ANL/EVS-16/3

Exploring the Application of Optical Remote Sensing as a Method to Estimate the Depth of Backwater Nursery Habitats of the Colorado Pikeminnow

Environmental Science Division

About Argonne National Laboratory
Argonne is a U.S. Department of Energy laboratory managed by UChicago Argonne, LLC under contract DE-AC02-06CH11357. The Laboratory's main facility is outside Chicago, at 9700 South Cass Avenue, Argonne, Illinois 60439. For information about Argonne and its pioneering science and technology programs, see www.anl.gov.

DOCUMENT AVAILABILITY

Online Access: U.S. Department of Energy (DOE) reports produced after 1991 and a growing number of pre-1991 documents are available free via DOE's SciTech Connect (http://www.osti.gov/scitech/).

Reports not in digital format may be purchased by the public from the National Technical Information Service (NTIS):
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Road
Alexandria, VA 22312
www.ntis.gov
Phone: (800) 553-NTIS (6847) or (703) 605-6000
Fax: (703) 605-6900
Email: **orders@ntis.gov**

Reports not in digital format are available to DOE and DOE contractors from the Office of Scientific and Technical Information (OSTI):
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831-0062
www.osti.gov
Phone: (865) 576-8401
Fax: (865) 576-5728
Email: **reports@osti.gov**

Disclaimer
This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor UChicago Argonne, LLC, nor any of their employees or officers, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of document authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof, Argonne National Laboratory, or UChicago Argonne, LLC.

ANL/EVS-16/3

Exploring the Application of Optical Remote Sensing as a Method to Estimate the Depth of Backwater Nursery Habitats of the Colorado Pikeminnow

prepared by

Yuki Hamada and Kirk E. LaGory
Environmental Science Division, Argonne National Laboratory

prepared for

Western Area Power Administration
Colorado River Storage Project Management Center
Salt Lake City, Utah

February 2016

CONTENTS

EXECUTIVE SUMMARY v

ACKNOWLEDGMENTS vi

NOTATION vii

1 INTRODUCTION 1

2 METHODS 2

2.1 Study Area 2
2.2 Modeling Backwater Depth 3
2.2.1 Data Used for Modeling Backwater Depth 3
2.2.2 Backwater Depth Modeling Approach 3
2.2.2.1 Preparing Image Data 5
2.2.2.2 Generating Predictor Variables and Calibration Dataset 6
2.2.3 Modeling for Water Depth Estimation 6

3 RESULTS 9

3.1 Model Development 9
3.1.1 Linear Regression Model 9
3.1.2 Multiple Regression Model 10
3.1.3 Partial Least Squares Regression Model 11
3.2 Qualitative Analysis of Modeled Water Depth 11
3.3 Quantitative Analysis of Modeled Water Depth 17

4 DISCUSSION 19

5 REFERENCES 22

FIGURES

1 Study Reach for Previous Backwater Surveys Located in Ouray National Wildlife Refuge Project Area within the Green River Subbasin 4

2 Empirical Reflectance Calibration Functions and Pixel Values Before and After Conversion. 5

3 Scatter Plots of Predicted Water Depth Using the Regression Models against Reference Water Depth... 9

Figures (Cont.)

4 Enlarged True-Color Image of Backwater 2 from National Agriculture Imagery Program, Reference Water Depth Map, and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares Regression Models. 13

5 Enlarged True-Color Image of Backwater 6 from National Agriculture Imagery Program, Reference Water Depth Map, and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares Regression Models. 14

6 Enlarged True-Color Image of Backwater 7a from National Agriculture Imagery Program, Reference Water Depth Map, and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares Regression Models 15

7 Enlarged True-Color Image of Backwater 13 from National Agriculture Imagery Program, Reference Water Depth Map, and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares Regression Models 16

8 Scatter Plots of Predicted Backwater Depth Using the Regression Models and Reference Water Depth.... 20

9 Classification Accuracy of Shallow and Deep Backwater Pixels for All Three Backswaters 21

TABLES

1 Remote-Sensing Variables Evaluated for Use as Predictors of Backwater Depth 7

2 Correlation Between the Reference Water Depth and Predictor Variables.. 8

3 Statistics of Regression Models Using Calibration Data.. 10

4 Intercept and Coefficients of Partial Least Squares Regression Model..... 12

5 Summary of Model Performance.. 19

EXECUTIVE SUMMARY

Low-velocity channel-margin habitats serve as important nursery habitats for the endangered Colorado pikeminnow (*Ptychocheilus lucius*) in the middle Green River between Jensen and Ouray, Utah. These habitats, known as backwaters, are associated with emergent sand bars, and are shaped and reformed annually by peak flows. A recent synthesis of information on backwater characteristics and the factors that influence inter-annual variability in those backwaters (Grippo et al. 2015) evaluated detailed survey information collected annually since 2003 on a relatively small sample of backwaters, as well as reach-wide evaluations of backwater surface area from aerial and satellite imagery. An approach is needed to bridge the gap between these detailed surveys, which estimate surface area, volume, and depth, and the reach-wide assessment of surface area to enable an assessment of the amount of habitat that meets the minimum depth requirements for suitable habitat.

Optical remote sensing is recognized as an effective tool for gathering river parameters across an extensive reach in a spatially continuous manner. Many existing studies tested remote sensing approaches for clear, gravel-bed rivers; however, these techniques may not be directly applicable for estimating depth of backwater habitats in relatively turbid rivers with sandy substrate such as the Green River. Argonne National Laboratory (Argonne) examined the effectiveness of optical remote sensing for estimating backwater depth using National Agriculture Imagery Program (NAIP) imagery that was collected in July 2006 for the Jensen-Ouray reach of the Green River. Three regression models—linear, multiple, and partial least square (PLS) regression models—were developed.

Among the three models, the multiple and PLS regression models showed moderately high correlation with the reference depth (R^2 = 0.669 and 0.689, respectively). The maps generated using the models share overall similarity in spatial patterns of relative depth (e.g., deep, moderate, and shallow) and correspond well to the reference maps. The model validation indicated that the PLS regression model had the most unbiased estimate of backwater depth, while the linear regression model has considerable bias across the backwater habitats. The multiple and PLS regression models appeared equally effective for estimating turbid water depth when pooling all backwaters. However, the PLS regression model had more unbiased and consistent depth estimates across backwaters than the multiple regression model.

While absolute depth estimates may be somewhat uncertain for backwaters in our study area that are deeper than approximately 40 cm, it would be possible to obtain reasonable estimates of the amount of habitat that exceed 30 cm in depth. Backwater habitats with these minimum depths are considered suitable for young-of-the-year Colorado pikeminnow (Muth et al. 2000), and are surveyed annually by the Upper Colorado River Endangered Fish Recovery Program to determine annual production of young. We consider our PLS regression approach supported by field survey data to be a viable option for developing cost-effective, reach-wide assessments of backwater habitat characteristics. The effect of interannual variability in suspended sediment concentrations on model performance should be evaluated to determine how the model could be adjusted among years to provide accurate results.

ACKNOWLEDGMENTS

We thank several Argonne National Laboratory (Argonne) staff members for their contributions to this report. Nick Haas developed the PLS regression model; Cory Weber developed software to generate digital elevation models of surveyed backwaters; and Lee Walston provided surface water elevation information that enabled determination of backwater depth on the imagery acquisition date. Jerry Wilhite of Western Area Power Administration and Mark Grippo of Argonne provided important review comments on an earlier draft. The work presented in this report was funded by the Western Area Power Administration under Contract No. DE-AC02-06CH11357.

NOTATION

B	blue spectral band
bw	backwater
cfs	cubic feet per second; 1 cfs = 0.028317 m^3/s
cm	centimeter(s)
DEM	digital elevation model
df	degree of freedom
DN	digital number
F	F-statistic
G	green spectral band
km	kilometer(s)
ln	natural log
log	log (base 10)
m	meter(s)
MAE	mean absolute error
ME	mean error
NAIP	National Agriculture Imaging Program
ND	normalized difference
ONWR	Ouray National Wildlife Refuge
p	statistical significance
PLS	partial least squares
r	Pearson correlation coefficient
R	red spectral band
R^2	coefficient of determination
SE	standard error
Sq	square
USGS	U.S. Geological Survey

EXPLORING THE APPLICATION OF OPTICAL REMOTE SENSING AS A METHOD TO ESTIMATE THE DEPTH OF BACKWATER NURSERY HABITATS OF THE COLORADO PIKEMINNOW

Prepared by

Yuki Hamada and Kirk E. LaGory
Environmental Science Division
Argonne National Laboratory

1 INTRODUCTION

Low-velocity channel-margin habitats serve as important nursery habitats for the endangered Colorado pikeminnow (*Ptychocheilus lucius*) in the middle Green River between Jensen and Ouray, Utah. These habitats, known as backwaters, are associated with emergent sand bars, and are shaped and reformed annually by peak flows. A recent synthesis of information on backwater characteristics and the factors that influence inter-annual variability in those backwaters (Grippo et al. 2015) evaluated detailed survey information collected annually since 2003 on a relatively small sample of backwaters, as well as reach-wide evaluations of backwater surface area from aerial and satellite imagery. An approach is needed to bridge the gap between these detailed surveys, which estimate surface area, volume, and depth, and the reach-wide assessment of surface area to enable an assessment of the amount of habitat that meets the minimum depth requirements for suitable habitat.

