Available online at www.sciencedirect.com

SCIENCE DIRECT®

Remote Sensing of Environment 95 (2005) 303–316

Remote Sensing of Environment

www.elsevier.com/locate/rse

Parametric (modified least squares) and non-parametric (Theil–Sen) linear regressions for predicting biophysical parameters in the presence of measurement errors

Richard Fernandes*, Sylvain G. Leblanc

Canada Centre for Remote Sensing, Environmental Monitoring Sec., Applications Division, Room 423 588 Booth Street, Ottawa, ON Canada K1A 0Y7

Received 27 February 2004; received in revised form 1 January 2005; accepted 8 January 2005

Abstract

Remote sensing often involves the estimation of in situ quantities from remote measurements. Linear regression, where there are no non-linear combinations of regressors, is a common approach to this prediction problem in the remote sensing community. A review of recent remote sensing articles using univariate linear regression indicates that in the majority of cases, ordinary least squares (OLS) linear regression has been applied, with approximately half the articles using the in situ observations as regressors and the other half using the inverse regression with remote measurements as regressors. OLS implicitly assume an underlying normal structural data model to arrive at unbiased estimates of the response. OLS regression can be a biased predictor in the presence of measurement errors when the regression problem is based on a functional rather than structural data model. Parametric (Modified Least Squares) and non-parametric (Theil-Sen) consistent predictors are given for linear regression in the presence of measurement errors together with analytical approximations of their prediction confidence intervals. Three case studies involving estimation of leaf area index from nadir reflectance estimates are used to compare these unbiased estimators with OLS linear regression. A comparison to Geometric Mean regression, a standardized version of Reduced Major Axis regression, is also performed. The Theil–Sen approach is suggested as a potential replacement of OLS for linear regression in remote sensing applications. It offers simplicity in computation, analytical estimates of confidence intervals, robustness to outliers, testable assumptions regarding residuals and requires limited a priori information regarding measurement errors.

Keywords: Regression; LAI; Biophysical parameters; Measurement errors; Reduced major axis; Theil-sen; Structural models; Functional models; Outliers

1. Introduction

1.1. Remote sensing

Quantitative remote sensing involves the prediction of in situ quantities based on remote measurements of radiation. This prediction problem relies on a model (either statistical or physically based) relating remote and in situ measurements.

1.2. Statistical models

Both the theoretical complexity in designing physically based models for radiative transfer in natural environments as well as uncertainties in specifying model parameters often lead to the use of simplified models derived to a lesser or greater degree using statistics based on paired remote and in situ measurements. Linear models relating in situ measurements to simple numerical transformations of remote measurements, together with available ancillary information, may be sufficient when the goal is prediction of in situ quantities rather than understanding radiative transfer processes. Using statistical terminology, this is a *regression* problem. The remote measurements are *regressor* variables and the in situ quantities are *response*

* Corresponding author.
E-mail address: Richard.Fernandes@CCRS.NRCan.gc.ca (R. Fernandes).

0034-4257/$ - see front matter.
doi:10.1016/j.rse.2005.01.005

variables. This nomenclature is used to distinguish the variable we wish to determine (*response*) based on the observations available (*regressors*) and does not necessarily specify a physically based relationship between these two sets of variables. The term *linear* implies that response variables are estimated as a linear combination of, in general, non-linear functions of individual regressors as long as no term in the regression equation includes more than one regressor. In this sense, knowledge of radiative transfer and measurement errors can still be used to define suitable regressors based on non-linear combinations of remote measurements. A widespread example is the use of spectral vegetation indices based on non-linear combinations of directional hemispherical reflectance in an attempt to provide regressors that may be related to vegetation structure while at the same time minimizing noise due to sensor calibration, atmospheric properties, view and illumination geometry and soil or understory reflectance.

1.3. Ordinary least squares linear regression

1.3.1. Basic definition and reference

Ordinary Least Squares (OLS) linear regression is widely used to infer linear regression model parameters in the remote sensing literature. Some studies fit model parameters with the in situ variable as a regressor and others do the reverse. We shall demonstrate that, in some cases, both of these approaches are incorrect.

OLS provides the unbiased minimum mean squared error estimate of response variables using a linear combination of regressors under a number of assumptions. The term *unbiased* implies that the prediction is equivalent to the expected value of the response even for a finite data set. This is more restrictive than the case of *consistency* where the predictor only tends to the expected value of the response as the sample size approaches infinity. While many of the OLS assumptions are rather subtle (Kendall & Stuart, 1967, Chapter 28; Sprent, 1969) there are two significant assumptions that we suggest are often not valid when calibrating remote sensing models (Cheng & Van Ness, 1999, pp. 70–71):

A1. The measurement errors of the regression model are random variables that are independent of both the regressors and each other.

A2. Data used to fit the regression line are independent and identically distributed normal random variables representative of the population over which prediction of the response variable is required.

The first assumption holds if the data are acquired without systematic measurement errors or if the second assumption holds and transformations are applied to the data to render the measurement errors uncorrelated with the observations. The second assumption is likely more restrictive since it implies that all the data are taken from the same underlying joint normal distribution. There are two problems with this latter assumption not commonly addressed in the remote sensing literature. Firstly, outliers in the sampling distribution relative to the population of interest will result in biased estimates of model parameters and hence biases in prediction of subsequent responses. Rigorous outlier detection and treatment of measurement errors in both regressors and response variables has only been performed in a few notable studies surveyed in the next section. Secondly, in many cases, data used to calibrate linear prediction models are either randomly selected from a population having a considerable range to the point that a joint normal distribution is not representative of the sampling distribution or are selected to span a range of values of regressor or response (e.g. stratified sampling). We shall provide both theoretical and empirical arguments that OLS regression is biased in these cases. Further, we shall show that there are other unbiased estimators in these cases.

1.3.2. Regression approaches that consider measurement errors

Only a few studies have considered the problem of errors in both regressors and response variables. There are two cases that are typically considered: errors due to systematic biases in measurement or theory and random errors in observations due to noise in measurement methods or collocation problems.

Systematic biases in measurements of regressors and response variables are often treated by calibration procedures. A typical example is the application of atmospheric correction to reduce temporal or spatial variability in vegetation indices (Carlson & Ripley, 1997; Song et al., 2001). Reducing the impact of location errors on regression estimators has also been treated in a rigorous fashion (Salvador, 1999). However, careful experimental design together with adequate calibration can reduce both of these error sources while still leaving a large random error component (Butson & Fernandes, 2004).

Random errors in both response and regressors have only been considered in a few articles reviewed. One approach has been to apply robust regressions that provide unbiased estimates of regression parameters in the face of either systematic or random errors (Rousseuw & Leroy, 1987). However, most of the robust estimators assume that random errors are due to measurement errors alone and are not also due to *equation errors* corresponding to a lack of fit to the chosen linear model even if there were no measurement errors. Therefore, they may not be appropriate where the amount of natural variability acceptable in observations is large relative to measurement errors and no clear cloud of outliers are present in the data. Furthermore, efficient confidence intervals of prediction are difficult to define for most robust regression approaches. One exception is the family of R-Estimators based on rank-ordered statistics (Huber, 1996). For example, Curran and Hay (1986), hereafter cited as CH86, apply Bartlett's method (Bartlett, 1949) to estimate Leaf Area Index (LAI) given the Simple Ratio (SR)

vegetation index. Here, LAI is defined as half the total green foliage surface area per unit ground area projected on the local horizontal datum (Fernandes et al., 2003). The SR (Jordan, 1969) is defined as the ratio of near-infrared to red directional hemispherical reflectance under a standardized illumination and view geometry (usually clear sky, nadir view).

CH86 describe a parametric regression (termed "least squares by error estimates", LSE) that considers measurement errors in both regressor and response variables. Davis et al. (1987) apply another parametric regression based on weighting residuals by measurement uncertainties (York, 1966). Murray (1994) used data from CH86 to show that York's method and two other parametric regressions that incorporate measurement errors (Ogren & Norton, 1992; Press et al., 1992) are equivalent to the "reduced major axis" (RMA) fit (CH86) when variances are constant with response and regressor levels. None of the fits tested by Murray (1994) distinguish between measurement error and intrinsic variability. Here, *intrinsic variability* is defined as the variation in the noise free data set that deviates from the exact linear model. Empirical calibration of remote sensing models may involve substantial intrinsic variability that result in deviations from assumption A2. As CH86 demonstrate, their LSE regression differs from the RMA regression because of the large amount of intrinsic variability in their data. Cohen et al. (2003) argue that Geometric Mean (GM) regression (they label it RMA but we use the conventional term as in Sokal & Rohlf, 1981) is suitable for calibrating remote sensing models in the absence of information on measurement errors. They state that, "Besides making no assumptions about errors in X and Y, RMA makes no assumptions about dependency."

