

Research and analysis **Methodology: How** childcare could be optimised across local areas

Published 14 May 2025

Applies to England

Contents

Introduction

Objective function

Genetic algorithm

Quadratic programming

Sensitivity analysis

Local authority case studies

Socioeconomic analysis

Limitations of the two-step approach

Appendix

Introduction

The aim of our approach is to achieve equality in access to childcare in England. In practice, this is equivalent to minimising the variability in childcare access across the country.

Our approach follows a two-step optimisation process. First, a genetic algorithm identifies the optimal locations for new childcare provision. Then, quadratic programming is applied to determine the optimal capacity of each provision, ensuring an efficient distribution of resources.

This method is based on the work by Xiang Li, Fahui Wang and Hong Yi, who proposed a two-step approach for facility planning to enhance equal accessibility. [footnote 1] The results from their paper showed that strategically locating new facilities has a greater impact on reducing accessibility inequality than optimising capacity alone.

We build on this and our previous work by applying their process to the context of childcare provision in England.

Objective function

Our objective is to improve equality in access to childcare by minimising variance in accessibility across different locations. To achieve this, we have defined an objective function that minimises Â, the total variation of accessibility at each demand location weighted by the amount of demand there. A smaller value indicates less disparity in accessibility, making this equivalent to maximising equality in access to childcare.

To quantify accessibility, we have extended the 2-step floating catchment area (2SFCA) method which we used previously.^[footnote 2] This 2SFCA method accounts for both the supply of childcare places and the demand from nearby families, producing an accessibility index (A_i). We also introduced a binary dummy variable (B_i) to indicate whether a candidate provision is selected to have childcare.

We used Census 2021 figures for children aged 0 to 7 to provide a measure of potential demand for childcare. More up-to-date population estimates are available, but we used Census 2021 in order to maintain consistency with our previous analysis.

As in our previous analysis, to calculate the accessibility index, we have used travel

time catchments of 15 minutes for travel by car and 25 minutes for public transport. This includes a maximum walking time of 15 minutes for users of public transport. We chose these times based on a reasonable estimate for the maximum time parents would travel to access childcare, but these times could be adjusted in the future with additional evidence to show how long parents are willing to travel.

The full derivation of the objective function can be found in our appendix, but the final form is:

$$\operatorname{Minimise}(\hat{A}) = \operatorname{Minimise} \sum_{i=1}^{m} (A_i - \bar{A})^2 D_i = \operatorname{Minimise} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} B_j C_j F_{ij} - \bar{A} \right)^2 D_i \qquad (1)$$

Equation shows an objective function where the objective is to minimise the total variation of accessibility at each demand location.

It reads: Minimise A-hat, which equals the sum from i = 1 to m of (Ai minus A-bar) squared times Di, which is equivalent to minimising the sum from i = 1 to m of the square of (the sum from j = 1 to n of Bj times Cj times Fij, minus A-bar), all multiplied by Di.

Where:

is the total variation of accessibility at each demand location

A_i is the accessibility index

Ā is the mean of accessibility

D_i is the demand of the location

B_i is the location of the childcare provision

Ci is the capacity of the provision

 F_{ij} is the normalised decay weight and a Gaussian function (see appendix)

m is the number of users

n is the number of candidate childcare provisions

i is the index of users from 1 to m

j is the index of candidate childcare provisions from 1 to n

Therefore, the optimisation problem is to search for optimal values of B_j and C_j that minimises the above equation.

This problem is solved in 2 steps. First, a uniform capacity is assumed for all provision and solves for the optimal values of B_j i.e. is there provision or not. This first step uses a genetic algorithm. Then with the locations of provision determined, the optimal capacity of each provision is found using quadratic programming.

Genetic algorithm

A genetic algorithm is a method to solve for both constrained and unconstrained optimisation processes. It is a metaheuristic algorithm. This is a type used in computer science and characterised by its ability to solve complex optimisation problems by mimicking natural phenomena.^[footnote 3] Genetic algorithms are inspired by natural selection and are used to produce high-quality solutions to complex problems using selection, crossover and mutation.^[footnote 4] Genetic algorithms modify a population of individual solutions, and at each step select individuals to produce 'children' to evolve to an optimal solution.

Important terms

[footnote 5]

Population: a population is a group of individuals (or chromosomes); each individual is a potential solution to the problem.

Individual/chromosome: a chromosome is an individual that contains a set of parameters that are called genes – in this case, a set of childcare provision locations.

