Testing Land Valuation Methodologies – Lot 4: Conventional valuation approaches

Final Report

Prepared by:
Axion Solutions Ltd

On behalf of:
The Welsh Government

Authors:

Henry Sanford MEng (Hons), BA MIET

David Harding BEng (Hons) AMIMechE

Date: March 2026

Executive Summary

This report presents the findings of Axion Solutions' work to develop, test and evaluate a transparent, open-data land valuation framework for Wales. Commissioned by the Welsh Government, the project forms part of a wider programme to strengthen the national evidence base on land values as a foundation for spatial analysis, policy development and potential fiscal reform.

The central objective was to determine whether land value can be robustly estimated at small-area scale using exclusively publicly available datasets. The study demonstrates that it can - with clear boundaries and qualifications.

Two established valuation approaches were operationalised:

- A **Residual Value Model (RVM)** reflecting development-led valuation practice;
- An **Income Capitalisation Model (ICM)** capturing existing income-generating potential.

Both approaches are grounded in professional and international valuation standards, but they capture fundamentally different economic mechanisms. The RVM estimates land value as a function of development feasibility, responding primarily to Gross Development Value (GDV), build costs, abnormal risk and developer return assumptions. The ICM derives land value from rental intensity and yield capitalisation, reflecting existing economic use rather than prospective redevelopment.

Applied across nine representative LSOAs spanning urban, suburban, rural, coastal and post-industrial contexts, the models reveal a clear and coherent spatial gradient across Wales. Highest land intensities are observed in strong-demand urban and commuter markets, particularly Cardiff, followed by Bridgend and Monmouthshire. Moderate land values are evident in mixed coastal and rural–suburban areas, while lower intensities are consistently observed in upland and structurally weaker post-industrial LSOAs such as Powys and parts of Gwynedd and Rhondda Cynon Taf.

A key finding is that development-led and income-led valuation signals frequently diverge. In several LSOAs, income capitalisation produces materially higher land values than development feasibility alone would suggest. Conversely, in rural areas with limited rental concentration, both approaches converge at structurally lower levels. This confirms that land value in Wales is shaped by both forward-looking development viability and existing income density - and that neither perspective alone is sufficient.

The introduction of a hybrid (combined) framework, integrating RVM and ICM outputs, proved analytically significant. The hybrid approach reduces distortion arising from reliance on a single method, dampens volatility in high-density urban cores, and prevents development-only modelling from understating land value in income-rich markets. The resulting combined land value per m² measure produces a more stable and policy-ready land intensity gradient across Wales.

The study confirms that open-data land valuation is feasible and robust at **LSOA scale**. A fully transparent, reproducible modelling architecture has been constructed, integrating transaction evidence, EPC attributes, rateable values and spatial datasets into a harmonised workflow aligned to 2021 LSOA geography. The data pipeline is technically scalable to Wales-wide implementation.

However, the findings are equally clear about limitations. While parcel-level modelling is technically possible, valuation precision at granular intra-LSOA scale is constrained by the quality and resolution of input data. The absence of an openly licensed cadastral dataset, detailed rental evidence, granular construction cost benchmarks and site-specific planning constraints limits parcel-level accuracy. The framework performs most reliably as a comparative, small-area statistical model rather than a substitute for site-specific professional appraisal.

Model sensitivity is concentrated in a limited number of parameters - principally GDV, capitalisation rates, construction costs and developer profit assumptions. The analysis shows that increasing structural complexity without improved data inputs would risk false precision. A transparent, bounded and auditable methodology is therefore more defensible than a highly complex model built upon incomplete information.

The approach is technically scalable across Wales. The computational architecture and modelling workflow can be extended to all LSOAs using the existing open-data ecosystem. However, if higher parcel-level precision or fiscal application is envisaged, access to improved datasets beyond open sources - particularly cadastral, rental and cost data - would be necessary.

In summary, this study demonstrates that:

- A transparent and reproducible open-data land valuation framework can be delivered at LSOA scale;
- Development-led and income-led approaches provide complementary insight into land value formation;
- A hybrid framework produces more stable and policy-relevant outcomes;
- Clear spatial land-value gradients across Wales reflect genuine economic structure;
- Model robustness is strongest where data density is high;
- Granular precision is fundamentally constrained by input data quality;
- The framework is scalable but bounded by structural data limitations.

The work provides a strong analytical foundation for Wales to develop a national land-value evidence base. It establishes both what is currently achievable using open data and what further data investment would be required to enhance precision and support future policy exploration.

Contents

List of Tables

List of Figures

Glossary

Term / Abbreviation	Definition
Abnormals	Development costs arising from site-specific constraints such as contamination, flood mitigation, unstable ground, utilities reinforcement or demolition. Treated as additional items in the Residual Value Model.
AVM (Automated Valuation Model)	A statistical or rule-based model that estimates property or land values using automated data inputs and standardised algorithms. Commonly used by lenders and assessment agencies.
BCIS (Building Cost Information Service)	A proprietary RICS service providing detailed construction cost benchmarks. Not used in this project due to licensing restrictions; open-data proxies are used instead.
Cap Rate (Capitalisation Rate)	The rate used to convert Net Operating Income (NOI) into capital value in income-based valuation. Represents the required rate of return for an investor.
CIL (Community Infrastructure Levy)	A planning charge used by some local planning authorities to fund infrastructure. Not directly modelled in this project.
Contingency	A percentage uplift applied to construction costs to account for unforeseen risks. Incorporated within the RVM.
DRC (Depreciated Replacement Cost)	A method of valuing buildings based on rebuild cost minus depreciation. Used within the ICM to separate building value from land value.
EPC (Energy Performance Certificate)	A statutory document containing property attributes such as floorspace, age, built form, heating type and performance metrics. An open dataset used extensively in this project.
FRAW (Flood Risk Assessment Wales)	National flood-risk mapping published by Natural Resources Wales. Considered but not used directly in valuation modelling.
Floorspace (FS)	The internal gross or net area of a dwelling or building, typically measured in square metres. Used to derive price-per-m² and rebuild-cost estimates.

GDV (Gross Development Value)	The total market value of a completed development project, calculated in the RVM using representative unit mixes and price-per-m² assumptions.
GIS (Geographic Information System)	A system for capturing, analysing and visualising spatial data. Used to map parcels, LSOAs and contextual features.
Hedonic Model	A regression-based valuation approach that decomposes price into constituent attributes. Considered but not adopted due to open-data limitations.
HPSSA (House Price Statistics for Small Areas)	An ONS dataset providing median and distributional property prices at LSOA and MSOA levels. Used for validation and temporal smoothing.
HPI (House Price Index)	The UK's official index of residential property prices, published by HM Land Registry. Used to normalise transaction values to a common date.
IAAO (International Association of Assessing Officers)	A professional body providing global standards for property assessment, particularly income and mass-appraisal methods.
ICM (Income Capitalisation Model)	A valuation model that estimates land value by capitalising land-attributable net operating income using an appropriate yield.
INSPIRE (Infrastructure for Spatial Information in Europe)	A directive establishing pan-European spatial data standards. INSPIRE Index Polygons are used as indicative freehold extents.
Land Value	The value of land excluding buildings or improvements. In this project, estimated via RVM and ICM using open data.
LSOA (Lower-layer Super Output Area)	A statistical geography used for census and administrative reporting in Wales and England. The primary analysis unit for this project.
MSOA (Middle-layer Super Output Area)	A larger statistical geography used alongside LSOAs for validation and smoothing where transaction volumes are low.
NDR (Non-Domestic Rates)	A tax on non-domestic properties. The RV and multiplier can be used to infer market rents for commercial valuation.
NOI (Net Operating Income)	The income remaining after deducting operating expenses from gross rent. Key input to the ICM.

OPEX (Operating Expenditure)	Costs associated with running a building, such as maintenance, insurance and management. Represented as ratios by use class in the ICM.
ONSPD (ONS Postcode Directory)	A quarterly lookup file linking postcodes to LSOAs, MSOAs and coordinates. Essential for geocoding PPD and EPC data.
Open Data	Data freely available for public use without proprietary restrictions. The exclusive data basis for this project.
Parcel	In this report, a representative land unit constructed from LSOA attributes and INSPIRE polygons, not a legal cadastral unit.
PPD (Price Paid Data)	HM Land Registry transaction records containing sale price, date, property type and location. Used for market benchmarking.
Profit (Developer Profit)	The return required by a developer on a scheme's GDV. Modelled as a percentage of GDV within the RVM.
RC (Replacement Cost)	The cost to rebuild a structure as new. Formed from floorspace and rebuild-cost-per-m² assumptions within the ICM.
Residual Land Value (RLV)	Land value derived from subtracting all development costs, fees, finance and profit from GDV in the residual valuation method.
RICS (Royal Institution of Chartered Surveyors)	The professional body setting valuation, development appraisal and land assessment standards used internationally.
RVM (Residual Value Model)	A development-led valuation method deriving land value as the residual after deducting costs and profit from GDV.
RV (Rateable Value)	The assessed rental value of a non-domestic property used to calculate NDR. Used in the ICM to infer rental income.
Sensitivity Analysis	Testing model outputs under alternative assumptions (e.g., ± shifts in GDV, costs or yields) to understand uncertainty and robustness.
Site Area (ha)	The gross area of a parcel, expressed in hectares. Used in both models for land-value-per-hectare calculation.

Spatial Harmonisation	The process of aligning datasets to a consistent spatial geography (e.g., LSOA), enabling integrated analysis.
Temporal Normalisation	Adjusting historic transaction data to a common valuation date using HPI or similar indices.
Unit Mix	The assumed configuration of dwelling types and sizes used to estimate GDV in the RVM.
Yield	The required rate of return used to capitalise NOI into capital value. Used within the ICM.

Section 1 - Introduction

This report provides a detailed account of the project activity and research conducted by Axion Solutions in developing and testing two land valuation methodologies to support the Welsh Government's programme of work on understanding land values across Wales. The report documents the analytical framework established for the two models, the data sources assembled and processed to support their operation, and the results and subsequent analysis undertaken to assess how both approaches function in practice.

The purpose of this document is to present a clear and transparent overview of the technical foundation for the valuation work, present results, and explore conclusions and further considerations. This includes the assumptions adopted within each model, the data treatments required to prepare inputs for valuation at the LSOA level, and a summary of the methodological choices made to ensure that the models are replicable, consistent and capable of being scaled. The report also identifies the practical constraints encountered.

1.1 Organisational Expertise and Capability

Axion Solutions brings a strong track record in delivering technically robust, policy-relevant research for UK public bodies, with particular strengths in analytical modelling, environmental assessment, spatial data analysis and evidence-based policy development. The organisation combines strategic consultancy capability with deep methodological expertise, enabling it to deliver complex research programmes that require transparency, reproducibility and rigorous quality assurance.

The team assigned to this project offers extensive experience in land-use analysis, environmental valuation, spatial modelling and the application of open-data methods to public-sector challenges. Collectively, the team has led multi-disciplinary research projects for national and local government, arm's-length bodies and academic partners, covering themes including green and blue infrastructure, climate adaptation, natural capital, land-use planning, and public-service innovation. The organisation's internal governance structure ensures clear accountability, with defined oversight mechanisms, structured risk management and a strong focus on delivering high-quality outputs aligned to client objectives and statutory requirements.

1.2 Background to the Project

The Welsh Government has, in recent years, undertaken extensive work to examine options for modernising local taxation, including the potential role of land value within a reformed system. Research commissioned as part of this programme has highlighted the potential advantages of a more consistent, transparent and evidence-based approach to land valuation across Wales (Welsh Government, 2020). While land value forms part of existing valuation and rating processes, there is currently no unified national methodology that can isolate land value independently of the built asset or operate exclusively using publicly available data.

Developing such an approach is important for several reasons. From a policy perspective, land value is increasingly recognised as a distinct economic variable, shaped by location, accessibility, public investment and planning decisions rather than the characteristics of buildings themselves. International tax literature points to the relative stability, neutrality and fairness of land-based taxation when compared with taxes levied on buildings or productive activity (Institute for Fiscal Studies, 2022). Reform programmes across the UK have repeatedly identified the absence of transparent, spatially consistent land-value evidence as a key barrier to evaluating or implementing alternative tax or valuation systems (Scottish Land Commission, 2025).

Against this backdrop, the Welsh Government commissioned the development and testing of two established valuation approaches that are widely recognised in professional and academic practice. The first is a residual valuation method, commonly used in development appraisal, where land value is derived from the balance remaining after deducting construction costs, professional fees, abnormal allowances, finance costs and an appropriate developer profit from the Gross Development Value of a notional scheme (RICS, 2019). This method reflects standard practice in assessing development viability and is well suited to areas where development potential is a key driver of land value.

The second approach is an income capitalisation method typically used for non-domestic and commercial land. Under this method, land (and, where relevant, buildings) is valued by estimating the net operating income generated by the property and dividing this by an appropriate capitalisation rate, with adjustments made to separate the return on land from the return on improvements. This approach aligns with income-based valuation principles used in international assessment standards and commercial property markets (International Association of Assessing Officers, 2017).

The objective of this project is to operationalise both approaches using only open data and to assess how they perform when applied across diverse geographies in Wales. The project includes LSOAs representing rural, urban, coastal and post-industrial contexts, enabling practical evaluation of methodological feasibility, data adequacy and spatial consistency. This project provides a foundation for assessing scalability, identifying limitations and informing the development of a reproducible, Wales-wide land valuation framework.

1.3 International Examples of Valuing Land

A number of jurisdictions have developed systematic approaches to valuing land as a distinct tax base, reflecting long-standing practices in land value taxation, mass appraisal, and property assessment. Although institutional frameworks differ between countries, international experience provides useful insight into the feasibility, benefits and challenges of large-scale land valuation using standardised methodologies and, increasingly, open data. The following examples illustrate relevant approaches.

New Zealand: National Land Value System (Rating Valuations)

New Zealand operates one of the most mature and transparent land valuation systems globally. Local authorities are required to carry out triennial mass valuations that separately estimate the value of land and improvements for every property. Valuation methods rely heavily on sales analysis, mass appraisal techniques and spatially consistent modelling, supported by a comprehensive national data infrastructure. The system demonstrates that national-scale land valuation is technically achievable and can be delivered using modernised data governance and modelling practices (New Zealand Valuer-General, 2021).

Denmark: Statutory Land Value Assessments (Ejendomsvurdering)

Denmark operates a statutory land valuation framework in which the taxable value of land is assessed independently of buildings. The Danish Property Assessment Agency uses mass appraisal techniques, spatial modelling and standardised market evidence to estimate land value at the parcel level. In recent reforms, Denmark has modernised its valuation system to improve transparency and public confidence, including open documentation of methodological assumptions and data sources (Danish Ministry of Taxation, 2020).

Estonia: Land Tax Valuation Model

Estonia operates a nationwide land valuation system used for municipal land taxation. Land value is assessed based on permitted land use, location, environmental characteristics and comparable transactions. Estonia’s approach is particularly relevant to Wales because it relies heavily on open geospatial datasets and automated modelling to achieve national coverage with limited manual valuation effort (Estonian Land Board, 2019).

United States: Land Residual and Income-Based Approaches

In the United States, several states and municipalities use mass appraisal models that explicitly separate land and improvements. Although most US jurisdictions tax property value rather than land alone, the International Association of Assessing Officers (IAAO) has long promoted the income, cost and sales-comparison approaches for deriving land value, and many county assessors use residual land valuation for development land or income capitalisation methods for commercial land. These methods are codified, widely used and align closely with the methods being tested in Wales (IAAO, 2017).

Australia: State Land Valuation Frameworks

Australian states such as New South Wales, Victoria and Queensland operate comprehensive land valuation systems for land tax and rating. Land is valued annually or biennially, typically using mass appraisal techniques combining sales analysis, spatial modelling and standardised valuation rules. The systems operate at very large scale and demonstrate the practicality of regular national land valuation cycles supported by open geographic data and transparent quality-assurance procedures (NSW Valuer General, 2022).

Scotland: Review of Land Value Approaches

Although Scotland has not implemented a land value tax, the Scottish Land Commission has examined international evidence and identified strong potential benefits from transparent, standardised approaches to land value modelling, emphasising the importance of open-data infrastructure and clear methodological documentation (Scottish Land Commission, 2020). This research highlights the relevance of reproducible, open-data-based approaches such as those being tested in Wales.

These international examples collectively demonstrate some interesting points:

- Land valuation at national scale is achievable and widely implemented;
- Mass appraisal techniques, including residual and income-based methods, are internationally standard practice;
- Open spatial data, consistent market evidence and transparent assumptions are essential for public confidence;
- The separation of land and improvement value is feasible and already routine in many systems.

The approaches adopted in this project – residual valuation and income capitalisation using open data – are therefore aligned with internationally recognised valuation practices and reflect the methodologies used in mature land valuation systems.

1.4 Scope and Structure of the Report

This report sets out the methodological and analytical work completed by Axion Solutions in developing the two valuation approaches. It provides an account of the modelling logic, data processing requirements and computational steps underpinning both methods, and summarises the results of tests undertaken to assess data coverage, modelling behaviour and the influence of key variables.

Section 2 presents a focused review of relevant literature and valuation practice relating to residual and income-based methods, limited to material that directly informed the design of the two approaches. This is followed by a full explanation of the methodology, including data architecture, construction of parcel- and LSOA-level inputs, and the mathematical formulation of the models. Section 3 summarises analytical outputs and findings. Sections 4 and 5 explore conclusions and the technical and practical considerations which warrant further investigation; including matters relating to data availability, feasibility, scalability and reproducibility.

1.5 Aims and Objectives

The aim of this work is to apply and test two established land valuation approaches in order to assess their practical suitability for producing consistent and transparent land value estimates across a selection of LSOAs in Wales using publicly accessible data and clearly documented modelling processes.

The objectives of the project are to:

1. Develop operational and replicable valuation models for both the residual valuation method and the income capitalisation method, ensuring that each is technically robust and suitable for application at the required spatial scale.
2. Assemble, prepare and integrate relevant open-source datasets, including transaction data, building and land-use information, dwelling mix statistics and non-domestic rating data, to support the operation of both models.
3. Apply the valuation models to the selected LSOAs to understand how each method performs across varied geographical and market contexts, and to assess sensitivity to key assumptions and data inputs.
4. Evaluate the transparency, comparability and scalability of the two approaches, considering how valuation outputs can be interpreted and how each method might function if extended to wider geographical coverage.
5. Identify methodological, operational and data-related considerations arising from model development and early testing, providing a basis for refining the valuation processes and completing the full analysis.

These aims and objectives guide the structure of the report and the technical work undertaken, supporting the Welsh Government's intention to better understand the practical application of land valuation methodologies in a Welsh context.

Section 2 – Methodology (Lot 4: Conventional Valuation Approaches)

The methodology for this project has been designed to develop, document and test two established valuation approaches using only publicly accessible datasets and clear, auditable modelling processes. The objective is to apply both methods consistently across a diverse set of Welsh LSOAs and assess their suitability for producing transparent and reproducible estimates of land value. The methodology integrates three components: data architecture, parcel and LSOA selection logic, and model formulation, into a unified analytical framework.

This section sets out the conceptual basis for the valuation approaches, the rationale for selecting the two models used in this project, and the way in which each method is operationalised. The focus is on ensuring that the methods can be applied systematically across a wide range of geographical, land-use and market contexts, while maintaining consistency with established valuation standards and ensuring clarity in how assumptions are defined, implemented and tested.

2.1 Literature Review

2.1.1 Key Literature Influencing Methodology

Research into land valuation methods provides several established foundations for the approaches adopted in this project. A central contribution is the comparative study by Bourassa and Hoesli (2022), which evaluates hedonic, residual and matching methods for estimating residential land value. Their findings emphasise that hedonic regression and residual approaches are the most operationally feasible where building-level data are available, while matching techniques, although accurate, require datasets rarely available at scale. This supports the adoption of a dual-method strategy in this project, combining residual and income-based methods where appropriate.

Work undertaken earlier in Wales as part of considerations for reforming local government finance also highlighted the conceptual challenges of separating land value from property value and the practical limitations in applying theoretical valuation approaches at the national level (Welsh Government, 2020). That analysis identified the lack of harmonised property and land-use datasets as a barrier to implementing valuation models consistently across LSOAs, a point echoed in wider UK appraisal literature.

International guidance has played an important role in shaping the methodology. The International Association of Assessing Officers (IAAO, 2017) provides detailed guidance on mass appraisal, income-capitalisation techniques and methods for distinguishing land from improvements. Similarly, RICS guidance on valuation and development appraisal sets out the accepted principles for residual valuation, including the treatment of developer return, finance and construction costs (RICS, 2020). These frameworks informed the structure of the residual valuation model used in this project, ensuring alignment with professional standards.

Beyond traditional valuation literature, emerging work on data-driven valuation provides insights into the opportunities and limitations of mass appraisal. Kolbe et al. (2019) demonstrate how integrated transaction, structural and spatial datasets can be used to estimate land values statistically, but highlight the dependency on high-quality, parcel-level data. Similar conclusions are drawn in international studies applying income-capitalisation approaches to urban land (Sugito et al., 2024), where the accuracy of results is shown to depend heavily on the availability of reliable income and yield data.

Recent work in machine-learning and property valuation (Tanashkina et al., 2025) provides additional context on the role of data availability and the trade-off between interpretability and predictive accuracy. Although advanced modelling techniques show promise, the literature stresses the need for transparency and explainability - requirements that align with the present project's emphasis on methods that can be audited, replicated and understood by non-technical audiences.

2.1.2 Gaps Identified in Prior Research

The reviewed literature identifies several persistent gaps that directly influence the methodological design of this project.

A recurring issue is the absence of integrated, parcel-level datasets that consistently link land characteristics, building attributes, transaction evidence and income information. While some datasets exist in isolation, they are often incomplete or poorly harmonised, which restricts the applicability of methods such as hedonic regression, matching or detailed residual appraisal at scale.

There is also limited research addressing the valuation of non-residential land using publicly available data. Much of the existing literature focuses on residential land or vacant sites, resulting in a lack of established methods for income-based valuation using open-source datasets for industrial, retail or office properties. This gap complicates attempts to apply income-capitalisation models consistently across mixed-use geographies.

A further limitation concerns the reproducibility of valuation methods in data-sparse environments. Several studies highlight how subjective input assumptions - particularly yields, depreciation rates, cost bases and developer returns - can materially affect valuation outcomes, reducing comparability across different locations or valuation exercises. This reinforces the need for clear documentation and transparent assumptions.

2.1.3 Implications for the Methodology

The methodology adopted in this project responds directly to the strengths and limitations identified in the literature.

Firstly, the dual approach - combining a residual valuation method with an income-capitalisation method - reflects evidence that no single model provides comprehensive coverage across land types. By applying both approaches, the project can generate valuation outputs that reflect the diversity of land use patterns across Wales.

Secondly, the development of a multi-source, harmonised dataset is necessary to mitigate the data-integration challenges highlighted in previous research. The methodology therefore integrates transaction data, building footprints, land-use classifications, rating data and environmental constraints into a consistent LSOA-level data architecture. This maximises the utility of available open data while maintaining transparency and reproducibility.

Thirdly, the modelling approach prioritises explainability and traceability, consistent with recommendations in valuation literature that highlight the risks associated with opaque methods or poorly justified assumptions. Where critical parameters - such as developer return - cannot be inferred directly from data, the assumptions are documented clearly to support later refinement and sensitivity analysis.

Taken together, these implications underpin a methodological framework that is transparent, replicable and suited to the practical realities of assessing land value across a representative set of Welsh LSOAs.

2.1.4 The Welsh Context and Diversity Across the Study LSOAs

Wales exhibits substantial spatial diversity that directly affects land markets and the feasibility of applying valuation models consistently across different areas. The ten LSOAs selected for this study represent a cross-section of Welsh settlement types, property markets and physical environments. Understanding this diversity is essential for interpreting valuation behaviour and selecting appropriate methodological assumptions. The following summary draws on evidence published by the Office for National Statistics (ONS), StatsWales, Welsh Government regional analyses and local authority evidence bases.

W01000255 – Flintshire 015A (North Wales)
Flintshire 015A presents a moderately populated, semi-urban context characterised by a mix of inter-war housing, post-war estate development and later suburban infill. Located in North Wales close to major employment corridors and market towns, the area typically benefits from good strategic connectivity while retaining a largely residential character. Settlement density is moderate, with established neighbourhood layouts and a blend of detached, semi-detached and terraced properties. Local services appear accessible, and environmental constraints are present but generally manageable, with no indication of extensive floodplain or protected-site encroachment.

The LSOA’s position within a relatively dynamic housing market provides a well-balanced test environment, with meaningful sales volumes, moderate EPC coverage and a stable mix of dwelling types. As a semi-urban area situated within economic commuting distance of Cheshire and Deeside, Flintshire 015A offers insight into how the valuation models respond to mixed market signals, transitional land-use patterns and development pressures typical of North Wales’ more populated communities.

W01000114 – Gwynedd 009D (Rural North West Wales)
Gwynedd 009D represents a sparsely populated rural LSOA situated within the upland landscapes characteristic of inland North-West Wales. Housing stock is typically dispersed, comprising older stone-built properties, small clusters of cottages

and farm-based dwellings. Settlement density is very low, with large tracts of agricultural land and open countryside defining the spatial character. Access to services, public transport and key employment centres is limited, which is reflected in socio-economic indicators showing pockets of deprivation linked primarily to remoteness rather than urban disadvantage.

The rural character of Gwynedd 009D makes it a highly distinctive modelling environment. Transaction volumes are low, dwelling types are heterogeneous, and EPC coverage is uneven, all of which influence the stability and comparability of valuation outputs. The area provides an important benchmark for understanding how the models behave when market evidence is thin, development potential is constrained and land-use patterns differ markedly from more urban or coastal LSOAs. This LSOA is particularly valuable for evaluating the adaptability of open-data valuation methods in sparsely settled regions.

W01001597 – Monmouthshire 006F (South East Wales)
Monmouthshire 006F typifies an affluent rural–commuter locality in South East Wales, with a settlement profile shaped by proximity to the Severn corridor and market-town catchments. Housing stock is predominantly owner-occupied, with above-average incomes, higher-value detached dwellings and modern suburban extensions. The local environment is characterised by rolling countryside, strong landscape appeal and significant planning constraints related to conservation and green-belt-style protections. Access to road networks is generally strong, reinforcing the area's commuter-oriented socio-economic profile.