This article augments CH86 by:

i. correcting the parametric structural regression solution documented in Curran and Hay (1986) to give the appropriate consistent parametric linear regression solution based on the modified least squares (MLS) approach;
ii. identifying consistent parametric regression solutions applicable to problems where the data are acquired due to a stratified sampling scheme or do not fit a bivariate normal distribution.
iii. introducing the Theil–Sen (Sen, 1968; Theil, 1950) regression to the remote sensing community as an extension of Bartlett's method and providing analytical prediction confidence intervals for the Theil–Sen (TS) regression;
iv. comparing the Theil–Sen robust linear regression solution to OLS, LSE and GM regression over selected data sets.

The first three objectives are addressed in a theoretical manner (Section 2) while the last objective is addressed using three case studies dealing with estimating LAI given the SR index (Section 3).

2. Theory

The mathematical basis of this section was originally drafted after ideas from Kendall and Stuart (1967) (hereafter KS67) but then revised to follow the style and updated concepts in Cheng and Van Ness (1999, hereafter CV99) based on the suggestions of one of the reviewers. The reader is encouraged to read both references since we have been brief in our treatment of the theory.

This section first defines the general problem of linear prediction that we address in this paper. The important distinction between functional and structural prediction problems as special cases of the ultra-structural prediction problem is introduced. The *ultra-structural problem* occurs, for example, when the data are sampled using a stratified random sampling scheme. For example, field sites could be pre-selected to span a range of a priori estimates of LAI based on land cover or forest survey data but the actual plots within each interval in this range were randomly selected. There are two special (extreme) cases of the ultra-structural problem: the structural regression problem and the functional regression problem. A regression method that is unbiased or consistent for the ultra-structural problem will also be unbiased or consistent for the structural and functional problems. The reverse is not necessarily true although there may be special cases, such as data that arise from a uniformly distributed population where other regression models will be consistent under stratified sampling. The *structural regression problem* corresponds to the case where data is randomly and independently sampled from a stationary underlying distribution. For example, the data could be sampled from one field with uniform growing conditions. The *functional regression problem* corresponds to the case where the data are sampled without randomization. For example, one could identify a set of target fields growing under various treatments that a priori should produce different LAI at a given sampling date. These terms are specific to linear estimation and prediction and should not be confused by other common statistical uses of the terms. For example, 'functional' is also often used to imply a functional equation that represents a physical law or a data transformation.

This section reports on consistent parametric solutions (modified least squares, MLS) to each of these three regression problems. The properties of other parametric solutions, such as GM and LSE, are also discussed in terms of consistency. Finally, a nonparametric (TS) solution to each of the three regression problems is also provided.

2.1. The general linear prediction problem

This section describes the problem of estimating an unobserved response η_0 (e.g. LAI) given both a calibration data set of noisy data $\{x,y\}$corresponding to noise free actual values $\{\xi\eta,\}$ (e.g. SR and LAI measurements) and either existing or new noisy observation x_0 (SR of a new

site) using a linear model. Using the terminology of CV99 (1999)(p. 6.) this problem corresponds to the linear *ultra-structural regression* problem with equation error.

Mathematically, the ultra-structural problem is represented as:

P1: Find $E(\eta_0|x_0,\{x,y\})$ where the underlying noise free data is linearly related together with some level of equation error, q_i, as

$$\eta_{i,j} = \beta_0 + \beta_1\xi_{i,j} + q_i \qquad i = 1,\ldots,n\ j = 1,\ldots,r; \quad (1)$$

where n corresponds to the number of levels of the regressor sampled and r corresponds to the number of samples taken for that level and data corresponding to observations $\{x,y\}$ with corresponding measurement errors $\{\delta,\varepsilon\}$:

$$x_{i,j} = \xi_{i,j} + \delta_{i,j} \quad y_{i,j} = \eta_{i,j} + \varepsilon_{i,j.} \quad (2)$$

The following assumptions are also made regarding the nature of the observations and the errors.

A3. The regressors are each assumed to be independent and identically distributed (IID) random variables with the same variance (the notation *IID(a,b)* implies an IID random variable with mean a and variance b):

$$E(\xi_{i,j}) \sim IID(\mu_i,\sigma^2). \quad (3)$$

This assumption assumes an underlying model for variability at each level of the regressor and does not necessarily imply that the actual sampling was performed with replication at each level.

A4. The equation and measurement errors are assumed to have zero mean and constant variance:

$$E(\delta_i) = E(\varepsilon_j) = E(q_j) = 0 \quad var(\delta_i) = \sigma_\delta^2$$
$$var(\varepsilon_j) = \sigma_\varepsilon^2 \quad var(q_i) = \sigma_q^2 \quad \forall i. \quad (4)$$

A5. The measurement and equation errors are assumed to be uncorrelated both between observations and between error terms.

Assumption A3 is the critical component that distinguishes the ultra-structural problem from purely functional and purely structural problems.

The functional and structural regression problems are commonly discussed in the regression literature (e.g. KS67 Chapter 28). Structural regression corresponds to the ultra-structural problem with assumption Assumption A3 replaced with the assumption that:

A6. the regressors are IID random variables and independent (not just uncorrelated) with the errors, i.e.

$$E(\xi_{i,j}) \sim IID(\mu,\sigma^2) \quad (5)$$

The requirement for independence of errors is difficult to support unless the underlying population is normally distributed. In this case, the only additional assumption is that the errors are uncorrelated with the observations.

Functional regression corresponds to the ultra-structural problem with Assumption A3 replaced by the assumption that:

A7. the regressors are unknown constants ξ_i.

There is no clear rule to determine when a problem transitions from purely functional to ultra-structural to purely structural (see for example CV99, p. 88; Carroll et al., 1995, p. 6). In general, the functional and ultra-structural problems arise because each new observation brings with it an additional unknown (μ_i). This is in contrast to the structural problem where the sampling distribution of new observations are completely specified by a few parameters. Fortunately, as we shall see in the next section consistent solutions to the functional and ultra-structural regression problems are identical and they also converge, although sometimes in the large sample limit, to the unbiased estimate when faced with structural regression problems.

2.2. Parametric solutions to the general linear regression problem

For brevity, we discuss the known solutions and their assumptions with reference to equations provided in Appendix A. The reader is encouraged to read the indicated references for further details.

2.2.1. Structural regression

Under the structural regression assumptions (A1–A6) and the additional assumption that,

A8. The regressors $\{\xi\}$ are normally distributed,

An unbiased maximum likelihood solution of the structural regression problem corresponds to the OLS estimators for the regression of y on x. (CV99, p.12). In other words, OLS gives the correct answer in this case. However, contrary to what is somewhat common practice, this predictor should not be directly inverted if one wishes to predict a later value of ξ given an observed y. Rather, the OLS estimator for the regression of x on y should be used. If the underlying distribution is not normal there is no unbiased or consistent linear predictor although the OLS solution is still the minimum mean squared error solution (CV99, p. 70). If the data diverge from normality it is also possible that the problem is really an ultra-structural or functional regression problem or there are substantial outliers in the data or both.

2.2.2. Functional and ultra-structural regression

The OLS predictor is biased for the functional and ultra-structural models (CV99, p. 71) since the assumption of a bivariate normal data distribution is no longer guaranteed. As argued by CV99 (p. 80) the functional

and ultra-structural problems cannot be directly solved without additional assumptions since each regressor observation, $x_{i,j}$, brings with it an additional underlying unknown μ_i. Consistent maximum likelihood solutions do not exist for this problem (Table 2.2 in CV99). In fact, the commonly used GM regression corresponds to an inconsistent and hence biased solution (CV99, p. 20–22,) and does not correspond to the maxima of the log likelihood function that describes the functional regression problem. CV99 (p. 44) states:

"the GM estimate is neither consistent nor does it possess desirable statistical properties. We do not recommend its use for the ME [Measurement Error] model."

Additonally, orthogonal regression solutions of which the LSE method is a simple case are also inconsistent since they do not include the equation error term (CV99, p. 89).

