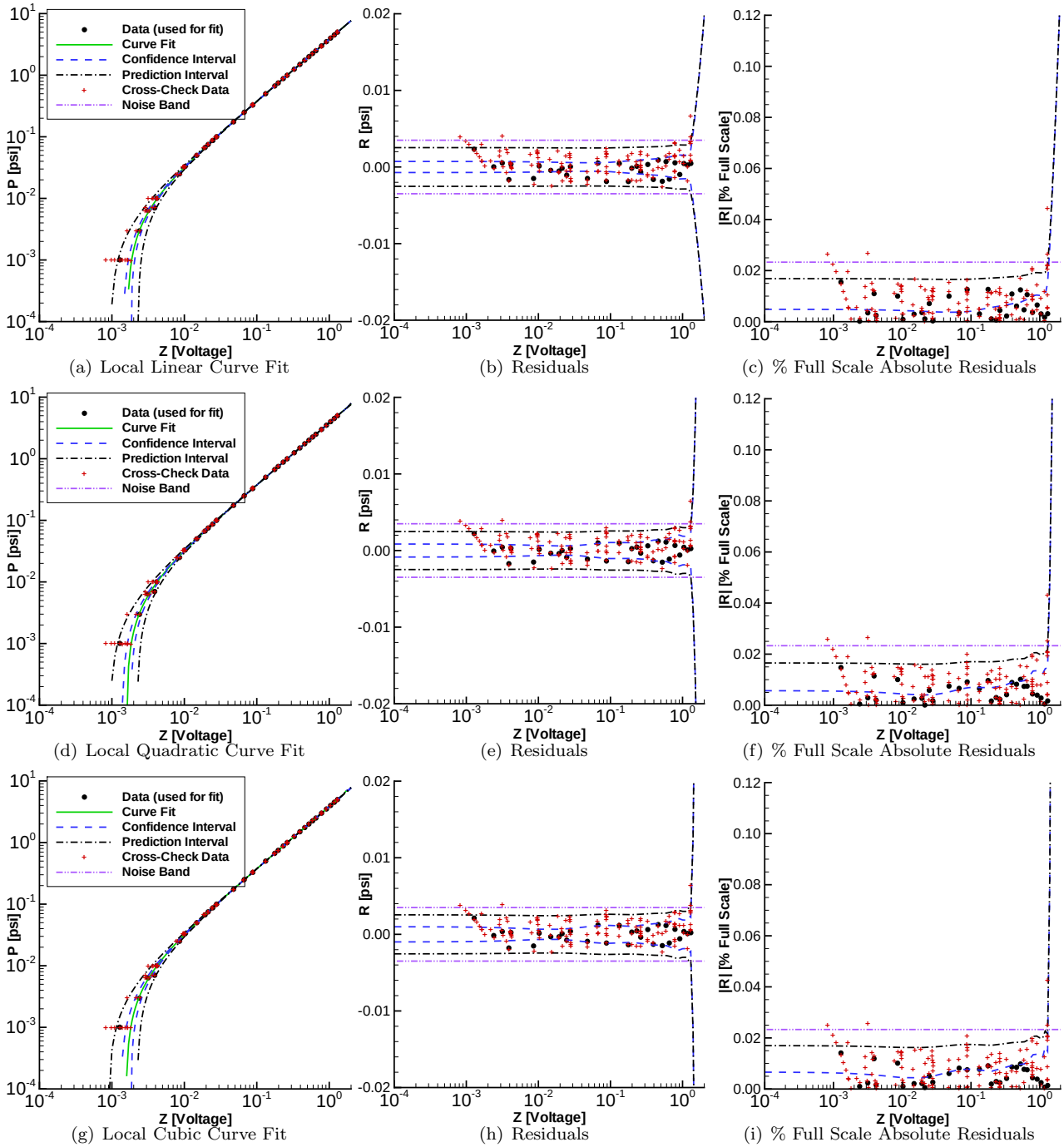


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 10. MRR2 Agumentation of the Everhart Model [15 psid ESP/32 point Calibration]