From 2003 to 2013, Argonne National Laboratory (Argonne) conducted backwater surveys for selected backwaters in the Ouray National Wildlife Refuge reach of the Green River to measure topographic features of backwater habitats during the base flow period (Grippo et al. 2015). The characteristics (location, surface area, volume, and depth) of backwaters surveyed were not consistent across years because of the dynamics of sandbars and associated backwater habitats. A total of 51 backwater habitats that represented 13 backwater locations were surveyed during the course of the 11 years. The data were used to quantify relationships between base flows and backwater habitat characteristics including surface area, volume, and depth (Grippo et al. 2015). Although surveys provide accurate topographic measurements, the process is costly, time consuming, and feasible for only a limited number of backwater habitats. These shortcomings of surveying impose a challenge to a comprehensive understanding of backwater habitat conditions for larger river reaches and multiple years.

Remote sensing offers simultaneous, spatially comprehensive data collection and potential for systematic, semiautomatic data processing over a large areal extent. It is a non-invasive data collection method that is also suitable for repeated surveys (Legleiter 2013). Optical remote sensing is perhaps the only method for estimating parameters over a large channel extent in a spatially contiguous manner (Marcus and Fonstad 2008). Sunlight in the optical spectral range that reaches the water surface is either reflected (or diffused), transmitted, or absorbed at the surface. The transmitted light attenuates as it travels through the water column

according to a specific function with depth. In theory, therefore, water depth at a given location can be estimated based on how much light is attenuated at that specific location of the river. Detailed theoretical background is provided by Legleiter et al. (2009).

Despite a long history of remote sensing research for river bathymetry, the application of remote sensing to estimating water depth has only become active over the last ten years. Legleiter et al. (2004) conducted a feasibility study for mapping channel bathymetry of a stream in Yellowstone National Park using optical imagery. They found that the natural log of the green and red spectral band ratio (ln[G/R]) was strongly correlated with the water depth (R^2 = 0.79). Legleiter et al. (2009) developed a systematic selection of optimal spectral bands for estimating water depth—optical band ratio analysis (OBRA)—and confirmed their earlier conclusion (Legleiter et al. 2004) that a linear regression model that used ln(G/R) is the optimal model for estimating water depth (R^2 = 0.79–0.98). Legleiter (2013) applied the technique for the National Agriculture Imagery Program (NAIP) image captured over the Laramie and Snake Rivers in Wyoming. The accuracy of water depth estimate was slightly lower than his earlier study (R^2 = 0.64). All studies above were conducted in clear-flowing, gravel-bed rivers that are very different from our study reach.

For turbid rivers, Legleiter and other (2011) found that an optical remote sensing method can reliably estimate water depth up to approximately 50 cm, and the estimates are subject to a significant degree of uncertainty if concentration of suspended sediment is greater than 60 mg/L, based on their field spectroscopy study. Javernick et al. (2014) attempted to map bathymetry of shallow braided rivers with substantial sediment load in the Ahuriri River, New Zealand using optical remote sensing in conjunction with photogrammetry-based 3D modeling. The study determined that the standard error in water surface elevation and river bed elevation were 11–28 cm and 25–41 cm, respectively. Although it would provide a cost-effective means for mapping and monitoring backwater habitats, true applicability of optical remote sensing for estimating water depth of backwater habitats in turbid rivers is currently unknown because of the limited number of relevant studies. Therefore, we explored the effectiveness of optical imagery for estimating backwater depth in the Green River using publicly available high resolution, multispectral NAIP imagery.

2 METHODS

2.1 STUDY AREA

We estimated water depth for backwaters located within the 177-km study reach between the Utah Route 149 bridge near Jensen, Utah, and the Utah Route 88 bridge at Ouray, Utah, where previous backwater surveys were conducted (Grippo et al. 2015). The reach has a substrate composed primarily of sand and numerous emergent sand bars and associated backwater habitats. For estimating water depth using remote sensing, four backwater habitats were selected for ground-truthing from those surveyed by Argonne since 2003 within a 15-km reach of the Green River within the Ouray National Wildlife Refuge Project (Figure 1).

2.2 MODELING BACKWATERH DEPTH

2.2.1 Data Used for Modeling Backwater Depth

The study required an aerial image and an estimate of bathymetry at the flow on the date the image was collected. A seamless NAIP image mosaic collected on July 14, 2006 was used to estimate water depth of the backwater habitats. The image mosaic consisted of three visible spectral bands (i.e., blue, green, and red) at a 1-m spatial resolution in an 8-bit radiometric data range. Although images for a region are typically collected from multiple dates, all frames that cover our study reach were collected on a single date (i.e., July 14). This single-date image mosaic made it possible for us to apply a single flow level for the day to estimate water depth. The image frames were color-balanced and georeferenced by the vendor, which resulted in a seamless single image available for our study. Within the seamless NAIP image mosaic, an area corresponding to our study reach and its vicinity was used for analysis.

Bathymetry of the backwater habitats was generated as part of our backwater habitat surveys (Grippo et al. 2015). Backwaters 2, 6, 7a, and 13 (Figure 1) were surveyed on September 26–28, 2006 using survey-grade equipment (e.g., Topcon GS GPS 34 Unit; Leica TC805 Ultra Total Station with +/- 1 cm horizontal and vertical accuracy). For each backwater habitat, a digital elevation model (DEM) was developed by interpolating elevation from the survey points using a natural nearest neighbor interpolation technique. This DEM was used to generate the reference water depths used in model calibration and validation. Flow on the day of the 2006 NAIP image collection was determined as 2,090 cfs based on mean daily flow at the USGS Jensen gauge station (USGS 09261000).

2.2.2 Backwater Depth Modeling Approach

Backwater depth modeling consisted of three steps: (1) image data preparation, which standardized the data values, (2) generation of predictor variables and a calibration dataset, which were derived from the 2006 NAIP image, and (3) regression model development, which used three types of regression analyses—linear, multiple, and partial least squares (PLS) regression analyses. Image preparation and transformation were performed using the ERDAS Imagine software (http://www.hexagongeospatial.com/products/producer-suite/erdas-imagine), calibration data generation was performed using the ArcGIS software (http://www.esri.com/software/arcgis), and statistical analysis and model development were performed using the R statistical language (https://www.r-project.org).

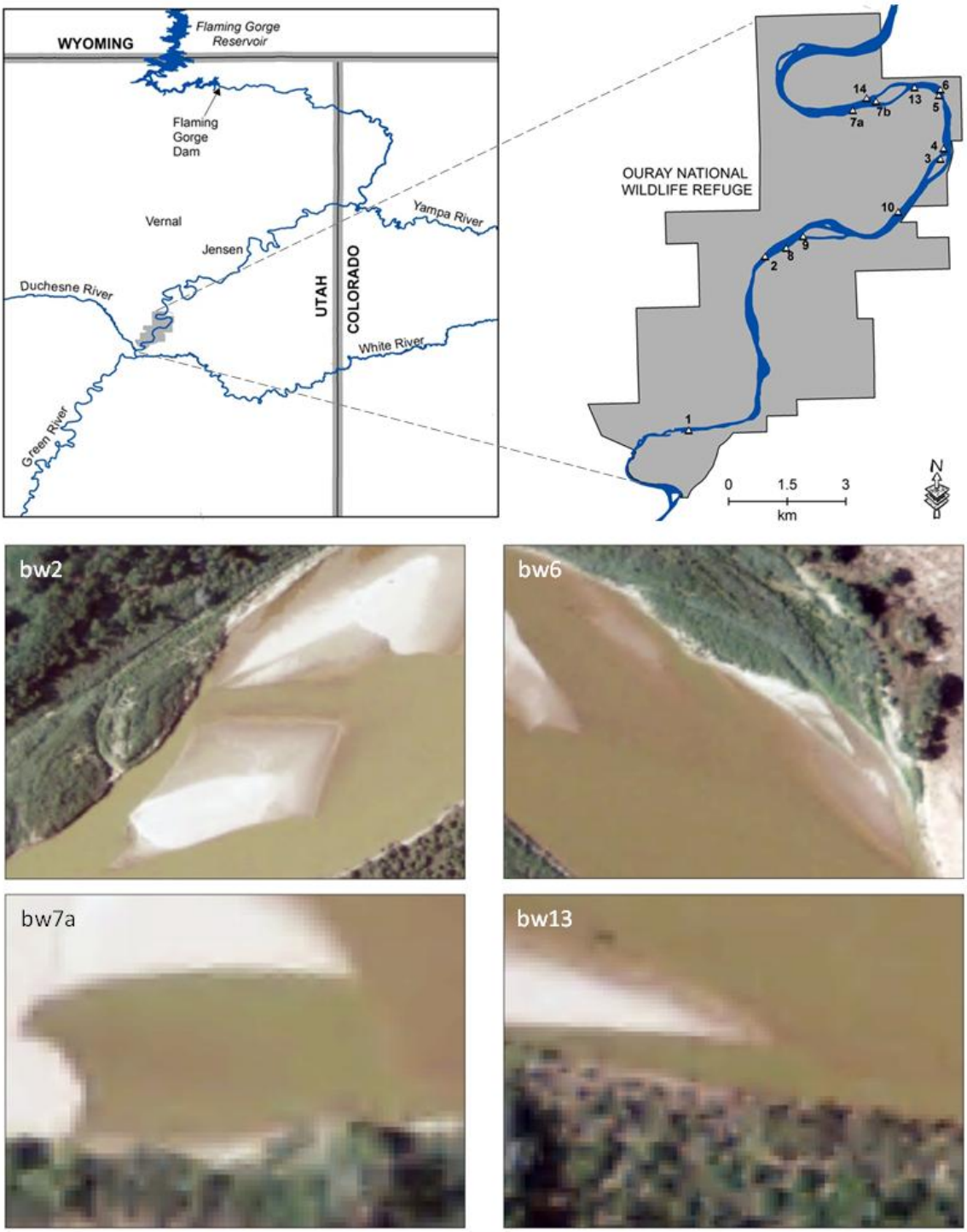

FIGURE 1 Study Reach for Previous Backwater Surveys Located in Ouray National Wildlife Refuge Project Area within the Green River Subbasin (top). Backwater habitats (bw) 2, 6, 7a, and 13 were used to estimate water depth using remote sensing. Aerial photos of the four backwater habitats are also shown.

