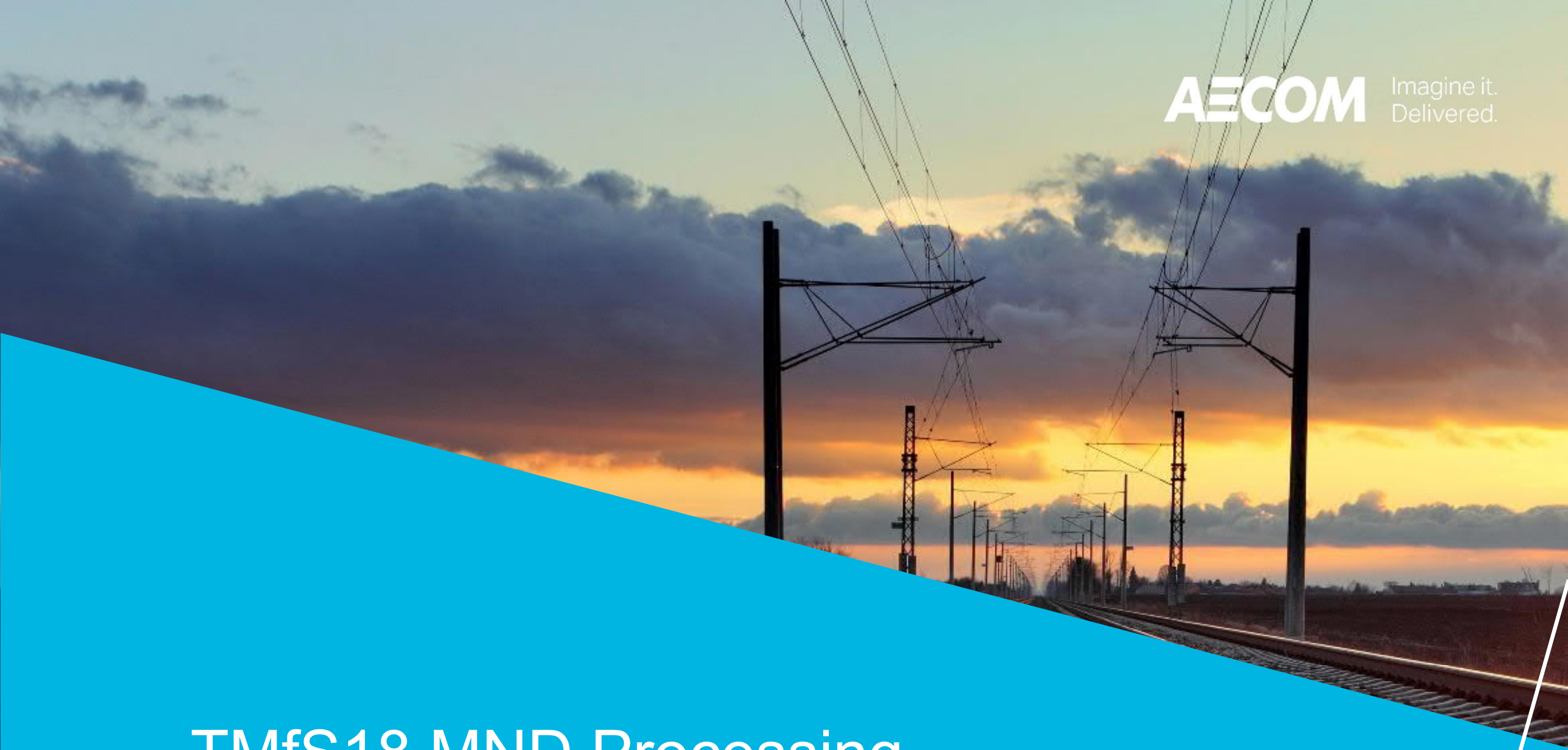



# TMfS18 MND Processing Technical Note

21 January 2020

## Quality information

| Prepared by | Checked by | Verified by | Approved by |
|---|---|---|---|
| Hugo Nilsson | Stefanos Psarras | Reza Tolouei | Richard Cann |
| Senior Consultant | Senior Consultant | Associate Director | Regional Director |

## Revision History

| Revision | Revision date | Details | Authorized | Name | Position |
|---|---|---|---|---|---|
| v1 | 21/01/2020 | | | | |
| V2 | 19/02/2020 | Map update to include Shetland Islands | INTERNAL | | |
| V3 | 6/8/2020 | Explanation of Area Classification | RJC | Richard Cann | Regional Director |
| | | | | | |

## Distribution List

| # Hard Copies | PDF Required | Association / Company Name |
|---|---|---|
| | | |
| | | |
| | | |
| | | |

Prepared for:

Transport Scotland

Prepared by:

Hugo Nilsson

Senior Consultant

E: hugo.nilsson1@aecom.com

T: 07747 214228

AECOM Limited
AECOM House
63-77 Victoria Street
St Albans
Hertfordshire
AL1 3ER

aecom.com

# Table of Contents



# List of Tables



# List of Figures

## List of Abbreviations

| Abbreviation | Description |
|---|---|
| CSRGT | Continuing Survey of Road Goods Transport |
| DfT | Department of Transport |
| DVLA | Driver and Vehicle Licensing Agency |
| HB | Home based (total of HBW & HBO) |
| HBO | Home based other |
| HBW | Home based work |
| HGV | Heavy goods vehicle |
| LGV | Large goods vehicle |
| MND | Mobile Network Data |
| MSOA | Middle Super Output Level |
| NHB | None home based |
| NTEM | National Trip End Model |
| NTS | National Travel Survey |
| NUTS | Nomenclature of Territorial Units for Statistics |
| OD | Origin Destination |
| PA | Production Attraction |
| POI | Point of interest score |
| RSI | Road side interview |
| SHS | Scottish Household Survey |
| TAG | Transport Analysis Guidance |
| TCD | Trip cost distribution |
| TLD | Trip length distribution |

# Section 1 – Introduction

## 1.1 Background

1.1.1 Transport Scotland have decided to purchase Mobile Network Data (MND) from Telefonica. The data is drawn from two months of MND records during 2018 and is provided in the form of Origin Destination trip matrices.

1.1.2 The timescales required by Transport Model for Scotland (TMfS18) precludes the use of MND in the demand matrix development, the MND is intended to be used for:

- Matrix validation in TMfS18, and hence in understanding the strengths and weaknesses of the demand representation;
- Matrix validation for the LATIS regional models, particularly the Aberdeen Sub Area Model, and the Tay Cities Regional Transport Model.

1.1.3 The matrices provided by Telefonica are subject to a number of biases and gaps that should be addressed before they can be used for the above purposes. In principle, this is through combining the MND with several other complimentary data sources.

1.1.4 Initial processing (referred to as 'core') of the MND will involve a review of the matrices and specification of a revised sectoring system for Telefonica, then establishing aggregate trip patterns. This will include processing secondary data sources, detailed verification, the development of synthetic matrices, vehicle split, and infilling of short trips. This is the current scope of this work commissioned by Transport Scotland, the result of which is presented in this report.

1.1.5 The output of this work is matrices of car demand prepared at the defined sectoring system, representing aggregate trip patterns. It is important to note these matrices will require significant further processing and refinements to be used as highway prior matrices for TMfS18. These processing steps, currently excluded from the scope of this work, include:

- Spatial disaggregation of demand into TMfS18 zones;
- Detailed segmentation / purpose split (i.e. home-based / non-home-based employers' business, other, etc.);
- Matrix merging and data fusion, using secondary data sources, to improve quality of matrices; and
- Detailed and iterative validation and refinement process to develop highway prior matrices for use in TMfS18 highway model.

## 1.2 Structure of the Report

1.2.1 The remainder of this report is structured as follows:

- Section 2 provides an overview of the overall MND processing and introduces the complimentary data sources used to inform this process;
- Section 3 sets out the process of developing synthetic car matrices, used as one of the key inputs in matrix development process;
- Section 4 describes the steps followed to process MND matrices and address their limitations;
- Section 5 sets out the verification tests undertaken to assess the suitability of MND matrices; and
- Section 6 provides a brief summary of the results and recommendations on suitability of the produced matrices for the intended use.

# Section 2 – Overall Approach to Matrix Development

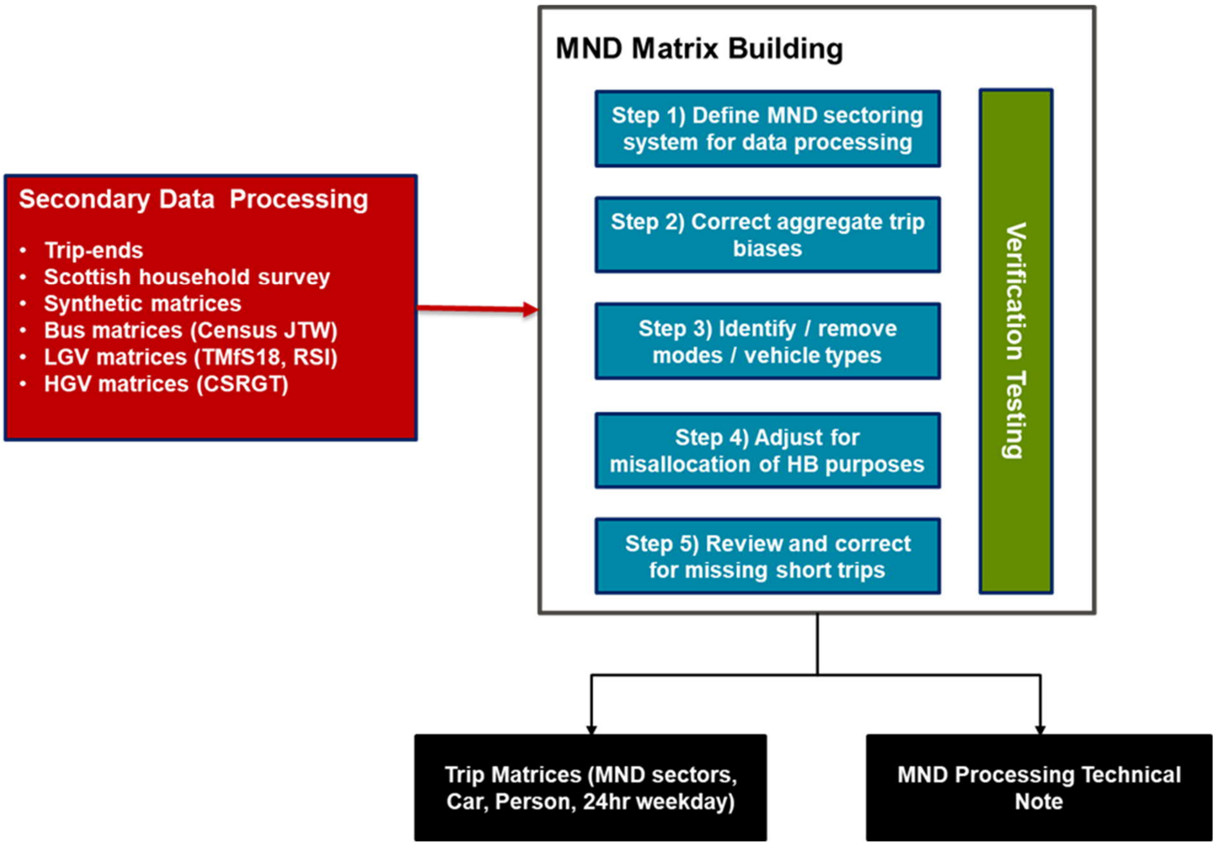
## 2.1 Methodology Outline

2.1.1 In principle, MND provides specific benefits which are typically lacking in alternative traditional matrix data sources; these include larger sample size, repeated sample over several says (day-to-day variability), and wide geographical coverage. The key limitations and challenges in the MND data include:

- definition of a single trip;
- identification of mode and vehicle type;
- identification of detailed trip purposes;
- detection of short trips;
- spatial resolution;
- expansion of the MND sample; and
- potential errors due to rounding or treatment of small values due to privacy reasons.

2.1.2 A step-wise methodology is defined to address the limitations of the MND matrices provided by the Mobile Network Operator (i.e. Telefonica), in order to develop aggregate matrices of travel patterns in Scotland; this is shown in Figure 2.1. It consists of five main steps with the processed secondary data feeding as inputs during the whole processes. Various forms of verification testing are carried out continually.

**Figure 2.1: Methodology Outline**




2.1.3 The following sets out why each step is required and what the main purpose of each step is. A more detailed description of each step is provided in Section 4. The final output of the matrix development process is the Car-Person 24-hour average weekday trip matrices at defined MND sector level.

2.1.4 The spatial accuracy of MND is limited; the first step therefore investigates if MND matrices at the zoning level where the data is provided include errors in allocation of trips to zones; this is expected to be the case. A sectoring system is defined through analysis of MND trip ends and comparisons with complimentary data sources and is used to re-expand and aggregate the data spatially.

2.1.5 In addition to a bias resulting from under-detection of short trips, the initial verification tests suggested that the MND matrices also suffer from a global bias likely to be resulting from inaccuracies in the expansion process carried out by the MND data provider. This bias causes an overall under-reporting of trips across all trip distances. To address this issue, global correction factors need to be applied.

2.1.6 One limitation of the MND processing for matrix development is the challenge of accurately detecting modes. Certain modes, such as train and ferry, will have very distinct patterns enabling their mode to be determined with more certainty. However, freight and bus modes cannot be easily separated from car through the processes used by the MND data provider. Therefore, these modes are combined with car to form an overall “Road” mode. The purpose of the third step is therefore to apply alternative methods for the identification and removal of bus and freight trips to form a residual car matrix.

2.1.7 A further limitation of the MND is the potential misclassification of trip purposes. While home-based and non-home-based trips can be identified based on inferred home locations, there is a potential risk of misclassification between home-based purposes (HBW and HBO). For this reason, an adjustment step has been included to ensure a more realistic proportion of HBW and HBO trips.

2.1.8 The main limitation of using MND data is perhaps its inability to fully detect short distance trips. This is due to a trip event only being recorded when signals from a mobile phone is detected by two or more masts. Hence, movements within the catchment area of a single mast will not be recorded. To address this issue, short distance trips are replaced with synthetic data.

## 2.2 Sectoring Systems

2.2.1 Throughout MND matrix development process, analysis and processing has been carried out at the TMfS model zone system or the defined MND sectoring system (see Section 4.2) whenever possible. However, this is not always feasible or desirable and therefore aggregate versions of these systems have been applied instead. The aggregate sectoring systems used during the development process are:

- Local Authorities (32 sectors within Scotland)
- Area Classification (Metropolitan, Urban and Rural)
- Level 3 NUTS Regions (25 sectors, of which 22 are in Scotland)
- TMfS Regions (12 sectors within Scotland)

2.2.2 The Area Classification is based on Scottish Government Urban Rural Classification 2016 (6 fold), summarised by 3 sectors:

1. Metro (populations of 125,000 or more) – UR6Fold (1)
2. Urban (populations of 10,000 to 124,999) – UR6Fold (2)
3. Rural (populations less than 10,000) – UR6Fold (3-6)

2.2.3 The classification is based on two criteria: population defined by National Records of Scotland, and accessibility based on drive time analysis to differentiate between accessible and remote areas in Scotland.