This LSOA provides a valuable counterpoint to post-industrial and remote rural areas. Sales values tend to be higher, supply relatively constrained, and development potential closely shaped by environmental and planning designations. The area therefore tests how the RVM responds to high GDV environments with elevated land quality, and how the ICM behaves in localities where commercial activity is limited but residential attractiveness is high. Monmouthshire 006F illustrates how land value models operate in mature, high-demand rural markets where landscape and planning context strongly influence valuation outcomes.

W01000449 – Powys 011C (Mid Wales)
Powys 011C exemplifies the sparsely populated agricultural core of Mid Wales. It is characterised by extensive farmland, dispersed hamlets, and a predominance of older detached and semi-rural dwellings. Service accessibility is limited, and travel distances to employment, health and retail centres are substantial. Density is extremely low, and the physical geography includes upland slopes, rivers and rolling countryside, with occasional floodplain interaction along minor watercourses. The area's economic profile is shaped largely by agriculture and primary industries.

Such rural conditions make Powys 011C one of the most demanding LSOAs for open-data modelling. Transaction evidence is intermittent, EPC coverage is inconsistent, and housing stock varies substantially in age, construction type and plot configuration. This LSOA provides an essential environment for stress-testing the models' ability to generate stable and meaningful outputs where market signals are weak and development viability is often marginal. Its inclusion strengthens the overall robustness and generalisability of the valuation methodology.

W01000617 – Pembrokeshire 002F (South West Wales)
Pembrokeshire 002F represents a coastal rural LSOA with a blend of permanent residential communities, tourism-linked housing, and older rural settlement structures. Housing density is low to moderate, with mixed-age dwellings set within a diverse landscape of coastal plains, farmland and proximity to designated environmental areas. Seasonal economic patterns influence housing demand and occupancy, and parts of the coastline may be subject to flood-risk and coastal-erosion considerations. Access to services can be variable, reflecting the dispersed rural geography typical of coastal Pembrokeshire.

This LSOA is significant for understanding how the models accommodate coastal constraints, variable land-use pressures and seasonal economic influences. Both RVM and ICM may encounter distinctive challenges here: GDV estimation may be impacted by the presence of holiday-let markets and atypical dwelling attributes, while income-based valuation may be influenced by low local commercial activity. Pembrokeshire 002F enriches the cross-section of test environments by combining rural characteristics with coastal pressures and conservation considerations.

W01001233 – Rhondda Cynon Taf 001F (Post-Industrial Valleys)
Rhondda Cynon Taf 001F is a post-industrial valley LSOA characterised by dense terraced housing, steep valley topography and strong historical links to coal mining and industrial employment. Settlement patterns follow the linear structure typical of South Wales Valleys communities, with housing concentrated along transport corridors and limited developable land on steep valley sides. Socio-economic indicators tend to reflect lower income levels, greater deprivation in certain domains and reduced access to services relative to more urbanised locations.

This LSOA provides a critical test of the models in areas where land value is strongly shaped by topographical constraints, legacy industrial land uses and lower market demand. Residential transaction prices are often modest, development opportunities constrained, and abnormal costs can be high due to topography or historic land conditions. Rhondda Cynon Taf 001F offers essential insight into how open-data valuation methods handle post-industrial landscapes and how cost-led and income-led approaches diverge under conditions of constrained viability.

W01002019 – Cardiff 032H (Capital City Urban LSOA)
Cardiff 032H reflects a highly urbanised setting within the capital city's established residential fabric. The LSOA is characterised by dense housing, significant rental markets, strong public transport provision and close proximity to employment hubs, services and amenities. Housing stock includes a mix of terraces, flats and subdivided Victorian or Edwardian properties, with high EPC coverage and strong transaction activity. Urban infrastructure, public realm investment and ongoing regeneration activity contribute to robust market dynamics.

As the most urbanised test area, Cardiff 032H provides a benchmark for valuation performance under high-density, high-turnover conditions. Strong market evidence supports more stable GDV estimation, while income-based valuation benefits from richer rental and commercial signals. This LSOA helps assess how the modelling framework reflects urban land value drivers and offers a point of comparison for understanding valuation gradients across Wales.

W01000517 – Ceredigion 002D (Rural/Coastal Mid Wales)
Ceredigion 002D provides a coastal rural context combining agricultural land, small settlements and proximity to sensitive natural environments. The area typically includes a mixture of traditional cottages, farm dwellings and low-density residential clusters, influenced by both rural economic activity and coastal amenity value. Flood risk and conservation designations may shape development potential in parts of the LSOA, and service accessibility varies substantially across the area.

This LSOA contributes important diversity to the test set. Coastal constraints, environmental designations and seasonal economic influences interact with thin transaction evidence and heterogeneous housing stock, making valuation particularly challenging. Ceredigion 002D allows for assessment of how open-data valuation methods operate in settings where land value is tied to landscape quality, environmental restrictions and limited development capacity.

W01001045 – Bridgend 019D (South Wales Growth Area)
Bridgend 019D is a suburban LSOA situated on the South Wales coastal plain, displaying moderate residential density and a mixture of older estates and more contemporary suburban development. The area benefits from strong connectivity via the M4 corridor, access to frequent public transport services and proximity to Bridgend town centre. Housing stock is varied and predominantly owner-occupied, reflecting stable suburban market conditions, with moderate deprivation levels and relatively consistent EPC and transaction coverage.

Bridgend 019D serves as a transitional case between the fully urban character of Cardiff and the more rural characteristics of Powys and Ceredigion. Its suburban profile provides a valuable midpoint for evaluating model behaviour, illustrating how residual and income-based valuations respond to moderate development pressure, balanced amenities and accessible transport infrastructure. This LSOA contributes significantly to assessing scalability and comparability across diverse settlement forms in Wales.

This spatial diversity has clear implications for selecting and testing valuation methods. Urban areas such as Cardiff provide more reliable data for income-based valuation due to richer rental evidence, while peri-urban and growth-area LSOAs such as Flintshire and Bridgend provide balanced datasets that support residual valuation. Deeply rural and post-industrial LSOAs Gwynedd, Powys, Pembrokeshire and Rhondda Cynon Taf pose challenges relating to low sales volumes, atypical dwelling stock, constrained geography and complex ground conditions, all of which affect both land value estimation and the stability of modelling assumptions.

These variations reinforce the importance of a dual-method approach and a harmonised open-data architecture capable of accommodating substantial differences in data availability, land-use structure and market conditions across Wales.

2.1.5 Data Sources Used

The following datasets form the core analytical foundation for the modelling process. Each dataset listed below was used directly in constructing variables, calibrating assumptions, or spatially contextualising valuation logic.

Market and Valuation Data

HM Land Registry – Price Paid Data (PPD)
HM Land Registry's Price Paid Data (PPD) provided the primary source of open-market residential transaction evidence used in this study (HM Land Registry, 2025a). The dataset contains information on sale price, transaction date, property type, new-build status and address identifiers, and formed the basis for deriving Gross Development Value benchmarks for the Residual Value Model. Prior to integration, the data were cleaned to remove duplicate entries, corporate bulk transfers and non–arm's-length transactions to ensure only representative market sales were retained. Postcode-to-LSOA lookup tables were used to assign transactions to 2021 LSOA boundaries, acknowledging that postcode-level geocoding introduces some spatial imprecision and that PPD contains no internal property attributes such as floorspace.

HM Land Registry – UK House Price Index (HPI)
The UK House Price Index (HPI) was used to adjust historical transaction values to a consistent valuation date (HM Land Registry, 2025b). The index supports temporal comparability across Local Super Output Areas, although its publication at Wales and regional level means it must be applied cautiously in small-area valuation.

Welsh Revenue Authority – Land Transaction Tax (LTT) Statistics
Land Transaction Tax statistics provided contextual information on the distribution and mix of property transactions within local authority areas (Welsh Revenue Authority, 2024). These data were used to validate the representativeness of PPD transaction patterns. Their aggregated spatial scale, however, limits their direct application within parcel-level valuation models.

StatsWales – Private Sector Rents
Private sector rent statistics supplied by StatsWales were used to support rental value assumptions within the Income Capitalisation Model (StatsWales, 2024a). Although these data help benchmark achievable rent levels, their coverage is limited to the 2019 release and selected ad hoc updates in 2024, reducing their suitability for time-series analysis.

ONS – Price Index of Private Rents (PIPR)
The Price Index of Private Rents provided a macro-level indicator of rental inflation across Wales (ONS, 2024a). It was used to contextualise rental market changes over time but was not applied directly in LSOA-scale modelling due to its aggregated spatial resolution.

ONS – House Price Statistics for Small Areas (HPSSA)

House Price Statistics for Small Areas (HPSSA) supplied annual median and distributional sales values at LSOA level (ONS, 2024b). These data were used to validate PPD-derived benchmarks and to provide temporal smoothing in LSOAs with limited transaction volumes. The annual update cycle limits its utility for capturing short-term market dynamics.

Spatial and Geographical Reference Data

ONS – LSOA 2021 Boundaries
The 2021 LSOA boundary dataset served as the unifying spatial framework for integrating all address-level and contextual datasets (ONS, 2022). PPD, EPC, census and contextual data were aligned to these boundaries using centroid-in-polygon or postcode-to-LSOA matching to ensure consistency across all stages of analysis.

HM Land Registry – INSPIRE Index Polygons
INSPIRE Index Polygons were used to approximate freehold extents and infer representative parcel footprints (HM Land Registry, 2025c). These polygons provided an important spatial reference in the absence of a complete cadastral dataset for Wales, although legal title boundaries do not always align with built-form footprints or development parcels and must therefore be interpreted with caution.

OS OpenMap Local and OS Open Zoomstack
Ordnance Survey's OpenMap Local and Open Zoomstack datasets were used to provide contextual spatial information including building footprints, road networks and settlement structure (Ordnance Survey, 2024a; Ordnance Survey, 2024b). These layers supported interpretation of urban morphology when developing representative parcel assumptions, although their generalised geometry limited their use as direct valuation inputs.

DataMap Wales
DataMap Wales provided access to a wide range of Welsh Government spatial datasets used to corroborate settlement form, environmental context and LSOA-level spatial characteristics (Welsh Government, 2025). These datasets supported contextual understanding rather than supplying quantitative valuation inputs.

EPC Open Data (England and Wales)
EPC Open Data contributed essential property attributes, including floorspace, age band and built form, enabling approximate price-per-square-metre indicators to be derived where PPD lacked floorspace information (DLUHC, 2025). The dataset was also used to understand dwelling size distributions within each LSOA. Its positional anonymisation and variation in coverage across property types were recognised limitations.

Census 2021 – Housing Variables
Census 2021 housing variables provided detailed information on dwelling types, tenure, occupancy patterns and structural characteristics (ONS, 2023). These data supported the construction of representative parcel assumptions and contextualised local housing stock profiles, although no valuation-related information is included.

LSOA Data (Doogal)
Doogal postcode-to-LSOA lookup tables were used to support spatial referencing where datasets lacked direct LSOA identifiers (Doogal, 2025). While not an official statistical product, the resource is widely used and accurate for geocoding purposes and was applied solely to support consistent spatial harmonisation.

ONS Postcode Directory (ONSPD)
The ONS Postcode Directory provides quarterly postcode-to-LSOA, MSOA and local authority lookups, including geocodes and positional quality indicators (ONS, 2024c). It offers a stable geocoding framework for joining otherwise unaligned datasets and is particularly useful for EPC records, rental datasets and administrative statistics that lack direct LSOA coding. ONSPD is robust, regularly updated and available under OGL terms.

2.1.6 Evaluation of Additional Data Sources Considered

OS Open Greenspace
OS Open Greenspace contains polygons for public open spaces, parks, sports facilities and allotments (Ordnance Survey, 2024c). These data could be used as amenity or accessibility covariates in hedonic valuation models. However, as the present study focuses on testing open-data feasibility for residual and income-based valuation methods - rather than hedonic modelling - the dataset was not incorporated into the modelling workflow. It remains a potentially useful contextual enhancement for future iterations.

WIMD 2019
The Welsh Index of Multiple Deprivation offers small-area measures of socio-economic deprivation across multiple domains (Welsh Government, 2019). Although WIMD provides valuable contextual information, especially in relation to accessibility and service provision, it was not incorporated into the RVM or ICM because these methods rely on market, income and cost relationships rather than socio-economic conditions. Its use would move the models towards hedonic price modelling, which falls outside the methodological scope for this phase.

Flood Risk Assessment Wales (FRAW)
The FRAW dataset provides detailed spatial layers for fluvial, tidal and surface water flood risk, including defended-area classifications (Natural Resources Wales, 2024a). These data are relevant for identifying abnormal development costs or feasibility constraints. However, integrating flood risk into an open-data residual valuation framework would require parcel-level risk scoring and cost uplifts, which cannot be derived reliably without a cadastral boundary dataset. The dataset was therefore not included in the current modelling phase.

Flood Map for Planning (TAN15)
TAN15 flood zones offer planning-specific classification of flood risk areas (Natural Resources Wales, 2024b). While highly relevant for development appraisal in practice, applying TAN15 designations requires site-specific parcel boundaries and development scoping. As the project instead uses representative parcel typologies rather than site-specific proposals, these zones were not applied.

Protected Sites (SSSI, SAC, SPA)
Protected site boundaries published by Natural Resources Wales and the Joint Nature Conservation Committee could contribute to constraint mapping (NRW/JNCC, 2024). However, without a corresponding parcel-level land and property cadastre, the geographic intersection with representative parcels cannot be meaningfully established. Incorporating these data would therefore introduce false precision and did not align with the open-data feasibility goals of the project.

Coal Authority – Mine Entries and Mining Risk Data
The Coal Authority dataset includes mine entry points and risk zones across Wales (Coal Authority, 2024). It is important for ground risk assessment and foundation abnormal costs in development appraisal. However, its value is contingent on parcel-level precision. Since the project operates at LSOA scale using representative parcels, rather than actual developable sites, mining risk data were not incorporated.

Predictive Agricultural Land Classification (ALC)
ALC grading identifies best and most versatile agricultural land and planning sensitivities (Welsh Government, 2024a). This classification would be useful when valuing edge-of-settlement or rural development land. The selected LSOAs in this study, however, primarily represent built environments and do not require agricultural land classification for valuation of residential parcels. As such, the dataset was not relevant to the modelling objectives.

LiDAR (Wales)
LiDAR digital terrain and surface models provide high-resolution elevation and topographic data (Natural Resources Wales, 2024c). These data could support analysis of slope, flood interaction or building-height estimation. However, the processing burden and lack of parcel specificity meant that LiDAR was not included in this feasibility-focused modelling phase.

BGS OpenGeoscience (Geology)
Geological layers for bedrock and superficial deposits could theoretically inform ground-risk uplifts in residual appraisal (BGS, 2024). Yet, as with mining data, their integration requires parcel-level specificity. Given the representative nature of parcels used in this study, geology was excluded to avoid implying a level of spatial accuracy not supported by the data model.

NaPTAN (DfT)
NaPTAN provides a national dataset of transport nodes, including bus stops, rail stations and ferry terminals (DfT, 2024). These data are frequently used for accessibility assessments in hedonic modelling. As the RVM and ICM do not incorporate accessibility variables and to ensure the modelling remains focused on valuation fundamentals rather than hedonic characteristics - NaPTAN was not included.

TrawsCymru / Transport for Wales (GTFS/TXC)
Timetable and route datasets in GTFS/TXC format allow calculation of public transport accessibility metrics (TfW, 2024). Such indicators are again relevant for

hedonic modelling but not for residual or income-capitalisation methods, and therefore were excluded.

Ofcom Connected Nations (Fixed Broadband and Mobile Coverage)
Ofcom's connectivity datasets provide spatial data on gigabit broadband, full-fibre availability and mobile network coverage (Ofcom, 2024). These variables can influence residential desirability but do not impact RVM or ICM calculations. They were therefore excluded from the modelling workflow.

ONS – Rental Affordability and PIPR (Extended Series)
ONS affordability and rent-inflation datasets offer national indicators of rental market conditions (ONS, 2024d). These would primarily enhance economic context and yield cross-checking rather than serve as direct modelling inputs. Since the core dataset (PIPR) was already used at contextual level, the extended affordability series were not included.

VOA Rating List and Summary Valuations
The VOA non-domestic rating list provides rental value proxies for commercial property (VOA, 2024). However, the licensing regime is not OGL, meaning the dataset cannot be used in an open-data valuation model. For this reason, it was excluded.

Police.uk Crime Data
Police.uk provides street-level crime and outcomes data that could act as a neighbourhood disamenity variable in hedonic price models (Home Office, 2024). As the project does not incorporate hedonic modelling techniques, crime data were not incorporated.

Air Quality in Wales – Open Data
Near-real-time air quality readings, together with pollutants (NO_2, $PM_{2.5}$ etc.), were considered (Air Quality Wales, 2024). These datasets are more relevant for environmental quality assessments and hedonic pricing studies and were therefore excluded from the valuation models.

UKCEH Land Cover Map – Open Statistics
The Land Cover Map provides national land-cover summaries and urban/rural classifications (UKCEH, 2024). While useful for contextual analysis, the dataset was not required for RVM or ICM inputs and was therefore excluded.

2.2 Conceptual Basis for Valuation

2.2.1 Site Value and the Distinction Between Land and Improvements

The valuation methods adopted in this project are grounded in the principle that the value of land can be separated from the value of structural improvements. In valuation theory, site value is defined as the value of land in its current serviced condition, reflecting its location, planning status, accessibility, infrastructure connections, and environmental constraints, but excluding the value of buildings, structures or other physical improvements (IAAO, 2017; RICS, 2020).

The distinction can be expressed formally as:

$$V_P = V_L + V_B$$

where:
V_P= market value of the property,
V_L= value of the land (site value),
V_B= value of buildings and improvements.

The practical challenge in land valuation is that land is rarely bought or sold separately from the property built upon it. Therefore, both valuation approaches operationalised in this project are designed to estimate $V_L$indirectly by accounting for the contribution of improvements to the overall property value.

In the residual valuation method, the land value is derived as the residual after deducting development-related costs and returns from the expected completed development value. In the income capitalisation method, the land value is derived as the capitalised value of the portion of net operating income that can be attributed to land rather than improvements.

These two approaches represent the principal internationally recognised methods for distinguishing land value from improvement value.

2.2.2 Capital and Rental Bases of Valuation

The capital value basis and the rental value basis represent two complementary approaches for estimating land value.

Capital Value Basis (Residual Valuation)
Under a capital value basis, the value of land is derived from the potential Gross Development Value (GDV) of a scheme, less all development, finance and return requirements. This method is widely used for residential and mixed-use development appraisal and reflects market practice where the end-value of a completed scheme is the primary determinant of land viability (Crosby et al., 2018).

Formally, residual land value is:

$$V_L = GDV - (C_{build} + C_{infra} + C_{fees} + C_{finance} + C_{abnormals} + \pi)$$

where:
C_{build}= construction cost,
C_{infra}= infrastructure and external works,
C_{fees}= professional fees,
$C_{finance}$= finance costs,
$C_{abnormals}$= abnormal or site-specific costs,
π= developer return or profit.

This formulation requires estimation of GDV from market evidence and detailed modelling of costs using open sources such as BCIS benchmarks, publicly available planning data, and environmental datasets.

Rental Value Basis (Income Capitalisation)
The income capitalisation method derives land value from the income-generating capacity of the site. This approach is widely used for non-domestic and commercial land, where tenancy, rent, and operating data provide a more consistent basis for valuation than sales evidence (IAAO, 2017).

The method is structured as follows:

$$NOI_{total} = R \times (1 - OER)$$
$$NOI_{land} = NOI_{total} - (V_B \times r)$$
$$V_L = \frac{NOI_{land}}{Y_L}$$

where:
R= annual rent or rental proxy (e.g. derived from Rateable Value),
OER= operating expense ratio,
V_B= depreciated building value,
r= required return on improvements,
Y_L= capitalisation rate applicable to land.

This method is particularly useful in urban and employment-dense areas where business rates and rental evidence offer clear indicators of economic activity.

2.2.3 Rationale for the Valuation Approaches Adopted

This project adopts two established approaches to valuing land: the Residual Value Model (RVM) and the Income Capitalisation Model (ICM). These methods were selected following an assessment of their technical robustness, compatibility with open datasets, and suitability for the full spectrum of Welsh land-use contexts.

Residual valuation is the standard method for assessing development land, particularly in residential and mixed-use contexts where land is valued on the basis of its capacity to support a viable development. Its formulation is explicit and auditable. Land value is derived by deducting construction costs, fees, abnormal allowances, finance, and developer profit from the Gross Development Value of a notional scheme. This approach is well aligned to the Welsh context, where direct evidence of land-only transactions is limited and where much of the housing market is characterised by second-hand sales rather than new-build transactions. The RVM therefore provides a transparent means of inferring land value from observable market activity, even in areas where market signals are thin.

The income capitalisation approach is widely used for non-domestic and income-producing property, reflecting the principle that land value can be inferred from the income it generates. By estimating net operating income and applying an appropriate capitalisation rate, this method enables the separation of land value from the value of improvements. The availability of Rateable Value as a publicly accessible, standardised proxy for rental income makes the ICM particularly suitable for Wales, where commercial rental data are otherwise limited. It is especially relevant for the more urbanised study areas and for locations with active commercial markets, while also offering insight into the value of employment land in mixed-use environments.

Taken together, the two methods form a complementary framework capable of responding to the diverse geographical and economic conditions across the selected LSOAs. The dual-model approach enhances methodological resilience by enabling land value to be inferred from both development viability and income-based perspectives. This allows the project to test how well open-data valuation techniques perform in contrasting environments - urban and rural, coastal and post-industrial - and provides a strong basis for evaluating the feasibility, scalability and transparency of open-data-driven land valuation across Wales.

2.3 Data Architecture

2.3.1 Overview of the Data Strategy

The data architecture for this project has been designed to ensure that both valuation are developed using open, reproducible, and transparently documented datasets. The emphasis on open-source data aligns with the core project requirement to test valuation methods that could be scaled Wales-wide without reliance on proprietary inputs.

The architecture is underpinned by five principles:

1. **Open-data compliance**: All datasets used are licensed under Open Government Licence (OGL), Creative Commons, or otherwise freely accessible.
2. **Reproducibility**: All transformations, spatial joins, and temporal adjustments are consistently documented to allow re-execution.
3. **Spatial harmonisation**: All data are referenced to the ONS 2021 LSOA boundaries, providing a consistent analytical geography.
4. **Traceability and auditability**: Each variable used in the models can be traced to a source dataset and a defined transformation step.
5. **Valuation standards alignment**: The structure ensures that inputs for RVM and ICM adhere to professional expectations for transparency in accordance with RICS Red Book Global Standards and IAAO mass appraisal guidance.

These principles ensure that the RVM and ICM can be evaluated not only on their technical performance but also on their feasibility for operational use by public bodies in Wales.

2.3.2 Dataset Integration

The integration of heterogeneous open datasets required a structured and standardised approach to ensure analytical coherence across all valuation inputs.

Because no individual dataset contained all attributes required for the Residual Value Model or the Income Capitalisation Model, a multi-stage processing pipeline was implemented to harmonise geographies, reconcile address-level data and maintain methodological transparency.

All datasets were first ingested in their native formats - including CSV, GeoPackage and Shapefile - and converted into consistent schemas. Spatial datasets were reprojected to the British National Grid (EPSG:27700), ensuring uniformity across all sources. Postcode-level datasets such as PPD and EPC were then matched to LSOA boundaries through the ONS Postcode Directory. Where Unique Property Reference Numbers (UPRNs) were available, they provided an additional layer of spatial precision; however, not all datasets included UPRNs, and EPC records in particular carry positional anonymisation that limits parcel-level specificity.

INSPIRE freehold polygons were linked to addresses using a combination of exact and fuzzy address-matching techniques, allowing transactions to be associated with indicative land parcel extents. While these polygons provide only an approximation of true parcel boundaries, they offered a valuable basis for contextualising typical plot sizes. Contextual spatial layers especially OS OpenMap Local were used to interpret local settlement form, street structure and surrounding land uses when developing representative parcel typologies.

Temporal harmonisation was undertaken to bring all historical market transactions to a common valuation date, thereby removing pure time effects from observed price variation. Because the Price Paid Data span multiple years, nominal sale prices reflect both underlying location-specific value and general market inflation. To isolate spatial differences in value and ensure comparability within and between LSOAs, all transactions were indexed to a single reference date using the UK House Price Index (HPI).

Sales values were adjusted using the standard index-ratio approach:

$$P_{t_0} = P_t \times \frac{HPI(t_0)}{HPI(t)}$$

where P_t is the observed sale price at transaction date t, $HPI(t)$ is the index value at that date, $HPI(t_0)$ is the index at the chosen valuation date, and P_{t_0} is the price restated in valuation-date terms.

The HPI measures average residential property price change over time and is constructed from completed transaction data using mix-adjusted statistical methods to control for variation in property types sold. By scaling historic prices in proportion to index movement, this procedure converts all transactions into constant-price terms at the valuation date. This removes distortion arising from market-wide inflation or contraction and ensures that observed variation across parcels reflects underlying market structure, settlement form and local demand conditions rather than differences in transaction timing.

This temporal adjustment is standard practice in mass appraisal and large-scale valuation modelling, particularly where transaction volumes are uneven across years. It enables transactions drawn from multiple time periods to be meaningfully compared within each LSOA and provides a consistent basis for GDV estimation and cross-area benchmarking.

Attribute augmentation was applied where EPC floorspace data could be linked to PPD transactions, enabling approximate price-per-square-metre indicators to be inferred. These relationships informed GDV estimation within the Residual Value Model and helped contextualise typical dwelling sizes within the study areas. PPD and EPC records were then aggregated to LSOA-level distributions such as medians and interquartile ranges to support representative parcel construction. Dwelling type proportions derived from Census 2021 further assisted in defining representative parcel assumptions in each LSOA. INSPIRE polygons contributed additional insight into typical parcel dimensions despite their known limitations.

Across all integration stages, strict data quality controls were applied following the principles of accuracy, completeness, consistency, timeliness, validity and uniqueness. Accuracy was supported through data cleaning and verification; completeness was monitored particularly in relation to EPC coverage; consistency was maintained through standard schema and coding practices; timeliness was ensured by aligning all datasets to their most recent release dates; validity was checked by excluding implausible transactions or records; and uniqueness was enforced through duplicate removal and reconciliation of inconsistent entries.

This systematic integration process provided a coherent, reproducible and auditable dataset suitable for delivering open-data-based residual and income capitalisation land valuation models. However, the use of solely open source data resulted in limitations to the granularity of parcel-level data achievable.

