International Association of Assessing Officers
76th Annual International Conference on Assessment Administration
August 29–September 1, 2010 • Walt Disney Dolphin Hotel • Orlando, Florida

ORLANDO 2010
IAAO
Building Strong Communities Through Assessment Innovation

IAAO

Building Strong Communities through Assessment Innovation

Using Geographical-Attribute Weighted Regression for CAMA Modeling

Wednesday, September 1, 2010 – 1:30-3:00PM

Joshua Myers, MS, Statistician, City of Norfolk
J. Wayne Moore, PhD, Independent Researcher

Overview

- PART 1 – Wayne Moore
  - Review of previous AVM performance research published in 2005/2006
  - Review new research that includes location coordinates from GIS and enhanced methodology for testing horizontal & vertical equity
  - Present horizontal & vertical equity test findings
- PART 2 – Joshua Myers
  - Describe Geographically Weighted Regression (GWR)
  - Describe Geographical-Attribute Weighted Regression (GAWR)
  - Explain GWR/GAWR modeling procedure
  - Describe the GAWR variable specification and how to practically interpret the model

Review of previous research*

- *The laboratory environment of the 2005 experiment:*
  - The experiment was blind to the participants
  - No knowledge of the data, other participants, etc.
- *Purpose:*
  - To contribute to the body of available knowledge on CAMA techniques used by assessors
  - To establish a benchmark (baseline) for further research to discover the factors that might help improve assessment accuracy as well as *horizontal and vertical property tax equity*

* Moore, J.W. 2006. Performance comparison of Automated Valuation Models. *Journal of Property Tax Assessment & Administration.* Volume 3. Issue 1, 43-60.

Review of previous research (cont'd)

- This experiment used following year sales (2004) as the test group instead of a random sample from the group of 2001-2003 sales used for model specification
- Reasons:
  - More closely resembles the reality faced by assessors
  - Could possibly uncover model stability problems
  - Purpose was comparative, not competitive
  - Eliminated concerns about sales chasing
- All 1,299 sales in the 2004 test group were used to evaluate the resultant value predictions – no outliers were eliminated for any participant

Review of previous research (cont'd)

- Twelve distinct sets of 1,299 value estimates provided 15,588 observations for sales ratio analysis and comparison of 4 methods:
  - Six sets from the MRA and AEP modelers
  - Two sets from cost calculations
  - Three sets from the TCM participants
  - One set from the jurisdiction itself – using TCM
- All went through exactly the same sales ratio study process
- Statistics calculated and used for each of the 12 sets of value estimates were:
  - Median A/S ratio
  - Price related differential (PRD)
  - Coefficient of dispersion (COD)