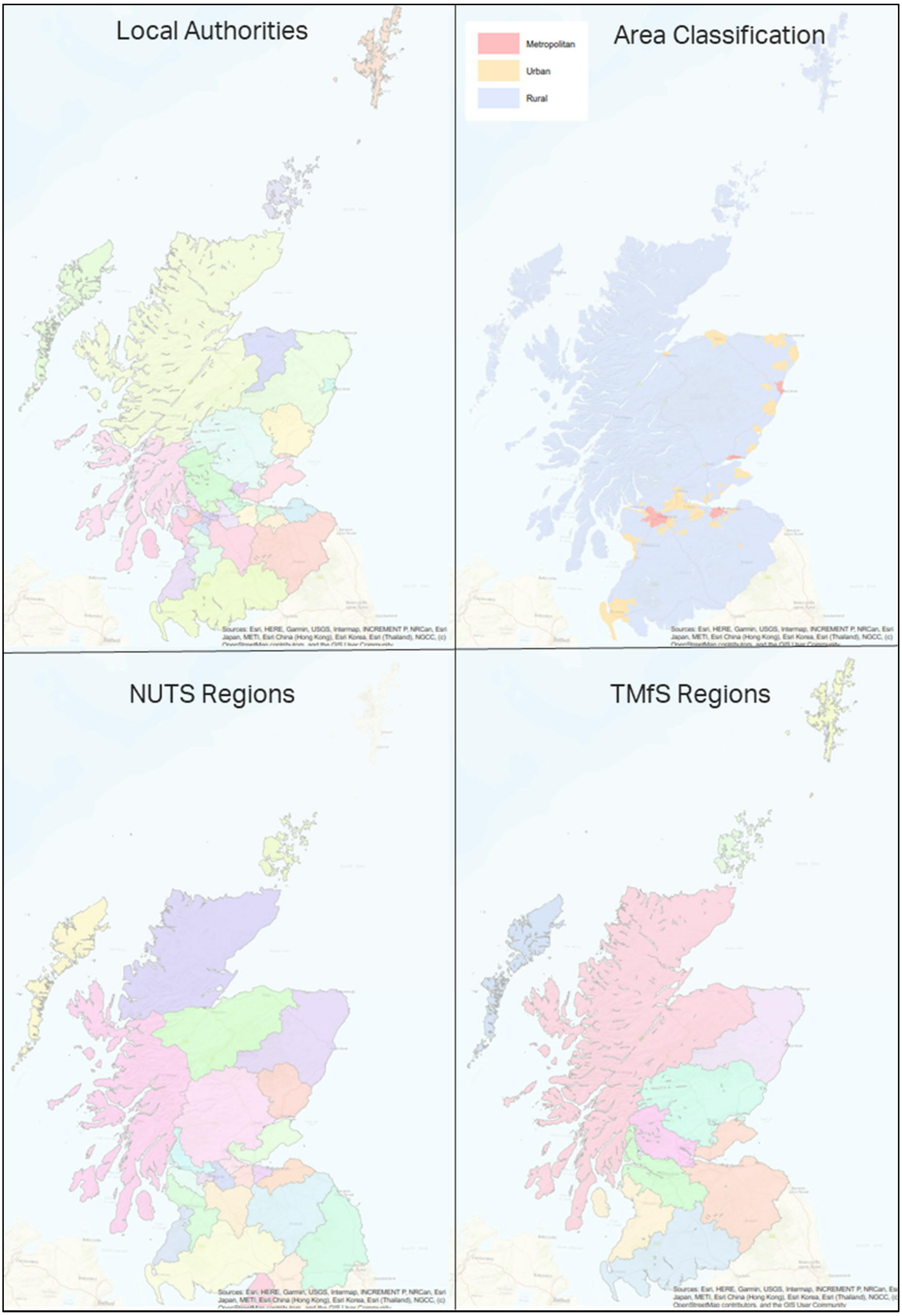
2.2.4 Where TMfS18 zones covered more than one sector the zone was defined as:

- Metro if it was part metro/rural; or
- Urban if it was part urban/rural.

2.2.5 With a few exceptions where the rural area was much larger than the metro/urban and the town already had one or more zones defined urban to represent it. Following this approach, the whole TMfS18 zone around Stranraer is urban because the town is classified as UR6Fold (2).

2.2.6 A visual representation of the sectoring systems is displayed in Figure 2.2.

**Figure 2.2: Sectoring Systems**




## 2.3 Processing Secondary Data

2.3.1 The MND development process and the synthetic matrix generation require a set of secondary data as inputs. The sources for these datasets include:

- 2011 Census;
- The Continuing Survey of Road Goods Transport (CSRGT);
- The National Trip End Model (NTEM);
- The National Travel Survey (NTS);
- Planning data;
- Road side interviews (RSI);
- The Scottish Household Survey (SHS);
- TMfS Trip End Model; and
- TMfS18TMfS Matrices

***2011 Census***

2.3.2 Travel to work data recorded during the 2011 census served as basis for approximating the distribution of bus and coach trips across Scotland.

2.3.3 This data records place of usual residence and place of work and thus approximates a travel pattern at PA-level. The travel to work data was gathered by local authority, of which there are 32 in Scotland (see Figure 2.2). Overall trip totals gathered from NTEM were utilised to estimate the overall trip totals made by purpose (HBW, HBO and NHB).

2.3.4 The resulting output was a 24-hour bus and coach matrix for each purpose at local authority level in PA-format for use in the mode identification and removal of bus trips in the MND development process (see Section 4.2).

***CSRGT***

2.3.5 The Continuing Survey of Road Goods Transport (CSRGT) surveys the UK activity of Great Britain registered heavy goods vehicles in a trip diary format. The level of granularity of the CSRGT data corresponds to level 3 NUTS regions developed for statistical purposes by the European Union. There are 22 level 3 NUTS regions in Scotland (see Figure 2.2).

2.3.6 A dataset for the year of 2017 was obtained and trips with an origin and destination outside Scotland had been removed from the analysis, forming the final sample size of 7106 trips. Each record has an associated grossing factor used to expand to population figures using population data for HGVs, for each quarter, from the DVLA licensing records.

2.3.7 This results in an annual matrix which was factored down by a factor 0.003375 to reflect an average weekday. This factor was derived using national combined automatic traffic count data from the DfT. An additional factor of 1.3 was applied to account for survey under-reporting as recommended by DfT guidance DfT, (2020b).

***NTEM***

2.3.8 Since the TMfS trip end model only estimates home-based trip ends, the NTEM was used to estimate non-home-based trip ends for Scotland. NTEM trip end data was gathered at local authority level. This data was disaggregated into TMfS zones using total attractions from the TMfS trip end model as weights.

2.3.9 NTEM was further utilised to estimate trip ends for external zones. Trip end data was gathered at MSOA-level and aggregated into TMfS external zones. Lastly, NTEM also formed as the basis to estimate the total number of trip ends for bus and coach across Scotland by purpose (HBW, HBO and NHB), to be used to factor the census-based bus matrices.

***NTS***

2.3.10 The National Travel Survey (NTS) was utilised to extract trip length distributions to be used for calibration of external trips during the synthetic matrix development.

***Planning data***

2.3.11 The planning data, which is collected from local authorities to support the TMfS trip end model, was used to extract population estimates. These estimates form as inputs to calculate the resulting trip rates for synthetic and MND matrices at the various aggregate levels during their verification processes.

***RSI***

2.3.12 Road side interviews (RSI) were carried out across 46 locations on the Scotland Road network and various survey dates, all occurring between Tuesday and Thursday, between the years 2015 and 2018. RSI data was used to inform analysis of LGV travel, and following removal of uncomplete records a dataset of 1645 LGV trips remained.

2.3.13 The process of identifying and removing LGV trips from the MND data has been carried out on a home-based and non-home-based basis. For this reason, proportions of home-based and non-home-based trips were derived from the RSI data.

2.3.14 The TMfS prior LGV matrix was used as basis for identifying and removing LGV trips from the MND matrices. However, due to quality concerns of the TMfS LGV matrix as evidenced by its unrealistic trip length distribution profile, insights from the RSI data were used to inform the adjustment of the TMfS LGV matrix.

2.3.15 Using a defined set of appropriate distance bins, a set of correction factors was calculated to align the trip length distribution of the TMfS matrix with the RSI data. In order to obtain a smooth set of correction factors, the initially calculated correction factors were fitted to an exponential function against the bin numbers.

***SHS***

2.3.16 The Scottish Household Survey (SHS) serves as one of the main data sources during the MND development process. The main SHS outputs are trip rates used as verification data during the MND development process, and trip length distributions for the generation of synthetic matrices.

2.3.17 Travel diary data during the period between the years 2012 and 2017 was obtained containing 57170 trip records made during neutral weekdays. Each record contained a weighting factor to account for bias associated with individual and trip sampling probability.

2.3.18 The TMfS origin and destination zone was identified for each trip record and based on these the corresponding distance appended from the TMfS distance matrix. Although the distance was recorded during the survey, the use of the TMfS model distance is conceptually more correct to use during the generation of the synthetic matrices as the distance matrix and trip length distribution should stem from the same data. It should also be mentioned that the intrazonal distances are set to half the distance to its nearest neighbouring zone.

2.3.19 Applying both individual and trip related weights, trip rates were calculated segmented by purpose (HBW, HBO and NHB), time period (AM, IP, PM and OP) and TMfS region. In order to obtain accurate estimates, samples from respondents indicating no travel during the surveyed date was included for the estimation. Using the weighted records within the internal model area, trip length distributions were extracted, segmented by purpose (HBW, HBO and NHB), area classification (metropolitan, urban and rural, see Figure 2.2), and time period (AM, IP, PM and OP). These were subsequently used for the generation of the synthetic matrices.

***Trip End Model***

2.3.20 The TMfS trip end model is one of the sub-components of the overall Transport Model for Scotland. It uses the most recently available planning data along with a set of trip rates to generate trip ends. For further details of the trip end model, refer to the model documentation.

2.3.21 The output from the trip end model consists of productions in the form of work-weekly tours and attractions in the form of weights. Both the productions and attractions are segmented into four home-based purposes (commute, employer’s business, education and other).

2.3.22 Productions were converted into trips by applying a factor of two on the assumption that a tour would on average consists of two trips, an additional factor of 0.2 to convert to from work-weekly to daily and finally aggregated into the HBW and HBO purposes. Similarly, the attractions were aggregated into HBW and HBO purposes and then constrained to production totals.

2.3.23 As greater trust has been placed on the SHS in terms of representing spatial variation in trip rates, the trip ends were adjusted to SHS trip rates by area types (i.e. urban, metropolitan, and rural). This ensured reliability in overall trip end totals while the spatially varying trip rates stemming from the detailed planning data within the trip end model being retained. The processed trip ends then served as inputs to the generation of the synthetic matrices.

### *TMfS18 Matrices*

2.3.24 The 2018 version of the TMfS LGV matrix served as the starting step for the identification of LGV trips within the MND matrices. Following aggregation into the MND sectoring system and before the vehicle removal step, RSI data were used to adjust the trip length distribution and to identify home-based and non-home-based trip proportions.

2.3.25 The TMfS distance skim was selected as the distance matrix during both the synthetic matrix development and the MND development process. Intrazonal distances were set to half the distance to its nearest neighbour. Before being used as input to the MND development process, the distance matrix was aggregated into the MND sectoring system using the TMfS18prior-estimation demand matrix as weight.

# Section 3 – Synthetic Matrix Development

## 3.1 Background

As noted in forthcoming TAG Unit on matrix development (DfT, 2020a), synthetic matrices are defined as estimated travel demand matrices based on some estimated statistical models, observed aggregate travel patterns and land-use information. The synthetic matrices are an integral part of the MND matrix development process, since they are used to fill in gaps and address limitations of the data.

3.1.1 Their importance is even more apparent when considering forecasting using strategic models. One of their key requirements is the consistency between trip ends and detailed land use and demographic data, and synthetic matrices often result in this desirable property, therefore this property should be introduced to the overall matrix development process.

3.1.2 However, for the purpose of the current study, synthetic matrices are being investigated, used and assessed for their suitability to represent and eventually infill the short distance trips, which are missing from the MND matrices.

3.1.3 The most common approach to develop synthetic matrices is based on gravity modelling techniques using deterrence functions. This approach has been introduced and described in various research papers (for example, see Feldman et al., 2012) and across many matrix development projects including the Highways England Regional Traffic Models.

However, the synthetic matrices include uncertainty and errors due to differences in local travel patterns, which have not been accounted for by the explanatory variables in the calibrated gravity models (DfT, 2020a). Methods have been developed to enhance the quality of the matrices by introducing into them further observed information on trip patterns, mainly through the matrix calibration process (Psarras et al, 2017).

3.1.4 Evidence is required to distinguish the segments in the matrix which have distinct travel patterns, in order to be depicted to the synthetic matrices. The sample size of source data also constrains the extent to which deterrence functions can be refined.

## 3.2 Hybrid Gravity Modelling

### *Theoretical Background*

3.2.1 The gravity model is often the preferred approach for solving the trip distribution problem in transport planning. It assumes that the interaction between two zones declines with increasing distance, cost and/or time of travel, but is positively correlated with the amount of activity (trips) at each zone. The rate of decline of the interaction is represented by an empirically-determined deterrence function, which must be calibrated so that the resulting synthesized trip cost distribution closely matches the trip cost distribution obtained from observed data.

3.2.2 Mathematically, the gravity model can be expressed as follows:

$$T_{ij} = A_i B_j O_i D_j f(\boldsymbol{a}, C_{ij}), \qquad \sum_i T_{ij} = D_j, \sum_j T_{ij} = O_i$$

and,

$$A_i = \frac{1}{\sum_j B_j D_j f(\boldsymbol{a}, C_{ij})}, \quad B_j = \frac{1}{\sum_i A_i O_i f(\boldsymbol{a}, C_{ij})}$$

where:

- $T_{ij}$ is the matrix of trips between origin zone i and destination zone j;
- $C_{ij}$ is the cost of travel between origin i and destination j;
- $O_i$ is the total number of trips originating at zone i;
- $D_j$ is the total number of trips destined for zone j;
- $A_i, B_j$ are balancing factors, solved for iteratively using a Furness process; and
- f is the deterrence function, for which the parameter vector $\boldsymbol{a}$ needs to be calibrated.

3.2.3 The conventional approach used to calibrate the parameter vector $\boldsymbol{\alpha}$ in a gravity model is based on solving a non-linear minimization problem. In principle, the parameters of the gravity model are estimated through an optimisation process, where a search algorithm is used to find a set of parameters which minimize the squared error between the synthesised and observed distributions, described as below:

$$e_{parms} = \sum_{c} (T_c^{obs} - T_c^{syn})^2$$

where

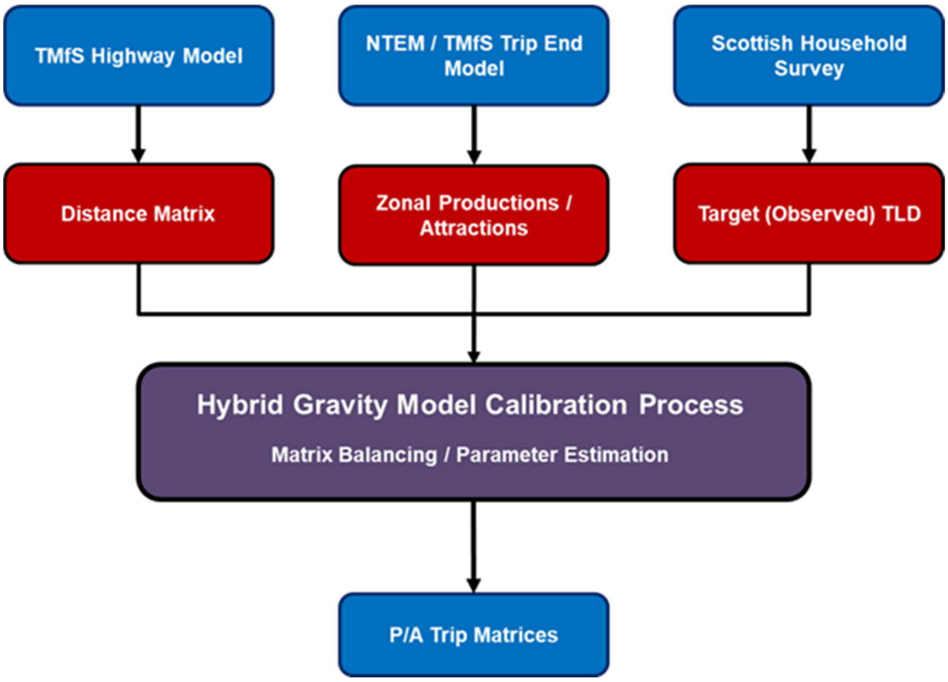
- $T_c^{obs}$ is the number of observed (target) trips in trip cost band $c$
- $T_c^{syn}$ is the number of synthesised trips in trip cost band $c$

3.2.4 It has been shown that the lognormal density function tends to give better description of travel pattern of individuals in comparison with other functional forms such as the Tanner function (Feldman et al., 2012).

3.2.5 Hybrid gravity model refers to gravity models that include additional parameters to allow for simultaneous calibration of multiple deterrence functions taking account of varying Trip Cost Distributions (TCDs). These intend to reflect significant variation in trip lengths and patterns between different geographical areas (e.g. urban vs. rural). A detailed description of the optimisation process where multiple deterrence functions are used is provided in Psarras, et al. (2017).

3.2.6 The process of synthetic matrix development and the inputs required are summarised in Figure 3.1.

**Figure 3.1: Synthetic Matrix Development**




***Consideration of Data Requirements***

3.2.7 The synthetic development process includes the following data requirements:

- Production and attraction trip ends from a trip end model;
- Cost skim matrix representing the cost for every origin-destination movement in the model; and
- Observed trip cost distributions (TCDs).

3.2.8 Trip end models, such as the TMfS Trip end model provides the basis for estimating trips produced and attracted in each model zone.

3.2.9 A key advantage of using the TMfS Trip end model, for example compared to using the National Trip end Model (NTEM), is that it allows for representation of spatial variation in trip rates underpinning trip end estimates. However, a limitation of the TMfS Trip end model is that it is lacking estimates for non-home-based trips. Therefore, NTEM has been applied for this segment as discussed in Section 2.3.

3.2.10 The most common approach to represent the travel cost for every movement in the matrix is the model cost skim. Therefore, the existing TMfS distance skims will be used to provide the cost skims.

3.2.11 Extra care should be given to the trip cost distributions selected for the calibration process. Given the limitations and uncertainty in the SHS trip diary data to develop a reliable and consistent estimate of generalised cost, trip length distributions (TLDs) can be developed as a measure of TCD and applied in the gravity model. To ensure consistency with the definition of cost in the cost matrix, distances from the TMfS skim matrix has been appended to the trip diary data and served as the basis for the TLDs.