2.3.3 Definition of Parcels

Parcels in this study are defined using address-linked property records and associated spatial data across both residential and non-domestic sectors. The 2021 Lower Layer Super Output Area (LSOA) geography provides the primary spatial framework for assigning and analysing parcels in a nationally consistent way. Each parcel corresponds to a property that can be reliably matched to an LSOA through postcode–geography lookups and linked to an indicative land footprint via INSPIRE Index Polygons. These polygons offer the best available open-data representation of freehold extents and enable parcel area to be incorporated into the valuation process for all relevant property categories, including commercial, retail and mixed-use sites.

Although INSPIRE footprints do not constitute legal cadastral boundaries and may not capture every nuance of ownership subdivision or built form, they reflect the general spatial pattern of landholdings across Wales sufficiently to support valuation at LSOA scale. Their integration with observed transactions and spatial assignments provides a coherent, scalable basis for parcel definition, consistent with mass-appraisal practice when authoritative parcel-level boundaries are not publicly available.

2.3.4 Criteria for Inclusion and Exclusion

Parcels are included in the analysis only where suitable information exists to support valuation under either the Residual Value Model or the Income Capitalisation Model. Inclusion requires that:

- the property can be spatially assigned to an LSOA using postcode-directory referencing;
- transaction, rental or rating evidence is available from open datasets such as Price Paid Data or the Rating List;
- an INSPIRE polygon or equivalent footprint is available to provide indicative parcel area;
- the land use and property type are typical of the local context, whether residential, commercial or mixed use.

Parcels are excluded where essential spatial or market attributes cannot be established for example, where properties lack traceable transaction or rating information, where addressing inconsistencies prevent reliable LSOA assignment, or where INSPIRE footprints clearly indicate non-comparable land uses (e.g., large agricultural holdings, industrial compounds or complex multi-title structures). This ensures that the parcel dataset reflects the predominant structural and land-use patterns within each LSOA.

2.3.5 Sampling Strategy

A stratified sampling approach is used to ensure that parcels selected within each LSOA reflect the diversity of land uses present. Residential strata are structured using Census 2021 dwelling-type distributions, while non-domestic strata draw on Rateable Value classifications, property use codes and local land-use indicators. Within each stratum, parcels are sampled to capture the observed distribution of transaction values, rating evidence and indicative parcel footprints.

This approach ensures that the analytical dataset mirrors the full range of development, income-producing and occupation-based landholdings encountered across Welsh communities. It supports consistent comparison across LSOAs with contrasting settlement forms - from high-density urban settings to rural, coastal and post-industrial areas - and reduces the influence of atypical parcels. The resulting sample provides a robust foundation for applying both valuation models and assessing their performance across diverse spatial contexts.

2.3.6 Justification for Parcel Count

A standard sample of 100 parcels per LSOA has been adopted in order to ensure robust, comparable and statistically meaningful land-value outputs across the study areas. This decision is supported by three methodological considerations.

First, the sample provides sufficient coverage relative to the dwelling stock and land uses typically present within an LSOA. Lower Layer Super Output Areas are designed to contain approximately 1,500 residents and between 600 and 700 households on average (ONS, 2022), though actual numbers vary by settlement type, with urban areas generally containing more dwellings and rural areas fewer. A sample size of 100 parcels therefore represents a substantial proportion of the local property stock without attempting a full census, enabling meaningful representation of the range of residential, non-domestic and mixed-use properties present in each area.

Second, a sample of this size provides greater stability in valuation outputs than smaller samples. Initial modelling iterations showed that samples below approximately 60 parcels per LSOA were more sensitive to outlier transaction values, small variations in dwelling-type composition and irregularities in spatial distribution. Increasing the sample size to 100 parcels consistently reduced variance in key metrics - particularly median land value and interquartile ranges - and improved the reliability of results in areas with thin transaction evidence, constrained housing markets or heterogeneous land-use patterns. The 100-parcel threshold therefore strikes a balance between capturing sufficient diversity and maintaining computational efficiency.

Third, the selected sample size aligns with established principles in mass-appraisal practice. Guidance for large-scale valuation emphasises the need for sample densities that reflect the heterogeneity of the appraisal unit and allow the statistical distribution of values to be represented rather than dominated by atypical properties. In contexts where parcel geometries cannot be derived from cadastral sources and representative sampling must be used, methodological transparency and consistency are critical; adopting the same sample size for each LSOA ensures comparability across contrasting geographies and avoids uneven model performance arising from localised data sparsity.

Taken together, these considerations support the use of 100 parcels per LSOA as a defensible and proportionate sampling approach. It offers a level of representativeness and statistical stability consistent with the needs of open-data modelling, while ensuring that outputs remain directly comparable across Wales' diverse settlement types.

2.3.7 Data Workflow

This section outlines the structured workflow used to transform raw spatial and market datasets into model-ready parcel-level inputs.

Step 1 – LSOA Boundary Ingestion

The 2021 Lower Layer Super Output Area (LSOA) boundary dataset was ingested as the primary spatial control layer. These boundaries define the nine study areas and serve as the geographic filter for all subsequent spatial operations.

Each study LSOA (e.g., W01000114 in Gwynedd, W01002019 in Cardiff 032H) is treated as a discrete analytical unit. It is important to note that each LSOA represents a small statistical geography nested within a wider local authority area (for example, Cardiff 032H lies within the City and County of Cardiff local authority). The modelling framework operates strictly at LSOA scale, not at local authority scale.

Step 2 – INSPIRE Polygon Filtering

The national INSPIRE Index Polygon dataset was spatially intersected with the LSOA 2021 boundaries.

For each of the nine study LSOAs:

- All INSPIRE polygons falling within the LSOA boundary were extracted.
- Polygons outside the boundary were excluded.
- This ensured that only parcels geographically located within the defined study areas were retained.

This step creates a filtered INSPIRE subset for each LSOA, reducing the national dataset to study-area-specific parcel geometries.

Step 3 – Unique Parcel Identification

Each INSPIRE polygon within each study LSOA was assigned a structured unique identifier.

The format used is:

$$\text{LSOA Code}\_PXXXX$$

For example:

- W01000114_P0001 (first parcel in Gwynedd 009D)
- W01002019_P0001 (first parcel in Cardiff 032H)

This naming convention serves several purposes:

- Enables clear traceability from output back to source geography
- Allows filtering by study LSOA independently of surrounding local authority areas
- Prevents duplication across study areas
- Supports reproducibility and re-execution of the workflow

The prefix ensures that parcels can be programmatically grouped by LSOA, while the numeric suffix provides a sequential internal parcel reference.

Step 4 – Postcode-Based Data Linking

Price Paid Data (PPD) and EPC records were filtered by postcode.

Postcodes were matched to parcels using the ONS Postcode Directory (ONSPD), which provides a consistent linkage between postcode, LSOA and geographic coordinates.

Where a postcode associated with a PPD or EPC record fell within a study LSOA:

- The record was assigned to the corresponding parcel using spatial matching.
- Multiple matches were resolved using spatial proximity and address consistency checks.

Rental data were sourced from StatsWales – Private Sector Rents (Market and Valuation Data). These data were used at contextual level to inform rental assumptions within the Income Capitalisation Model.

Step 5 – Input Parcel Selection

Parcels were included in the modelling dataset only where a postcode match yielded associated PPD and/or EPC data sufficient to support valuation inputs.

Parcels that:

- Had matched postcode data, and
- Contained usable transaction and/or building attribute information

were collated as input parcels for the Residual Value Model and/or Income Capitalisation Model.

Parcels with no matched market or attribute data were discarded from the modelling dataset. This reflects the current limitations of open-data granularity and highlights a structural constraint in large-scale parcel-level modelling where comprehensive cadastral and property-level datasets are not openly available. The resulting filtered parcel dataset formed the direct input to both valuation models.

2.3.8 Potential Biases and Controls

A range of potential biases may arise when constructing parcel datasets from open-source property records and spatial references. Recognising and managing these risks is essential to ensuring that the resulting valuation dataset is robust, transparent and suitable for large-scale comparative analysis. The following considerations and controls are built into the parcel-definition and sampling process.

Transaction mix bias
Price Paid Data (PPD) reflects only completed residential transactions and may under-represent certain property types - particularly social housing, long-term rentals, or areas with historically low market turnover. To reduce this bias, transaction-derived indicators are supplemented with broader contextual signals at LSOA and MSOA level, including cross-checks against distributional measures from House Price Statistics for Small Areas (HPSSA). Outlier filtering and the use of medians rather than means further limit the influence of atypical transactions.

Rating and commercial-market bias
Rateable Value (RV) is used as the principal open-data proxy for commercial and non-domestic income potential. Because rating assessments may lag market conditions or vary by valuation tone, parcel sampling incorporates a range of commercial property types rather than relying on a narrow subset of highly active use classes. Using strata informed by RV classifications ensures that non-domestic parcels form a balanced and representative proportion of the LSOA sample.

Spatial footprint misalignment
INSPIRE Index Polygons provide the most consistent open-data representation of freehold extents, but they do not correspond exactly to legal cadastral boundaries or true developable parcels. To avoid over-interpreting parcel geometry, footprints are used only as indicative measures of relative size and spatial configuration. Parcel-level comparisons do not assume cadastral precision, and no attempt is made to infer site-specific developability from the polygons.

Urban density and built-form effects
In high-density urban areas, flats and subdivided buildings may have limited or no meaningful land component associated with individual units. To avoid overstating land value for such properties, sampling rules ensure that land attribution is consistent with typical plot structures at building level, rather than attempting to model individual dwellings. This avoids unrealistic parcel geometries in urban centres while preserving representativeness.

Small-area variability
LSOAs with low transaction frequency or sparse commercial activity may exhibit elevated volatility in price- or rent-based indicators. Temporal normalisation using the House Price Index (HPI) and cross-validation against HPSSA reduce sensitivity to single-period market fluctuations. Sampling strata ensure that sufficient parcels are drawn from each major use type to avoid results being driven by very small numbers of transactions or rating entries.

Address-matching and geocoding bias
Parcel assignment relies on postcode and address linkages which, in some cases, may introduce error (e.g., historic postcodes, merged records, incomplete addresses). Robust postcode-to-LSOA matching via ONS Postcode Directory, combined with exclusion of properties that cannot be reliably matched, minimises misclassification. The sampling frame includes only parcels with clear spatial referencing to avoid distortions arising from ambiguous or incomplete address records.

Non-comparable land uses
Certain land uses, such as agricultural holdings, large industrial compounds or complex multi-title structures, are excluded where their characteristics fall outside the scope of the valuation models. This prevents atypical parcels from exerting disproportionate influence on LSOA-level estimates and maintains internal consistency across study areas.

Taken together, these controls ensure that parcel definition and sampling remain methodologically sound, minimise the risk of structural bias, and uphold the standards of transparency and reproducibility expected in large-scale, open-data valuation exercises.

2.4 Modelling Methodology

This section provides a detailed explanation of the modelling methodology used to estimate land values using open data. Two complementary valuation models were developed and implemented in Python:

- **Residual Value Model (RVM)** – a development-led approach to estimating land value based on market GDV and development economics.
- **Income Capitalisation Model (ICM)** – an income-based approach to estimating land value by capitalising net income attributable to land.

Both models are fully auditable, use transparent parameter sets, and operate exclusively with open-source datasets.

2.4.1 Overview of Approaches Considered

A range of modelling approaches was assessed during the scoping phase. Methodologies such as hedonic regression, multilevel regression and hybrid transaction–covariate models were considered but not adopted. These methods require detailed property attributes (e.g., condition, internal specification, neighbourhood amenities) which are not consistently available at national open-data level in Wales. They also rely on assumptions regarding functional forms and covariate interactions that reduce transparency.

2.4.2 Residual Value Model (RVM)

The RVM estimates the value of land as a residual after accounting for all development costs, finance and an appropriate developer return. The method answers a simple economic question:

What is the maximum price a rational developer could pay for the land while achieving a normal profit margin?

The model expresses this as:

$$\text{Residual Land Value} = \text{GDV} - \text{Total Development Costs.}$$

The implementation incorporates detailed components reflecting typical development economics.

2.4.2.1 Gross Development Value (GDV)

GDV is simulated using a typical scheme configuration for each parcel. It is calculated as the sum of the value of all unit types expected to be delivered:

$$GDV = \sum_{u} N_u \cdot S_u \cdot P_{m^2} \cdot F_u$$

where:

- N_u: number of units of dwelling type u,

- S_u: average unit size (m²),
- P_{m^2}: base price per m²,
- F_u: dwelling-type price adjustment factor.

This approach allows GDV to vary systematically with dwelling type, scale and location, ensuring internal consistency across Wales and avoiding reliance on proprietary scheme assumptions.

2.4.2.2 Build Costs

Build costs are estimated using open-data-compatible assumptions that mirror BCIS-style cost structures:

$$C_{build} = \sum_{u} N_u \cdot S_u \cdot BCIS_u \cdot LF$$

where:

- $BCIS_u$: assumed £/m² for each unit type,
- LF: regional or LSOA-level location factor,
- N_u: number of units of dwelling type u,
- S_u: average unit size (m²).

This reflects differences in construction costs associated with different dwelling forms.

2.4.2.3 Infrastructure and External Works

Infrastructure and external works are modelled as a fixed proportion of GDV:

$$C_{infra} = GDV \times 0.15$$

This assumption reflects early-stage feasibility practice where detailed site-specific civil engineering data are unavailable.

2.4.2.4 Professional Fees and Contingency

Professional fees:

$$C_{fees} = 0.10 \cdot (C_{build} + C_{infra})$$

Contingency:

$$C_{cont} = 0.05 \cdot C_{build}$$

This reflects typical allowances used in early-stage development appraisals.

2.4.2.5 Abnormal Costs

The model incorporates explicit abnormal cost allowances associated with:

- contamination,
- flood mitigation,
- adverse ground conditions.

Parcel-level values override defaults where provided.

Total abnormals:

$$C_{abn} = C_{contamination} + C_{flood} + C_{ground}$$

This ensures that representative parcels incorporate realistic viability constraints.

2.4.2.6 Finance Model (Monthly Compounding)

Finance costs are modelled using a monthly interest mechanism. Infrastructure and abnormal costs are assumed to be incurred at the start of the programme, with build, fees and contingency spread evenly across the development period.

For each month m:

$$C_{finance} = \sum_{m=1}^{T} Outstanding_m \cdot r_{monthly}$$

Profit is added at the end of the project without additional interest.
This provides a time-sensitive measure of finance exposure, reflecting the impact of development duration and monthly borrowing patterns.

2.4.2.7 Developer Profit

Developer return is calculated as:

$$C_{profit} = GDV \times 0.17$$

A 17% return on GDV has been adopted as it sits within the standard range used in UK residential development viability testing. Professional guidance from the Royal Institution of Chartered Surveyors (RICS) indicates that residential developer returns typically fall between 15% and 20% of GDV, depending on scheme risk and market conditions (Crosby et al. (2018)).

2.4.2.8 Residual Land Value

The land value is calculated as:

$$RLV = GDV - (C_{build} + C_{infra} + C_{fees} + C_{cont} + C_{abn} + C_{finance} + C_{profit})$$

A negative value indicates that a development of this representative type is not financially viable on the parcel.

2.4.2.9 Land Value Intensity Measures

The model outputs consistent land-intensity measures:

$$LV_{m^2} = \frac{RLV}{\text{Developable Area (m}^2\text{)}}$$

$$LV_{ha} = \frac{RLV}{\text{Site Area (ha)}}$$

These enable consistent comparison across parcels and LSOAs.

2.4.3 Income Capitalisation Model (ICM)

The ICM calculates land value by converting income generated by the property into a capitalised land value. It is well-established in commercial valuation practice and is adapted here to operate using open data.

2.4.3.1 Rent Estimation from Rateable Value

Annual rental income is proxied directly using Rateable Value (RV):

$$\text{Rent} \approx RV$$

Under UK rating legislation, Rateable Value represents the assessor’s estimate of the annual rent at which a non-domestic property might reasonably be expected to let on the open market at the valuation date. It therefore provides a nationally consistent proxy for market rent.

In this framework, the RV-derived rent is carried forward directly into the next stage of the Income Capitalisation Model, where operating expenses are deducted to calculate Net Operating Income (NOI), which is then capitalised to estimate land value.

2.4.3.2 Operating Expenses and Net Operating Income

Net operating income (NOI) is:

$$NOI_{total} = Rent \cdot (1 - OPEX)$$

Use-class-specific OPEX ratios (industrial, office, retail, leisure) are used, with parcel-level overrides permitted.

2.4.3.3 Building Value (Replacement Cost New – Depreciation)

Replacement cost is estimated as:

$$RCN = BuiltArea_{(m^2)} \times RebuildCost_{(m^2)}$$

Depreciation is then applied according to age bands:

- 0–10 years → 5%
- 10–30 years → 20%
- 30–50 years → 40%
- 50+ years → 60%

$$BV = RCN \times (1 - Dep)$$

These depreciation bands reflect a simplified economic life approach consistent with standard Depreciated Replacement Cost (DRC) methodology used in UK valuation practice. In DRC frameworks, buildings are assumed to depreciate progressively over their economic life due to physical deterioration, functional obsolescence and market ageing.

These stepped bands provide a transparent and auditable proxy for age-related value loss within an open-data modelling environment, where detailed condition surveys, refurbishment histories and asset-level assessments are not available.

While simplified, the approach ensures internal consistency across LSOAs and avoids introducing subjective, site-specific adjustments that would undermine reproducibility. Sensitivity testing around depreciation assumptions can be undertaken in future phases to assess the impact on Income Capitalisation Model outputs.

2.4.3.4 Return on Improvements

A return is applied to the value of the improvements:

$$ImprovementReturn = BV \times r_{imp}$$

Default $r_{imp} = 3\%$.

The 3% rate reflects a conservative capital return attributable to the building element only, separate from land. In income-based valuation, total property yields typically range from 5–7% (or higher), incorporating both land and building risk. Applying a lower 3% return to improvements ensures that income attributed to the structure reflects its relatively stable service potential, while allowing the residual income to be correctly allocated to land.

This parameter supports a transparent separation of land and improvement value and will be tested through sensitivity analysis in the next phase.

2.4.3.5 Land NOI

$$NOI_{land} = NOI_{total} - ImprovementReturn$$

This isolates the income attributable to land after accounting for the return on buildings.

2.4.3.6 Capitalisation

$$LV = \frac{NOI_{land}}{CapRate}$$

Capitalisation rates reflect the market-required return for income streams of differing risk profiles. Sector-specific rates are applied to align with observed UK investment yield evidence (e.g. MSCI, CBRE, JLL, Savills).

For modelling purposes, mid-range rates are adopted: 5% industrial; 6% office/retail. These values sit within observed market bands and reflect relative sector risk.

2.4.3.7 Land Value Intensity Metrics

$$LV_{m^2} = \frac{LV}{\text{Developable Area (m}^2\text{)}}$$

$$LV_{ha} = \frac{LV}{\text{Site Area (ha)}}$$

These allow direct comparison with RVM outputs.

Parcel-level outputs are aggregated to LSOA-level median, mean, and percentile measures (5th and 95th percentiles). The 5th (p5) and 95th (p95) percentiles indicate the lower and upper bounds within which the central 90% of parcel-level values fall. Together, these measures capture both central tendency and dispersion, enabling robust comparison across the 9 study LSOAs.

2.4.4 Normalisation Across Regions

Ensuring comparability of land-value outputs across Wales requires a consistent modelling framework, combined with procedures that control for regional differences in market structure, settlement form and data availability. Both the Residual Value Model (RVM) and the Income Capitalisation Model (ICM) apply a unified set of assumptions and valuation parameters across all study LSOAs. This allows regional variation to emerge from underlying empirical data rather than from model configuration.

Normalisation is achieved through several mechanisms. First, all properties are modelled using LSOA-level inputs derived from open datasets, such as transaction values, rating assessments and observed parcel characteristics. This ensures that regional differences reflect genuine market conditions - for example, higher Rateable

Values in Cardiff, lower transaction density in Gwynedd, or distinctive coastal patterns in Ceredigion - rather than varying methodological choices.

Second, the models incorporate temporal normalisation, with transaction evidence indexed to a common valuation date using the House Price Index. This removes timing effects that would otherwise distort comparisons between regions with different sales frequencies or market cycles. The use of medians, interquartile ranges and robust filtering techniques further reduces the influence of outliers or atypical transactions that might otherwise skew regional comparisons.

Third, normalisation accounts for differences in land-use structure, including dwelling-type mix, density, parcel size distributions and the relative presence of commercial versus residential properties. Because each LSOA's parcel sample reflects its actual property composition, regional outputs fairly capture contrasts between dense urban centres, mixed suburban environments, coastal communities and sparsely populated rural areas. This approach ensures that differences between regions are rooted in real spatial and market characteristics, not in inconsistent sampling.

Finally, outputs from both models are summarised at LSOA level using consistent statistical measures - median land value, mean, 5th and 95th percentiles. This provides a standardised basis for cross-region comparison, avoids bias arising from differing sample sizes and supports aggregation to regional and national scales.

Taken together, these normalisation procedures ensure that regional variation in land-value outputs arises from the distinct socio-economic, physical and market characteristics of Welsh communities, rather than from differences in modelling assumptions or data handling. This provides a defensible basis for comparing land-value dynamics across Wales's diverse regions.

2.5 Key Assumptions

The valuation models developed under Lot 4 were shaped by a series of explicit assumptions embedded within both the modelling framework and the open-data environment. These assumptions were necessary to ensure national consistency, support reproducibility, and enable both the Residual Value Model (RVM) and the Income Capitalisation Model (ICM) to operate coherently across diverse Welsh geographies.

For analytical clarity, assumptions were structured into two overarching categories:

- **Structural assumptions** — analytical boundary conditions arising from data availability, licensing constraints and agreed scope.
- **Calibrated assumptions** — value-shaping parameters grounded in valuation practice and capable of refinement through evidence, sensitivity testing and scenario modelling.

Within this hierarchy, assumptions are presented below under four operational groupings: data and input assumptions; parcel-construction assumptions; methodological assumptions; and model-specific assumptions.

2.5.1 Data and Input Assumptions

Because the project relied exclusively on open datasets, a number of simplifying assumptions were required to accommodate inconsistencies, aggregation effects and data gaps in publicly available sources.

These represent primarily structural assumptions arising from the open-data constraint.

LSOA-Level Spatial Assignment
All properties were matched to 2021 LSOA boundaries through postcode-directory linkage (ONSPD). This approach assumes that postcode centroids provide a sufficiently accurate representation of local geography for LSOA-scale analysis. While appropriate for statistical modelling, postcode-based assignment may introduce minor spatial imprecision at parcel level.

Use of Transaction and Rating Data as Market Proxies
Residential sale prices from HM Land Registry Price Paid Data were assumed to represent observable market conditions.

For non-domestic property, **Rateable Value (RV)** was assumed to be equivalent to annual market rent at the valuation date, consistent with UK rating legislation. RV was therefore used as a proxy for gross rental income within the ICM. This assumes that rating assessments reasonably reflect market rental levels, acknowledging potential lag effects between revaluation cycles.

Indicative Parcel Area from INSPIRE Polygons
Parcel footprints were derived from INSPIRE Index Polygons. These approximate freehold extents but do not constitute legal cadastral boundaries. They were assumed to be sufficiently representative for LSOA-scale modelling while recognising that development parcels, title subdivisions and easements cannot be fully captured within an open-data-only framework.

Temporal Normalisation Using HPI
All residential transactions were indexed to a common valuation date using the UK House Price Index (HPI). This assumes that regional or Wales-level HPI movements reasonably capture price change dynamics at LSOA level, even where local transaction volumes are limited.

These data assumptions ensured coherence across LSOAs but necessarily introduce generalisation, particularly in rural or low-transaction environments.

2.5.2 Parcel Construction and Sampling Assumptions

The parcel dataset integrated residential and non-domestic properties and was shaped by several structural modelling decisions.

Stratified Parcel Sampling
A consistent sample of 100 parcels per LSOA was drawn using stratified selection. This approach assumed that a representative subset of parcels could capture local variation in land-use, geometry and market evidence. A full census of land was not feasible within the open-data environment.

Representativeness Across Land Uses
Sampling strata were constructed using Census dwelling-type proportions (for residential land) and Rateable Value classifications (for commercial land). This assumes that these distributions reasonably reflect functional land-use composition within each LSOA.

Exclusion of Non-Comparable Land Uses
Large agricultural holdings, industrial compounds and complex multi-title sites were excluded where open-data attributes were insufficient to support reliable valuation. This assumption prioritised comparability and methodological stability over exhaustive coverage.

Developable Land Ratio
A uniform developable land ratio of 0.70 was applied to parcels to approximate the proportion of land capable of supporting built development after allowing for infrastructure and non-developable areas.

This ratio was selected as a conservative mid-range modelling assumption, broadly consistent with typical UK site efficiency benchmarks (where 25–35% of gross area is allocated to access, servicing and non-developable uses). However this value should be taken as a simplification for modelling efficiency rather than a definitive value. Additional modelling complexity would be needed to reflect the significant diversity of developable ratios seen in practice.

2.5.3 Methodological Assumptions Applied Across Both Models

Both RVM and ICM operated under a consistent national modelling architecture. These assumptions include both structural constraints and calibrated parameters.

Uniform Modelling Architecture
Core valuation parameters — including profit margins, finance logic, capitalisation rates and operating expense ratios — were applied uniformly across study areas. This prevented artificial regional bias and ensured comparability.

Median-Based Aggregation
Outputs were summarised using medians and percentile bands to reduce the influence of outliers. This reflects mass appraisal practice and strengthens stability in thin markets.

Open-Data Cost and Yield Benchmarks
Construction cost and yield assumptions were derived from open-data proxies rather than proprietary BCIS or market leasing databases. This ensured transparency but limits local calibration precision.

Absence of Site-Specific Planning and Legal Constraints
The models did not incorporate detailed planning permissions, development agreements, biodiversity requirements, hope value, or title encumbrances. These omissions reflect structural data constraints rather than analytical oversight.

Together, these assumptions ensured transparency and reproducibility but introduce bounded uncertainty, particularly in areas strongly shaped by planning policy or market atypicality.

2.5.4 Model-Specific Assumptions: Residual Value Model (RVM)

The RVM relies on a set of calibrated assumptions that materially influence viability outputs.