Fortunately, consistent solutions to the ultra-structural equation error regression problem can be found by optimizing criteria other than maximum likelihood. These solutions will necessarily also satisfy the structural and functional equation error regression problems as they are special cases of the ultra-structural problem. CV99 discuss three general approaches to the solution: higher order moments, modified least squares (MLS), and grouping methods (including TS).

If the underlying regressors are not normal (A8 does not hold) then one can use the method of higher order moments to arrive at unbiased predictors (CV99, p. 120). However, this method requires estimation of third and fourth order moments and thus requires both large sample sets for precise estimates of these moments while also being highly susceptible to biases due to outliers. For example, CV99 (p. 128) show a typical problem where the method of higher order moments results in slope estimators with approximately 44 times the variance than alternative approaches.

The MLS method provides a minimum root mean squared error (RMSE) unbiased predictor of the true, measurement error free, response variable (CV99, p. 84) given a priori knowledge of either both measurement errors ($\sigma_\delta^2, \sigma_\varepsilon^2$), the measurement error of the response (σ_δ^2) or the intercept (β_0). Appendix A summarizes the MLS predictors and their confidence intervals. It should be noted that there is no consistent (and unbiased) estimator if only σ_ε^2 is known since the equation error must still be separated from the error in measuring the response variable. All of the MLS regressions are based on estimates of sample variances and covariances and are thus subject to bias in the presence of even a single outlier. The use of robust estimates of these variances and covariances is discussed in (CV99, p. 221) but the properties of the predictor based on these robust estimators are currently not well known. MLS provides the most efficient prediction confidence intervals in a minimum mean squared error sense of any linear predictor conditional on the known a priori data. MLS is equivalent to the unbiased OLS method in the limit when the measurement error of the regressor tends to zero. Hence, the MLS prediction also tends to the maximum likelihood regression solution in this case.

In many cases a priori knowledge is lacking regarding measurement errors or intercepts required for MLS. An alternate constraint on the ultra-structural regression problem is specification of the rank-order of groups of observations such as the TS approach. This approach is feasible if an upper bound on the measurement error can be estimated such that only data that are pair wise separated by more than this bound are ranked. For example, if we can assume that the measurement error in LAI is 1 unit and 1 unit in the simple ratio index then we could state that a site with an LAI of 4 and simple ratio of 6 has a higher rank than a site with an LAI of 2 and a simple ratio of 4. Theil (1950) developed a grouping estimator for the predictor slope, β_1, based on the assumption that the ranks are known. The slope estimator corresponds to the median of the pair wise slopes of all data points that meet the ranking criteria. The intercept is then given by using the same form as the OLS estimator but with the TS slope. Sen (1968) later extended this estimator, now called Theil–Sen (TS) regression, to deal with the presence of ties in the list of pairwise slopes (see Appendix A).

Sen (1968) proved that TS regression is unbiased for the functional regression model while CV99 (Theorem 4.2) show that it is consistent for the purely structural regression model. Peng and Xueqin Wang (2004) recently proved TS regression is also consistent for the ultra-structural model as long as the residuals between the response and prediction form a continuous distribution. This requirement is usually satisfied when measurements are performed by the same instrument and the data follow an underlying model that is close to linear. Perhaps more importantly, the TS estimator is robust to up to 28.9% outliers in the data (Theil, 1950) so that large deviations from the equation error model or mistakes in the assumed ranking of data can be tolerated to some extent.

One concern with a grouping method is the potential increase in the width of prediction confidence intervals since optimality criteria such as maximum likelihood or minimum mean square error are not directly applied. Peng and Xueqin Wang (2004) report on the asymptotic relative efficiency (A.R.E.) of the slope estimated by the Theil–Sen method for cases where the slope estimate based on unbiased minimum mean square error solutions are known. The A.R.E. is the ratio of the variance of the minimum RMSE estimator of the regression slope to the variance of the Theil–Sen regression slope. In summary, the A.R.E is always better than 86.4%, approaches 95.5% for normally distributed data (OLS), and will exceed 100% (i.e. TS is more efficient) for data that has large tails such as double exponential or Cauchy distributions. Appendix A provides analytical formulae for prediction confidence intervals of the TS method.

3. Application

3.1. Problem of LAI mapping and why empirical methods are useful

LAI is a quantitative measure of foliage quantity and status (e.g. Sellers et al., 1994.) and is a fundamental input to numerical models of carbon, water and energy cycles (Monteith & Unsworth, 1990; Running & Coughlan, 1988). There have been numerous previous studies relating LAI to broad-band spectral measurements from satellites (e.g. Chen, 1996; Hall et al., 1995). Physically based models relating canopy reflectance to LAI have had some success and are operationally used for mapping global LAI patterns (Myneni et al., 2002). However, nadir spectral measurements tend to saturate at moderate levels of LAI so that retrieval uncertainty for physical models can be large unless *a priori* information is applied (Knyazikhin et al., 1998). In contrast, empirical models based on regressions between in situ LAI and remote satellite measurements allow LAI estimates to be constrained by the sampling distribution (Chen & Cihlar, 1995). As such, ecophysiological constraints on LAI are incorporated to make predictions past the point of saturation in nadir reflectance.

The purpose of this section is to evaluate the impact of departures from assumptions in linear regression models for empirical estimation of LAI given simple spectral vegetation indices.

3.2. Summary of three data sets

Three data sets are used in our study. All three data sets consist of ordered pairs of observed LAI and SR. We chose data sets using the SR, since it is available from a wide variety of sensors and does not require site-specific

Table 1
Descriptions of data sets used in case studies

Data set identification			
Data set	Data Set 1	Data Set 2	Data Set 3
Description	Synthetic, variable measurement and intrinsic error	Low measurement error; local	High measurement error; national
Reference	Sellers 1985	Curran & Hay 1986	Chen et al., 2002
Cover	Typical agricultural crop	Semi-natural grassland	Typical conifer stands
Location	Not applicable	Latkhill Dale, Derbyshire, UK	Canada south of treeline
Sampling domain			
Spatial extent	Not applicable	5km^2	26x10^{6}km^2
Temporal extent	Not applicable	September 08–15, 1982	In situ: summer 1998 Remote: summer 1998
Sampling design			
Stratification	*N*=51 Functional, *N*=74 structural, Single population	*N*=26, likely random.	Stratified by biome type and across 6 Landsat TM frames. Sampled from local permanent sample plots. Agriculture only in Ottawa region.
Separation	Not applicable	Within single pasture.	99 stands>1km
In situ sampling			
Method	Specified	Specified destructive	Optical gap fraction
Spatial Support	Not applicable	2m×2m	100 m square
Measurement error	Various treatments	0.17	20%
Sample mean	Not applicable	1.57	4.36
Sample stdev.	Not applicable	0.68	2.58
Sample max	3.4	2.97	10.32
Sample min	0	0=0.	0.38
Remote sampling			
Method	2-stream radiative transfer model	Milton radiometer	Landsat TM 5+6S
Spatial support	infinite	~1m–2m	~100m square
Measurement error	Various treatments	0.10	12%
Sample mean	3	3.25	7.61
Sample Stdev.	Not applicable	0.90	3.59
Sample max	5	4.87	15.52
Sample min	1	1.45	2.45

Data Set 3 includes three columns when reporting in situ and remote summary statistics corresponding, from left to right, to confiner, deciduous and agriculture sub-sets.

parameters. Nevertheless, our results could be extended to other vegetation indices. For brevity, we summarise the data sets in Table 1. This section discusses aspects of each data set of interest to the objectives of our study.

3.2.1. Simulated problem

A synthetic data set was used to produce a set of simulated noisy measurements using the same underlying physical process relating LAI and SR but within two different sampling schemes to represent both structural and functional problems. The data sets was generated based on a two stream radiative transfer model (Sellers, 1985) to reproduce the curve of SR versus LAI, shown in Fig. 1, for a crop canopy of spherically distributed leaves. A soil single scattering albedo of 0.2 in both visible and NIR and a leaf single scattering albedo of 0.2 and 0.95 in visible and NIR, respectively, were used. These parameters produced an asymptotic SR of 6.5 at infinite LAI. Only data with the range of SR 1 and SR 5 were used to minimize the impact of the asymptotic non-linearity on the comparisons.