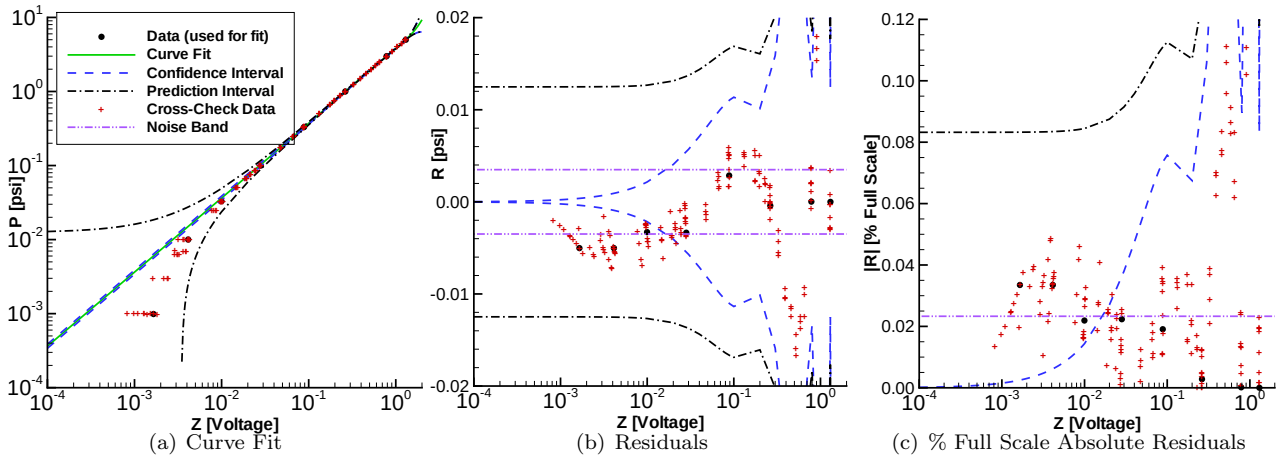


(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

(d) Local Quadratic Curve Fit (e) Residuals (f) % Full Scale Absolute Residuals

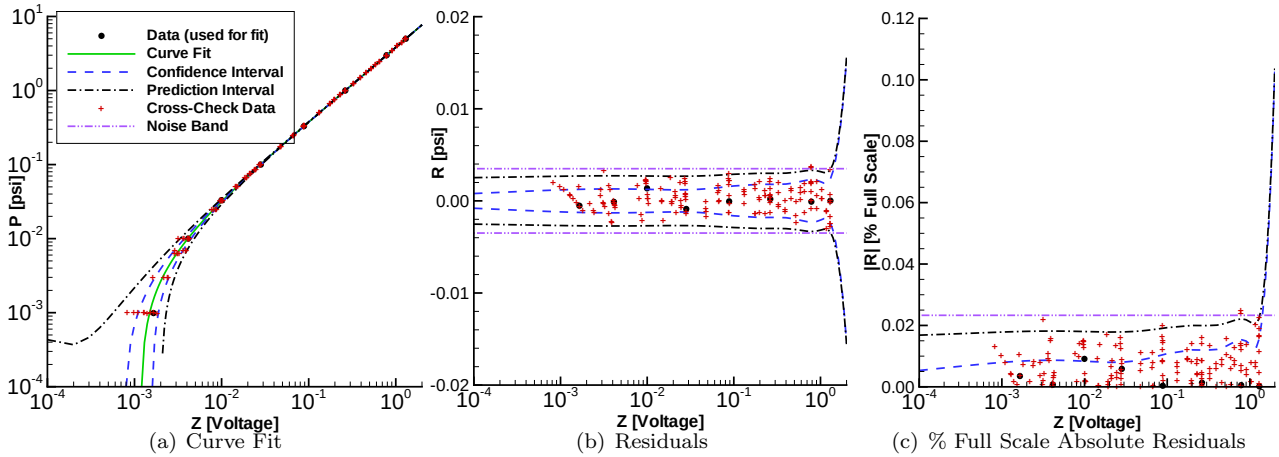
(g) Local Cubic Curve Fit (h) Residuals (i) % Full Scale Absolute Residuals