3.2.12 Especially in the synthetic matrix development, the TLDs and the selection of the distinct areas for which travel patterns are described by these TLDs is a key determinant of the matrix quality. These reflect systematic variations in TLDs that are not explained by variations in land use proximity.

3.2.13 Hence, the decision for these sub-segments should be drawn on the spatial variation which could exist in the model internal area, along with the available sample size in the observed data sources used to build the TLDs. The spatial variation in travel pattern is reflected in the observed variation in the TLDs by area.

3.2.14 The most straightforward criteria to select the areas and the associated TLDs is the distinct area types which are included in the model area. For example, metropolitan, urban and rural areas exist in the model and these areas tend to have different travel patterns. Using distinct TLDs associated with the different area types, as inputs in the hybrid gravity modelling, the synthetic matrices would represent the different travel patterns.

## 3.3 Implementation

***Matrix Development Process***

3.3.1 The hybrid gravity modelling described in section 3.2 has been applied to develop synthetic matrices, within the TMfS internal area. Accordingly, TMfS highway assignment model has been used to extract distance skims, which represent the cost skim in the synthetic matrix development.

3.3.2 The production and attraction trip ends have been extracted from the TMfS Trip end model. The trip ends have been generated at all-day production-attraction (PA) level, segmented by trip purpose: Home-Based Work (HBW), Home-Based Other (HBO) and Non-Home-Based (NHB).

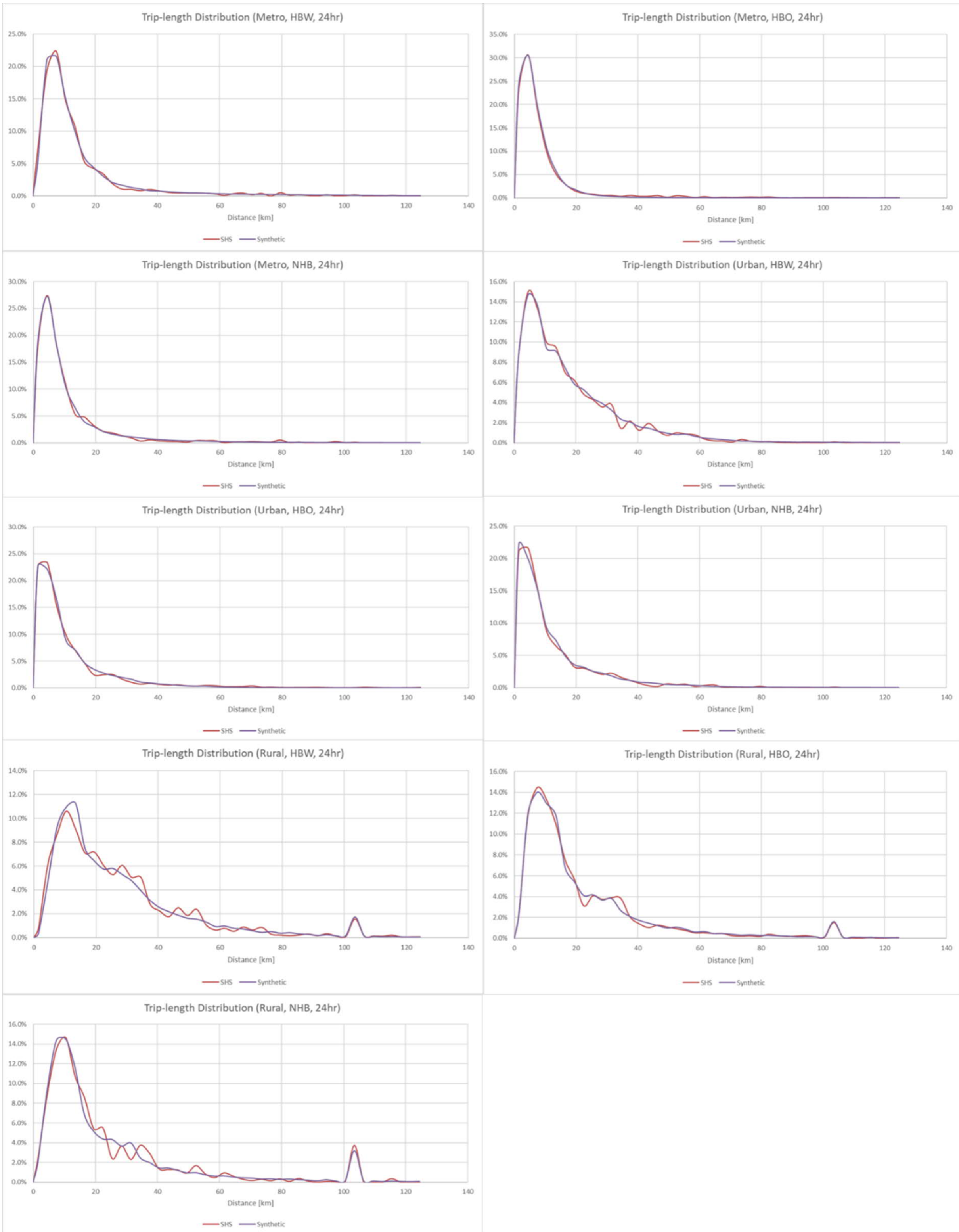
3.3.3 The areas being calibrated in the hybrid synthetic matrix development are based on the TMfS zone classification into metropolitan, urban and rural.

***Results***

3.3.4 One of the key indicators for the performance of the synthetic matrices is the goodness of fit of the modelled TLDs against the observed ones originated from the SHS. Figure 3.2 demonstrates how modelled TLDs from the hybrid gravity models are compared with the observed TLDs for trips produced in metropolitan and urban areas for HBW, HBO and NHB trip purposes, respectively. A good fit is observed in all cases.

3.3.5 In the case of the rural trip length distributions, a hump is observed just above the hundred-kilometre threshold. An investigation was carried out, which revealed that due to the assignment of intrazonal distances on island zones being very large (due to having a far nearest neighbour), intra-zonal island trips had been assigned inflated distances. For this reason, these trips have been excluded from the MND verification process.

## Figure 3.2 TLD Comparisons between Modelled (Synthetic) and Observed (SHS)



# Section 4 – MND Matrix Development

## 4.1 Specification of Mobile Network Data

4.1.1 Matrices of mobile phone data were commissioned from Telefonica, which provided expanded data in the TMfS zoning system and at agreed demand segmentation and geographical coverage.

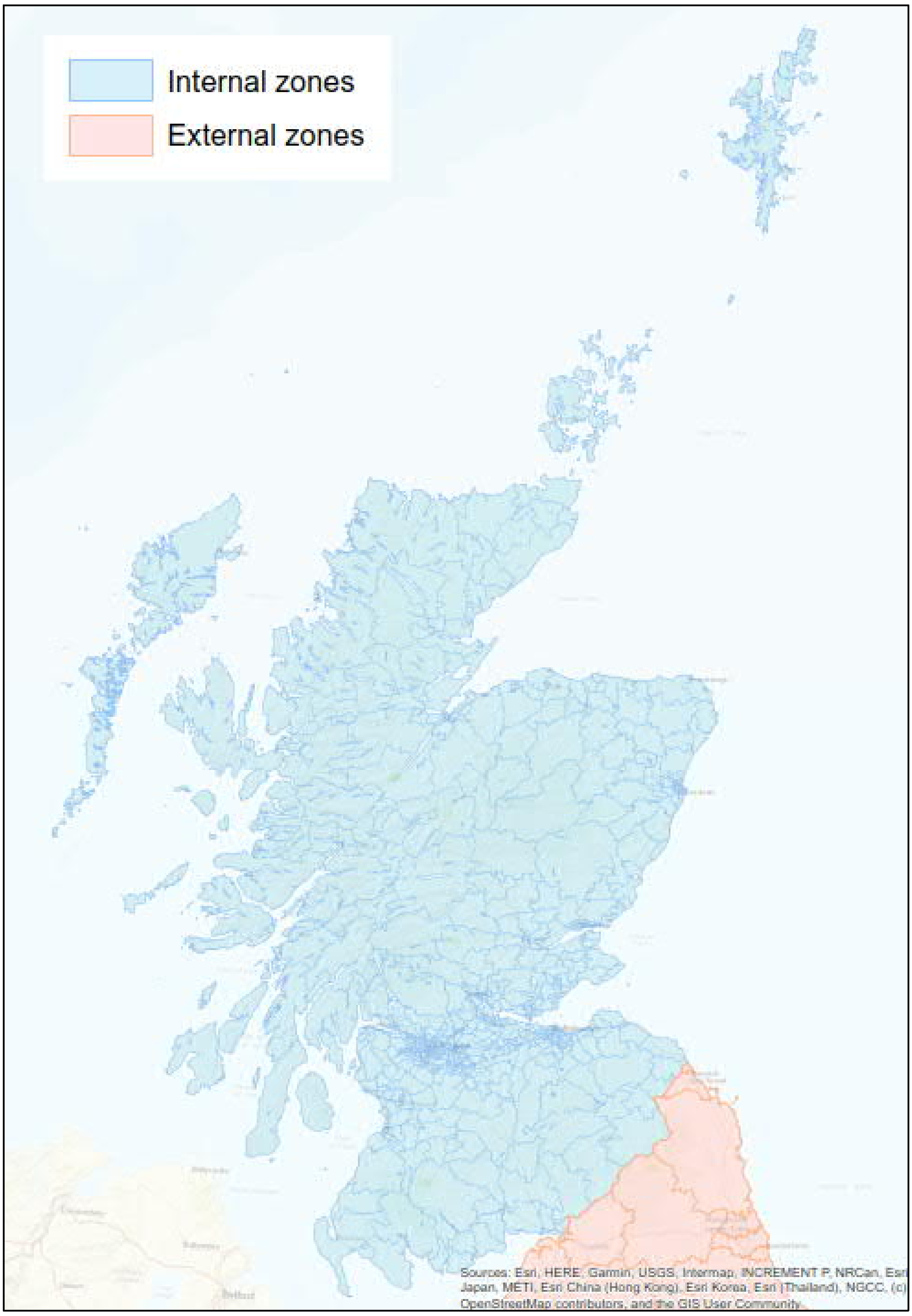
4.1.2 The matrices are based on records of mobile phone movements between 1st March 2019 and 30th April 2019.

4.1.3 Origin Destination (OD) trip records were generated for an average working day (derived from Tuesday-Thursday records) segmented by the following:

- six time periods: AM period (0700-1000), AM peak hour (0800 – 0900), Inter-peak period (1000-1600), PM period (1600-1900), PM peak hour (1700 – 1800), Off-peak period (1900-0700);
- four modes: ferry, flight, road, and rail;
- three trip purposes: home-based-work, home-based-other, and non-home-based; and
- two directions (for home-based trips): from-home and to-home.

4.1.4 The geographical coverage for data collections is shown in Figure 4.1, where the data collection area corresponds to the internal TMfS zones.

**Figure 4.1 Data Collection Area**




4.1.5 All movements that start, finish, or travel through the internal area were included in the processed OD matrices. The initial MND matrices were provided in detail TMfS zoning system, with a further set of matrices re-expanded and reproduced at a more aggregate sectoring level later defined; this is the spatial level the MND matrices are judged to be reliable at and should be used.

## 4.2 MND Matrix Development Methodology

4.2.1 The MND development methodology is summarised in Figure 2.1. It consists of five main steps with the processed secondary data feeding as inputs during the whole processes. The five steps are:

1. Define MND sectoring system
2. Correct aggregate trip biases
3. Identification and removal of non-car trips from the MND data

4. Review and correct for missing short-distance trips
5. Correcting for purpose misallocation of home-based trips

***Step 1: Define MND Sectoring System***

4.2.2 One of the limitations with the use of MND data is the uncertainty in the spatial allocation of MND trips to model zones. The more disaggregate the zone system is, the more uncertainty is attached to the allocation of trip origins and destinations. On the other hand, the use of a more aggregate system lead to a loss of granularity and the information content being reduced. Therefore, it is vital to strike a balance between the two extremes. Figure 4.2 compares the TMfS18 zoning and the final defined MND sectoring system. The process to define this sectoring system is described below.

**Figure 4.2 Comparison of the TMfS Zoning and MND Sectoring Systems at the Central Belt Region (above) and for Central Glasgow (below)**


4.2.3 An analysis was carried out to investigate the reliability in the zone allocation of the MND trip ends by comparing them to the trip end from the TMfS trip end model and looking for suspicious allocation patterns, for example neighbouring zones having significantly different trip rates despite having similar land use. To aid this analysis, the Point of Interest (POI) score, which is a measure developed by the data provider to reflect the zonal allocation certainty, served as an additional measure to identify zones with low certainty.

4.2.4 Using an iterative process, a sectoring system was defined for which analysis and processing of the MND data was deemed to be satisfactory in terms of trip end allocation, but still being at such a disaggregate level as to be useful for the intended use. The defined sectoring system resulted in 524 sectors, compared to the 803 TMfS zones. Figure 4.3 and Figure 4.4 highlight the significant increase in the correlation between the MND and TMfS Trip end model trip ends following the aggregation process.

**Figure 4.3 Trip End Comparison Before Aggregation**

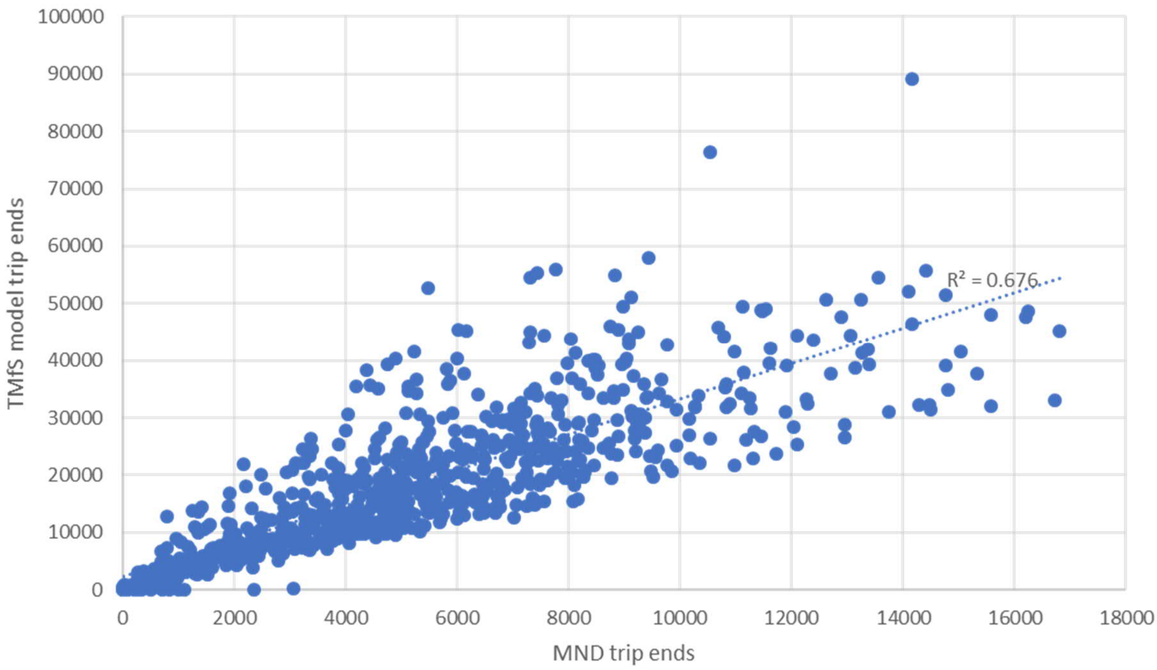
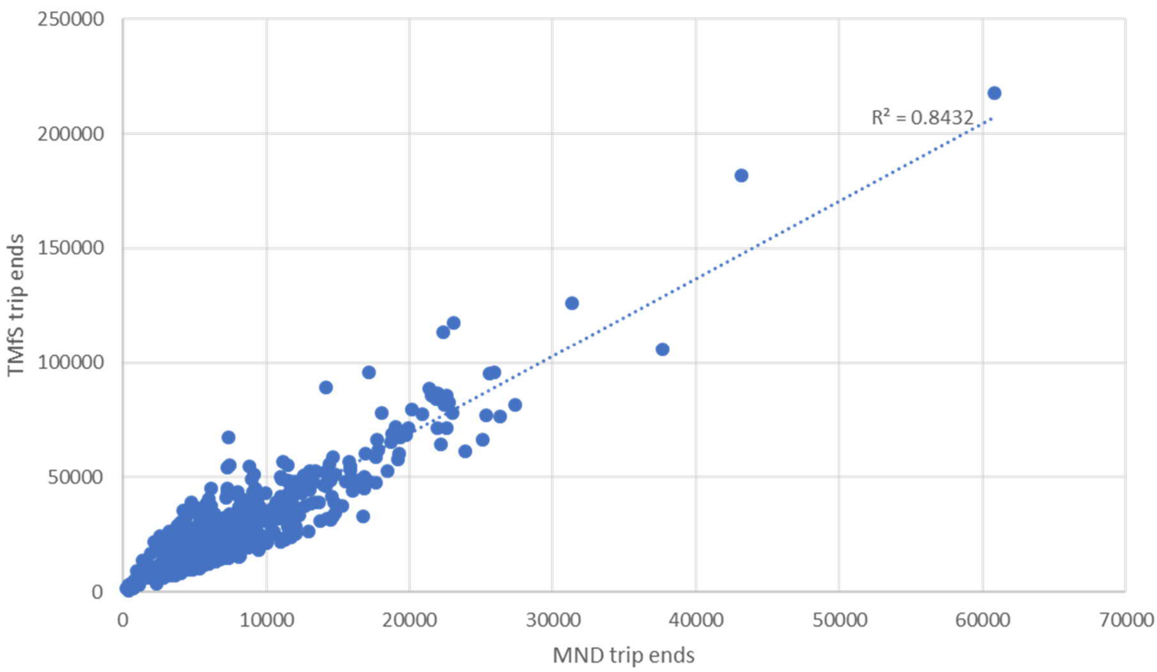



**Figure 4.4 Trip end Comparison After Aggregation**




4.2.5 The OD matrices were reprocessed and re-expanded by the MND data provider into these defined sectors, hereafter referred to as “MND sectors”. The following processing steps and adjustments of mobile data were undertaken at the above sector level.