For the structural problem, a total of 74 noise free points were sampled based on a normal SR distribution with mean 3 and variance 1.5. For the functional regression problem, a total of 51 noise free points were sampled at a regular SR interval of 0.1 from SR 1 to SR 6. Fig. 1 shows the noise free sample for both problems super-imposed on the underlying noise-free radiative transfer model curve. Additive zero mean, constant variance normally distributed noise was applied to each data set. Noise levels corresponding to variances of 0.2, 0.4, 0.6, 0.8, 1.0 and 1.2 units in LAI and SR were applied. Each noisy data set was replicated 100 times using a randomly selected noise process with the same underlying noise free data. OLS (with both SR and LAI as regressors), GM, MLS and TS regression methods were applied to the noise free and noisy data sets.

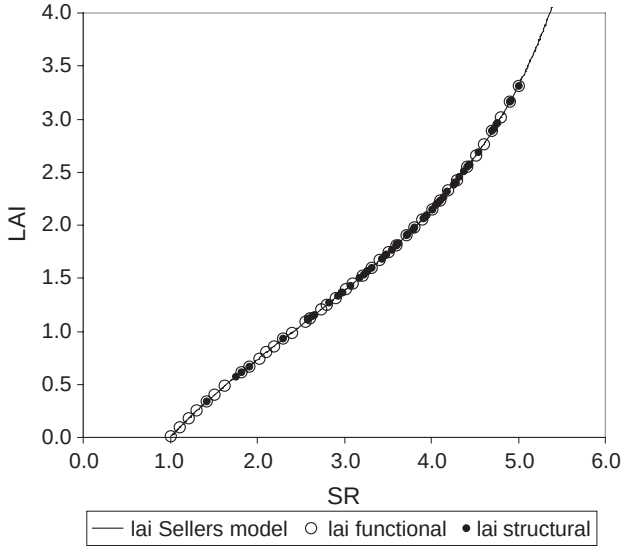

Fig. 1. Modelled relationships between LAI and SR for a noise free synthetic data set. The modelled relationship was generated using the 2-stream turbid media radiative transfer model of Sellers (1985) for a typical broadleaf crop canopy. Only values in the linear range between SR of 1 and 5 were used. The sample sets for functional and structural regression problems are indicated as hollow and filled circles, respectively.

3.2.2. Structural regression problem (Data Set 2)

CH86 provide a set of 26 points of LAI measured by destructive sampling within a single field together with the SR derived from spatially co-incident tripod mounted radiometer measurements. Each plot was a 2 m square of semi-natural grassland. No information was provided on view and illumination geometry or calibration to apparent reflectance at surface. However, it is likely atmospheric correction errors are negligible and the target is likely close to Lambertian so illumination geometry should not be a significant factor as nadir view was used. Additive measurement errors of σ_δ=0.10 and σ_ϵ=0.17 are specified in CH86. These errors are less than 3% of the observed range in SR and 6% of the observed range in LAI. The problem can therefore be assumed to be a structural regression problem with a relatively normal distribution of regressors and low measurement errors.

3.2.3. Ultra-structural regression problem (Data Set 3)

This data set, taken from, Chen et al. (2002), hereafter referred to as C02, consists of in-situ LAI measurements based on optical gap fraction methods and SR derived from Landsat 5 Thematic Mapper atmospherically corrected reflectance taken across Canada during mid-growing season conditions. Seven World Reference System 2 frames, corresponding to Landsat 5 Thematic Mapper image regions, were sampled across Canada. These data form the basis of linear regression algorithms used to generate initial Canada-wide LAI estimates (C02). Butson and Fernandes (2004) found measurement errors on the order of 10% for the SR and Chen and Cihlar, (1995) suggest one standard deviation errors on the order of at least 20% in LAI. The in-situ LAI sites within and between Landsat scenes were stratified as much as possible to span the observed range in LAI and land cover type with a few (~3) replicates performed at each category in the stratification. This data set therefore corresponds to an ultra-structural regression problem with likely a greater affinity to functional regression due to the strong stratification of plots along perceived vegetation density gradients.

3.3. Results

3.3.1. Synthetic data set (Data Set 1)

We first focus on the case where the response, LAI, has no measurement error while the regressor, SR, has high measurement error to illustrate the pitfalls of OLS. Fig. 2 illustrates sample model fits for Data Set 1 with high levels of measurement error ($\sigma^2{}_\delta$=1.2 units) in SR and no measurement error in LAI. For the structural regression

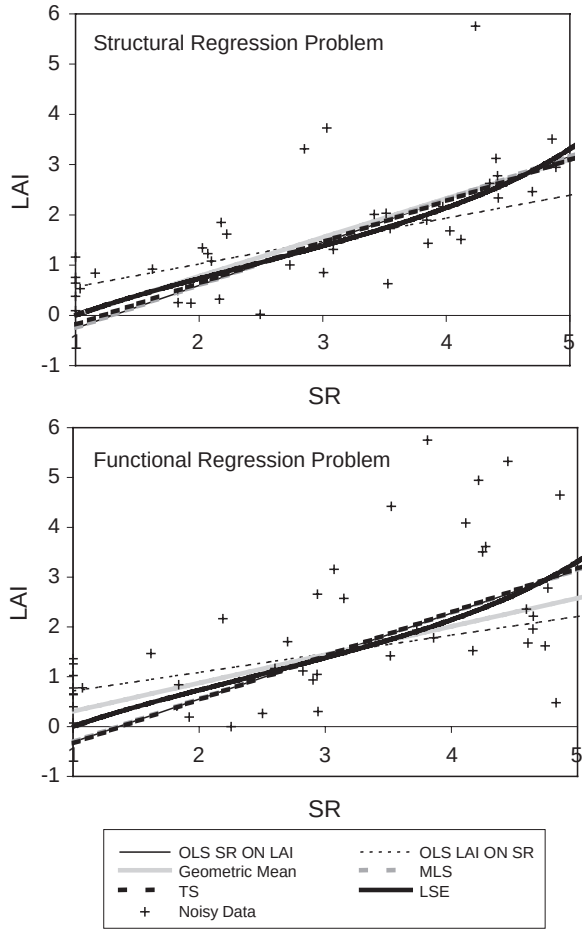

Fig. 2. Linear regression fits to both structural and functional model synthetic datasets of LAI versus SR with zero measurement error in LAI and a 1.2 variance measurement in measurement error in SR. The Sellers (1985) radiative transfer model was used to define the underlying population of LAI and SR values.

problem, the theoretically unbiased linear solution corresponds to predict LAI from the OLS regression of SR on LAI. As expected, both the MLS and TS regressions closely match this unbiased estimate. The GM regression is also very close to the unbiased OLS solution with a slight overestimate at mid-level SR values. In contrast, the OLS regression of LAI on SR differs substantially for all the other solutions. This example clearly indicates the danger of blindly using OLS regressions. This example also demonstrates a case where the GM regression is comparable to the unbiased approached. However, as shown in the functional regression example in the second panel of Fig. 2, the GM Regression can be biased when the sampling approach differs from the normal structural model. In this functional counterpart to the structural problem just discussed, the GM regression lies approximately half way in between the two OLS regressions while the MLS and TS functional regressions closely match the OLS of SR on LAI.

Fig. 3 summarizes RMSE results for 100 repetitions of regression fits for both structural and functional regressions at a range of noise levels. The MLS regression is mathematically equivalent to OLS regressions in the case of noise in one variable only and to LSE in the case of equal noise in both variables so it is not separately labeled in the charts provided. Generally speaking all methods perform similarly at low noise levels (1 sigma ≤ 0.2). The difference in RMSE increases between the two OLS methods as increasing additive noise is applied to only one of LAI or SR. This is expected since these methods have opposite hypotheses regarding measurement errors. The OLS of SR on LAI assumes no measurement errors in LAI and vice versa for the OLS of LAI on SR. With the structural regression trials, the OLS methods are the minimum mean squared error unbiased estimates when their assumptions hold. This theoretical aspect is reflected in that they empirically produce the lowest RMSE for the structural problems where only one of LAI or SR has measurement error. Additionally, the TS method closely approximates the performance of the minimum RMSE OLS method in accordance with the theoretical property that the TS solution approaches an OLS solution for structural regression problems. In contrast, GM regression only exhibits a consistent minimum RMSE with equal additive measurement errors to both LAI and SR. The fact that the GM approach actually attains a minimum RMSE is partially a co-incidence due to the fact that the ratio of variances of noise free LAI and SR (1.18) is approximately the same as the ratio of variance of noise processes (set at 1). GM approach is equivalent to the MLS approach when the ratio of sample variances of the data equals the ratio of the measurement error variances (Draper & Smith, 1981, p 91). We expect that GM will not perform as well with data sets where these two ratios are not close. At noise levels of 0.6 or larger, TS consistently placed second behind GM regression in terms of minimum RMSE for the structural problem with equal noise in LAI and SR. The LSE method usually fell behind the TS and optimum OLS approaches in terms of minimum RMSE for estimating the true, measurement error free, response variable.