Figure 11. MRR2 Agumentation of the Simple Linear Fit [15 psid ESP/32 point Calibration]



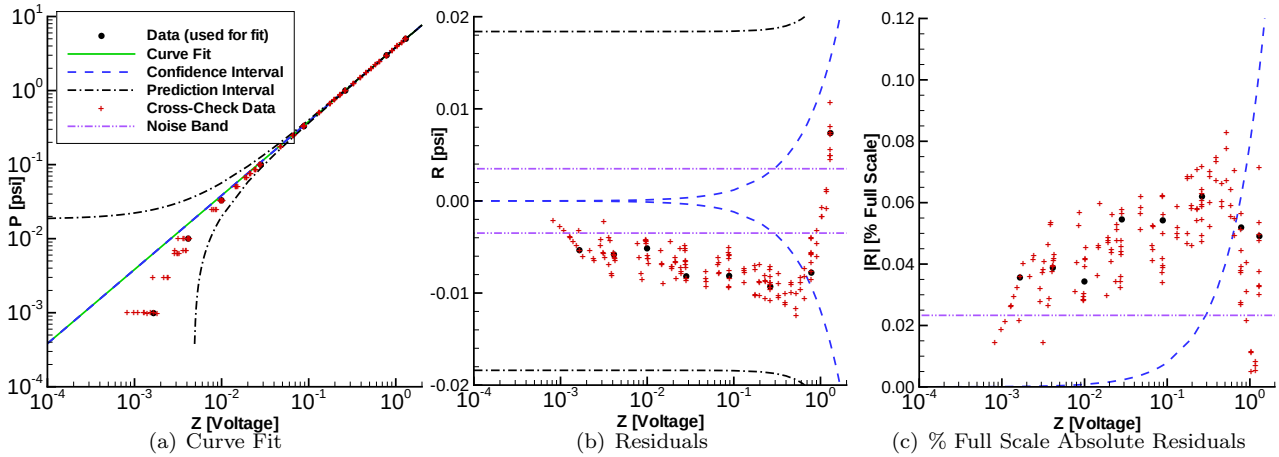
(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 12. 4th Order Polynomial Curve Fit [15 psid ESP/9 point Calibration]



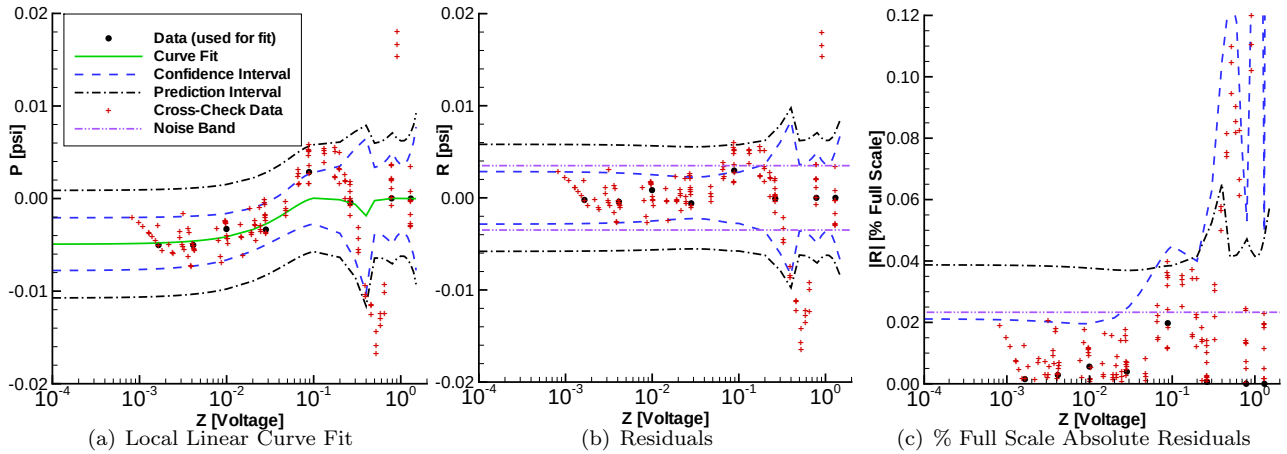
(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 13. Everhart Model Curve Fit [15 psid ESP/9 point Calibration]



