Scientific
Research
Publishing

# Comparison of Geographically Weighted Regression of Benthic Substrate Modeling Accuracy on Large and Small Wadeable Streams

## Ken R. Sheehan[1], Stuart A. Welsh[2]

[1]KnoWhere LLC, 969 Ocean Blvd #3, Hampton, NH, USA
[2]U.S. Geological Survey, West Virginia Cooperative Fish and Wildlife Research Unit, West Virginia University, Morgantown, West Virginia, USA
Email: swelsh@wvu.edu
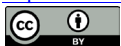
## Abstract

Aquatic habitat assessments encompass large and small wadeable streams which vary from many meters wide to ephemeral. Differences in stream sizes within or across watersheds, however, may lead to incompatibility of data at varying spatial scales. Specifically, issues caused by moving between scales on large and small streams are not typically addressed by many forms of statistical analysis, making the comparison of large (>30 m wetted width) and small stream (<10 m wetted width) habitat assessments difficult. Geographically weighted regression (GWR) may provide avenues for efficiency and needed insight into stream habitat data by addressing issues caused by moving between scales. This study examined the ability of GWR to consistently model stream substrate on both large and small wadeable streams at an equivalent resolution. We performed GWR on two groups of 60 randomly selected substrate patches from large and small streams and used depth measurements to model substrate. Our large and small stream substrate models responded equally well to GWR. Results showed no statistically significant difference between GWR $R^2$ values of large and small stream streams. Results also provided a much needed method for comparison of large and small wadeable streams. Our results have merit for aquatic resource managers, because they demonstrate ability to spatially model and compare substrate on large and small streams. Using depth to guide substrate modeling by geographically weighted regression has a variety of applications which may help manage, monitor stream health, and interpret substrate change over time.

## Keywords

Stream Habitat Modeling, Geographically Weighted Regression, Spatial Scale,

Habitat Interpolation, Geographic Information System

## 1. Introduction

Wadeable stream habitat is monitored and studied across stream size, from ephemeral to many meters wide, to manage for various aspects of stream ecology including fish population dynamics and species occurrence [1] [2] [3]. Because of ties between habitat and population dynamics, wadeable stream assessment and monitoring protocols focus on quantifying key abiotic variables, such as substrate and depth. Their assessments are used to create maps, monitor change, and categorize streams based on the information from those assessments. Differences in stream sizes within or across watersheds, however, may lead to incompatibility of data at varying spatial scales [4] [5]. Issues caused by moving between scales on large and small wadeable streams are not typically addressed by many forms of statistical analysis, making the comparison of large (>30 m wetted width) and small stream (<10 m wetted width) habitat assessments difficult [6]. Geographically weighted regression (GWR) is a spatial modeling technique which may address issues of scale compatibility for important variables in stream habitat models.

Benthic substrate is a key variable in aquatic habitat assessments because of its biotic and abiotic importance. Therefore, efforts to maximize effectiveness of substrate assessment are important, because substrate data collection is an integral activity which guides management of those streams [7] [8]. Water depth is also an important and frequently measured stream habitat variable, and is closely correlated with sizes of benthic substrate [9] [10] [11]. Water depth is measured within aquatic habitat assessments, because it drives a variety of ecologic processes within the stream including location, food abundance, predator prey relationships, fish size and reproductive success [12] [13] [14]. However, connections among fish and habitat management protocols and current science often lag [15]. To close that gap, analysis and modeling of stream habitat have begun to incorporate spatial qualities of stream variables like benthic substrate and depth [16] [17].

To monitor and assess substrate on large and small streams, multiple data analysis methods are often used to effectively capture and convey information across stream size. However, due to structural differences associated with changes in stream size and incompatibility of data at varying spatial scales, modeling techniques successful in predicting substrate on a small stream a few meters wide may not have application for a larger lotic body [4] [18] [19]. Therefore, determination of whether or not a single type of spatial modeling is capable of accurately mapping substrate on wadeable streams of both large and small size is a worthwhile undertaking. Further, such a determination would provide an appropriate basis for comparison of streams of varying size within and among stu-

dies if a method was successful in doing so.

Ecological data interact spatially with the environment and it is important to address those qualities or partition them out of the dataset [20]. The importance of depth and benthic substrate to streams is closely tied to their location in the stream, and removal of spatial information would limit analysis and conclusion. Therefore, to accurately model stream habitat variables such as substrate and depth, the spatial qualities in stream data must be addressed within the analysis framework. For example, spatial non-stationarity, auto-correlation, and multi-collinearity are qualities of spatial data which may obscure meaningful results of statistical analysis such as regressions, one of the most commonly used types of analysis of stream habitat data [21] [22]. In addition to the pitfalls caused by spatial auto-correlation and multi-collinearity, there are issues caused by moving between scale though cross scale analysis is important for ecological analysis [4] [23] [24] [25].

Geographically weighted regression has specific qualities which may adequately address data modeling issues caused by moving between scales. Geographically weighted regression is a modeling method which has been successfully used for analysis in fisheries studies by incorporating spatial attributes of data including non-stationarity and spatial autocorrelation [26]. The method addresses spatial non-stationarity by removing constraints of a global model and allowing for local variance to be calculated at each data location [27]. Theoretically, GWR has similarities to ordinary least squares regression (OLS) as both are statistical models that use a spatial non-stationarity approach [28] [29] [30]. For GWR, the standard formula of OLS as shown in Equation (1) is reparameterized as Equation (2) where $(u_i, v_i)$ can be thought of as $x$ $y$ coordinates in association with both $\beta_0$ and $\beta_k$ [29]. In Equations (1) and (2), $y_i$ is the value of the dependent variable for observation unit $i$; $I = 1, 2, \ldots, n$, $n$ is sample size, $\beta_0$ is the regression intercept, $k = 1, 2, \ldots, K$, $K$ is the number of predictors, $\beta_k$ denotes the coefficients for $k$, $x_{ik}$ indicates the value of the $k^{th}$ variable for observation unit $i$, and $\varepsilon$ denotes the residuals.

