



Subregional Pupil Projections

Population Statistics User Group

23rd July 2024

Dónal Ring



Presentation outline

1. Introduction and aims

- Existing outputs and what's missing
- Outline of planned output

2. Methodology

- Methodology for projections
- Methodology for uncertainty estimates

3. Results

- Projections
- Accuracy testing

4. Summary and next steps



Section 1 – Introduction and aims

1. Introduction – (previously) existing outputs

- DfE's National Pupil Projections
- GLA's School Roll Projection Service
- GLA's ward-level Projected Demand for School Places (discontinued)



Department
for Education



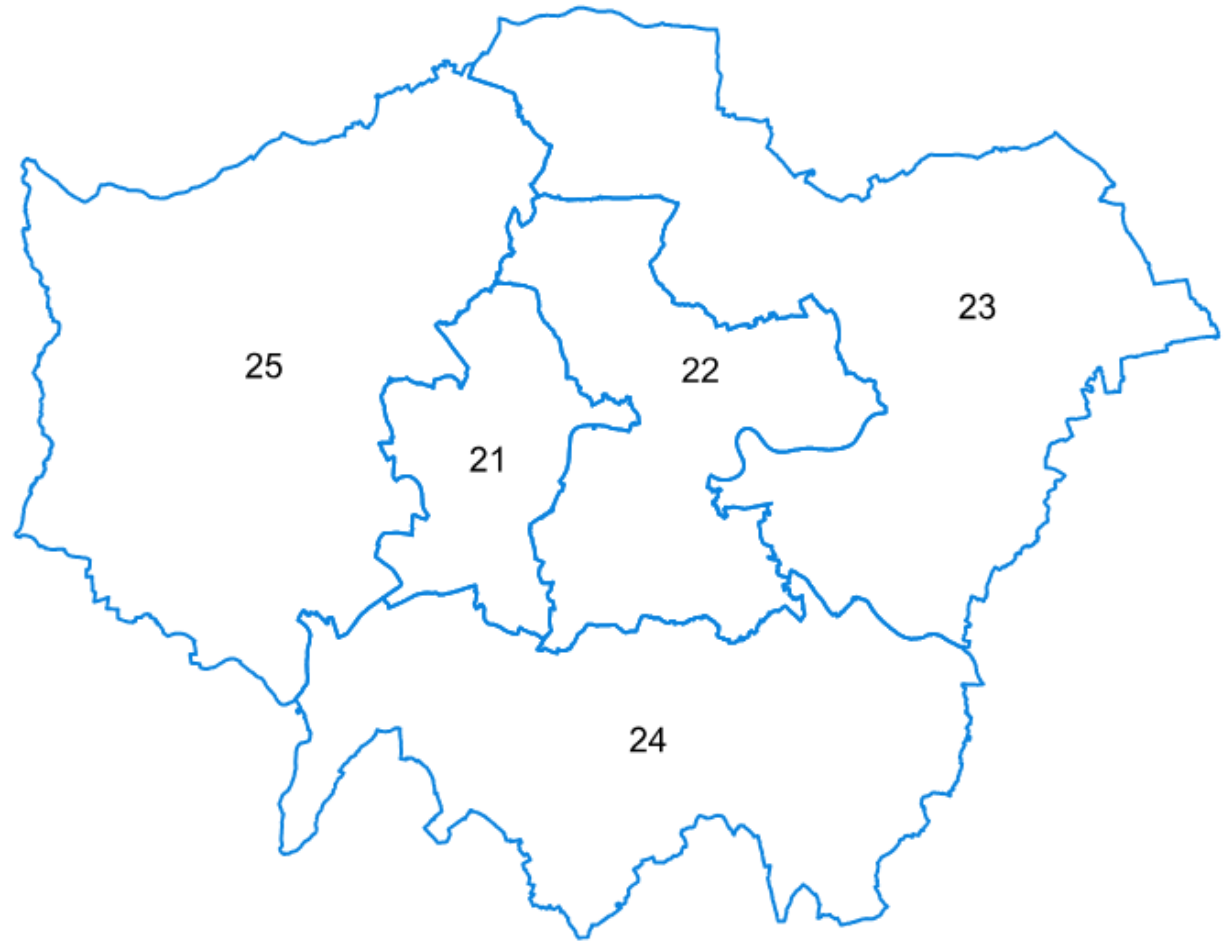


1. Introduction – what's missing

- DfE's projections don't go below national level
- When modelling at lower geographies, time series are too messy and with too many distortions for statistical modelling
- No prediction/uncertainty intervals
- School Rolls Projection Service can't be made transparent and publicly available
- Reliance on population projections for GLA's output causes difficulties with scheduling

1. Introduction – what we aim to produce

- 10-year projections
- Reception to year 11
- ITL2 geography (all in England, not only London)
- Only inputs are births and pupil numbers
- Released with 95% prediction intervals





