FAME
BASE YEAR MATRICES THINK PIECE

FINAL VERSION

By

John Bates (lead)

Ian Williams

For

Department for Transport

Version	*Date*	*Author*	*Notes*
1.0	22/1/2026	Ian Williams	Partial early Draft
2.0	26/01/2026	Ian Williams	Expanded draft now with TNR font.
3.0	02/02/2026	John Bates	Further expansions with equations
4.0	09/02/2026	Ian Williams	Chapter 5 expanded
5.0	10/02/2026	John Bates	Draft Final
6.0	26/02/2026	Ian Williams & John Bates	Final

FAME – BASE YEAR MATRICES THINK PIECE

Contents

Tables

FIGURES

FAME – BASE YEAR MATRICES THINK PIECE

1 INTRODUCTION

1.1.1 This think piece was commissioned by DfT to make recommendations on the development of base year matrices for a new national freight transport model, to support the Department in the preparation of specifications for the Freight Analysis and Modelling Environment (FAME), a new cross-modal freight analysis environment. It is a joint effort by Ian Williams and John Bates, both independent consultants.

1.1.2 The recommendations will concentrate throughout on providing further detail or alternative suggestions to what is provided within the FAME Feasibility Report (FFR) by MDST (2024) that already covers many aspects well. It will cover additional items as well as those previously included within Ian's FAME Peer Review (FPR) of early 2024, updating and fleshing out key aspects on methodology and data improvements that had been introduced there.

1.1.3 In contrast to the FFR recommendations, for the reasons discussed in Chapter 2 below we believe that it is essential to begin with the concept of PC matrices when estimating base year OD matrices. Providing an explicit representation of the supply chains that move shipments of goods from producers to consumers enables a consistent representation of the underlying economic logic that underpins the sequence of linked delivery legs that generate the observed OD pattern of goods movements. This supply chain representation is in addition to the need to explicitly represent intermodal chains wherever an OD movement may transfer between road, rail or maritime at intermediate terminals/ports traversed between its origin and destination. We direct our thoughts firstly towards the estimation of the P and C "ends", together with the appropriate level of spatial and commodity segmentation, and then the connections between them to establish the matrices, covering both domestic and international flows.

1.1.4 It is expected that the Base Matrices will be an important element in the overall model forecasting procedure, which is the subject of a separate think piece. In addition, however, there will need to be an interface with what we will refer to as the National Passenger Transport Model [NPTM] in the form of freight vehicle O-D matrices, since – unlike other modes – travel times on the highway network are dependent on **all** components of road traffic.

1.1.5 In Chapter 8, the main conclusions from this think piece are summarised and the priorities are listed for the tasks that address the various data gaps that have been determined.

2 THE NEED FOR A SUPPLY CHAIN BASED MODEL

2.1.1 The methodology for the design of the National Freight Transport Model [NFTM] in FFR Section 9 recommends to avoid the use of agent-based modelling. This seems appropriate, given the large study area size at which the NFTM must operate, which would create long model run times to ensure that stable results are generated. Furthermore, the risk factor involved in developing an agent-based model could be substantial.

2.1.2 The proposed adoption of a four-stage model structure seems a more appropriate approach, though within this structure the specific stages of the O-D distribution model would benefit from substantial refinement beyond the relatively simplified treatment provided in FFR, for the reasons now discussed.

2.1.3 It is crucial when modelling freight transport demand that the main behavioural-economic factors that influence the scale and the spatial pattern of freight modal OD matrices can be represented appropriately. The main influences on these modal OD matrices for a specific type of commodity are:

a) The volume of the commodity that is produced in the factory, quarry, farms in the zone
b) The volume of the commodity that is consumed by those living or working in the zone (Final Demand), as well as in inputs to its local industries (Intermediate Demand)
c) The volume of the commodity that is warehoused temporarily within the zone as a stage within its shipment from production to consumption
d) The volume of the commodity that is transhipped between vehicles or modes at specific ports, rail terminals or container yards.

The elasticity of demand with respect to the generalised cost of transport is quite different across this set of influences.

2.1.4 The transhipment relocation elasticity is relatively high, because for shippers it may be fairly easy to switch their transport providers between successive shipments if there are cost reductions to be obtained. For rail and maritime movements some contracts may extend over a year or more due to the high vehicle/vessel costs involved but in general major transhipment location changes are feasible within a few years at most.

2.1.5 The distribution centre relocation elasticity is intermediate. As can be observed in recent decades in the UK, it is relatively quick and not that expensive to construct major distribution centres where they are in current demand. New concentrations of distribution centres are created in locations that suit the current configuration of the road, rail and port networks based on optimising overall distribution costs across supply chains. Obsolete warehousing is then discarded.

2.1.6 The production and consumption relocation elasticities are relatively low for most commodity types, except in those few cases where transport cost is a major contributor to the overall consumption cost of the good. This is most typical for relatively low value bulk goods, such as cement works and power stations that are located close to quarries and to raw material sources. Otherwise for most other types of activity the rate of relocation due just to transport costs will tend to be much less important than that due to various other external influences. Many factories will require a local skilled or experienced workforce plus a network of local service suppliers, so that major industrial relocations tend to be complex and expensive.

2.1.7 Accordingly, we are not convinced that the proposed gravity model approach in FFR Section 9.3 for the estimation of O-D all-mode matrices would be adequate. We would strongly recommend the adoption instead of an alternative supply chain based approach to generate the required OD matrices of goods transport. This is necessarily a somewhat more complex model structure because it provides an explicit representation for each individual commodity type of

- Production-Consumption (P-C) supply chains that trace goods along distribution channels each comprising distinct logistic O-D legs from the original production zone through a sequence of national, regional and local distribution centres up to the final consumption zone,
- and of intermodal chains, where an O-D leg may transfer between road, rail or maritime at intermediate terminals/ports traversed between its origin and destination.