2.2.2.1 Preparing Image Data

The original data values (known as digital numbers or DNs) of the NAIP image were converted to reflectance values using an empirical line method adopted from Smith and Milton (1999). This method is commonly applied when field-based reflectance data and detailed atmospheric parameters at the time of image collection are not available. We first located the darkest (i.e., shaded water) and brightest (i.e., rooftops) features in the image using a principal component analysis technique. Within those features, 36 to 60 pixels having extreme values were selected, and their pixel values were extracted. Those values for the darkest and brightest pixels were assumed to represent reflectance values of 1% and 99% across the spectral bands, respectively. Based on this assumption, we developed a linear function for each spectral band to covert the DNs to reflectance values in the image (Figure 2, left). The reflectance values that were obtained using the functions show consistency across the spectral band such that the darkest pixels have reflectance values of approximately 1% and the brightest pixels have reflectance values of approximately 99% (Figure 2, right).

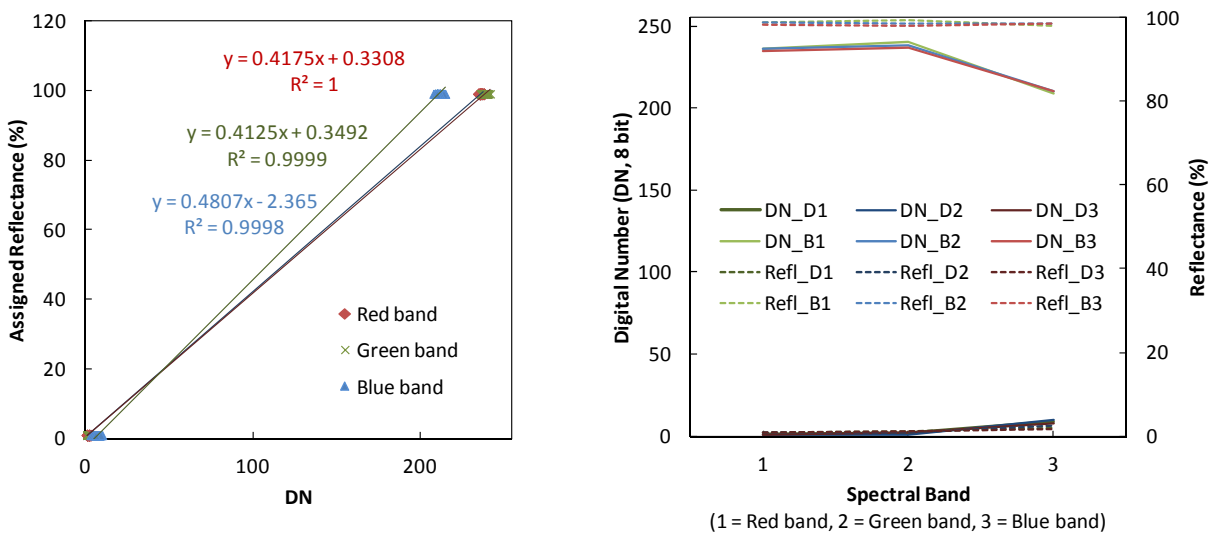

FIGURE 2 Empirical Reflectance Calibration Functions (Left) and Pixel Values Before and After Conversion (Right). D = dark, B = bright, Refl = reflectance.

In general, an optical image represents how much sunlight was reflected by the surface at the time of image collection. For a river, the image indicates the amount of light scattered on the surface, attenuated through water column, and absorbed by substrates, which provide useful information for estimating depth of water and suspended sediment.

The mean daily flow at the USGS Jensen gage station on the date of the 2006 NAIP image collection (July 14), was 2,090 cfs. The reference water depth map was generated from our DEM of each backwater using the estimated water surface elevation at a flow of 2,090 cfs. As stated earlier, the DEMs were generated from our 2006 field survey data of subject backwater habitats. The water surface elevation was subtracted from the DEM to generate the reference map representing water depth for each backwater habitat.

2.2.2.2 Generating Predictor Variables and Calibration Dataset

To estimate water depth of backwater habitats using optical remote sensing, nine types of predictor variables were tested using the NAIP image (Table 1). In addition to reflectance of the blue, green, and red bands, square and log terms were calculated for each band. Products and ratios were computed for all possible band combinations to account for interactions between bands. Simple ratios and normalized difference ratios, which are known to compensate for spectral noise associated with incoming light variability, were also computed for all possible band combinations. Natural log of simple band ratios was included because the transformation characterizes exponential attenuation of light in water column (Legleiter et al 2004, 2009).

The generation of the calibration dataset consisted of four steps: (1) sampling, (2) outlier analysis, (3) Pearson correlation analysis, and (4) collinearity analysis. We randomly selected 400 points consisting of 100 points for each of the four backwater habitats. These points were selected only from locations that are expected to be submerged at the flow of 2,090 cfs. For the 400 points, we extracted values of a total of 33 variables to generate a calibration dataset to model water depth. All calibration samples were examined using an outlier test and QQplot. Five samples that equate to 1.25% of the calibration data were determined as outliers and eliminated from the calibration set without replacement. Correlation between the reference water depth and the predictor variables in the calibration set were computed (Table 2), and predictors with $r < |0.3|$ were eliminated from the calibration set. A total of 20 predictors were carried forward at this step for further consideration.

Lastly, collinearity of the predictor variables within the calibration set was examined. Because of the highly correlated nature of spectral bands of the optical imagery, we applied a very generous criterion for assessment. When multiple predictors had perfect or extremely strong collinearity ($r > 0.90$), those that have lower correlation with the reference water depth were excluded from the calibration set. Based on the criteria, five variables (B+G, B+R, G+R, G-B, and R-G) were excluded from the calibration data. A total of 15 predictors were in the final calibration (Table 2, italics), and they were carried forward for regression modeling approaches described below.

2.2.3 Modeling for Water Depth Estimation

To estimate backwater depth, linear, multiple, and PLS regression models were applied. A linear regression model using a natural log of the green-red band ratio (ln[G/R]) was first developed for estimating water depth of backwater habitats. This predictor was found to be an effective predictor for estimating water depth for clear flowing rivers with gravel substrates (Legleiter et al. 2004, 2009). Although our study backwaters do not have ideal properties for application of a linear regression model, we developed such a model to provide a reference for comparison of the performance of multiple and PLS regression models.

TABLE 1 Remote-Sensing Variables Evaluated for Use as Predictors of Backwater Depth

Category	Remote-Sensing Variable			Description
Spectral Band	Blue (B)	Green(G)	Red(R)	Reflectance values of the NAIP 2006 imagery. Input for the rest of the variables.
Sum	B+G	B+R	G+R	Sum of reflectance values of two spectral bands.
Difference	G-B	R-B	R-G	Difference of reflectance values between two spectral bands.
Squared	BSq	GSq	RSq	Squared values of reflectance of the spectral bands.
Log	LogB	LogG	LogR	Log (base 10) of reflectance values of the spectral bands.
Product	B*G	B*R	G*R	Represent interactions between two spectral bands.
Ratio	B/G G/B	B/R R/B	G/R R/G	The most basic spectral index type used for remote sensing studies. Mitigates the effects of variability in spectral data.
Normalized Difference (ND) Ratio	NDBG	NDBR	NDGR	The most commonly used spectral index type using remote sensing studies. Calculated as the difference in reflectance values of two spectral bands divided by their sum. More effective for mitigating variability and noise in spectral data.
Natural Log (ln)	ln(B/G) ln(G/B)	ln(B/R) ln(R/B)	ln(G/R) ln(G/R)	Natural log of reflectance value ratios. The form tested for depth estimation for clear-flowing, gravel-bed rivers. Applies attenuation of lights passing through water column.

A multiple regression model was first developed using all 15 predictors, and their significance for estimating water depth was examined. The predictors that were determined to contribute significantly to the model ($p < 0.001$) were used to develop the final multiple regression model.

A PLS regression model is a statistical technique that is capable of handling highly redundant (or collinear) predictor variables (as well as response variables, when applicable) by transforming them into a smaller set of uncorrelated variables by achieving the maximum covariance prior to regression analysis (Wold et al. 1984). Thus, the technique is suitable when you need to analyze more than a few predictor variables that are highly collinear and/or how response variables are related to predictor variables. Because of its unique capability, a PLS regression model is widely utilized for analytical chemistry such as chemometrics or hyperspectral remote sensing (Wold et al. 2001), bioinformatics (Land et al. 2011), and neurosciences (Zhao et al. 2013), as an alternative technique to multiple regression. While the technique facilitates prediction of the responses, the PLS regression does not describe underlying relationship between the predictor and response variables (Wold et al. 1984).