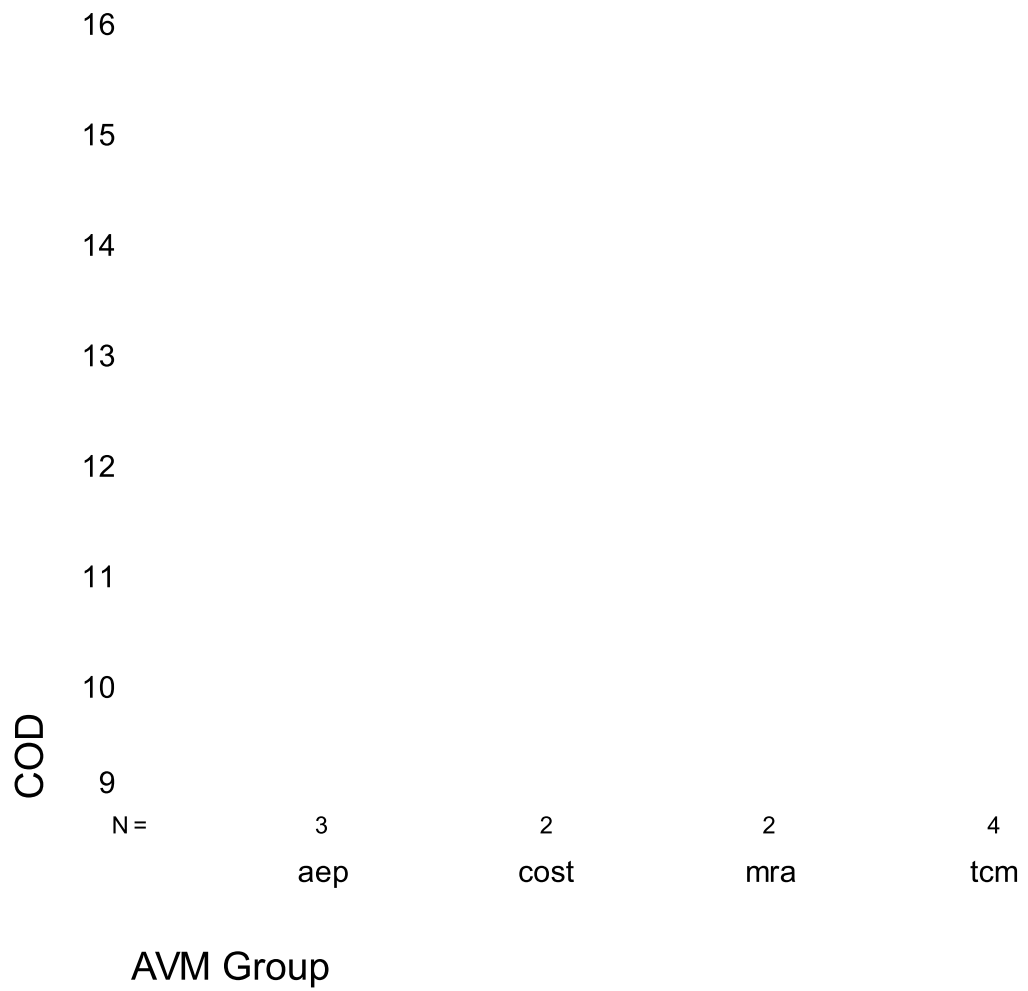
Review of previous research (cont'd)

Results

AVM Type	COD	Median Ratio	PRD
AEP	10.2	94	1.01
AEP	10.9	96	1.03
AEP	12.0	102	1.04
AEP	13.8	90	1.06
COST	14.4	98	0.97
COST	14.9	94	0.98
MRA	10.0	99	1.03
MRA	10.5	99	1.03
TCM	10.1	94	1.00
TCM	10.1	95	1.02
TCM Jurisdiction	10.2	89	1.01
TCM	11.3	96	1.01

Review of previous research (cont'd)

Note: The *transportable cost-specified market* (TCM) AVM is *market-calibrated cost*



New AVM Research with X-Y Coordinates

- Same data set as used for the 2005 AVM research was used
- Only Lat-Long parcel coordinates were added to the data set
- Test methodology was improved compared to 2005
  - Horizontal & vertical equity tested separately
  - *Random sampling and assignment of 31 COD test groups of 25 parcels each*
  - Thus, 775 parcels randomly selected from 1299 blind testing sales in 2004
  - Quintile groups of 155 sale ratios each created from the same 775 sale sample
    - 775 sorted from low to high sale price and divided into five equal groups (quintiles) of 155 each
    - Average (mean) of 155 ratios computed for each vertical stratum and labeled *Quintile Mean Ratio*
  - Horizontal equity tested by comparing the mean CODs (M_{COD}) of the 31 test groups in each AVM using analysis of variance (ANOVA) hypothesis testing
  - Vertical equity tested by comparing the *quintile mean ratios* (*QMRs*) of the five strata - lowest to highest - within each AVM; unequal *QMRs* when tested with ANOVA would indicate vertical inequality
  - New hypothesis tests were conducted at the alpha = 0.01 level instead of the alpha = 0.05 level (i.e. for the new testing, $p < 0.01$; 1 in 100 chance of error)

Significance of the New Research

- Josh Myers' work has been accomplished under the same strict experimental controls as the 2005 research.
- As the baseline, Josh Myers' GAWR/GWR work was blindly applied to the same test dataset used in the 2005 experiment.
- Josh Myers' GAWR model performed significantly better than all other models developed by some of the best modelers in the country and tested in 2005 using exactly the same data set and experimental controls, **but with no x-y coordinates**.
- Josh's Geographical Attribute Weighted Regression (GAWR) model incorporates the property x-y coordinates - this research confirms that location, location, and location *are in fact* the three most important characteristics in real estate value.