Section 2 – Methodology for producing new projections

2. Methodology – projections

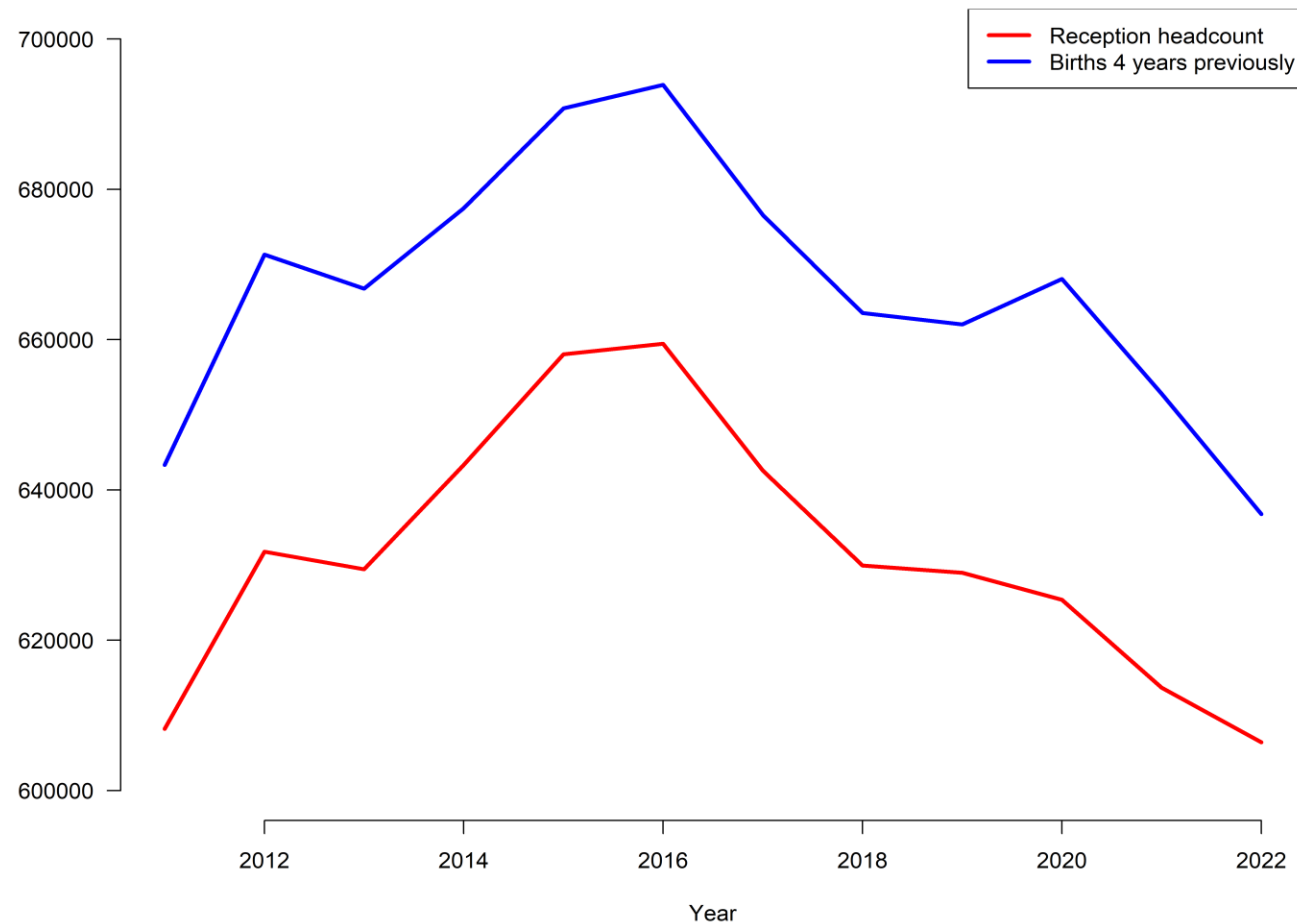
- Reception and year 1 predicted from births 4 and 5 years previously, respectively

$\text{predicted_reception} = \text{births_4_years} * \text{ratio_births_to_reception}$

$\text{predicted_year_one} = \text{births_5_years} * \text{ratio_births_to_year_one}$