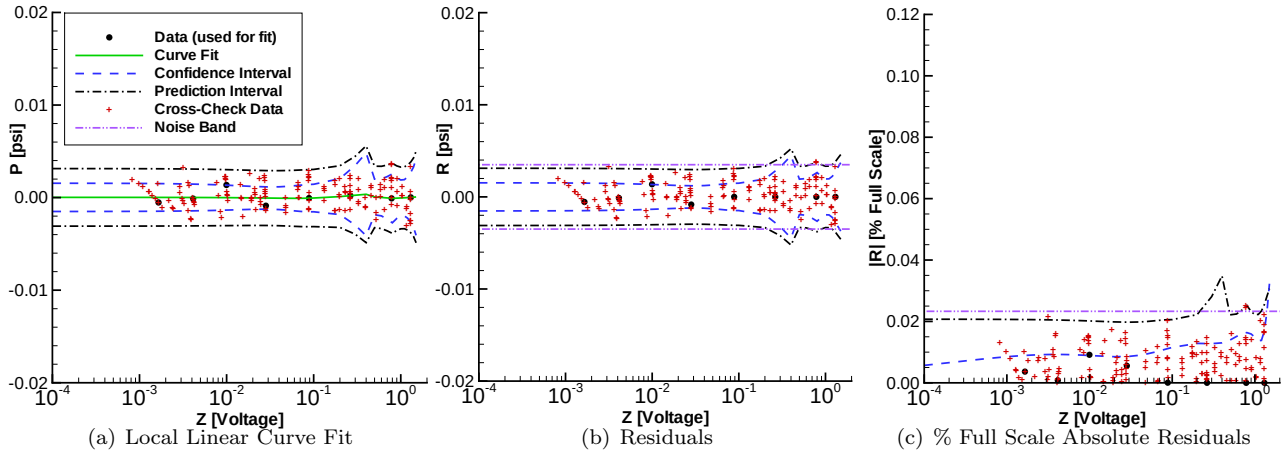
(a) Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 14. Simple Linear Fit [15 psid ESP/9 point Calibration]