Gene: the binary building blocks of the individual are represented as either 0 (no childcare provision) or 1 (childcare provision). The number of genes is decided by how many potential site locations there are in the local area. We have defined a potential site as one output area (OA), which is the level of output area used in the census.^[footnote 6]

Parents: in the context of a genetic algorithm, parents are existing potential solutions selected based on their performance.

Children: children are new potential solutions created by combining features of the parent solutions through operations like crossover and mutation.

Figure 1: Population containing 5 individuals and 10 genes within the individual



Figure 1 shows an illustration of important genetic algorithm terms. An individual is depicted as a vertical column divided into smaller units, each representing a gene.

How does the algorithm work?

The genetic algorithm selects individuals from the current population to be parents and uses those to produce children for the next generation. Over multiple generations the population evolves towards an optimal solution.

There are 4 main rules used as part of the genetic algorithm.

- 1. Selection rules: these select the individuals (parents) that contribute to the population of the next generation. This is determined by the fitness function.
- 2. Crossover: combine 2 parents to form children for the next generation.
- 3. Mutation: apply random changes to the parents to form additional children.
- 4. Termination: reproduction is repeated until a termination condition has been reached.

The genetic algorithm process is summarised in the following figure.





Figure 2 is a flowchart illustrating the process of a genetic algorithm. It begins with Start, followed by Initial population, then Create fitness values, Selection, Crossover, and Mutation. After Mutation, the algorithm checks a condition: 'Termination: Number of generations below required threshold?' If the answer is No, the process loops back to Create fitness values. If it is Yes, the process proceeds to End.

Calculate fitness value

In each generation, we assess the fitness of each solution and use this to evaluate the potential solutions. The fitness function outputs the quality of the solution, its fitness value. In our example, the fitness function is the calculation of accessibility scores for each individual and its resulting variance. The calculation of the accessibility score is explained in our methodology document.^[footnote 7]

Selection

Selection is where the parents are selected to produce a child. There are several selection methods available.^[footnote 8] We have chosen to use a combination of 'elitism' and 'tournament selection' as, in our test cases, this produces outputs with the least variance. In tournament selection, to choose parents, a subset of individuals is chosen at random, and the fittest of them is selected. We tested multiple tournament sizes (that is, the number of individuals in the subsets), and chose to use 3. Our implementation of elitism always carries the best individual from the previous generation into the next, bypassing crossover and selection, ensuring high-quality solutions are maintained.

Crossover

After parents are selected, crossover 'reproduces' new individuals from these parents. We are using a one-point crossover, where the parent 1 and parent 2's genes copied onto child 1 and child 2 respectively, until a random point, when which parent's are copied onto which child's is swapped, so from that point parent 1's genes are copied onto child 2, and parent 2's are copied onto child 1's. The crossover rate ranges from 0 to 1 and can be optimised. A rate of 1 means all offspring are produced through crossover, while a rate of 0 results in the new generation being direct copies of the parents, with changes occurring only through mutation. We performed a global sensitivity analysis to determine the optimal value. See <u>our section below</u> on this.

Figure 3: Crossover from Parent 1 and Parent 2 to form Child 1 and Child 2



Figure 3 shows an illustration of a genetic algorithm crossover: 2 parent populations – one light grey, one dark grey – swap genetic material. An arrow leads to 2 new populations where individuals are half light grey and half dark grey, representing the mixing of genes.

Mutation

The mutation rate determines how many individuals should be mutated in one generation. The mutation rate maintains genetic diversity between generations and prevents the algorithm getting stuck in local optima. The mutation rate can be between 0 and 1. A higher mutation rate increases randomness in the algorithm, while too small a mutation rate may cause it to converge prematurely and get stuck in local optima.

Figure 4: Mutation for Child 1 and Child 2



Figure 4 is an illustration of a genetic algorithm mutation: 2 child populations are shown with random light grey and dark grey genes.

Termination criteria

The termination criteria determine when the genetic algorithm should stop. There are several different common termination criteria, but we chose to end the genetic algorithm after a specific number of generations. We tested for the optimal number and found that the fitness value tended to stabilise around 20 generations for the parameters we were using.

Quadratic programming

In the first stage, we that assumed the capacity of each childcare provision is the same. Once we determined the optimum provision locations of the candidate sites, we applied quadratic programming to determine the optimum capacity of each

selected provision. This should further reduce the variance in the accessibility scores.