- Each forecasted using exponential smoothing

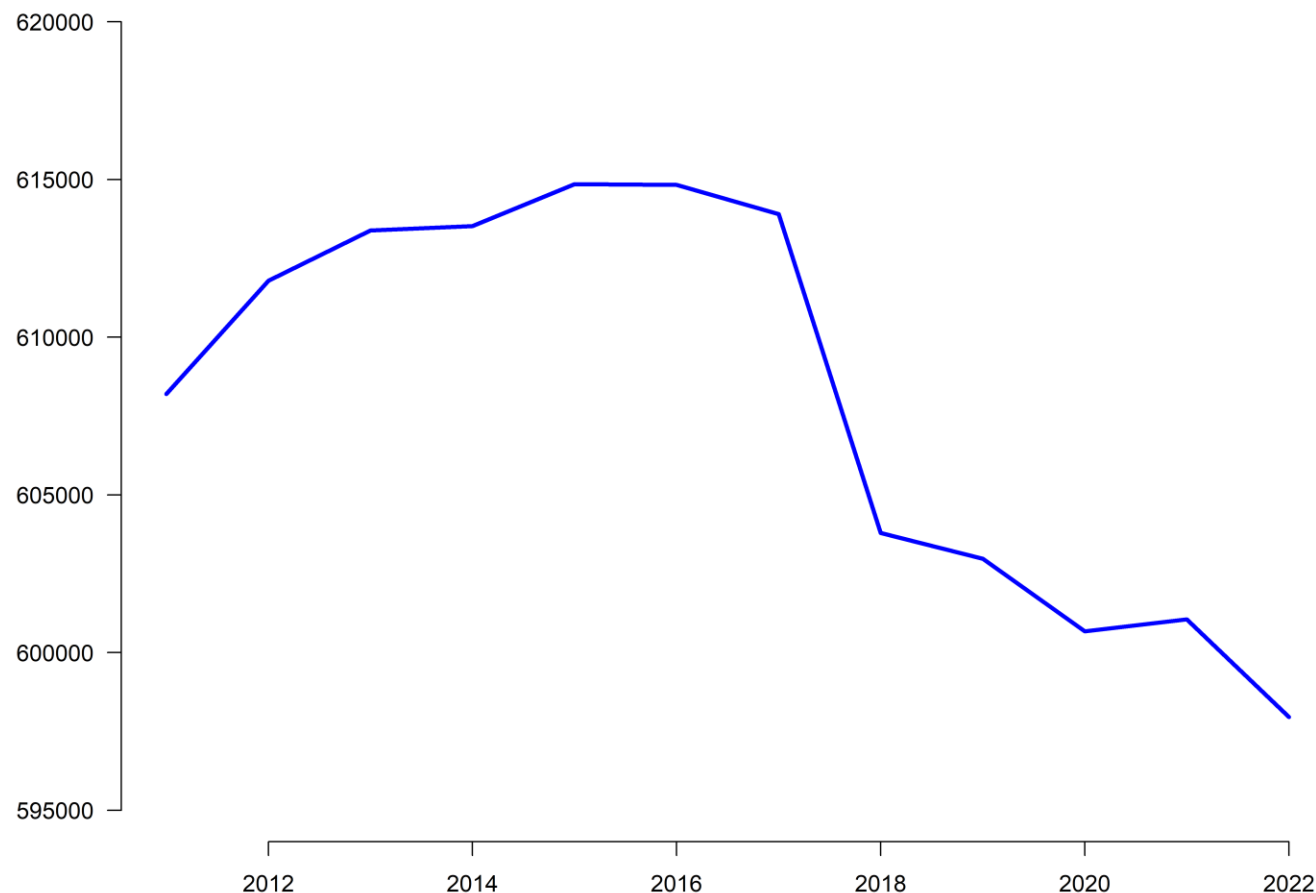
Comparison of UK-wide reception with births 4 years previously