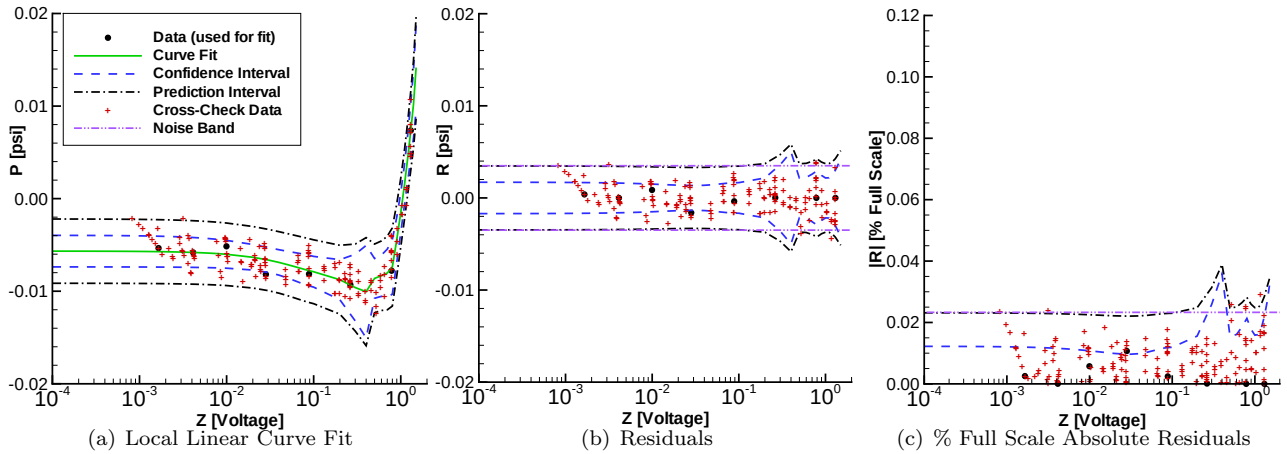
(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 15. Nonparametric Curve Fit of 4th Order Polynomial Fit Residuals [15 psid ESP/9 point Calibration]



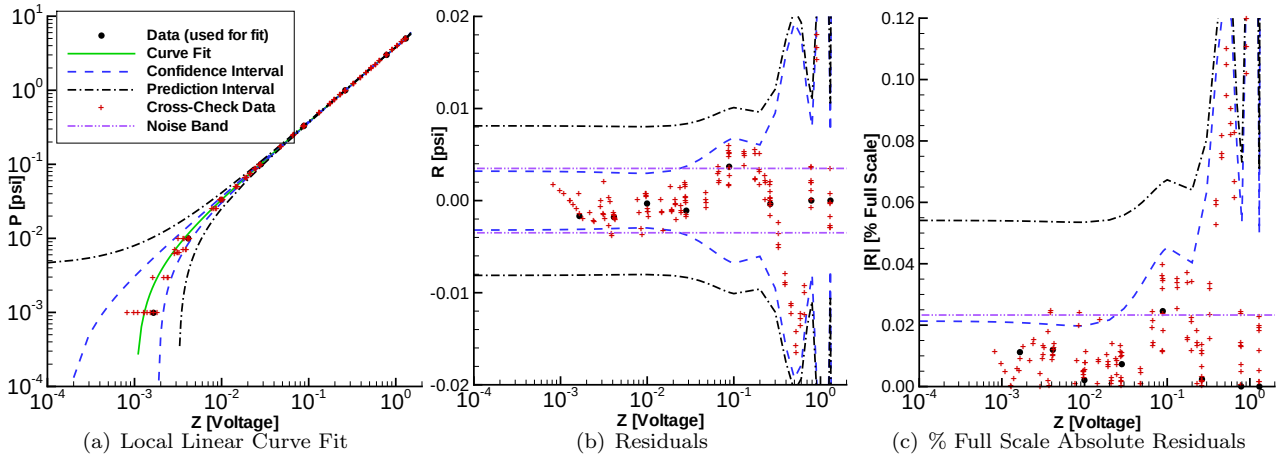
(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 16. Nonparametric Curve Fit of Everhart Model Residuals [15 psid ESP/9 point Calibration]