New AVM Research: Findings - 1

Summary of Equity Findings for AVM Types with the Research Data Set

	Overall findings				*QMR* findings by quintile					COD findings by quintile				
AVM	Med	COD	PRD*	VEI	1	2	3	4	5	1	2	3	4	5
AEP	0.93	9.95	1.01	6.3	0.91	0.94	0.97	0.96	0.91	10.05	9.34	9.33	9.19	11.11
COST	0.93	14.91	0.99	10.8	0.88	0.92	0.98	0.98	0.96	12.10	14.62	15.00	14.40	15.56
GAWR	0.98	7.67	1.01	4.9	1.01	0.97	0.97	0.98	0.97	8.48	6.51	6.55	6.13	10.14
GWR	0.99	8.39	1.01	6.6	1.03	0.98	0.98	0.98	0.96	8.71	7.93	7.44	7.17	9.72
MRA	0.99	9.91	1.03	13.7	1.00	1.01	1.04	0.99	0.90	8.93	8.94	8.92	9.11	11.71
NoRCN	0.98	7.93	1.01	5.5	1.02	0.98	0.97	0.98	0.96	8.39	6.98	6.70	6.45	10.39
TCM	0.94	10.06	1.01	4.5	0.95	0.95	0.97	0.93	0.92	10.44	8.95	8.78	9.50	12.29

COD - the average absolute deviation of calculated sale ratios from their median expressed as a percentage of the median (IAAO 1997). Larger COD values indicate diminished uniformity.

PRD - the mean assessment to sale ratio of a sales sample divided by the weighted mean ratio of the same sample, computed as the total assessed value divided by the total of the sale prices.

VEI - the absolute value of the difference between the highest and lowest of the five QMRs within a study group divided by the mean of the five QMRs, expressed as a percentage. Larger VEI values indicate less vertical equity. The VEI was developed by Moore for his 2008 dissertation.

* Jensen, D.L. 2009. The effects of heterogeneous variance on the detection of regressivity and progressivity. *Journal of Property Tax Assessment & Administration.* Volume 6. Issue 3, p. 5. Jensen outlined the problems with using PRD as a measure.