GDV Benchmarks
GDV was derived from transaction evidence and floorspace indicators. While expressed as central benchmarks, development recognised that GDV behaves as a distribution rather than a single fixed value, particularly across heterogeneous markets.

Construction Costs
Construction costs were estimated using BCIS-style proxies and a national location factor. These assume homogeneity within dwelling types and do not reflect contractor-specific or specification-level variation.

Infrastructure, Fees and Contingency
Fixed percentage allowances were applied to reflect early-stage feasibility practice. These simplify modelling but may not reflect localised delivery conditions.

Abnormal Costs
Abnormal costs were incorporated where identified but treated using structured allowances rather than fully risk-weighted probabilistic modelling.

Finance Structure
Finance was modelled using a fully debt-funded structure with evenly distributed expenditure. This differs from phased or equity-blended real-world arrangements.

Developer Profit
Developer return was modelled at 17% of GDV, consistent with accepted UK residential viability ranges (15–20%). Profit was recognised as a key calibrated parameter and major driver of residual value sensitivity.

These assumptions influence RVM outputs most strongly in rural, coastal and post-industrial contexts where viability margins are thin.

2.5.5 Model-Specific Assumptions: Income Capitalisation Model (ICM)

The ICM similarly relies on calibrated parameters affecting income separation and capitalisation.

Rent Derived from Rateable Value
Rent was assumed equivalent to Rateable Value at the valuation date. This presumes rating assessments reasonably approximate market rent across use classes and locations.

Standardised OPEX Ratios
Operating-expense ratios were assigned by use class and applied uniformly across Wales. This does not capture building-specific management intensity or geographic variation.

Replacement Cost New and Depreciation
Building value was estimated using rebuild-cost-per-m² assumptions and age-band depreciation reflecting a simplified economic life model. Functional obsolescence was recognised but not separately modelled.

Return on Improvements
A 3% return on improvements was applied to separate building income from land income. This reflects a conservative capital return on built assets relative to total property yields.

Capitalisation Rates
Land capitalisation rates were assigned by use class (typically 5–7%). These represent broad market ranges and do not incorporate LSOA-level yield variation.

These assumptions ensure coherent income attribution but limit the ability of the model to reflect nuanced local commercial property dynamics.

2.5.6 Analytical Implications

Across both models, assumptions were deliberately structured to:

- Preserve transparency and auditability;
- Avoid false precision from spurious calibration;
- Enable sensitivity testing of high-impact parameters;
- Clearly distinguish between structural constraints and value-driving variables.

The structured assumption framework strengthens interpretability and provides a robust foundation for further refinement in the final phase. It also clarifies that model outputs represent bounded, policy-facing estimates rather than site-specific market appraisals.

2.6 Limitations

The modelling framework developed for this project provides transparent, reproducible land-value estimates using only open-source datasets. However, several important limitations constrain the precision and site-specific applicability of the outputs. These limitations reflect challenges inherent in the underlying datasets, the structural design of the models, and the open-data-only constraint specified for this work. They should be taken into account when interpreting results or considering future enhancements.

A major constraint is the absence of an authoritative, openly licensed cadastral dataset for Wales. The INSPIRE Index Polygons used in this project provide indicative freehold extents but do not represent legal parcels, development plots or subdivided titles. As a result, valuation outputs are estimated for representative not actual parcels, limiting the ability to incorporate parcel-specific constraints such as site shape, access, easements, historic land use or ground conditions. Similarly, EPC data provide only partial coverage of dwelling floorspace, with under-representation of older properties and certain tenures. Where EPC attributes are unavailable, representative size distributions must be inferred, introducing uncertainty into GDV estimates (in the RVM) and building-value estimates (in the ICM).

Transaction thinness represents another source of uncertainty. Many LSOAs record few annual residential transactions, making it difficult to infer robust price distributions for certain dwelling types. Temporal smoothing using HPSSA or rolling transaction windows can mitigate volatility, but cannot replace granular market evidence. Build-cost estimation also relies on open-data proxies rather than BCIS benchmarks, leading to potential divergence from actual construction conditions, regional cost pressures or specification trends.

Within the model structure itself, several simplifying assumptions are necessary to enable national-scale, open-data-based valuation. The representative parcel approach provides analytical consistency across Wales but does not reflect the unique development constraints or opportunities associated with specific sites. Planning policy considerations such as density limits, affordable housing requirements, developer contributions, biodiversity net gain, or heritage constraints are not embedded in the modelling logic, even though these factors strongly influence real-world land values and viability.

In the Residual Value Model, GDV is estimated using a simulated scheme structure with assumed dwelling mixes, sizes and hedonic price adjustments. While this approach maintains national consistency and avoids dependence on local development briefs, it may not represent optimal or permissible development patterns for all areas. The finance model assumes a single-phase, debt-funded programme with evenly distributed construction expenditure, whereas actual development cashflows may involve phasing, staged receipts, blended financing or strategic land assembly, all of which materially affect viability. Negative residual values are interpreted as non-viable outcomes under the assumed representative scheme, rather than as literal market land prices.

The Income Capitalisation Model is subject to analogous constraints. The estimation of rent from Rateable Value is a practical method, but rating assessments do not always reflect current market rents, particularly in areas where revaluations lag market change. Operating expense ratios are represented through generalised use-class profiles and do not vary with building condition, specification or management intensity. Rebuild-cost assumptions reflect broad benchmarks and depreciation bands rather than detailed construction specifications. Capitalisation rates are defined at use-class level and do not incorporate local yield evidence or market sentiment. These factors collectively limit the precision of income-based land values.

Across both models, there is no representation of developer behaviour, negotiation dynamics, land assembly costs, hope value, or premiums associated with strategic or adjoining land. The outputs therefore demonstrate the feasibility of using open data for broad, comparative land-value modelling, rather than providing site-specific valuations. They should be interpreted as indicative estimates illustrating methodological performance, not as substitutes for full professional valuation advice.

In summary, the limitations identified here are primarily driven by the open-data-only remit, the absence of detailed proprietary datasets, and the representative-parcel modelling approach. Although the models offer consistent, transparent and scalable valuation estimates across Wales, the results are best understood as policy-facing evidence of methodological feasibility rather than precise predictions of market land values for individual sites.

2.6.1 Missing Data

Despite the strong availability of open datasets in Wales, several structural and methodological limitations constrain the extent to which land valuation models can achieve parcel-level accuracy using open-source data alone.

A key limitation was the inability to utilise the SAIL Databank. Although SAIL contains high-quality, property-linked and longitudinal microdata, its access restrictions and licensing conditions mean that it cannot be used in open-data modelling. Excluding SAIL necessarily limits the granularity at which socio-demographic, health or tenure attributes can be analysed within the valuation context.

The absence of a complete, authoritative Welsh cadastral dataset also remains a fundamental constraint. While INSPIRE Index Polygons provide indicative freehold extents, they do not represent development parcels, ownership subdivisions, easements or land-use constraints. As a result, parcel-level modelling of abnormal development costs, site constraints or buildability - standard practice in professional valuation - cannot be undertaken reliably within an open-data-only framework.

In the non-residential domain, significant data gaps remain. The VOA Rating List contains rich information on hereditaments, floor areas and rateable values, which would be highly relevant for income capitalisation modelling of commercial land uses. However, as these data are not licensed under the Open Government Licence, they were excluded.

Other limitations stem from the nature of specific datasets. EPC records provide vital property attributes but suffer from positional anonymisation and variable coverage across dwelling types. PPD provides transaction-level evidence but lacks floorspace and interior characteristics. HPSSA provides only annual updates, limiting sensitivity to short-term market shifts. OS OpenMap Local supplies contextual mapping but lacks cadastral precision.

Taken together, these limitations highlight that while open data are sufficiently robust for testing the feasibility and reproducibility of RVM and ICM at LSOA scale, further accuracy - particularly relating to site-specific valuation, abnormal costs and precise parcel geometry - cannot be achieved without access to proprietary or licensed datasets.

Section 3 - Findings

This section presents the analytical findings arising from the application of the Residual Value Model (RVM) and the Income Capitalisation Model (ICM) to the parcel data assembled for the nine study LSOAs. The analysis examines how the models behave across different geographical and market contexts, and how land values vary when derived from development-led and income-led approaches. The outputs presented here represent the valuation results for the parcels included in this analysis, aggregated to LSOA level to support consistent comparison across the study areas.

The structure of this section reflects three analytical components. First, it reviews the characteristics of the parcel datasets in each LSOA and considers how variations in land-use composition, dwelling stock, market evidence and spatial form influence the behaviour of the two models. Second, it presents and interprets the valuation outputs generated by the RVM and ICM, identifying patterns, contrasts and the underlying drivers of variation across the ten study areas. Finally, the section summarises key findings from the analysis and outlines the next steps required to refine the models, extend the valuation to full parcel coverage and prepare for the development of a combined or hybrid valuation framework in the final phase.

3.1 Analysis of the Parcel Data

The parcel dataset assembled for this analysis contains 900 parcels, with exactly 100 parcels sampled within each of the nine study LSOAs. This structured sampling ensures that each area contributes an equal volume of valuation evidence, enabling like-for-like comparison across contrasting geographies. The dataset incorporates the full range of variables needed to operate both the Residual Value Model (RVM) and the Income Capitalisation Model (ICM), covering site characteristics, market attributes, abnormal cost indicators and income-based parameters.

Figure 2, below, shows the breakdown in land use across the 900 parcels. From the chart, we see the majority of the land parcels represent residential use, with industrial and retail being the next largest categories.

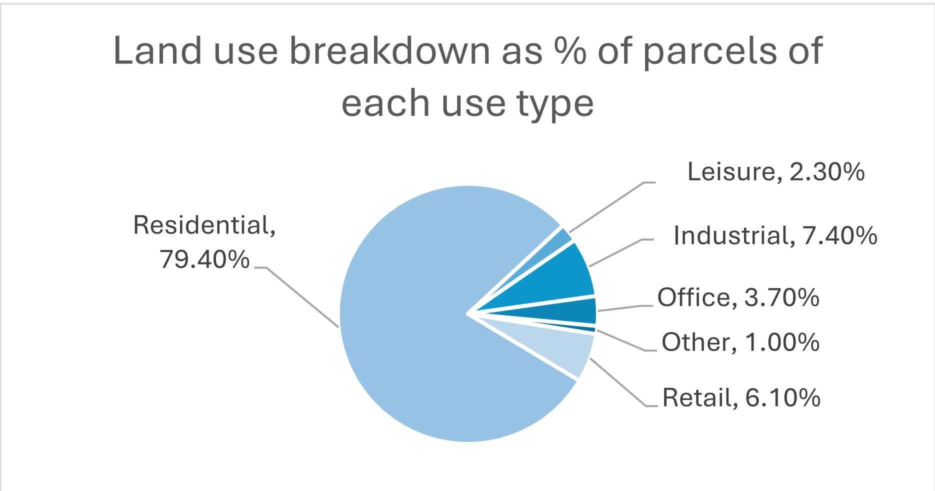

Figure 1: Breakdown of property by type

Across the dataset, there is clear and intentional variation in the attributes that influence the behaviour of both valuation methods. Parcel sizes range from approximately 0.06 ha to 0.39 ha, reflecting the differing settlement forms represented in the study areas. Smaller parcels tend to dominate in more urban LSOAs such as Cardiff and Bridgend, while larger and more irregular parcels are characteristic of rural LSOAs including Powys and Gwynedd. Developable land ratios are typically between 0.65 and 0.75, and programme durations range from 12 to 36 months, providing realistic variation in development exposure for RVM financing assumptions.

RVM-specific inputs vary meaningfully across the sample. Base prices per square metre, where present, span from around £1,300/m² to more than £2,400/m², corresponding to expected market-strength differences between higher-demand areas in South East Wales and more rural or coastal locations. Build-cost location factors also exhibit variation, with values ranging from around 1.02 to 1.08, reflecting regional cost differences. Abnormal cost allowances for contamination, flooding and ground conditions range from £0 to more than £30,000 per parcel, creating a realistic distribution of development-risk profiles that materially influence residual land value outcomes.

For parcels included in the ICM, income-based variables display similarly wide variation. Rateable values range from below £1,000 to more than £60,000, highlighting the contrast between the robust commercial activity in Cardiff and the limited rating evidence in rural or small-settlement LSOAs. Operating-expense ratios fall between 0.25 and 0.35 depending on use class, while rebuild-cost assumptions typically lie between £1,200/m² and £1,800/m². Building ages range from new-build stock to properties more than 50 years old, with associated differences in depreciation and improvement-return calculations.

The dataset therefore captures a wide spectrum of land-use, property and market conditions across the study areas. This diversity is essential for testing the valuation models under realistic conditions. Urban LSOAs contribute parcels with stronger market signals, higher rating density and more consistent dwelling structures; rural and coastal areas contribute parcels with larger footprints, lower turnover and more heterogeneous property characteristics; while post-industrial areas contribute parcels with elevated abnormal costs and mixed economic conditions.

Taken together, the parcel dataset provides a sufficiently rich and internally varied foundation for analysing land-value patterns across the nine LSOAs. It ensures that the valuation models respond to real differences in property characteristics and local market context, rather than artificial variation introduced through modelling assumptions. The breadth and distribution of inputs observed across the 900 parcels underpin the results presented in Section 3.2 and support a robust assessment of the feasibility and stability of the valuation methods.

3.2 Analysis of the Residual Value Model (RVM) Results

Within all nine study LSOAs, Residual Value Model (RVM) outputs return positive mean and median residual land values. This indicates that, under central assumptions, representative residential development schemes are economically viable across all study areas, albeit with varying margins and degrees of volatility.

The results reflect GDV calibration, cost benchmarking and profit assumptions, and provide a representation of spatial variation in development-led land value across Wales.

Table 1: LSOA Scale Outputs from Residual Value Modelling

LSOA Name	Mean RLV (£)	Median RLV (£)	P5 (£)	P95 (£)	Parcel Count
Gwynedd 009D W01000114	119,811.50	96,597.73	10,812.85	302,343.90	85
Flintshire 015A W01000255	85,984.34	76,843.86	-42,753.40	249,753.80	70
Powys 011C W01000449	121,241.50	113,061.60	-13,255.60	293,064.90	90
Ceredigion 002D W01000517	176,484.10	165,189.80	26,501.96	397,521.00	85
Pembrokeshire 002F W01000617	216,811.90	196,119.90	54,562.39	469,058.60	80
Bridgend 019D W01001045	186,353.90	169,311.50	49,625.82	410,492.60	75
Rhondda Cynon Taf 001F W01001233	30,735.45	16,912.74	-60,294.20	151,279.10	85
Monmouthshire 006F W01001597	385,600.00	349,796.40	151,127.70	717,470.10	80
Cardiff 032H W01002019	147,597.90	137,777.00	28,645.35	345,316.50	65

Central Tendency and Overall Viability

Mean residual land values (RLVs) range from approximately **£30,735** in Rhondda Cynon Taf 001F (W01001233) to **£385,600** in Monmouthshire 006F (W01001597). Median values follow a similar hierarchy, ranging from **£16,913** to **£349,796**.

Several key observations emerge:

- All LSOAs exhibit positive central RLVs, indicating broad viability under baseline assumptions.
- The difference between mean and median values is generally modest, suggesting relatively stable parcel distributions in most markets.
- Higher-value residential markets produce materially stronger residual outcomes.

Monmouthshire 006F generates the highest residual values, reflecting strong GDV performance and favourable market conditions typical of affluent commuter geographies. Pembrokeshire 002F, Bridgend 019D and Ceredigion 002D also demonstrate robust central viability, with mean RLVs exceeding £170,000–£210,000.

By contrast, Rhondda Cynon Taf 001F returns materially lower residual values. Although positive on average, the margins are thinner, reflecting weaker market demand, legacy land-use constraints and heightened cost sensitivity.

Distributional Profile and Downside Risk

The percentile distribution provides insight into the variability and downside exposure within each LSOA.

Negative 5th percentile values are present in:

- Flintshire 015A (–£42,753),
- Powys 011C (–£13,256),
- Rhondda Cynon Taf 001F (–£60,294).

This indicates that a subset of parcels within these areas are marginal or unviable under stress conditions. Such outcomes are consistent with:

- Larger or less efficient parcel geometry,
- Elevated abnormal-cost exposure,
- Greater sensitivity to GDV variation in thin markets.

Rhondda Cynon Taf 001F exhibits the most pronounced downside risk, reinforcing its structural vulnerability to development-led valuation.

In contrast, stronger markets such as Monmouthshire 006F and Pembrokeshire 002F have positive lower-bound values, suggesting greater resilience to parameter variation.

Upside Dispersion and Market Asymmetry

The 95th percentile values highlight the asymmetric nature of development viability in higher-demand areas. Monmouthshire 006F reaches a 95th percentile RLV exceeding **£717,000**, while Pembrokeshire 002F and Bridgend 019D approach **£410,000–£470,000**.

This dispersion reflects the amplification effect of high GDV environments. Where strong sales values combine with efficient parcel characteristics and limited abnormal constraints, residual margins expand materially.

The wider spread in affluent and suburban markets demonstrates that development-led land value is inherently non-linear, with upside sensitivity magnified in high-demand contexts.

Spatial Pattern and Economic Coherence

The spatial gradient observed across the RVM outputs aligns closely with known Welsh market dynamics:

- Strongest residual values occur in commuter belt and coastal-suburban markets.
- Moderate viability is observed in mixed rural–urban LSOAs.
- Structurally weaker and post-industrial areas exhibit thinner margins and greater volatility.

This pattern reflects the interaction between achievable GDV, build-cost proxies, abnormal exposure and market evidence strength. The coherence of this gradient provides confidence that the model is functioning logically under the refined parameter framework.

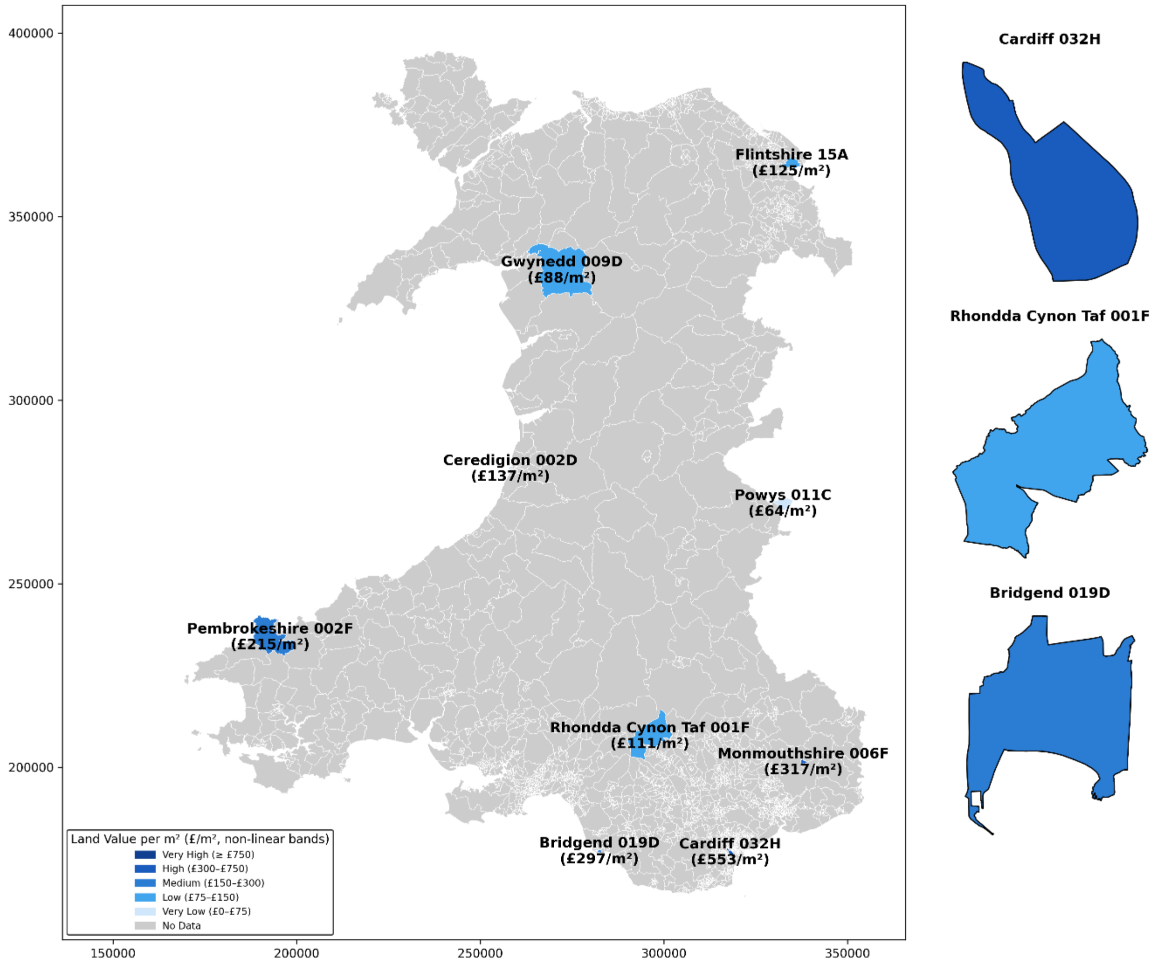

Figure 2: Mapped LSOA Outputs of the RVM

Sensitivity and Parameter Influence

Despite positive central outcomes, the RVM is structurally sensitive to several calibrated parameters:

- Gross Development Value assumptions,
- Developer profit margins,
- Construction and infrastructure cost proxies,
- Abnormal cost allowances,
- Programme duration and finance rates.

The persistence of negative lower-bound values in selected LSOAs confirms that residual valuation is inherently high-variance, particularly in thinner markets. Small shifts in GDV or cost inputs materially alter residual outcomes.

This reinforces the importance of structured sensitivity testing in the next phase.

Implications for Interpretation

The RVM results indicate that development-led land value is broadly positive across the study areas under current central assumptions. However:

- Viability margins vary substantially across geographies.
- Downside risk persists in structurally weaker markets.
- Residual valuation remains highly assumption-sensitive.

Accordingly, the RVM should be interpreted as:

- A measure of development viability under defined parameter sets;
- A bounded estimate of compliant residential land value;
- One component of a broader hybrid land-value framework.

The results confirm that the RVM performs coherently within the open-data modelling environment, but they also demonstrate why reliance on development-led valuation alone may produce unstable fiscal signals in certain contexts.

3.2.1 Intra-LSOA RVM Results

This section analyses parcel-level variation in Land Value per m² (LV/m²) within each study LSOA. Expressing results in £/m² allows direct comparison of land intensity across parcels of differing size and improves interpretability for spatial and fiscal analysis.

For each LSOA, the minimum, maximum, average and range of parcel-level LV/m² were assessed. The results demonstrate significant intra-LSOA heterogeneity, confirming that development-led land value is highly sensitive to parcel characteristics and localised GDV–cost relationships.

It is important to note that negatively valued parcels (where present) are not displayed in the LSOA maps shown in sections 3.2.1 and 3.3.1.

Gwynedd 009D (W01000114)

- **Average LV/m²:** £87.97
- **Minimum LV/m²:** –£6.85
- **Maximum LV/m²:** £669.30
- **Range:** £676.14

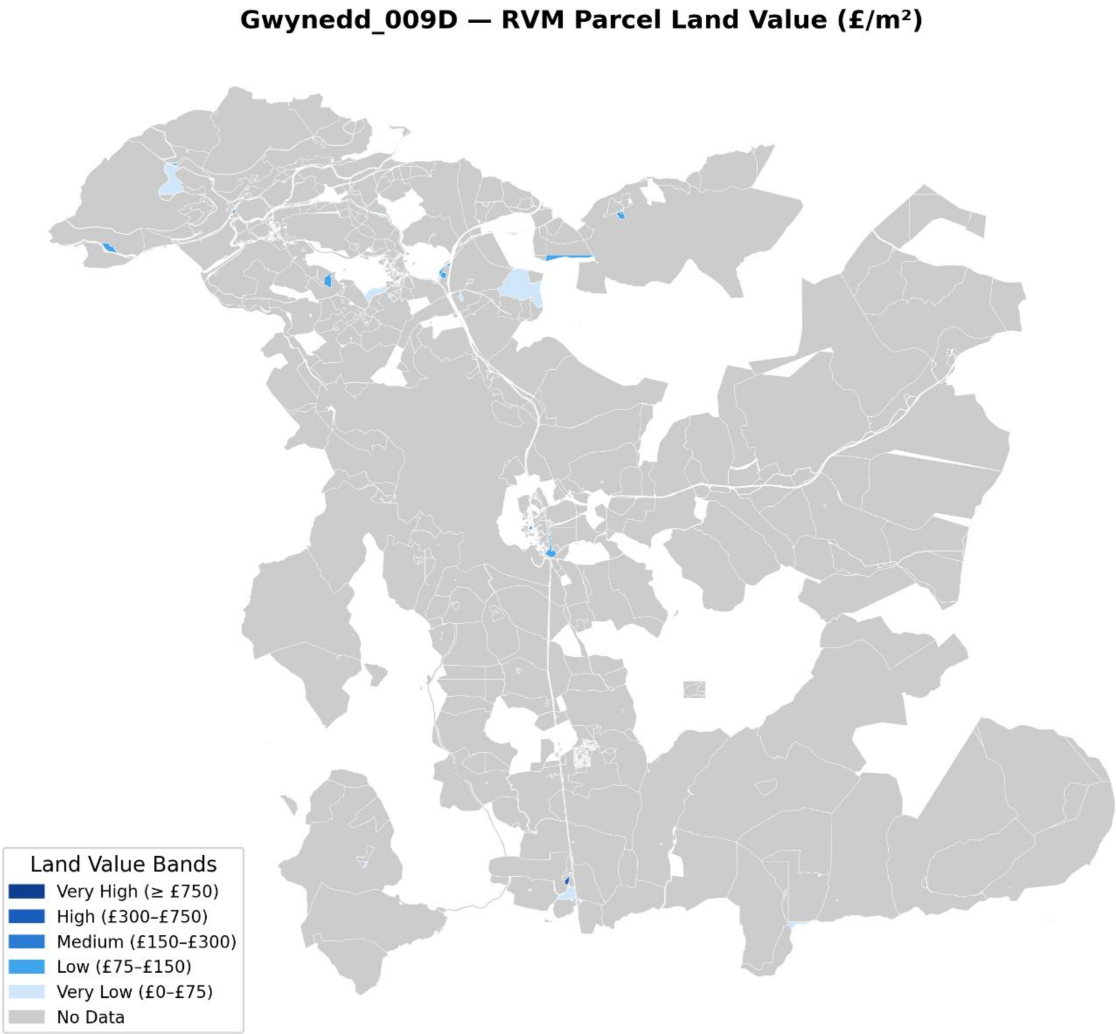

Figure 3: Gwynedd 009D RVM Results

Gwynedd exhibits moderate average land intensity with limited downside exposure. The minimum value is only marginally negative, suggesting that most parcels are broadly viable under central assumptions. The substantial upper bound reflects strong performance in higher-GDV or more efficient parcels. The range indicates rural heterogeneity but relatively contained structural risk.