2.1.8 This would greatly increase the logical coherence of the demand estimation and would improve the subsequent stages of mode choice and of vehicle type choice. It would increase its ability to respond plausibly to policy tests and to provide improved future forecasts of the pattern of demand change by O-D. It is the approach that had been used in the Department's BYFM model (WSP, 2011) and in the TRIMODE model of freight transport across Europe (Noekel et al., 2018).

2.1.9 Introducing this supply chain and intermodal chain functionality may require some software development but in the age of Python and of other software development aids, this software development, if needed, should not be a huge or a very high risk task. Alternatively, the Department could use the MEPLAN software package that it has used for many years for passenger demand estimation within NTMv2. MEPLAN includes all of the functionality required for supply chain and intermodal modelling, as previously demonstrated within the BYFM project.

3 OVERALL CHARACTERISTICS OF THE MATRICES

3.1 OVERVIEW

3.1.1 This chapter focuses on the segmentation characteristics that will be used in the various stages of the construction of the base year OD matrices. The majority of these characteristics match closely to the segmentation details proposed originally in the FAME Feasibility Report, as summarised below in Table 1. In those cases where alternative or extended approaches are recommended, they are discussed further in the sections below to explain why they provide a preferred approach.

TABLE 1 SUMMARY OF FFR RECOMMENDATIONS ON MATRIX CHARACTERISTICS

FFR Section	Type	Comments / links to Sections in this report
	Base year	See 3.2
7.2	Exclude LGV non-delivery trips	OK see 3.3
	Include tonnage for all goods	OK
7.3	Set of modes	OK see 3.4
7.4	Modelled area	OK see Northern Ireland discussion in 3.5
7.5	Zoning	OK see 3.6
7.6	Time period	OK see 3.7
7.7	Road vehicle types	No – model should use 5 HGV types, see 3.8
	Other mode vehicle/vessel types	OK
9.3	NST 20 commodity types	OK see 3.9

3.2 BASE YEAR

3.2.1 The base year to adopt should be one where the transport patterns are relatively stable and where the main statistical data sources are not in flux.

3.2.2 Accordingly, the COVID period should be ruled out, which would also remove 2021, the year in which the DfT's Continuing Survey of Road Goods Transport (CSRGT) data collection methodology was switched from paper to online, which introduced a break in the data series.

3.2.3 It should also avoid the turmoil created by the various stages of the BREXIT introduction of the full rule changes governing imports and exports to Europe. For example, BREXIT appears to have substantially reduced the tonnage carried on accompanied HGVs to and from Europe, leading to a gradual change in the spatial pattern of international road goods traffic, switching tonnage from accompanied HGVs through the Dover Strait to use instead unaccompanied trailers via the North Sea ports.

3.2.4 These issues suggest that base years of 2024 or 2025 would be most suitable, though it should be noted that there are substantial delays in the availability to the UK DfT of the foreign registered HGV OD movement data that is assembled by the EC.

3.3 REPRESENTING LGVS

3.3.1 It is assumed that the estimation of LGV trips associated with the delivery of goods is to be part of the NFTM, whereas other non-goods delivery movements in LGVs should be part of the National Passenger Transport Model (NPTM). Accordingly, it will be important to ensure that the LGV non-delivery O-D (passenger) matrix that is input to the NTMv5 does not have gaps or double counting with respect to the LGV delivery O-D (freight) matrix that is produced by the NFTM. It should also be noted that the representation of LGV trips within the NTS[1] has a number of serious gaps that have been identified by the excellent analysis provided by Le Vine et al. (2013). Their insights may help in minimising gaps and overlaps within the representation of LGV trips for individual purposes.

3.4 SET OF MODES

3.4.1 We agree with the FFR recommendation with respect to modes

> "The FAME should be multimodal (to allow competition between modes) and intermodal (to allow, where appropriate, cooperation mainly between road and other modes). It should actively model road (HGVs and LGVs), rail, shipping and inland waterway freight transport and also model air freight and pipeline traffic at the point at which it generates road and rail movements." (FFR, p.51)

3.5 REPRESENTING NORTHERN IRELAND

3.5.1 The FFR correctly identifies the need to include all RORO ferry services to the Island of Ireland, from GB and elsewhere, as part of including all UK freight movements within, to, from and through NI, as well as between NI and the rest of the world

3.5.2 Section 2.7 of FFR briefly discusses the issue of whether the model could also cover Northern Ireland on an equal basis with the rest of the UK. A full representation of internal traffic within NI implies that it would further be necessary to represent zones within Ireland (Irl) in considerable detail in areas within the northern half of that country in order to represent road freight between Irl and NI and through NI to/from Scotland. It will also be important to represent the Irish domestic road freight movements that transit through NI, particularly along the A5 Western Transport Corridor, for which funding of €600m has in February 2024 been

[1] Offsetting the incomplete and variable representation of LGV trips by purpose and in trends through the years within the NTS is important for the NPTM, because the NTS is the data source underpinning the NTEM trip rates by trip purpose that support the NPTM and other UK passenger models.

committed by the Dublin Government towards the long awaited A5 55-mile upgrade within NI, though subsequent legal challenges have slowed its progress.

3.5.3 The complexity with extending the model to include NI fully within the internal study area, lies in the very substantial existing differences in many trade data sources, freight data sources and data definitions between each of GB, NI and Irl. Accordingly, the effort required to assemble all of the required data in a suitably consistent form for the relatively small number of zones in and around NI, is likely to be comparable to the effort originally required to model GB as a whole, rather than being proportional to the number of zones used on the Island of Ireland.