New AVM Research: Findings - 2

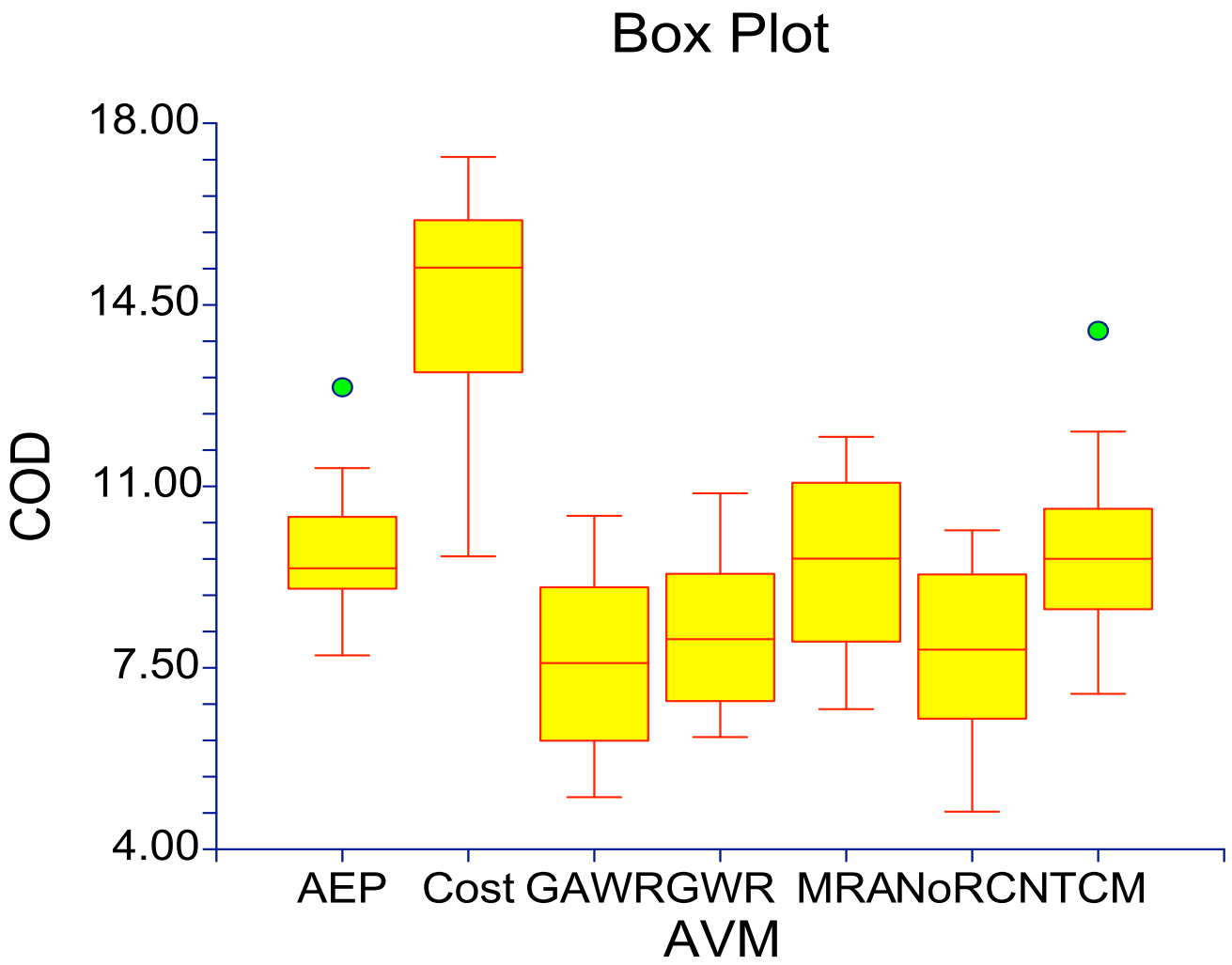
The ANOVA test statistically confirms that at least one AVM's mean COD is different, providing evidence against the null hypothesis that all AVMs produce the same COD performance. The Tukey-Kramer multiple comparison tests for all pair-wise differences between the COD means tells the actual differences.

Comparison of Horizontal Equity Performance between AVMs using the Research Data Set

AVM Test group	Test group sample size	M_{COD}	Tukey-Kramer multiple-comparison test results COD performance is different from:
GAWR	31	7.559	MRA, AEP, TCM, COST
NoRCN	31	7.780	MRA, AEP, TCM, COST
GWR	31	8.239	AEP, TCM, COST
MRA	31	9.549	GAWR, NoRCN, COST
AEP	31	9.641	GAWR, NoRCN, GWR, COST
TCM	31	9.731	GAWR, NoRCN, GWR, COST
COST	31	14.447	GAWR, NoRCN, GWR, MRA, AEP, TCM

New AVM Research: Findings - 3

The comparative AVM COD test results can be illustrated graphically with a box plot



