Protocol Calibration in the National Resources Inventory

Cindy Yu
Jason C. Legg

August 23, 2007

ABSTRACT

The National Resources Inventory (NRI) is a large-scale longitudinal survey conducted by the National Resource Conservation Service in cooperation with the Center for Survey Statistics and Methodology (CSSM) at Iowa State University since 1982. A key NRI estimate is year-to-year change in acres of developed land, where developed land includes roads and urban areas. Since 2004, a fully digital data collection procedure has been used. The data from the NRI calibration experiment are used to estimate the relationship between data collected under the old and new protocols. A measurement error model is postulated for the relationship, where the duplicate measurements are used to estimate the error variances. If any significant discrepancy is detected between new and old measures, some parameters that govern the algorithm under new protocol can be changed to alter the relationship. The data analyses suggest that the relationship is a line with an intercept of zero and a slope of one, therefore the parameters in current use are acceptable. The paper also provides a way to model the measurement error variances as functions of the proportion of developed land, which is essential for estimating the effect of measurement error for the whole NRI data.

1 Introduction

During a long-term monitoring study, advances in theory and methodology for collecting data occur. Changing data collection procedures can reduce

measurement error and other nonsampling errors. For example, the introduction of computer assisted self-administered interviewing (CASI) has been shown to increase reporting accuracy for studies involving sensitive subjects (Tourangeau and Smith, 1996). A survey begun in the 1980s likely did not use CASI, but rather relied on telephone or in-person interviewing. Converting the survey from in-person interviewing to CASI could reduce bias in estimators. However, since measuring change is one of the primary objectives of longitudinal surveys, the effect of a change in survey mode needs to be measured. The data or survey instrument can be calibrated so that comparisons between data collected under current and prior modes are possible.

The National Resources Inventory (NRI) is a large-scale monitoring program designed to assess status, condition, and trends of soil, water, and related resources (Nusser and Goebel, 1997). Much of the NRI data are observed via photograph interpretation. Prior to 2000, photograph interpretation was performed on a transparent overlay on an aerial photograph. Now, the photographs are digitized and photograph interpretation is conducted on a computer. Along with the change to digital imagery, a new protocol has been created for determining area devoted to developed land in a land segment. A calibration study was conducted using 2003 data to assess the impact of the protocol change and whether adjustments are needed.

The NRI survey has a stratified two-stage design. For central (excluding Texas) and western states, the strata are defined by the Public Land Survey (PLS) System. For states under the PLS, a stratum is defined to be a two mile by six mile block, which is one-third of a township. Typically, two half-mile by half-mile blocks, called segments, are selected within a stratum. Within each selected segment, three points are selected using a restricted randomization procedure to ensure geographic spread. Segment level observations are made on the areas devoted to built-up zones, roads, streams, and small water bodies. We refer to structures and the maintained area around structures as urban land, roads as roads or railroads, and developed land as the combination of urban land and roads. Land use, composition, and erosion data are observed at the point level. The NRI protocol change is related to segment level observations.

The NRI longitudinal observation scheme was a pure panel from 1982 through 1997. Observations were made every five years. In 2000, the NRI began using a supplemented panel design with yearly data collection. A supplemented panel design has a pure panel component, called the core, and a rotating panel component composed of supplemental panels or supplements.

Segments were selected into the core and supplements using sampling rates determined by 1997 segment compositions. Segments containing points associated with wetlands, high erosion, or adjacent to urban areas were more likely to be selected into the core than other segments. We only considered segments in the core for the calibration study.

The change in the developed land observation protocol deals with assigning area to residences. The protocol for residential areas from 1982 through 2003 involved the data gatherers delineating the area around residences considered as urban. The delineation process involves tracing the boundary of a polygon using a hand planimeter on a transparent overlay on top of an aerial photograph. Under the new protocol, data gatherers place a cross using a mouse click on the roof of all of the residences on a digital photograph displayed on a computer monitor. A computer program places a hexagon centered on each cross on the digitized photograph. Two hexagons are linked if the distance between their boundaries is below a specified threshold. If four or more polygons are linked, the area of the polygons are considered developed land. An area entirely closed in by linked hexagons or other delineated built-up areas is considered built-up if the enclosed area is below another specified threshold. Roads are delineated by choosing a line thickness and tracing the road or by delineating the area around the road boundary. Non-residential urban areas are delineated using the vertex method. small water bodies are delineated like non-residential urban areas and small streams are delineated like roads. The protocols for collecting road, non-residential urban area, small water body, and small stream data are the same for previous and current data collection except that delineation is done on a computer rather than on a transparent overlay. In 2003, data were collected using both the new and old protocols, but only the new protocol has been used in data collection beginning in 2004.