3.5.4 It will be important when funding the estimation of base year matrices within FAME to think through the resource implications for model development of including NI as an integral part of the internal modelled study area. The strategies for inclusion of NI within the NFTM have been discussed further in Section 6 of the FPR.

3.6 ZONING SYSTEM

3.6.1 The FFR recommendation for a two tier zoning system seems sound from the viewpoint of achieving reasonable run times.

3.6.2 The upper tier 2 which is proposed to be at the local authority district level would need to also append the point zones that are needed for major freight nodes, such as ports, airports, warehouse concentrations, inland tank storage sites that are linked to pipelines, etc. This would add in another 100 to 150 point zones in addition to the specified total of 374 zones on the tier 2 fast turnaround model. This is the zoning level at which the earlier BYFM model had been successfully developed and run.

3.6.3 The overseas zones should be as in FRR:

> "Geographic zoning outside the UK should be at a less detailed geographic level, based on ITL 3 for the Republic of Ireland, ITL 1 for European regions (plus point zones for ports), and world regions for locations beyond Europe" (FRR, p. 54)

3.6.4 The CSRGT[2] indicates for 2024 that the average length of haul by HGV type varies from 55 kms for large rigids up to 134 kms for articulated HGVs. This indicates that even at the LA district zoning level, most HGV trips are likely to be interzonal, rather than being intrazonal or between adjacent zones. Accordingly, when using the HGV OD matrices as input to a local model, any disaggregation of the HGV OD matrices from their original LA district based zoning system to the required finer spatial zoning system could be carried out relatively simply, through spatially disaggregating the origin zones and the destination zones with suitable proportional factors for local Os and for Ds. There is unlikely to be a major need to rerun a local gravity model to take account of local accessibility impacts because their influence is likely to be relatively minor.

3.6.5 In contrast, the conversion of the OD matrix for LGV delivery trips from the national to a finer scale local zoning system is likely to be less straightforward, due to the much more localised trip patterns that are typical for LGVs.

[2] CSRGT Table RFS0108.

3.7 TIME PERIOD

3.7.1 The basic time unit of yearly freight OD flows seems appropriate. However, blanket conversion factors from annual to period in an average day for HGVs and LGVs are not appropriate because each type of road vehicle and associated usage type has distinctive characteristics.

3.7.2 For HGVs, it will be more productive to use detailed factors segmented by vehicle type and size. The time of day and day of week patterns may be very different across the HGV size/type vehicle classes listed in the next section. The factors further differ within each vehicle class by road type, so this may also be helpful to apply within congested assignment equilibrations. Annex D provides evidence of some of these differentiations.

3.7.3 It should be feasible for DfT statistics to create sets of current year tables (analogous to the historic 2018 data illustrated in Table 6 of Appendix D) that are segmented by the combination of the set of modelled goods vehicle types by the set of standard road types. Each set of tables would contain factors to convert from annual vehicle flows to, say: (1) specific months; (2) average week; (3) AADF; (4) individual hour of day for each individual day of week; (5) etc. There would be separate sets of tables for each of the road type segments and for each of the vehicle type segments, as well as for each pair of these two types of segments (e.g. artics on motorways). These tables could be used, as appropriate, to scale traffic to the periods of interest. The factors segmented by vehicle type could be applied to the annual OD matrices, which are segmented by vehicle types. The factors segmented by vehicle type by road type could be applied to the GV traffic assigned to an individual link.

3.7.4 The daily pattern of LGV movements for the parcel delivery purpose will be spread relatively evenly across the working day, though with an early start and may be reduced at weekends. In contrast, for grocery deliveries movements are over two shifts, continuing into the late evening and across weekends. Turning to passenger LGV trips, commuter and to a lesser extent in course of business van trips will be more concentrated in the morning and evening peaks. Leisure trips in vans are more common outside the working day or at weekends.

3.8 VEHICLE TYPES, COSTS AND NETWORKS

3.8.1 The proposed usage of just two HGV types (OGV1 and OGV2) within FRR, in addition to LGVs, does not appear to be appropriate. Instead at least the 5 HGV classes (4 rigids[3] by size, plus all artics) system used in GBFMv5 and in BYFM and available within DfT road freight statistics, should be adopted within the construction of the base year OD matrices. This distinction may also be helpful later within the assignment, where vehicles above a certain size may have different costs or access restrictions associated with them on specific types of links.

3.8.2 The advantages to this more detailed HGV classification approach include:

- Strong association between logistic leg types and vehicle class
- Improved conversion from tonnes to vehicles
- Better estimates of fuel consumption volumes

[3] These rigid vehicles are: Over 3.5t to 7.5t", Over 7.5t to 17t", Over 17t to 25t", Over 25t"

- Better representation that the usage of large artics in inner city areas is small, so the larger number of smaller GVs required to carry a given tonnage, gets picked up in the link congestion measurements.
- Improved representation of alternative fuel options. Small rigids moving mainly on short journeys currently are better suited to battery electric propulsion options than large rigids used for long distances.

3.8.3 The FFR suggests that the use of
"(say) five different types of HGV, would be possible based on the baseline data from CSRGT, but would make the NFTM more complex and may lead to relatively spurious or even inexplicable responses." (p.56)

3.8.4 We disagree that this is necessarily the case. Provided that the appropriate vehicle operating cost functions are applied and that there is a coherent representation of the different logistic leg types within the supply chain for individual commodity types, then the representation of the split across vehicle types can be represented in an explicable fashion. There are various suitable sources[4] for current data on the set of individual component vehicle operating cost elements (e.g. fuel, drivers' wages, insurance, depreciation, etc.) for a cross-section of goods vehicles of different sizes.