$$y_i = \beta_0 + \sum_k \beta_k x_{ik} + \varepsilon_i \tag{1}$$

$$y_i = \beta_0(u_i, v_i) + \sum_k \beta_k(u_i, v_i) x_{ik} + \varepsilon_i \tag{2}$$

Though still a type of linear regression, GWR allows for the often unaddressed variable of location to be addressed directly within the dataset. Spatial statistical methodologies such as GWR provide an appropriate framework to address whether or not they are able to perform equally as well on large or small wadeable streams. Geographically weighted regression may provide avenues for efficiency and needed insight into stream habitat such as substrate and depth data as well [31]. More specifically, GWR was created to model data with heterogeneity, which stream habitat variables often exhibit [28]. The method was first introduced in the mid 1990's [28] [29] [32] and later applied to ecological studies [30] [33], and oceanic fisheries research and management [26] [34]. Researchers

have examined and compared the applicability of GWR for analysis of spatial data relative to that of other regression methods [35] [36] [37] [38]. However, such analyses have not extended far into the field of fisheries science [26] [31]. The present study provides a much needed method for statistical modeling and habitat comparison of large and small streams, where results can be presented as practical, easy to interpret stream habitat maps.

## 2. Materials and Method

Depth and substrate were collected from four streams for this study. Two streams were located in the Greater Yellowstone Ecosystem, Gallatin National Forest bordering the western edge of Yellowstone National Park and two were located in West Virginia; one each in Monongalia and Kanawha Counties (Figure 1). A combined total of 17,040 $x$, $y$ coordinate points, for a sum total of 1583.07 square meters were recorded; each coordinate point represented a 0.3048 m × 0.3048 m (0.093 m$^2$) cell on the stream site. Depth and dominant substrate were recorded individually at all coordinate points, which represented the centroid of each cell. Sites were chosen because they contained at least one pool and riffle interface.

The first of the two western sites was located in the Gallatin National Forest on Little Wapiti Creek, Montana, (111°16'53.546"W, 45°2'20.639"N). The second study site was located in another tract of Gallatin National Forest on Grayling Creek, Montana (111°6'16.407"W, 44°48'16.878"N). Sampling occurred between July 18, 2008 through August 4, 2008 during daylight hours and periods of normal stream flow. The Wapiti Creek site measured 33.5 meters long by 10 meters
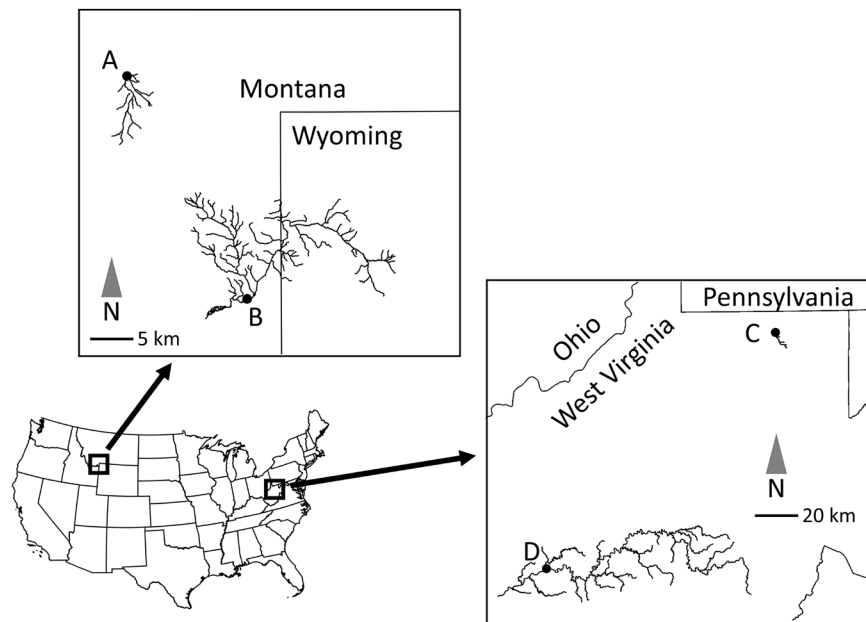


**Figure 1.** Geographic locations of study sites including Little Wapiti Creek (A) and Grayling Creek (B) in Montana, and Aarons Creek (C) and Elk River (D) in West Virginia, USA.

wide, for a total of 3630 cells (335 m$^2$), each representing depth, and dominant substrate type. The Grayling Creek site measured 27.5 meters long by 18 meters wide, for a total of 4950 cells (495 m$^2$), each representing depth and dominant substrate size. The study sites were selected due to relatively remote, though easily accessible and undisturbed locations in the Greater Yellowstone Ecosystem, a mountainous region including characteristic precipitation regime of heavy winter snows and dry summers. The two sites were 28.2 km apart. Little Wapiti creek feeds the Gallatin River Watershed, and Grayling Creek drains directly into Hebgen Lake. Free range cattle use the Little Wapiti watershed during summer, though both sites were relatively undisturbed and stream bed alteration was not attributable to cattle or other unnatural disturbances at the time of sampling.

Clendenin Shoals on Elk River is just downstream of the effluent of Big Sandy Creek in Kanawha County, West Virginia (81˚21'3.857"W, 38˚29'20.73"N). The site measured 22 meters wide, by 27 meters long. Wetted width of the river varied between 30 and 80 meters wide. Unique among the four study sites, this site was located within the town limits of Clendenin. Housing and moderate urban development occur along the banks or the river, though flow and river dynamics are relatively undisturbed due to lack of urban, suburban, or farming development upstream of the study site. The second eastern study site was located on Aarons Creek, which lies within the Monongahela River system in Monongalia County, West Virginia (79˚56'0.625"W 39˚37'8.69"N). The site measured 24.33 × 8.66 m and contained 2268 cells (210.69 m$^2$). There is sparse to moderate urban and suburban development along approximately 70% of this 13.5-km stream. The riparian area of the stream (5 - 50 m wide) is a mixture of field, lawn, and mixed hardwood forest.

