Unravelling the dynamics behind the urban morphology of port-cities using a LUTI model based on cellular automata

Aditya Tafta Nugraha[a], Ben J. Waterson[a], Simon P. Blainey[a], and Frederick, J. Nash[b].

[a]Transportation Research Group, School of Engineering, University of Southampton

[b]Agents, Interactions, and Complexity Research Group, School of Electronics and Computer Science, University of Southampton, Southampton, United Kingdom

Abstract

The urban morphology is characterised by self-organisation where interactions of multiple agents produce emerging patterns on the urban form. Port-urban relationship added to the complexity of port cities' urban form. Most urban cellular automata (CA) models simulate land-use evolution through transition rules representing multi-factored local interactions. However, calibration of CA-based urban land use and transport interaction (LUTI) models often utilise manual methods due to complexity of the process. This limits insights on urban interactions to a few explored settlements and prevents applications for planning and assessment of transport policies in other contexts. This paper, therefore, addresses three main points. The paper (i) demonstrates an improved method for the calibration of CA-based LUTI models, (ii) contributes to a better understanding of the urban dynamics in port city systems by quantifying generalizable interactions from a wide range of port-urban settlements, and (iii) illustrates how the use of these interactions in a simulation model can allow long-term impact predictions of planning interventions.

These were done by formulating a model in a similar structure as a neural network model to enable automatic calibration using an application of the gradient-descent algorithm. The model was then used to quantify the dynamics between land-use, geographic, and transport factors in 46 port-based and 10 non-port settlements across Great Britain, thus enabling cross-sectional analysis. Cluster analysis of the calibrated interactions in the study areas was conducted to examine the variations of these interactions. This produced two main groups. In the first group, consisting larger settlements, connections between ports and other urban activities were weaker than in the second group which consisted of smaller port-settlements. Overall, the findings of the research are consistent with existing evidence in the port-cities literature but go further in quantifying the interaction between urban agents within port-urban systems of various sizes and types. These quantified interactions will enable planners to better predict the longer-term consequences of their interventions.

1. Introduction

The urban morphology is a complex system characterised by self-organisation where interactions of multiple agents produce emerging patterns on the urban form (Batty, 2007). In port cities, this is made more complex by the port-urban interactions, which could be mutually beneficial in some parts and antagonistic in other parts (Hall and Jacobs, 2012). The effect of the expansion of both the port and the urban elements could be a “double-edged” sword to the port city system (Lu et al., 2021). The successes of both port and urban developments in port cities are linked through the spatial organisation (land-use) and spatial interaction (transport) between the different functions in the cities (Merk, 2013). Indeed, tensions between port and urban functions often materialise in the forms of transport (congestion) and land use (space limitations and competitions) problems (Ducruet and Lee, 2006). In order to plan successful long-term interventions, transport and urban planners working in port cities context must have a better understanding of the dynamics between different factors within the port-urban system. Generalised insights into these individual interactions are therefore critical in successful transport and land-use planning of urban settlements as they would allow planners to better predict the long-term outcomes of their plans.

Cellular automata (CA) models are often used for modelling urban land-use due to their capability in replicating self-organisation behaviour (Batty, 1997; Batty, 2007). Most urban CA models simulate land-use evolution through transition rules with consideration of multiple factors representing the local interactions (Santé et al., 2010). Calibration of urban CA models could be seen as a process to measure the effect of each factor. However, measuring these individual effects from empirical data is complicated due to their autocorrelation and to the effects of specific events occurring in their history (known as path-dependence) on their urban forms (Van Vliet et al, 2013b). Due to these complications, manual methods are prevalent in the calibration of urban land-use and transport interaction (LUTI) models based on CA (Aljoufie et al., 2013). This limits the insights into urban interactions to a few explored settlements. Consequently, the interactions measured are often specific to urban dynamics of the few explored settlements as the presence of path-dependence prevents generalisation of results from one study area to other areas which may not share the same history.

This limitation presents a considerable hindrance in applying a knowledge gained in certain port cities to assess the impacts of certain transport and urban policies in other port cities. This is even more problematic when the variability between port cities is considered. Ducruet and Lee (2006) proposed a framework with 9 categories of port cities depending on the balance of their centrality and their intermediacy. These are concepts used to describe characteristics of transportation hubs (Fleming and Hayuth, 1994) which were contextualised as the intensity of their urban (centrality) and port (intermediacy) activities compared to the rest of the regions. Beyond this framework, there could be variations in the spatial organisations between urban and port functions. Hoyle (2000) proposed that western port cities often began with both functions being in proximity to each other but the link between them become weaker as the two functions grow. However, other port cities may follow the East Asian spatial growth model (Lee et al., 2008) where strong links between port and city are maintained as the functions grow, or the Middle Eastern model (Akhavan, 2017), where the development of one of the functions lags behind the other function. Furthermore, the effect of ports on the urban areas would also vary depending on the nature of the traffic carried which would have different impacts on the urban economy through the industry interlinkage (Yochum and Agarwal, 1988; Kwak et al., 2005). Due to these variabilities of port city systems, it is important for port cities studies to expand beyond specific case studies and put more effort into testing the applicability of port city evolutionary models in the more generalised context of port cities (Woo et al., 2011; Ng et al., 2014). In particular, in the context of port cities’ urban form there is a need to examine the contributions of geographic conditions, transport access and proximity to other land uses to the potentials of urban areas in attracting certain types of land use activities within different port city settings.

The aims of this paper are therefore; to introduce an effective automated calibration method of urban CA-LUTI models, to quantify the interactions between urban agents over a wide range of port settlements, and to illustrate how these interactions can be used in a simulation model to predict long-term impacts of policy interventions.

2. Materials and Method: measurement of urban dynamics

2.1. Study Area Selection and Data Collection

In order to measure urban dynamics of port cities, 56 settlements of various sizes in Great Britain were chosen as study areas as shown in Figure 1. These consist of 46 port settlements and 10 non-port settlements as comparison. The population and area sizes of these study areas range from a settlement with a population of 650 over an area of 1 km^2 (Kyle of Lochalsh) to one with over 1.5 million population with a 563 km^2 area (Tyneside metropolitan area). Note, however, that these selections excluded larger settlements such as Greater London due to computational resources limitations.