The intent of the protocol change is to reduce the measurement error in urban area determinations. Marking residences is a more repeatable process than delineation, because the boundary of a delineated area is subject to the data gatherer's discretion. Roads and non-residential urban areas involve a decision on what portion of the land is maintained. Therefore, roads and non-residential determinations remain at the discretion of the data gatherer. Any change in the measurement error distribution for delineations of roads and non-residential determinations is due to changes in the quality of data collection materials.

The data from the NRI calibration experiment are used to estimate the

relationship between data collected under the old and new protocols. If the relationship is not a line with an intercept of zero and a slope of one, parameters in the program that translates crosses into areas will be modified. The size of the hexagons and linking rules can be changed to alter the relationship between observations made under the different protocols. A second objective of the experiment is to provide an estimate of the relative contribution of the measurement error variance to the total variance of the estimator. The calibration study involves repeated observations, which allows estimates of measurement error variances for developed land. We will model the measurement error variances as a function of developed land.

2 Experiment Design

NRI observations are taken with measurement error. Therefore, the regression of measurements under the new protocol on measurements under the prior protocol is biased (Fuller 1987). In order to correct the bias, estimators of the measurement error variances under the two protocols are necessary. The calibration experiment was designed with replicates for measurement error variance estimation. Individual segments were selected based on geography and 2003 measurements under the previous protocol.

The NRI data gatherers have access to previously collected data. Therefore, the measurement error is assumed to be correlated over time. The data collection procedure was designed to reduce the correlation between two observations made on the same segment in 2003. Four people are involved in data collection under the new protocol. The first two people make 2001 observations using the available 1997 materials. The third and fourth person make 2003 observations, where the third person uses materials from the first 2001 data gatherer and the fourth person uses materials from the second 2001 data gatherer. A fifth person has made a determination for 2003 under the old protocol previously. Eight data collectors are grouped together. For each eight segments, four data collectors are randomly selected to work on the first four segments and the complement set of data collectors are assigned to work on the second set of four segments. A Latin square design assigns the four segments to the four data collectors such that each data collector performs each of the four observation types once. Some control is made across groups of eight segments to ensure mixing of data collectors into the groups of four. A working assumption under this design is that the two observations under

the new protocol made in 2003 are independent and are also independent of the observation made under the old protocol. The independence assumption is justified by the inclusion of the intermediate data collector between the 2003 data collector and the original 1997 data collector.

Photograph interpretation occurs at three Remote Sensing Laboratories (RSL). The RSLs are termed West, Central, and East. Each RSL collects data on states in the region of the RSL. Data gatherers at each RSL receive support first from leaders within the RSL. Training of data gatherers also occurs at each RSL. Differences between data collection techniques can arise due to differences in leadership at the RSLs. Therefore, the segment selection occurred within RSL regions.

Segments were divided based on the area of developed land and area of small water and small streams. Segments completely covered with water, federal land, or developed land are not highly interesting for the experiment because the residence protocol will not need to be applied. Therefore, segments classified as 100% urban, 100% federal, or 100% water were not included in the study. Alaska and Hawaii were not included in the selection. Segments were selected from the remaining pool to have a spread of land features. Segments with a change in urban, water, or road determinations from 2001 to 2003 under the old protocol were selected with certainty for the Central and West RSLs. A subset of segments in these categories were selected with certainty for the East RSL. The remaining segments were divided into strata defined by presence of developed land with no change, presence of water with no change and no developed land, and no water or developed land in segment with no change from 2003. Within each category, segments were sorted by a geographic code and a systematic sample was selected. A total of 2699 segments were selected into the study. The West RSL had 608 segments, the Central RSL had 1055 segments, and the East RSL had 1036 segments.

The sample selection is biased, since segments were selected on the basis of 2003 observations. Segments without developed land in 2003 under the old protocol were not selected at a high rate and those selected contained water features. Therefore, the occurrence of a segment without developed land under the old protocol and developed land under the new protocol is much less in our sample than the occurrence of a segment with developed land under the old protocol and no developed land under the new protocol. This sampling bias affects the estimated relationship between old and new protocol observations near a true developed land value of zero acres.