3.8.5 Annex D below assembles data that identifies some of the important differences between HGV types in their road usage characteristics by road type, time of day and day of week, as well as by trip length and in their average loads across types of goods.

3.8.6 For rail mode, two vehicle types would suffice, intermodal and bulk.

3.8.7 For accompanied HGVs, all domestic and international ferries plus Eurotunnel should be treated as part of the road network, rather than as a maritime vehicle type. Similarly, international rail freight services through Eurotunnel are part of the rail network.

3.8.8 For the maritime transport mode, distinguish vessel types between:
- Other domestic coastal shipping
- Short sea unaccompanied RORO
- Short sea LOLO
- Deep sea unitised

3.8.9 The typical vessel sizes and vessel route organisation are very different between short sea container services to Europe and deep sea services to further afield, which is why these benefit from being distinguished.

3.8.10 We agree with the FFR set of recommendations: on inland waterway, to explicitly represent short sea unitised and deep sea container services, while for bulk goods, to model only their inland distribution from the UK ports relevant to that type of bulk.

3.8.11 These recommendations would require the development of cost models by vehicle type, as explained in the FFR, for each of the required modal vehicle/vessel types that have been

---

[4] For GV operating cost components see for example:
MotorTransport (https://motortransport.co.uk/knowledge-zone/cost-tables);
RHA (https://www.rha.uk.net/Membership/Annual-Cost-Movement-and-Pay-Surveys);

included above, together with the appropriate road, rail, inland waterway and maritime networks for use by these vehicles/vessels. For road, the existing NTM network would be appropriate.

3.9 COMMODITY TYPES

3.9.1 It would be beneficial to distinguish PC matrices for a commodity type by whether they are destined to intermediate consumption or to final demand[5], because these will have separate economic trends, geographical distributions and supply chain options, due to their likely difference in typical shipment size and in annual door-to-door shipment volume. This latter characteristic is a major influence on the cost-effectiveness of freight rail service provision.

3.9.2 For similar reasons, distinguishing PC matrices between domestic movements and international movements (e.g. imports and exports) is helpful.

3.9.3 As recommended in FFR, we should use the NST2 commodity classification (20 commodity groups) that is collected within the CSRGT. This segmentation is appropriate for use within the demand model stage, prior to the eventual conversion of OD matrices from units of tonnes to units of vehicles for use within the assignment stage. The typical trip lengths and types of modes and of vehicles used differs between commodity types, as does the likely future growth rate in demand for each commodity type. It is for such reasons that this segmentation is necessary within the demand model.

3.9.4 In summary, this recommendation would generate a maximum of 80 PC matrices (commodity type 20 x intermediate/final 2 x domestic/international 2).

3.10 LOGISTIC LEGS

3.10.1 Although some bulk primary products, such as ores or grains, may be transported directly without intermediate storage from where they are produced to where they are consumed or processed, this direct transfer is not the norm for most goods. Instead, foodstuffs and consumer goods usually pass through many intermediate distribution centres within the supply chain that connects the initial producer of the goods through to their final consumer.

3.10.2 To represent the succession of transport movements from the producer to the consumer within supply chains, we distinguish three logistic leg types (primary, secondary and tertiary) as in BYFM. These leg types occur within a variety of supply chains that may contain different numbers and combinations of logistic leg types. Further context on supply chains and logistic legs is provided below in Annex C.

4 POTENTIAL DATA SOURCES

4.1 INTRODUCTION

4.1.1 The biggest challenge in freight demand model development is not the model design and construction itself but rather the assembling of sufficiently consistent datasets that cover all

[5] Within standard economic Input-Output tables, *intermediate consumption* is where the goods are consumed by industries as inputs to the production of other goods (e.g. iron ore to make steel), whereas *final demand* is where the goods are consumed directly by the population or in investments, etc.

modes to a level of detail that then enables the model to be successfully implemented and calibrated. While freight data availability has improved somewhat through time in Great Britain, it is still not to a level that is adequate to provide really strong support for freight demand model development nor for insightful analysis of potential impacts of freight transport policy measures. Recent improvements in AI functionality could potentially reduce the cost of assembling suitable supply chain data, through providing suitable linkages between existing data sources, but this is an avenue that may still require substantial further development.

4.1.2 This chapter focuses on the more general challenges associated with obtaining comprehensive and consistent freight and logistics data. Data sources relating to the volume of production and consumption are discussed in more detail in Chapter 5.

4.2 THE PHILOSOPHY OF DATA FUSION

4.2.1 Before discussing the data, it will be useful to consider the underlying philosophy regarding data fusion. To give a simple example, we will briefly describe the much more straightforward example of constructing person trip matrices on a Production/Attraction [P/A] basis. We note that in any data fusion process, it is essential, when dealing with sample data, to take account of the underlying statistical properties: while these will often be quite complicated, it is not our intention to discuss them in any detail.

4.2.2 From sources such as household surveys and roadside/on-train interviews, it is possible to obtain estimates of cells of the matrix, typically segmented by trip purpose, but these will rarely be sufficient to fill the complete matrix. It therefore becomes necessary to estimate the missing cells, and the standard way to do this is to construct a model which attempts to explain the underlying matrix structure. Once this is achieved, the accepted scientific approach would be to look for independent data against which the matrix could be validated.

4.2.3 There are many possible sources of such data, but a commonly chosen one relates to vehicle counts on the network. This of course involves a further "model", relating to how vehicles choose paths through the highway network. It will also be noted that count data contains no information about purpose.