***Step 2: Correct Aggregate Trip Biases***

4.2.6 As suggested by Table 4.1, the MND data is displaying a global pattern of fewer trips compared to the complimentary data sources for home-based trips when trips produced in the internal area are compared. This pattern is above and beyond the expected lack of short-distance trips.

**Table 4.1 MND Trips Compared to Trips Derived from Complimentary Data Sources per Distance Band (Home-Based Trip Purposes)**

| From (km) | To (km) | MND Road | Synthetic Car, HGV, LGV & Bus | MND Road to Synthetic, HGV & LGV |
|---|---|---|---|---|
| 0 | 2 | 13,472 | 462,045 | 3% |
| 2 | 4 | 89,298 | 992,694 | 9% |
| 4 | 6 | 196,240 | 830,013 | 24% |
| 6 | 8 | 316,798 | 778,393 | 41% |
| 8 | 10 | 356,188 | 701,080 | 51% |
| 10 | 30 | 1,688,708 | 2,344,216 | 72% |
| 30 | 100 | 746,411 | 970,349 | 77% |
| 100 | 500 | 2,603,271 | 3,634,674 | 72% |
| **10** | **500** | **2,519,195** | **3,474,620** | **73%** |

4.2.7 Two distinct biases are at play within the MND data. It is missing short-distance trips, but gradually less so until a threshold of roughly 10 kilometres. After this distance, the proportion of MND trips compared to the synthetic data remains somewhat constant (within the expected uncertainty in the data sources). This suggests that the MND data suffers from a systematic under-detection of trips. A reason for this is likely to relate to the expansion process carried out by the MND data provider.

4.2.8 Of all trips longer than 10 kilometres, the proportion of MND trips compared with synthetic data is 73% for home-based trips. Therefore, an adjustment factor of 1.37 was applied to all home-based MND trips to align the trip end total to that of the Synthetic data. No evidence for such bias was found for non-home-based trips.

***Step 3: Identification and Removal of Non-Car Road Trips from the MND Data***

4.2.9 In order to produce a car matrix, it is necessary to identify and remove other modes of travel from the MND Road matrices. These modes consist largely of freight and bus/coach trips. Freight trips have been estimated separately for HGV and LGV as outlined in sections 4.2.10 - 4.2.14. Identification of bus trips is based on census travel-to-work data as outlined in sections 4.2.15 - 4.2.16. The removal of non-car modes from the MND matrices was carried out in the following order:

1. HGV;
2. LGV;
3. Bus.

**HGV**

4.2.10 The HGV matrices are sourced from CSRGT data as explained in Section 2.3.The removal of HGV trips was subsequently carried out at all-day level on the assumption that all HGV trips are non-home-based. Therefore, the HGV matrix was subtracted from the non-home-based MND road matrix. The HGV matrix was estimated at NUTS region level and the removal process were therefore carried out at this level of aggregation.

4.2.11 A cap of 50% was applied for the removal with the assumption that it is highly unlikely that HGV would constitute more than 50% of all modes for any sector-to-sector movement.

**LGV**

4.2.12 The removal of LGV trips was carried out at all-day level and separately for home-based and non-home-based purposes. The basis for the LGV matrix is the TMfS18 LGV prior matrix, however it was adjusted using insights from the RSI data to correct the bias observed in its trip length distribution, as outlined in Section 2.3.

4.2.13 A removal cap of 50% was also applied to the removal of LGV trips under the assumptions that LGV trips would not constitute more than 50% of the combined car, bus and LGV trips for any aggregate sector-to-sector movement.

4.2.14 As the LGV matrix is a fully specified matrix at the MND sector level it allowed for removal of trips to be explored at various levels of aggregation. Removal at a very disaggregate level could risk the chance of the removal cap being reached due to increases uncertainty in trip numbers for many movements and thus reducing the overall LGV trips being removed. On the other hand, removal at a very aggregate system would reduce the accuracy of the removal process. Following a comparison of the removal at various levels of aggregation, it was determined that removal on TMfS region level (see Figure 2.2) would yield the most realistic results.

**Bus**

4.2.15 The removal of bus trips was carried out at all-day level, separately for the HBW, HBO and NHB purposes. The bus matrix is estimated at local authority-level, which therefore also determined the aggregation level for the removal of bus trips.

4.2.16 A removal cap of 50% was also applied to the removal of bus trips under the assumptions that bus trips would not constitute more than 50% of the combined car and bus trips for any movement.

4.2.17 Table 4.2 highlights the change in overall trip ends. As expected, trips are reduced across all purposes following the removal of freight and bus trips. This is especially true for NHB trips, as freight trips forms a large proportion of non-home-based trips.

**Table 4.2 Trip End Totals Before and After Removal of Bus and Freight Trips**

| Purpose | Synthetic Car | MND - Before removal | MND - After removal |
|---|---|---|---|
| **NHB** | 674,010 | 916,705 | 702,539 |
| **HBW** | 2,147,213 | 1,548,094 | 1,438,492 |
| **HBO** | 3,860,311 | 3,239,741 | 2,945,636 |

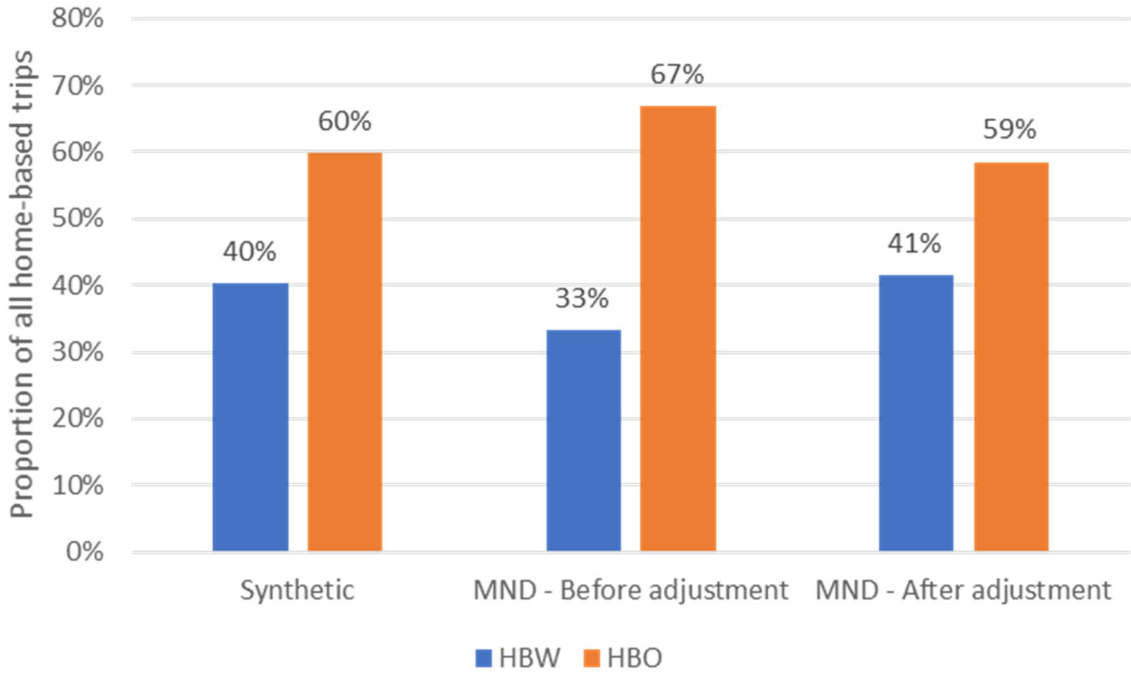
***Step 4: Correcting for Purpose Misallocation of Home-Based Trips***

4.2.18 Following the identification and removal of non-car trips from the MND data a clear pattern of under-estimation of HBW and, conversely, over-estimation of HBO trips was revealed.

4.2.19 This effect could be explained by misallocation of home-based trips between HBW and HBO purposes within the MND data processing when trip purposes are inferred. In order to correct for this bias, the purposes were reallocated from HBO to HBW for each OD movement when comparison with purpose split from synthetic matrices suggested that this was required. This process took into consideration the estimation error within the purpose split of the synthetic matrices (derived from SHS data); a 10% error range was assumed and the MND purposes were only corrected where the difference was beyond this range.

4.2.20 The resulting purpose split for trips longer than 5 kilometres is displayed in Figure 4.5.

**Figure 4.5 Purpose Split for Trips Longer than 5km**




***Step 5: Review and Correct for missing Short-distance Trips***

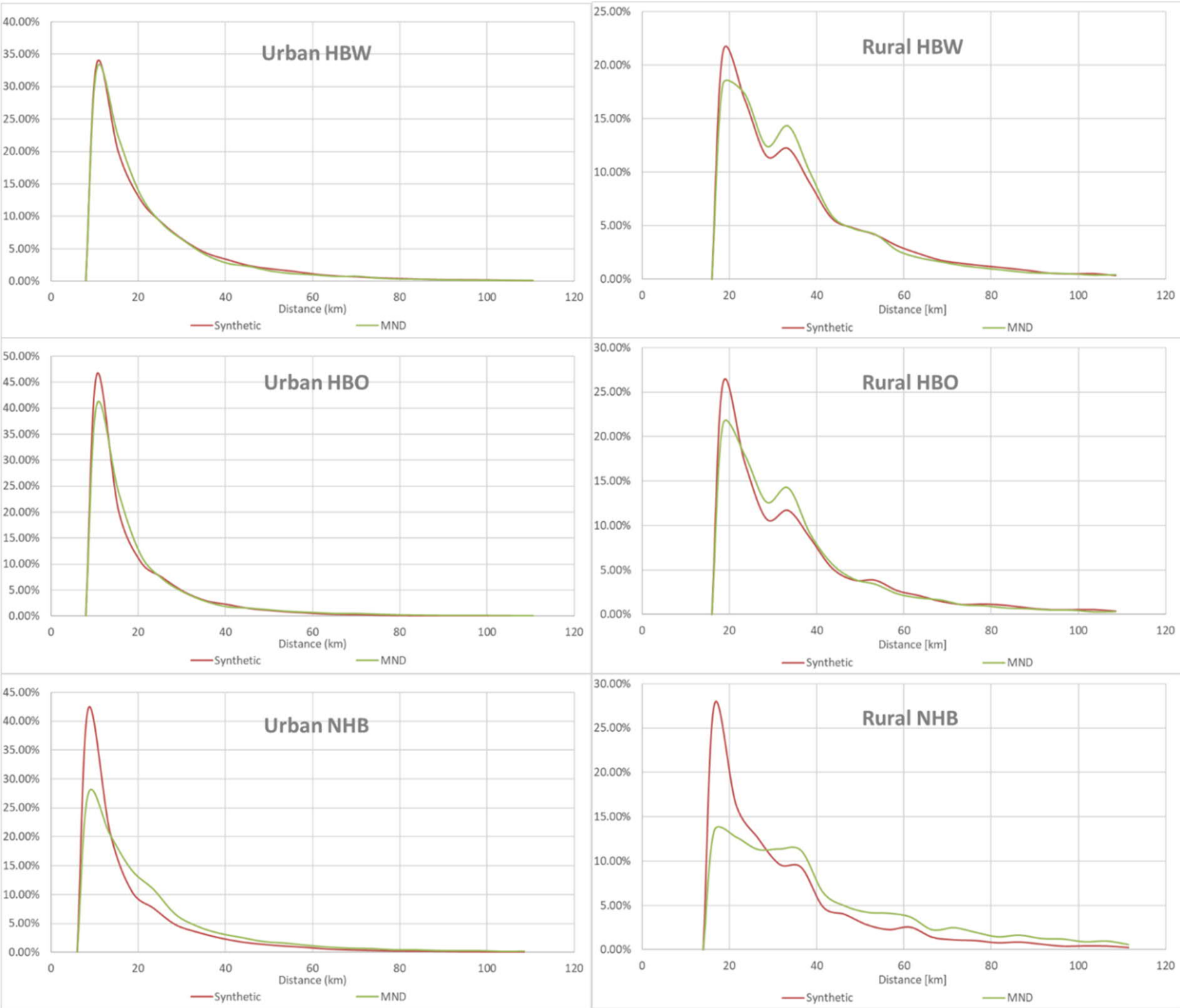
4.2.21 One key limitation of MND data is that it lacks short trips. This is due to short trips not being recorded by more than one mobile phone mast and thus they cannot be detected. In order to address this deficiency, short trips from MND are replaced with synthetic trips or re-expanded, depending on the trip distance, up to a distance threshold above which the MND data is deemed to be sufficiently able to capture all trips. This threshold has been determined by comparison of trip length distributions between MND and synthetic matrices.

4.2.22 As the MND trip detection process is dependent on trip frequency and inferred home-location, the reliability of the trip detection varies with trip purpose. Similarly, the detection of MND trip data varies between area types, due to the mast network density varying between urban and rural areas. Therefore, comparisons of trip length distributions were carried out separately for each purpose and separately for areas classified as either metropolitan and urban (collectively referred to as “urban”) and rural trips.

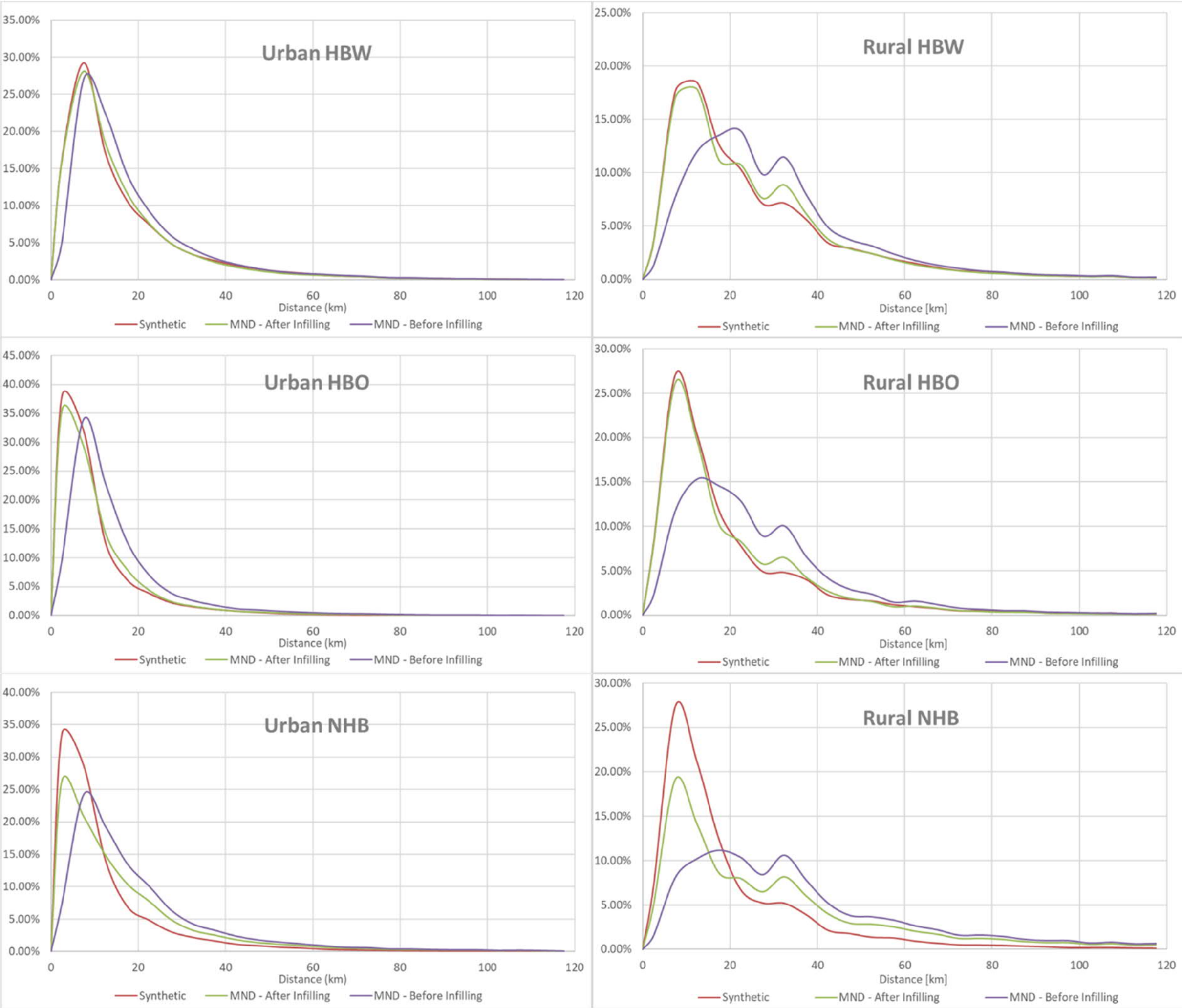
4.2.23 Distance bins of two kilometres was examined to determine the exact thresholds for infilling. For distance bins where MND formed a very low proportion (less than 50%) compared to synthetic data, synthetic data was used to replace the MND data. For distance bins where trip proportions were still significantly smaller than synthetic data, but more than 50%, the MND was uplifted to align with the corresponding number of synthetic trips.