The data used in this paper are observations made using 2003 photography from the West RSL. Developed land areas were converted into proportions by dividing built-up determinations by digitized segment size. Segments containing federal land were removed from the analysis dataset, because the boundary of federal land within a segment cannot be determined. Twenty-seven calibration segments contain some federal land. An additional seventy-seven segments where all three 2003 observations have no developed land and one segment with 100% developed land were removed from the analysis dataset. Including segments with no or all developed land would increase the evidence that observations under the new and old protocol estimate the same quantity, possibly masking some match departures away from the extremes. Some data were modified after initial data collection through a review process. The review process ensures that the new protocol is calibrated to a properly applied old protocol. The dataset used for this analysis contains 503 segments.

3 Estimation of the mean function

We define developed land to be the sum of segment areas reported for large urban, small urban, and public roads. The variable of interest is the proportion of developed land in a segment in 2003. Let $B_{j,i,03}$ be the developed land area observation for segment i made by observer j using 2003 materials, where j can be 1 or 2 for observations made under the new protocol or o for the observation made under the old protocol. Let S_i be the digitized size of segment i. The three estimated proportions of developed land in segment i are

$$Y_{1i} = B_{1,i,03}/S_i, \tag{1}$$

$$Y_{2i} = B_{2,i,03}/S_i, \tag{2}$$

and

$$X_i = B_{o,i,03}/S_i. \tag{3}$$

The mean of X_i is 0.231 (0.011) and the mean of $2^{-1}(Y_{1i} + Y_{2i})$ is 0.232 (0.011). The correlation between X_{1i} and $2^{-1}(Y_{1i} + Y_{2i})$ is 0.954.

A proposed model for the 2003 data is

$$X_i = x_i + u_i, \tag{4}$$

$$Y_{ji} = \beta_0 + \beta_1 x_i + e_{ji}, \tag{5}$$

$$x_i \sim (\mu_x, \sigma_x^2), \tag{6}$$

and

$$\begin{bmatrix} u_i \\ e_{1i} \\ e_{2i} \end{bmatrix} | x_i \sim \left[0, \begin{bmatrix} \sigma_{ui}^2 & 0 & 0 \\ 0 & \sigma_{ei}^2 & 0 \\ 0 & 0 & \sigma_{ei}^2 \end{bmatrix} \right] \tag{7}$$

for all i and $j = 1, 2$, where x_i is the true proportion for segment i, and u_i and e_{ji} are measurement errors on segment i under the old and the new protocols, respectively. The errors are assumed independent across different segments. By the experiment design, we assume e_{1i}, e_{2i} and u_i are conditionally independent from each other for each segment i.

To estimate the parameters, we define the observation vector

$$\boldsymbol{Z} = (Z_1, Z_2, Z_3) = (X_i, 2^{-1}[Y_{1i} + Y_{2i}], \sqrt{2}^{-1}[Y_{1i} - Y_{2i}]). \tag{8}$$

Let the sample covariance matrix of $\boldsymbol{Z}$ be

$$\boldsymbol{m} = (n-1)^{-1} \sum_{i \in A} (\boldsymbol{Z}_i - \overline{\boldsymbol{Z}})'(\boldsymbol{Z}_i - \overline{\boldsymbol{Z}}). \tag{9}$$

Under the model, the sample covariance matrix has expectation

$$E(\boldsymbol{m}) = \begin{bmatrix} \sigma_x^2 + \sigma_{a,u}^2 & \beta_1 \sigma_x^2 & 0 \\ \beta_1 \sigma_x^2 & \beta_1^2 \sigma_x^2 + 0.5\sigma_{a,e}^2 & 0 \\ 0 & 0 & \sigma_{a,e}^2 \end{bmatrix} \tag{10}$$

where $\sigma_{a,u}^2$ and $\sigma_{a,e}^2$ denote the averages of σ_{ui}^2 and σ_{ei}^2, respectively, and σ_x^2 is the variance of the targets of the old procedure for the calibration sample and does not have an interpretation for the whole NRI. Method of moments estimators are

$$\widehat{\beta}_0 = \overline{Z}_2 - \widehat{\beta}_1 \overline{Z}_1, \tag{11}$$

$$\widehat{\beta}_1 = m_{12}^{-1}(m_{22} - 0.5m_{33}), \tag{12}$$

$$\widehat{\sigma}_x^2 = (m_{22} - 0.5m_{33})^{-1} m_{12}^2, \tag{13}$$

$$\widehat{\sigma}_{a,e}^2 = m_{33}, \tag{14}$$

$$\widehat{\sigma}_{a,u}^2 = m_{11} - \widehat{\sigma}_x^2, \tag{15}$$

and

$$\widehat{\theta} = \frac{\widehat{\sigma}_{a,u}^2}{\widehat{\sigma}_{a,e}^2}. \tag{16}$$