4.2.4 Suppose we have estimated highway Production/Attraction [P/A] matrices T_{ij}^{p} by purpose p. From an assignment (which will normally use a User Equilibrium approach based on Wardrop's Principle) we get a set of paths $\varepsilon^{*}_{a|ij}$ where a is a link, and i and j are zones. In practice, this means that a conversion from P/A to O-D is carried out[6], and the resulting highway matrix A_{ij}^{p} assigned to the network. The outcome flows V_a on any link a can be compared with observed counts. Based on the above description, we can write: $V_a = \sum_{ij} \sum_{p} A_{ij}^{p} . \varepsilon^{*p}_{a|ij}$ where ε^{*} gives the optimum proportion[7] of A_{ij} allocated to link a.

[6] See the discussion in 2.5 of TAG Unit M2.1

[7] It should be noted that while in Wardrop equilibrium the link flows are unique, the path flows are not. Hence the ε^{*} factors are just one possible solution out of an infinite number.

4.2.5 In practice, the reliability of ε^* depends on: a) the underlying theory of user equilibrium, b) the link cost functions, c) the coding of the network, and d) the level of demand as represented by A_{ij}, since this affects the level of congestion. In addition, we have simplified here for the purposes of presentation by omitting any reference to user class (which relates both to the demand matrix, the link costs, and in some cases differential access to the network).

4.2.6 Given all this, including the generality of the modelling processes, it would be unwise to expect a high degree of consistency between the outcome flows and the observed counts, even if there was a substantial level of confidence in both the demand matrices and the counts themselves. WebTAG Unit M3.1 recognises this in principle, and proposes three separate tasks (Network Data, Coding & Checking; Network Calibration & Validation; and Route Choice Calibration & Validation) with the aim of removing possible errors. However, these tasks are unlikely to impact much on either a) or b).

4.2.7 Problems occur when, after reasonable adjustments to the network, it is concluded that significant validation differences remain which are essentially attributable to the matrix. Ideally, further data should be introduced to the whole modelling procedure in such a way that the base P/A matrix is modified to be consistent with any changes to its O-D counterpart. Unfortunately, there is very little documented experience of how to do this.

4.2.8 An alternative approach, widely used, is to modify the assignment (O-D) matrix, using various techniques that can be referred to as 'matrix estimation from counts' [ME]. At its most general, this consists of a model form for the (estimated) matrix $\tilde{A}_{ij}$, the parameters of which are estimated under a selection of constraints, which may relate to row and column totals, totals within different sub-matrices or total traffic flows across defined screenlines (thus introducing network elements relating to paths). Furthermore, they may range from 'hard' equality constraints, which must be satisfied, to 'soft' inequality constraints, which may be downgraded by assigning them lower weights or bounds (taken to reflect the modeller's estimate of confidence in them).

4.2.9 Early applications of this process were relatively unconstrained, and led to suspicions that it could produce quite unrealistic matrices purely to align with the traffic counts. For this reason, the existing Guidance on Highway Assignment (TAG M3.1) imposes strict limits on the amount of change to the matrices brought about by ME[8] . In our view, while recognising that the ME procedure needs to be carefully controlled, the guidance is far too strict, and we incline strongly to the viewpoint expressed in the Arup (2016)[9] Report [p 30]:

> "Overall, the recommended perspective on validation is that one should report on how well the model relates to a whole range of mutually incompatible measures and then consider the results in the context of the objectives for the model. The objectives will only rarely be about use of the model to give good traffic flows in the base year, so the times when count validation is paramount should be few. In general, it is not recommended that some count data should be held back for independent validation. Thus, model calibration and validation can be considered to be a combined process, with reporting on validation still a key element."

---

[8] "8.3.4 Matrix estimation should not be used solely to achieve the flow validation criteria at the expense of significantly changing the demand matrix patterns" [TAG M3.1]. Similar remarks apply to Public Transport in §8.2 of unit M3.2.

[9] Provision of technical advice and support for matrix building guidance

[p 51] "Ideally, given an appropriate model form based on theory, the parameters would be populated by a holistic estimation process which takes account of all the available data, with due recognition for any inconsistencies, biases and the statistical accuracy of the different sources. While some attempts have been made along these lines…, they involve both theoretical and computation effort which is currently too advanced for most cases of standard practice. There is therefore a need to develop approaches which are practical but which nevertheless maintain some of the necessary rigour."

4.2.10 The conclusion is that we should downplay the requirements for validation, and find appropriate ways to make use of as much of the varied data sources as possible. We accept, of course, that the lack of validation means that other approaches are required to build confidence in the matrices.

4.2.11 It is therefore necessary to prepare clear and convincing documentation, to show how the matrices have been derived, including the way in which any model parameters have been estimated and their level of statistical confidence, the statistical properties of the various types of data which has been used, and the care with which the matrix construction has been carried out. Sufficient description is required to enable a thorough audit of the procedure, which should, in principle, be repeatable.

4.2.12 In addition, the essential "reasonableness" of the matrices will need to be demonstrated, in terms of, for example, average trip length by mode, concentration of movements with areas of identifiable relevance (factories, warehouses etc.). Overall, the aim should be to show that the matrices are consistent with the available evidence and that any extrapolation to parts of the matrices where data is not available can be rationalised.

4.3 IMPROVING DATA TO QUANTIFYING SUPPLY CHAIN LINKAGES

4.3.1 The belief of the authors is that it is better to construct an appropriately structured freight demand model, even if that may need to be implemented in a form that exhibits some data weaknesses, rather than constructing a misleadingly over-simplified model using only the available restricted datasets. More specifically, a freight model that does **not** have an explicit representation of the freight supply chain system and of the linkages between successive logistic legs, is **unlikely** to be helpful for policy testing for many of the aspects of interest to the Department.