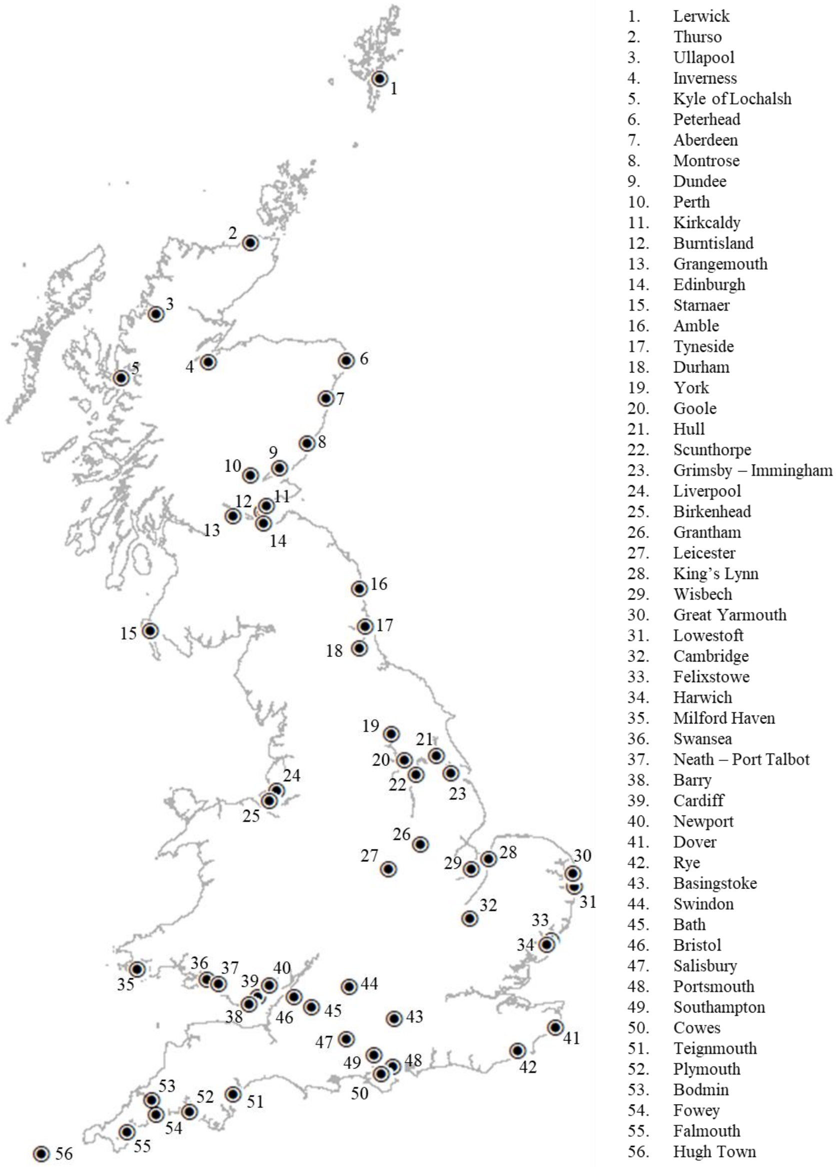


Figure 1 Map of study areas across Great Britain

As mentioned, these study areas also included settlements without ports such as Cambridge and settlements where port activity has ceased such as York. Types of ports included highly containerised port such as Felixstowe, non-container freight ports such as Grimsby and Immingham, mainly passenger ports such as Harwich, fishing ports such as Peterhead, and those with mixed traffic such as Dover (ferry passenger and non-containerised freight) and Southampton (cruise passenger and containerised freight). These study areas also vary in terms of their form of waterfront from coastal settlements such as Dover, estuarial such as Southampton, inland such as Wisbech, and smaller island settlements such as Hugh Town. The general patterns of urban dynamics are uncovered by comparing results from calibrations of these study areas.

For these study areas, land use, geographic, transport, and planning restrictions data were collected. Land use data were obtained from Ordnance Survey's AddressBase® Plus dataset[1] for the year 2018 which contained classed and geo-located addressable properties. Data on geographic features and locations of transport infrastructures were obtained from Ordnance Survey OpenMap dataset[2]. These include geographic factors such as terrain types and transport factors such as location of roads and railway infrastructures. These datasets are complemented with public transport service data from the Traveline National Dataset provided under the Open Government Licence[3] and terrain slope data obtained from LIDAR Composite Digital Terrain Model (DTM)[4]. High level planning restrictions data were obtained from the Joint Nature Conservation Committee's (JNCC) website[5].

2.2. Agent-based Urban Cellular Automata Model

This paper made use of an agent-based CA model to represent between urban agents. CA models are spatially explicit land use models utilising a lattice of regular and uniformly sized cells to represent geographic location (Van Vliet et al., 2012). They simulated the states of cells as the result of their own states and the states of their neighbouring cells through a set of transition rules (Batty, 1997). Work by Nugraha and colleagues (2020) showed that urban CA models with hexagonal cells have consistency advantages over those with square cells. Therefore, the model in this paper used hexagonal cells with each cell having an area size of 22,500 m^2, which is below the maximum recommended size for urban CA models according to Samat (2006). While urban CA models can represent self-organisation behaviour (Batty, 1997) by simulating individual interactions, calibrating these individual effects from empirical data is complicated due to their autocorrelation (Van Vliet et al, 2013b). The neural network research field has developed automated calibration approaches for identifying patterns in large data, so it seems beneficial to marry a neural network approach and the calibration of urban CA models.

The approach started with a model loosely based on Metronamica (RIKS, 2010), which uses cell-potential-based transition rules to simulate cell states (Santé et al., 2010). Rather than employing a fully cell-based approach where cells have categorical states, the model used a multi-agent approach (as in Van Vliet et al., 2012) where each cell contains information on the levels of different land uses agents co-existing within it. The calibration process, which is covered in subsection 2.3 of this paper, therefore measured the interactions between these agents, as well

---

[1] Available at a cost from: bit.ly/3kb6nzs
[2] Available freely from: bit.ly/3D31wt2
[3] Available freely from: bit.ly/3APHgJk
[4] Available freely from: bit.ly/3mi4Zh9
[5] Available freely from: bit.ly/3gzES1V

as the way these agents react to geographic and transport conditions, which resulted in their locational decision within the urban system. The model considered four main components to predict the distribution of urban agents into cells. As shown in Equation 1, the potential of a cell j in attracting land use activity *v* ($P_{v,j}$) is calculated from land use ($N_{v,j}$), geographic ($\boldsymbol{G}_{v,j}$), transport ($\boldsymbol{T}_{v,j}$) and institutional (planning policy) ($\boldsymbol{I}_{v,j}$) factors.

Cell potential $$\boldsymbol{P}_{v,j} = \boldsymbol{I}_{v,j}\ .\ \boldsymbol{G}_{v,j}\ .\ \boldsymbol{T}_{v,j}\ .\ \boldsymbol{N}_{v,j}$$ *(Equation - 1)*

Land use component

Within the model, proximity to activities was represented by the land use component. As the main aim of model calibration in this paper was to investigate interactions between urban agents, more emphasis was given to the land use component as it controlled interactions between urban agents.

In this agent-based model, the agents were defined as individual unit of activities found in port-city systems. Land use data from the AddressBase® Plus dataset were projected into cells. These points were then classified into five main representations; 'residential' (such as houses and flats), 'port' (such as harbour and terminal facilities), 'manufacturing' (such as industrial complexes), 'consumer-services' (such as restaurants and shops), and 'business-services' (such as offices). The number of agents within each cell were then aggregated to measure the levels of different urban activities for that cell (Figure 2).

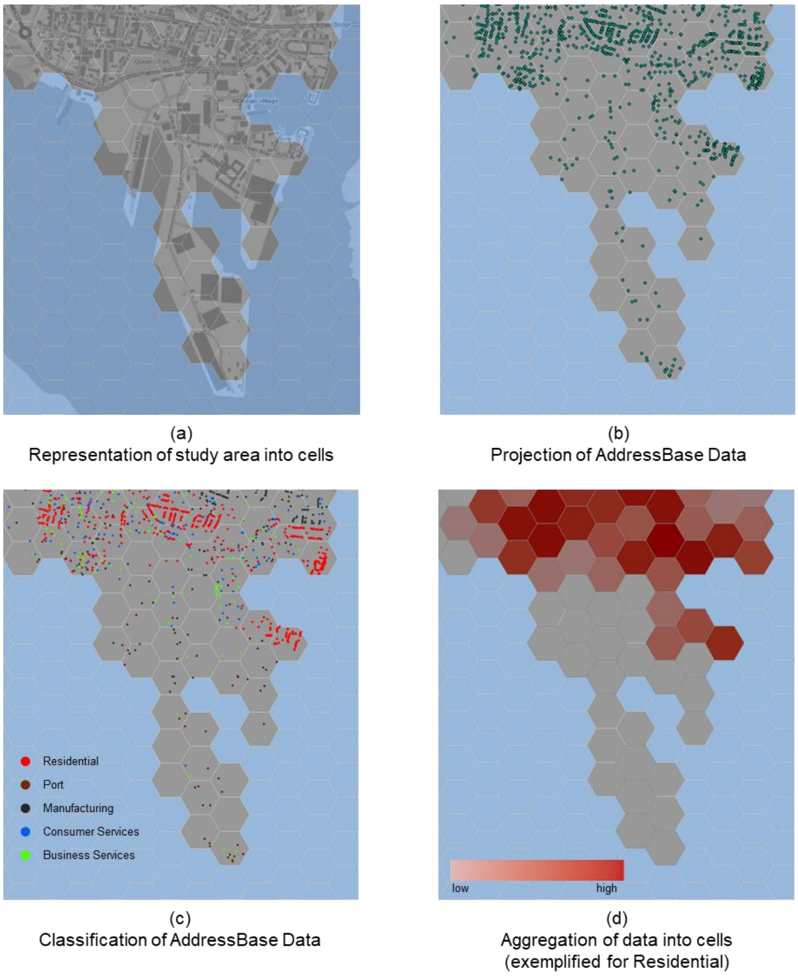


Figure 2 Processing of Land Use Data

In this way, the model explicitly represented multi-purpose developments and developments at varying density levels. More importantly, the near-continuous cell states of this approach allowed an easier adoption of the gradient descent algorithm, which dealt with continuous solution space. The land use component allowed the model to measure neighbourhood effect which represents the way agents react to the effect of proximity to other land use. This was governed by Equations 2, 3a, and 3b.