Quadratic programming is the process of solving optimisation problems involving quadratic functions. In our case, the variance is a quadratic function and minimising this becomes a quadratic programming problem.

We minimised equation (1) using Sequential Least Squares Quadratic Programming (SLSQP).^[footnote 9] SLSQP is an iterative process for bound-constrained non-linear optimisation, and our problem involves bounds and constraints. SLSQP uses a gradient-based approach, iteratively updating the solution using a quadratic approximation of the objective function while ensuring feasibility within the given constraints and bounds. Other methods, such as the Limited-memory Broyden–Fletcher–Goldfarb–Shanno Bound algorithm (L-BFGS-B) and trust-region approaches, were considered. However, testing showed that SLSQP provided smaller variance and converged more reliably than the alternative methods.

What are the bounds and constraints?

Unlike the genetic algorithm, the quadratic programming is a deterministic process. This means that as we add in additional bounds and constraints to the equation, the variance will be greater.

We determined that our bounds are the upper and lower limits of a childcare provision's capacity. We have 2 constraints in our model:

- 1. The capacity of each provision is a positive number
- 2. The sum of all provision capacities is the weighted sum by child population

Sensitivity analysis

To determine the optimal values for the genetic algorithm's population size, mutation rate and crossover rate we performed a global sensitivity analysis. [footnote 10] The values were based on those used in previously published studies.[footnote 11] The values we chose to analyse were as follows:

Population size: 5, 10, 20, 40

Mutation rate: 0.001, 0.005, 0.01, 0.05, 0.1 Crossover rate: 0, 0.2, 0.5, 0.8, 1

The sensitivity analysis identified the smallest variation across both the genetic algorithm and quadratic programming. This approach was necessary because optimising the genetic algorithm in isolation could lead to poorer convergence during the quadratic programming step. We repeated the model twice for a small local authority to reduce computation time.

We found the optimal parameters for the genetic algorithm were:

- Population size: 20
- Mutation rate: 0.005
- Crossover rate: 0.2
- Number of generations: 20

Local authority case studies

We have produced 2 case studies – the first looking at the local authority of Lincolnshire, and the second at Sandwell and Walsall combined.

Lincolnshire is a local authority with a mix of rural and urban areas. It has a higherthan-average level of childcare accessibility, but also has a high variability in childcare accessibility across areas. This makes it a suitable case study for testing whether our method can help create a more balanced distribution of childcare provision.

Walsall and Sandwell both rank below the 4th percentile for childcare accessibility. In both cases, this holds whether we only consider Ofsted-registered provision or include school-based provision regulated by the Department for Education. In our analysis, we only consider childcare places in providers registered on Ofsted's Early Years Register (EYR). However, even when including school-based provision, as demonstrated in a Victoria University study, the accessibility ranking remains unchanged.^[footnote 12]

Due to the computational demands of our optimisation approach, we have not been able to replicate the methodology from our previous publication at the same scale. Instead, we have applied our method to a subset of local authorities to demonstrate its effectiveness. For details on our previous methodology and how accessibility scores were originally calculated, please refer to our first publication. [footnote 2]

Since these case studies focus on specific local authorities in isolation, the accessibility scores produced in this publication cannot be directly compared to those in our previous publications. In our earlier work, accessibility scores were calculated at the OA level across England and provided national-level comparisons. However, because we are looking at local authority case studies in this publication, our figures may not include families who cross local authority borders to access childcare provision, either by travelling from outside our selected local authority to use provision within them, or by leaving to access provision in neighbouring local authorities.

We have aggregated results and present them at the lower-layer super output area (LSOA) level in this publication. To show potential movement, we have plotted confidence boundaries indicating which OAs may access provision in another local authority. This does not change the underlying data, but simply provides a visual representation of areas where access to childcare across local authorities may be likely.

Case Study 1: LincoInshire

Figure 5: Child population in Lincolnshire, Census 2021



Figure 5 is a map showing areas of Lincolnshire shaded from pale yellow to dark red to represent childcare population, with dark red indicating a larger population. Place names are labelled in black. Key towns like Lincoln, Boston and Grantham are marked.

The variability in accessibility decreased at each step of our process. Initially, accessibility variation in Lincolnshire was 14 childcare places accessible per 100 children, ranging between the lowest LSOA having 0.1 places per 100 children and the highest having 132 places per 100 children. This large difference and variation highlight the significant disparity in childcare accessibility, with some areas facing severe shortages while others have a surplus of available places.