Flintshire 015A (W01000255)

- **Average LV/m²:** £124.77
- **Minimum LV/m²:** –£124.98
- **Maximum LV/m²:** £651.18
- **Range:** £776.16

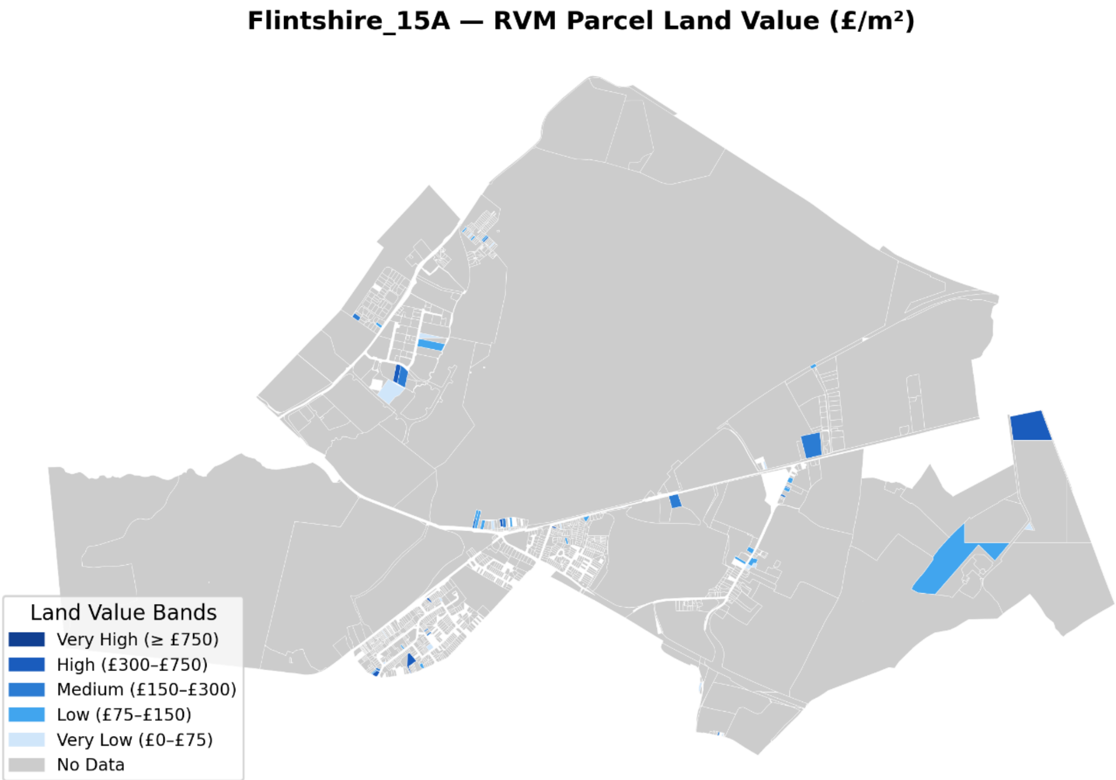

Figure 4: Flintshire 015A RVM Results

Flintshire shows stronger central land intensity than Gwynedd but with deeper downside exposure. The negative lower bound (–£124.98/m²) indicates that certain parcels remain materially unviable under stress conditions. The upper bound suggests strong performance in favourable sub-markets. Overall dispersion reflects transitional semi-urban market dynamics.

Powys 011C (W01000449)

- **Average LV/m²:** £63.79
- **Minimum LV/m²:** –£29.17
- **Maximum LV/m²:** £314.00
- **Range:** £343.17

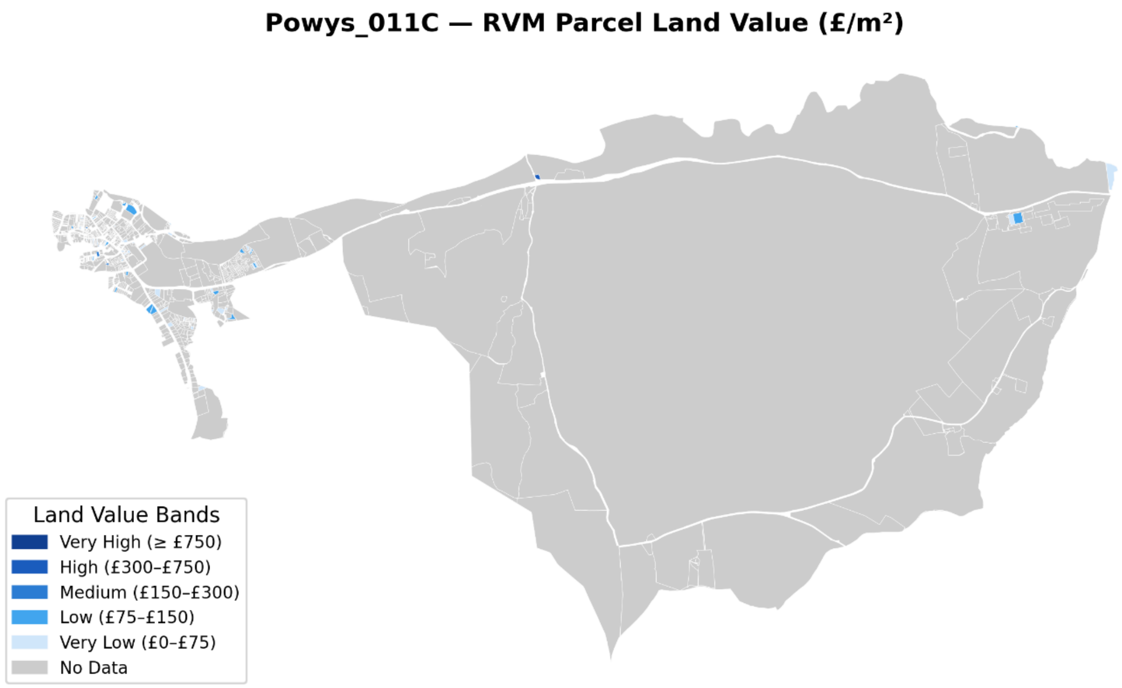

Figure 5: Powys 011C RVM Results

Powys returns relatively modest average land intensity and moderate dispersion. While some parcels fall below viability threshold, negative exposure is not extreme. The narrower range relative to other LSOAs suggests more constrained GDV amplification effects in rural agricultural contexts.

Ceredigion 002D (W01000517)

- **Average LV/m²:** £136.73
- **Minimum LV/m²:** –£4.06
- **Maximum LV/m²:** £1,013.41
- **Range:** £1,017.47

Figure 6: Ceredigion 002D RVM Results

Ceredigion demonstrates strong average land intensity with limited downside exposure. The upper-bound value exceeding £1,000/m² indicates highly favourable parcels benefiting from coastal or amenity-driven GDV uplift. The asymmetric distribution reflects significant upside potential relative to modest downside risk.

Pembrokeshire 002F (W01000617)

- **Average LV/m²:** £215.26
- **Minimum LV/m²:** £2.03

- **Maximum LV/m²:** £1,056.69
- **Range:** £1,054.66

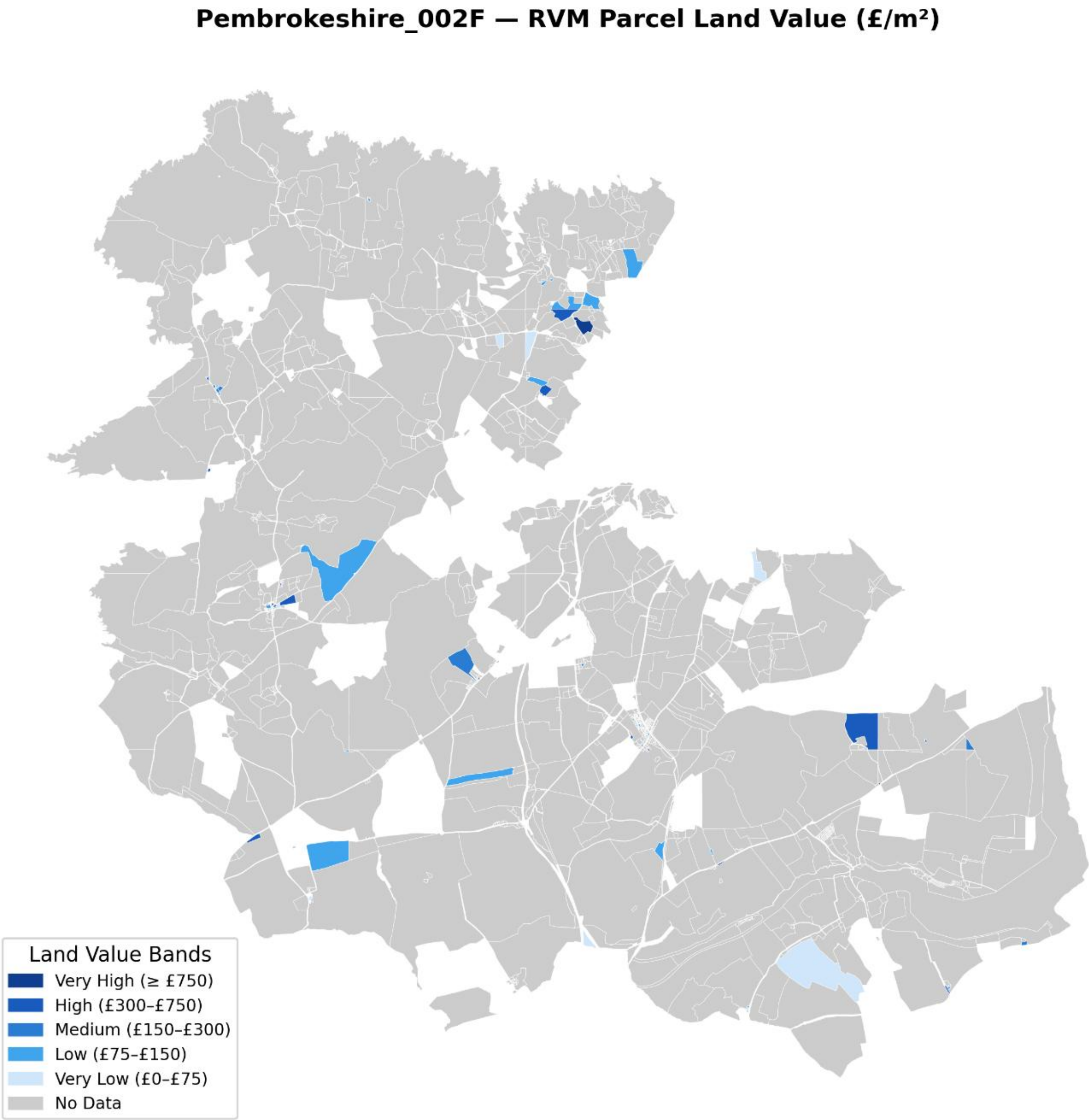

Figure 7: Pembrokeshire 002F RVM Results

Pembrokeshire is notable for having no negative parcel-level LV/m² values. Even the lowest parcel remains positive, indicating structurally resilient development viability under current assumptions. The high maximum suggests strong amplification effects in premium coastal parcels.

Bridgend 019D (W01001045)

- **Average LV/m²:** £297.06
- **Minimum LV/m²:** £18.49
- **Maximum LV/m²:** £883.07

- **Range:** £864.57

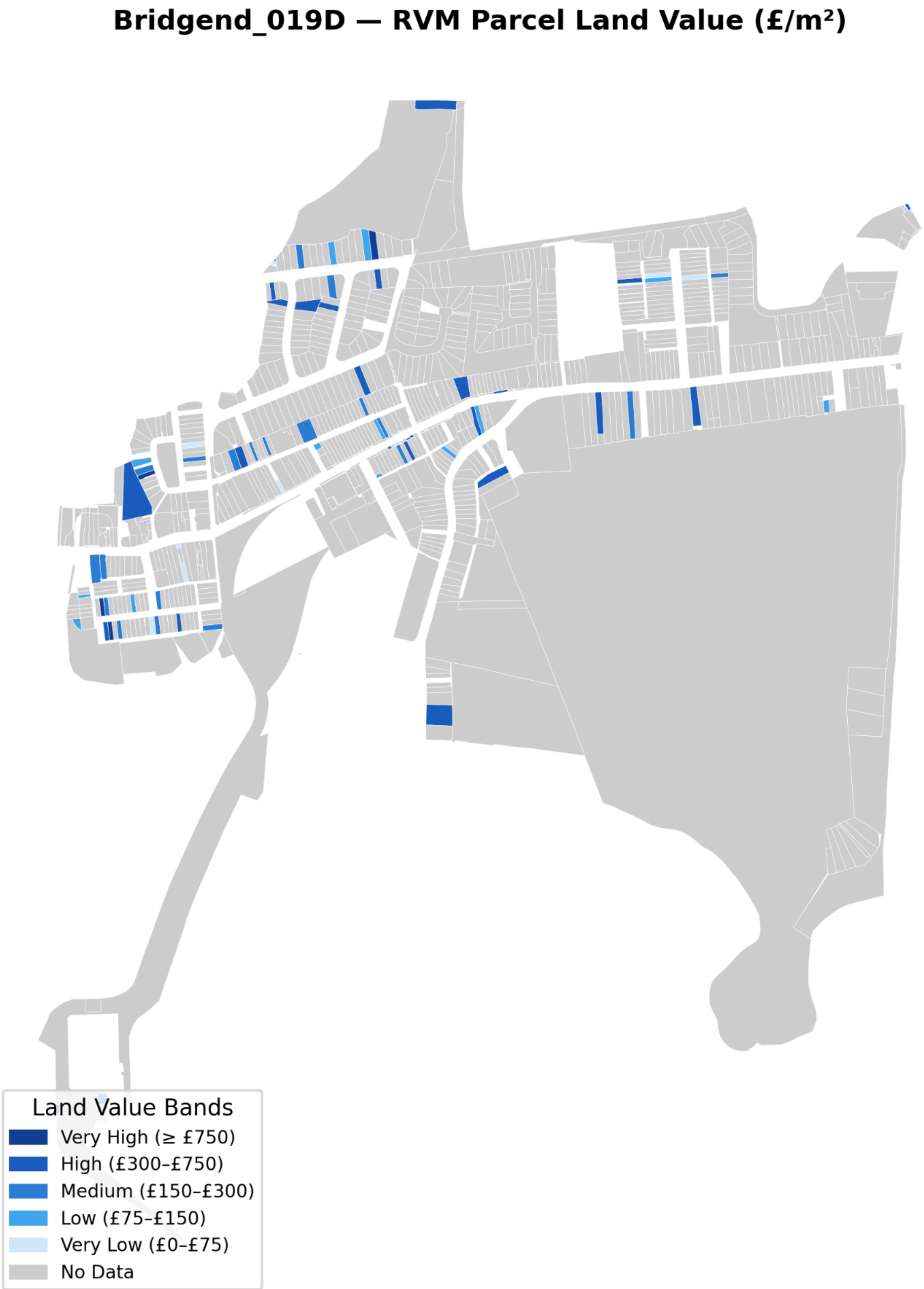

Figure 8: Bridgend 019D RVM Results

Bridgend exhibits strong and stable land intensity. The strictly positive minimum confirms consistent viability across parcels. The high upper bound reflects suburban growth dynamics and efficient development density. Dispersion is significant but structurally positive.

Rhondda Cynon Taf 001F (W01001233)

- **Average LV/m²:** £110.74
- **Minimum LV/m²:** –£343.81
- **Maximum LV/m²:** £1,571.63
- **Range:** £1,915.44

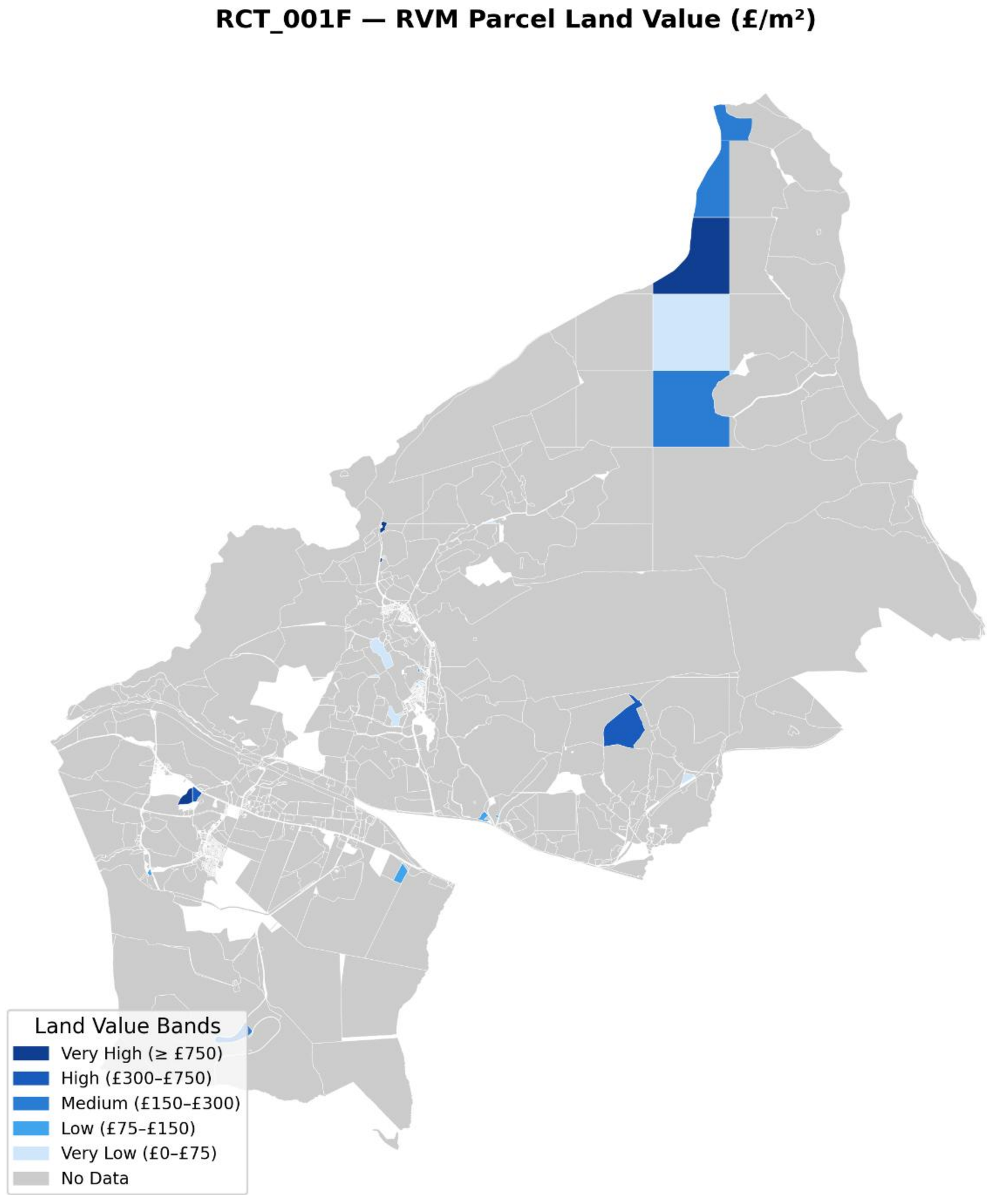

Figure 9: Rhondda Cynon Taf 001F RVM Results

RCT displays the greatest intra-LSOA volatility. The minimum value (–£343.81/m²) represents the deepest negative exposure across all study areas, indicating significant development risk in certain parcels. Conversely, the maximum value exceeds £1,500/m², demonstrating extreme sensitivity to parcel-specific GDV–cost alignment. This very wide range confirms structural heterogeneity and heightened risk concentration in post-industrial contexts.

Monmouthshire 006F (W01001597)

- **Average LV/m²:** £316.60
- **Minimum LV/m²:** £67.85
- **Maximum LV/m²:** £1,387.13
- **Range:** £1,319.27

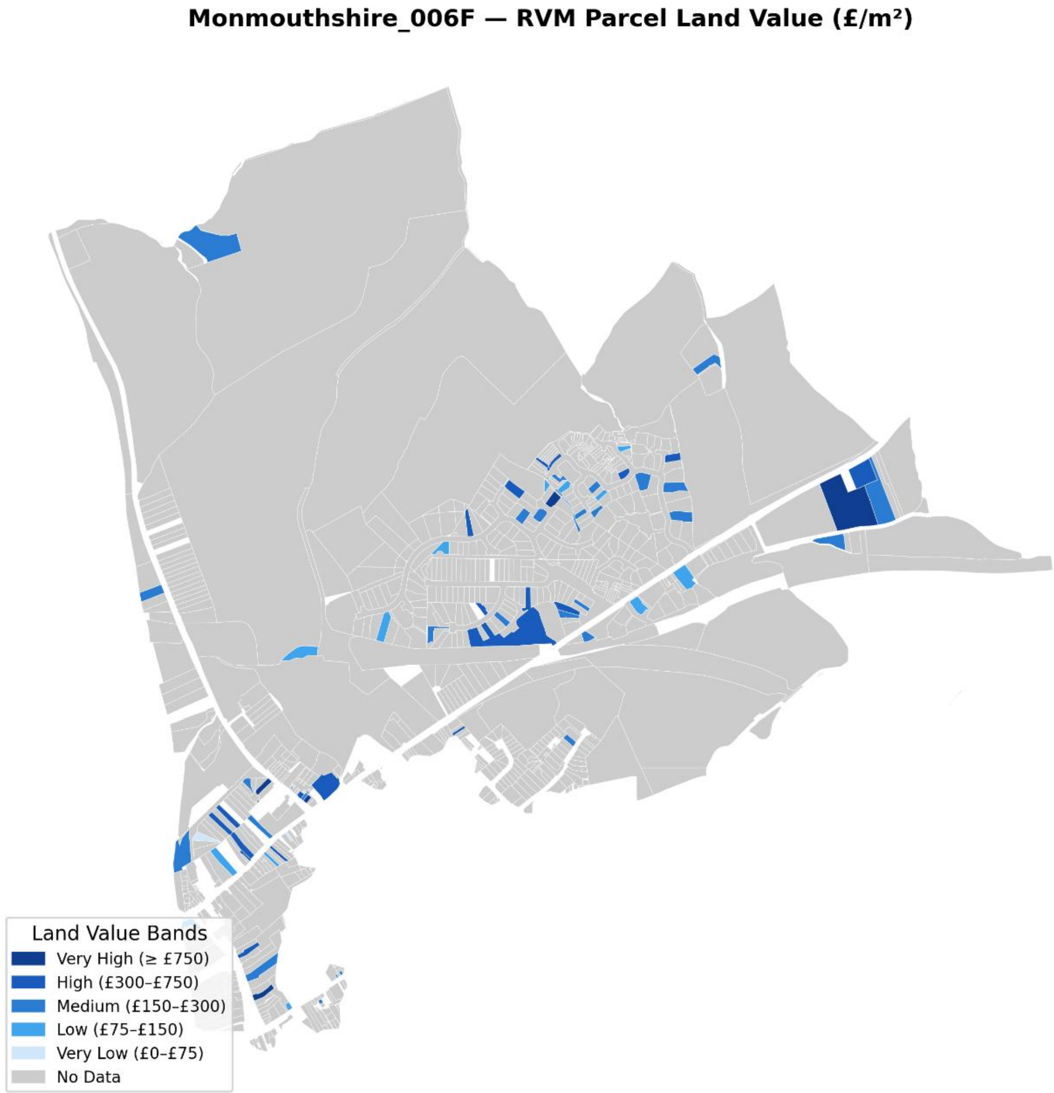

Figure 10: Monmouthshire 006F RVM Results

Monmouthshire produces the strongest central land intensity among the study areas, with all parcels remaining positive. The high upper bound reflects strong commuter-market GDV performance. The large range is driven primarily by upside amplification rather than downside risk.

Cardiff 032H (W01002019)

- **Average LV/m²:** £553.37
- **Minimum LV/m²:** £0.93
- **Maximum LV/m²:** £2,396.47
- **Range:** £2,395.55

Figure 11: Cardiff 032H RVM Results

Cardiff exhibits the highest average land intensity across all LSOAs. The minimum value is effectively neutral but positive, indicating near-breakeven parcels under conservative assumptions. The maximum value of £2,396/m² demonstrates substantial GDV-driven amplification in high-density urban locations. The very wide range reflects strong vertical density efficiencies and urban market stratification.

Cross-LSOA Interpretation

Several clear patterns emerge:

- **Highest average land intensity:** Cardiff (£553/m²), followed by Monmouthshire (£317/m²) and Bridgend (£297/m²).
- **Strong coastal/suburban performance:** Pembrokeshire and Ceredigion show high upper-bound values with limited downside exposure.
- **Greatest volatility:** RCT exhibits the widest range (£1,915/m²), reflecting extreme parcel-level sensitivity.
- **Deepest downside risk:** Flintshire and RCT show materially negative lower bounds.
- **Most structurally resilient:** Pembrokeshire and Monmouthshire show no negative LV/m² parcels.

The intra-LSOA LV/m² results confirm that development-led land value is highly parcel-specific and spatially heterogeneous. Stronger commuter and urban markets demonstrate amplified land intensity, while post-industrial and transitional markets exhibit greater volatility and downside risk.

These findings reinforce the importance of bounded scenario testing and support the continued use of a hybrid valuation framework to stabilise fiscal interpretation.

3.3 Analysis of the Income Capitalisation Model (ICM) Results

The Income Capitalisation Model (ICM) estimates land value by isolating land-attributable net operating income and capitalising it using use-class-specific yield assumptions. Unlike the Residual Value Model, which reflects forward-looking development viability, the ICM captures the value of existing income-generating capacity within each LSOA.

The ICM results show materially higher land values than the RVM in most study areas, alongside significantly wider intra-LSOA dispersion. This reflects the leverage effect of capitalisation, whereby relatively small differences in income or yield assumptions can generate large swings in capital value.