New AVM Research: Findings - 4

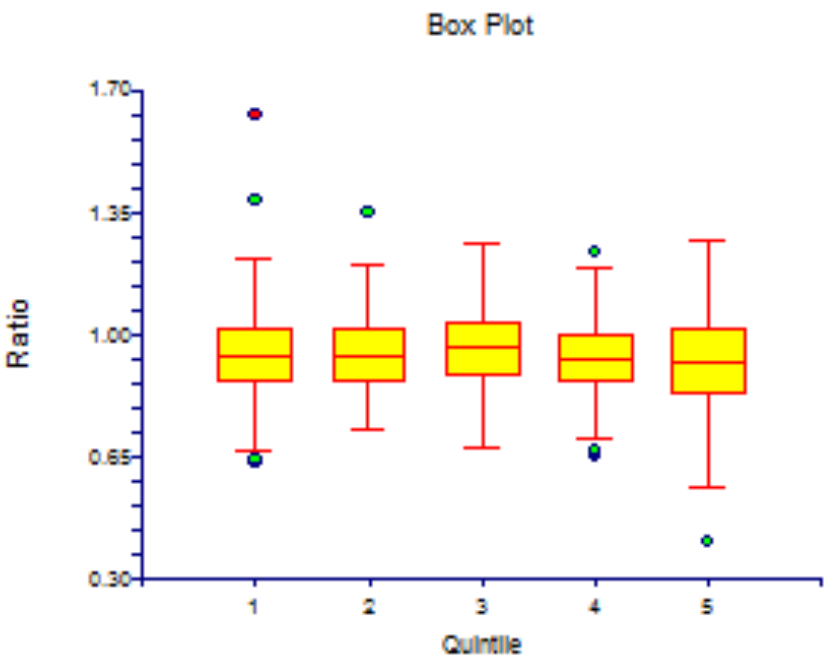
Hypothesis testing for vertical equity was performed using the five groups of 155 ratios in each quintile. A separate analysis was conducted for each AVM test group. Six of the seven AVMs tested ***failed*** the vertical equity test.

ANOVA Test Results of AVM Group Quintile Mean Ratios (QMRs) for Vertical Equity

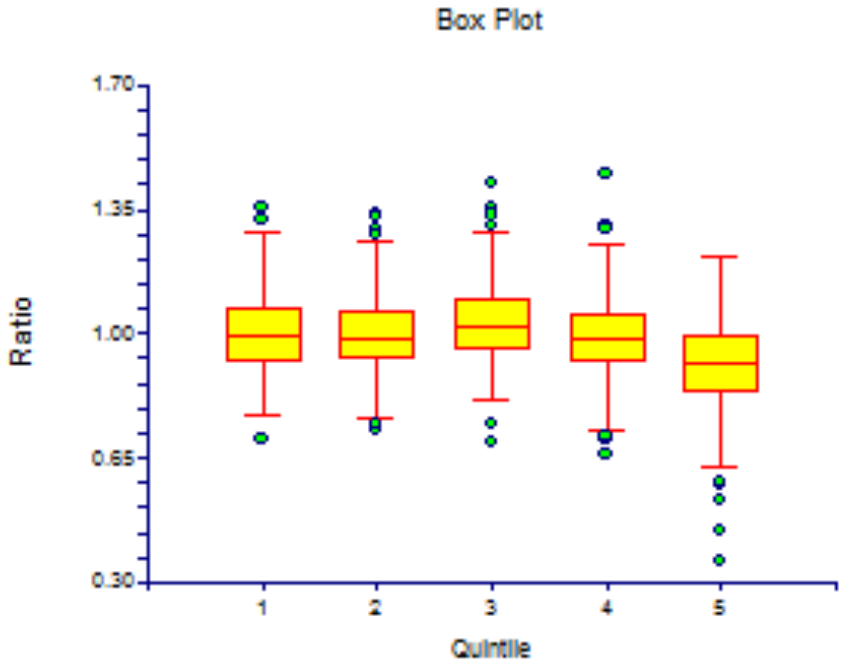
Group	Mean ratio	|Max -Min|	MSE	*F*-ratio[a]	*p*	ANOVA H_0 decision[b]	Kruskal-Wallis H_0 decision[c]	Ratios different?[d]
AEP	0.935	0.059	0.0141	8.32	<.001**	Reject	Reject	Yes
COST	0.942	0.102	0.0293	10.02	<.001**	Reject	Reject	Yes
GAWR	0.983	0.048	0.0112	5.12	<.001**	Reject	Reject	Yes
GWR	0.985	0.065	0.0126	7.46	<.001**	Reject	Reject	Yes
MRA	0.989	0.136	0.0154	26.11	<.001**	Reject	Reject	Yes
NoRCN	0.982	0.053	0.0117	5.65	<.001**	Reject	Reject	Yes
TCM	0.943	0.043	0.0151	2.73	0.02813	Accept	Accept	No