Table 1: Parameter Estimates from the Original Model for West RSL Data

$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\sigma}^2_x$	$\widehat{\sigma}^2_{a,e}$	$\widehat{\sigma}^2_{a,u}$	$\widehat{\theta}$
0.0011	0.998	0.0556	0.00086	0.00498	5.8
(0.0032)	(0.015)	(0.0044)	(0.00022)	(0.00058)	(2.0)

The function $\widehat{\theta}$ estimates the ratio of the average error variance under the old and new protocols. Parameter estimates using estimators (11)-(16) are in Table 1 for data from the West RSL. Standard errors are estimated using a delete-1 jackknife, where the jackknife weights are those of simple random sampling. From the global parameter estimates, the intercept is not statistically significant from 0 and the slope is not statistically significantly different from 1. The error variance under the old protocol is estimated to be 5.8 times the error variance under the new protocol.

Nonlinearity of the fixed relationship between Y and x needs to be investigated in case the actual relationship is not a straight line. Therefore, we propose a second model that allows for a shift in the slope. Our new model specifies a slope change at $x = 0.5$. [1]

Because x_i is not directly observed, we construct an estimator for x_i using an estimated generalized least squares (EGLS) procedure. That is, $(Z_{1i}, Z_{2i} - \widehat{\beta}_0)$ is regressed on $(1, \widehat{\beta}_1)'$ using relative weights equal to the inverses of the estimated average variances, where $(\widehat{\beta}_0, \widehat{\beta}_1)$ are the estimated coefficients in Table 1. That is

$$\widehat{x}_i = w_1 Z_{1i} + w_2(Z_{2i} - \widehat{\beta}_0)/\widehat{\beta}_1 \tag{17}$$

where $w_1 = \widehat{\sigma}^{-2}_{a,u}/(\widehat{\sigma}^{-2}_{a,u} + 2\widehat{\beta}^2_1\widehat{\sigma}^{-2}_{a,e})$, and $w_2 = 2\widehat{\beta}^2_1\widehat{\sigma}^{-2}_{a,e}/(\widehat{\sigma}^{-2}_{a,u} + 2\widehat{\beta}^2_1\widehat{\sigma}^{-2}_{a,e})$.

To ease the computation of parameter estimates, we write the inverted model as a regression of X_i on y_i, the true measurement on segment i under the new protocol. The model with a slope change at $x = 0.5$ is

$$Y_{ji} = y_i + e_{ji}, \tag{18}$$

$$X_i = \delta_0 + \delta_1 y_i v_i + 0.5\delta_1(1 - v_i) + \delta_2(y_i - 0.5)(1 - v_i) + u_i, \tag{19}$$

[1] We could choose other points to split the line. However, plots of the average of new measurements versus the old suggest $x = 0.5$ gives a good chance to detect a trajectory difference.

$$v_i = \begin{cases} 1 & x_i < 0.5 \\ 0 & x_i \geq 0.5, \end{cases} \tag{20}$$

and the moments of the errors, e_{ji} and u_i, are defined as before. The indicator variable, v_i, is estimated by $\widehat{v}_i$ by replacing x_i with $\widehat{x}_i$ in (20). To estimate δ_0, δ_1 and δ_2, we regress Z_{1i} (i.e. X_i) on $G_i = (1, Z_{2i}\widehat{v}_i + 0.5(1-\widehat{v}_i), (Z_{2i}-0.5)(1-\widehat{v}_i))$. Since Z_{2i} are taken with measurement error, the above regression is biased (Fuller 1987). We adjust the regression estimators to account for the effect of measurement error. By writing the model as regression of X_i on y_i, the complexity in computing the correction matrices is reduced. Let A_1 denote the part of the sample where $\widehat{v}_i = 1$ and A_2 denote the part of the sample where $\widehat{v}_i = 0$. The bias corrected regression estimator is