Across the nine study LSOAs, mean land values range from **£84,524** (Powys 011C) to **£1,541,422** (Cardiff 032H), confirming that income-driven land value is highly sensitive to urban intensity, commercial activity and rental benchmarks.

Table 2: LSOA Level Outputs for Income Capitalisation Modelling

LSOA Name	Mean RLV (£)	Median RLV (£)	P5 (£)	P95 (£)	Parcel Count
Gwynedd 009D W01000114	128,228.70	145,518.40	-143,089.00	400,829.70	15
Flintshire 015A W01000255	368,989.60	392,062.60	-296,384.00	1,204,292.00	30
Powys 011C W01000449	84,524.12	15,399.39	-98,227.10	408,553.80	10
Ceredigion 002D W01000517	150,657.00	93,502.22	-295,132.00	637,738.00	15
Pembrokeshire 002F W01000617	234,036.60	152,555.40	-120,698.00	754,413.30	20
Bridgend 019D W01001045	536,549.40	588,227.40	-231,305.00	1,303,570.00	25
Rhondda Cynon Taf 001F W01001233	311,054.10	314,432.80	-47,501.00	697,217.60	15
Monmouthshire 006F W01001597	324,451.80	251,676.10	-253,022.00	836,850.60	20
Cardiff 032H W01002019	1,541,422.00	1,731,178.00	-131,608.00	3,107,259.00	35

Central Tendency and Market Hierarchy

The highest mean and median land values occur in **Cardiff 032H**, with:

- **Mean LV:** £1,541,422
- **Median LV:** £1,731,178

This reflects strong rateable value density, significant income-producing floorspace and capitalisation amplification in a high-demand urban centre.

The next tier of high-performing areas includes:

- **Bridgend 019D** – Mean £536,549; Median £588,227
- **Flintshire 015A** – Mean £368,990; Median £392,063
- **Monmouthshire 006F** – Mean £324,452; Median £251,676
- **Rhondda Cynon Taf 001F** – Mean £311,054; Median £314,433

These areas demonstrate substantial income-based land values, indicating that even where development-led viability (RVM) may be moderate, income-generating activity produces stronger capitalised land value signals.

Lower average values are observed in:

- **Powys 011C** – Mean £84,524
- **Gwynedd 009D** – Mean £128,229
- **Ceredigion 002D** – Mean £150,657

These LSOAs reflect weaker commercial intensity and lower aggregate rental bases.

Distributional Profile and Volatility

The ICM results display substantially wider dispersion than the RVM results, as demonstrated by the 5th and 95th percentile values.

Downside Risk (p5)

All nine LSOAs exhibit negative 5th percentile values, indicating that some parcels produce negative land-attributable income after accounting for OPEX, depreciation and return on improvements.

The most pronounced lower-bound exposures are observed in:

- **Flintshire 015A:** –£296,384
- **Ceredigion 002D:** –£295,132
- **Monmouthshire 006F:** –£253,022
- **Bridgend 019D:** –£231,305

Even high-performing markets exhibit negative lower-bound values. This reflects the mechanical sensitivity of income capitalisation to:

- Low or marginal rental income,
- Higher operating expenses,
- Depreciation impacts,
- Yield amplification effects.

Unlike the RVM, where negative residuals reflect development unviability, negative ICM lower-bound values typically arise where income is insufficient to support positive land-attributable NOI under assumed parameters.

Upside Amplification (p95)

The upper-bound values are materially higher than those observed under the RVM:

- **Cardiff 032H:** £3,107,259
- **Bridgend 019D:** £1,303,570
- **Flintshire 015A:** £1,204,292
- **Monmouthshire 006F:** £836,851
- **Pembrokeshire 002F:** £754,413

These high upper-bound values illustrate the leverage effect of yield capitalisation. In parcels where:

- Rateable values are high,
- OPEX ratios are moderate,
- Building depreciation is lower,
- Capitalisation rates are favourable,

land value expands rapidly.

This asymmetric dispersion is a defining characteristic of income-based valuation.

Intra-LSOA Variability

The spread between p5 and p95 values highlights extreme intra-LSOA heterogeneity:

- **Cardiff:** Range exceeding £3.2 million
- **Flintshire:** Range exceeding £1.5 million
- **Bridgend:** Range exceeding £1.5 million

Even rural LSOAs such as Powys and Gwynedd show substantial spreads (approximately £500,000 range), reflecting variability in non-domestic income concentration and parcel-level rental exposure.

Notably, **Powys 011C** has a very low median (£15,399) relative to its mean (£84,524), indicating skewed distribution driven by a small number of higher-income parcels. This suggests that income-based land value in Powys is concentrated rather than evenly distributed.

Spatial and Economic Interpretation

The ICM results broadly align with expected economic geography:

- **Urban centres generate the strongest income-capitalised land values.**
- **Mixed suburban and commuter markets exhibit robust secondary performance.**
- **Rural areas show lower central values but still meaningful income-derived land signals.**

Interestingly, some LSOAs (e.g., RCT 001F) display significantly stronger ICM results than RVM results. This indicates that, while development-led viability may be constrained, existing income-generating activity still supports substantial capitalised land value.

This divergence reinforces the complementary nature of the two modelling approaches.

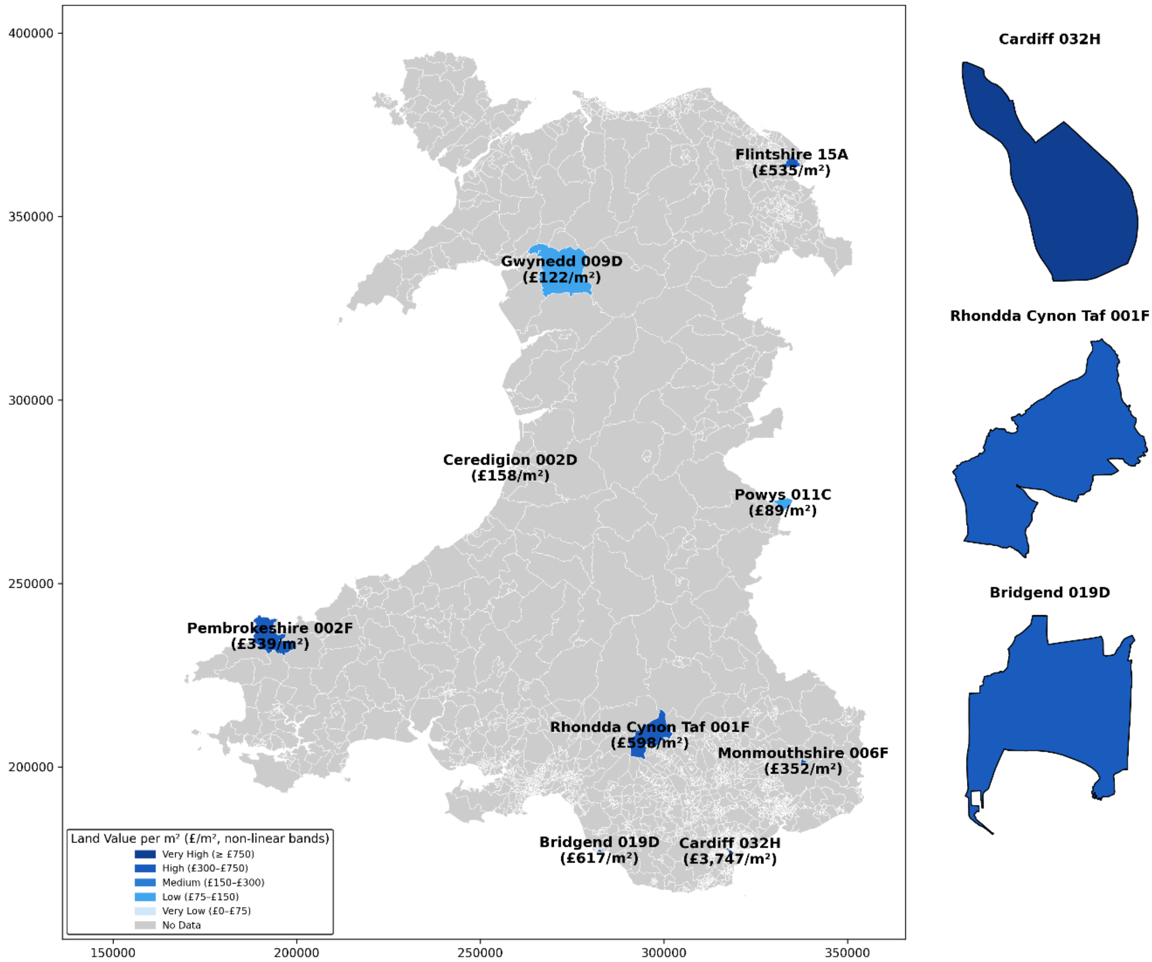

Figure 12: LSOA Outputs Mapped for ICM

Comparative Observations Relative to RVM

Relative to the RVM:

- ICM land values are materially higher in most LSOAs.
- Volatility is greater due to capitalisation amplification.
- Downside risk appears in all LSOAs, even strong markets.
- Urban markets exhibit extreme upside dispersion.

The ICM captures a fundamentally different economic signal: **existing income strength rather than development feasibility**. In urban and employment-dense areas, this produces substantially larger land value estimates.

Analytical Implications

The ICM results confirm:

1. Income-driven land value in Wales is highly concentrated in urban centres.
2. Yield assumptions materially influence capitalised outcomes.
3. Negative lower-bound values arise even in high-performing markets due to income variability.
4. The model is inherently high-variance and sensitive to OPEX, depreciation and yield parameters.

These characteristics make ICM outputs powerful but volatile indicators of land value.

Overall Conclusion on ICM Performance

The Income Capitalisation Model produces:

- Strong central land value signals in urban and commuter markets;
- Meaningful income-based land value even in weaker markets;
- Substantial intra-LSOA dispersion driven by income concentration;
- Higher absolute values and volatility relative to the RVM.

The results confirm that the ICM is highly effective in capturing income-based land intensity but must be interpreted within a bounded scenario framework due to its sensitivity to rental and yield assumptions.

The divergence between RVM and ICM outputs further supports the strategic importance of a hybrid valuation framework in the final phase.

3.3.1 Intra-LSOA ICM Results

This section examines parcel-level variation in **Land Value per m² (LV/m²)** derived from the Income Capitalisation Model (ICM) within each study LSOA. Expressing ICM results in £/m² enables consistent comparison across parcels and highlights the intensity and volatility of income-driven land value at micro-spatial scale.

Unlike the RVM, which exhibited moderate dispersion, the ICM results show substantially wider intra-LSOA ranges, reflecting the leverage effect of capitalisation and the sensitivity of land value to rental density, OPEX assumptions and yield parameters.

Gwynedd 009D (W01000114)

- **Average LV/m²:** £121.52
- **Minimum LV/m²:** –£358.07
- **Maximum LV/m²:** £869.60
- **Range:** £1,227.67

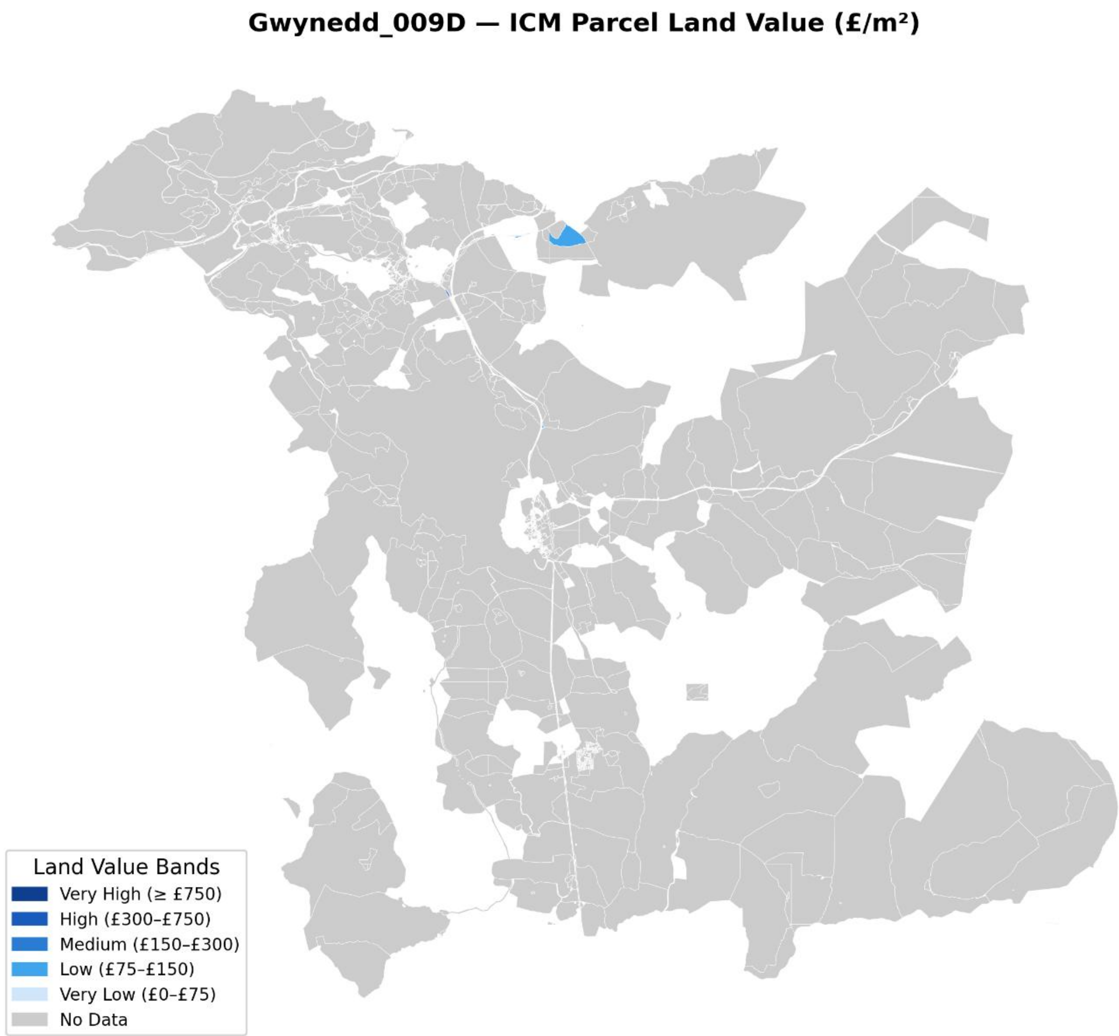

Figure 13: Gwynedd 009D ICM Results

Gwynedd exhibits moderate average land intensity but with material downside exposure. The negative lower bound suggests that certain parcels generate insufficient land-attributable income after building return and OPEX deductions. However, the upper bound demonstrates meaningful income-driven value in stronger sub-locations. The dispersion reflects rural income concentration effects and heterogeneous non-domestic activity.

Flintshire 015A (W01000255)

- **Average LV/m²:** £534.64
- **Minimum LV/m²:** –£374.52
- **Maximum LV/m²:** £4,035.52
- **Range:** £4,410.04

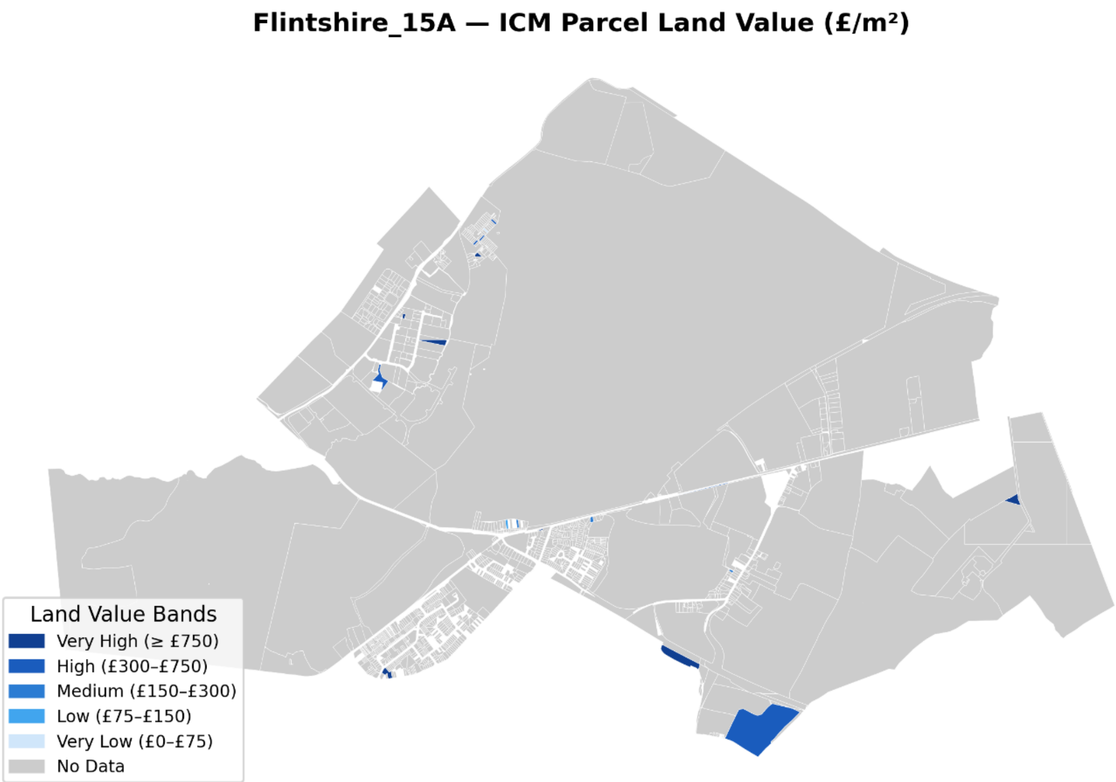

Figure 14: Flintshire 015A ICM Results

Flintshire displays strong average land intensity with extremely wide dispersion. The upper bound exceeding £4,000/m² reflects parcels with high rateable values capitalised at relatively tight yields. Conversely, the negative lower bound indicates substantial variability in income strength across parcels. This LSOA shows one of the widest income-based volatility profiles outside Cardiff.

Powys 011C (W01000449)

- **Average LV/m²:** £88.72
- **Minimum LV/m²:** –£118.99
- **Maximum LV/m²:** £788.04
- **Range:** £907.04

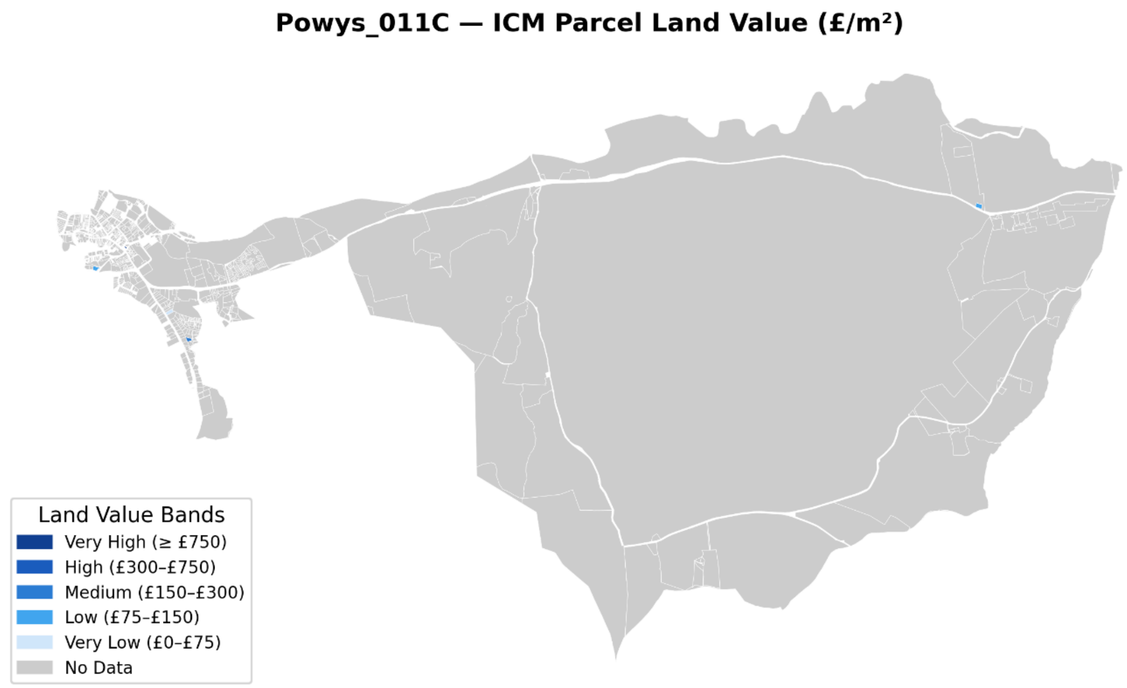

Figure 15: Powys 011C ICM Results

Powys demonstrates relatively modest income-driven land intensity with moderate volatility. The negative lower bound suggests marginal income viability in some parcels, while the upper bound indicates concentrated commercial income pockets. Compared to urban LSOAs, the income signal is weaker and more uneven.

Ceredigion 002D (W01000517)

- **Average LV/m²:** £157.89
- **Minimum LV/m²:** –£701.02
- **Maximum LV/m²:** £766.82
- **Range:** £1,467.85

Figure 16: Ceredigion 002D ICM Results

Ceredigion shows significant volatility relative to its average intensity. The lower bound (–£701/m²) is materially negative, indicating high sensitivity to income assumptions in certain parcels. However, the upper bound is strong. This asymmetric distribution reflects seasonal and concentrated income streams typical of coastal markets.

Pembrokeshire 002F (W01000617)

- **Average LV/m²:** £338.60
- **Minimum LV/m²:** –£174.40
- **Maximum LV/m²:** £3,462.09
- **Range:** £3,636.49

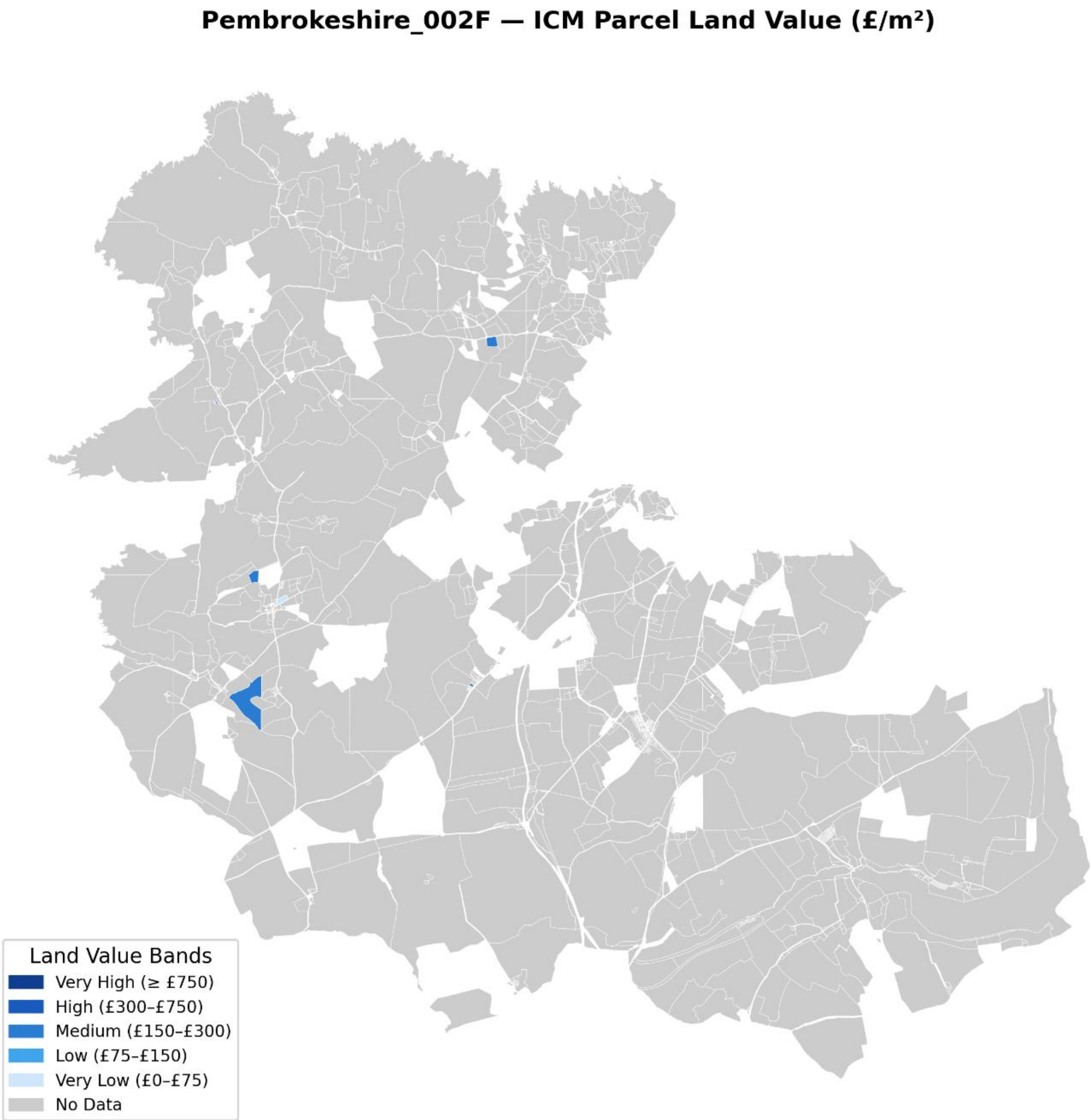

Figure 17: Pembrokeshire 002F ICM Results

Pembrokeshire demonstrates strong income-based land intensity with substantial upside dispersion. The upper bound above £3,400/m² indicates highly favourable income-yield alignment in select parcels. The negative lower bound confirms variability but overall central performance is robust.