We expected the OLS methods to fare worse with the functional regression trials since, according to theory, they are no longer unbiased predictors of the underlying true LAI. In fact, the best OLS approach for the structural regression case of error in only one variable was also the best approach for the corresponding functional problem. This may in part be due to the fact that the sampling distribution used for our synthetic functional regression problem could also be parameterized as a two parameter uniform distribution and hence could be interpreted as a non-normal structural regression problem. As mentioned previously the OLS methods also minimize RMSE for non-normal structural regressions. In hindsight, a multimodal sampling scheme could have been applied to demonstrate a very extreme functional regression problem. However, it should also be emphasized that the good performance of the OLS methods only applies to the synthetic trials with error in one variable. The error in both variable trials clearly

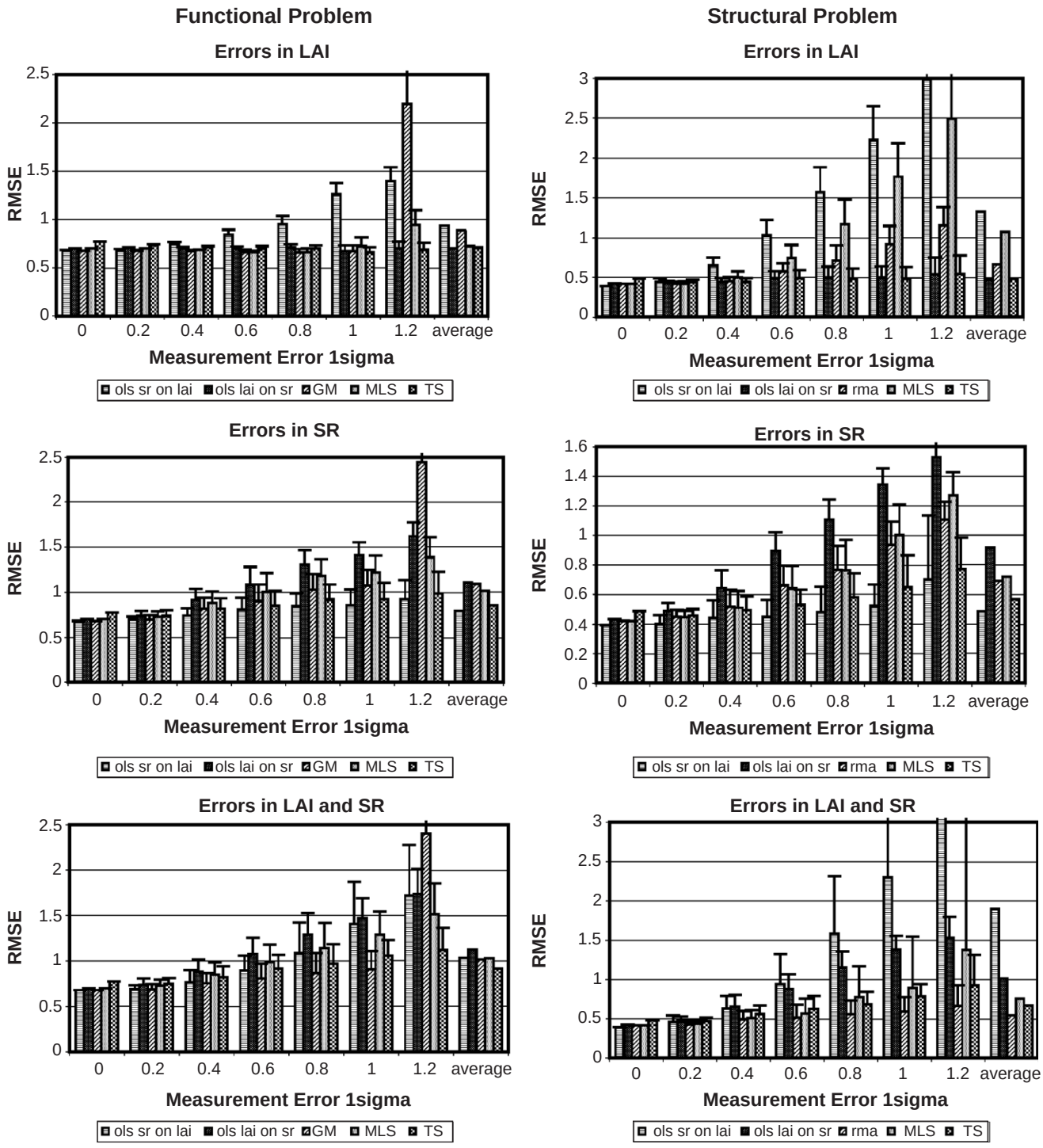

Fig. 3. Summary charts of average (bars) and standard deviation (vertical whiskers) of root mean square error of different regression fits to various synthetic datasets generated using the Sellers (1985) model with additive Gaussian measurement errors. Each bar corresponds to 100 trials. The charts are separated according to both underlying noise free data (either functional or structural problems) and presence of measurement error in LAI only, SR only or equal amounts in LAI and SR. The average performance for each regression method over all levels of measurement error is also reported in each chart.

shows that OLS is not optimal in the common case where there is measurement error in both variables.

The TS approach typically followed the minimum RMSE OLS approach for trials with error in only one variable and was usually close to the GM regression in terms of the minimum RMSE method for errors in both LAI and SR trials. The exception being with the highest levels of noise where the TS approach was superior to all others. In this

case, the noise process could be interpreted as producing outliers from the assumed underlying linear model. The TS robustness property likely resulted in its superior performance. The MLS estimator typically fell in the middle of the other methods in terms of minimum RMSE.

3.3.2. Additive measurement errors (Data Set 2)

Fig. 4 compares regression estimators assuming the additive noise model of CH86. A natural log transform was applied to the observed SR regressors to improve the linearity of the relationship. Fig. 4 clearly demonstrates that OLS will give different results depending on the specification of the response and regressor. The relative differences between the two OLS estimates are on average 15% with differences over 50% for LAI below 2. These differences can have detrimental impacts on large area LAI estimates since they apply to the mean estimated LAI for a given SR and therefore cannot be expected to cancel out due to intrinsic variability (i.e. variance in observations not related to measurement errors). In this case, the measurement errors in SR are approximately three times smaller than those in LAI. Hence, the OLS of LAI on SR should provide the minimum RMSE estimate assuming this regression problem is structural and will also be unbiased to the extent the noise free data is normally distributed.

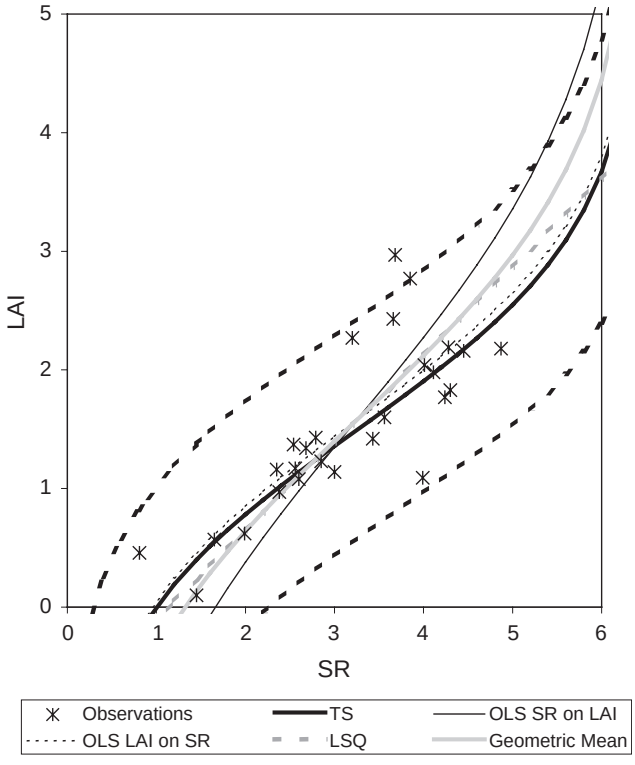

Fig. 4. Comparison of TS, OLS of LAI on SR, OLS of SR on LAI and LSE regressions for estimation of LAI given SR with the CH86 data set. The fitted MLS regression was virtually identical to the OLS of SR on LAI and is not shown on the current figure. Prediction confidence intervals at the 67.5% level are included for the TS regression.