Neighbourhood Effect

$$\boldsymbol{N_{v,j}} = \boldsymbol{S}\left(\sum_{\boldsymbol{u}}\sum_{\boldsymbol{i}\,|\,\boldsymbol{i}\in N(\boldsymbol{j})}\left(\left(\boldsymbol{Z_{u,i}}\right)^{\lambda_{A,u,v}}.\ \boldsymbol{A_{u,v,i,j}} - \left(\boldsymbol{Z_{u,i}}\right)^{\lambda_{R\,u,v}}.\ \boldsymbol{R_{u,v,i,j}}\right)\right) \quad \textit{(Equation - 2)}$$

Attraction

$$\boldsymbol{A_{u,v,i,j}} = \boldsymbol{\alpha_{A,\ u,v}} - \frac{\boldsymbol{\alpha_{A,\ u,v}}}{\left(\boldsymbol{1} + \boldsymbol{Exp}\left(-\boldsymbol{\beta_{A,\ u,v}}.\left(\boldsymbol{D_{i,j}} - \boldsymbol{\gamma_{A,\ u,v}}\right)\right)\right)} \quad \textit{(Equation – 3a)}$$

Repulsion

$$\boldsymbol{R_{u,v,i,j}} = \boldsymbol{\alpha_{R,\ u,v}} - \frac{\boldsymbol{\alpha_{R,\ u,v}}}{\left(\boldsymbol{1} + \boldsymbol{Exp}\left(-\boldsymbol{\beta_{R,\ u,v}}.\left(\boldsymbol{D_{i,j}} - \boldsymbol{\gamma_{R,\ u,v}}\right)\right)\right)} \quad \textit{(Equation – 3b)}$$

Neighbourhood effect of cell *j* in attracting land use activity *v* ($\boldsymbol{N_{v,j}}$) was calculated from the attraction and repulsion effects from other activities. A land use activity of type *u* located at cell *i* would have some potential in both attracting and repulsing a land use type *v* at cell *j* ($\boldsymbol{A_{u,v,i,j}}$ and $R_{\boldsymbol{u,v,i,j}}$). The strength of these attraction and repulsion were dependent on the suitability between land uses *u* and *v* which were determined by the logistic decay parameters $\boldsymbol{\alpha_{u,v}}$, $\boldsymbol{\beta_{u,v}}$, and $\boldsymbol{\gamma_{u,v}}$ as well as the distance between cells *i* and *j* ($\boldsymbol{D_{i,j}}$). The $\boldsymbol{\alpha}$ parameters controlled the general potency of the attraction or repulsion, the $\boldsymbol{\beta}$ parameters controlled how sensitive that potency changes with distance and the $\boldsymbol{\gamma}$ parameters controlled the potency at immediate distance. Attraction and repulsion were also reliant on the intensity of land use *u* in cell *i* ($Z_{u,i}$). This scale effect was governed by the parameters $\boldsymbol{\lambda_{A,u,v}}$ for attraction and $\boldsymbol{\lambda_{R,u,v}}$ for repulsion.

Geographic and transport components

Within the model, physical suitability of cell *j* for land use development *v* was represented by geographic component ($\boldsymbol{G_{v,j}}$) and accessibility potential of cell *j* for land use development *v* was represented by transport component ($T_{\boldsymbol{v,j}}$). Transport component was divided into static transport ($ST_{\boldsymbol{v,j}}$) to represent proximity to transport facilities and dynamic transport ($DT_{\boldsymbol{v,j}}$) to represent ease of accessing activities in other cells.

Geographic and static transport components were calculated using weighted geometric mean of geographic and static transport factor scores. Dynamic transport component was calculated using an economic potential transport accessibility measurement as the sum of attraction of all other nodes over the friction between a pair of cells (Gutiérrez, 2001). The frictions between cells were represented by travel time between them given the transport infrastructure. Origin-destination journey time matrices for dynamic transport calculation were obtained using the OpenTripPlanner software as in Young (2016). Changes in dynamic transport due to change in activity distribution were simulated using a high-level 4-stage transport model. This transport model simulated flows and delays based on speed-flow curves along the links. Link classification and speed flow curves were based on a regional model for England as shown in Atkins (2009).

Institutional component

The institutional component in the model could be used to represent the effect of planning policies on the attractiveness of cells. When a cell *j* falls under the area where development for land use type *v* was not permitted, $\boldsymbol{I_{v,j}}$ equalled to 0. In that way, regardless of the potentials from neighbourhood, geographic, or transport components, the total cell potential would be equal to 0 due to the multiplicative structure of the equation (Equation 1). Where development was permitted or uncontrolled, $\boldsymbol{I}_{v,j}$ equalled to 1. A value between 0 and 1 was used to represent policies to discourage developments while a value larger than 1 represented policies to promote certain types of development within the cell.

2.3. Model Calibration using Gradient Descent Algorithm

To enable the use of gradient descent algorithm, model calibration was represented as an optimisation problem with an objective function that minimises the disagreement between actual land use distribution and the predicted cell potentials. Cell potentials were, in turn, calculated from a network of differentiable functions. The gradient descent algorithm used partial differentiations of the objective function with regards to calibrated parameters to update solutions towards an optima. The model has been built and calibrated using the PyTorch machine learning library which is available at an online repository[6].

Model representation as optimisation problem

Table 1 in the next page presents the mathematical model containing the four components and their intermediate calculations. The equations in Table 1 are tabulated to show the hierarchy of the calculation with equations at the higher hierarchy taking inputs from equations in the lower hierarchy. Variables colour-coded in red are those requiring calibration. Their ranges are defined at the end of Table 1. Meanwhile, variables colour-coded in blue represent input data. Table 2 presents a glossary for model data inputs.

As detailed in Section 2.2, the neighbourhood effect was the most complex interaction regulated by the most variables. An individual neighbourhood interaction between two land use activities was governed by two main constituents, attraction and repulsion. Each was in turn governed by the parameters $\alpha_{u,v}$ representing the general potency of the effect, $\beta_{u,v}$ representing how sensitive that potency changes with distance, $\gamma_{u,v}$ representing the potency at immediate distance, and $\lambda_{u,v}$ representing how quickly that potency grows as the neighbour's presence intensify.

The effect of a geographic feature x on a land use activity v was quantified by $\mu_{x,v}$, regulating the range of input values $\eta_{x,j}$ where the effect was most sensitive, and $\omega_{x,v}$, regulating how sensitive the land use v is to the change in geographic feature x. The same mathematical structure applied to static transport factors y with the parameters $\upsilon_{y,v}$ and $\varphi_{y,v}$ regulating the effect of input values $\rho_{y,j}$. The effect of dynamic accessibility to a land use activity u, measured by journey time on a specific mode of transport m, on cell's attractiveness to land use activity v is regulated by $\kappa_{u,v,m}$, which scaled the attractiveness linearly to the intensity of land use activity u at the destination, and $\partial_{u,v,m}$, which controlled attractiveness inverse-log-linearly to access time.

[6] https://anonymous.4open.science/r/CA_Based_LUTI_Model-1336/readme.md
The repository has been anonymised in line with double-blind policy. Some data has been modified due to data confidentiality agreement.

Table 1. Formulation of model into optimization problem		
Objective function	$min \sum_j \left(\frac{(Z_{v,j} - Predicted_{v,j})^2}{Predicted_{v,j}} \right)$	
where:		
Predicted values	$Predicted_{v,j} = \frac{P_{v,j}}{\sum_{j \mid j \epsilon SA} P_{v,j}} . GTotal_{v,SA}$	
Cell potential	$P_{v,j} = I_{v,j} \,.\, G_{v,j} \,.\, T_{v,j} \,.\, N_{v,j}$	
Geographic effect	$G_{v,j} = \prod_x B\left(F(C(\eta_{x,j}), \mu_{x,v})\right)^{\omega_{x,v}}$	
Bounding function	$B(\theta) = \frac{log(1+e^{10\theta})}{log(1+e^{10})}$	
Shifting function	$F(\theta, \phi) = \left(1 - \left(1/_{\phi}\right)(1-\theta)\right)$	
Projection function	$C(\theta) = c + (1-2c)\left(0.98\left(\frac{\theta - min_\theta}{max_\theta - min_\theta}\right) + 0.01\right)$	
Transport effect	$T_{v,j} = ST_{v,j} \,.\, DT_{v,j}$	
Static transport effect*	$ST_{v,j} = \prod_y B\left(F(C(\rho_{y,j}), \nu_{y,v})\right)^{\varphi_{y,v}}$	**this equation also unfolds to bounding, shifting, & projection functions as Geographic Effect*
Zone to cell mapping	$DT_{v,j} = DT_{v,jj} \quad for\, j \in jj$	
Dynamic transport effect	$DT_{v,jj} = \sum_m \sum_u \sum_{ii} \left(\frac{\kappa_{u,v,m} . STotal_{u,ii}}{(1+ODTM_{jj,ii,m})^{\partial_{u,v,m}}} \right)$	
Neighbourhood effect	$N_{v,j} = S\left(\sum_u \sum_{i \mid i \in N(j)} \left((Z_{u,i})^{\lambda_{A,u,v}} . \ A_{u,v,i,j} - (Z_{u,i})^{\lambda_{R\,u,v}} . R_{u,v,i,j} \right) \right)$	
Rectifier	$S(\theta) = log(1 + e^{\theta}) + 0.00001$	
Masker	$Z_{u,i} = 0 \quad where \quad (u = v) \cap (i = j)$	
Attraction effect	$A_{u,v,i,j} = \alpha_{A,\ u,v} - \frac{\alpha_{A,\ u,v}}{\left(1 + Exp\left(-\beta_{A,\ u,v} . (D_{i,j} - \gamma_{A,\ u,v})\right)\right)}$	
Repulsion effect	$R_{u,v,i,j} = \alpha_{R,\ u,v} - \frac{\alpha_{R,\ u,v}}{\left(1 + Exp\left(-\beta_{R,\ u,v} . (D_{i,j} - \gamma_{R,\ u,v})\right)\right)}$	