TABLE 2 Correlation Between the Reference Water Depth and Remote-Sensing Variables

Category	Remote-Sensing Variable[a] and Pearson Correlation Coefficient [*r*]		
Spectral Band	*Blue (B)* *[0.37]*	*Green(G)* *[0.58]*	*Red(R)* *[0.35]*
Sum	B+G [0.49]	B+R [0.36]	G+R [0.48]
Difference	G-B [0.51]	R-B [-0.12]	R-G [-0.54]
Squared	*BSq* *[0.38]*	*GSq* *[0.59]*	*RSq* *[0.34]*
Log	*LogB* *[0.36]*	*LogG* *[0.58]*	*LogR* *[0.34]*
Product	*B*G* *[0.48]*	*B*R* *[0.37]*	*G*R* *[0.49]*
Ratio	B/G [-0.24]	B/R [0.28]	*G/R* *[0.53]*
	G/B [0.24]	R/B [-0.28]	R/G [-0.28]
Normalized Difference (ND) Ratio	NDGB [0.24]	NDRB [-0.28]	*NDRG* *[-0.53]*
Natural Log (ln)	ln(B/G) [-0.24]	ln(B/R) [0.28]	*ln(G/R)* *[0.53]*
	ln(G/B) [0.24]	ln(R/B) [-0.28]	ln(R/G) [-0.28]

[a] Remaining predictor variables after collinearity analysis, which better represented the relationship with water depth, are shown in *italics*. These variables were used to develop a multiple regression model.

A PLS regression model was developed using all 33 predictors regardless of collinearity because it is able to handle highly collinear predictor variables as discussed above. A quasi-principal component analysis was performed on the 33 predictors and yielded 33 components having the maximum covariance. We selected the first 11 components for further regression analysis because the 12th component only improved R^2 between predicted and reference depth by 0.001. Once the analysis was complete, the output from the 11 components was reverted to the original 33 predictor variables. The intercept and coefficients for the PLS regression model in the predictor variable domain are summarized in Table 4.

Each regression model was qualitatively and quantitatively examined using scatter plots and descriptive statistics. The models were applied to the spectral data to estimate water depth, and maps of predicted water depth were generated.

Backwater depths predicted using the linear, multiple, and PLS regression models were validated using a sample of 400 points consisting of 100 randomly selected points from each of the four backwater habitats. The sampled points were examined using an outlier test and QQplot. Three samples (0.75% of the validation data) were determined outliers and eliminated from the validation set without replacement.

3 RESULTS

3.1 MODEL DEVELOPMENT

3.1.1 Linear Regression Model

The linear regression model was as follows:

$$Water\ depth\ (m) = 10.753 \times \ln(G/R) + 0.492 \quad (1)$$

The scatter plot of water depth that was predicted using the linear regression model (Equation 1) and the reference depths calculated from the DEM used for the calibration data is shown in Figure 3 (left). For comparison to the data scatter, the plot shows a perfect one-to-one relationship between predicted and reference water depth. The reference depth ranged up to approximately 2 m. However, the model was not able to predict water depths greater than 1 m. The model indicated difficulty accurately predicting water depth beyond approximately 30 cm using the calibration data.

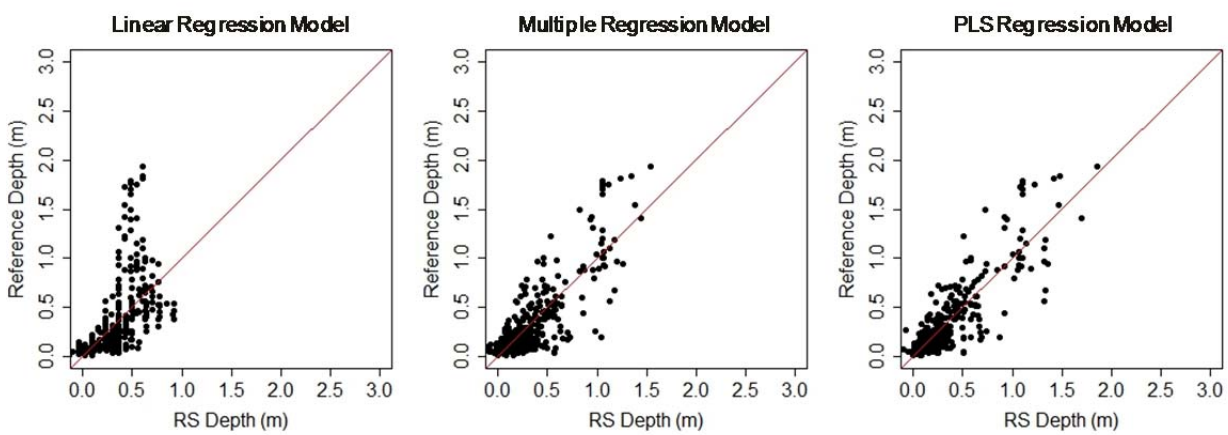

FIGURE 3 Scatter Plots of Predicted Water Depth Using the Regression Models against Reference Water Depth. Regression lines correspond to a hypothetical one-to-one relationship.

Descriptive statistics of the linear regression model using the calibration data is shown in Table 3. The model produced an *F*-value ($F = 156.7$) that was statistically significant ($p < 0.001$) and the coefficient of determination ($R^2 = 0.283$) indicates that approximately 28% of water depth variability can be explained by ln(G/R) alone. The standard error of the model was 0.319 m. A simple average of residuals that estimates bias in model output (i.e., mean error [ME]) indicated there was no positive or negative bias in water depth prediction (ME = 0.000). An average of absolute values of residuals that estimates an overall magnitude of error (i.e., mean absolute error [MAE])–was 0.155 m. The linear regression model was applied to the 2006 NAIP image to generate a map of predicted water depth.

TABLE 3 Statistics of Regression Models Using Calibration Data[a]

	R^2	df	SE	F	p	ME	MAE
Linear regression model	0.283	393	0.319	156.7	< 0.0001	0.000	0.155
Multiple regression model	0.675	388	0.216	134.6	< 0.0001	0.000	0.096
PLS regression model	0.737	362	0.016	690.9	< 0.0001	0.000	0.083

[a] PLS = partial least squares; R^2 = coefficient of determination, df = degree of freedom, SE = standard error, F = F-statistic, p = significance, ME = mean error, MAE = mean absolute error

3.1.2 Multiple Regression Model

A multiple regression model was developed using the 15 predictor variables described in Section 3.2.2. The only statistically significant ($p < 0.01$) predictors were included in the model. The final multiple regression model was as follows:

$$Water\ depth\ (m) = 1.845 \times Blue - 4.883 \times Red - 0.028 \times (B \times G) + 0.048 \times (G \times R) - 76061.32 \times NDRG - 38132.21 \times \ln(G/R) + 114.446 \quad (2)$$

The scatter plot of water depth that was predicted using the multiple regression model (Equation 2) and the reference depth in the calibration data is shown in Figure 3 (center). The plot shows that points are clustered near the one-to-one line between predicted and reference water depth. The model predicted water depth up to approximately 1.5 m, which is 50 cm shallower than the maximum depth in the reference data. Although it showed greater promise than the linear regression model for predicting backwater depth, the multiple regression model had low accuracy when predicting water depth beyond approximately 30 cm using the calibration data.

Descriptive statistics of the multiple regression model using the calibration data is shown in Table 3. The model showed a smaller F-value (F = 134.6) than that of the linear regression model (F = 156.7). The multiple regression model had a moderately high coefficient of determination (R^2 = 0.675), which was considerably higher than that of the linear regression model (R^2 = 0.283). The standard error of the multiple regression model was 0.216 m, which is noticeably smaller than that of the linear regression model (SE = 0.319). There was no positive or negative bias in water depth prediction (ME = 0.000), and MAE was 0.096 m. The multiple regression model was applied to the 2006 NAIP image to generate a map of predicted water depth.

3.1.3 Partial Least Squares Regression Model

The scatter plot of water depth that was predicted using the PLS regression model (Table 4) and the reference depth in the calibration data is shown in Figure 3 (right). The model predicted water depth up to approximately 2 m, which is nearly the entire depth range in the reference data. While the plot shows some similarity to the one for the multiple regression model (Figure 3, center), points were more tightly clustered along the one-to-one line than those of the multiple regression model, particularly for the depth around 30–40 cm. Thus, we could anticipate that the PLS regression model would have a greater predictive power for backwater depth than the multiple regression model.

Descriptive statistics of the PLS regression model (Table 4) using the calibration data is shown in Table 3. The coefficient of determination of the model (R^2 = 0.757) was slightly lower than that of the multiple regression model (R^2 = 0.675) and considerably higher than that of the linear regression model (R^2 = 0.283). On average, the standard error in water depth prediction using the PLS regression model was 0.016 m, which is noticeably lower than that of the multiple regression model (SE = 0.216) and the linear regression model (SE = 0.319). The PLS regression model showed negative bias in water depth prediction (ME = 0.000) like the other two regression models. MAE of the PLS regression model was comparable to that of the multiple regression model (ME = 0.096). The PLS regression model was applied to the 2006 NAIP image to generate a map of predicted water depth.

3.2 QUALITATIVE ANALYSIS OF MODELED WATER DEPTH

The water depth maps of the four backwater habitats (i.e., bw2, bw6, bw7a, and bw13) that were generated by applying the linear, multiple, and PLS regression models are shown in Figures 4–7. The maps showed similar color gradation representing overall distributions of relative water depth across the backwater habitats. Although the models did show some differences in predicted depth, there was overall similarity in areas categorized as deep and shallow.

For backwater 2, the three models predicted deeper water in the southwestern part of the habitat and at the center of the cross channel through the sandbar (Figure 4). This basic pattern can also be seen in the reference water depth map. The range of water depth predicted by the linear regression model was generally smaller, as indicted by the lighter color overall, than that from the other two models. The multiple and PLS regression models yielded very similar overall patterns of water depth. Both maps indicated the deepest water in the southwestern part of the habitat, which showed similarity to the reference water depth estimated from the survey data.