The TS regression closely matches the OLS of LAI on SR as expected theoretically and as seen in the structural regression trials in Data Set 1. The TS prediction confidence intervals are quite large due to the relatively small range of the data in the SR axis relative to the scatter of data in the LAI axis. Both the LSE and GM regression lines fall close to the mid-point of the optimal OLS of LAI on SR and the worst case error of the OLS of SR on LAI. The MLS structural regression solution was not shown since it essentially coincided with the OLS of LAI on SR due to the substantially larger LAI measurement errors.

3.3.3. Large area domain with substantial variability in LAI and SR (Data Set 3)

In this example we are faced with the difficult task of using pooled data that is exposed to substantial intrinsic variability and measurement error. As such, one should expect large confidence intervals. The goal therefore is to develop unbiased regression fits with the hope that random errors will cancel over large samples. Fig. 5 compares GM, MLS and TS regressions (with 67.5% confidence intervals) applied to SR and LAI values, transformed after applying a logarithmic transform on SR and a 4th root transform on LAI. These transformations were selected as they provided homeoscedastic regression residuals. OLS regressions of SR on LAI using the linear model of Baret and Guyot (1991), as published in C02 is included for comparison.

The MLS and TS lines and their confidence intervals agree very closely in all three cases. This result is in part due to the large amount of random variation induced by measurement error and in part due to the mixing of measurements from different stands to increase intrinsic variability. The large confidence intervals reflect this high level of variability and measurement error relative to the range of the data. Furthermore, the large cloud of points at a SR of 7 for conifers is not characteristic of the rest of the data and suggests exceptional measurement errors (possibly outliers). Examination of the data found that they were acquired using insufficient spatial sampling (Fernandes et al., 2003). Clearly, in developing regression to predict LAI one should sample over the entire range of possible observed LAI. As with the current data set, stratified sampling across a range of LAI is often implemented to reduce field measurement demands. This sort of stratification may reduce prediction errors but one should always map the residuals in LAI estimates to determine the extent to which they contribute to bias errors due versus spatially random errors.

The OLS fit based on the physically based equation taken from C02 differs substantially in comparison to MLS and TS regressions. For example, the fit from C02 reaches an LAI of 0 at a SR of 2.71 while the other regressions predict an LAI of 1. C02 argue that this is due to the SR ratio of background without overstory foliage. It is evident

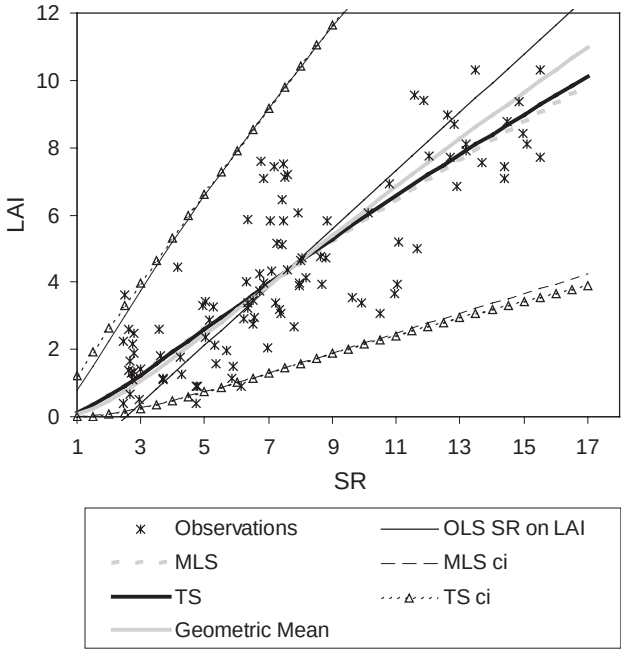

Fig. 5. Comparison of MLS and TS regressions with GM and the functional OLS regression reported in C02 over conifer sites sampled across Canada. Prediction confidence intervals at the 67.5% level are included for parametric and TS regressions.

from the data that SR may vary substantially at low LAI and that specifying a single background SR may bias the functional regression fit at other levels of LAI. Examination of Fig. 5 shows a difference in estimated LAI of 2 units at a SR of 15 between OLS and either MLS or TS regressions. Given that the MLS and TS regressions are based on a large number of high LAI values it may be possible that this difference is due to biases in the OLS regression. There is insufficient data to test this hypothesis and in the case of this data set the difference is small in relative terms.

The GM regression line also differs from the parametric and TS regression lines. The OLS of LAI on SR (not shown) is close to the parametric fit given that LAI measurement errors were estimated to be twice those of the SR measurement errors. This suggests that, as with data set 2, GM regression selects a predictor close to the bisector of the two OLS line fits. Given the lack of theoretical arguments to support the consistency of the GM regression and the strong arguments in support of the consistency of the parametric and TS regressions we suggest that the GM regression is likely biased in this example.

4. Conclusions

Linear regression is a widespread tool for prediction of in situ quantities given remote sensing measurements. The regression problem is in general an ultra-structural error in equation problem given that data used to produce the predictor are usually sampled with some sort of stratification and that exact linearity is not expected a priori. Two extreme cases of this ultra-structural model are the functional and structural models. The structural model assumes that all of the remote sensing measurements are drawn from a population that can be statistically described using a few parameters (e.g. random sampling of similar land surfaces). The functional model, in contrast, assumes each remotely sensed observation has a unique expected value that cannot be inferred given knowledge of other observations without some sort of linear or non-linear model (e.g. stratified sampling). Thus each new observation brings with it a new unknown to the regression problem.

Theoretical results based largely on KS67 and CV99 were provided for the minimum mean square error regression for both structural and functional models. For structural problems, the unbiased parametric predictor corresponds to the OLS with the quantity being predicted as the response variable. For functional problems, the MLS method provided an unbiased regression but requires additional information related to either measurement errors or the intercept. This TS regression is also unbiased for functional models and converges on the unbiased minimum mean squared error OLS solution for structural models. Significantly the TS predictor requires only knowledge of an upper bound on measurement errors to allow an accurate ranking of the data. Additionally, the TS method is robust to 29.8% outliers in the data. This robustness also applies to errors in ranking of data so that the knowledge of measurement errors need only be correct for ~70% of the data. Approximate confidence intervals of prediction of the TS and structural approaches were given. The TS estimator allows for relatively robust regressions while providing analytical confidence intervals. TS may result in slightly lower asymptotic confidence intervals compared to MLS. If this is a concern then one could should both the MLS and TS regression when measurement errors are known, as differences between these estimators can be indicative of violations in the more restrictive MLS regression assumptions.

Three test data sets of measured LAI and SR were used to explore the performance of the parametric and TS approaches as well as the GM and LSE methods advocated elsewhere. A synthetic data set showed that the TS approach was usually the minimum RMSE method or ranked close to the minimum method. The OLS approach was the RMSE method only with errors in one rather than both variables. A second data set corresponding to a structural regression problem again showed that the parametric and TS approaches were similar to the optimal OLS approach. However, the GM and LSE approaches fell in between the two OLS regression lines and were likely biased predictors. The third data set showed that both

parametric and TS regressions tend to arrive at similar solutions for ultra-structural problems. This is in contrast to the OLS, GM and LSE methods that, theoretically, should be biased.

Although remote sensing technology and theory have advanced significantly since KS67 and CH86, the majority of the studies we surveyed have been using potentially biased OLS linear regression estimators. Both OLS and GM (including RMA as a special case) are sub-optimal estimators of linear structural relationships and should not be used without very strong justification. Rather, the Theil–Sen estimator should be considered for univariate linear regression. MLS regression is a useful back-up solution when data are known to be outlier free and a priori information on measurement errors or the intercept are known. At a minimum, researchers considering linear regression should identify the potential uncertainty in predictors based on the extreme cases of OLS. Finally, we caution that we have only addressed some of the issues related to regression in this paper. Our study did not consider the spatial pattern of predictions and their residuals or the appropriateness of the sample population for the domain over which the regressions were applied. We expect that users of regression in remote sensing will continue to pay close attention to these aspects.

Acknowledgements

The authors acknowledge Andrew Davidson, Bert Guindon, Ian Olthof and Darren Puilot and two anonymous reviewers for their constructive comments.