4.2.24 Figure 4.6 highlights the distances where the truncated TLD for synthetic and MND trips begin to align and thereby indicates the threshold below which infilling, or uplifting, should be carried out. Figure 4.7 demonstrates the resulting shifts in the TLDs before and after the infilling and uplifting. As shown, there is a better match with observed data for home-based purposes; this is discussed further in Section 5.5.

## Figure 4.6: Comparisons of Truncated TLDs



## Figure 4.7: TLD Shifts Before and After Infilling




4.2.25 The thresholds for infilling and uplifting of the MND matrices are summarised in Table 4.3.

## Table 4.3: Distance Thresholds for Infilling

| | HBW | | HBO | | NHB | |
|---|---|---|---|---|---|---|
| | Urban | Rural | Urban | Rural | Urban | Rural |
| **Replaced with Synthetic** | <6 km | <10 km | <6 km | <10 km | <6 km | <14 km |
| **Uplifted MND** | 6-8 km | 10-16 km | 6-8 km | 10-16 km | | |
| **Original MND** | >8 km | >16 km | >8 km | >16 km | >6 km | >14 km |

# Section 5 – Matrix Verification

## 5.1 Verification Methodology

5.1.1 Whilst there is little control over the issues in the MND matrices as the data processing is undertaken independently by Telefonica, it is necessary to understand and identify any issues or biases in the data, so that these could be addressed during the MND development process. It is also important to ensure that, after adjustments, the data is as unbiased as possible and is consistent with reliable secondary data sources. This is the main purpose of the verification stage.

5.1.2 The following key elements of the mobile phone data were reviewed and verified through comparisons with independent data sources.

- matrix symmetry;
- trip ends and trip rates;
- pattern of trips; and
- allocation of trips to purposes.

5.1.3 The verification process was undertaken iteratively, following each processing step, in order to assess the impact of each step on the outturn matrices and whether the outcome is consistent with external evidence and general expectation. The following describes the verification tests undertaken and the key results.

## 5.2 Matrix Symmetry

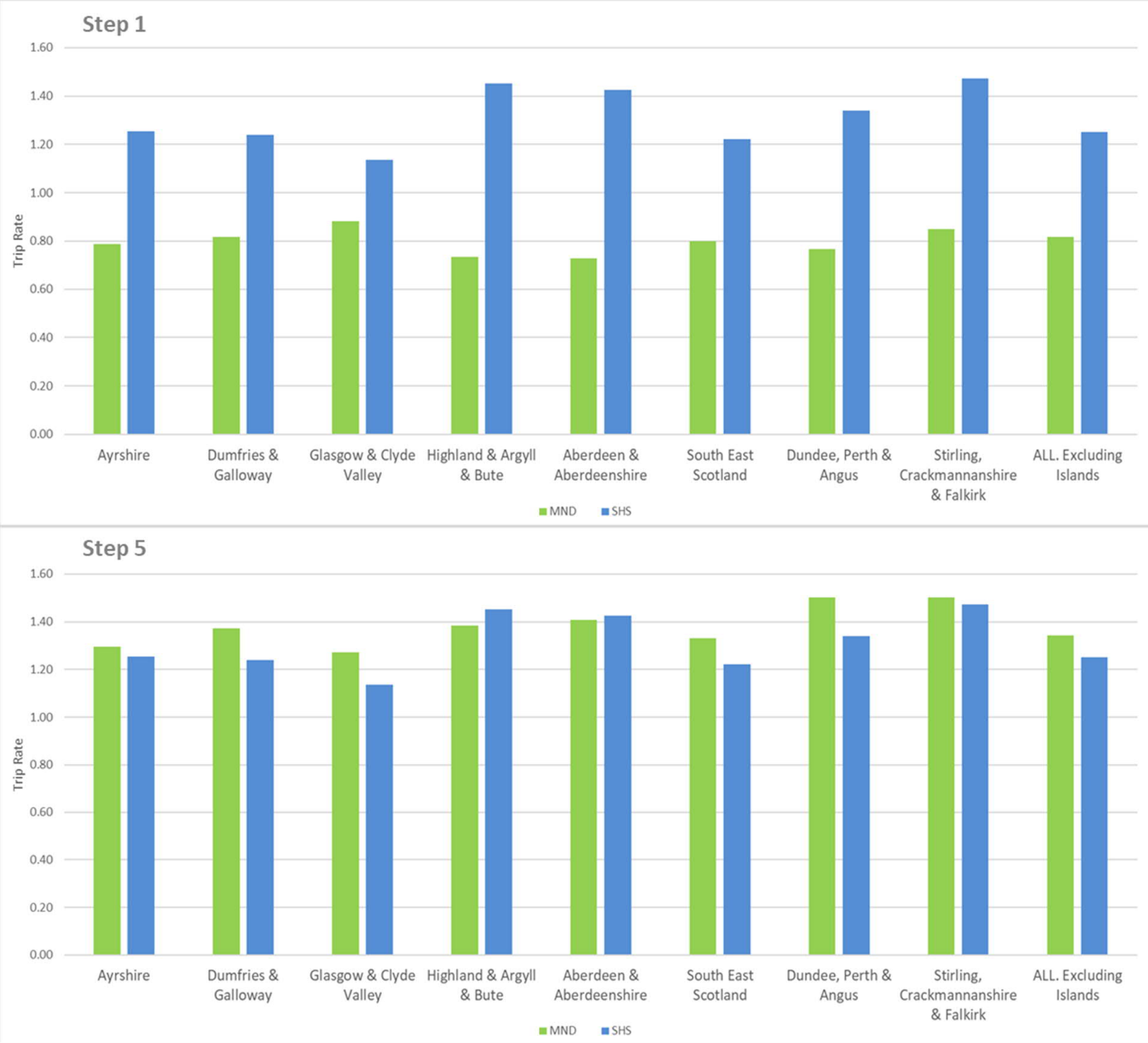
5.2.1 Figure 5.1 provides a comparison of all-day trip origins and destination at MND sectoring system. As suggested by the slope and the high value of $R^2$, there is a high degree of symmetry within the MND matrices as most trip makers seem to return to their place of origin at the end of an average day. A high degree of correlation is maintained before and after the MND matrix development process, as demonstrated by the scatter plots for steps 1 and 5.

**Figure 5.1 Comparison of Internal Trip Origins and Destinations in MND Matrices**


## 5.3 Trip Rates

5.3.1 Figure 5.2 compares MND and SHS trip rates across the TMfS regions (see Figure 2.2). Initially, trip rates are consistently under-reported, mainly due to the inability to fully detect short-distance trips in the MND matrices. Following the MND development process the MND trip rates are aligned with the SHS data, while continuing to reflect geographical variation.

Figure 5.2 Comparison of MND and SHS Trip Rates at Region Level at Step 1 and Step 5




## 5.4 Trip Ends

5.4.1 Table 5.1 provides an overview of the total number of internal trip productions for each trip purpose at each development step. Compared to the unprocessed MND data (Step 1), trip end totals are generally more aligned with the synthetic data (i.e. trip end model). The exception is for NHB trips, where MND totals are larger. These have deliberately not been adjusted as there is more uncertainty in the estimation of synthetic NHB trips; they are derived from NTEM rather than the TMfS trip end model.

Table 5.1 Total Trip End Productions by Purpose for Each Step

| | | MND Productions | | | | |
|---|---|---|---|---|---|---|
| **Purpose** | **Synthetic Productions** | **Step 1** | **Step 2** | **Step 3** | **Step 4** | **Step 5** |
| **HBW** | 2,147,213 | 1,121,807 | 1,548,094 | 1,438,492 | 1,783,814 | 2,205,900 |
| **HBO** | 3,860,311 | 2,347,640 | 3,239,741 | 2,945,636 | 2,600,314 | 4,095,478 |
| **NHB** | 674,010 | 916,705 | 916,705 | 702,539 | 702,539 | 905,064 |
| **All** | 6,681,535 | 4,386,152 | 5,704,540 | 5,086,668 | 5,086,668 | 7,206,442 |

## 5.5 Trip Length Distribution

5.5.1 Figure 5.3 to Figure 5.8 display comparisons between the trip length distribution of the MND and synthetic data for each purpose and area type, in each MND matrix development step. Figure 5.9 highlights the change in average trip length, per purpose and area type, for each step. Note that step 2 (correcting for aggregate trip biases) has been omitted from these figures as this step has no impact on the trip length distribution.

5.5.2 One key feature is how the unprocessed MND data (step 1) is lacking short trips compared to the synthetic data, with this short distance bias being more prevalent in rural areas and for NHB trips. This can likely be attributed to increased mast density in built up areas and that HBW trips, and to a lesser degree HBO trips, are made more frequently than NHB trips, which increase the likelihood of detention.

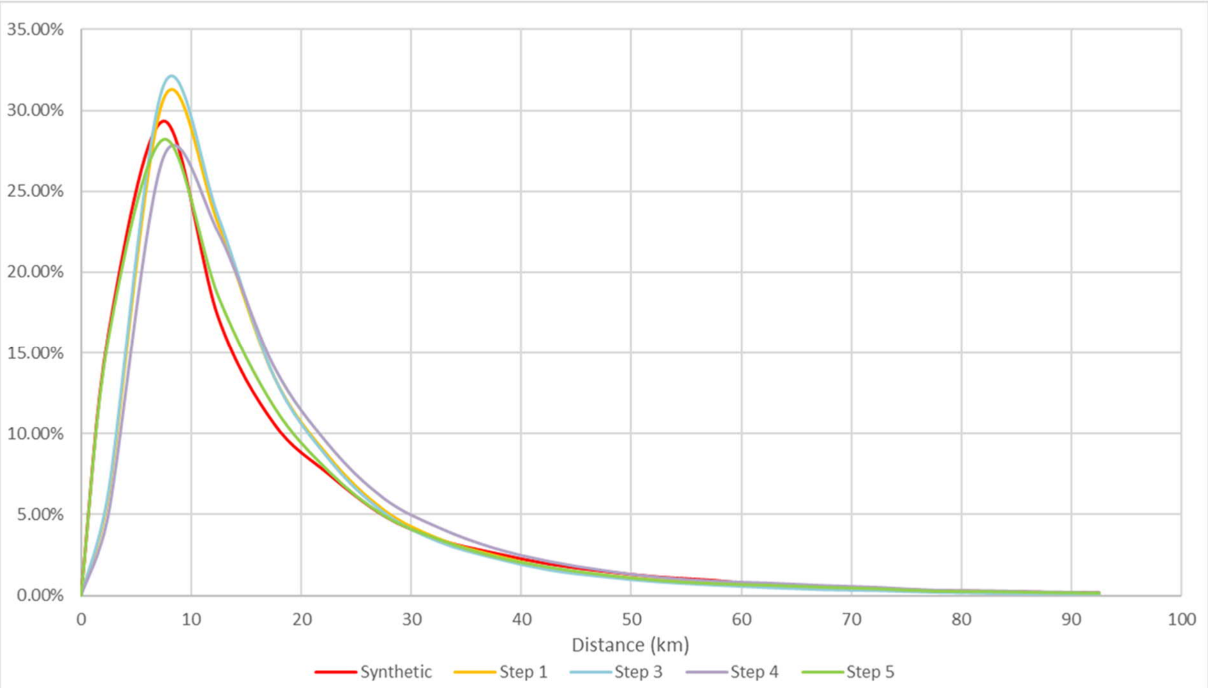
5.5.3 Following the removal of bus and freight trips (step 3) the distribution shifts towards shorter trips for all purposes and area types, as evidenced by the TLD patterns. The expectation would be to see a larger shift towards shorter trips for the non-home-based trip purpose than the home-base one. This is because freight trips are generally long trips and are normally found to be mostly non-home-based. However, this was not a clear finding in the TLD shift for NHB rural trips, although it was revealed through a reduction in its average trip length.

5.5.4 Although an expected reduction in trip length took place across all purposes and area types, the reduction was not as large as anticipated for non-home-based trips. This result can likely be attributed to the quality of the freight matrices. As freight trips are likely to constitute a large part of the unfiltered MND data for NHB trips, poor quality freight data has impacted the ability to remove these trips from the MND data.

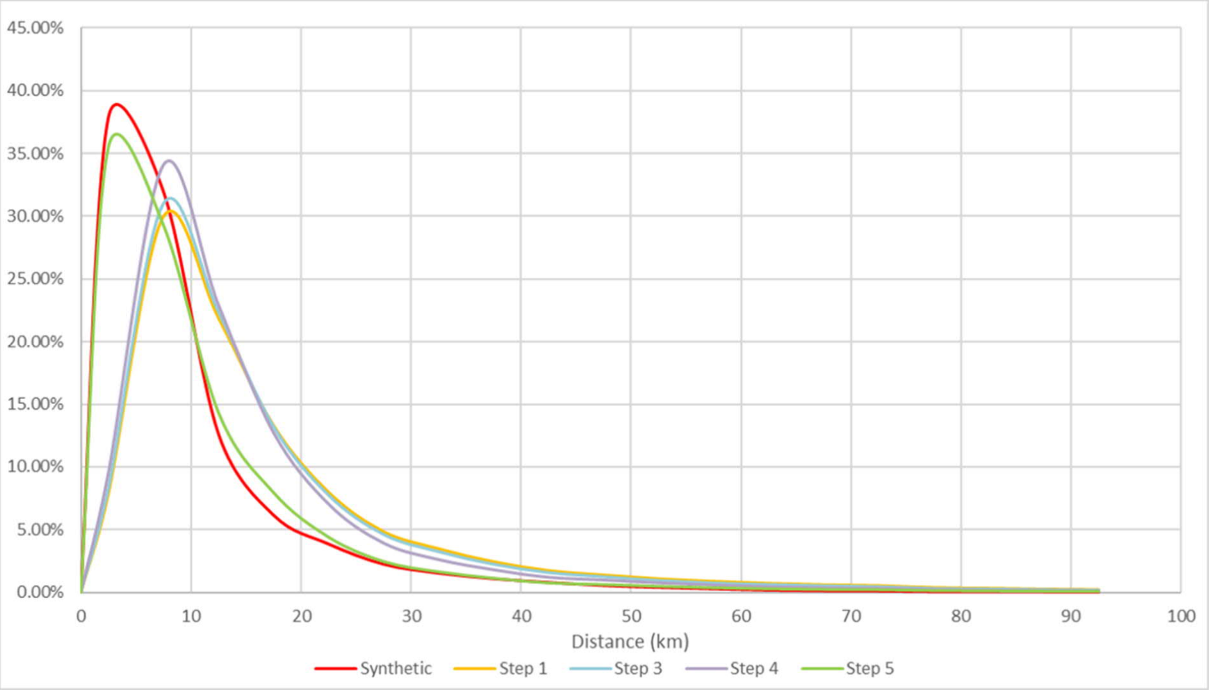
5.5.5 Following the adjustment of home-based purpose proportions (step 4), trip length distributions and averages for HBW and HBO nears the ones from the synthetic matrices, as is expected. During this step, non-home-based trips remain unaffected.

5.5.6 Following the final step of infilling short distance trips (step 5), the MND and synthetic trip length distributions and average trip lengths closely match for the home-based purposes. Bigger discrepancies are observed for NHB trips, which can be explained by two previously discussed reasons. Firstly, there is more uncertainty within the synthetic NHB matrices, stemming from being generated from NTEM rather than the trip end model and less confidence in representation of NHB trips within SHS, used to develop observed TLDs for synthetic matrix development. Furthermore, due to the low quality of the freight data, the removal process of the freight trips includes a high level of uncertainty.