Bridgend 019D (W01001045)

- **Average LV/m²:** £617.42
- **Minimum LV/m²:** –£599.81
- **Maximum LV/m²:** £2,617.71
- **Range:** £3,217.52

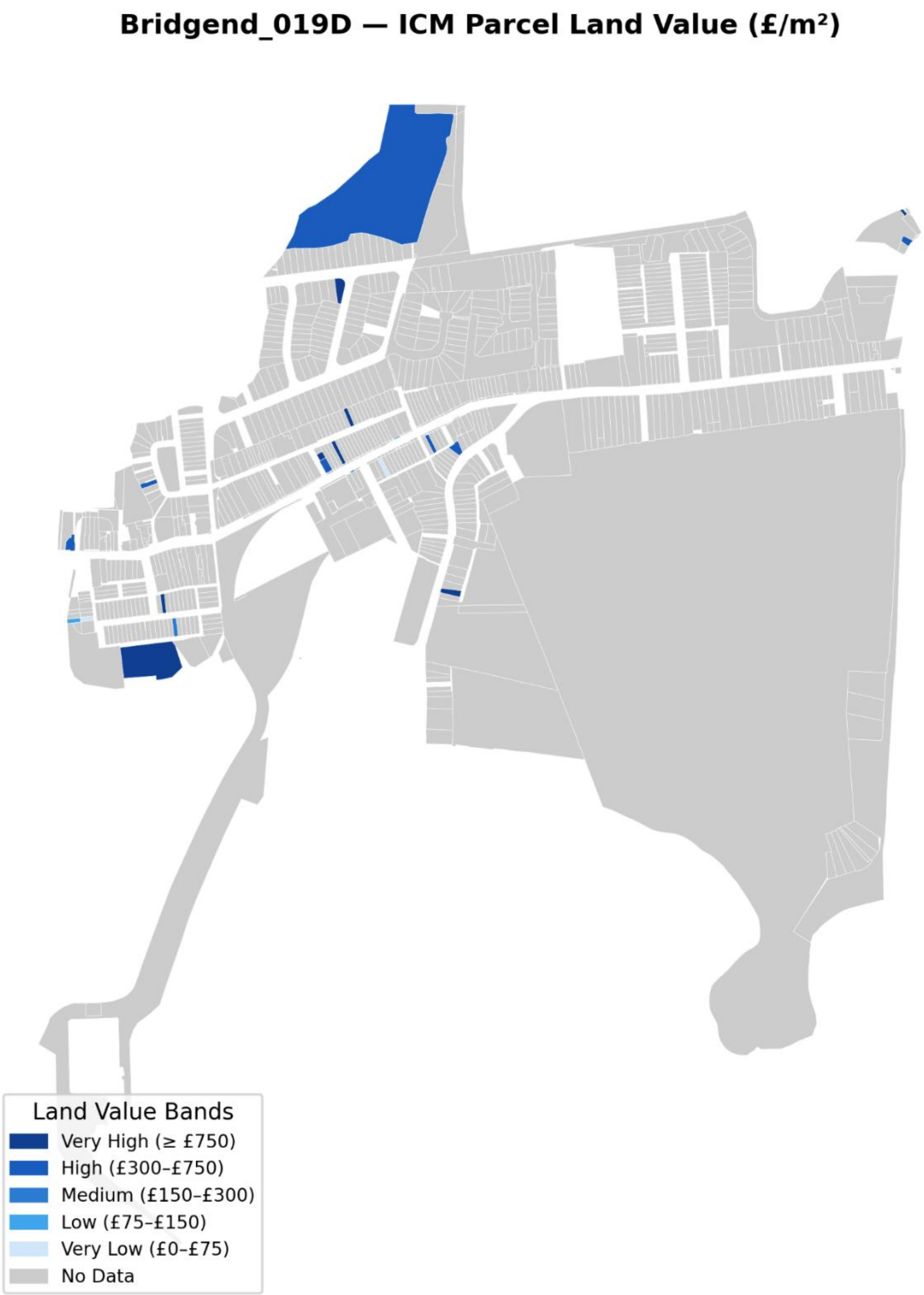

Figure 18: Bridgend 019D ICM Results

Bridgend exhibits high average land intensity and substantial volatility. The deep negative minimum reflects parcels where operating income fails to support land value after improvement return adjustments. The strong upper bound demonstrates pronounced yield amplification in higher-income parcels. Overall, income concentration effects are significant.

Rhondda Cynon Taf 001F (W01001233)

- **Average LV/m²:** £598.21
- **Minimum LV/m²:** –£145.00
- **Maximum LV/m²:** £3,570.85
- **Range:** £3,715.86

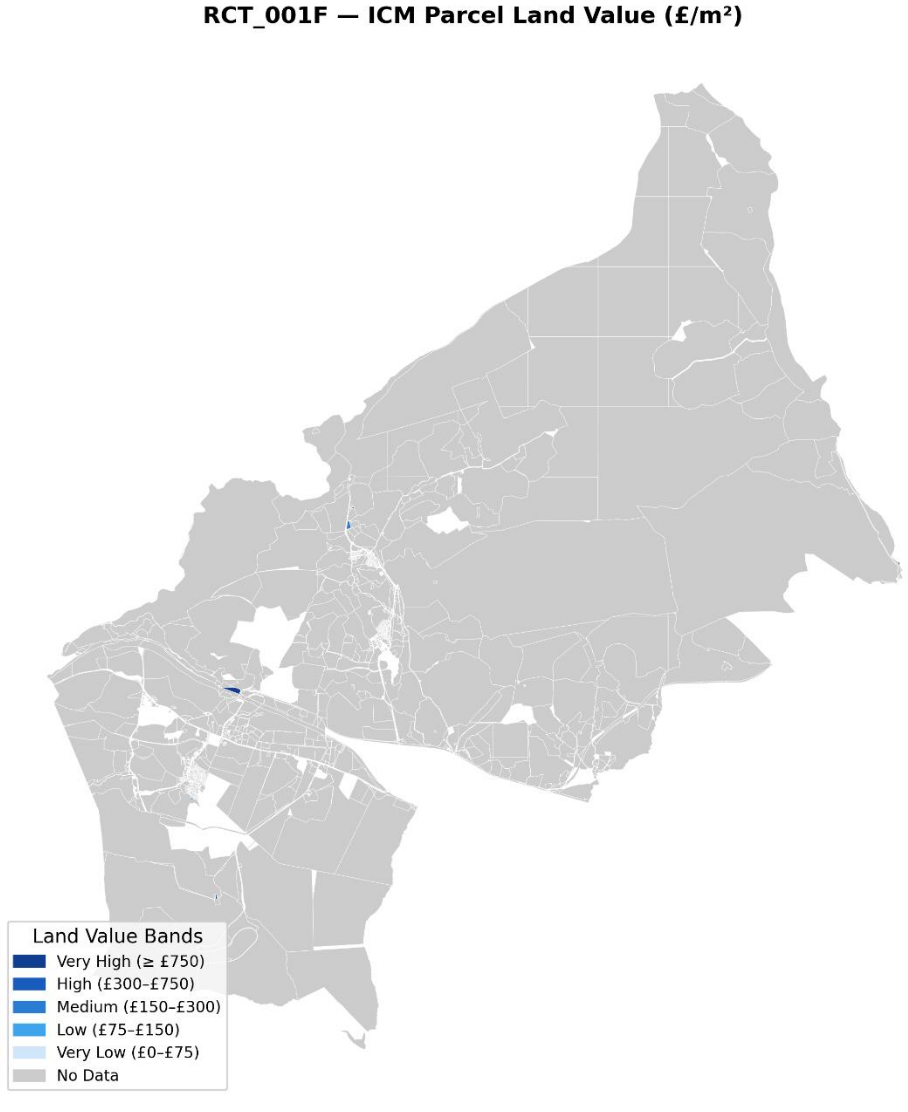

Figure 19: Rhondda Cynon Taf 001F ICM Results

RCT shows strong average income-based land intensity, significantly exceeding its RVM intensity profile. This indicates that while development viability may be thinner, existing income streams generate substantial capitalised value. The wide dispersion confirms income concentration effects and parcel-level heterogeneity.

Monmouthshire 006F (W01001597)

- **Average LV/m²:** £352.35
- **Minimum LV/m²:** –£595.64
- **Maximum LV/m²:** £1,930.99
- **Range:** £2,526.63

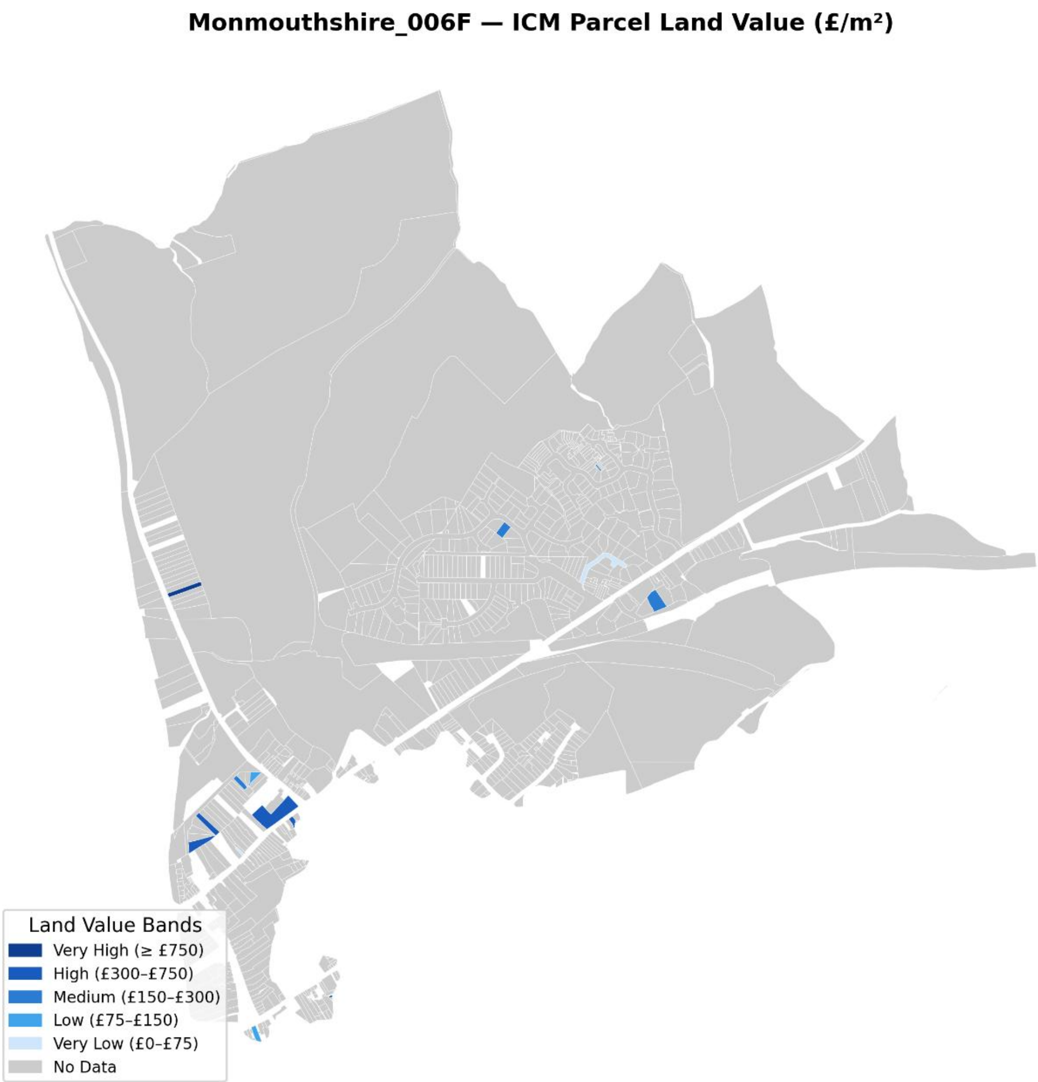

Figure 20: Monmouthshire 006F ICM Results

Monmouthshire exhibits strong central performance with significant upper-bound amplification. Negative lower-bound values reflect variability in income durability and building-adjusted NOI calculations. Overall, income-driven land value is materially positive and resilient.

Cardiff 032H (W01002019)

- **Average LV/m²:** £3,747.34
- **Minimum LV/m²:** –£1,081.93
- **Maximum LV/m²:** £16,813.32
- **Range:** £17,895.24

Figure 21: Cardiff 032H ICM Results

Cardiff displays by far the highest income-driven land intensity across all study areas. The average of £3,747/m² reflects concentrated commercial rental density and capitalisation effects in a high-demand urban core.

The extreme upper bound (£16,813/m²) demonstrates strong yield leverage where high rateable values intersect with favourable cap rates. However, the presence of negative lower-bound values indicates that not all parcels sustain strong land-attributable NOI after improvement return adjustments.

Cardiff's range of nearly £18,000/m² highlights the amplification effect of income capitalisation in dense urban environments.

Cross-LSOA Observations

Several structural patterns emerge from the parcel-level ICM analysis:

- **Income-based land intensity is highest in Cardiff**, followed by Bridgend, RCT and Flintshire.
- **Volatility is materially greater under ICM than RVM**, reflecting yield amplification.
- **All LSOAs exhibit negative lower-bound parcels**, confirming sensitivity to income durability and cost allocation.
- **Urban and suburban areas show extreme upside dispersion**, driven by rental density.
- **Rural areas show moderate averages but still significant range**, reflecting income concentration rather than uniform activity.

Compared to the RVM intra-LSOA results, ICM dispersion is substantially wider. This confirms that income capitalisation produces higher absolute land values but also greater parcel-level volatility.

Analytical Implications

The intra-LSOA ICM results confirm that income-driven land value in Wales is:

- Highly concentrated in urban and commercial cores;
- Amplified by yield sensitivity;
- Subject to significant parcel-level volatility;
- Structurally higher in absolute terms than development-led land value in most areas.

These findings reinforce the importance of interpreting ICM outputs within bounded scenario corridors and further support the development of a hybrid valuation framework to stabilise fiscal interpretation.

3.4 Analysis of the Combined Results

The combined land value results integrate the outputs of the Residual Value Model (RVM) and the Income Capitalisation Model (ICM) into a single hybrid land value indicator expressed in **£ per m²**. Because both models were applied to an unequal number of parcels within each LSOA, the weighted combined land value per m² represents the combination of development-led and income-led land valuation.

This hybrid measure was designed to address the structural limitations inherent in relying exclusively on either model. The RVM captures forward-looking development feasibility under defined cost and profit assumptions, while the ICM reflects existing income-generating capacity capitalised at market yields. Individually, each model provides a partial perspective. Combined, they produce a more stable and policy-relevant land value signal.

Table 3: Combined LSOA-scale outputs from the RVM and ICM

LSOA Name	RVM Land Value (£/m^2)	ICM Land Value (£/m^2)	Weighted Combined Land Value (£/m^2)
Gwynedd 009D W01000114	87.97	121.52	104.74
Flintshire 015A W01000255	124.77	534.64	329.71
Powys 011C W01000449	63.79	88.72	76.26
Ceredigion 002D W01000517	136.73	157.89	147.31
Pembrokeshire 002F W01000617	215.26	338.60	276.93
Bridgend 019D W01001045	297.06	617.42	457.24
Rhondda Cynon Taf 001F W01001233	110.74	598.21	354.48
Monmouthshire 006F W01001597	316.60	352.35	334.47
Cardiff 032H W01002019	553.37	3,747.34	2,150.36

Overall Land Value Gradient Across Study Areas

The combined land value per m² results exhibit a clear and economically intuitive spatial gradient across the nine study LSOAs.

At the upper extreme, **Cardiff 032H** produces a combined land value intensity substantially above all other study areas. This reflects the interaction of strong development viability (RVM) and extremely high income capitalisation (ICM) in a dense metropolitan core. Even after moderating the ICM amplification effect through hybridisation, Cardiff remains structurally distinct in terms of land intensity.

A second tier of LSOAs demonstrates strong but materially lower combined values. This includes:

- **Bridgend 019D**,
- **Rhondda Cynon Taf 001F**,
- **Monmouthshire 006F**,

- **Flintshire 015A**.

These areas benefit from a combination of viable development conditions and meaningful income-generating activity. In several of these LSOAs, the income signal significantly elevates the hybrid value relative to RVM alone, indicating that existing rental and commercial intensity materially supports land value beyond pure development feasibility.

A third tier comprises:

- **Pembrokeshire 002F**,
- **Ceredigion 002D**,
- **Gwynedd 009D**,
- **Powys 011C**.

These areas exhibit more modest land intensity. In rural and peripheral markets, both development viability and income capitalisation produce lower central values. The hybrid measure therefore reflects structural economic constraints rather than modelling artefacts.

Moderating Effect of Hybridisation

A key analytical objective of the combined approach was to stabilise volatility observed in the standalone models.

Under the RVM, several LSOAs exhibited moderate land intensity but relatively narrow upside potential. Under the ICM, by contrast, urban and income-rich areas displayed extreme amplification due to capitalisation leverage.

The hybrid measure reduces this amplification effect by averaging the two signals. For example:

- In **Rhondda Cynon Taf 001F**, RVM results suggested relatively modest development-led land intensity, whereas ICM results were materially higher. The hybrid value sits between these two extremes, reflecting both constrained development viability and meaningful income concentration.
- In **Flintshire 015A**, the income signal substantially exceeded the development signal. The combined result moderates this divergence, preventing overstatement of land value driven solely by yield compression.
- In **Cardiff 032H**, ICM values were exceptionally high due to dense rateable value concentration and capitalisation effects. The hybrid measure tempers this by incorporating the more moderate RVM result, while still preserving Cardiff's structural dominance.

In rural LSOAs such as **Powys 011C**, where both RVM and ICM means are modest, the combined result confirms structural economic limitation rather than model instability.

The hybridisation therefore achieves two important stabilising functions:

1. It reduces the volatility inherent in income capitalisation in high-density markets.
2. It prevents development-only modelling from understating land value in income-active areas.

Divergence Between Development and Income Signals

The combined results also illuminate the structural divergence between development feasibility and income-generating capacity.

In several LSOAs - notably **RCT 001F**, **Flintshire 015A**, and **Bridgend 019D** - the ICM mean materially exceeds the RVM mean. This indicates that existing income streams are sustaining land value beyond what would be inferred from development appraisal alone.

Conversely, in certain rural areas, the RVM and ICM means are closer in magnitude, suggesting alignment between limited development potential and modest income concentration.

This divergence confirms that land value is not a singular construct. It may be driven by:

- Future development potential (RVM),
- Existing income capacity (ICM),
- Or a combination of both.

The hybrid measure acknowledges this duality.

Relative Positioning of LSOAs

When assessed comparatively, three structural tiers emerge:

Metropolitan Core:
Cardiff 032H — exceptionally high combined land intensity reflecting concentrated rental density and capitalised income streams.

Strong Secondary Markets:
Bridgend 019D, RCT 001F, Monmouthshire 006F, Flintshire 015A — balanced markets with both viable development and income concentration.

Peripheral and Rural Markets:
Pembrokeshire 002F, Ceredigion 002D, Gwynedd 009D, Powys 011C — structurally lower land intensity reflecting thinner rental bases and lower GDV benchmarks.

Importantly, the combined results preserve differentiation without producing artificial clustering or volatility spikes.

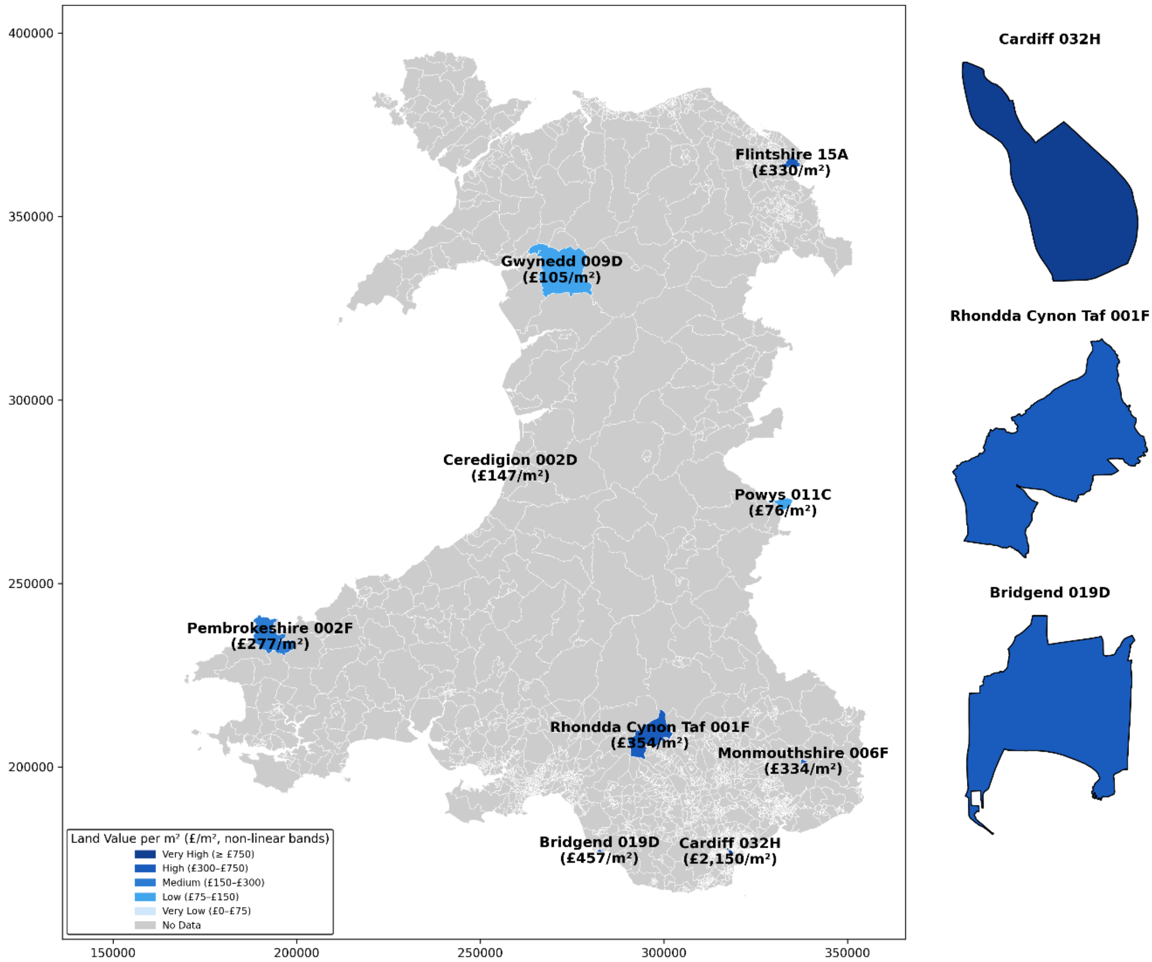

Figure 22: Mapped Combined Outputs of the RVM and ICM

Implications for Interpretation and Policy

The hybrid land value per m² results provide a more stable and defensible indicator of land intensity than either standalone model.

From a fiscal or policy perspective, the combined measure:

- Mitigates over-amplification from yield-driven income capitalisation;
- Avoids underestimation of land value in income-active areas;
- Produces a coherent Wales-wide gradient; and
- Reflects both structural economic strength and development feasibility.

It therefore offers a more suitable foundation for:

- Comparative spatial benchmarking;

- Scenario modelling;
- Sensitivity analysis; and
- Potential fiscal reform exploration.

The combined outputs are not merely arithmetic averages; they represent a structural reconciliation of two distinct valuation logics.

The combined land value per m² analysis confirms that:

- A hybrid valuation framework produces coherent, economically intuitive results;
- Development-led and income-led signals diverge materially in certain LSOAs;
- Hybridisation moderates volatility while preserving structural hierarchy;
- Cardiff remains structurally distinct, but other LSOAs exhibit meaningful differentiation; and
- The resulting gradient reflects genuine spatial-economic variation rather than modelling artefact.

The combined framework therefore represents the most analytically balanced and policy-ready interpretation of land value within the constraints of the open-data modelling environment.

Section 4 – Conclusions

This section draws together the findings from the application of the Residual Value Model (RVM), the Income Capitalisation Model (ICM), and the combined hybrid framework across the nine study LSOAs. It synthesises the analytical results presented in Sections 3.2–3.4 and interprets their implications for understanding land value across Wales within an open-data modelling environment.

The conclusions assess the performance, strengths and limitations of the modelling framework, highlight the structural drivers of variation observed across Welsh geographies, and evaluate the robustness and scalability of the approach. Particular attention is given to the feasibility of open-data land valuation at LSOA scale, the complementary insights generated by development-led and income-led valuation methods, and the stabilising role of the hybrid framework. The section concludes by identifying key sensitivities, data constraints and considerations relevant to potential national implementation.

4.1 Summary of Key Findings and Lessons Learned

The work completed demonstrates that a transparent, reproducible land valuation framework can be constructed using exclusively open datasets and applied consistently across Wales at LSOA scale. The Residual Value Model (RVM) and Income Capitalisation Model (ICM) were successfully implemented at parcel level and aggregated to LSOA level, producing coherent land-value gradients across urban, suburban, rural, coastal and post-industrial contexts.

The analysis confirms that open-data valuation is feasible at LSOA scale and that a nationally consistent, auditable modelling workflow can be delivered within current public data constraints. However, valuation precision becomes materially constrained at granular intra-LSOA and parcel scale. While parcel-level modelling is technically achievable, its reliability is bounded by the quality, completeness and resolution of input datasets.

In particular, the absence of:

- An openly licensed cadastral dataset,
- Detailed and current rental evidence,
- Site-specific planning constraints,
- Granular construction cost data,
- Parcel-level tenure and legal information,

limits the degree to which outputs can be interpreted as site-accurate market valuations. The framework therefore performs most robustly as a small-area statistical model rather than as a substitute for professional parcel-level appraisal.

A central lesson from the project is therefore clear. Valuation accuracy at granular scale is fundamentally limited by the quality and availability of input data. If enhanced parcel-level precision is required, access to more detailed administrative or proprietary datasets - beyond the current open-data ecosystem - would be necessary.

Dual-Method Framework: Complementary Insight

A major analytical finding is that the dual-model approach provides complementary and structurally distinct insight into land value formation.

- The **Residual Value Model (RVM)** captures development viability and responds primarily to GDV assumptions, construction costs, abnormal risk exposure and developer profit margins.
- The **Income Capitalisation Model (ICM)** captures existing income-generating capacity and responds to rental density, operating cost structures, depreciation assumptions and capitalisation rates.

In several LSOAs, these models produced materially different signals. In transitional or post-industrial markets, income-based land values exceeded development-led viability estimates, suggesting that existing income streams may sustain land value even where new development margins are thinner. In rural markets, both models converged at lower levels, reflecting structural economic constraints.

The divergence between RVM and ICM was not an inconsistency but an analytical strength. It demonstrated that land value is not driven by a single mechanism, but by the interaction of development feasibility and income concentration.

The weighted integration of RVM and ICM into a hybrid land value per m² measure proved critical to producing stable and interpretable results.