For backwater 6, all three models predicted that the central part of the backwater was deeper than surrounding areas, and this pattern corresponded well with the reference water depth map (Figure 5). The linear regression model predicted deeper water at the central part of the habitat particularly near the vegetated sandbar, which shows greater similarity to the reference map than the maps generated from the multiple and PLS regression models. The multiple

TABLE 4 Intercept and Coefficients of Partial Least Squares Regression Model[a]

Non-Standardized Coefficients

Intercept	Blue (B)	Green (G)	Red (R)	BSq	GSq	RSq	LogB	LogG	LogR
-11880.61	-8.51	-19.59	-17.42	-0.61	0.45	0.27	2052.20	3541.82	3499.54
	B+G	B+R	G+R	G-B	R-B	R-G	B*G	B*R	G*R
	-7.31	-6.55	-9.78	-68.89	-80.28	16.90	-0.06	0.11	0.08
	B/G	B/R	G/R	G/B	R/B	R/G	NDGB	NDRB	NDRG
	2899.72	-3707.51	536.58	-1894.29	1298.79	1298.79	4202.08	-1723.49	-5314.58
	ln(B/G)	ln(B/R)	ln(G/R)	ln(G/B)	ln(R/B)	ln(R/G)			
	-634.53	-118.17	681.06	634.53	118.17	118.17			

Standardized Coefficients

Intercept	Blue (B)	Green (G)	Red (R)	BSq	GSq	RSq	LogB	LogG	LogR
N/A[b]	-63.11	-144.75	-124.11	-500.80	454.20	261.49	119.04	165.70	156.80
	B+G	B+R	G+R	G-B	R-B	R-G	B*G	B*R	G*R
	-106.00	-94.13	-138.26	-203.60	-182.75	56.04	-56.45	99.74	77.02
	B/G	B/R	G/R	G/B	R/B	R/G	NDGB	NDRB	NDRG
	130.08	-142.49	26.33	-129.91	78.17	78.17	115.19	-40.92	-131.93
	ln(B/G)	ln(B/R)	ln(G/R)	ln(G/B)	ln(R/B)	ln(R/G)			
	-35.18	-5.68	33.82	35.18	5.68	5.68			

[a] BSq = square of blue band; GSq = square of green band; RSq = square of red band; LogB = log of blue band; LogG = log of green band; LogR = log of red band; NDGB = normalized difference of green and red bands; NDRB = normalized difference of red and blue bands; NDRG = normalized difference of red and green bands

[b] There is no intercept with a standardized partial least squares model because the standardization results in the distribution of output values having a zero intercept.

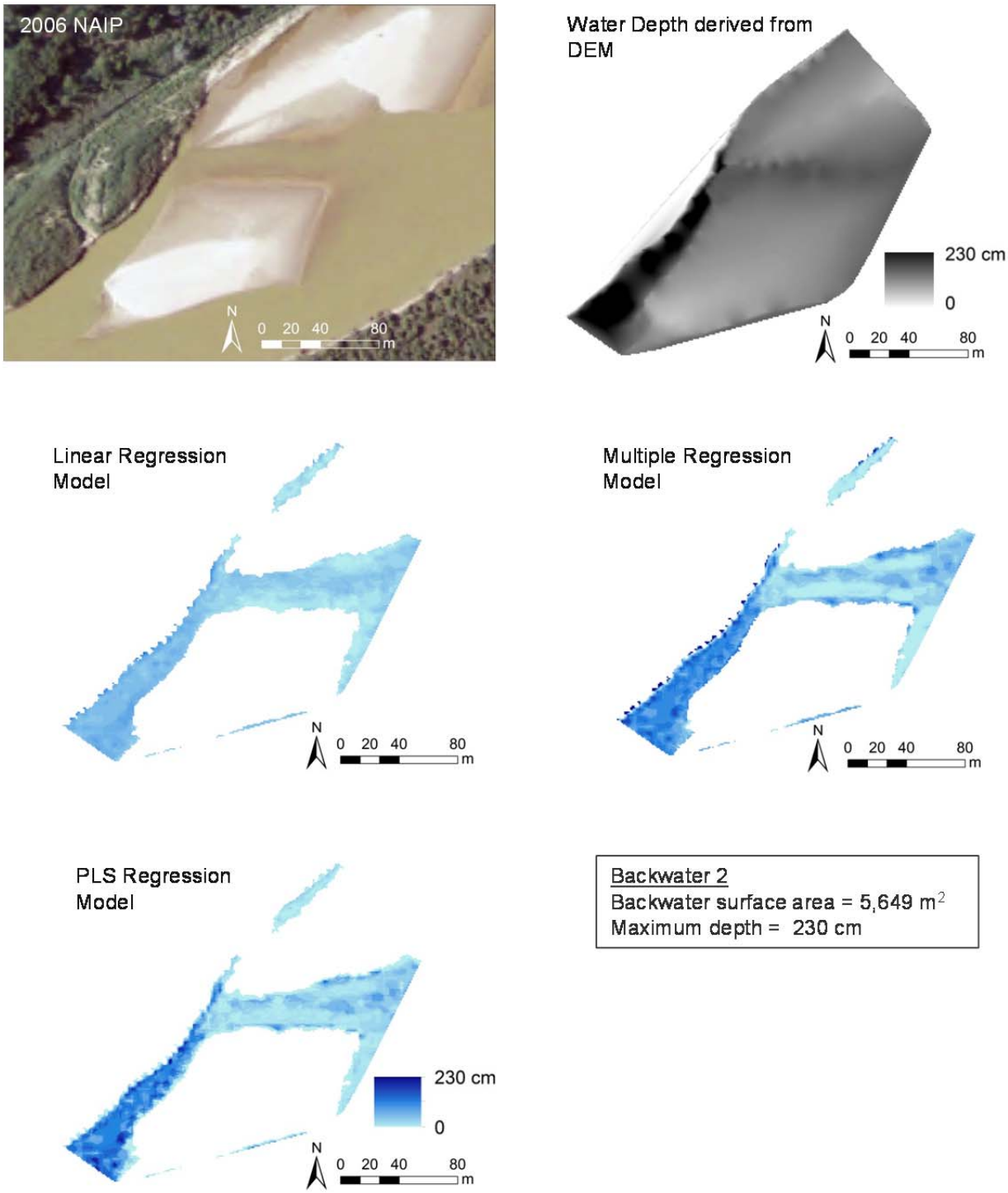

FIGURE 4 Enlarged True-Color Image of Backwater 2 from National Agriculture Imagery Program (NAIP) (top left), Reference Water Depth Map (top right), and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares (PLS) Regression Models (bottom). The reference map was generated using the digital elevation model (DEM) developed from the survey data.

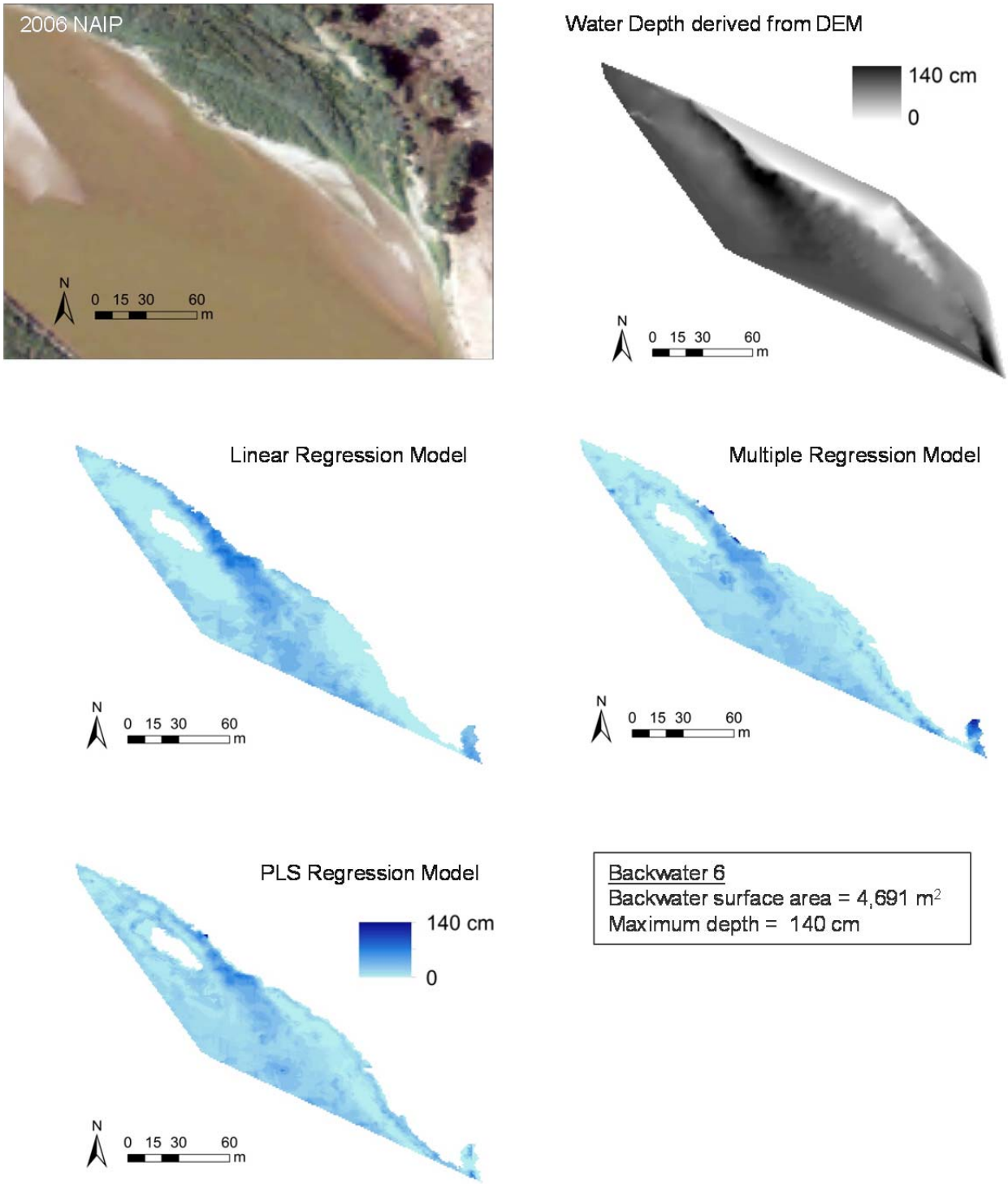

FIGURE 5 Enlarged True-Color Image of Backwater 6 from National Agriculture Imagery Program (NAIP) (top left), Reference Water Depth Map (top right), and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares (PLS) Regression Models (bottom). The reference map was generated using the digital elevation model (DEM) developed from the survey data.