Appendix A. Solutions to the general linear regression problem

This appendix provides formulae for unbiased solutions to the general linear regression problem with both measurement errors and equation errors as defined in Eqs. (1–5) and their associated assumptions. The tilde overbar is used to denote estimators based on sampled data.

A1. Structural regression problem

An unbiased maximum likelihood solution only exists if the noise free regressors are normally distributed and the measurement errors in response and regressors are independent. In this case the solution is simply the appropriate OLS predictor:

$$\hat{\eta}_0 = E(\eta_0|x_0,\{x,y\}) = \tilde{\alpha}_0 + \tilde{\alpha}_1 x_0 \tag{AI.1}$$

$$\tilde{\alpha}_0 = \bar{y} - \tilde{\alpha}_1\bar{x} \quad \tilde{\alpha}_1 = s_{xy}/s_{xx}. \tag{AI.2}$$

Here, s_{xy} is the sample covariance and s_{xx} the sample variance of the regressor. The confidence intervals of the OLS predictor are well known (KS67, p. 363).

We are not aware of a consistent linear regression solution for non-normal structural regression problems. However, the appropriate OLS solution is the minimum RMSE solution for non-normal structural regression problems. We note also that the ultra-structural and functional regressions cited in the next section are also consistent predictors of the normal structural regression problem and tend to the minimum RMSE estimator for non-normal structural regression.

A2. Functional regression problem

When the assumptions required for a functional regression problem hold, an unbiased predictor is given by:

$$\hat{\eta}_0 = E(\eta_0|x_0,\{x,y\}) = \widetilde{\beta}_0 + \widetilde{\beta}_1 x_0. \tag{AI.3}$$

Here $\widetilde{\beta}_0$ and $\widetilde{\beta}_1$are estimates of the parameters of the linear functional relationship between ξ and η. Unbiased estimates for these parameters can only be arrived at given additional information.

A2.1. Parametric functional parameter estimates

Table A1 summarizes results from CV99 (pp. 54–55) giving parametric minimum mean square unbiased estimates of the parameters in Eq. (AI.3) for both the equation error and no equation error cases. The slope estimate depends on the type of a priori knowledge. These known parameters must be specified a priori and strictly speaking not estimated; although in practice they are usually estimated using ancillary data. In some cases the problem is unidentified in that there is no consistent or unbiased estimator that can be arrived at without additional knowledge.

Table A1
Linear functional relation parameter estimates

Known parameter	No Equation error	Equation error
λ	$\widetilde{\beta}_0,\widetilde{\beta}_{1,1},\widetilde{\sigma}^2_{\delta,1}$	Unidentified
σ^2_{ε} , σ^2_{δ}	$\widetilde{\beta}_0,\widetilde{\beta}_{1,1}$	$\widetilde{\beta}_0,\widetilde{\beta}_{1,2},\widetilde{\sigma}^2_{\eta}$
σ^2_{δ}	$\widetilde{\beta}_0,\widetilde{\beta}_{1,2},\widetilde{\sigma}^2_{\varepsilon,1}$	$\widetilde{\beta}_0,\widetilde{\beta}_{1,2},\widetilde{\sigma}^2_{\varepsilon,1}$
σ^2_{ε}	$\widetilde{\beta}_0,\widetilde{\beta}_{1,3},\widetilde{\sigma}^2_{\delta,2}$	Unidentified
β_0	$\widetilde{\beta}_{1,4},\widetilde{\sigma}^2_{\delta,3},\widetilde{\sigma}^2_{\delta,2}$	$\widetilde{\beta}_0,\widetilde{\beta}_{1,4},\widetilde{\sigma}^2_{\varepsilon,2}$

The various parameter estimates for the slope and intercept are defined below:

$$\widetilde{\beta}_0 = \bar{y} - \widetilde{\beta}_{1,.}\bar{x} \tag{AI.4}$$

$$\widetilde{\beta}_{1,1} = \frac{s_{yy} - \lambda s_{xx} + \sqrt{(s_{yy}-\lambda s_{xx})^2 + 4\lambda s^2_{xy}}}{2s_{xy}} \tag{AI.5}$$

$$\widetilde{\beta}_{1,2} = s_{xy}/(s_{xx} - \sigma^2_{\delta}) \tag{AI.6}$$

$$\widetilde{\beta}_{1,3} = (s_{yy} - \sigma^2_{\varepsilon})/s_{xy} \tag{AI.7}$$

$$\tilde{\beta}_{1,4} = (\bar{y} - \beta_0)/\bar{x} \tag{AI.8}$$

Here s_{yy} is the sample variance of the response. The measurement error variance estimates indicated in Table A1 are given in CV99 (p. 55) and not included here for brevity. Parametric regressions based on estimates in Table A1 and Equation (AI.3) are also consistent predictors if the data turns out to be from a normal structural model as they coincide asymptotically with the unbiased OLS predictor (CV99, p. 71).

KS67 (p. 389) gives the confidence interval of $\tilde{\beta}_{1,1}$ estimated from Eq. (AI.5):

$$\tan\left(\tan^{-1}\tilde{\beta}_{1,1} \pm \sin^{-1}\left[2t\left\{\frac{s_{xx}s_{yy} - s_{xy}^2}{(n-2)\left[\left(s_{xx}-s_{yy}\right) + 4s_{xy}^2\right]}\right\}^{0.5}\right]\right) \tag{AI.9}$$

where t is the Student t statistic and n is the number of samples in the data set. This confidence interval is exact for structural relationships between response and regressor. A more general version exists (CV99, p. 64) that applies to structural and functional regression problems but our experience is that it produces similar intervals to Eq. (AI.9) for data with typical measurement error levels. Different, asymptotic confidence intervals must be applied for other cases where only one measurement error variance or the slope is given:

$$\pm z_{\gamma/2}\,\tilde{\sigma}_{\beta}^2 n^{-1/2} \tag{AI.10}$$

where $z_{\gamma/2}$ is the 100(1-γ/2) percentile of the standard normal distribution and the variance of the slope is estimated using

$$\tilde{\sigma}_{\beta}^2 = \frac{\tilde{\sigma}_{\delta}^2}{\tilde{\sigma}^2}\left(\frac{\tilde{\sigma}_{\varepsilon}^2}{\tilde{\sigma}_{\delta}^2} + \tilde{\beta}_1^2\right)\tilde{\beta}_{1,\cdot} + \frac{\tilde{\sigma}_{\varepsilon}^2\tilde{\sigma}_{\delta}^2}{\tilde{\sigma}^4_{\cdot}} \tag{AI.11}$$

This estimator is only useful when the noise to signal ratio given by $\frac{\tilde{\sigma}_{\varepsilon}^2\tilde{\sigma}_{\delta}^2}{\tilde{\sigma}^4}$ is small (CV99, p. 74).

The prediction confidence interval for subsequent responses is then derived by shifting the data so the origin coincides with $\bar{x}$:

$$\hat{\eta}_0\begin{cases} -ts_{yy}\dfrac{(n+1)}{n} - x^2\hat{C}_{\alpha_1}^{-} \\ +ts_{yy}\dfrac{(n+1)}{n} - x^2\hat{C}_{\alpha_1}^{+} \end{cases} \tag{AI.12}$$

where $\hat{C}_{\alpha_1}^{-}$ and $\hat{C}_{\alpha_1}^{+}$ are respectively the positive and negative confidence intervals estimated by Eq. (AI.9) or (AI.10) for the value of t specified.