Study sites on all four water bodies were delineated by grid cells (1/3 m$^2$ resolution per cell, an area of 0.093 m$^2$) using a fifty-meter tape measure, laser rangefinder, and flagging (later removed). Habitat variables depth and dominant substrate were measured at the centroid of each cell along a secured tape measure crossing the entire site horizontally. A single piece of rebar was inserted into the bank material on each streambank and high tensile line was secured to the rebar to guide the tape measure. As each row of data collection was finished the rebar was repositioned upstream to provide support for the next. Starting at the downstream left of each site, values for depth and benthic substrate were recorded for each $x$, $y$ coordinate. When each row was completed, the tape measure was moved upstream for completion of the next row. This was repeated until the site (minimum of one pool riffle interface) was captured in a complete grid of $x$, $y$ coordinate points with habitat variables at each point.

Stream depth (cm, top-setting wading rod) was measured at the center of each cell. The dominant substrate size was also recorded in each cell along a continuous scale in millimeters from 0.05 to 300 mm based on longest axis diameter. Thus, actual values of water depth and substrate size were recorded for each 0.093 m$^2$ cell for each study site. Four data sets (one for each site) were created electronically using ArcMap 10 (ESRI, 2010) and Microsoft Excel 2010. In Arc-

map, corner points for each study site were geo-referenced and exported to Microsoft Excel. Next, $x$, $y$ coordinates were calculated for the remainder of cells in the site grid and appended to the initial dataset of water depth and substrate size. The final datasets were imported back to ArcMap 10 for analysis.

Prior to the use of geographically weighted regression, Ordinary Least Squares (OLS) regression was run on each site to establish appropriate need for geographically weighted regression [29] [39]. Output from OLS regression, the Koenker-BP test statistic [40] was examined to establish spatial non-stationarity variance significance. Moran's I was also run on all sites to indicate spatial pattern of the data [41]. If spatial autocorrelation was discovered in Moran's I values and residual maps of OLS regression indicated broad patterns of over and under estimation of values, then geographically weighted regressions were run.

Once appropriate use of GWR was established at each study site, GWR was run on all sites in entirety. In addition, 30 random points were generated within the stream boundaries at each site, resulting in a total of 120 random sample points (Figure 1, Figure 2). Each of the 120 points had a buffer of 2.5 meters applied to it creating a sampling quadrat, and all points captured within that quadrat were used for GWR regressions. Geographically weighted regression on Aarons Creek, Elk River, Grayling Creek, and Little Wapiti Creek sites produced 30 $R^2$ values each (one for each sampling quadrat). Thus, 60 $R^2$ values were recorded for two categories of streams; large (Elk and Grayling sites) and small (Little Wapiti and Aarons creek sites).

Additional details concerning GWR tests were as follow: substrate size was set as the dependent variable and depth was set as the explanatory variable. In order to create a standard comparable result across sites, regressions were run on each site using the same kernel type and bandwidth method. Each regression was observed to assure that no model misspecification occurred for any features. Specifically, GWR parameters for each pool riffle complex were run using an adaptive kernel type and bandwidth parameter.

Search radius was explored to provide maximum $R^2$ values and no model miss-specification. Decreasing the search radius from the default (30) yielded higher overall $R^2$ values due to substrate depositional pattern which often occurs in highly localized regions and transitions abruptly. After exploring search radius, it was set to eight points for Grayling, Little Wapiti, and Aarons Creek and 16 points for Elk River, the largest site. For reference, values of r-squared were recorded for each site as a whole.

Results would provide two data sets for analysis; a population of 60 $R^2$ variables from large streams and a population of 60 $R^2$ variables from small streams. Variance between the data sets was examined using an f-test. Results from the f-test would determine the appropriate type of t-test. Comparison between large and small streams was accomplished by performing Welch's t-test assuming unequal variance on the two populations under the null hypothesis that the means of the two populations were not significantly different at an $\alpha$-level of 0.01.

## 3. Results and Discussion

Results from OLS for all sites had significant p-values for the Koenker-BP test statistic, indicating that non-stationary variance has made standard error of the regressions unreliable. Moran's I tests for all sites demonstrated a clustered pattern in the residuals from OLS regression; all Moran's I values were significant at an $a$ level of 0.01, indicating there was less than a 1% chance this pattern was due to chance, and likely due to spatial non-stationarity qualities in the data. Per described methodology, GWR was run on all sites in entirety, and on all random samples of both large and small streams after appropriate use of the GWR statistic was established. Visual inspection of OLS regressions shows large areas of over and under estimation in prediction values of substrate along heterogeneous areas, and zones of substrate transition (**Figure 2**, **Figure 3**).
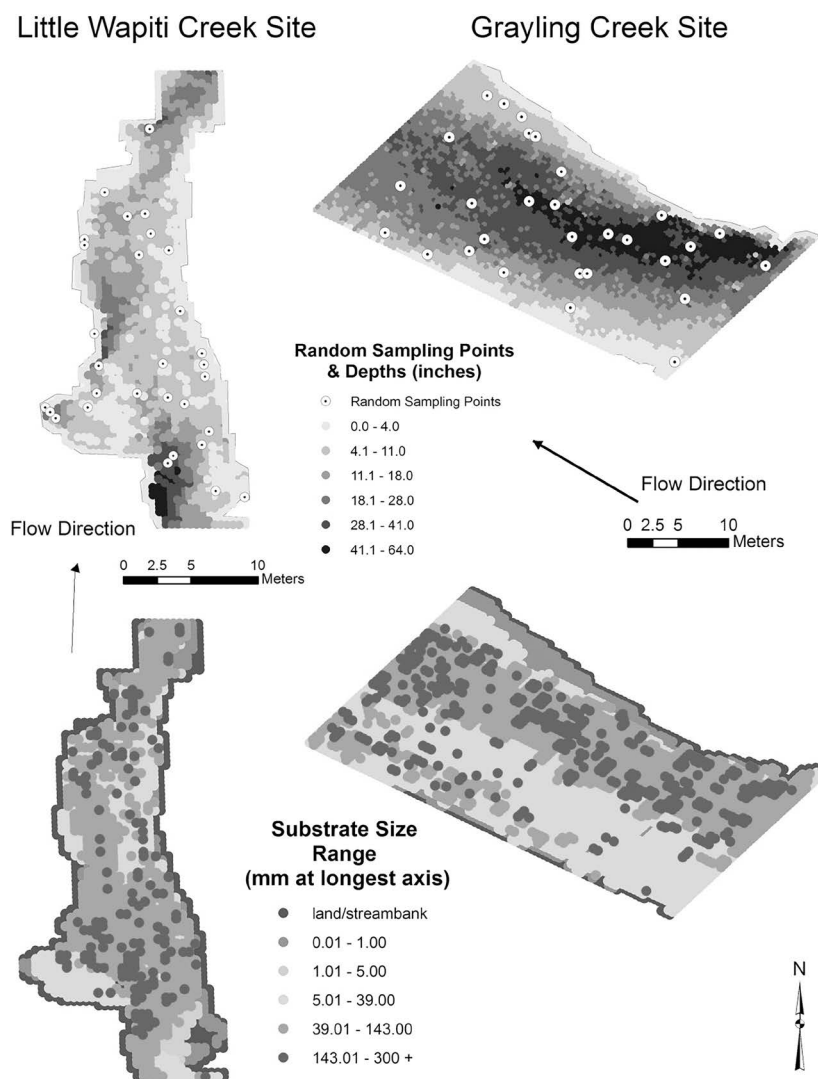