$$\begin{bmatrix} \widehat{\delta}_0 \\ \widehat{\delta}_1 \\ \widehat{\delta}_2 \end{bmatrix} = (G'G - C)^{-1}(G'Z_1), \tag{21}$$

where

$$C = \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_1 & 0 \\ 0 & 0 & C_2 \end{bmatrix} \tag{22}$$

$$C_1 = \sum_{i \in A_1} 0.25(Y_{1i} - Y_{2i})^2, \tag{23}$$

and

$$C_2 = \sum_{i \in A_2} 0.25(Y_{1i} - Y_{2i})^2. \tag{24}$$

Under the model described in (18) and (19), the sample covariance of $\boldsymbol{Z}$ can be used to estimate the average error variances once the regression coefficients are obtained. The estimator for $\sigma^2_{a,e}$ is m_{33} from (14). An estimator for $\sigma^2_{a,u}$ is obtained by combining estimators from the two parts of the data set. Let m_v be the sample covariance matrix of (Z_1, Z_2) for data with $\widehat{v}_i = 1$ and m_{1-v} be the covariance sample covariance matrix of (Z_1, Z_2) for data with $\widehat{v}_i = 0$. The expectation of the mean squares is

$$E\{m_v\} = \begin{bmatrix} \delta_1^2\sigma^2_{vy} + n_1^{-1} \sum_{i \in A_1} \sigma^2_{ui} & \delta_1\sigma^2_{vy} \\ \delta_1\sigma^2_{vy} & \sigma^2_{vy} + \sum_{i \in A_1} (2n_1)^{-1}\sigma^2_{ei} \end{bmatrix} \tag{25}$$

Table 2: Parameter Estimates from the Split Line Model for West RSL Data

$\widehat{\delta}_0$	$\widehat{\delta}_1$	$\widehat{\delta}_2$	$\widehat{\sigma}^2_{a,e}$	$\widehat{\sigma}^2_{a,u}$	$\widehat{\theta}$
0.0012	0.985	1.045	0.00086	0.00454	5.3
(0.0033)	(0.027)	(0.064)	(0.00022)	(0.00077)	(1.7)

and

$$E\{m_{1-v}\} = \begin{bmatrix} \delta_2^2\sigma^2_{1-v,y} + n_2^{-1}\sum_{i\in A_2}\sigma^2_{ui} & \delta_2\sigma^2_{1-v,y} \\ \delta_2\sigma^2_{1-v,y} & \sigma^2_{1-v,y} + \sum_{i\in A_2}(2n_2)^{-1}\sigma^2_{ei} \end{bmatrix} \tag{26}$$

where n_1 is the size of A_1, n_2 is the size of A_2, σ^2_{vy} is the variance of y_i in A_1, and $\sigma^2_{1-v,y}$ is the variance of y_i in A_2. Method-of-moments estimators for σ^2_{vy} and $\sigma^2_{1-v,y}$ are

$$\widehat{\sigma}^2_{vy} = \widehat{\delta}_1^{-1} m_{v,12} \tag{27}$$

and

$$\widehat{\sigma}^2_{1-v,y} = \widehat{\delta}_2^{-1} m_{1-v,12}. \tag{28}$$

An estimator for $\sigma^2_{a,u}$ is

$$\widehat{\sigma}^2_{a,u} = (n_1+n_2)^{-1}(n_1\{m_{v,11} - \widehat{\delta}_1^2\widehat{\sigma}^2_{vy}\} + n_2\{m_{1-v,11} - \widehat{\delta}_2^2\widehat{\sigma}^2_{1-v,y}\}). \tag{29}$$

Estimators for the parameters of model (18)-(19) are in Table 2. Standard errors were computed using a delete-1 jackknife.

The intercept is not statistically significantly different from zero and both the slopes before $x = 0.5$ and after $x = 0.5$ are not statistically significantly different from 1 (Figure 1). The values of estimated $\widehat{\sigma}^2_{a,u}$ and $\widehat{\theta}$ are smaller than the corresponding estimates from model (4)-(5), but the difference in estimates is not large.