The hybrid framework:

- Moderated extreme yield-driven amplification observed in ICM outputs in dense urban cores;
- Prevented development-only modelling from understating land value in income-rich but development-constrained areas;
- Reduced volatility associated with reliance on a single valuation logic;
- Produced a coherent Wales-wide land intensity gradient.

The combined approach therefore represents the most balanced and policy-ready interpretation of land value within the constraints of open-data modelling.

4.2 Key Emerging Patterns

Several consistent structural patterns emerged across the study areas when comparing the Residual Value Model (RVM), the Income Capitalisation Model (ICM) and the hybrid land value per m² outputs.

A Clear Spatial Gradient Across Wales

Across all modelling approaches, a coherent and economically intuitive spatial hierarchy is evident.

Land values are consistently highest in strong-demand urban and commuter markets - most notably Cardiff, followed by Bridgend, Monmouthshire and parts of Flintshire. These areas benefit from a combination of strong GDV performance, concentrated rental density and development efficiency.

Moderate land values are observed in mixed coastal and rural–suburban environments such as Pembrokeshire and Ceredigion. These LSOAs exhibit meaningful income signals but at lower density and with less GDV amplification than core urban markets.

Lower land values are consistently found in rural, upland and structurally weaker post-industrial LSOAs such as Powys and parts of Gwynedd and RCT. In these areas, thinner markets, dispersed settlement patterns and lower rental concentration constrain both development viability and income capitalisation.

Importantly, this gradient is observed across RVM, ICM and hybrid outputs. While the absolute magnitude differs between models, the relative positioning of LSOAs remains structurally consistent. This convergence reinforces confidence that the modelling framework is capturing genuine spatial-economic differences rather than artefacts of parameter selection.

Data Density Shapes Stability and Confidence

A second consistent pattern concerns the relationship between data density and model stability.

In urban and suburban LSOAs with robust transaction volumes and well-populated rating bases, model outputs are comparatively stable and internally coherent. Development-led and income-led signals tend to align more predictably, and sensitivity to parameter variation is moderated by stronger empirical anchors.

By contrast, in rural and structurally weaker markets, thinner transaction evidence and concentrated income streams produce wider dispersion and greater sensitivity to core assumptions such as GDV and yields. Volatility observed in these areas reflects underlying market thinness rather than modelling inconsistency.

The fundamental conclusion is that model stability is directly correlated with the strength and granularity of the underlying data environment. Where market signals are dense and consistent, outputs converge more reliably. Where evidence is sparse, outputs remain bounded but exhibit greater sensitivity.

This observation underscores a crucial point. Improvements in granular data availability - particularly cadastral precision, rental evidence and construction cost transparency - would materially enhance intra-LSOA valuation confidence.

Structural Drivers of Variation

Across all study areas, variation in land value appears to be driven by a combination of:

- Market demand intensity,

- Income concentration and rental density,
- Development feasibility and build efficiency,
- Accessibility and employment proximity,
- Historic land-use and abnormal risk exposure.

These drivers operate consistently across both RVM and ICM logic. The hybrid framework confirms that neither development viability nor income potential alone fully explains land value patterns; rather, land value emerges from their interaction.

4.3 Advantages and Limitations of the Approach

Advantages

The principal strengths of the framework are as follows:

Full transparency and reproducibility
All modelling steps, data transformations and parameter assumptions are explicitly documented and traceable. This ensures that results can be independently replicated, scrutinised and updated over time. Transparency reduces reliance on opaque calibration and strengthens confidence in outputs for public-sector use.

Exclusive reliance on openly licensed datasets
The framework operates entirely within the Open Government Licence (OGL) ecosystem. This removes legal or licensing barriers to implementation, supports accessibility, and ensures that the methodology can be maintained and refined without dependence on proprietary data providers.

Explicit and auditable modelling assumptions
All high-impact parameters — including GDV benchmarks, yields, build costs, OPEX ratios and profit margins — are clearly defined. This makes sensitivity testing straightforward and ensures that stakeholders can understand how outputs are derived. The model avoids hidden calibration or black-box statistical methods.

Capacity to operate consistently across varied Welsh geographies
The modelling architecture functions across dense urban cores, suburban commuter markets, coastal settlements, rural uplands and post-industrial valleys. This geographic flexibility demonstrates structural robustness and ensures comparability across diverse settlement types.

Complementary dual-method logic
The combination of the Residual Value Model (development viability) and the Income Capitalisation Model (income-generating capacity) captures distinct economic drivers of land value. Their integration mitigates over-reliance on a single valuation mechanism and enhances interpretability.

Scalable data architecture
The parcel workflow, spatial harmonisation and aggregation processes are technically scalable to Wales-wide implementation. The computational framework

can process large datasets and generate consistent outputs across all LSOAs with minimal structural modification.

Collectively, these strengths produce a defensible evidence base suitable for strategic planning, comparative spatial analysis and potential fiscal exploration. The framework avoids opaque proprietary calibration and aligns with public-sector standards of accountability and auditability.

Limitations

The principal limitations of the approach arise from structural data constraints rather than weaknesses in modelling logic.

Absence of parcel-level cadastral precision
INSPIRE polygons provide indicative land extents but do not constitute legally defined development parcels. This limits the ability to incorporate subdivision patterns, access constraints, easements and other site-specific characteristics that materially affect real-world land value.

Generalised cost, yield and OPEX assumptions
Construction costs, operating expense ratios and capitalisation rates are applied using bounded, standardised assumptions. While transparent, these do not fully capture localised specification differences, contractor markets or investor sentiment at micro-spatial scale.

Limited rental and tenure granularity
Rateable Value provides a proxy for rental income, but detailed lease structures, covenant strength, incentive packages and tenure duration are not captured. This constrains precision in income attribution within the ICM.

Inability to explicitly model planning policy, land assembly or developer behaviour
The framework does not incorporate affordable housing requirements, biodiversity net gain obligations, phased delivery strategies, strategic land assembly costs or developer risk appetite. These factors significantly influence land transactions in practice but fall outside the open-data environment.

Sensitivity to GDV and yield assumptions in thin markets
In rural and post-industrial LSOAs with limited transaction evidence or concentrated income streams, model outputs are more sensitive to parameter adjustments. This reflects underlying data thinness rather than instability in the modelling structure.

Crucially, the analysis indicates that increasing structural complexity within an incomplete data environment does not necessarily improve robustness. Introducing highly granular or probabilistic modelling on partial or inconsistent datasets risks creating false precision. In such contexts, apparent sophistication may obscure rather than clarify uncertainty.

The findings therefore support a deliberate methodological choice. A simpler, transparent and bounded framework is more defensible than a structurally complex model dependent on unavailable granular inputs. Robustness is achieved through clarity of assumptions and sensitivity testing, not through opacity.

4.4 Drivers of Land Value Variation

The variation in land values observed across Wales is not random, nor solely a function of modelling assumptions. Rather, it reflects the interaction of structural economic forces, spatial constraints, market depth and institutional context. The modelling results - across RVM, ICM and hybrid outputs - reveal that land value emerges from the convergence of multiple reinforcing mechanisms rather than a single dominant variable.

Market Demand Intensity as the Primary Amplifier
At its core, land value is an expression of demand intensity relative to supply constraint. In high-demand urban and commuter environments, such as Cardiff and key suburban markets, land values are amplified by strong housing demand, employment accessibility and liquidity in both sales and rental markets.

In these areas, small changes in GDV or rental benchmarks translate into disproportionately large changes in land intensity, because the residual component of value (after costs or capitalisation adjustments) expands rapidly once threshold viability is exceeded.

By contrast, in structurally weaker or peripheral markets, demand intensity is lower and more volatile. In such environments, viability margins are thinner and income capitalisation is more sensitive to yield shifts, producing flatter land-value gradients.

Rental and Income Concentration as a Capitalisation Multiplier
The ICM results demonstrate that land value is highly responsive to income concentration — not simply average rent levels, but the density and durability of rental streams within a spatial unit.

Urban cores and employment nodes generate concentrated rental flows which, when capitalised, produce substantial land-value amplification. This explains why Cardiff exhibits structurally higher income-driven land values even after hybrid moderation.

Importantly, income concentration effects can operate independently of development feasibility. In certain LSOAs, income-based land values materially exceed development-led residual values, indicating that existing economic use may support land intensity beyond what new-build viability alone would suggest.

This divergence highlights that land can derive value from current economic function as much as from prospective redevelopment.

Economic Structure and Employment Base
Land value intensity reflects underlying economic structure. Areas with diversified employment bases, service-sector concentration and commuter connectivity exhibit stronger and more resilient land-value signals.

Post-industrial areas may retain pockets of income concentration but often display greater dispersion and volatility, reflecting economic transition rather than stable growth dynamics.

In rural areas, land value is more closely tied to amenity, accessibility and limited commercial clustering rather than dense employment nodes.

Historic Land-Use Patterns and Path Dependency
Historic land-use patterns shape present-day land intensity through path dependency. Post-industrial landscapes, fragmented ownership structures, legacy contamination and irregular parcel geometry continue to influence viability and capitalisation outcomes.

Conversely, historically planned suburban expansions with coherent parcel geometry and infrastructure provision exhibit more predictable development-led outcomes.

Land value therefore reflects accumulated structural conditions as much as current market signals.

Accessibility and Service Provision
Accessibility - to employment centres, transport networks and services - consistently correlates with higher land intensity. Commuter markets benefit from dual effects: strong residential demand and stable income signals.

Service provision, infrastructure quality and connectivity operate as silent multipliers within both RVM and ICM logic, shaping GDV benchmarks and rental sustainability.

In summary, land value variation across Wales emerges from the interplay between market demand, income density, development feasibility and spatial-economic structure. The modelling framework has revealed these dynamics with clarity, reinforcing the analytical validity of the dual-method and hybrid approach.

4.5 Robustness of the Approach and Applicability Across Land Types

Confidence in the Robustness of the Values and Methodology

Overall confidence in the robustness of the values is moderate to strong at LSOA scale, and more qualified at granular parcel level.

The modelling framework performs consistently in data-rich environments and produces coherent land-value gradients aligned with Welsh economic geography. The dual-model structure (RVM and ICM), together with the hybrid integration, reduces reliance on a single valuation mechanism and mitigates model-specific distortions. The explicit documentation of assumptions and the use of sensitivity testing further strengthen the transparency and defensibility of the approach.

However, robustness is not uniform across all contexts. The precision of outputs is constrained by the quality, completeness and resolution of the underlying datasets. In areas with thin transaction evidence, limited rental data or irregular parcel structures, results exhibit greater dispersion and sensitivity to key parameters such as GDV and capitalisation rates.

It is therefore important to distinguish between comparative robustness and absolute precision. The framework is robust for comparative analysis across LSOAs and for identifying structural gradients in land intensity. It is less robust as a tool for precise parcel-level valuation in the absence of more granular cadastral, rental and planning data. This limitation reflects structural constraints in the data ecosystem rather than weaknesses in the modelling logic itself.

Land That Is Easier or Harder to Value Using This Approach

The framework performs most reliably in environments where market signals are dense, consistent and well-supported by open data.

Easier to value:

- Urban and suburban areas with high transaction volumes;
- LSOAs with strong and well-distributed Rateable Value data;
- Locations with relatively regular parcel geometry and standardised built form;
- Markets with stable rental and sales evidence.

In these contexts, both RVM and ICM operate predictably, and hybrid outputs exhibit relatively stable central tendencies.

Harder to value:

- Rural and upland areas with low transaction frequency;
- Coastal settlements with seasonal or atypical occupancy patterns;
- Post-industrial landscapes with fragmented ownership and legacy land-use constraints;
- Mixed-use or atypical parcels where open-data attributes are limited;
- Areas where income is highly concentrated in a small number of properties.

In such environments, thin market evidence and structural heterogeneity increase model sensitivity and widen intra-LSOA dispersion. Development viability may be highly parcel-specific, and income capitalisation may amplify small variations in rental assumptions.

Importantly, these challenges do not invalidate the approach; rather, they highlight the structural limits of open-data modelling at fine spatial resolution. Enhanced parcel-level precision would require access to more granular administrative or proprietary datasets.

In summary, the methodology is robust and defensible as a transparent, LSOA-scale land-value framework. Its limitations are primarily data-driven and are clearly identifiable. The approach performs best where data density is strong and should be interpreted cautiously where granular evidence is thin.

4.6 Scalability and National Implementation

From a technical and architectural perspective, the approach is scalable across Wales. The data pipeline, spatial harmonisation process and modelling logic are capable of being extended to all LSOAs using the same open-data inputs and parameter framework. The workflow has demonstrated that national datasets can be systematically filtered, integrated and processed to produce consistent outputs at small-area scale.

However, scalability in a technical sense does not automatically imply scalability in terms of precision or evidential strength. The modelling framework is robust for comparative, LSOA-scale analysis across Wales. It can identify structural gradients in land value intensity and provide a consistent analytical baseline. In this sense, the approach is operationally scalable.

What cannot be assumed is that parcel-level precision will improve simply through national rollout. In fact, scaling the framework without improved data inputs would extend the same structural limitations observed in thinner markets across a larger geography. Therefore, the approach is scalable in methodology and computation, but bounded in accuracy by the quality and granularity of available data.

What Wider Requirements Would Be Needed to Value the Whole of Wales?

To move from a scalable analytical framework to a higher-precision national land valuation system, several wider requirements would be necessary.

1. Authoritative Parcel-Level Cadastral Data
An openly licensed, legally defined cadastral dataset would significantly improve spatial precision. Current reliance on indicative INSPIRE polygons limits the ability to incorporate subdivision patterns, access constraints and parcel-specific development potential.

2. Enhanced Rental and Income Data
More granular and current rental evidence - including lease length, incentives, covenant strength and sector-specific yield benchmarks - would materially improve income capitalisation accuracy, particularly in employment-led markets.

3. Granular Construction Cost Data
Access to more detailed, regionally differentiated construction cost benchmarks would reduce sensitivity in development-led modelling and improve the realism of abnormal-cost assumptions.

4. Planning and Policy Constraint Layers
Integration of site-specific planning designations, affordable housing requirements, environmental constraints and infrastructure obligations would strengthen development viability modelling and reduce abstraction.

5. Market Behaviour and Transaction Context
Land assembly costs, developer strategy, phasing risk and behavioural considerations influence real-world land transactions but are not captured within open datasets.

6. Data Governance and Update Mechanisms
A Wales-wide system would require structured update cycles aligned with transaction data releases, rating revaluations and economic change.

Without improvements in input data quality and granularity, scaling the current framework across Wales would produce a consistent but structurally bounded national land-value model.

It would be capable of:

- Comparative spatial benchmarking,
- Identifying structural gradients,
- Supporting high-level fiscal or planning analysis.

It would not, however, provide parcel-accurate valuation precision or replace professional appraisal in site-specific contexts.

Accordingly, national scalability is technically feasible but analytically constrained. If greater valuation precision is required - particularly for fiscal implementation or regulatory application - access to higher-quality, more granular datasets beyond the current open-data environment would be essential.

Section 5 – Further Considerations

5.1 Areas for Further Research and Methodological Refinement

The framework developed in this study establishes a transparent and reproducible foundation for LSOA-scale land valuation using open data. However, several avenues for further research and refinement could materially enhance analytical precision and policy relevance.

Data Enhancement as the Primary Refinement Lever
The most significant limitation identified throughout the modelling process is not methodological but structural - namely, the granularity and completeness of available input data. Future work should prioritise improvements in data quality before increasing modelling complexity.

In particular, access to an openly licensed cadastral dataset would allow:

- Accurate identification of legally defined development parcels;
- Improved alignment between parcel geometry and buildable area;
- Better integration of land subdivision patterns and access constraints.

Similarly, more granular rental evidence - including lease lengths, incentive structures, vacancy rates and covenant strength - would significantly strengthen income capitalisation modelling. Current reliance on Rateable Value as a proxy is appropriate within open-data constraints but does not capture the full dynamics of commercial leasing markets.

Construction cost modelling would also benefit from access to regionally differentiated benchmarks reflecting specification, contractor markets and abnormal site conditions. Open-data cost proxies are sufficient for feasibility testing but introduce unavoidable abstraction in areas with atypical build requirements.

The central lesson is that analytical refinement is primarily a data challenge rather than a modelling challenge.

Structured Scenario and Sensitivity Frameworks
While this study has incorporated bounded parameter assumptions, further work could formalise structured scenario corridors. This could include:

- Low, central and high GDV trajectories under different market conditions;
- Yield sensitivity under tightening or loosening investor sentiment;
- Cost inflation scenarios under different construction market pressures;
- Developer return bands reflecting varying risk appetite.

Such scenario modelling would enhance the framework's utility for long-term strategic planning and fiscal stress-testing, while preserving transparency.

However, increasing the number of scenarios should not come at the expense of interpretability. The value of the current framework lies in its clarity; additional layers should be introduced only where supported by stronger data inputs.

Refinement of the Hybrid Integration Method
The hybrid integration of RVM and ICM has demonstrated stabilising benefits. Future work could explore:

- Confidence-weighted hybridisation, where model weight varies according to data density;
- Development of “land value envelopes” showing upper and lower bounds rather than single-point outputs;
- Temporal smoothing across rolling data windows to reduce short-term volatility.

These refinements would further strengthen robustness without fundamentally altering the modelling logic.

5.2 Public Understanding and Communication

If the framework were to be used beyond technical analysis - particularly in strategic planning or fiscal discussions - public communication would be critical.

Land valuation, particularly when separated from building value, is not intuitively understood by most members of the public. Common misunderstandings may include:

- Confusion between land value and total property value;
- Perception that modelled values represent immediate sale prices;
- Concern that modelling assumptions introduce arbitrariness;
- Misinterpretation of LSOA-level averages as binding parcel-level determinations.

Clear explanation would need to emphasise that:

- Land value reflects location, accessibility and economic context rather than the building itself;
- Two complementary methods are used to triangulate value;
- Outputs are comparative and statistical rather than site-specific valuations;
- Sensitivity to assumptions is explicitly recognised and bounded.

Visual mapping, explanatory diagrams and plain-language summaries would be essential tools for engagement.

A further communication challenge relates to volatility. Because both development feasibility and income capitalisation are sensitive to market conditions, outputs may vary over time. Explaining this as a structural feature of land markets rather than model instability would be important.

Transparency remains the strongest asset of the framework in supporting public understanding.

5.3 Considerations for Use in a Land-Based Tax Context

The modelling framework developed in this study could, in principle, inform exploration of land-based fiscal reform. It demonstrates that land value can be estimated separately from building value using transparent and reproducible methods.

However, several important considerations would apply before such use.

Precision and Fairness

For taxation purposes, parcel-level precision would need to be significantly higher than that required for comparative spatial analysis. Current limitations in cadastral precision, rental granularity and cost benchmarking would need to be addressed to ensure fairness and consistency.

Governance and Updating

A land-value-based system would require:

- Clear parameter governance (e.g., how GDV, yields and cost assumptions are set);
- Regular update cycles aligned with market data releases;
- Appeals mechanisms for contested valuations;
- Transparent publication of assumptions and methodologies.

Without these structures, technical modelling alone would not be sufficient for implementation.

Distributional and Behavioural Effects

Land taxation interacts with housing markets, development incentives and investment behaviour. While the modelling framework can estimate land intensity, broader policy analysis would be required to assess:

- Distributional impacts across regions and property types;
- Behavioural responses in development timing and land supply;
- Transitional effects on existing landowners.

The framework is well suited to scenario modelling of these questions but is not, in its current form, a complete fiscal system design.

5.4 Strategic Perspective

The overall findings suggest that Wales now has the foundation for a transparent, scalable and analytically coherent understanding of land value at small-area scale.

The modelling demonstrates:

- Structural differentiation in land intensity across Welsh geographies;
- Complementary insight from development-led and income-led approaches;
- Stabilisation through hybrid integration;
- Clear sensitivity boundaries;
- Technical scalability to national coverage.

However, it also makes clear that greater parcel-level precision depends fundamentally on improved data inputs. Enhancing the data ecosystem - particularly cadastral, rental and cost datasets - would deliver greater marginal benefit than increasing structural modelling complexity.

In strategic terms, the framework represents a substantial step forward in evidence capability, while clearly identifying the data and governance conditions required for future evolution.

References

1. International Association of Assessing Officers (2017) *Property Assessment Valuation*. 3rd edn. Kansas City: IAAO.

2. RICS (2020) *RICS Valuation – Global Standards (Red Book Global)*. London: Royal Institution of Chartered Surveyors.

3. Welsh Government (2020) *Reforming Local Government Finance in Wales: Summary of Findings*. Cardiff: Welsh Government.

4. Bourassa, S.C. and Hoesli, M. (2022) 'Hedonic, Residual, and Matching Methods for Residential Land Valuation'. Swiss Finance Institute Research Paper No. 22-55. Available at: papers.ssrn.com/sol3/papers.cfm?abstract_id=4148824

5. Kolbe, J. et al. (2019) *Land Value Appraisal Using Statistical Methods*. FORLand Working Paper 2019/07. Available at: econstor.eu/bitstream/10419/213061/1/FORLand-2019-07.pdf

6. Sugito, N.T. et al. (2024) 'Urban Land Value Estimation Model Based on Income Capitalisation Method'. *E3S Web of Conferences*, 473, 06002. Available at: e3s-conferences.org/articles/e3sconf/pdf/2024/130/e3sconf_igeos2024_06002.pdf

7. Trojanek, M. (2010) 'The Application of Income Approach in Property Valuation in Poland'. *Economics and Sociology*, 3(2), pp. 35–47. Available at: researchgate.net/publication/274184012

8. Bridgend County Borough Council (2023) *Local Development Plan Evidence Base*. Available at: bridgend.gov.uk

9. Ceredigion County Council (2022) *Local Housing Market Assessment*. Available at: ceredigion.gov.uk

10. JLL (2023) *Cardiff Property Market Update*. Available at: jll.co.uk

11. Monmouthshire County Council (2022) *Local Development Plan and Housing Evidence*. Available at: monmouthshire.gov.uk

12. Natural Resources Wales (2023) *Flood Risk Assessment Wales (FRAW)*. Available at: naturalresources.wales

13. Office for National Statistics (2020) *Welsh Index of Multiple Deprivation*. Available at: ons.gov.uk

14. Office for National Statistics (2021) *Census 2021: Housing Characteristics*. Available at: ons.gov.uk

15. Office for National Statistics (2021) *Rural-Urban Classification for Small Areas*. Available at: ons.gov.uk

16. Pemberton, S. and Goodwin, M. (2021) ‘Cross-Border Economic Interaction in North Wales and Cheshire’. *Regional Studies*, 55(4), pp. 611–627.

17. Penn, A. and Jones, M. (2020) ‘Post-Industrial Transition in the South Wales Valleys’. *Town Planning Review*, 91(6), pp. 663–684.

18. StatsWales (2024) *Dwelling Stock and Housing Market Indicators*. Available at: statswales.gov.wales

19. Visit Wales (2023) *Visitor Economy Overview*. Available at: visitwales.com

20. Welsh Government (2023) *Regional Economic and Labour Market Profiles*. Available at: gov.wales

21. Crosby, N., McAllister, P. and Wyatt, P. (2018) *Residual Land Values: Measuring Performance and Investigating Viability*. University of Reading / Investment Property Forum.

22. Census 2021 (2023) *Housing Variables, England and Wales*. Office for National Statistics. Available at: ons.gov.uk

23. DataMap Wales (2025) *Spatial Datasets and Mapping Services*. Welsh Government. Available at: datamap.gov.wales

24. EPC Open Data (2025) *Energy Performance of Buildings Certificates – Open Data (England and Wales)*. Department for Levelling Up, Housing & Communities. Available at: epc.opendatacommunities.org

25. HM Land Registry (2025a) *Price Paid Data*. Available at: landregistry.data.gov.uk/app/ppd

26. HM Land Registry (2025b) *UK House Price Index*. Available at: landregistry.data.gov.uk/app/ukhpi

27. HM Land Registry (2025c) *INSPIRE Index Polygons*. Available at: use-land-property-data.service.gov.uk/datasets/inspire

28. Office for National Statistics (2024a) *House Price Statistics for Small Areas (HPSSA)*. Available at: ons.gov.uk/peoplepopulationandcommunity/housing/datasets/housepricestatisticsforsmallareas

29. Office for National Statistics (2024b) *Price Index of Private Rents (PIPR)*. Available at: ons.gov.uk/economy/inflationandpriceindices/datasets/priceindexofprivaterentsukhistoricalseries

30. Office for National Statistics (2022) *Lower Layer Super Output Area Boundaries 2021 – Wales*. Available at: datamap.gov.wales/layers/geonode:lsoa_2021_w_hwm

31. Ordnance Survey (2024a) *OS OpenMap Local*. Available at: ordnancesurvey.co.uk/products/os-open-map-local

32. Ordnance Survey (2024b) *OS Open Zoomstack*. Available at: ordnancesurvey.co.uk

33. StatsWales (2024a) *Private Sector Rents by Local Authority*. Available at: statswales.gov.wales/Catalogue/Housing/Private-Sector-Rents

34. StatsWales (2024b) *Land Transaction Tax Statistics*. Available at: stats.gov.wales/topic/35/land-transaction-tax

35. Doogal (2025) *Postcode to LSOA Lookup*. Available at: doogal.co.uk

36. Natural Resources Wales (2024a) *Flood Risk Assessment Wales (FRAW)*. Available at: datamap.gov.wales

37. Natural Resources Wales (2024b) *Flood Map for Planning (TAN15)*. Available at: datamap.gov.wales

38. Natural Resources Wales (2024c) *LiDAR Datasets*. Available at: datamap.gov.wales

39. NRW/JNCC (2024) *Protected Sites Datasets*. Available at: datamap.gov.wales

40. BGS (2024) *OpenGeoscience: Geology Datasets*. British Geological Survey. Available at: bgs.ac.uk

41. Coal Authority (2024) *Mine Entries and Mining Risk Data*. Available at: data.gov.uk

42. Ofcom (2024) *Connected Nations Data*. Available at: ofcom.org.uk

43. ONS (2024c) *ONS Postcode Directory (ONSPD)*. Available at: geoportal.statistics.gov.uk

44. TfW (2024) *Transport for Wales GTFS/TXC Datasets*. Available at: traws.cymru

45. UKCEH (2024) *Land Cover Map – Open Statistics*. Available at: data.gov.uk

46. Welsh Government (2019) *Welsh Index of Multiple Deprivation (WIMD)*. Available at: statswales.gov.wales

47. Welsh Government (2024a) *Predictive Agricultural Land Classification*. Available at: datamap.gov.wales