**Figure 2.** Little Wapiti and Grayling Creek locations with 30 random sample points applied within each site. Actual substrate and depth information is also included for reference. Geographically weighted regressions were performed in a 2.5 meter radius surrounding each random sampling point.

Aaron's Creek Site                    Elk River Site
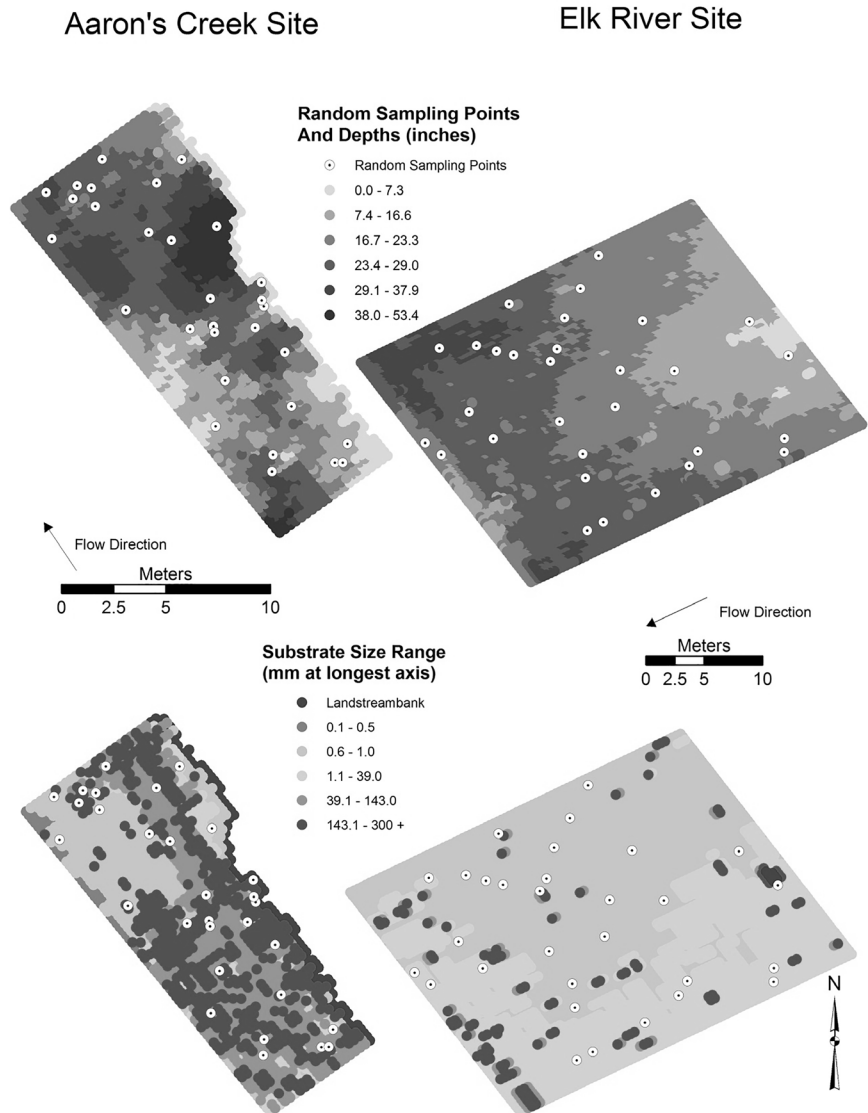


**Figure 3.** Aarons Creek and Elk River locations with 30 random sample points applied within each site. Actual substrate and depth information is also included for reference. Geographically weighted regressions were performed in a 2.5 meter radius surrounding each random sampling point. Elk River displays greater areas of substrate homogeneity than other sites.

Geographically weighted regression $R^2$ values of full study sites were 0.754 on Elk River, 0.839 on Aarons Creek, 0.871 on Grayling Creek, and 0.912 on Little Wapiti Creek. Random sampling and regression at 60 larger order and 60 smaller order stream quadrats produced $R^2$ values ranging from 0.170 to 0.998 on Elk River, 0.698 to 0.930 on Aarons Creek, 0.699 to 0.990 on Grayling Creek, and 0.701 to 0.932 on Little Wapiti Creek. Test for variance indicated the variance of the two populations was not equal with a p-value < 0.01. Means for $R^2$ of random quadrats were 0.83 for small streams and 0.79 for large streams. Results from the Welch's t-test demonstrated differences in means from substrate models of random samples of large and small streams were not significant as shown

by a p-value of 0.22 at $\alpha$ level of 0.01.