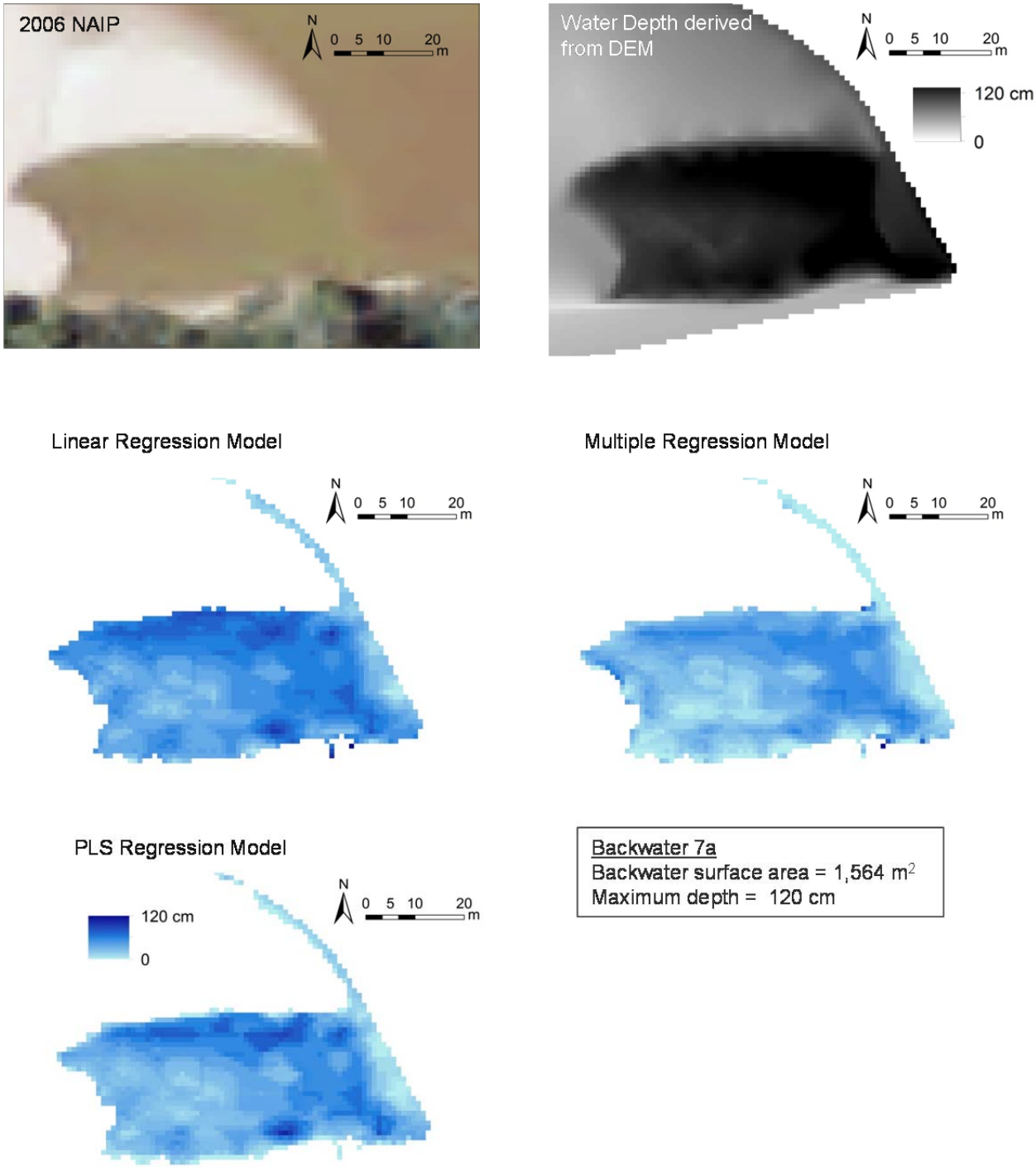

FIGURE 6 Enlarged True-Color Image of Backwater 7a from National Agriculture Imagery Program (NAIP) (top left), Reference Water Depth Map (top right), and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares (PLS) Regression Models (bottom). The reference map was generated using the digital elevation model (DEM) developed from the survey data.

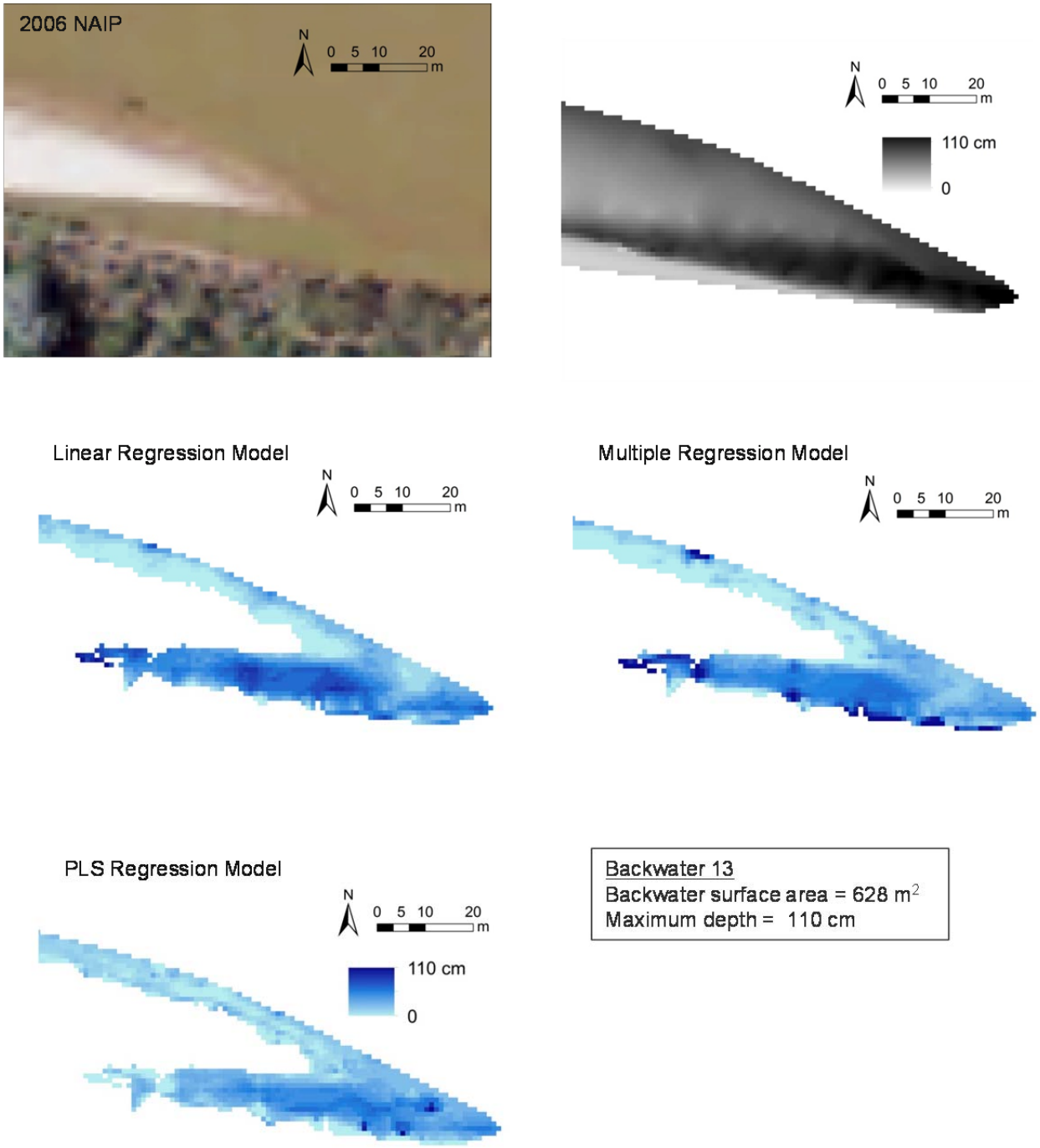

FIGURE 7 Enlarged True-Color Image of Backwater 13 from National Agriculture Imagery Program (NAIP) (top left), Reference Water Depth Map (top right), and Water Depth Maps Derived from Linear, Multiple, and Partial Least Squares (PLS) Regression Models (bottom). The reference map was generated using the digital elevation model (DEM) developed from the survey data.

regression model, however, appears to capture deep water at the southeastern tip of the habitat more accurately than the other models.

For backwater 7a, all three maps show strong similarity with the reference map (Figure 6). The three models yielded similar patterns of water depth: (1) a ridge of sand at the mouth of the backwater that created a shallow area that crossed the mouth (east end of backwater), (2) an area of deeper water just west of this ridge, and (3) relatively shallow water on the west end of the backwater. These patterns corresponded well with the reference water depth map (Figure 6). Deeper water along the edge of the sandbar was better predicted by the linear and PLS regression models than the multiple regression model. The linear and PLS regression models also showed several deep spots near the mouth, which were not as noticeable in the multiple regression model output.