4.3.2 In contrast to passenger modelling, for which adequate data is generally available in the UK, freight modelling in both the UK and most other countries has struggled in recent decades to evolve to represent the major logistical changes that have been induced by the economic shift away from traditional heavy industry to an economy focussed instead on producing or importing a huge range of finished consumer products.

4.3.3 For freight modelling, the design, specification and implementation of the model structure is a significant task. Nevertheless in the UK, as noted earlier, it is not as great a challenge as obtaining access to good consistent data that is segmented by mode, supply chain stage and commodity type, with which to construct and calibrate a comprehensive, flexible freight demand model. The virtually complete absence of explicit data to quantify current supply chain structures within the UK (and within most, though not all other European countries) is by far the biggest data gap for the development of a robust spatial demand model that is underpinned by behavioural-economic responses.

4.3.4 Accordingly, without improved data this gap would generate a large amount of work in order to fill this void in a meaningful fashion. For example, a major consumption of resources within each of the BYFM (WSP, 2011) and TRIMODE (Noekel, Williams & Fiorello, 2018) projects that explicitly included supply chain linkages within their freight demand models, has been the research and analysis required in order to create for each individual commodity type some coarse estimates of the actual split of observed road movements between: primary; secondary; and tertiary logistic legs; as well as to distinguish the HGV trips that act as feeder legs to or from other non-road modes. This quasi-observed, processed approximation is required for the demand model calibration procedure in order to ensure for each individual commodity type

- that the split between the alternative distribution channels within a supply chain is plausible for the goods movements from the producer zones to the consumer zones
- that the influence of changes in goods flows on rail, ship or air will impact correctly on road feeder flows to and from their modal terminals.

4.3.5 This data analysis and production task unfortunately is one where the actual research and analysis work carried out in previous projects, such as BYFM, would need to be largely reworked with the current CSRGT data. The earlier work can provide some pointers as to what to look for and on the more productive avenues to explore but nevertheless a considerable volume of careful and skilled analysis work would be unavoidable. This task would add significantly to the resource cost, timescale and risk of quality reduction of this model development project.

4.3.6 There are various options now summarised on how improved required supply chain data might be assembled. These increase in their resourcing costs and correspondingly in the quality and scale of supply chain data coverage that they would provide.

4.3.7 **Reprocess CSRGT -** The simplest and quickest initiative would be to reanalyse the CSRGT data for a few representative recent years so as to attribute a land use type (e.g. quarry, farm, distribution centre, retail, household, etc.) at each loading and unloading point. Provided that a postcode or address has been provided for this point, an automated procedure could be developed to match this post code or address to suitable domestic and non-domestic datasets of properties and their usage.

4.3.8 **Extend CSRGT** – Better information could be obtained by adjusting future CSRGT survey forms to request information on the type of activity at each loading and unloading point for every surveyed journey.

4.3.9 The detailed methodology for these two CSRGT supply chain data capture initiatives are discussed in greater detail in Annex B:11.1 below.

4.3.10 A number of alternative options are then outlined in Annex B:11.2 below, for more resource intensive, **innovative supply chain surveys**, including the examples of: the automated shipment-based truck trip diary collection exercise in the Netherlands (de Bok, et al., 2022); the ECHO shippers survey of entire supply chains in France (Guilbault et al., 2008); and shipment transport chains in Sweden, using the recurring Commodity Flow Survey.

4.3.11 In summary, there is a critical need to provide access to informative data that can identify and quantify the logistic leg linkages within supply chains. The only uncertainty relates to which of the above approaches is best placed to address this need in an effective fashion.

4.4 SURVEY OF INLAND MODE AND ORIGIN OR DESTINATION OF PORT TRAFFIC

4.4.1 The freight flow data that is collected by the three main modes of road, rail and port/maritime is collected in an entirely independent fashion by each mode. Annex A below further discusses some of the oddities in measurement units and interpretations that this independence encourages. Accordingly, it would be very informative to carry out on a common basis across modes a survey of the inland transport characteristics to and from each of the main UK ports.

4.4.2 The published official UK port statistics for the water side of UK ports have through the decades provided a high level of statistical detail and of continuity. In contrast, port statistics are almost non-existent relating to transport activity on the inland side of the port. Because ports are the locations where much of the UK competition between road and rail (and waterborne) modes occurs, they could provide an effective means for identifying those contexts where each of the inland modes is more successful.

4.4.3 The key data items to collect for a suitable sample of import and of export consignments passing through the UK ports (plus Eurotunnel) would include:

- The inland mode to or from the port
- The UK inland origin (export) or destination (import) ultimate location and postcode, ideally also the activity type (factory, quarry, warehouse, etc.) of this inland location
- The location of any intermediate intermodal terminal at which the mode is switched, plus the mode of this switch
- The commodity type (NST)
- The mode of appearance at the port
- The weight of the consignment[10] and if possible its volume m^3.
- The consignment dwell time at or around the port, between being on the inland mode and on the vessel.

4.4.4 If, as is suggested on FFR p.103, this survey can be based on a suitably selected sample of custom returns, then the precise set of data items from the list above that can easily be extracted from such returns will need to be confirmed.

4.4.5 The resulting port inland movement data would be informative on many fronts. It would be of use for each individual commodity type in informing the calibration of its mode choice model. It would help to identify the characteristics of a large proportion of the primary logistic legs for imported goods for each commodity type. The modal flows could be cross-checked against the corresponding flow estimates from the official modal statistics for each individual mode to highlight any inconsistencies in measurements or data definitions, that are peculiar to the published statistics for that mode.