After applying the genetic algorithm, this variation reduced to 10 places per 100 children, as the genetic algorithm optimised the location of childcare provision to improve overall access.

The final step changed the allocation of places within the provision, reducing

variability to 6 places per 100 children. Now, the accessibility ranged between 0 places per 100 children and 35 places per 100 children. This shows that our method has reduced the inequality in childcare accessibility, making provision more evenly distributed across Lincolnshire.

Due to the potential for families to cross local authority borders to access childcare, we are only confident in the childcare accessibility calculations for the areas where this cannot happen within our predetermined travel time. In Lincolnshire, 78% of the OAs' accessible childcare provision falls within the same local authority. There are no OAs which can only access areas outside of Lincolnshire.

Figure 6: Confidence boundaries in Lincolnshire



Figure 6 shows an outline of Lincolnshire shaded from white to black to represent confidence boundaries. The darker areas around the borders shows that families can access childcare provision in a neighbouring local authority. 78% of output areas in Lincolnshire can only access provision within Lincolnshire, making this figure largely white in appearance.

Case Study 2: Sandwell and Walsall

Figure 7: Child population in Sandwell and Walsall, Census 2021



Figure 7 is a map showing areas of Sandwell and Walsall shaded from pale yellow to dark red to represent childcare population, with dark red indicating a larger population. Place names are labelled in black. Key towns like Walsall, Smethwick and West Bromwich are marked.

Initially, accessibility variation in Sandwell and Walsall combined was at 4.8 childcare

places accessible per 100 children, ranging between the lowest LSOA with 2.5 places per 100 children and the highest with 25 places per 100 children. After applying the genetic algorithm, this variation reduced to 2.4 places per 100 children. The final step changed the allocation of places within the provision, reducing variability to 1.9 places per 100 children. Now, the accessibility ranged between 11 places per 100 children and 19 places per 100 children.

Due to the potential for families to cross local authority borders to access childcare, we are only fully confident in the childcare accessibility calculation in 2% of OAs in Sandwell and Walsall combined. This 2% can only access provision within Sandwell and Walsall within the predefined travel time. The reason for this is that Sandwell and Walsall are predominantly urban areas with short travel distances and well-connected public transport. These factors make it easier for parents and carers to cross local authority boundaries to access childcare provision in neighbouring areas. There are no OAs which can only access areas outside of Sandwell and Walsall.

Figure 8: Confidence boundaries in Sandwell and Walsall



Figure 8 shows an outline of Sandwell and Walsall shaded from white to black to represent confidence boundaries. The darker area around the borders shows that families can access childcare provision in a neighbouring local authority. There is a gradual transition. The centre of Sandwell and Walsall is white and fades to black as proximity to the border increases.

Potential sites

As described above, the number of genes in our genetic algorithm is decided by how many potential site locations there are in the local authority. We have defined a potential site as one OA which contains 40 to 250 households, with an average of 28 children aged 7 and under as of 2021. Therefore, the number of potential sites in Lincolnshire is 2,460. Sandwell and Walsall have a combined total of 1,830.

Constraints and bounds

We are constrained by the number of EYR childcare places as of 31 December 2024: [footnote 13]

- Lincolnshire has a total of 16,595 places
- Sandwell and Walsall combined have 9,662 places

Our optimisation ensures that the total number of places remains the same as the current provision. Additionally, capacity bounds, dictated by current capacity, are applied to individual provision:

- Each output area in Lincolnshire with places must have between 3 and 197 places
- Each output area in Sandwell and Walsall with places must have between 3 and 168 places

Socioeconomic analysis

We computed the changes in childcare places and childcare accessibility between the original 31 December 2024 data and our optimised scenarios. We then tested for statistically significant correlations between these changes and several socioeconomic variables at a significance level of 0.05. We used multiple models and ensured results were consistently significant across all of them.

In cases where we compared to aggregated changes in childcare accessibility scores, these were aggregated by producing a weighted average of accessibility scores, weighted by their population of 0- to 7-year-olds from Census 2021 data.

In cases where we tested multiple socioeconomic variables against the same change in places or accessibility, we corrected for the increased likelihood of false positives (finding a correlation when there was not one present) by applying Bonferroni corrections, dividing our significance level by the number of variables tested.