2. Methodology – projections

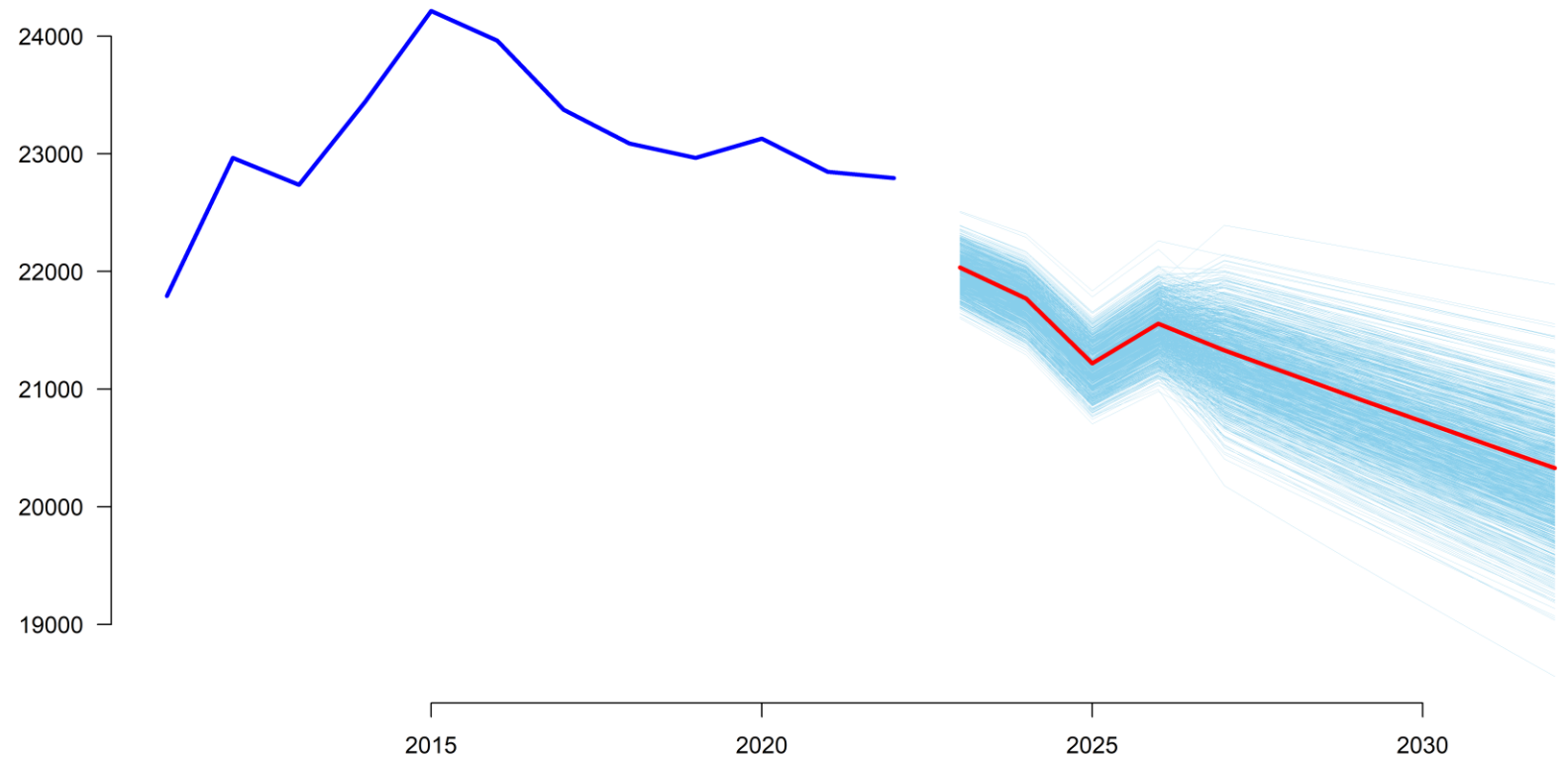
- Years 2 onwards predicted from proportion continuing from year before
- E.g. 1.02 from year 1 to year 2, 1.01 from year 2 to year 3, etc...