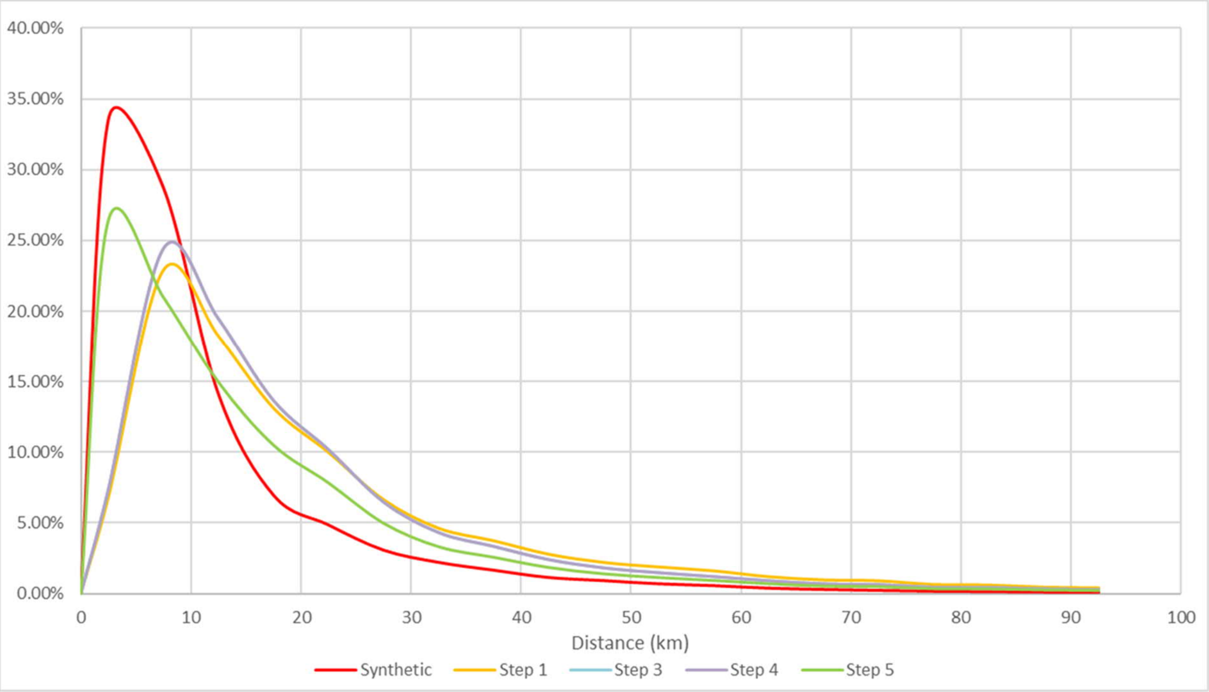
**Figure 5.3 Change in TLDs for HBW Urban Trips**



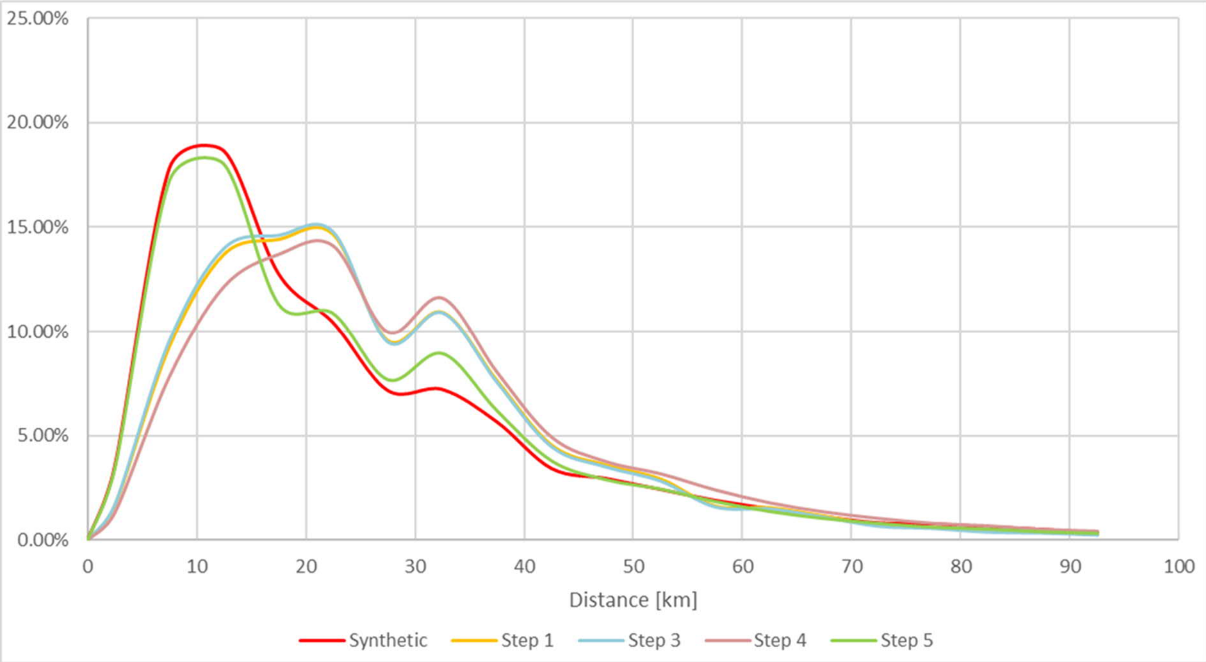
## Figure 5.4 Change in TLDs for HBO Urban Trips




## Figure 5.5 Change in TLDs for NHB Urban Trips



## Figure 5.6 Change in TLDs for HBW Rural Trips




## Figure 5.7 Change in TLDs for HBO Rural Trips

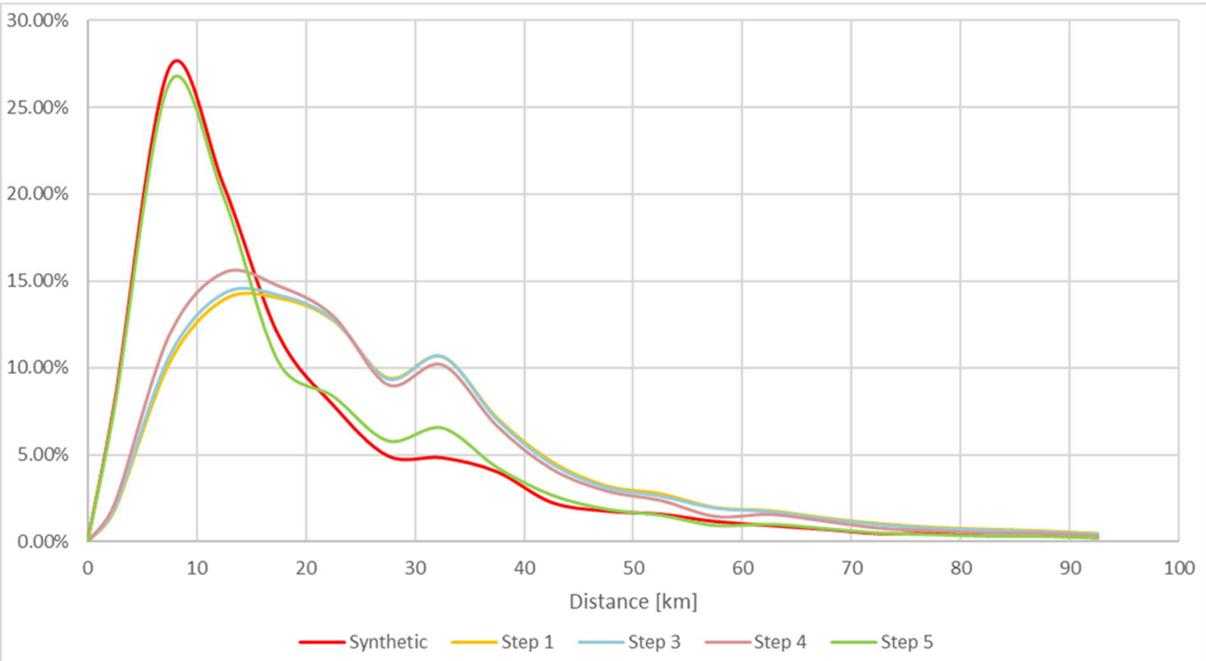


**Figure 5.8 Change in TLDs for NHB Rual Trips**

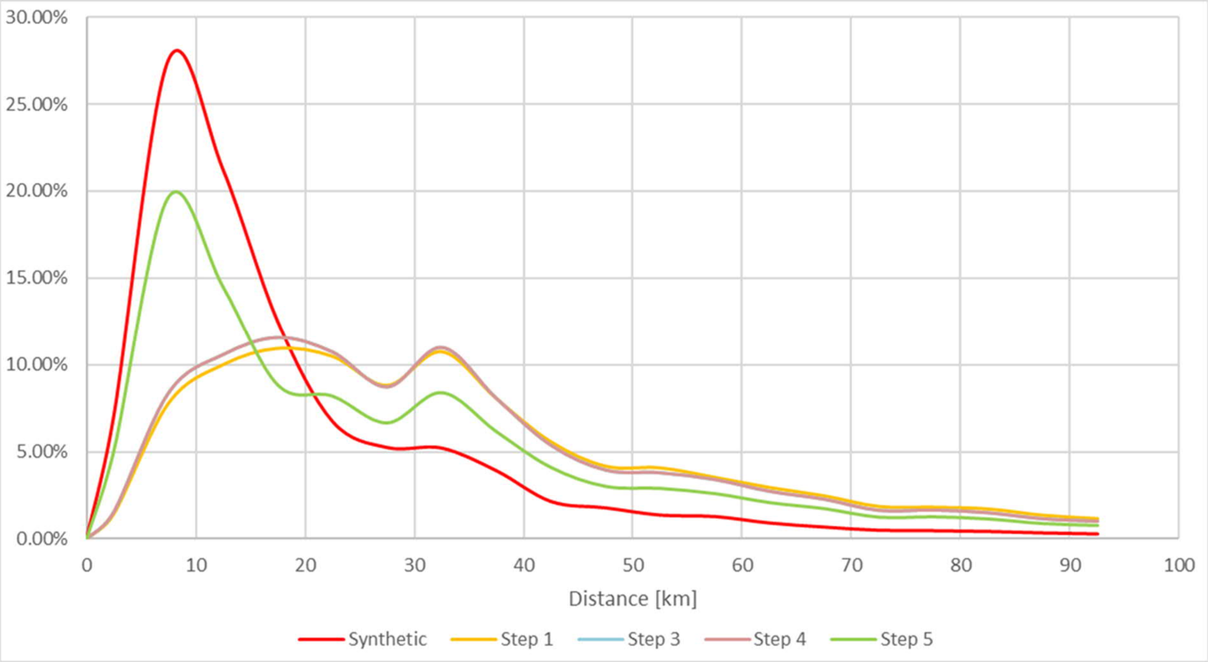
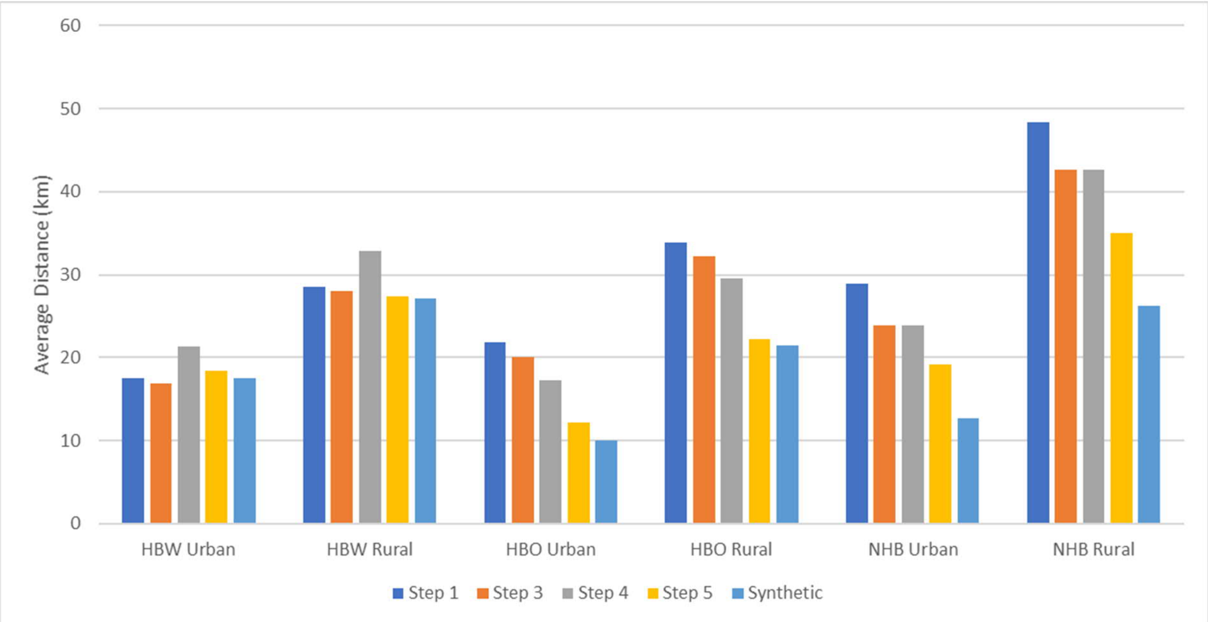
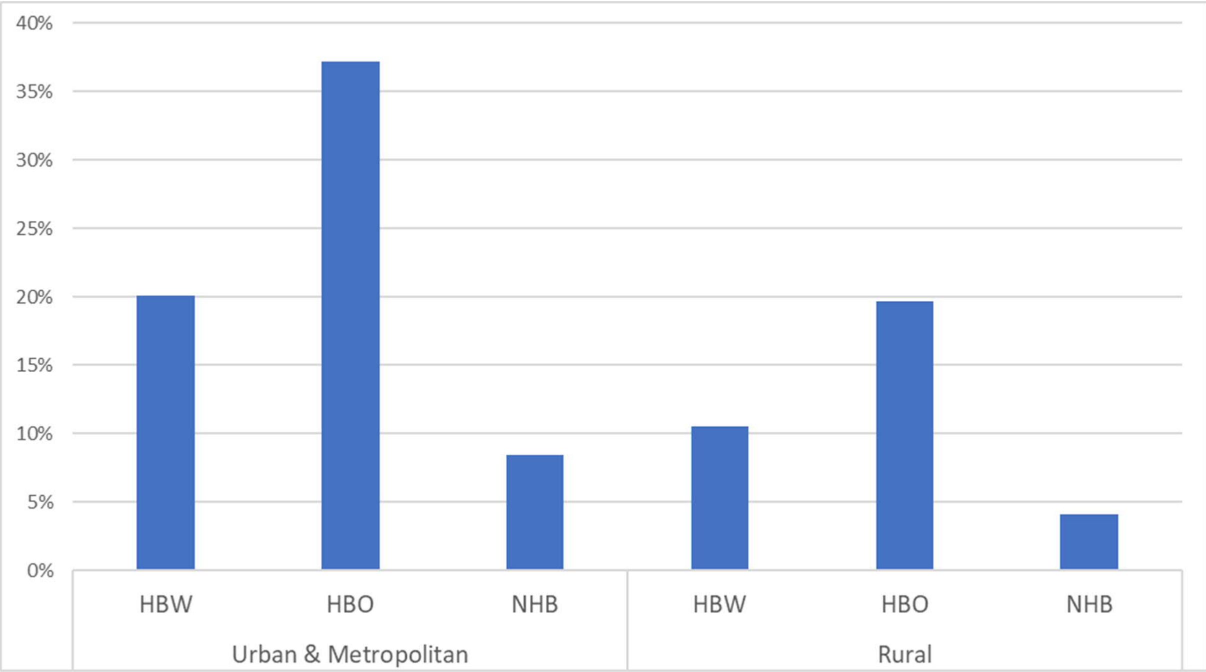



**Figure 5.9 Change in Averge Trip Distance**




5.5.7 To put into context any impact resulting from the low quality freight data and the subsequently affected MND car matrix as a whole, the relative size of each matrix segment is highlighted in Figure 5.10. Any issues related with the removal of freight trips from the non-home-based matrices is deemed to have a small impact as they represent a small proportion of all trips.