Deprivation

We used ranks of the Income Deprivation Affecting Children Index (IDACI) from the English indices of deprivation 2019 – LSOA Level to gather data on the proportion of children aged 0 to 15 in income deprived households.^[footnote 14] We used a lookup to link the 2021 LSOA data to the IDACI data, which was measured for 2011 LSOAs.

We compared this data to changes in childcare accessibility score and places aggregated to LSOA level. We found significant negative correlations between IDACI rank and change in childcare accessibility score in both case studies, with p-values of less than 0.001 for both Lincolnshire, and Sandwell and Walsall combined.

We also computed the proportion of LSOAs in the most deprived 10% of England in each ward (following recommendations for summary measures from the Indices of Deprivation 2019 statistics release) and compared this with aggregated changes in accessibility scores and places at ward level.^[footnote 15] We found significant negative correlations between this variable and changes in the childcare accessibility score in both case studies, with p-values of less than 0.01 both for Lincolnshire and for Sandwell and Walsall combined.

Example plots displaying these trends can be seen in Figures 9 and 10.

Figure 9: Scatter plots of IDACI data against changes in aggregated childcare accessibility in Lincolnshire



Figure 9 shows 2 scatter plots showing the relationship between deprivation and changes in childcare accessibility scores in Lincolnshire. The left chart shows that LSOAs with higher deprivation (lower IDACI rank) tend to have slightly larger declines in accessibility, indicated by a downward trend line. The right chart shows that wards with a higher percentage of LSOAs in the most deprived 10% of England tend to have greater increases in accessibility, indicated by an upward trend line.

Figure 10: Scatter plots of IDACI data against changes in aggregated childcare accessibility in Sandwell and Walsall



Figure 10 shows 2 scatter plots showing the relationship between deprivation and changes in childcare accessibility scores in Sandwell and Walsall combined. The left chart shows that LSOAs with higher

deprivation (lower IDACI rank) tend to have slightly larger declines in accessibility, indicated by a downward trend line. The right chart shows that wards with a higher percentage of LSOAs in the most deprived 10% of England tend to have greater increases in accessibility, indicated by an upward trend line.

Census 2021 data

We used the Census 2021 to gather data on the economic activity and education levels of women in households that have a dependent child aged 0 to 4 years. This data is only available at middle super output area (MSOA) level. Therefore, our comparisons were made against changes in childcare accessibility scores and childcare places aggregated to MSOA level.

At MSOA level we also tested correlations for household income.

In both case studies, we found a significant negative correlation between the percentage of women in households with a dependent child aged 0 to 4 who have level 4 or higher qualifications and change in childcare accessibility score. Both case studies found significance at a 0.01 threshold.

Example plots of comparisons between change in accessibility score and socioeconomic variables measured at MSOA level can be seen in Figures 11 and 12.

Figure 11: Scatter plots of socioeconomic variables against changes in aggregated childcare accessibility in Lincolnshire, by MSOA



Figure 11 shows 6 scatter plots showing the relationship between different socioeconomic variables and changes in childcare accessibility scores in Lincolnshire.

The top left chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who have level 4 or higher qualifications tend to have larger declines in accessibility, indicated by a downward trend line.

The top centre chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who have no qualifications tend to have greater increases in accessibility, indicated by an upward trend line.

The top right chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are economically inactive due to looking after home or family tend to have greater increases in accessibility, indicated by an upward trend line.

The bottom left chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are in employment tend to have slightly larger declines in accessibility, indicated by a

slight downward trend line.

The bottom centre chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are working full time tend to have slightly larger declines in accessibility, indicated by a slight downward trend line.

The bottom right chart shows that MSOAs with a higher household income tend to have larger declines in accessibility, indicated by a downward trend line.

Figure 12: Scatter plots of socioeconomic variables against changes in aggregated childcare accessibility in Sandwell and Walsall, by MSOA



Figure 12 shows 6 scatter plots showing the relationship between different socioeconomic variables and changes in childcare accessibility scores in Sandwell and Walsall combined.

The top left chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who have level 4 or higher qualifications tend to have larger declines in accessibility,

indicated by a downward trend line.

The top centre chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who have no qualifications don't have any change in accessibility, indicated by a straight trend line.

The top right chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are economically inactive due to looking after home or family tend to have greater increases in accessibility, indicated by an upward trend line.

The bottom left chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are in employment tend to have small declines in accessibility, indicated by a slight downward trend line.

The bottom centre chart shows that MSOAs with a higher percentage of women in households with a dependent child aged 0 to 4 who are working full time tend to have larger declines in accessibility, indicated by a downward trend line.