For backwater 13, the three models predicted deeper water between the bank and sandbar and shallow water on the river side of the sandbar (Figure 7). These patterns are also seen in the reference water depth map. The maps developed from the linear and multiple regression models share many similarities in water depth across the habitat. Water depth predicted by the PLS regression model showed greater spatial variability compared to the prediction yielded by the other two models, which had gradual depth changes over large areas.

3.3 QUANTITATIVE ANALYSIS OF MODELED WATER DEPTH

Figure 8 shows scatter plots of water depth predicted using the three regression models against reference water depth by model and backwater habitat. In general, the PLS regression model predicted water depth that were the most consistent with the reference depth of the three models. The linear regression model had the most inconsistent water depth prediction and considerable bias regardless of the backwater compared to other two models. The multiple regression model produced relatively unbiased water depth prediction, except for backwater 7a.

When pooling all backwater habitats, the predicted water depth showed moderately high correlations with the reference depth for the multiple and PLS regression models (R^2 = 0.669 and 0.689, respectively). The water depth predicted using the linear regression model had limited correlation with the reference depth (R^2 = 0.279). All three models had regression intercepts very close to 0 (-0.012–0.006) and slopes that approximate 1 (1.032–1.194), which suggest consistent and unbiased prediction of water depth (Figure 8, pooled). However, when analyzing each backwater independently, the agreement of water depth between the reference and remote sensing prediction varied across the backwaters.

For backwater 2, the predicted water depth showed moderate correlations with the reference depth for all three models (R^2 = 0.376–0.406) (Table 5). The multiple and PLS regression models had intercepts close to 0 (-0.083 and 0.069, respectively) indicating unbiased prediction, but the linear regression model had an intercept of -0.344 (Figure 8, bw2). The multiple and PLS regression models had slopes close to 1 (1.167 and 0.988, respectively) indicating that there is no noticeable trend (e.g., overestimate or underestimate) in depth prediction. The linear regression model had a slope of 2.633 suggesting consistent

underestimation of water depth overall. While the linear regression model nearly consistently underestimated water depth, the multiple regression model tended to overestimate water depths less than approximately 60 cm and tended to underestimate water depths greater than 60 cm.

For backwater 6, the predicted depth showed greater correlations with the reference depth than that of backwater 2 for all three models (R^2 = 0.592–0.695) (Table 5). All three models had intercepts very close to 0 (-0.001–0.004) and slopes less than 1 (0.651–0.866), indicating that water depth tended to be consistently overestimated throughout the depth range (Figure 8, bw6).

For backwater 7a, the predicted depth showed lower correlations with the reference depth than that of backwaters 2 and 6 for all three models (R^2 = 0.109 and 0.249, respectively) (Table 5). All models had intercepts greater than 0 (0.087–0.132) indicating unbiased prediction, and slopes less than 1 (0.449–0.698) suggesting a tendency of overestimation of water depth (Figure 9, bw7a). The linear regression model tended to underestimate shallow water less than approximately 20 cm and tended to overestimate depths greater than 20 cm. The multiple regression model tended to underestimate water depths shallower than 25 cm and tended to overestimate depths greater than 25 cm. The PLS regression model tended to underestimate shallow water less than 30 cm and overestimate depths greater than 30 cm.

Backwater 13, correlations between the reference and predicted water depth show noticeable variation (R^2 = 0.268–0.514) (Table 5). The linear and multiple regression models had positive intercepts (0.070 and 0.067, respectively) with slopes less than 1 (0.650 and 0.797, respectively) indicating a negative bias with underestimation in modeled water depth (Figure 8, bw13). The PLS regression model, on the other than, had an intercept of -0.080 with a slope 1.222 indicating a positive bias with overestimation in predicted water depth. The linear regression model tended to underestimate shallow water less than 20 cm and tended to overestimate depths greater than 20 cm. The multiple regression model tended to underestimate depths less than 30 cm and tended to overestimate depths greater than 30 cm. The PLS regression model had the opposite pattern to that of other two models. The PLS regression tended to overestimate depths less than 40 cm and tended to underestimate depths greater than 40 cm.

Figure 9 shows classification accuracy of backwater depth for all three backwaters when applying a 30 cm threshold. When a model predicted < 30 cm for a pixel with a reference depth < 30 cm, the prediction is considered success. When a model predicted $\geq$ 30 cm for a pixel having its reference depth $\geq$ 30 cm, the prediction is also considered success.

All three models differentiated shallow and deep backwater pixels with accuracy higher than 76% (Figure 9). The PLS regression model yielded the highest overall accuracy (84.1%) followed by the multiple regression model (80.6%). The linear regression models detected backwater pixels $\geq$ 30 cm more accurately (88.1%) than the two models (77.6– 9.9%). The linear regression model, however, did not detect backwater pixels shallower than 30 cm (70.2%) as accurately as those for the multiple and PLS regression models (82.1% and 86.3%, respectively). the PLS regression model had high accuracy for detecting both shallow and deep backwater pixels, which resulted in the highest overall accuracy (84.1%).

TABLE 5 Summary of Model Performance

Correlation between Reference and Modeled Water Depth

	R^2	F	p	Slope	p	Intercept	p
All							
LR	0.279	152.4	< 0.001	1.194	< 0.001	-0.043	0.234
MR	0.669	796.9	< 0.001	1.061	< 0.001	-0.012	0.501
PLS	0.689	872.4	< 0.001	1.032	< 0.001	0.006	0.700
BW2							
LR	0.376	57.2	< 0.001	2.633	< 0.001	-0.344	0.021
MR	0.390	168.7	< 0.001	1.167	< 0.001	-0.083	0.245
PLS	0.406	148.3	< 0.001	0.988	< 0.001	0.069	0.295
BW6							
LR	0.695	218.5	< 0.001	0.651	< 0.001	-0.001	0.954
MR	0.592	139.0	< 0.001	0.866	< 0.001	0.003	0.858
PLS	0.635	166.7	< 0.001	0.812	< 0.001	0.004	0.782
BW7a							
LR	0.232	29.7	< 0.001	0.477	< 0.001	0.087	0.003
MR	0.109	12.0	< 0.001	0.449	< 0.001	0.132	< 0.001
PLS	0.249	32.5	< 0.001	0.698	< 0.001	0.091	< 0.001
BW13							
LR	0.267	36.0	< 0.001	0.650	< 0.001	0.070	0.151
MR	0.333	49.6	< 0.001	0.797	< 0.001	0.067	0.114
PLS	0.514	104.7	< 0.001	1.222	< 0.001	-0.080	0.068

BW = backwater; LR = linear regression model; MR = multiple regression model; PLS = partial least square regression model

4 DISCUSSION

Among the linear, multiple, and PLS regression models, the multiple and PLS regression models showed moderately high correlation between the reference depth and the calibration data (R^2 = 0.675 and 0.737, respectively). The multiple and PLS regression models also had unbiased prediction of water depth as indicated by mean estimates of error (ME = 0.000 and MAE = 8.3–9.6 cm; Table 3). The model validation also showed that those two models had moderately high correlation (R^2 = 0.669 and 0.689, respectively; Table 5).

Water depth maps derived from the three regression models generally showed similar spatial patterns of relative depth that corresponded relatively well to the reference water depth. The variability in the range of actual to predicted depth varied depending on the model and/or backwater habitat. For backwater 2, for example, the linear regression model indicted a small depth variation across the habitat, which is shown as a small range of blue tone in the map, compared to the maps of the multiple and PLS regression models (Figure 4).