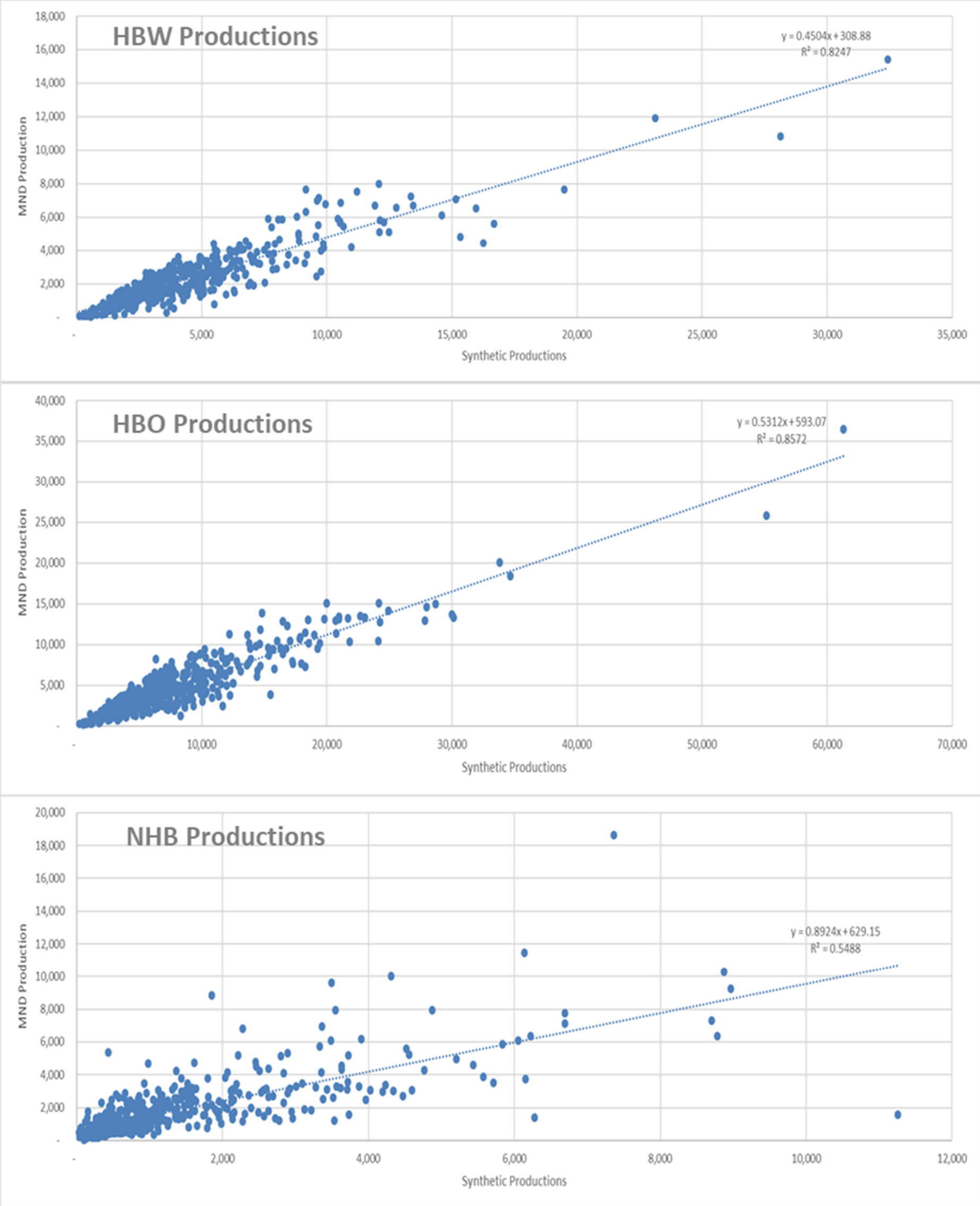
**Figure 5.10 Proporportion of Trips by Area and Purpose**



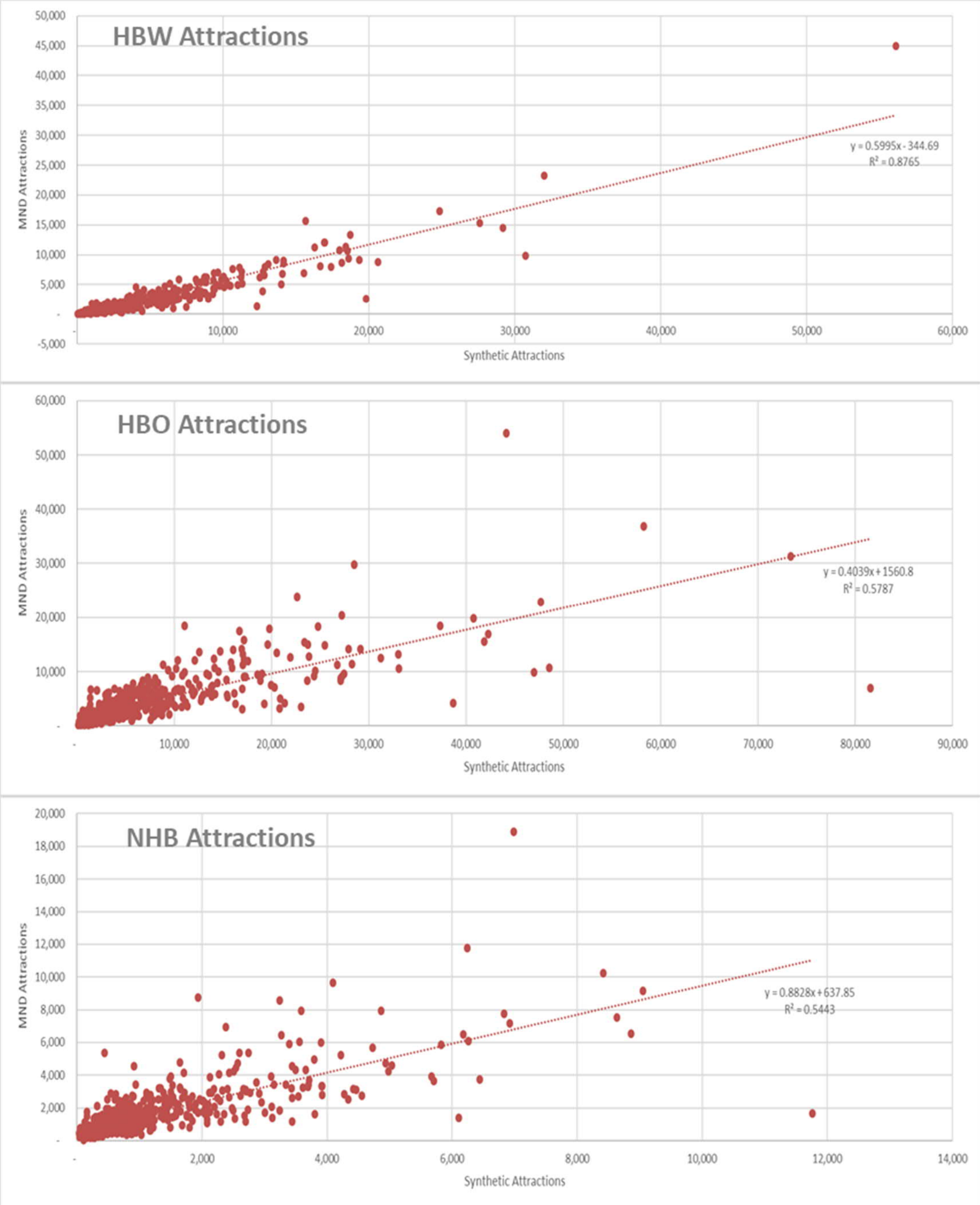

## 5.6 Trip End Scatter Plots

5.6.1 Figure 5.11 and Figure 5.12 highlight the relationship between internal trip end productions and attractions at the initial stage between the MND and synthetic matrices. This can serve as an indicator on the accuracy of the distribution of trip ends in the MND data. While there is a high degree of similarity for HBW trips, attractions for HBO and NHB trips in general have a lower level of agreement, where there is also more uncertainty in synthetic trip attractions (sourced from the trip end model).

## Figure 5.11 Comparison of Synthetic and MND Internal Productions (Step 1)



## Figure 5.12 Comparison of Synthetic and MND Internal Attractions (Step 1)




5.6.2 Figure 5.13 and Figure 5.14 highlight the relationship between internal trip end productions and attractions between the MND and synthetic matrices, following the MND matrix development process. The high degree of similarity observed in the unprocessed MND data for HBW trips and HBO attractions has remained. Furthermore, there is a marked increase in the relationship between MND and synthetic trip ends for HBO attractions and NHB trip ends. This is an indication that the MND development process has managed to adjust the trip end distribution in order to better reflect the demographic characteristics of the MND sectors.

## Figure 5.13 Comparison of Synthetic and MND Internal Productions (Step 5)

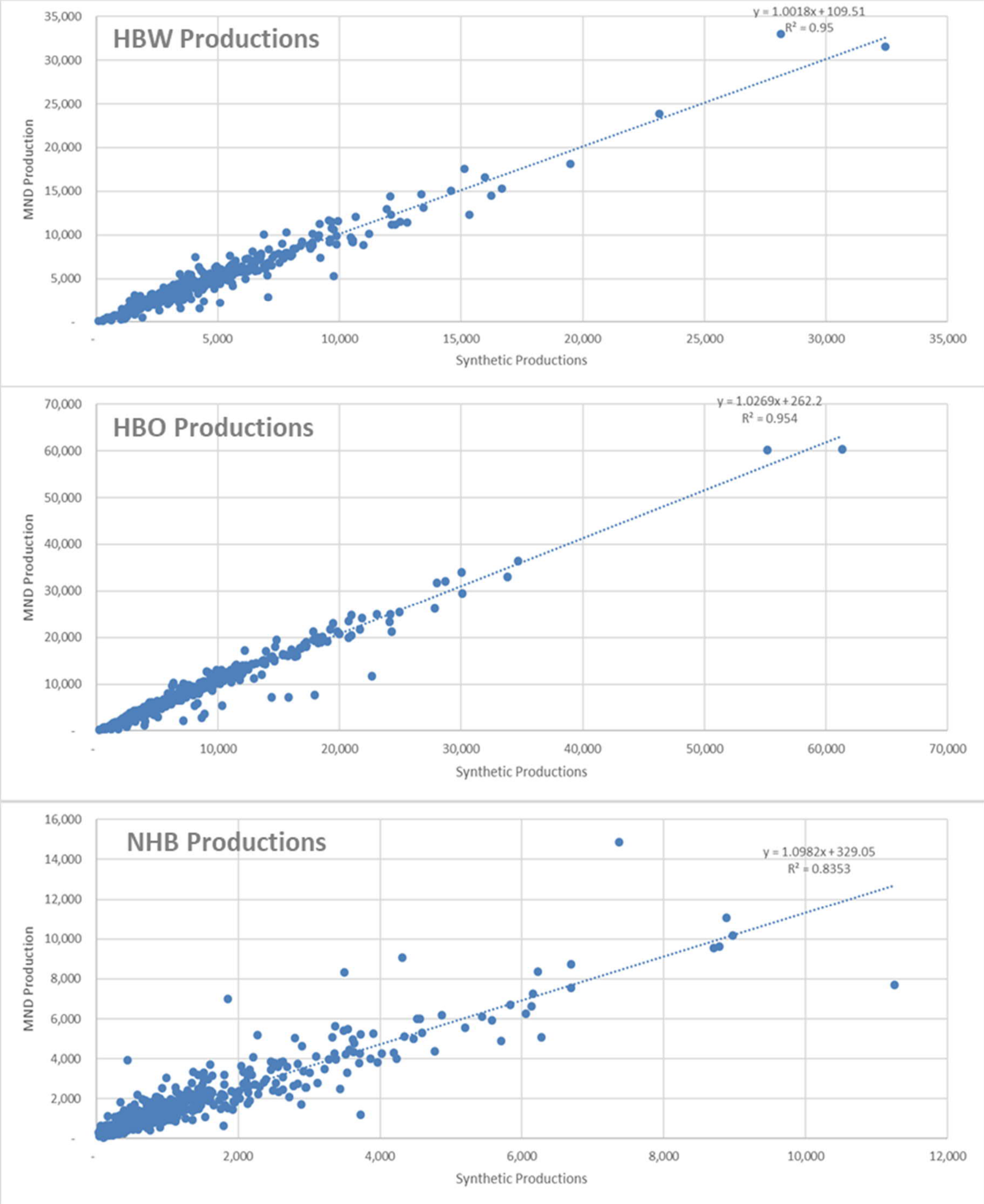
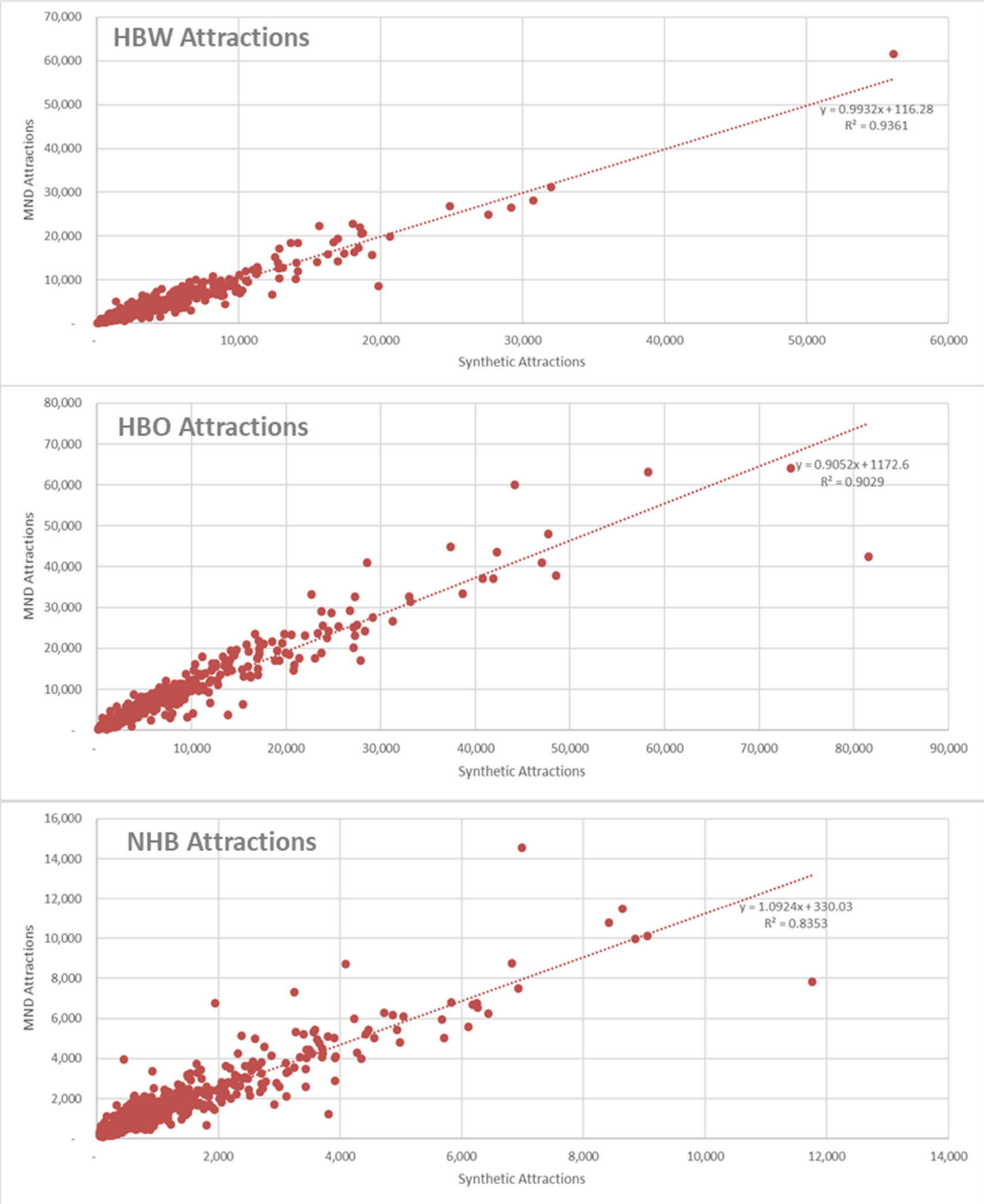


**Figure 5.14 Comparison of Synthetic and MND Internal Attractions (Step 5)**




## 5.7 Sector – Sector Trips

5.7.1 Table 5.2 provides an overview of the sector-to-sector OD movements at TMfS region-level. It highlights that a significant majority of trips are concentrated between a few numbers of regions in Scotland, mostly corresponding to the central belt area (South East Scotland along with Glasgow and the Clyde Valley). Table 5.3 provides the corresponding movements in the synthetic matrices sharing a similar pattern.