We compute an approximate F-test of

$$H_o : (\delta_0, \delta_1, \delta_2) = (0, 1, 1) \tag{30}$$

versus

$$H_a : (\delta_0, \delta_1, \delta_2) \neq (0, 1, 1). \tag{31}$$

The F statistic is 0.52, which when compared to F distribution with 3 and 497 degrees of freedom results in a p-value of 0.67. Therefore, we accept the reduced model of

$$Y_{ji} = y_i + e_{ji}, \tag{32}$$

Figure 1: Fitted split line model with binned Z1 and Z2 means

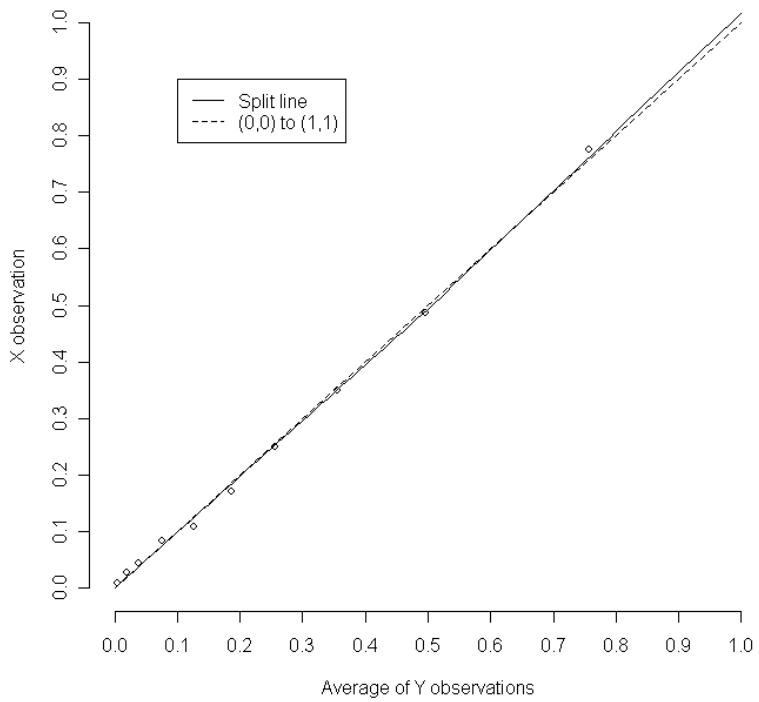


Table 3: Parameter Estimates from the Reduced Model for West RSL Data

$\widehat{\sigma}_x^2$	$\widehat{\sigma}_{a,e}^2$	$\widehat{\sigma}_{a,u}^2$	$\widehat{\theta}$
0.0555	0.00086	0.00498	5.8
(0.0042)	(0.00022)	(0.00058)	(2.0)

$$X_i = y_i, \tag{33}$$

where the moment and independence assumptions of e_{ji} and u_i are the same as (7).

Figure 1 contains both the fitted split line (dashed) and the $(0,0)$ to $(1,1)$ reference line (solid). In order to asses the model fit, we divide the data set into 10 bins with equal number of observations up to rounding from data sorted by $\widehat{x_i}$ values. Figure 1 shows the mean of Z_2 versus Z_1 in each bin. The binned means lie closely around the lines, indicating the reduced model $((\delta_0, \delta_1, \delta_2) = (0, 1, 1))$ suffices for describing the data.

Collectively, the result of the F-test and the evidence in Figure 1 suggest that the relationship between data collected under the old and new protocols is a line with an intercept of zero and a slope of one. Therefore, no additional modifications to the program need to be made for the West. Under the reduced model, the estimated average error variances can be obtained using Equations (13) to (15), where the coefficients β_0 and β_1 in (10) are replaced by 0 and 1, respectively (Table 3).

In addition, we test whether the mean of Z_1 is statistically significantly different from the mean of Z_2 within each bin using approximate t-test (Table 4). The t-statistics are constructed as bias adjusted Beale ratios to account for skewness (Tin 1965). The t-tests provide evidence that the observations under the new and old protocol do not correspond well for segments with little developed land. However, misfitting the actual trend near $x_i = 0$ will not result in large changes in total estimates. Part of the difference between new and old protocols is attributable to the bias in the calibration sample selection. Further, several of the differences in the small bins are due to differences in delineation of small road pieces. The new protocol is the same as the old protocol for road measurements. Changing the protocol will not affect the differences due to roads. Therefore, parameters in the program that translates crosses into areas under new protocol are accepted for the West RSL.