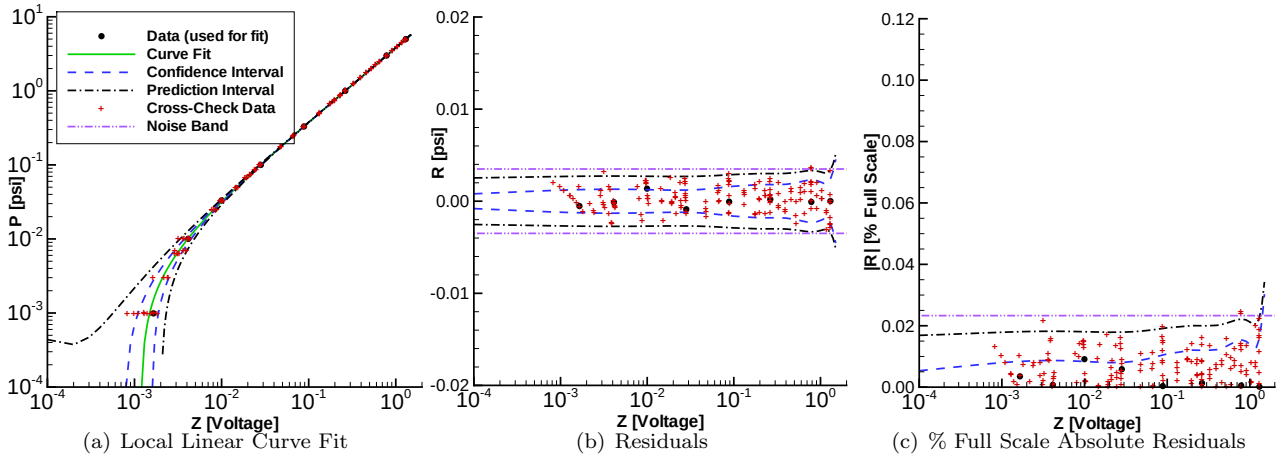
(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 17. Nonparametric Curve Fit of Simple Linear Fit Residuals [15 psid ESP/9 point Calibration]



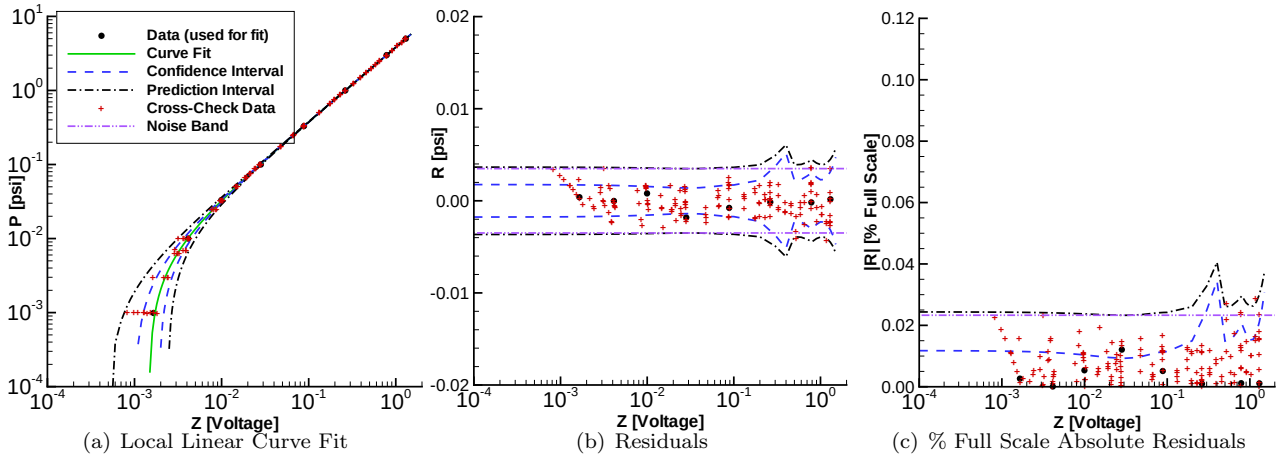
(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 18. MRR2 Agumentation of the 4th Order Polynomial Fit [15 psid ESP/9 point Calibration]



(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 19. MRR2 Agumentation of the Everhart Model [15 psid ESP/9 point Calibration]



(a) Local Linear Curve Fit (b) Residuals (c) % Full Scale Absolute Residuals

Figure 20. MRR2 Agumentation of the Simple Linear Fit [15 psid ESP/9 point Calibration]