Cohort progression from Reception-2011 to Year Eleven-2022



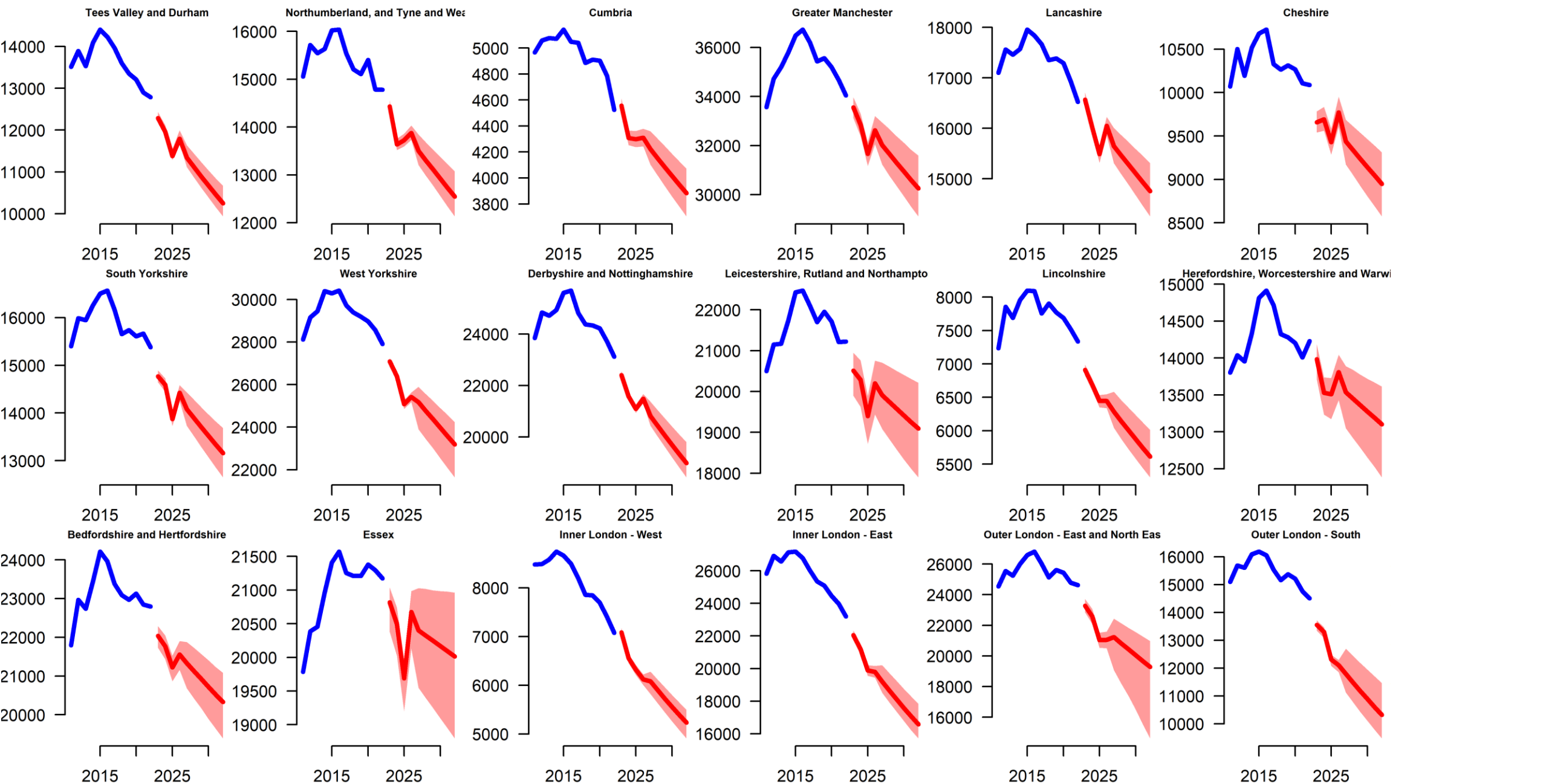
2. Methodology - uncertainty

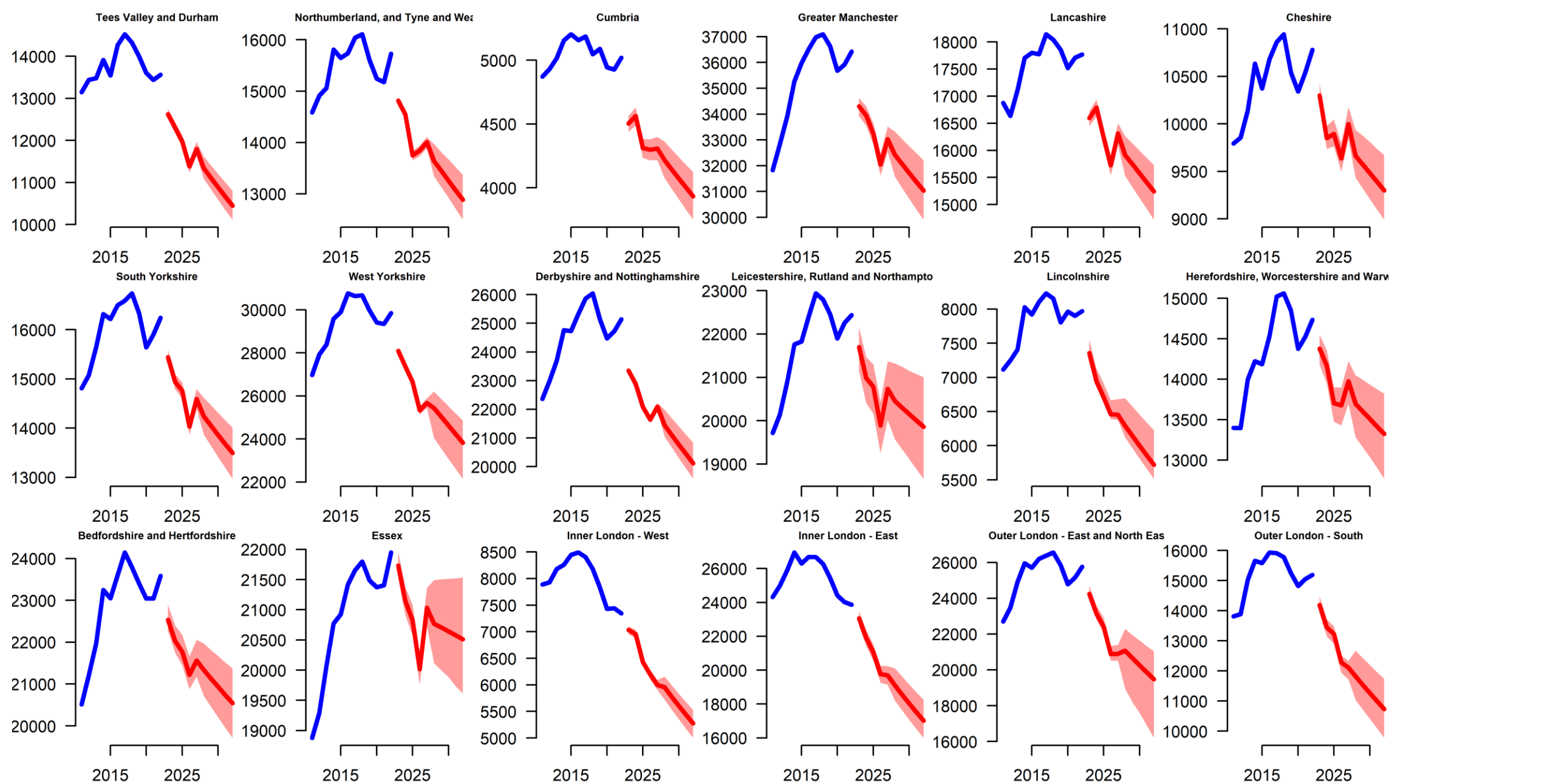
- Difficulty in measuring uncertainty when combining forecasts
- Bootstrapping used
- Simulate the process 1,000 times, based on resamples of the original series
- Choose the 2.5th and 97.5th percentiles (for 95 percent prediction intervals)