Notes. [a] The *F*-ratio was for F(4,770). [b] H_0: No difference in mean *QMRs*. [c] H_0: No difference in median *QMRs*. [d] Yes indicates vertical inequity because the Quintile Mean Ratios (QMRs) are different by a statistically significant amount.
** Term significant at $p < 0.01$.

New AVM Research: Findings – 5



(a) TCM - accepted test result



(b) MRA - rejected test result

Vertical Equity ANOVA Test Result Box Plots for TCM & MRA

(Similar box plots exist for AEP, COST, GAWR, GWR, and NoRCN)

New AVM Research: Findings – 6

City of Norfolk dataset findings

Descriptive Statistics for the City of Norfolk Dataset for GWR, GAWR, and Sales

	Descriptive statistic								
Group	Count	Median	Mean	Std dev	Min	Max	Med a/s	COD	PRD
Sales	1575	206500	248692	146209	60300	2000000	n/a	n/a	n/a
GWR	1575	212814	255409	144114	68370	1460350	1.029	7.956	1.010
GAWR	1575	213235	255452	144639	69312	1454034	1.029	7.489	1.012

Equity Findings with the City of Norfolk Dataset

	Overall findings				*QMR* findings by quintile					COD findings by quintile				
Model	Med	COD	PRD	VEI	1	2	3	4	5	1	2	3	4	5
GAWR	1.03	7.49	1.01	8.0	1.09	1.05	1.03	1.02	1.01	8.07	6.54	6.53	7.13	7.28
GWR	1.03	7.96	1.01	6.8	1.08	1.05	1.03	1.01	1.01	8.67	7.10	7.22	7.38	7.98

New AVM Research: Findings – 7

Fairfax County dataset findings

GAWR was also applied to the Fairfax County, Virginia data from the 2006 initiative sponsored by the IAAO (Clapp and O'Connor, 2008)*. The sales history dataset was comprised of 51,190 valid residential sales between January 1967 and December 1991. The prediction hold-out sample was comprised of 5000 valid residential sales between January 1972 and June 1991. The table below compares the predictive power of GAWR against the two best practice results and the default OLS (MRA) result reported by Clapp and O'Connor. The GAWR model produced better results than any of the alternative models.

Equity Findings with the Fairfax County Data (January 1972 to June 1991)

		Absolute Value Percentage Error		
Model	Mean	25th Percentile	Median	75th Percentile
Myers GAWR	*10.6*	*3.0*	*6.4*	*11.8*
Gloudemans	11.8	3.7	7.8	14.1
Case	11.8	3.7	8.0	14.1
OLS (MRA)	12.6	4.0	8.4	15.8

New AVM Research: Conclusion

- The original research questions developed were:
  - To what extent do measures of equity/accuracy differ among available methodologies for estimating the market value of single-family homes?
  - Does the use of spatial attributes from GIS in the market value estimating model specification improve the measure of equity/accuracy by a statistically significant amount?
  - Does the addition of attribute weighting improve the measure of equity/accuracy of the market value estimating model by a statistically significant amount over standard GWR?
- The research conclusions are:
  - Horizontal equity does differ among available AVM methodologies
  - Spatial attributes from GIS do improve the COD measures of horizontal equity/ accuracy by a statistically significant amount when incorporated in a geographical-attribute weighted regression (GAWR) model; however, attribute weighting does not produce a statistically significant difference over GWR
  - Vertical equity improvement is not supported by the evidence and needs to be the subject of future research