Table 4: Approximate t-test for the Differences Between Z1 and Z2 Over Ten Bins

Bins	1	2	3	4	5
mean of Z_1	0.008	0.027	0.045	0.083	0.109
mean of Z_2	0.004	0.020	0.037	0.076	0.127
t-value	4.31	2.70	1.82	1.17	-2.21
Bins	6	7	8	9	10
mean of Z_1	0.172	0.249	0.350	0.487	0.775
mean of Z_2	0.187	0.255	0.355	0.495	0.756
t-value	-1.49	-0.32	-0.25	-0.55	1.45

4 Estimation of the variance function

The calibration experiment provides the opportunity to estimate the effect of measurement error on NRI estimators. In order to extend the variance results to a larger set of data than the calibration data set, we need a functional form for the measurement error variance. The reason for this requirement is that the calibration experiment is a biased sample of the NRI segments. If the measurement error variance is a function of x_i, then the estimates of the average variances depends on the set of x_i chosen for the calibration experiment. Modeling the variance functions is difficult due to a few extreme differences between observations made on the same segment. Model assumptions presented below are made to construct estimators of the measurement error variance functions. However, the assumptions are not believed to be true for all of the data, nor would many standard diagnostic procedures be possible to check the validity of assumptions.

The expectation of the squared deviations Z_{3i}^2 and $(Z_{1i} - Z_{2i})^2$ are estimated as a function of the true proportion of developed land. As with the estimation of the mean function before, the estimated true proportion is used as a proxy for the truth. Two constraints are put on the functional form of variances. One constraint is that the functions be symmetric around 0.5. The underlying assumption is that delineations of developed lands when the true proportion are 40% and 60% are associated with the same level of difficulty. In other words, the delineation of an area in a particular segment has same effort as the delineation of the complement of the area. The second constraint is that the variance of Z_{3i} is proportional to the variance of $(Z_{1i} - Z_{2i})$. A plot of $Z_{3i}^{-2}(Z_{1i} - Z_{2i})^2$ versus x_i is flat except near zero and

one, providing evidence for the second modeling constraint.

An examination of the data shows that the variance of the squared deviations increases as the squared deviations increase. A working assumption for modeling is that the variances of centered Z_{3i}^2 and $(Z_{1i} - Z_{2i})^2$ are proportional to the square of the expectations. This working assumption is that of a constant coefficient of variation model, which is commonly used to model data with increasing variances. One argument for the constant coefficient of variation is that the deviations possess binomial distribution properties as they are constructed from proportions. The second moment of a binomial distribution is a quadratic function of the true proportion.

Initial models were fit to the squared deviations. However, due to the skewness of the data, the fitted functions were poorly estimated. The skewness of the data on the squared scale caused estimators to be determined by only a few segments. Transformations of the data were explored to find a transformation suitable for using a least squares estimator. The square root transformation decreased the effect of skewness in the data enough to make the least squares solution reasonable. The working model on the transformed scale is

$$E|Z_{3i}| = \gamma_0 + \gamma_1(0.5^{2.5} - |x_i - 0.5|^{2.5}) := g_i, \tag{34}$$

$$E|Z_{1i} - Z_{2i}| = \kappa(\gamma_0 + \gamma_1(0.5^{2.5} - |x_i - 0.5|^{2.5})) = \kappa g_i. \tag{35}$$

Since the model is not linear in coefficients, the Gauss-Newton algorithm is used to obtain the non-linear generalized least squares fit. The estimating equations were weighted by an initial estimate of g_i and κ. The 2.5 power was determined by comparing the fit and the mean squared errors across several different powers. The distributions of the absolute deviations are well approximated by the distributions of a multiples of a χ_1^2 random variables. Therefore, we compared the mean squared errors to 2, the variance of a χ_1^2 random variable. The mean squared error from the 2.5 power model fit is 2.06. Table 5 gives the estimated coefficients of the variance function and their delete-1 jackknife standard errors.

In order to estimate $\widehat{\theta}$, the ratio of the variance of error in the previous protocol to the one in the current protocol, the variance functions in (34) and (35) need to be converted back to square scale. We ratio adjust the squared fitted functions so that the average of the squared fitted functions is

Table 5: Parameter Estimates from the Final Model with 2.5 Power for West RSL Data