subject to the calibrated parameters:	
Geographic parameters	$0 < \mu_{x,v} \leq 1 \, , \, \omega_{x,v} \geq 0$
Static transport parameters	$0 < \upsilon_{y,v} \leq 1 \, , \, \varphi_{y,v} \geq 0$
Dynamic transport parameters	$\kappa_{u,v,m} \geq 0 \, , \, \partial_{u,v,m} \geq 0$
Neighbourhood effect parameters	$\alpha_{A,\ u,v} \geq 0 \, , \ \beta_{A,\ u,v} \geq 0 \, , \ \gamma_{A,\ u,v} \geq 0 \, , \ \lambda_{A,\ u,v} \geq 0$ $\alpha_{R,\ u,v} \geq 0 \, , \ \beta_{R,\ u,v} \geq 0 \, , \ \gamma_{R,\ u,v} \geq 0 \, , \ \lambda_{R,\ u,v} \geq 0$

Table 2. Model inputs glossary	
$\boldsymbol{Z_{v,j}}$	Observed land use activity v in cell j
$\boldsymbol{GTotal_{v,SA}}$	Grand total amount of activity v in study area SA
$\boldsymbol{STotal_{v,ii}}$	Sub-total amount of activity v within the zone ii
$\boldsymbol{\eta_{x,j}}$	Suitability of geographic factor x for development in cell j
$\boldsymbol{\rho_{y,j}}$	Distance to transport infrastructure y from cell j
$\boldsymbol{ODTM_{jj,ii,m}}$	Journey time from origin zone ii to destination zone jj using mode m
$\boldsymbol{D_{i,j}}$	Distance between cells i and j

Solving the optimisation problem

Calibration for each study area was done by minimising the disagreement between actual and predicted land use distributions using an adaptation of the gradient descent algorithm with momentum and multiple start points. Table 3 summarises the mathematical functions involved in the gradient descent algorithm with momentum (see Ruder, 2016 for a more detailed discussion).

Table 3. The gradient descent algorithm with momentum	
Objective function to optimise	$\boldsymbol{min\, f}([\boldsymbol{x}]_t)$
Solution at iteration t	$[\boldsymbol{x}]_t = [\boldsymbol{x_1} \quad \boldsymbol{x_2} \quad \cdots \quad \boldsymbol{x_n}]_t$
Gradient at $[\boldsymbol{x}]_t$	$[\boldsymbol{g}]_t = \left[\frac{\partial f([x]_t)}{x_1} \quad \frac{\partial f([x]_t)}{x_2} \quad \cdots \quad \frac{\partial f([x]_t)}{x_n}\right]_t$
Solution update at iteration t	$[\boldsymbol{v}]_t = \boldsymbol{r}[\boldsymbol{g}]_t + \boldsymbol{p}[\boldsymbol{v}]_{t-1}$, $\quad$ $\boldsymbol{where\ 0 < p < 1}$
Solution at iteration $t+1$	$[\boldsymbol{x}]_{t+1} = [\boldsymbol{x}]_t - [\boldsymbol{v}]_t$

The parameter r was the learning rate dictating the length of steps taken at every solution update. The momentum factor, p, dictated the weight with which updates in the previous iterations are considered in current iteration. Each calibration started from 20 randomised initial solutions and the best solution at the end of the 2,000th iteration was used as the final solution. After conducting a pilot study, a learning rate of 10^{-6} and a momentum factor of 0.1 were used in model calibrations. For a small number of study areas, a finer calibration with a learning rate of 10^{-7} starting from informed solution points was necessary as the model diverged at the higher learning rate.

Clustering of study areas

Once the model has been calibrated against each study areas, clustering was done to identify groups of study areas with similar patterns of urban interactions. The calibration approach has been effective in untangling urban form into the individual interactions, which described how an average unit of land use development reacted to individual factors (e.g. neighbourhood reaction to each type of land use activities, each geographic suitability factors). These individual factors, in theory, could be decomposed even further to the effect of individual parameters. The calibration approach described in this paper stopped short of untangling these individual parameters' effect. Such endeavour would require a more microscopic analysis likely consisting primarily stated preference approach asking individual urban agents how they would react in a series of hypothetical situations. While such microscopic examination might contribute to theoretical study of urban system, it would not improve the model's practical use in planning and would requires much more effort, especially considering the number of study areas.

In other words, the calibration approach in this paper by examining locational revealed preference of agents in the urban form was not sensitive to the noise around the exact values of individual parameters effect. It was, however, sensitive to the individual interactions which were the combined effect of the constituent parameters. Therefore, curves formed by the individual parameters for each study area better described these individual interactions than the individual parameters' values. Cluster analysis was therefore conducted using the graph clustering method. The graph clustering method grouped similar curves together by taking gauge points along an axis. Curves were clustered based on their scores across the clustering variables. In the examples given in Figure 3, the 2nd and 3rd curves are more similar and therefore more likely to be clustered into the same group than the curve in the 1st instance. Clusters were formed using principal component analysis to extract curve(s) that were linearly uncorrelated to one another. Each of these principal components represented unique group of urban dynamics.

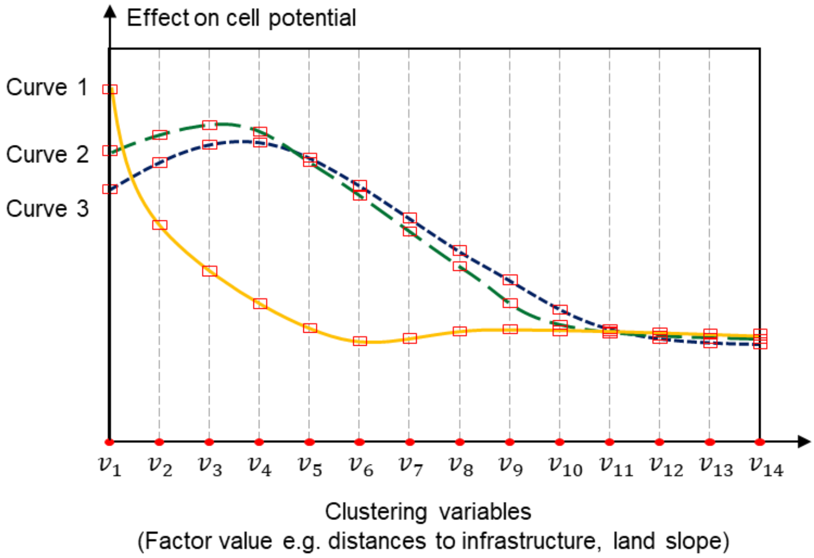


Figure 3 Illustration of the graph clustering method

3. Results

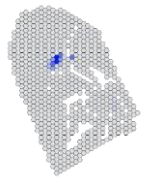
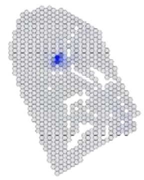
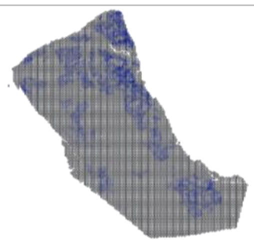
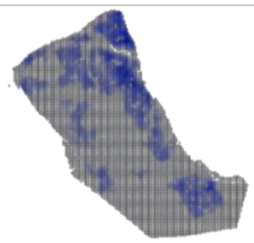
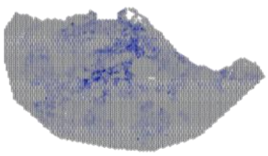
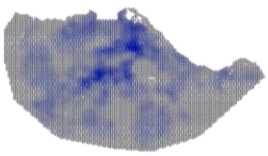
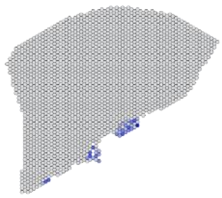
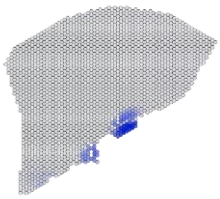
3.1. Individual Study Area Calibration Performance