Visual inspection of residual maps of GWR regressions indicated the statistic responded similarly to substrate for all streams, regardless of size, by minimizing over and under estimation of predicted values of substrate. This can be directly observed in GWR regressions by noticing the minimization of locations returning standard deviation errors in the −1.5 to −0.5 and the 0.5 to 1.5 category in comparison to those from OLS regressions on all streams (**Figure 4**, **Figure 5**). The majority of predicted values which fell into those two categories were adjusted into the −0.5 to 0.5 standard error range, showing a marked improvement in overall site models for substrate. Overestimation and underestimation of values was limited to highly heterogeneous localized regions in GWR regressions (**Figure 4**, **Figure 5**).
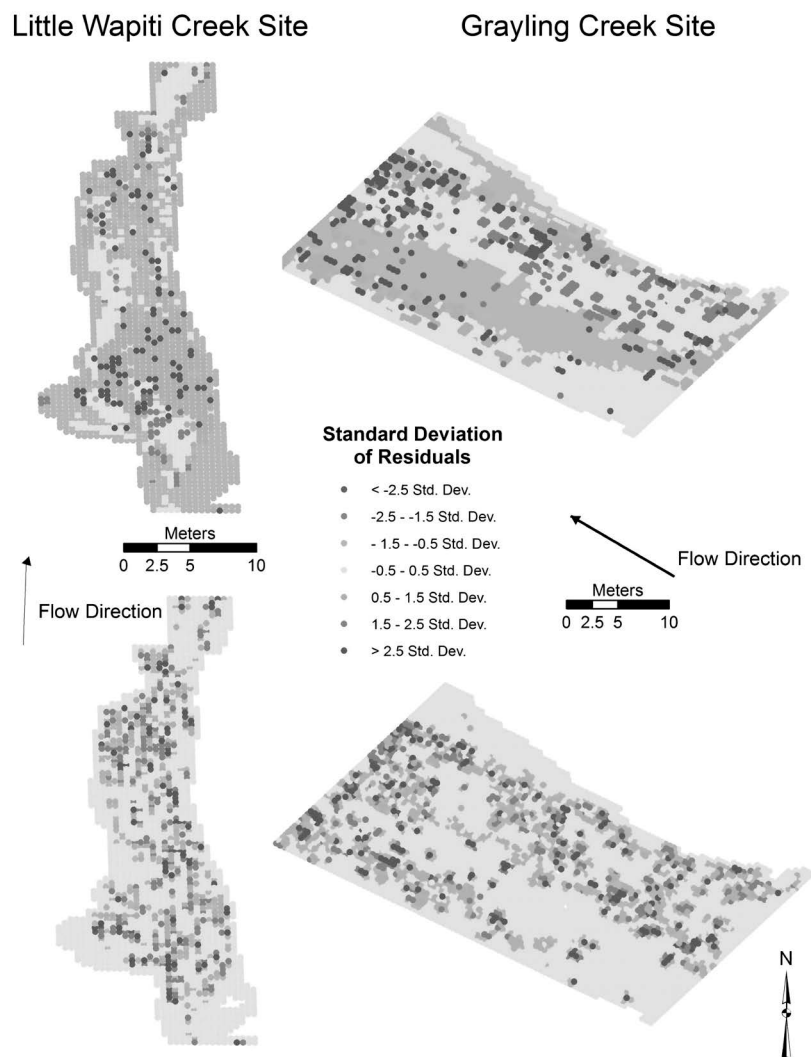


**Figure 4.** Residuals of ordinary least squares regression (top of figure) and geographically weighted regression residuals (bottom of figure) of Little Wapiti and Grayling creeks. Visual inspection of standard deviations of geographically weighted regressions contain a much larger amount of results within 0.5 standard deviations, the result of non-static variance applied at individual locations.
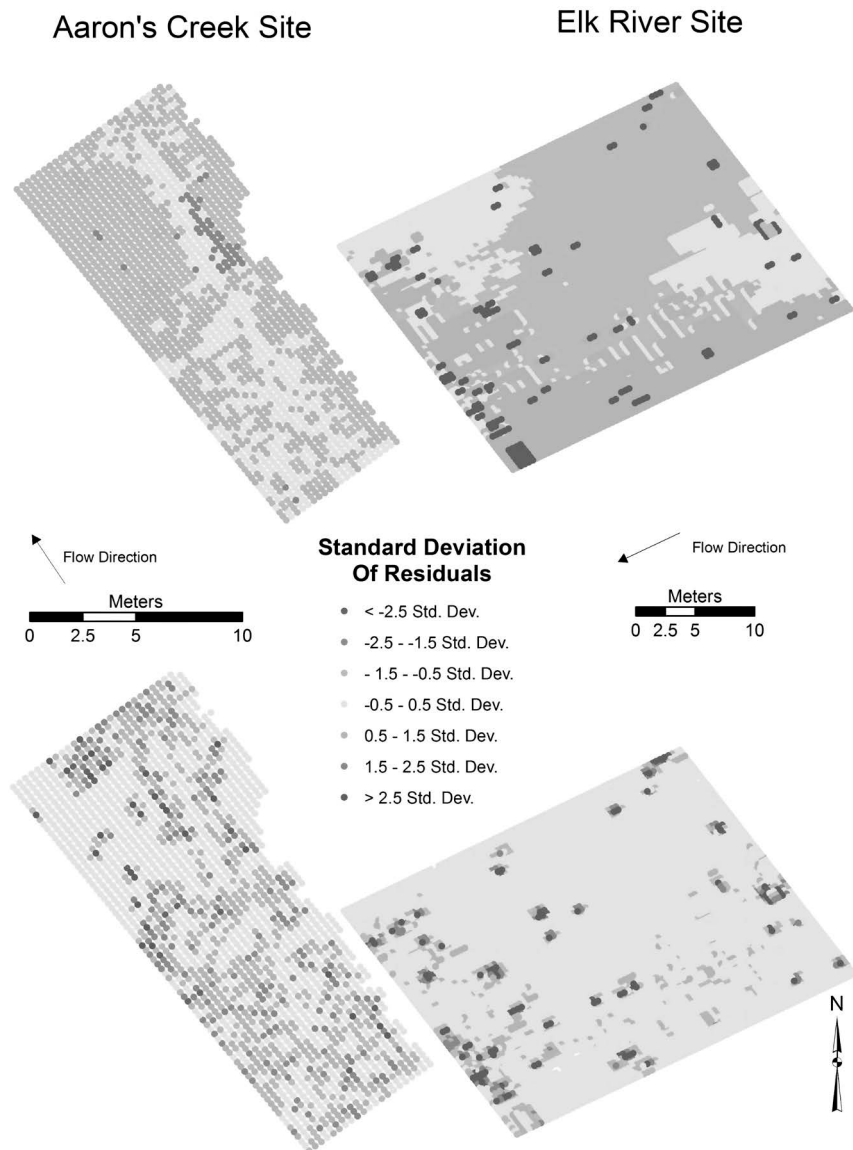
**Figure 5.** Residuals of ordinary least squares regression (top of figure) and geographically weighted regression residuals (bottom of figure) of Aarons Creek and Elk River. Elk River has large areas of more homogeneous substrate, which GWR was better able to model using local variance calculations.