$\widehat{\kappa}$	$\widehat{\gamma_0}$	$\widehat{\gamma_1}$	$\widehat{\theta}$
3.55	0.00212	0.129	5.8
(0.31)	(0.00030)	(0.012)	(2.7)

the same as the average of Z_{3i}^2 and $(Z_{1i} - Z_{2i})^2$. Let

$$R_1 = \left(\sum_{i=1}^{n} \widehat{g}_i^2 \right)^{-1} \sum_{i=1}^{n} Z_{3i}^2 \tag{36}$$

and

$$R_2 = \left(\sum_{i=1}^{n} \widehat{\kappa}^2 \widehat{g}_i^2 \right)^{-1} \sum_{i=1}^{n} (Z_{1i} - Z_{2i})^2. \tag{37}$$

Estimators for the mean of the squared deviations are

$$\widehat{E}(Z_{3i}^2) = R_1 \widehat{g}_i^2 \tag{38}$$

and

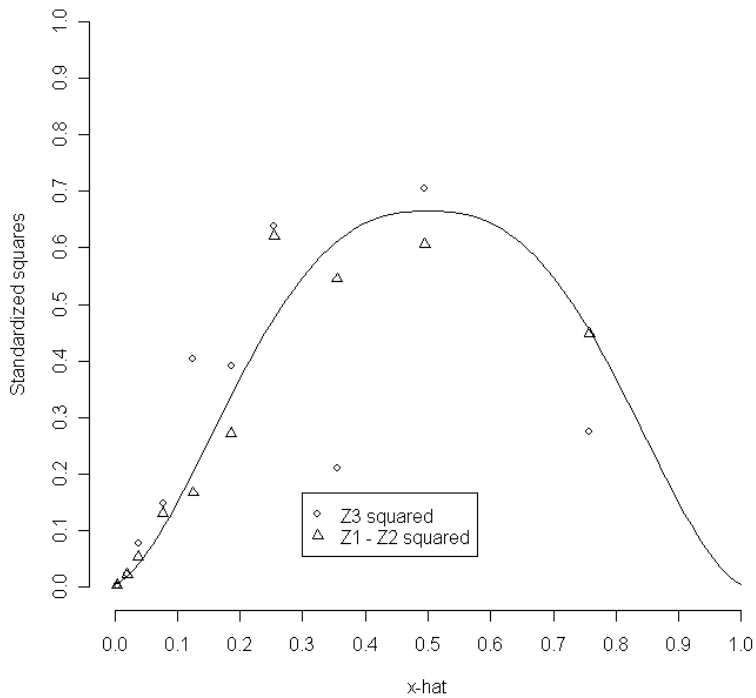
$$\widehat{E}(Z_{1i} - Z_{2i})^2 = R_2 \widehat{\kappa}^2 \widehat{g}_i^2. \tag{39}$$

An estimator of the ratio of measurement error variances is

$$\widehat{\theta} = R_1^{-1} R_2 \widehat{\kappa}^2 - 0.5. \tag{40}$$

Standard errors are computed using a delete-1 jackknife variance estimator (Table 5). The estimated θ of 5.8 is near the estimate using the average variances when fitting the mean function earlier (Table 3). The estimated variance of the ratio of variances is not well estimated in any of our results due to the skewness in the distribution, which explains the discrepancy between standard errors of θ estimators. In order to see the two fits of (38) and (39) on the square scale, we plot the fitted functions of squares and standardize them to the same scale (Figure 2). The model fits the data well on the square scale, indicating that the model furnishes adequate results.

Figure 2: Fitted Variance Function of 2.5 power with binned Z_3^2 and $(Z_1 - Z_2)^2$ means



5 Discussion

The parameters used to translate marked residences into developed area have been adjusted during data collection and analysis. The parameters used in this article for the West RSL provide encouraging results that estimators under the new and old protocol coincide within an acceptable tolerance. Some adjustments to the protocol have been attempted for the discrepancy between measurements when the proportion of developed land is very low. However, the adjustments did not solve the lack of fit problem. Further analysis related to the effect of segment size and regional differences has been considered. Similar procedures are used to examine the relationship between observations under the new and old protocols for the Central and East RSLs.

6 References

Fuller, W. A. (1987). *Measurement Error Models*. New York: John Wiley and Sons Inc.

Nusser, S. M. and J. J. Goebel (1997). The National Resources Inventory: a long-term multi-resource monitoring programme. *Environmental and Ecological Statistics* 4, 181-204.

Tin, M. (1965). Comparison of some ratio estimators. *Journal of the American Statistical Association* 60, 294-307.

Tourangeau, R. and T. W. Smith (1996). Asking sensitive questions: the impact of data collection mode, question format, and question content. *Public Opinion Quarterly* 60, 275-304.