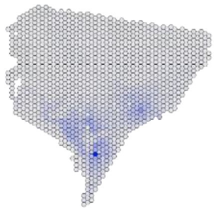
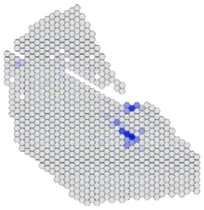
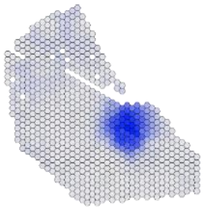
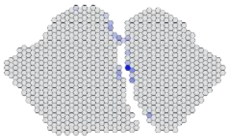
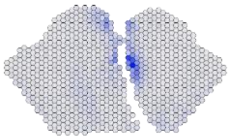
Performance of model calibration against the actual land use distribution in each study area was measured by Pearson correlation index as well as through visual comparison. Table 4 shows the Pearson correlation indices for the 56 study areas and Table 5 provides examples of the of visual comparison between actual and predicted distributions. The number of cells information are provided to indicate geographical size of the study area.

Table 4. Pearson correlation indices for the study areas						
Study Areas	Number of Cells	Residential	Port	Manufacturing	Consumer Services	Business Services
Aberdeen	12,747	0.776	0.572	0.466	0.634	0.572
Amble	727	0.846	0.512	0.445	0.877	0.415
Barry	1,193	0.791	0.540	0.414	0.409	0.227
Basingstoke	6,732	0.822	-	0.372	0.502	0.494
Bath	3,356	0.820	-	0.592	0.749	0.661
Birkenhead	11,259	0.877	0.479	0.463	0.563	0.362
Bodmin	708	0.777	-	0.349	0.892	0.333
Bristol	25,100	0.767	0.257	0.306	0.650	0.582
Burntisland	265	0.799	0.656	0.282	0.561	0.372
Cambridge	18,862	0.806	-	0.381	0.650	0.433
Cardiff	7,516	0.721	0.476	0.309	0.567	0.425
Cowes	895	0.817	0.246	0.399	0.537	0.840
Dover	2,262	0.836	0.622	0.332	0.626	0.378
Dundee	4,977	0.819	0.183	0.393	0.627	0.514
Durham	4,228	0.807	-	0.387	0.600	0.426
Edinburgh	6,692	0.773	0.315	0.265	0.450	0.688
Falmouth	1,937	0.820	0.410	0.246	0.496	0.327
Felixstowe	1,343	0.821	0.524	0.328	0.437	0.530
Fowey	426	0.780	0.156	0.353	0.960	0.368
Goole	3,313	0.832	0.522	0.390	0.714	0.392
Grangemouth	5,351	0.817	0.482	0.394	0.567	0.270
Grantham	2,071	0.790	-	0.446	0.684	0.431
Great Yarmouth	2,761	0.865	0.402	0.389	0.590	0.683
Grimsby-Immingham	5,500	0.872	0.558	0.467	0.603	0.361
Harwich	801	0.762	0.453	0.318	0.382	0.295
Hugh Town	292	0.736	0.239	0.402	0.618	0.359
Hull	8,427	0.836	0.454	0.458	0.556	0.442
Inverness	3,179	0.809	0.188	0.382	0.709	0.682
King's Lynn	1,602	0.776	0.297	0.253	0.716	0.426
Kirkcaldy	1,299	0.755	0.240	0.368	0.635	0.440
Kyle of Lochalsh	695	0.783	0.388	0.426	0.620	0.532
Leicester	34,842	0.726	-	0.540	0.718	0.505
Lerwick	1,291	0.787	0.218	0.393	0.433	0.467
Liverpool	11,247	0.689	0.389	0.444	0.551	0.689
Lowestoft	2,636	0.801	0.258	0.318	0.582	0.464
Milford Haven	3,213	0.838	0.408	0.454	0.651	0.311
Montrose	500	0.752	0.270	0.342	0.469	0.480
Newport	5,730	0.821	0.349	0.510	0.564	0.378
Perth	2,400	0.764	0.216	0.327	0.750	0.802
Peterhead	1,365	0.807	0.361	0.367	0.620	0.606
Plymouth	8,671	0.792	0.230	0.380	0.542	0.358

Study Areas	Number of Cells	Residential	Port	Manufacturing	Consumer Services	Business Services
Port Talbot	6,602	0.857	0.391	0.271	0.518	0.666
Portsmouth	4,317	0.812	0.324	0.333	0.457	0.242
Rye	715	0.685	0.576	0.975	0.901	0.400
Salisbury	2,428	0.784	-	0.480	0.785	0.612
Scunthorpe	4,948	0.885	0.431	0.486	0.519	0.578
Stranraer	633	0.776	0.193	0.337	0.595	0.556
Southampton	11,233	0.799	0.499	0.431	0.589	0.425
Swansea	10,795	0.675	0.342	0.332	0.594	0.490
Swindon	4,906	0.850	-	0.482	0.539	0.414
Teignmouth	809	0.866	0.424	0.315	0.734	0.563
Thurso	561	0.829	0.321	0.301	0.691	0.439
Tyneside	30,647	0.743	0.406	0.462	0.657	0.471
Ullapool	196	0.731	0.173	0.369	0.543	0.352
Wisbech	2,249	0.771	0.180	0.457	0.794	0.598
York	6,905	0.812	-	0.165	0.659	0.477

Table 5. Comparisons of model predictions with actual distributions

Study Areas*	Actual Distribution	Predicted Distribution	Correlation Index
Rye Consumer Services Cells: 715			*R = 0.901*
Birkenhead (Wirral) Residential Cells: 11,259			*R = 0.877*
Edinburgh Residential Cells: 6,692			*R = 0.772*
Dover Port Cells: 2,262			*R = 0.622*

Felixstowe Business Services Cells: 1,343			*R = 0.530*
Amble Manufacturing Cells: 772			*R = 0.445*
Cowes Port Cells: 895			*R = 0.246*
*Maps are not to scale			

3.2. Settlement Clusters and Their Urban Dynamics

Cluster analysis discovered two main groups of port cities in the study areas which differed especially in their interactions between their urban and port activities: “General Cities” and “Port-Dependent Cities”. These main groups could be further subdivided into two subgroups. The distinctions between the subgroups were attributed more to the effects of manufacturing and services activities. Additionally, Cardiff and Felixstowe were singled out as outliers. Figure 4 describes the overall hierarchy of these clusters.