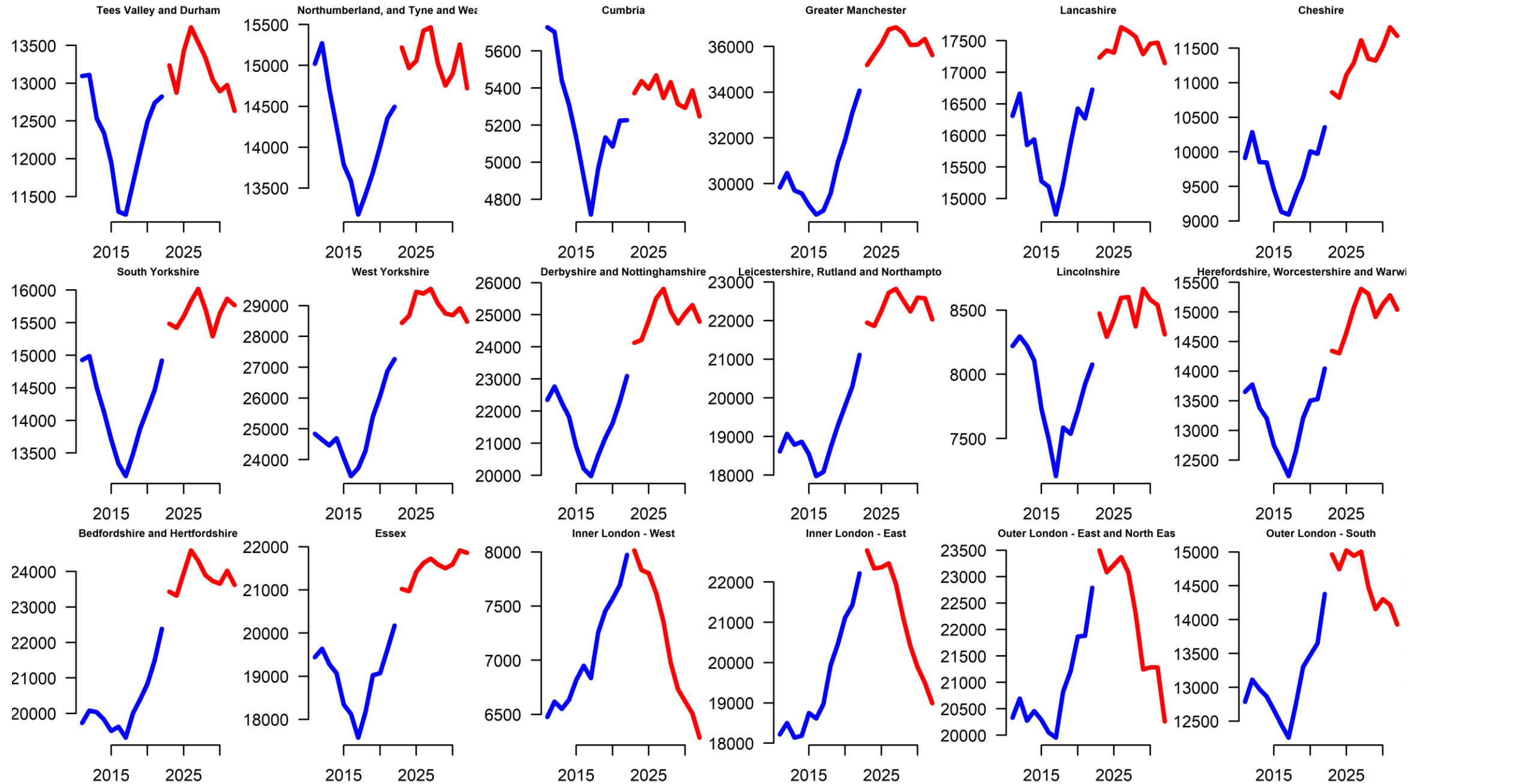




Section 3 – Results









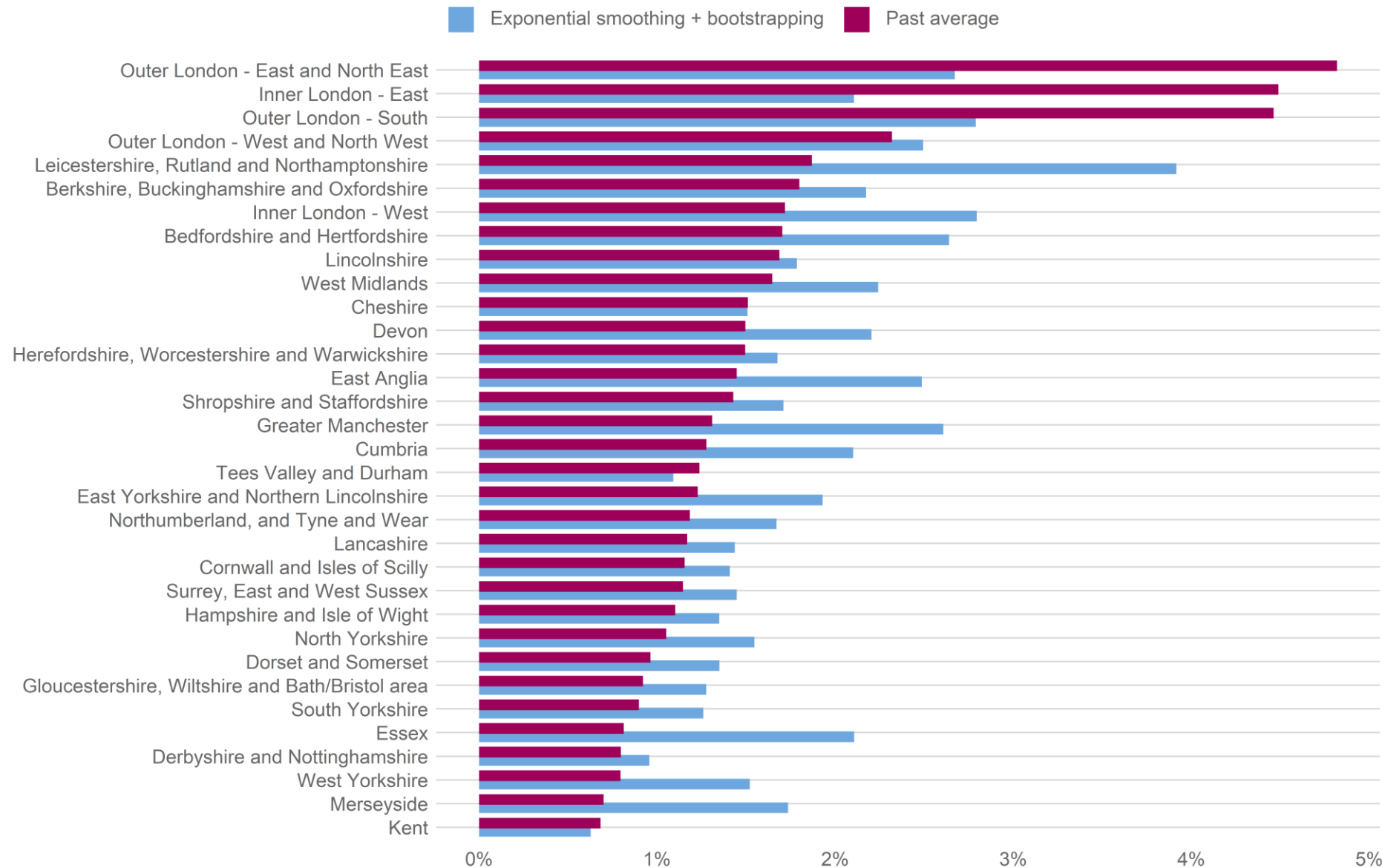
3. Results - accuracy

- Tested predictions against real data to measure accuracy
- Have only tested Reception so far
- Compared against an alternative method of prediction and quantifying uncertainty – using past average of ratios, based on 1, 3, and 5 years of past data. Births series remain the same
- Metrics used were (1) mean average percentage error (MAPE) and (2) percentage of real points falling within the prediction intervals

3. Results accuracy

- Total MAPE of 1.9% for exponential smoothing, and 1.6% for past average
- Both values quite low
- Suggests method works and is reasonably accurate
- Less accurate for London ITLs