### Table 5.2 Mobile Network Data Sector to Sector Trips at TMfS Regional Level

| Sector | Ayrshire | Dumfries & Galloway | Glasgow & Clyde Valley | Highland & Argyll & Bute | Aberdeen & Aberdeenshire | South East Scotland | Dundee, Perth & Angus | Stirling, Crackmannanshire & Falkirk | Totals | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Ayrshire** | 414,454 | 2,367 | 32,435 | 1,055 | 111 | 998 | 253 | 529 | **452,201** | 6.30% |
| **Dumfries & Galloway** | 2,224 | 179,223 | 3,950 | 63 | 42 | 1,336 | 137 | 245 | **187,220** | 2.60% |
| **Glasgow & Clyde Valley** | 40,989 | 3,910 | 2,260,621 | 10,016 | 705 | 31,707 | 3,478 | 34,276 | **2,385,701** | 33.30% |
| **Highland & Argyll & Bute** | 1,209 | 79 | 9,436 | 516,245 | 8,278 | 836 | 1,452 | 747 | **538,282** | 7.50% |
| **Aberdeen & Aberdeenshire** | 225 | 48 | 881 | 8,374 | 666,365 | 1,197 | 11,781 | 261 | **689,132** | 9.60% |
| **South East Scotland** | 1,341 | 1,410 | 45,582 | 709 | 1,051 | 1,716,602 | 37,043 | 37,443 | **1,841,181** | 25.70% |
| **Dundee, Perth & Angus** | 364 | 133 | 5,163 | 1,511 | 10,651 | 39,892 | 577,127 | 7,944 | **642,785** | 9.00% |
| **Stirling, Crackmannanshire & Falkirk** | 634 | 206 | 34,007 | 780 | 263 | 35,518 | 7,764 | 348,219 | **427,390** | 6.00% |
| **Totals** | **461,441** | **187,376** | **2,392,075** | **538,753** | **687,465** | **1,828,086** | **639,035** | **429,662** | **7,163,893** | |
| | 6.40% | 2.60% | 33.40% | 7.50% | 9.60% | 25.50% | 8.90% | 6.00% | | |

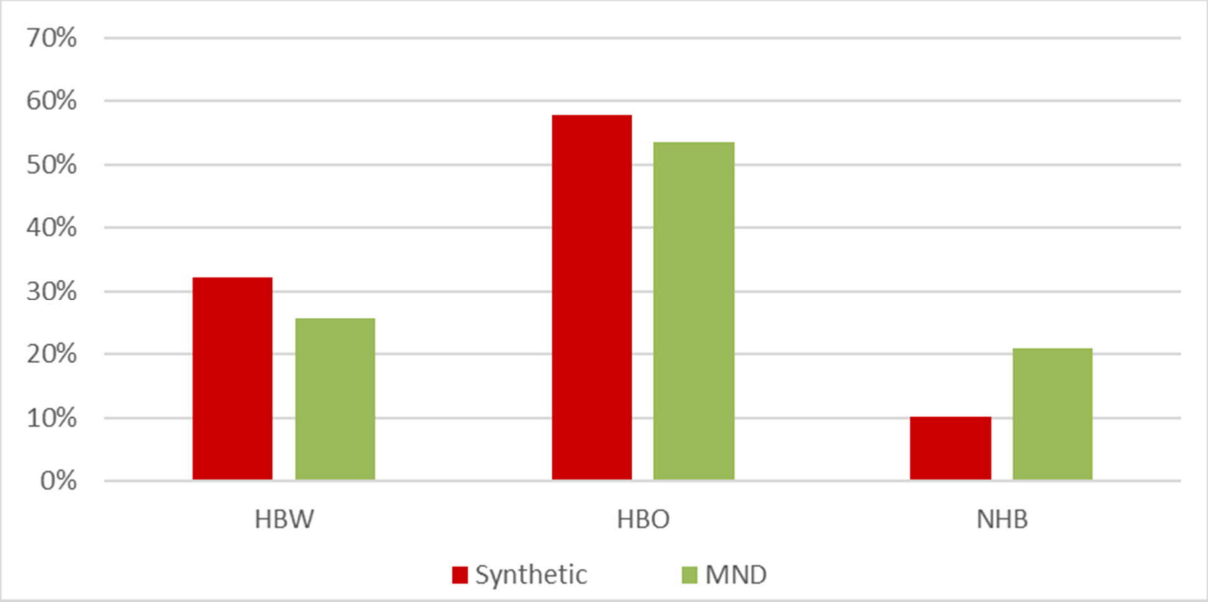
## Table 5.3 Synthetic Sector to Sector Trips at TMfS Regional Level

| Sector | Ayrshire | Dumfries & Galloway | Glasgow & Clyde Valley | Highland & Argyll & Bute | Aberdeen & Aberdeenshire | South East Scotland | Dundee, Perth & Angus | Stirling, Crackmannanshire & Falkirk | Totals | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Ayrshire** | 354,236 | 2,659 | 58,774 | 1,853 | 16 | 1,532 | 149 | 970 | **420,188** | 5.9% |
| **Dumfries & Galloway** | 2,873 | 154,367 | 4,332 | 26 | 6 | 2,167 | 45 | 180 | **163,995** | 2.3% |
| **Glasgow & Clyde Valley** | 53,977 | 3,839 | 2,029,304 | 6,328 | 233 | 41,404 | 3,085 | 38,782 | **2,176,952** | 30.4% |
| **Highland & Argyll & Bute** | 1,978 | 26 | 7,047 | 528,301 | 10,885 | 943 | 1,557 | 761 | **551,496** | 7.7% |
| **Aberdeen & Aberdeenshire** | 17 | 6 | 245 | 10,384 | 673,514 | 801 | 6,630 | 157 | **691,755** | 9.7% |
| **South East Scotland** | 1,465 | 2,013 | 42,481 | 840 | 742 | 1,553,529 | 31,294 | 37,473 | **1,669,837** | 23.3% |
| **Dundee, Perth & Angus** | 149 | 42 | 3,157 | 1,373 | 6,105 | 30,219 | 517,412 | 7,429 | **565,887** | 7.9% |
| **Stirling, Crackmannanshire & Falkirk** | 889 | 153 | 39,067 | 667 | 140 | 36,225 | 6,847 | 335,500 | **419,489** | 5.9% |
| **Totals** | **415,584** | **163,105** | **2,184,407** | **549,773** | **691,641** | **1,666,819** | **567,020** | **421,252** | **6,659,601** | |
| | 5.8% | 2.3% | 30.5% | 7.7% | 9.7% | 23.3% | 7.9% | 5.9% | | |

## 5.8 Purpose Split

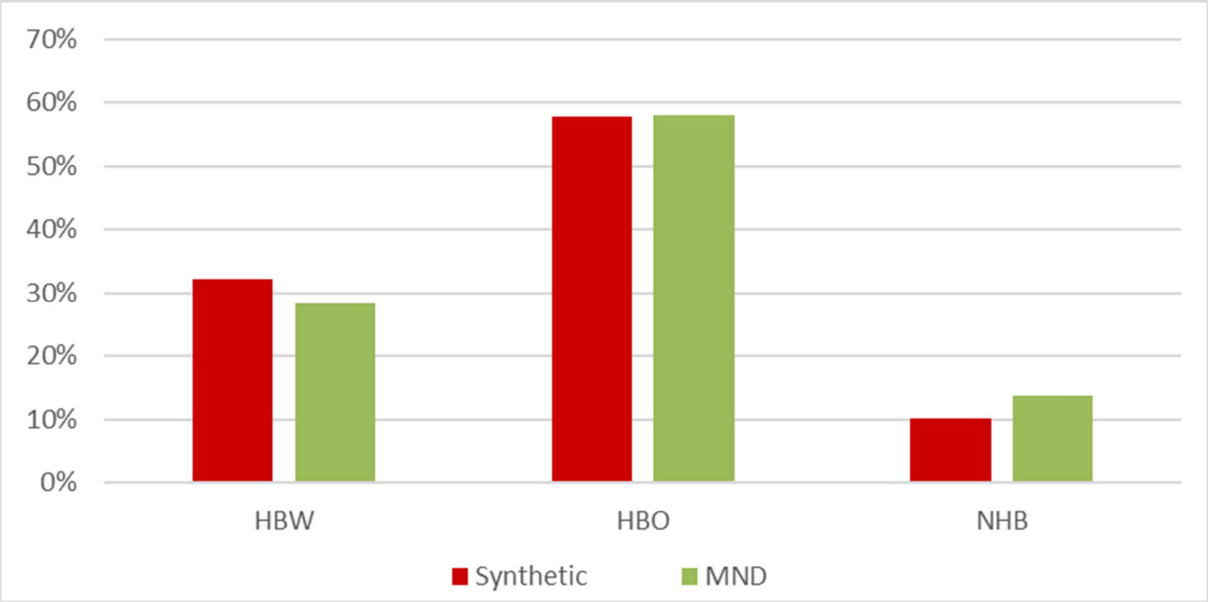
5.8.1 Figure 5.15 compares the purpose split for the internal productions at the initial stage (Step 1). In relation to the synthetic matrices, the MND data has more home-based trips and less NHB trips.

Figure 5.15 Comparison of Purpose Split at Step 1



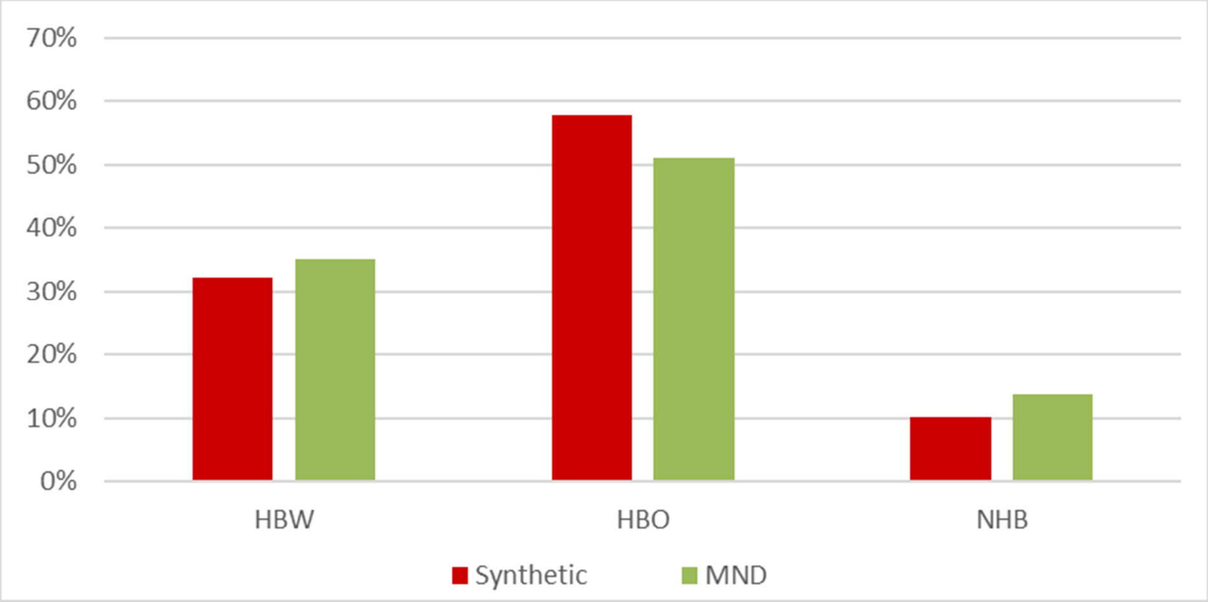

5.8.2 Figure 5.16 compares the same proportions following the removal of freight and bus trips (Step 3) and demonstrates how the purpose split has aligned itself closer to the synthetic purpose split, as a result of a higher proportion of NHB trips having been removed. However, it is noted that the HBW proportion among MND data is somewhat lower than the corresponding proportion for synthetic data.

Figure 5.16 Comparison of Purpose Split at Step 3



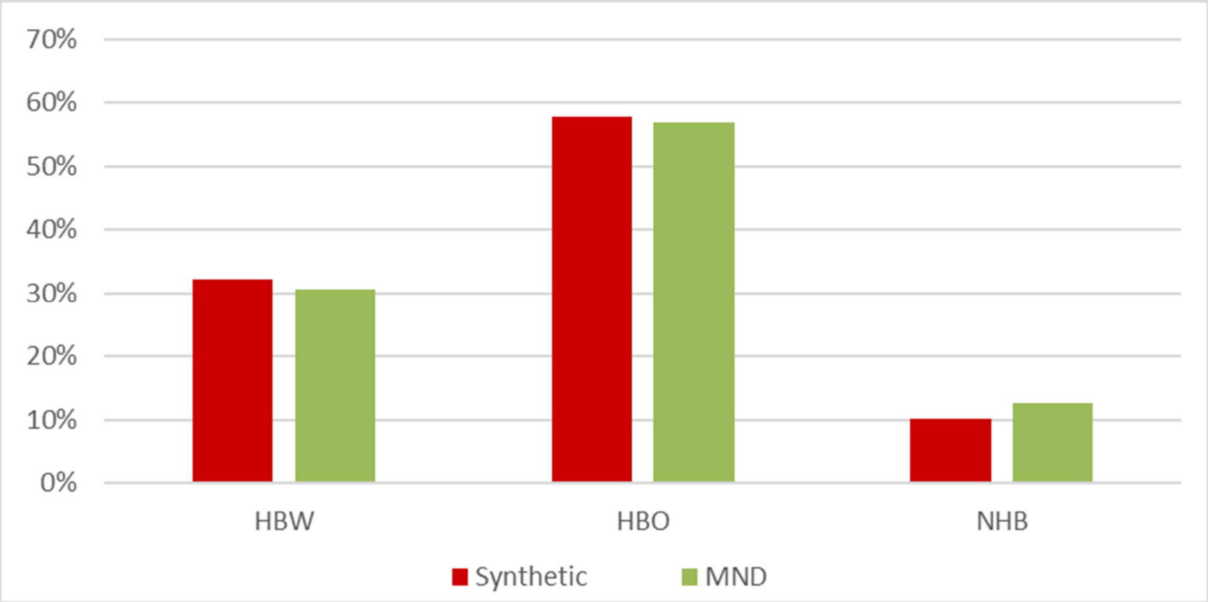

5.8.3 Figure 5.17 compares the purpose split following the adjustment of home-based purposes. While the proportion of HBW trips has been increased as intended, the graph suggests that it has exceeded the HBW proportion for synthetic matrices. This is due to the adjustment being carried out excluding short trips, which are under-detected in MND data and that they will be replaced in the following step.

**Figure 5.17 Comparison of Purpose Split at Step 4**




5.8.4 Figure 5.18 compares the purpose split for the internal productions in the final step. In contrast to the initial MND data, MND purpose split proportions are more aligned to the synthetic ones, with the exception of a slight over-representation of NHB trips. This, as discussed previously, can be related to the uncertainties in the NHB trip end data and the low quality of freight secondary matrices.

**Figure 5.18 Comparison of Purpose Split at Step 5**



# Section 6 – Summary and Recommendations

6.1.1 The matrix development process presented in the earlier sections has achieved two main objectives:

- verification of initial MND matrices (as received by Telefonica) and specification of the appropriate MND sectoring system for reprocessing of the MND matrices; and
- establishment of aggregate trip patterns.

6.1.2 The latter included processing of secondary data sources, detailed verification process, synthetic matrix development, vehicle split, and short-distance trip infilling. This process, as it has been described in detail, produced matrices by aggregate purposes, at all-day level, in the MND sectoring system. These matrices reflect trip patterns which match reasonably well the observed trip length profiles sourced from SHS by area type (i.e. metropolitan, urban and rural), and their associated trip rates are statistically the same as the observed trip rates from the SHS, and the Scottish Trip End model.

6.1.3 Reflecting on the outcome of verification tests and processing of the data, the produced matrices are particularly expected to be more reliable, and provide more confidence, in representation of longer distance, inter-urban aggregate trip patterns, particularly for home-based car person trips.

6.1.4 It should be noted, however, that some weaknesses in the matrices are expected to exist, partly reflecting lack of more reliable complimentary data sources, and partly a reflection of the reduced scope of this exercise. Further processing and adjustment of the matrices and better-quality secondary datasets are recommended to be used to address these. In particular, use of more reliable freight and bus matrices, if available, in the vehicle split process is expected to result in a more accurate exclusion of freight trips from the MND road matrices and in particular better representation of the NHB car trips, especially in the Scottish rural areas.

6.1.5 In addition to the above, further processing of the matrices and extra steps are recommended in order to further reduce residual errors in the matrices, introduce further spatial and temporal granularity and demand segmentation to the matrices, and produce highway prior matrices suitable for use in the TMfS highway assignment and demand model (see Paragraph 1.1.5).

# Section 7 – Bibliography

DfT, (2020b). *Discrepancies between Road Freight and Road Traffic HBV traffic estimates.* London: Department for Transport
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/527810/hgv-traffic-estimates-report.pdf [Accessed 20 Jan. 2020].

DfT, (2020a). *TAG Unit M2.2: Base Year Matrix Development*. London: Department for Transport.
Available at:
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/841080/tag-unit-m2-2-base-year-matrix-development-november-2019-forthcoming-change.pdf [Accessed 20 January 2020].

Feldman, O., Forero-Matrinez, J., Coombe, D. (2012). Alternative Gravity Modelling Approaches for Trip Matrix Synthesis. In: *Proceedings of the European Transport Conference*. Glasgow: European Transport Conference.

Psarras, S., Prince, R., Tolouei, R. and Tolofari, S. (2017). A Hybrid Gravity Modelling Approach for Trip Matrix Synthesis. In: *Proceedings of the European Transport Conference*. Barcelona: European Transport Conference.