Our results establish the effectiveness of GWR as a modeling tool for wadeable streams of disparate size when analysis of spatial habitat data is required. Specifically, our results show substrate is modeled in an equally effective manner on large or small wadeable streams when using GWR. This is supported by comparison of two large, and two small wadeable streams and values of $R^2$ from their geographically weighted regressions. Our results have implications to fisheries biologists and managers wishing to provide consistent, comparable assessment, analysis, and mapping results of stream habitat variables when their studies span multiple stream sizes. Implications include using a single regression type (GWR) to model stream habitat variables, potential for more reliable management of

stream habitat when substrate is a key decision variable, and use of visual residual maps to gain insight into the results of substrate modeling.

The ability to use a single regression model to address spatial mapping and assessment of both large and small wadeable streams demonstrates geographically weighted regression's utility to stream assessments and thus to fisheries science. It accomplishes this by minimizing the amount of statistical procedures necessary to properly map stream habitat variables such as substrate. The statistical analysis type used does not need to change in order to analyze and compare streams of varying size.

Our study indicates GWR's strength as a statistical procedure for modeling frequently assessed stream habitat spatial variables such as substrate. This is important in part because failure to address spatial qualities in a dataset when performing regressions may create ambiguous or erroneous results due to spatial autocorrelation [21] [22]. Using geographically weighted regression to model stream habitat data with acknowledged spatial qualities directly addresses the issues posed by spatial auto-correlation. It performs this task by removing a single global variance value used to calculate individual predictions within the model with local variance which is calculated using a specified search radius or parameter [28] [32]. This change allows for more accurate representation of structure found within data exhibiting spatial non-stationarity (spatial variance), such as the depositional pattern of substrate. The usefulness of GWR to fisheries science and management can be seen in this example of spatial variance; there are many frequently assessed variables found within a stream with spatial non-stationarity including flow velocity, temperature, location of large woody debris, and seasonal location of fish.

When modeling variables relevant to stream management such as substrate the ability of an analysis method to produce consistent results is particularly important. Consistent, unambiguous results lend themselves towards use in comparison because of confidence in the methodology and the added complexity of comparing results from different methods being removed. As stream size varies, our study indicates mean values of $R^2$ are consistent and do not differ statistically when an appropriate GWR model is applied. Without consistent results the conclusions and management decisions drawn from them would be suspect, and at the very least, somewhat weaker.

Results of the study are in line with prior literature discussing fluvial geomorphologic relationships between depth and substrate [42] [43]. In our results the correlation between depth and substrate depositional pattern is shown by relatively high $R^2$ values between substrate and depth. We provide evidence that substrate may be effectively modeled using depth as the independent, or influencing variable when using geographically weighted regression to model wadeable stream habitat data. This statement is supported by accuracy levels of geographically weighted regression models of substrate on small and large wadeable streams are not significantly different as indicated by GWR $R^2$ values and t-test results.

Practical maps (maps which maintain both utility towards the named purpose and ease of use) are important aspects of interpreting results of stream habitat assessments, but are not often provided by statistical analysis as they are by GWR. Maps can be a valuable tool and provide insight which non-visual results would not provide [44]. Visual analysis of habitat type, amount, and fish populations on small streams does successful precedent [45]. In addition to practical maps of habitat data already produced by mapping stream habitat data in a geographic information system, GWR provides maps in the form of residuals and predicted values for each coordinate location in the dataset. Therefore, visual interpretation of data and results is a useful benefit to the decision making process provided by GWR when used in a program such as ArcMap 10.

An integral visual step in the GWR process for evaluating substrate in streams is establishment of the appropriate use of spatial regression rather than traditional ordinary least squares regression. Ordinary least squares regression is always run to demonstrate non-stationarity in the data, without which GWR would not be an appropriate course of action [28] [32]. A visual byproduct of OLS regression is the residual map created to show amount of standard deviation of each predicted value at each data coordinate. Besides showing non-stationarity, substrate was clearly over and underestimated in a greater proportion of the study area than with GWR residual maps on all four main study sites. In this instance, visual comparison of GWR and OLS regression residual maps helped to illustrate the benefits of applied local variance for spatial stream data such as substrate, rather than a single variance. Specific to the comparison of large and small streams, visual inspection of OLS and GWR residual maps show that GWR reacts similarly to both large and small streams.

As seen in maps of large streams Elk River, and Grayling Creek maintained sizeable areas with high levels of homogeneity of substrate in comparison to smaller stream sites Aarons and Little Wapiti creeks (Figure 2, Figure 3). However, Elk River substrate deposition pattern, while relatively heterogeneous in some areas, does have a large area of sand interspersed with boulders and cobble. That model $R^2$ means of large and small stream groups were not significantly different even with marked difference of substrate depositional pattern is a strong indication of the effectiveness of GWR regression for modeling stream substrate.

## 4. Conclusions

There are several future considerations brought into focus as a result of this study. Because of the correlation between depth and substrate and the ability of GWR to model it, future exploration could examine the potential of GWR for predictive modeling of substrate based on depth structure of the stream. Such an exploration would potentially remove the need for in depth substrate data collection once a baseline for modeling was created for the stream in question. It will be interesting to examine other frequently assessed variables and observe

their response to similar methodology. This study also provides evidence that OLS residuals may guide map creation by demonstrating where models over or underestimate values.