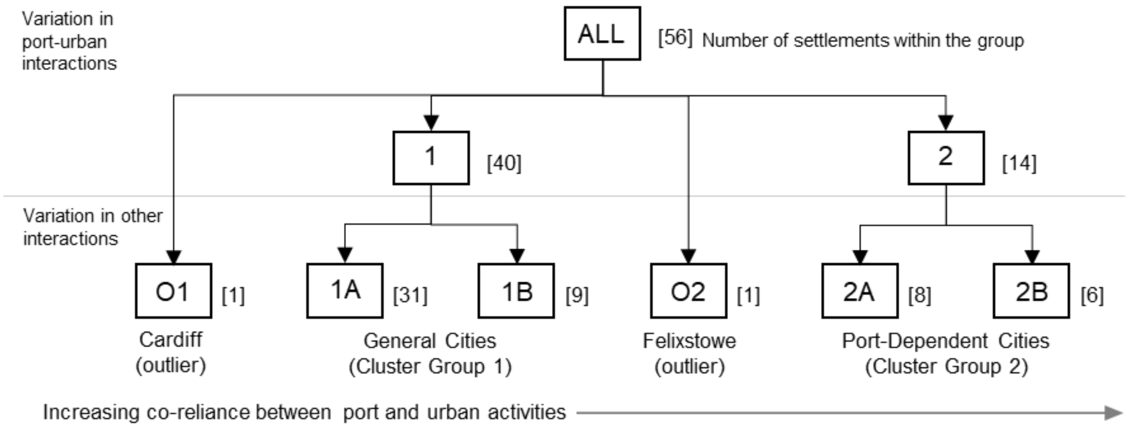
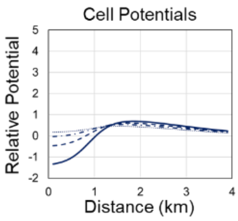
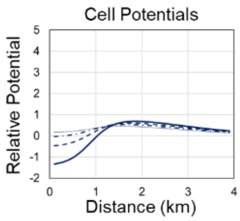
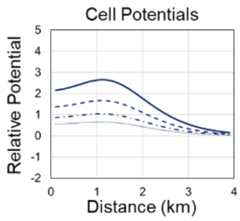
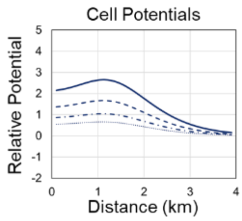
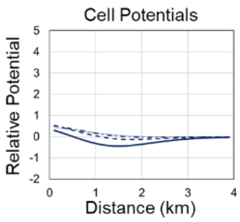
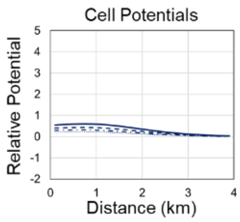
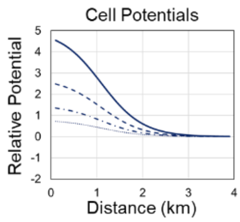
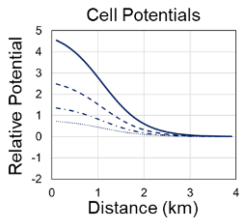
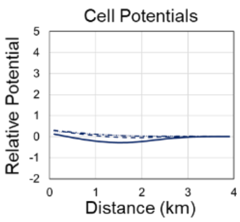
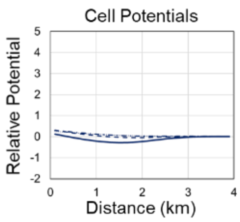
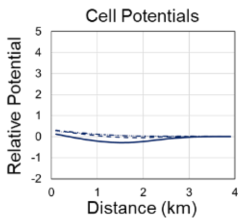
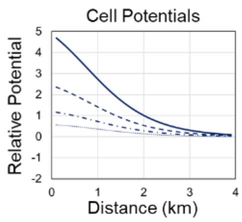


Figure 4. Clusters hierarchy from variations in neighbourhood effects

Neighbourhood effect

Some examples of the neighbourhood effect between port and other urban activities within the main clusters are presented in Table 6. In the mainstream subgroup, 1A, the biggest subgroup with 31 out of 56 settlements and containing all non-port settlements, the net effects of manufacturing on housing are negative in small distances and positive around a distance of about 2km. In the 9 settlements grouped as 1B, while the peak attraction occurred at about the same distance, housing developments are not repulsed by being too close to manufacturing activities. This was more typical of the urban dynamics between manufacturing and housing in smaller settlements in the second group. Meanwhile, the attraction of retail establishments on housing is stronger in subgroup 1A than in 1B. The subgroups of the second group, 2A and 2B, were more similarly sized at 8 and 6 settlements. In settlements in subgroup 2A, which were of relatively larger sizes than subgroup 2B, self-agglomeration behaviour of services activities was more prevalent, mimicking the behaviour of services activities in group 1.

Table 6. Examples port-urban interactions observed in the main clusters			
Cardiff	*General cities*	*Felixstowe*	*Port-dependent cities*
The effect of proximity to ports on housing			
			
The effect of proximity to ports on business services			
			
The effect of proximity to ports on consumer services			
			
individual level — low level — medium level — high level			

Some interactions, however, were found to be similar across all study areas including the effects of port on manufacturing as shown in Figure 5. Finally, Figure 6 summarises neighbourhood interactions observed in the model calibration. These star diagrams indicate the overall nature of interactions using green line and plus symbol (+) to represent attraction and red line and backslash symbol (\) for repulsion. Unbroken lines indicate interactions that are sensitive to distance while dashed lines indicate those that are not. The width of the lines indicates the strength of these interactions.

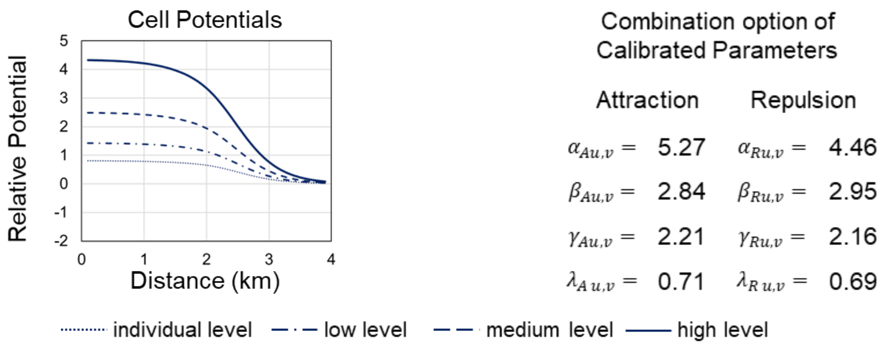


Figure 5. The effect of proximity to ports on manufacturing

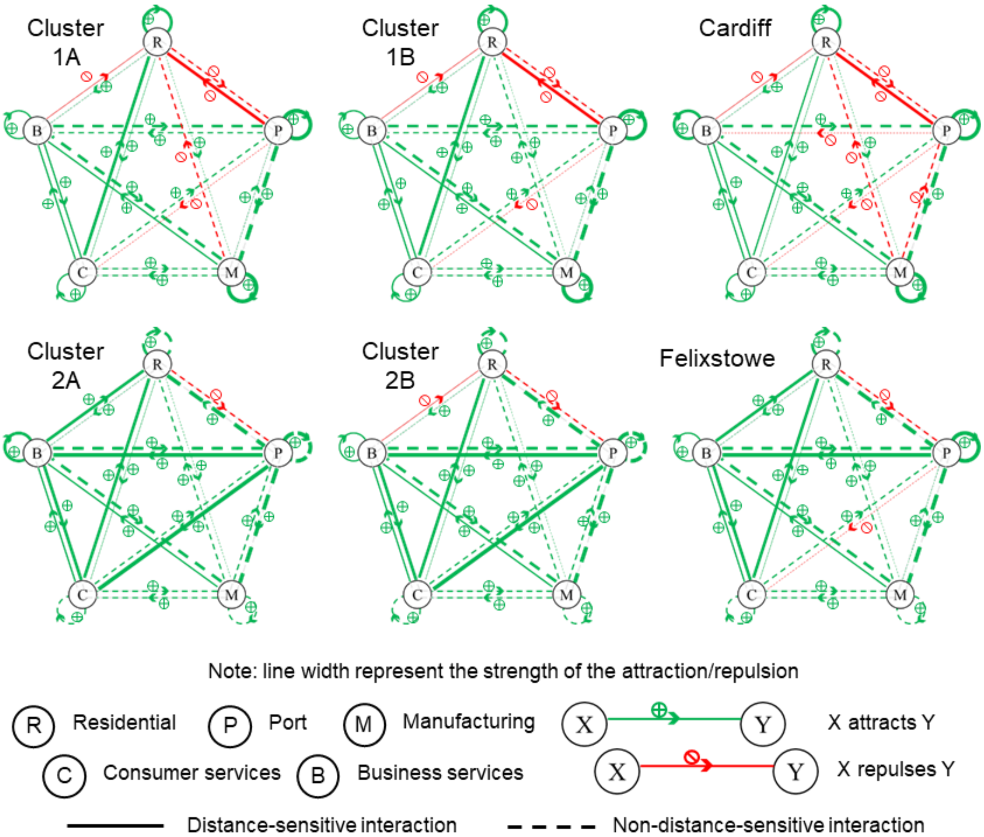


Figure 6. Summary of urban interactions in the settlements observed

Geographic and transport effects

Moving on to transport and geographic effects, these were generally found to be similar across the study areas. One of the key findings was that employment activities benefitted from access to housing developments although the neighbourhood effects of housing on other land uses were found to be generally weak. Consequently, areas with high levels of employment activities were generally accessible (due to sensitivity to transport effect), but not necessarily nearby (due to insensitivity to neighbourhood effect), from residential areas. Further, access to rail connection was found to have a stronger attraction effect on employment activities rather than housing developments. An exception to this was on smaller and more compact settlements where housing developments clustered more closely to rail stations. This is shown in Figure 7 with the dashed line being more sensitive to distance than the unbroken line.