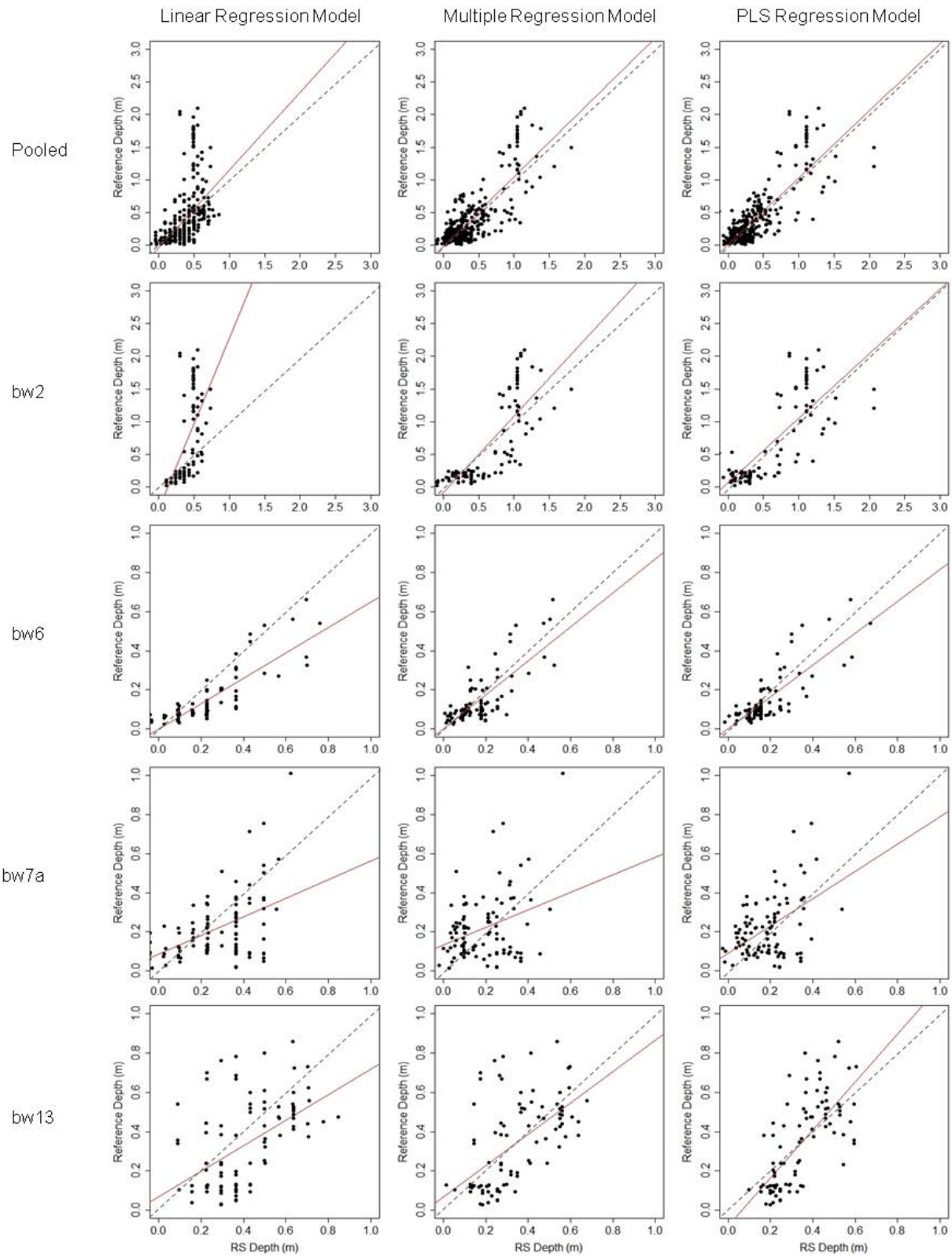

FIGURE 8 Scatter Plots of Predicted Backwater (bw) Depth Using the Regression Models and Reference Water Depth. Model regression lines are solid, and one-to-one lines are dashed. PLS = partial least square. RS = remote sensing prediction

Linear Regression Model

Overall accuracy = 76.3%

		Reference	
		< 30 cm	≥ 30 cm
Predicted	< 30 cm	184	16
	≥ 30 cm	78	118
		70.2%	88.1%

Multiple Regression Model

Overall accuracy = 80.6%

		Reference	
		< 30 cm	≥ 30 cm
Predicted	< 30 cm	215	30
	≥ 30 cm	47	104
		82.1%	77.6%

PLS Regression Model

Overall accuracy = 84.1%

		Reference	
		< 30 cm	≥ 30 cm
Predicted	< 30 cm	226	27
	≥ 30 cm	36	107
		86.3%	79.9%

FIGURE 9 Classification Accuracy of Shallow (< 30 cm) and Deep (≥ 30 cm) Backwater Pixels for All Three Backwaters. PLS = partial least square.

Water depth predicted using the PLS regression model showed the highest correlation with the reference depth in general of three models tested. It was apparent that backwaters 7a and 13 had more biased water depth prediction than backwaters 2 and 6 (Figure 8 and Table 5).

When applying a threshold of 30 cm to classify shallow and deep backwater pixels, the multiple and PLS regression models differentiated deep waters pixels from shallow water pixels more accurately (80.6% and 84.1%, respectively) than the linear regression model did (76.3%).

When accurately detecting areas of deep backwater is important, the linear regression model would provide a highest accuracy (88.1%), but it would come at the expense of a low accuracy of shallow water detection (70.2%) (Figure 9).

On the basis of our analysis, the multiple and PLS regression models are expected to be equally effective for predicting water depth in turbid rivers (e.g., backwater habitats) when pooling all backwaters. However, the PLS regression model had more unbiased and consistent depth prediction across the backwaters than the multiple regression model.

Similar to results presented by Legleiter et al. (2011), the linear regression model using ln(G/R) did not perform well in predicting water depth for our study reach. The other two models may also have difficulty predicting water depth greater than approximately 40 cm. The PLS regression model tended to have a greater depth range for reliably predicting backwater depth (from 30 cm for backwater 6 to 2 m for backwater 2) than the multiple regression model (30 cm for backwater 7a to 1.5 m for backwater 2) (Figure 8).

Lower regression coefficients for the relationship between predicted and reference water depths for backwaters 7a and 13 compared to backwaters 2 and 6 may be associated with the backwater surface areas and the depth range. Backwaters 7a and 13 were considerably smaller (1,564 m^2 and 628 m^2, respectively; Figures 6 and 7) than backwaters 2 and 6 (5,649 m^2 and 4,691 m^2, respectively; Figures 4 and 5), yet these small backwaters had a water depth range approximately 1.1 m that was comparable to that of backwater 7a (1.4 m). Such a great depth change within a small area likely caused considerable depth change or variation within pixels. This situation may have resulted in unexpected complexity in relationship between spectral reflectance values (or pixel values) and water depth that the models developed using a 1-m resolution image could not capture. Using a finer-resolution image (e.g., 0.5 cm resolution) could mitigate the problem.

Inclusion of a large number of edge pixels in the calibration and valuation datasets due to the size of backwater area would be another source of uncertainty for backwaters 7a and 13. Edge pixels that are located near boundaries between water and a sandbar or shore could include non-water surface within a pixel. That means that values of those pixels are average of spectral reflectance signatures containing both water and non-water surfaces. Thus, the pixel values do not have the same relationship with water depth as those of non-edge pixels. This source of uncertainty could be reduced by excluding areas near edges from the selection for calibration and valuation data.

While absolute depth estimates for backwaters may contain uncertainty, knowledge about the distributions of suitable backwater habitats for the endangered Colorado pikeminnow (i.e., backwaters with a maximum depth greater than 30 cm) would be beneficial for river management and hydropower operation planning. Such information could be obtained using the regression models, particularly the PLS regression model that is well calibrated using the field survey data. Based on our finding from the study, we recommend applying a PLS regression model for estimating and monitoring backwater depth in the Green River. Because bathymetry of turbid rivers with sandy substrates is likely dynamic, field surveys should be conducted based on a statistically sound sampling design for model calibration and validation in order to reduce uncertainty in backwater depth estimates. It is likely that regression equations developed for one year will not be directly applicable to other years with different suspended sediment levels. Further exploration of the sensitivity of the approach to different sediment conditions is recommended.

5 REFERENCES

Grippo, M., LaGory, K.E., Hayse, J.W., Walston, L.J., Weber, C.C., Waterman, D., Magnuson, A.K., and Jiang, X.H. (2015). *Relationships between Flow and the Physical Characteristics of Colorado Pikeminnow Backwater Nursery Habitat in the Middle Green River, Utah*., final report to Upper Colorado River Endangered Fish Recovery Program, Denver, Colorado, U.S. Fish and Wildlife Service, Denver, Colorado.

Javernick, L., Brasington, J., and Caruso, B. (2014). Modeling the topography of shallow braided rivers using Structure-from-Motion photogrammetry. *Geomorphology*, *213*, 166-182.

Land, W. H., Ford, W., Park, J. W., Mathur, R., Hotchkiss, N., Heine, J., Eschrich, S., Qiao, X., and Yeatman, T. (2011). Partial least squares (PLS) applied to medical bioinformatics. *Procedia Computer Science*, *6*, 273-278.

Legleiter, C. J. (2013). Mapping river depth from publicly available aerial images. *River Research and Applications*, *29*(6), 760-780.

Legleiter, C. J. (2015). Calibrating remotely sensed river bathymetry in the absence of field measurements: Flow REsistance Equation-Based Imaging of River Depths (FREEBIRD). *Water Resources Research*, *51*(4), 2865-2884.

Legleiter, C. J., Kinzel, P. J., and Overstreet, B. T. (2011). Evaluating the potential for remote bathymetric mapping of a turbid, sand-bed river: 2. Application to hyperspectral image data from the Platte River. *Water Resources Research*, *47*(9).

Legleiter, C. J., Roberts, D. A., and Lawrence, R. L. (2009). Spectrally based remote sensing of river bathymetry. *Earth Surface Processes and Landforms*, *34*(8), 1039-1059.

Legleiter, C. J., Roberts, D. A., Marcus, W. A., and Fonstad, M. A. (2004). Passive optical remote sensing of river channel morphology and in-stream habitat: Physical basis and feasibility. *Remote Sensing of Environment*, *93*(4), 493-510.

Marcus, W. A., and Fonstad, M. A. (2008). Optical remote mapping of rivers at sub-meter resolutions and watershed extents. *Earth Surface Processes and Landforms*, *33*(1), 4-24.

Smith, G. M., and Milton, E. J. (1999). The use of the empirical line method to calibrate remotely sensed data to reflectance. *International Journal of remote sensing*, *20*(13), 2653-2662.

Wold, S., Ruhe, A., Wold, H., and Dunn, III, W. J. (1984). The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. *SIAM Journal on Scientific and Statistical Computing*, *5*(3), 735-743.

Wold, S., Sjöström, M., and Eriksson, L. (2001). PLS-regression: a basic tool of chemometrics. *Chemometrics and intelligent laboratory systems*, *58*(2), 109-130.

Zhao, Q., Caiafa, C. F., Mandic, D. P., Chao, Z. C., Nagasaka, Y., Fujii, N., Zhang, L. and Cichocki, A. (2013). Higher order partial least squares (HOPLS): a generalized multilinear regression method. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *35*(7), 1660-1673.

Environmental Science Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 240
Argonne, IL 60439-4854

www.anl.gov

Argonne National Laboratory is a U.S. Department of Energy
laboratory managed by UChicago Argonne, LLC