In sum, consistent, accurate, and comparable spatial modeling of substrate on both large and small size wadeable streams are possible when using geographically weighted regression. Large and small stream substrate models responded equally as well to GWR while providing practical, easy to interpret maps of the data and analysis results. Further, the outcome eliminates the need for multiple types of statistics to be used to model streams of different sizes. It additionally provides a much needed method for comparison of large and small streams in the form of an $R^2$ value (by the same procedure) which gives proportional accuracy of models. Our methods have merit for fisheries managers, because they provide comparable, clear results which may be used to visualize substrate, a key habitat variable, useful for management of fish populations in large and small wadeable streams.

## Acknowledgements

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

[1] Gido, K.B., Falke, J.A., Oakes, R.M. and Hase, K.J. (2006) Fish-Habitat Relations across Spatial Scales in Prairie Streams. In: Hughes, R.M., Wang, L. and Seelbach, P.W., Eds., *Landscape Influences on Stream Habitats and Biological Assemblages*, American Fisheries Society Symposium 48, Bethesda, 265-285.

[2] Rosenfeld, J. (2003) Assessing the Habitat Requirements of Stream Fishes: An Overview and Evaluation of Different Approaches. *Transactions of the American Fisheries Society*, **132**, 953-968. https://doi.org/10.1577/T01-126

[3] Winemiller, K.O., Flecker, A.S. and Hoeinghaus, D.J. (2010) Patch Dynamics and Environmental Heterogeneity in Lotic Ecosystems. *Journal of the North American Benthological Society*, **29**, 84-99. https://doi.org/10.1899/08-048.1

[4] Gotway, C.A. and Young, L.J. (2002) Combining Incompatible Spatial Data. *Journal of the American Statistical Association*, **97**, 632-648.
https://doi.org/10.1198/016214502760047140

[5] Scott, J., Heglund, P.J. and Morrison, M.L. (2002) Predicting Species Occurrences. Issues of Accuracy and Scale. Island Press, Washington DC.

[6] Sheehan, K.R. and Welsh, S.A. (2013) Accuracy of Stream Habitat Interpolations across Spatial Scales. *Journal of Geographic Information System*, **5**, 602-612. https://doi.org/10.4236/jgis.2013.56057

[7]    Hansen, W.F. (2001) Identifying Stream Types and Management Implications. *Forest Ecology and Management*, **143**, 39-46.
https://doi.org/10.1016/S0378-1127(00)00503-X

[8]    Ian, M. (1999) The Importance of Physical Habitat Assessment for Evaluating River Health. *Fresh Water Biology*, **41**, 373-391.
https://doi.org/10.1046/j.1365-2427.1999.00437.x

[9]    Allan, J.D. and Castillo, M.M. (2007) Stream Ecology: Structure and Function of Running Waters. Springer Verlag, Berlin.

[10]   Gorman, O.T. and Karr, J.R. (1978) Habitat Structure and Stream Fish Communities. *Ecology*, **59**, 507-515. https://doi.org/10.2307/1936581

[11]   Knighton, D. (1998) Fluvial Forms and Processes: A New Perspective. Oxford University Press, New York.

[12]   Bradford, M.J. and Heinonen, J.S. (2008) Low Flows, Instream Flow Needs and Fish Ecology in Small Streams. *Canadian Water Resources Journal*, **33**, 165-180.
https://doi.org/10.4296/cwrj3302165

[13]   Chisholm, I.M., Hubert, W.A. and Wesche, T.A. (1987) Winter Stream Conditions and Use of Habitat by Brook Trout in High-Elevation Wyoming Streams. *Transactions of the American Fisheries Society*, **116**, 176-184.
https://doi.org/10.1577/1548-8659(1987)116<176:WSCAUO>2.0.CO;2

[14]   Harvey, B. and Stewart, A. (1991) Fish size and Habitat Depth Relationships in Headwater Streams. *Oecologia*, **87**, 336-342. https://doi.org/10.1007/BF00634588

[15]   Minns, C.K. (2001) Science for Freshwater Fish Habitat Management in Canada: Current Status and Future Prospects. *Aquatic Ecosystem Health & Management*, **4**, 423-436. https://doi.org/10.1080/146349801317276099

[16]   Tilman, D., Fargione, J., Wolff, B., D'Antonio, C., Dobson, A., Howarth, R., *et al*. (2001) Forecasting Agriculturally Driven Global Environmental Change. *Science*, **292**, 281. https://doi.org/10.1126/science.1057544

[17]   Vitousek, P.M., Mooney, H.A., Lubchenco, J. and Melillo, J.M. (1997) Human Domination of Earth's Ecosystems. *Science*, **277**, 494.
https://doi.org/10.1126/science.277.5325.494

[18]   Rastetter, E.B., King, A.W., Cosby, B.J., Hornberger, G.M., O'Neill, R.V. and Hobbie, J.E. (1992) Aggregating Fine-Scale Ecological Knowledge to Model Coarser-Scale Attributes of Ecosystems. *Ecological Applications*, **2**, 55-70.
https://doi.org/10.2307/1941889

[19]   Wheatley, M. and Johnson, C. (2009) Factors Limiting Our Understanding of Ecological Scale. *Ecological Complexity*, **6**, 150-159.
https://doi.org/10.1016/j.ecocom.2008.10.011

[20]   Borcard, D., Legendre, P. and Drapeau, P. (1992) Partialling out the Spatial Component of Ecological Variation. *Ecology*, **73**, 1045-1055.
https://doi.org/10.2307/1940179

[21]   González-Megías, A., Gómez, J.M. and Sánchez-Piñero, F. (2005) Consequences of Spatial Autocorrelation for the Analysis of Metapopulation Dynamics. *Ecology*, **86**, 3264-3271. https://doi.org/10.1890/05-0387

[22]   Legendre, P. (1993) Spatial Autocorrelation: Trouble or New Paradigm? *Ecology*, **74**, 1659-1673. https://doi.org/10.2307/1939924