The bottom right chart shows that MSOAs with a higher household income tend to have larger declines in accessibility, indicated by a downward trend line.

Limitations of the two-step approach

Metaheuristic approaches are widely used for large-scale complex optimisation problems, but they come with inherent challenges. One major issue is that they cannot guarantee finding a globally optimal solution. Genetic algorithms and other metaheuristic approaches rely on searching over a large set of feasible solutions, meaning the solution depends on multiple factors including the initial population, selection choice, crossover and mutation rates, and termination criteria. In our model, the genetic algorithm starts with a randomly generated initial population of where provision could be located, meaning different runs can result in different solutions. Furthermore, genetic algorithms can become trapped in local optima and additional refinement strategies would be necessary to overcome this limitation.

The computational demands of the quadratic programming, the second step of our approach, are significant. Quadratic programming involves solving a constrained optimisation problem with a large number of variables, which grows significantly as the number of childcare provision and number of constraints increase. Scaling for an

entire region or nation would be extremely challenging due to the large increase in computation time. We applied our model to a large local authority with 24,000 places, while England has just under 1.3 million places. Using our high-powered machines, the quadratic programmer took upwards of 6 days to converge, with each iteration taking 420 seconds. Given these computational costs, we explored parallelisation and alternative optimisation methods. However, these methods did not result in significant improvements in efficiency or solution quality. Since our problem involves strict constraints and bounds, we required a sequential method which limited our attempts to parallelise.

In our previous analysis, we showed childcare accessibility at the OA level across the whole of England. In contrast, here we have produced only 2 case studies focused on specific areas, restricting the ability to compare between our analyses.

A limitation when comparing these 2 publications lies in the regional and national borders. Our first analysis only counted places from Ofsted-registered provision in England, although it was likely that some parents living on the borders between England and Wales and England and Scotland were crossing the border to use childcare in a different country from where they lived. However, in this analysis the issue is amplified by the salient borders being between neighbouring local authorities. Therefore, it is hard to estimate what the true potential capacity and demand for these services is. We have tried to demonstrate our confidence around the borders of our case studies to address the potential variations in accessibility across different local authorities. However, despite these efforts, caution should still be exercised when interpreting the results near the borders.

In addition, some of the limitations of this study are consistent with those outlined in our previous publication. Mainly, this analysis only considers childcare places in providers registered on Ofsted's Early Years Register (EYR). A limitation of using the EYR is that it records the number of places offered but does not indicate how many are occupied.

Furthermore, the EYR does not include:

- parents who are using informal childcare, such as care from grandparents
- places in some schools early years provision in about 10,000 schools (around 300,000 occupied childcare places) is exempt from registration with Ofsted and recorded differently, although provision for under-twos in schools is already captured in the EYR [footnote 16]
- a small number of providers of childcare on non-domestic premises who have been redacted for providers' personal safety and for national safety

Parents face additional barriers to accessing childcare beyond geographical proximity. Other factors, including affordability, availability, personal preferences and trust in providers, can influence a parent's or carer's decision on choice of childcare provision. In some areas, there might be a greater demand for informal childcare, such as provided by relatives. It would be good to investigate what take-up is like in these areas to help ensure that childcare provision aligns with the actual needs of families, rather than just their geographical accessibility.

Appendix

Full derivation of the objective function

This derivation is taken from a paper by Xiang Li, Fahui Wang and Hong Yi.^[footnote 1] The accessibility index A_i is given by:

$$\mathbb{P}_{\mathbb{P}} = \sum_{\mathbb{P}=1}^{\mathbb{P}} \mathbb{P}_{\mathbb{P}} \mathbb{P}_{\mathbb{P}}$$

Equation shows the accessibility index which is a weighted sum calculation of the capacity and normalised decay weight with a dummy variable to indicate whether a facility is selected or not. It reads: Ai equals sum from j = 1 to n of (Bj times Cj times Fij)

– where B_j is the location of the provision, C_j is the capacity of the provision, n is the number of candidate provisions, j is the index of the candidate provision from 1 to n and F_{ij} is the normalised decay weight and a Gaussian function given by:

$$F_{ij} = \begin{cases} e^{-\frac{1}{2}} \left(\frac{d_{ij}}{d_m}\right)^2 - e^{-\frac{1}{2}} \\ \frac{1 - e^{-\frac{1}{2}}}{1 - e^{-\frac{1}{2}}}, d_{ij} \le d_m \end{cases}$$

Equation for normalised decay weight using a Gaussian distance decay function.