A2.2. Rank based functional parameter estimates

A rank-based unbiased functional regression slope estimate was given by Theil (1950) as the median of pairwise slopes of the observations. Sen (1968) refined this estimate to deal with the case of ties in the ranking:

$$\tilde{\beta}_1 = median\{\tilde{B}\} \tag{AI.13}$$

where

$$\tilde{B} = median\left\{b_{ij}\middle| b_{ij} = \frac{y_j - y_i}{x_j - x_i}, x_i \neq x_j, 1 \le i < j < n\right\} \tag{AI.14}$$

and

$$\tilde{\beta}_0 = \bar{y} - \tilde{\beta}_1\bar{x} \tag{AI.15}$$

Ideally, to estimate the confidence interval of the predictor one would use the joint confidence interval of the slope and intercept. A joint confidence interval for the slope and intercept has not been derived (CV99, p. 120). The upper and lower confidence interval for the slope parameter β_1 from the Theil–Sen estimator follows by application of the Mann–Kendall (Kendall & Gibbons, 1990) statistic:

$$\tilde{C}_{\tilde{\beta}_1}^{+} = \underline{\tilde{\beta}}_1\left(0.5\left\{N' + \frac{t}{18}\left[n(n-1)(2n+5) + \sum_{p=1}^{q} T_p(T_p-1)(2T_p+5)\right]^{0.5}\right\}\right)$$

$$\tilde{C}_{\tilde{\beta}_1}^{-} = \underline{\tilde{\beta}}_1\left(0.5\left\{N' - \frac{t}{18}\left[n(n-1)(2n+5) + \sum_{p=1}^{q} T_p(T_p-1)(2T_p+5)\right]^{0.5}\right\} + 1\right) \tag{AI.15}$$

where $\underline{\tilde{\beta}}_1$ is the vector of estimated slopes sorted in ascending order; $\overline{N}$' represents the index of the median value of $\underline{\tilde{\beta}}_1$; $\boldsymbol{q}$ is the number of subsets of tied values in $\underline{\tilde{\beta}}_1$; $\boldsymbol{T_p}$ is the cardinality of the $\boldsymbol{p}$th sub-set of tied values; and rounding to the nearest integer rank is applied.

This estimate assumes that the values of noise free regressors, ξ, are sufficiently separated in comparison to their measurement error σ_{δ}^2 that the x_i's used in the median are in the same order as their counterpart, unobserved ξ_i's.

A2.3. Ultrasctuctural regression problem

Both the parametric and TS regressions for the functional model are also unbiased for the ultra-structural model and tend asymptotically to the OLS regression solution when provided with data taken from a structural regression problem (CV99, p. 55 and pp. 117–118). This property of the parametric and TS regression is especially attractive since it does not require a priori knowledge of the type of regression problem to arrive at unbiased estimators. The price paid is in increased confidence intervals (i.e. decreased efficiency) when using these estimators instead of the OLS estimator for a true normal structural model.

There has been some criticism of the TS method when faced with structural regression problems due to the

assumption that the data are correctly ranked. CV99 (p. 118) argue that the data will not be correctly ranked when there is a large number of samples taken from a normal structural model when the measurement error σ_δ^2 is not small. However, there is no a priori reason to use all pairwise combinations of observations when constructing the TS estimator. Censoring combinations that fall close to each other will reduce the efficiency of the estimator but as long as the censoring is based on a rule independent of the data (e.g. enforcing a minimum separation between x_i's used in pair wise slopes) the TS will still remain unbiased. On the other hand, the TS method requires only knowledge of an upper bound on the measurement error rather than the exact ratio or magnitude of measurement error typically needed for the parametric functional methods.

References

Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APARA assessment. *Remote Sensing of Environment*, *35*, 161–173.

Bartlett, M. S. (1949). Fitting a straight line when both variables are subject to error. *Biometrics*, *5*, 207–211.

Butson, C., & Fernandes, R. A. (2004). A consistency analysis of surface reflectance and leaf area index retrieval of overlapping clear sky Landsat ETM+_ imagery. *Remote Sensing of Environment*, on-line proof, 369–380.

Carlson, T. N., & Ripley, D. A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, *62*, 241–252.

Carroll, R. J., Ruppert, D., & Stefanski, L. A. (1995). *Measurement Error in Nonlinear Models*. London: Chapman and Hall.

Chen, J. M. (1996). Evaluation of vegetation indices and a modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, *22*, 229–242.

Chen, J. M., & Cihlar, J. (1995). Algorithms for retrieving LAI and FPAR of boreal conifer forests using Landsat TM images. *Digest of 17Th Canadian Symposium on Remote Sensing*, Saskatoon, Canada, 235–242.

Chen, J. M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S. G., White, H. P., et al. (2002). Validation of Canada-wide leaf area index maps using ground measurements and high and moderate resolution satellite imagery. *Remote Sensing of the Environment*, *80*, 165–184.

Cheng, C. -L., & Van Ness, J. W. (1999). *Statistical Regression with Measurement Error*. New York: Oxford University Press Inc.

Cohen, W. B., Maiersperger, T. K., Stith, T., Gower, S. T., & Turner, D. P. (2003). An improved strategy for regression of biophysical variables and Landsat ETM+data. *Remote Sensing of Environment*, *84*, 561–571.

Curran, P. J., & Hay, A. M. (1986). The importance of measurement error for certain procedures in remote sensing at optical wavelengths. *Photogrammetric Engineering and Remote Sensing*, *52*, 229–242.

Davis, P. A., Berlin, G. L. P. S., & Chavez, J. R. (1987). Discrimination of altered basaltic rocks in the southwest United States by analysis of Landsat thematic mapper data. *Photogrammetric Engineering and Remote Sensing*, *53*, 45–55.

Draper, N. R., & Smith, H. (1981). *Applied Regression Analysis*, 2nd ed. New York: Wiley.

Fernandes, R., Butson, C.G, Leblanc, S., & Latifovic, R. (2003). Landsat-5 TM and Landsat-7 ETM+based accuracy assessment of leaf area index products for Canada derived from SPOT-4 VEGETATION data. *Canadian Journal of Remote Sensing*, *29*, 241–258.

Hall, F. G., Townshend, J. R., & Engmann, E. T. (1995). Status of remote sensing algorithms for estimation of land surface state parameters. *Remote Sensing of Environment*, *51*, 138–156.

Huber, P. J. (1996). *Robust Statistical Procedures*, 2nd ed. CBMS-NSF Regional Conference Series in Applied Mathematics, 68. New York: Society for Industrial and Applied Mathematics (SIAM).

Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, *50*, 663–666.

Kendall, M. G., & Gibbons, J. D. (1990). *Rank correlation methods* (5th ed.). London: Arnold. pp. 248.

Kendall, M. G., & Stuart, A. S. (1967). *Advanced theory of statistics*, *vol. 2*. London: Charles Griffin and Company.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Diner, D. J., & Running, S. W. (1998). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, *103*, 32257–32275.

Monteith, J. L., & Unsworth, M. H. (1990). *Principles of Environmental Physics*. London: Edward Arnold.

Murray, R. (1994) Weighted least-squares curve fitting with errors in all variables. *1994 ASPRS/ACSM Annual Convention and Exposition ASPRS Technical Papers* (pp. 460–470). Reno, Nevada.

Myneni, R. B., Knyazikhin, Y., & Privette, J. (2002). Global products of leaf area index and fraction absorbed PAR from year one of MODIS data, submitted to. *Remote Sensing of Environment*, *83*(1–2), 214–231.

Ogren, P. J., & Norton, J. R. (1992). Applying a simple least squares algorithm to data with uncertainties in both variables. *Journal of Chemical Education*, *69*(4), A130–A131.

Peng, H., & Xueqin Wang, X. (2004). *Asymptotics of the Theil–Sen Estimator in Simple Linear Regression Model with a Non-random Covariate*. Journal of American Statistical Association in review.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. R. (1992). *Numerical Recipes: The Art of Scientific Computing*, Second ed. New York: Cambridge Press.

Rousseuw, P. J., & Leroy, A. M. (1987). *Robust Regression and Outlier Detection*. New York: Wiley.

Running, S. W., & Coughlan, J. C. (1988). A general model of forest ecosystem processes for regional applications: I. Hydrologic balance, canopy gas exchange and primary production processes. *Ecological Modelling*, *42*, 125–154.

Salvador, R. (1999). A parametric model for estimating relations between unprecisely located field measurements and remotely sensed data. *Remote Sensing of Environment*, *67*, 99–107.

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, *6*, 1335–1372.

Sellers, P. J., Los, S. O., Tucker, C. J., Justice, C. O., Dazlich, D. A., Collatz, G. J., et al. (1994). A global 1° by 1° NDVI data set for climatic studies, pt.2: The adjustment of the NDVI and generation of global fields of terrestrial biophysical parameters. *International Journal of Remote Sensing*, *17*, 3519–3546.

Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association*, *63*, 1379–1389.

Sokal, R. R., & Rohlf, F. J. (1981). *Biometry*, 2nd ed. Birkenhead: W.H. Freeman and Company.

Song, C., Woodcock, C. E., Seto, K. C., Pax Lenney, M., & Macomber, S. A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, *75*, 230–244.

Sprent, P. (1969). *Models in Regression*. Birkenhead: Willmer Brothers Limited.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Nederlandse Akademie Wetenchappen Series A.*, *53*, 386–392.

York, D. (1966). Least-squares fitting of a straight line. *Canadian Journal of Physics*, *44*, 1079–1086.