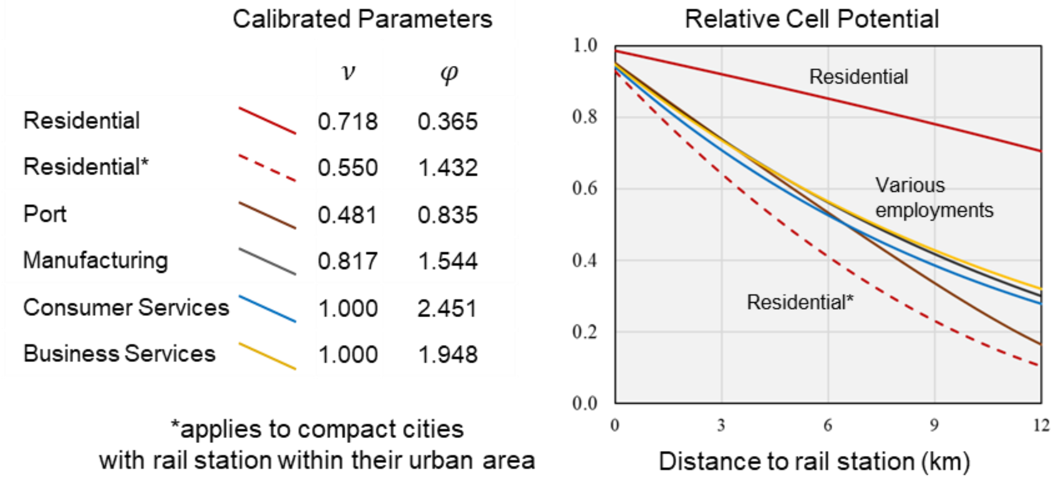


Figure 7. The effect of proximity to rail station on cell potentials

4. Discussion

4.1. Model Performance

The model generally performed better for predicting residential and consumer services distributions which were more dispersed across the study areas and had more cell-to-cell variation. However, it should be noted that the Pearson correlation calculation penalised misallocation of activities to nearby cells as severe as they do misallocation to further cells. As the urban CA model predicts 'whereabouts' activities are located rather than their exact location (Van Vliet et al., 2013a), misallocation of activities by 1 or 2 cells should have better agreement than misallocation by 20 cells. Fuzzy Kappa simulation (Van Vliet et al., 2013a) considers the geographical nuances of urban CA model to examine the agreement between two land use maps but this was based on categorical cell states rather than the numerical cell states containing information of activity levels, as used in the model in this paper. Therefore, in the absence of a more suitable index, this paper opted for the Pearson correlation index for its simplicity but complemented it with visual comparison.

The comparisons in Table 5 show that while higher Pearson correlation indices would indicate a more precise match between actual and predicted distribution, lower values still indicated an approximate match of the activity distributions. This observation supports that the model could provide a sensible approximation of the urban interactions within the study areas.

4.2. Settlement Clusters and Their Urban Interactions

The efficient calibration process has allowed quantification of urban dynamics in the general context of port cities, as opposed to case specific examinations currently dominating the field (Ng et al., 2014). The clustering process summarised the calibration results for individual study areas by grouping them into clusters. The general cities cluster was populated by non-port settlements and relatively larger port settlements. Meanwhile, the port-dependent cities cluster consisted of smaller port settlements. In the first group, the attractions of ports to other urban land-use activities were either small or negative, while these effects were more positive in settlements in the second group. Most cities within the second group were coincidently smaller than general cities. Consequently, competition for land between different activities may be less severe.

In other words, the presence of port within port-dependent cities attracts more urban development than they do in general cities. Additionally, this indicated that the presence of ports was a distinguishing factor in the urban dynamics of small settlements, while less prominent differences were observed between the urban dynamics of port and non-port settlements in larger study areas.

Overall, these findings are consistent with existing evidence in the port-cities literature but go further in quantifying the interaction between urban agents within port-urban systems of various sizes and types. In this regard, the majority of the study areas were classified into the general cities cluster while port-dependent cities consisted mainly of smaller port settlements. This supports Hall and Jacobs' (2012) observations that in the system where both port and urban components are small, the urban economy relies heavily on activities directly related to the port. As the system grows, however, this reliance seems to weaken. This was true even when the urban growth has been driven by the port such as in Dover. One explanation is that as ports grow, they induce the growth of interlinked industries (Kwak et al., 2005) which may not be directly related to port activities such as retail and leisure to accommodate the needs of ports' labour. These less impactful activities, in turn, become more attractive for housing and other forms of urban development rather than the port, thus weakening the port-urban interactions in terms of wider urban form.

Felixstowe and Cardiff were found to be outliers. Although, it is possible that there will be other global port cities not included in this study that would have similar urban activities to Felixstowe or Cardiff.

Felixstowe appeared to show similar port-urban interactions as port dependent cities but had some elements of general cities' behaviours such as the weaker relationships between port and consumer services activities. Port activities in Felixstowe took off in 1967 as the first purpose-built container terminal in the UK (Hutchison Ports, n.d.). The nature of container trading and the advancement of motoring and rail industry around this time allowed cargo to be easily transported to manufacturing or retail facilities elsewhere in the country. Therefore, industrial and retail activities relying on Port of Felixstowe's trade might choose to locate in more attractive centres such as the Greater London or the Midlands regions leaving businesses more closely related to ports in Felixstowe. This would have alleviated some of the land use competition which Ducruet and Lee (2006) argued to cause tension between port and urban developments.

Meanwhile, Cardiff showed an even weaker port-urban connection than general cities. The circumstances of port development in Felixstowe and Cardiff might have contributed to their uniqueness. This could be due to the weaker maritime history in Cardiff. Generally, port settlements within the general cities cluster have had port in the centre of their urban development. Meanwhile, prior to the discovery of coal, Cardiff was a small non-maritime village. The port was built during the coal boom in the 19th century away from the then-village-centre and therefore the city and the port developed as separate entities (Jenkins, 2007).

4.3. Predicting Policy Impacts in Port City

In order to demonstrate the application of the urban interaction insights from model calibration as tool to assist in planning, the model has been used to simulate urban growth in Southampton under different port development policy scenarios. The Associated British Port (ABP), Port of Southampton's operator, has been planning to turn Dibden Bay, a site on the western bank of River Test opposite the current port site, into a container terminal to alleviate port capacity and congestion problems in the current container terminal closer to the city centre. However, these plans have been rejected due to the potentially detrimental impacts of such development on the environment.

The long-term effects of such policies are often not straight forward. Lu and colleagues (2021) have demonstrated that, expansion of cities can either alleviate or worsen traffic-related air pollution depending on the distribution of activities and the available transport infrastructure.

A simulation LUTI model has used the urban dynamics described in the previous section to predict long-term land use and transport impacts of different policy options. The model took Southampton's 2018 activity distribution and simulated 50-year's growth under restriction policy where port development is not allowed within Dibden Bay and promotion policy where port development in Dibden Bay is encouraged. A constant annual growth rate was assumed based on Southampton's average growth rate from 2010 to 2020. Some of the model's results are presented in Figure 8.

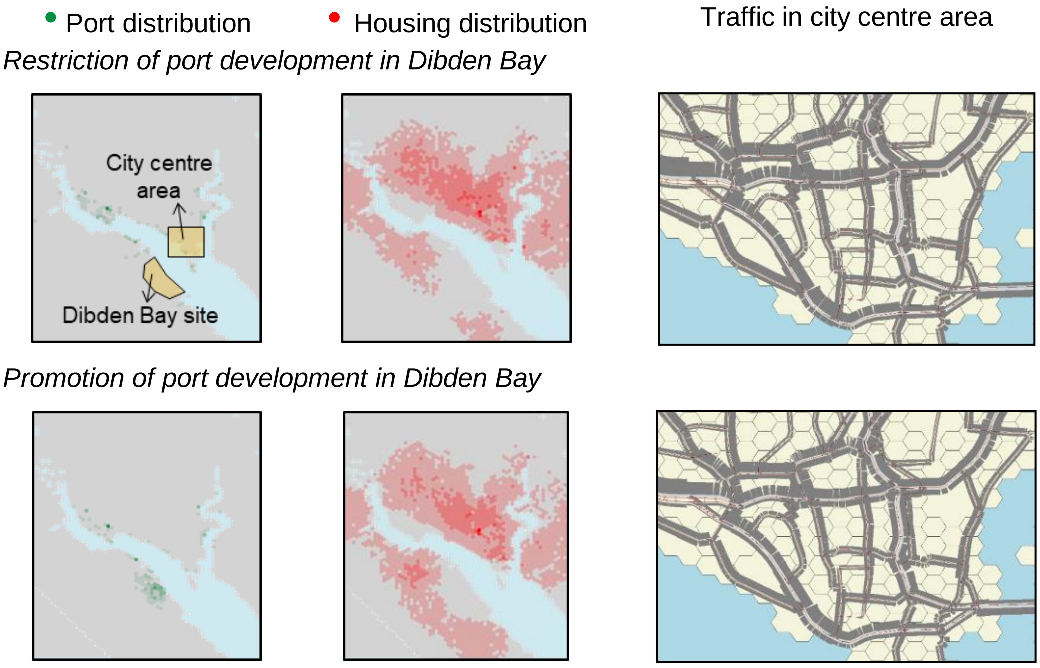


Figure 6. Simulations of long-term development policy impact in Southampton

Promotion policy predictably resulted in the rise of port development in Dibden Bay. However, this would also trigger housing developments around the site. This is due to the attraction of access to port employment in Dibden Bay, but noticeably this occurred at some distance away rather than immediately next to the port as proximity to port is not necessarily attractive for housing. Mainly, this resulted in housing densification in existing sub-centres north (Marchwood) and south (Hythe) of Dibden Bay, but they also sprawl to greenfield site enveloping Dibden Bay. While this may alleviate housing needs near the city centre, there could be additional loss of green spaces nearby Dibden Bay. The model could also be used to evaluate complementary policies to prevent this such as designating greenbelt around Dibden Bay or more potent high-density development policies where further development could only happen within currently developed areas

