

Report: Review of Automated Valuation Model (AVM) Methodology for Wales Revaluation (Report 2 of 2)

For: Valuation Office Agency, United Kingdom

By: Peadar Davis, Ulster University and Centre for Appraisal Research and Technology

Project Background

The United Kingdom's Valuation Office Agency (VOA) are developing automated valuation models (AVMs) to revalue domestic hereditaments of Wales. In support of their research and development of the AVMs, VOA sought academic critical friends with relevant expertise to provide support and assurance through a quality control review. Specifically, the critical friends were expected to:

1. Advise on statistical and modeling techniques employed by VOA to ensure appropriate use, specifically, of geographically weighted regression (GWR) and its extension for comparability weighting (GCWR), with regard to published academic research and international best practices among government property tax/valuation authorities
2. Advise on alignment of VOA modeling procedures with international standards, such as those published by the International Association of Assessing Officers (IAAO)
3. Recommend any additional statistical test and analyses VOA should incorporate in their research and modelling procedures

Paul Bidanset and Peadar Davis, PhD were asked to serve as critical friends based on their authorship of the paper "*[a]ccounting for locational, temporal, and physical similarity of residential sales in mass appraisal modeling: the development and application of geographically, temporally, and characteristically weighted regression*"¹. Project deliverables include a 2-page report by each summarizing project findings and recommendations.

Summary of Findings:

1. I would concur and strongly endorse the findings and recommendations of Paul Bidanset in the accompanying Report. Given his detailed discussion of appraisal modelling and practical Statistics package options and solutions, this report will attempt to focus more on issues of development and application to the wider challenges likely to face any subsequent policy and the implications of that on this exercise.
2. Use of a variety of cutting edge approaches (albeit within certain constraints (eg SRF restricted to certain dwelling types) has provided a deep and broad analysis which provides an excellent evidence base for policy making.
3. The analysis has been carried out to a standard that is publishable in its own right – the combination of novel, rich datasets and the expert application of robust and to an extent innovative modelling approaches would be of great interest to the wider academic and appraisal community. This is important both in terms of confidence in the process moving forward and in terms of supporting any future policy – as the level of interest and scrutiny from a wide range of interested / concerned parties would be likely to be unprecedented.
4. The caveats and considerations adopted, such as restricting application of certain models and approaches to certain property types or to certain geographical areas generally appears to be logical and appropriate to the task in hand – it is not exhaustive but arguments with regards to limits are well made.
5. Despite the complexity of the material, the communication style is very clear and easy to follow for appropriately knowledgeable persons. The summaries are very succinct and to the point and establish clearly the processes followed and the reasons for decisions taken.

¹ Bidanset, P., McCord, M., Lombard, J. A., Davis, P., & McCluskey, W. (2018). Accounting for locational, temporal, and physical similarity of residential sales in mass appraisal modeling: the development and application of geographically, temporally, and characteristically weighted regression. *Journal of Property Tax Assessment and Administration*, 14(2), 5-13.

6. The utilisation of comparable analysis is an excellent development within the broader modelling framework, with the potential to add some explanatory power, but also to link conceptually both to the rich Comp Sale AVM methodologies prevalent in North America and also perhaps more pertinently to the more traditional valuation approaches used by professional valuers and the house buying public.

Recommendations:

1. Consider some wider flexible use of comparables across type – for example, some terrace houses are market substitutes for some apartments (Flats!) across some market areas. Also, some fine tuning to delineate within type classifications may be useful (for example ‘terraces’ which are really ‘townhouses’ and probably more of a market substitute to semi or even detached houses – this may be captured by the age classification but not always).
2. Boundary effects – focus on identification / restrictions on value change across administrative/ modelling boundaries - at least highlighting changes beyond certain limits.
3. Undertake a range of cluster analysis to inform any banding exercise.
4. The link back to comparable sales is likely to prove essential in defending the AVM process. In terms of further work / approaches for when the models are deployed against a population of unsold properties - one perspective could be to reverse the comparable selection process to answer the question – where/how many properties do we *not* have a solid comparable for? This would allow a reverse engineering of ‘beacons’ and the establishment of a scoring metric to ‘red line’ properties falling out with a certain tolerance, for example. The selected comparable/s would then be associated with a population of properties for the purposes of tribunal defence. Ideally this would be geographical zones, but these need not be tessellated
5. Building on the above point, at some stage conceptualisation will need to delineate between what is useful (or perhaps essential) for a discrete valuation exercise, in contrast with what is necessary (against cost and time constraints) for a banded approach. Absolute overall accuracy is of less importance in this regard and the challenge changes - from perhaps identifying valuation changes over physical valuation boundaries to more theoretical band boundaries.
6. Some further thought given to how the methodologies vary valuations over relatively short distances for anything other than physical characteristic differences. The issue is that the methodologies can be ‘too good’ to an extent and will ascribe different values to ostensibly identical properties in close proximity. Whilst this effect would naturally be blunted in a banded system (where the AVM values are ascribed to bands), it is probably necessary in a discrete system also, as was found to be the case in the Northern Ireland residential valuation exercise.
7. These factors tend to raise the influence of robustness, efficiency (eg limiting requirement to capture new data on the unsold properties) and explainability relative to absolute statistical accuracy in terms of deciding on the final nature of the models.