[23]   Levin, S.A. (1992) The Problem of Pattern and Scale in Ecology: The Robert H. MacArthur Award Lecture. *Ecology*, **73**, 1943-1967.
https://doi.org/10.2307/1941447

[24] Townsend, C.R., Dolédec, S., Norris, R., Peacock, K. and Arbuckle, C. (2003) The Influence of Scale and Geography on Relationships between Stream Community Composition and Landscape Variables: Description and Prediction. *Freshwater Biology*, **48**, 768-785. https://doi.org/10.1046/j.1365-2427.2003.01043.x

[25] Urban, D.L. (2005) Modeling Ecological Processes across Scales. *Ecology*, **86**, 1996-2006. https://doi.org/10.1890/04-0918

[26] Windle, M.J.S., Rose, G.A., Devillers, R. and Fortin, M.J. (2009) Exploring Spatial Non-Stationarity of Fisheries Survey Data Using Geographically Weighted Regression (GWR): An Example from the Northwest Atlantic. *ICES Journal of Marine Science*, **67**, 145-154. https://doi.org/10.1093/icesjms/fsp224

[27] An, K.-J., Lee, S.-W., Hwang, S.-J., Park, S.-R. and Hwang, S.-A. (2016) Exploring the Non-Stationary Effects of Forests and Developed Land within Watersheds on Biological Indicators of Streams Using Geographically-Weighted Regression. *Water*, **8**, 120. https://doi.org/10.3390/w8040120

[28] Charlton, M., Fotheringham, S. and Brunsdon, C. (2005) Geographically Weighted Regression. NCRM Methods Review Papers (NCRM/006).

[29] Fotheringham, A.S., Charlton, M. and Brunsdon, C. (1996) The Geography of Parameter Space: An Investigation of Spatial Non-Stationarity. *International Journal of Geographical Information Science*, **10**, 605-627. https://doi.org/10.1080/026937996137909

[30] Kupfer, J. and Farris, C. (2007) Incorporating Spatial Non-Stationarity of Regression Coefficients into Predictive Vegetation Models. *Landscape Ecology*, **22**, 837-852. https://doi.org/10.1007/s10980-006-9058-2

[31] Sheehan, K.R., Strager, M.P. and Welsh, S.A. (2013) Advantages of Geographically Weighted Regression for Modeling Benthic Substrate in Two Greater Yellowstone Ecosystem Streams. *Environmental Modeling and Assessment*, **18**, 209-219. https://doi.org/10.1007/s10666-012-9334-2

[32] Brunsdon, C. (1995) Estimating Probability Surfaces for Geographical Point Data: An Adaptive Kernel Algorithm. *Computers & Geosciences*, **21**, 877-894. https://doi.org/10.1016/0098-3004(95)00020-9

[33] Austin, M. (2007) Species Distribution Models and Ecological Theory: A Critical Assessment and Some Possible New Approaches. *Ecological Modelling*, **200**, 1-19. https://doi.org/10.1016/j.ecolmodel.2006.07.005

[34] Wang, L., Lyons, J., Kanehl, P. and Bannerman, R. (2001) Impacts of Urbanization on Stream Habitat and Fish across Multiple Spatial Scales. *Environmental Management*, **28**, 255-266. https://doi.org/10.1007/s0026702409

[35] Ali, K., Partridge, M.D. and Olfert, M.R. (2007) Can Geographically Weighted Regressions Improve Regional Analysis and Policy Making? *International Regional Science Review*, **30**, 300. https://doi.org/10.1177/0160017607301609

[36] Gao, X., Asami, Y. and Chung, C.J.F. (2006) An Empirical Evaluation of Spatial Regression Models. *Computers and Geosciences*, **32**, 1040-1051. https://doi.org/10.1016/j.cageo.2006.02.010

[37] LeSage, J.P. (2001) A Family of Geographically Weighted Regression Models. In: Anselin, L., Florax, R.J. and Rey, S.J., Eds., *Advances in Spatial Econometrics*, Springer, Berlin, 241-263. https://doi.org/10.1007/978-3-662-05617-2_11

[38] Zhang, L. and Gove, J.H. (2005) Spatial Assessment of Model Errors from Four Regression Techniques. *Forest Science*, **51**, 334-346.

[39] Brunsdon, C., Fotheringham, S. and Charlton, M. (2000) Geographically Weighted

Regression as a Statistical Model. Unpublished Working Paper, Department of Geography, University of Newcastle, Newcastle.

[40] Koenker, R. (1981) A Note on Studentizing a Test for Heteroscedasticity. *Journal of Econometrics*, **17**, 107-112. https://doi.org/10.1016/0304-4076(81)90062-2

[41] Moran, P.A.P. (1953) The Statistical Analysis of the Canadian Lynx Cycle. I. Structure and Prediction. *Australian Journal of Zoology*, **1**, 163-173. https://doi.org/10.1071/ZO9530163

[42] Smith, G.H.S. and Ferguson, R.I. (1995) The Gravel-Sand Transition along River Channels. *Journal of Sediment Research*, **65**, 423-430. https://doi.org/10.1306/D42680E0-2B26-11D7-8648000102C1865D

[43] Kratzer, J.F., Hayes, D.B. and Thompson, B.E. (2006) Methods for Interpolating Stream width, Depth, and Current Velocity. *Ecological Modelling*, **196**, 256-264. https://doi.org/10.1016/j.ecolmodel.2006.02.004

[44] Gergel, S., Stange, Y., Coops, N., Johansen, K. and Kirby, K. (2007) What Is the Value of a Good Map? An Example Using High Spatial Resolution Imagery to Aid Riparian Restoration. *Ecosystems*, **10**, 688-702. https://doi.org/10.1007/s10021-007-9040-0

[45] Hankin, D.G. and Reeves, G.H. (1988) Estimating Total Fish Abundance and Total Habitat Area in Small Streams Based on Visual Estimation Methods. *Canadian Journal of Fisheries and Aquatic Sciences*, **45**, 834-844. https://doi.org/10.1139/f88-101