Promotion policy also resulted in the current port site nearby the city centre being overtaken by retail (not pictured) and residential development. This indicated a similar trend that Liverpool has seen in the historic docks and indeed would be continuing an existing trend in Southampton.

The transport sub-component of the model, based on the traditional 4-stage transport model, also detected some changes in traffic conditions between the two policies. By promoting port development in Dibden Bay, city centre traffic was alleviated by the removal of port traffic away from the city, but this was replaced by currently suppressed urban traffic, resulted only in minor reduction of traffic around the city centre. This meant that additional measures to promote public transport and non-motorised transport would be required to lock-in the traffic benefits of the port's migration to Dibden Bay. Additionally, transport infrastructure connecting the new port site to the rest of the region would likely need improvement to handle the increased demand. The model therefore could contribute to the policy decision-making process.

5. Conclusions

This paper has presented an automated calibration approach to calibrate an urban CA model, thus allowing quantification of urban dynamics in the general context of port cities. The findings of this paper built on existing port-cities literature but went further in quantifying interactions within port-urban system of various sizes and types.

This paper has demonstrated improved methods for the calibration of LUTI models based on CA which provide insights into the behaviours of individual interactions between urban agents. Such findings provide important insights for future research by contributing to a better understanding of the dynamics between port and urban agents within a port city system. By doing so, this paper provides the tool to better-predict how urban activities in port cities will react to different policy scenarios. These are also invaluable for the transport and urban planning of port-cities as they enable planners to better predict the longer-term consequences of their interventions.

By using a technique from computer science, this paper has examined a large and varied set of port-urban settlements and quantified the interactions between urban agents within them. This paper, therefore, has expanded what is currently known about the port-urban system both in terms of breadth and of depth.

Acknowledgements
This work made use of Ordnance Survey's AddressBase® Plus dataset, supplied in 2019, under the authorisation given (Contract Number: 40133005). © Crown copyright (2019) OS.

This work was supported by Indonesia's Ministry of Finance through The Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan – LPDP) Program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funder.

References

Akhavan, M. (2017). Development dynamics of port-cities interface in the Arab Middle Eastern world – The case of Dubai global hub port-city. *Cities,* 60, pp.343-352.

Aljoufie, M., Zuidgeest, M., Brussel, M., van Vliet, J., van Maarseveen, M. (2013). A cellular automata-based land use and transport interaction model applied to Jeddah, Saudi Arabia. *Landscape and Urban Planning*, 112, 89-99.

Atkins (2009). *Joint Local Transport Plan – The Greater Bristol Modelling Framework (JLTP GBMF) Modelling Support: G-BATS3 v2.3 Highway Local Model Validation Report. Appendix 2.*

Batty, M. (1997). Cellular automata and urban form: A primer. *Journal of the American Planning Association*, 63(2), 266-274.

Batty, M. (2007). *Cities and Complexity: understanding cities with cellular automata, agent-based models, and fractals*. The MIT Press.

Ducruet C., and Lee, S. (2006). Frontline soldiers of globalisation: port-city evolution and regional competition. *Geojournal,* 67(2), pp.107-122.

Fleming, D.K. and Hayuth, Y. (1994). Spatial characteristics of transportation hubs: Centrality and Intermediacy. *Journal of Transport Geography,* 2(1), pp.3-18.
Gutiérrez, J. (2001). Location, economic potential and daily accessibility: An analysis of the accessibility impact of the high-speed line Madrid-Barcelona-French border. *Journal of Transport Geography*, 9(4), pp.229-242.
Hall, P.V. and Jacobs, W. (2012). Why are maritime ports (still) urban, and why should policy-makers care? *Maritime Policy & Management*, 39(2), pp.189-206.
Hutchison Ports (n.d.). *Port of Felixstowe*. Available from: https://www.portoffelixstowe.co.uk/
Hoyle, B.S. (2000). Global and local change on the port-city waterfront. *The Geographical Review*, 90(3), 395-417.
Jenkins, G.H. (2007). *A concise history of Wales*. Cambridge University Press: Cambridge.
Kwak, S.J., Yoo, S.H., and Chang, J.I. (2005). The role of the maritime industry in the Korean national economy: an input—output analysis. *Marine Policy,* 29 (4), pp.371-383.
Lee, S.W., Song, D.W., & Ducruet, C. (2008). A tale of Asia's world ports: the spatial evolution in global hub port cities. *Geoforum,* 39(1), pp.372-385.
Lu, J., Li, B., Li, H., and Al-Bakarani, A. (2021). Expansion of city scale, traffic modes, traffic congestion, and air pollution. *Cities,* 108.
Merk, O. (2013). The Competitiveness of Global Port-Cities: Synthesis Report. *OECD Regional Development Working Papers*, 2013/13, OECD Publishing.
Ng, A.K.Y., and Ducruet, C., Jacobs, W., Monios, J., Notteboom, T.E., Rodrigue, J., Slack, B., Tam, K., and Wilmsmeier, G. (2014). Port geography at the crossroads with human geography: between flows and spaces. *Journal of Transport Geography,* 41, pp.84-96.
Nugraha, A.T, Waterson, B.J, and Blainey, S.P. (2018). A modelling framework of land-use and transport interaction to assess co-evolution of port and urban development. 50th Annual Conference of UTSG, UK.
Nugraha, A.T., Waterson, B.J., Blainey, S.P., and Nash, F.J. (2020). On the consistency of urban cellular automata models based on hexagonal and square cells. *Environment and Planning B.*
RIKS (2010). *Metronamica Documentation.* RIKS BV: Maastricht.
Ruder, S. (2016). An overview of gradient descent optimization algorithms, *arXiv preprint* arXiv:1609.04747.
Samat, N. (2006). Characterizing the scale sensitivity of the cellular automata simulated urban growth: a case study of the Seberang Perai Region, Penang State, Malaysia. *Computers, Environment, and Urban Systems,* 30, pp.906-920.
Santé, I., García, A.M., Miranda, D., and Crecente, R. (2010). Cellular Automata Models for the Simulation of Real-World Urban Processes: A review and analysis. *Landscape and Urban Planning,* 96, pp.108-122.
Van Vliet, J., Hagen-Zanker, A., Hurkens, J., Van Delden, H. (2013). A fuzzy set approach to assess the predictive accuracy of land use simulations. *Ecological Modelling*, 261–262 (2013), pp.32-42.
Van Vliet, J., Hurkens, J., White, R., and Van Delden, H. (2012). An activity based cellular automaton model to simulate land-use changes. *Environment and Planning B,* 39, pp.198-212.
Van Vliet, J., Naus, N., van Lammeren, R.J.A., Bregt, A.K., Hurkens, J., and Van Delden, H. (2013b). Measuring the neighbourhood effect to calibrate land use models. *Computers, Environment, and Urban Systems,* 41, pp.55-64.
Woo, S., Pettit, S.J., Kwak, D., and Beresford, K.C. (2011). Seaport research: A structured literature review on methodological issues since the 1980s. *Transportation Research Part A*, 45, pp.667-685.
Yochum, G.R., and Agarwal, V.B. (1988). Static and changing port economic impacts. *Maritime Policy & Management,* 15(2). Pp.157-171
Young, M. (2016). An Automated Framework to Derive Model Variables from Open Transport Data using R, PostgreSQL and OpenTripPlanner. *Paper presented at 24th GIS Research UK Conference*, London, March.