It reads: Fij equals (e to the power of one-half times (dij divided by dm) squared minus e to the power of negative one-half) divided by (1 minus e to the power of negative one-half), provided that dij is less than or equal to dm

– where i is the demand node (OA population weighted centroid), j is the location of supply (childcare providers), d_{ij} is the travel time from demand node i to supply node j, and d_m is the size of the catchment area (varying size for public transport and driving).

The weighted average of the accessibility index must equal the ratio of total supply capacities to total demand (that is, the total number of users). Therefore, the total accessibility (A) is written as:

$$P = \sum_{p=1}^{P} P_{p} P_{p} = \sum_{p=1}^{P} P_{p} P_{p}$$

Equation shows the weighted average of the accessibility index.

It reads: A equals sum from i = 1 to m of (Ai times Di) is equal to sum from j = 1 to n of (Bj times Cj)

– where D_i is the demand of the location, m is the number of users, and i is the index of users from 1 to m.

Thus, the mean of accessibility (\bar{A}) is:

$$\bar{\mathbb{P}} = \frac{\mathbb{P}}{\sum_{\mathbb{P}=1}^{\mathbb{P}} \mathbb{P}_{\mathbb{P}}} = \frac{\sum_{\mathbb{P}=1}^{\mathbb{P}} \mathbb{P}_{\mathbb{P}}}{\sum_{\mathbb{P}=1}^{\mathbb{P}} \mathbb{P}_{\mathbb{P}}}$$

Equation shows the mean of accessibility.

It reads: A bar equals A divided by the sum from i = 1 to m of Di is equal to the sum from j = 1 to n of (Bj times Cj) divided by the sum from i = 1 to m of Di

And the standard deviation (Â), also known as the total variation of accessibility at each demand location, is:

$$\hat{A} = \sqrt{\frac{\sum_{i=1}^{m} (A_i - \bar{A})^2 D_i}{\sum_{i=1}^{m} D_i}}$$

Equation shows the standard deviation which captures the total deviation of accessibility at each demand location weighted by the amount of demand there.

It reads: A-hat equals the square root of the sum from i = 1 to m of (Ai minus A-bar) squared times Di, divided by the sum from i = 1 to m of Di

We are looking to minimise the total variation in accessibility. Therefore, we minimise subject to:

 $\sum_{j=1}^{n} B_j = p, \quad B_j = 0 \text{ or } 1 \ (B_j = 1 \text{ for any existing provision})$

 $\mathbb{P}_{\mathbb{P}} = 0$, if $\mathbb{P}_{\mathbb{P}} = 0$, and otherwise $\mathbb{P}_{\mathbb{P}} \le \mathbb{P}_{\mathbb{P}} \le \mathbb{P}_{\mathbb{P}} \le \mathbb{P}_{\mathbb{P}}$

$$\sum_{j=1}^{n} B_j C_j = \sum_{j=1}^{n} C_j = C$$
$$\sum_{i=1}^{m} D_i = D$$

Equation is four lines showing the bounds and constraints of the model.

Line 1 reads: The sum of j = 1 to n of Bj is equal to p, Bj equals 0 or 1, where Bj equals 1 for any existing provision

Line 2 reads: Cj equals 0 if Bj equals 0, and otherwise Cmin is less than or equal to Cj and Cj is less than or equal to Cmax

Line 3 reads: The sum of j =1 to n of (Bj times Cj) is equal to the sum of j = 1 to n of Cj is equal to C Line 4 reads: The sum of i =1 to m of Dj is equal to DA-hat equals the square root of the sum from i = 1 to m of (Ai minus A-bar) squared times Di, divided by the sum from i = 1 to m of Di

– where p is the number of provision selected from candidates, C_{min} and C_{max} are the lower and upper bounds of the provision's capacity, C is the sum of capacities of all provision and D is the sum of demands from all users.

In this planning scenario, the number and the sum of capacities of each provision is known from Ofsted's EYR. The range of each provision's capacity is given by the minimum and maximum capacity of the provision from the EYR. The number of users and demands are also known, and finally so are the locations of the provision and users. Therefore, the objective function can be written as:

$$\operatorname{Minimise}(\hat{A}) = \operatorname{Minimise} \sum_{i=1}^{m} (A_i - \bar{A})^2 D_i = \operatorname{Minimise} \sum_{i=1}^{m} \left(\sum_{j=1}^{n} B_j C_j F_{ij} - \bar{A} \right)^2 D_i \qquad (1)$$

Equation shows an objective function where the objective is to minimise the total variation of accessibility at each demand location.