[10] Clarify what is included in this consignment weight measurement in addition to the goods, e.g. any packaging of the goods themselves, pallets, ISO containers, trailers for RORO, etc.

4.5 DEFINITIONAL CONSISTENCY REQUIREMENTS ACROSS DATA SOURCES

4.5.1 It is essential to examine the consistency of data definitions across the main economic and modal freight data sources[11] and across the data sources within each mode, because addressing definitional inconsistencies is a key task when assembling sound data on which to support comprehensive, consistent freight models. Strategies for mitigating the impacts of such definitional inconsistencies will need to be developed.

4.5.2 The statistical data that is collected for each freight mode is generally designed in a form that is closely suited to the requirements and traditions of that specific mode. For those mainly working within a specific mode this approach is helpful and avoids unnecessary complications. In contrast, for those charged with assessing policies that impact across modes, it is a source of inconsistency in information that needs to be recognised and understood, if errors in data interpretation are to be avoided.

4.5.3 A small selection of examples is discussed below in Annex A, illustrating how inconsistencies in freight data definitions arise, and the confusion that they can cause for analysts unless they are resolved in a systematic and standardised fashion.

5 ESTIMATING ZONAL PS AND CS

5.1 OVERVIEW

5.1.1 This initial step calculates the zonal production and consumption – the Ps and Cs. More specifically it assembles for the base year suitably consistent data to estimate for each commodity type the annual volume, within each internal zone, of production, intermediate consumption and final consumption. For each external zone, the volumes of imports to and of exports from the UK are similarly estimated by commodity type.

5.1.2 There are two parts to this zonal estimation task.

- Estimate the monetary value of the production and of the consumption for each commodity type in each domestic and foreign zone. Within the domestic zones, this monetary value for consumption is further segmented between intermediate demand and final consumption.
- Estimate the net tonnage[12] of the production and of the consumption for each commodity type in each domestic and foreign zone.

5.1.3 The monetary value unit is important because monetary value is likely to be the more appropriate unit to use to support forecasting of freight demand in future years, through usage of future estimates of changes in GDP and international trade. It also helps in distinguishing between the intermediate and the final consumption totals within domestic zones. The use of existing UK economic Input-Output (I-O) tables or of more general Supply and Use Tables from

[11] For example, the definitions used within most statistics on road, rail and maritime goods flows are those appropriate to the individual mode and so are not consistent between individual modes, either in their classification by type of goods or in their definition of measurement weight of the shipment.
[12] "Net" denotes here the tonnage of the goods themselves, excluding the weight of any packaging, containers or of the vehicles carrying the goods.

ONS[13] provides a means of connecting the level of production of a commodity in a zone with the level of consumption of the commodities required for its production.

5.1.4 The tonnage unit is necessary because the OD matrices required as output from the demand model need to be in units of tonnes rather than money to enable them to be converted to the vehicle units that are assigned to modal networks within the transport supply model.

5.2 ESTIMATING IMPORTS AND EXPORTS

5.2.1 Since the implementation of BREXIT, data has become more readily available on the volume of trade to and from individual EU countries. Now it is feasible to download[14] HMRC annual data on each of import and export trade in a form that is segmented by

- UK region
- Trading country
- Commodity type – 2-digit SIC code 00 to 96
- Units - £ and tonnes

5.2.2 This 2-digit SITC code is a product-based classification that will need to be matched to the transport oriented NST 2007 classification used in freight transport statistics and in the model. The conversion can be derived based on the Eurostat correspondence tables[15] enabling the transposition of detailed data collected according to the Combined Nomenclature into other classifications like the Standard International Trade Classification (SITC) or the standard goods classification for transport statistics (NST 2007).

5.2.3 It is now feasible to disaggregate the trade in goods[16] down to the ITL2 spatial scale with 41 sub-regions covering the UK, but only at the SITC 1-digit level. This additional spatial detail may be helpful as a step towards obtaining the required zonal detail in the data. This regional data is compiled by merging trade data collected by HMRC with employment data from the Inter-Departmental Business Register (IDBR).

5.2.4 A business' trade is allocated to a region based on the proportion of its employees employed in that region. Accordingly, it would be prudent to check the estimated goods trade flow levels to or from locations such as Central London to ensure that this methodology is not biased by issues relating to the locations of the head offices or administrative centres of trading businesses.

5.2.5 These trade flows, particularly for imports will in many cases not be direct PC flows. Especially in the case of final demand these imported goods may just be moved to a national distribution centre, from where they enter the UK domestic supply chain.

---

13 https://www.ons.gov.uk/economy/nationalaccounts/supplyandusetables/articles/inputoutputanalyticaltables/guidanceforuse

14 https://www.uktradeinfo.com/trade-data/rts-custom-table/

15 https://ec.europa.eu/eurostat/cache/metadata/EN/ei_et_esms.htm#:~:text=Broad%20product%20categories%20are%20defined,the%20Broad%20Economic%20Categories%20(BEC).

16 https://www.gov.uk/government/statistics/uk-regional-trade-in-goods-statistics-disaggregated-by-smaller-geographical-areas-2023/uk-regional-trade-in-goods-statistics-disaggregated-by-smaller-geographical-areas-2023-explanatory-notes

5.2.6 It would also be beneficial to explore the ONS official statistics currently in development that present the estimated value of exports and imports of goods for 2023 for International Territory Levels 1, 2 and 3, and city regions[17], as this provides a finer geographical split down to 182 areas across the UK. The Inter-Departmental Business Register (IDBR) is the primary data source used here to apportion trade value to the level of the local unit of a business from the level of the reporting unit. It would, however, require careful examination of this methodology to ensure that it matches to the data requirements of the model.