Part 2
Description of Method

Geographically Weighted Regression

GWR Background and Theory

- Standard Multiple Regression (MRA) yields one set of regression coefficients for the entire study area.
  - These coefficients indicate the estimated unit value of each independent variable.
- GWR runs one separate weighted MRA model for every subject property.
  - Therefore, GWR has a separate set of local regression coefficients for every subject property.

GWR Background and Theory

- The basic GWR relationship is as follows, where k is the number of independent variables, (x_i,y_i) are the coordinates for the i^{th} subject property, $\beta_h(x_i,y_i)$ is the regression coefficient for the h^{th} independent variable for the i^{th} subject property, X_{ih} is the value of the h^{th} independent variable for the i^{th} subject property, Y_i is the sale price for the i^{th} subject property, W_i is the weights matrix for the i^{th} subject property, and ε_i is the error term for the i^{th} subject property.

$$Y_i = \beta_0(x_i, y_i) + \sum_{h=1}^{k} \beta_h(x_i, y_i) X_{ih} + \varepsilon_i$$

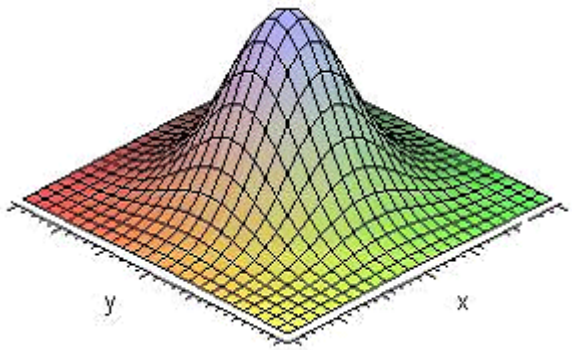
$$\hat{\beta}_i = (X^T W_i X)^{-1} X^T W_i Y$$

GWR Background and Theory

- GWR should be viewed as a series of locally weighted regressions with weighting done by geographic distance.
  - Sale properties that are closer to the subject property have more influence than those that are farther away.
  - These weights are established by a weight function.
  - An adaptive bandwidth, the best number of nearest-neighbor sale properties to include in each sliding neighborhood, is used. Past a certain distance, all sale properties are given zero weight.

GWR Background and Theory

- The weight function used in this research is the bi-square function, given below. Here, d_{ij} is the distance between the i^{th} subject property and the j^{th} sale property and b is the adaptive bandwidth.



$$w_{ij} = (1 - (d_{ij} / b)^2)^2, d_{ij} \leq b$$

$$w_{ij} = 0, otherwise$$

Image Source: Richard Borst presentation at the 2009 IAAO Conference, used with permission.

GWR Background and Theory

- As should be apparent, GWR is fundamentally different from other methods in how it accounts for location.
  - Other methods tend to account for location with neighborhood adjustments, whereas GWR accounts for location using parcel centroid coordinates from GIS and sliding neighborhoods.
  - These neighborhood adjustments used by other methods create a boundary value problem where parcels that meet on the boundary of a two neighborhoods may be given very different adjustments. Also, other methods usually assume neighborhoods to be homogenous when often they are not.
- GWR methodology has been applied to housing datasets from Toronto, Montreal, Sarasota, Fl, Catawba, NC, Fairfax, Va, and Milwaukee, WI, but, to our knowledge, hasn't yet been put into practice inside of an actual assessor's office or tested in a controlled experiment.

GAWR

- GWR is weighted according to distance, but GAWR is weighted according to distance ***and*** attribute similarity.
- Say, you have three properties: a subject property, a smaller sale property two doors down to the left of the subject property, and an exact replica of the subject property that sold two doors down to the right. Under GWR, the smaller property would be given about the same weight as the replica, but under GAWR the replica would be given more weight than the smaller property.
- **GAWR is an analytic implementation of the sales comparison method, which is the #1 recommended method for single-family residential**.