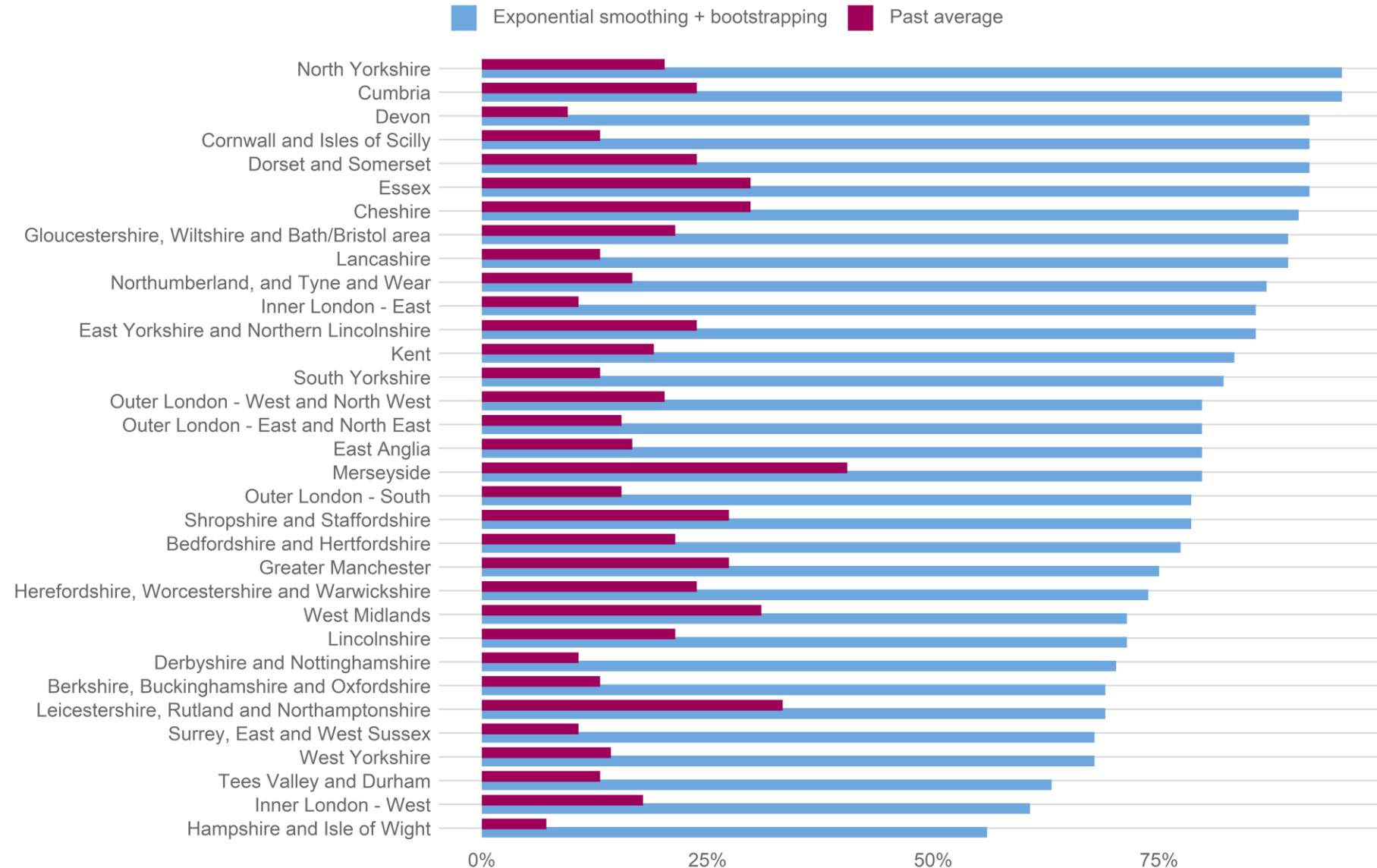
– Mean average percentage error, by ITL2 and modelling method



3. Results - accuracy

- Total percentage within intervals of 79% for bootstrapped estimates, compared to 20% for past average method
- 95% prediction intervals used for bootstrapping. 79% is a reasonably good performance
- 20% quite low – most of the time, real values will be outside of the range presented

Percentage within prediction intervals, by ITL2 and modelling method





Section 4 – Summary and next steps

4. Summary

- Full (provisional) range of school projections now produced, from reception to year 11
- Simple, publicly available input data sources
- ITL2 geography and with uncertainty estimates attached
- Projections seem accurate and uncertainty seems reasonable



4. Next steps

- Further accuracy testing
- User engagement and methodology review
- Uncertainty estimates for Year 2 to Year 11
- Publish on GitHub
- Aim for publication late summer



Contact

donal.ring@london.gov.uk

demography@london.gov.uk