It reads: Minimise A-hat, which equals the sum from i = 1 to m of (Ai minus A-bar) squared times Di, which is equivalent to minimising the sum from i = 1 to m of the square of (the sum from j = 1 to n of Bj times Cj times Fij, minus A-bar), all multiplied by Di

- X Li, F Wang and H Yi, '<u>A two-step approach to planning new facilities towards</u> <u>equal accessibility</u>', in 'Environment and Planning B,' volume 44(6), 2016, pages 994 to 1011. <u>←</u> <u>←</u>²
- <u>Childcare accessibility by neighbourhood</u>', Office for National Statistics, June 2024;

'<u>Childcare accessibility by neighbourhood, England: Ofsted's methodology</u>', Ofsted, June 2024. $\underline{\leftarrow} \underline{\leftarrow}^{\underline{2}}$

- 3. '<u>Metaheuristic algorithm an overview</u>,' ScienceDirect Topics, 2023 *←*
- 4. '<u>What is the genetic algorithm</u>?', The MathWorks, Inc. *←*
- 5. A Brital, '<u>Genetic algorithm explained: everything you need to know about genetic</u> <u>algorithm</u>', Medium, October 2021. <u>←</u>
- 6. 'Statistical geographies', Office for National Statistics, 2021. ←
- '<u>Childcare accessibility by neighbourhood, England: Ofsted's methodology</u>', June 2024. <u>←</u>
- D Goldberg and K Deb, '<u>A Comparative Analysis of Selection Schemes Used in</u> <u>Genetic Algorithms</u>', in 'Foundations of Genetic Algorithms', Volume 1, 1991, pages 69 to 93. <u>←</u>
- P Boggs and J Tolle, <u>'Sequential quadratic programming</u>', Acta Numerica, Volume 4(4), 1996, pages 1 to 51. <u>←</u>
- 10. '<u>Sensitivity analysis an overview</u>', ScienceDirect topics, 2005. <u>←</u>
- 11. O Roeva, S Fidanova and M Paprzycki 'Influence of the population size on the

genetic algorithm performance in case of cultivation process modelling', in 'Proceedings of the 2013 Federated Conference on Computer Science and Information Systems', January, 2013, pages 371 to 376;

A Hassanat, K Almohammadi, E Alkafaween, E Abunawas, A Hammouri and V B Surya Prasath, '<u>Choosing mutation and crossover ratios for genetic algorithms – a</u> review with a new dynamic approach', in 'Information', Volume 10(12), 2019, page 390. <u>←</u>

- 12. P Hurley, M Tham and H Nguyen, '<u>Childcare accessibility in the United Kingdom</u>', Victoria University, September 2024. <u>←</u>
- 13. '<u>Childcare providers and inspections: management information</u>', Ofsted, February 2022 (updated January 2025). <u>←</u>
- 14. 'English indices of deprivation 2019 LSOA Level', Ministry of Housing, Communities and Local Government, September 2019. <u>←</u>
- 15. '<u>The English indices of deprivation 2019</u>', Ministry of Housing, Communities and Local Government, September 2019. <u>←</u>
- 16. For more detailed statistics on early years provision, please refer to: '<u>Provision for</u> <u>children under 5 in England</u>', Department for Education, July 2024. <u>←</u>
 - ↑ <u>Back to top</u>

Help us improve GOV.UK

To help us improve GOV.UK, we'd like to know more about your visit today. <u>Please fill in this survey (opens in a new tab)</u>.

Services and information

Benefits

Births, death, marriages and care

Government activity

<u>Departments</u>

Business and self-employed News Childcare and parenting Guidance and regulation Citizenship and living in the UK **Research and statistics** Crime, justice and the law Policy papers and consultations Disabled people **Transparency** Driving and transport How government works Education and learning Get involved Employing people Environment and countryside Housing and local services Money and tax Passports, travel and living abroad Visas and immigration Working, jobs and pensions

HelpPrivacyCookiesAccessibility statementContactTerms and conditionsRhestr o Wasanaethau CymraegGovernment Digital Service

OGL

All content is available under the <u>Open Government Licence v3.0</u>, except where otherwise stated



© Crown copyright