GAWR Weight Function

- The general form of the spatial-attribute weight function for the GAWR model is given below. Here, f(τ) is an exponential function that changes the weight according to the difference (τ) between attributes of the i^{th} subject property, A_i, and its j^{th} neighboring sale property, A_j.

$$w_{ij} = ((1-(d_{ij}/b)^2) * f(\tau))^2,\ d_{ij} \le b$$

$$w_{ij} = 0,\ otherwise$$

$$f(\tau) = e^{-\left|1-(A_j/A_i)\right|}$$

GWR/GAWR Modeling Procedure

- A forward step-wise procedure was employed in the model building phase using the sales history dataset.
  - For the research data, the sales history dataset was the valid sales from 2001 to 2003. For the Norfolk data, the sales history dataset was the valid sales from 07/01/2005 to 06/30/2007.
- At each step, leave-one-out cross-validation was used to determine the COD and the corresponding optimum bandwidth for each of the possible variable additions.
- In general, the variable that most lowered the COD at each step was added to the model.
- It is also important to determine the best weight function and the optimum time period to use in the sales history dataset.

GWR/GAWR Modeling Procedure

- Taking into account the principle of parsimony, the quality of the local coefficients, and the quality of the estimates, the set of variables that yields the lowest cross-validated COD is chosen and the adaptive bandwidth for the set is recorded.
- The chosen set of variables and the adaptive bandwidth are then used to run the prediction of the subject properties using the sales history dataset.
- R, a statistical software program, was used in the analysis.

GWR/GAWR Variable Specification

Norfolk Variable Set	*Research Variable Set*
Intercept	*Intercept*
RCNLD for Dwelling and Garage	*RCN for Dwelling and Garage*
Squared Reverse Month of Sale (Time)	*Reverse Month of Sale (Time)*
Pre-determined Land Market Value as an offset	*Pre-determined Land Market Value*
RCNLD for Other Improvements as an offset	*Total Other Area (sum of attic and basement*
	Total Garage Area
	Total Living Area
	Neighborhood Indicator Variable (coded 1 if the sale property is in the same neighborhood as the subject property, and 0 otherwise)

The Use of Replacement Cost in GAWR

- This goes against the conventional wisdom.
- While being a poor representation of market value, the replacement cost acts as a good index of the property's attributes and how it compares to nearby properties.
- GAWR becomes much more explainable to taxpayers and appraisers when replacement cost is used as the primary variable, because replacement cost is easily understood by many people.
- This clears the major hurdle for regression-based AVMs of being able to explain the model to the masses. The interpretation is now simple: the estimate of property value is made using an implementation of the sales comparison method to make adjustments to the replacement cost.

Conclusions and Future Work

- The 2005 research demonstrated that no statistically significant difference exists between TCM, MRA, & AEP, but that they are each significantly better than Cost.
- The 2010 research demonstrates that GAWR performs significantly better than other commonly used appraisal methods that do not use parcel location X-Y coordinates.
- This new research has shown evidence that if a statistically significant improvement is desired, parcel centroid coordinates must be incorporated using a model like GAWR.
- The Fairfax Dataset results show that not all models that use parcel X-Y coordinates are created equal. GAWR performed the best among a host of models that also used X-Y coordinates.

Conclusions and Future Work

- As demonstrated in Norfolk, GAWR can be applied to local jurisdictions and is not just some “academic” method.
  - The replacement cost is used as one of the main variables in the GAWR model and makes the model much more explainable to the masses.
  - GAWR can also result in potentially large time-savings over manual methods.
- Our future plan is to continue to make the GAWR model better and try to improve vertical equity results.
- This work will be has been accepted for publication in *Property Tax Assessment & Administration,* 2010 Q3

Questions
&
Discussion