5.2.7 The import and export tonnages that have been determined from trade statistics should be compared by commodity type against those from transport terminal tonnages through airports, ports and Eurotunnel, since all of these entry points provide comprehensive data, with the exception of the land boundary with the Republic of Ireland.

5.3 ESTIMATING ZONAL PRODUCTION

5.3.1 For domestic production the sub-national economic data sources are only available in monetary units. The most useful source appears to be the Regional gross value added (balanced) by industry: local authorities by ITL1 region, dataset that is produced annually by ONS[18]. The GVA data is available annually, segmented by local authority by SITC 2-digit code. This provides much of the spatial detail required for the model.

5.3.2 There do not appear to be comprehensive sub-national sources of data on production in units of tonnage, only data at the UK level within the annual ONS dataset for UK Manufacturers' sales by product data[19] (previously termed PRODCON). While this has no spatial differentiation, it does have a detailed CN classification code that could be matched to SITC or NST classifications. It has both the total value and tonnage for UK production for most manufactured goods, though not for agricultural output, for which the data are collected instead by DEFRA.

5.3.3 This product dataset could be used to produce national *volume to value* ratios for application to the zonal data on economic production levels by SITC. These calculations would need to be applied initially at the most detailed SITC code level by zonal level available and only subsequently be suitably aggregated up to the required more aggregate NST codes to generate the production tonnage estimates by zone.

5.3.4 This general approach has some uses but is likely to be relatively coarse, as opposed to the more specific information for the location of major tonnage generators from individual industrial facilities, such as cement works, quarries, car plants, fuel refineries, steel plants, etc. that is potentially available from trawling through industry specific directories, such as the

---

17 https://www.ons.gov.uk/businessindustryandtrade/internationaltrade/bulletins/internationaltradeinuknationsregionsandcities/latest#data-sources-and-quality

18 https://www.ons.gov.uk/economy/grossdomesticproductgdp/datasets/regionalgrossvalueaddedbalancedbyindustrylocalauthoritiesbyitl1region

19 https://www.ons.gov.uk/businessindustryandtrade/manufacturingandproductionindustry/datasets/ukmanufacturerssalesbyproductprodcom

British Geological Survey for non-fuel mineral production[20], construction building materials[21] , etc. Examining the full range of available industry directories is a reasonably tedious task but it is likely to be the most effective way to ensure that there is a sound information base for the spatial differentiation of production tonnage for the main heavy industry sectors. It was the approach that was adopted within the BYFM model development (WSP, 2011) to minimise the effects from the coarseness of the use of national level value to volume ratios.

5.3.5 A further potential set of sources of production tonnages is the set of modal freight statistics from the Department, derived from the CSRGT for road, from ORR statistics for rail and from individual port statistics. However, for the reasons illustrated below in Section 5.6 and in Annex C, the current zonal estimates of goods tonnage lifted that are based on these transport data sources may contain an indeterminate mixture of genuine goods production,; warehousing activity and intermodal activity. Annex B identifies improvements to the CSRGT and to its processing that would improve its usefulness in distinguishing production and consumption from warehousing activity.

5.3.6 The other issue with the CSRGT data is that its current sample size is far from that which would be required to ensure that it could produce accurate zonal estimates of production and consumption, segmented by NST. The sample size increase that would be required to guarantee such small sampling errors throughout is likely to be prohibitively costly to manage as well as being very unpopular, unless there is some technological change that greatly simplifies the response task for the transport industry.

5.3.7 For the non-commodity NST categories 14 to 18 indicated in Table 2 as having significant vehicle kilometres in 2024, the economic statistics would not provide zonal estimates, so the CSRGT may need to be spatially disaggregated and used for them, while noting the likelihood that some of these movements are likely to be to/from intermodal or distribution centre facilities, rather than production or consumption zones.

[20] https://www.bgs.ac.uk/mineralsuk/statistics/uk/
[21] https://www.gov.uk/government/statistics/building-materials-and-components-statistics-july-2025

TABLE 2 GOODS MOVED BY ROAD (MILLION TONNE-KILOMETRES) BY COMMODITY TYPE, GB

NST	Commodity group	2024
	All products of agriculture, forestry, raw materials	**28,653**
1	Agricultural products	11,658
2	Coal and lignite	263
3	Metal ore and other mining and quarrying	16,732
	All food products, including beverages and tobacco	**30,945**
4	Food, including beverages and tobacco	30,945
	All textile, leather and wood products	**5,568**
5	Textiles and textile products; leather and leather products	785
6	Wood products	4,783
	All metal, mineral and chemical products	**21,493**
7	Coke and refined petroleum products	4,546
8	Chemical products	5,577
9	Glass, cement and other non-metallic mineral products	7,720
10	Metal products	3,650
	All machinery and equipment, consumer durables	**6,322**
11	Machinery and equipment	2,525
12	Transport equipment	2,778
13	Furniture and other manufactured goods	1,019
	All other products	**75,269**
14	Waste related products	13,978
15	Mail, parcels	1,917
16	Empty containers, pallets and other packaging	11,611
17	Household and office removals	2,139
18	Groupage	44,831
19	Unidentifiable goods	789
20	Other goods	[u]
	All commodities	**168,250**

Source: Road Freight Statistics Table RFS0105

5.4 ESTIMATING ZONAL CONSUMPTION

5.4.1 The first step is to separate the consumption between the final demand by the population and the intermediate consumption by domestic industries that consume goods as an input to the production of other goods. The national total values are available from ONS National Accounts.

5.4.2 The zonal value of intermediate consumption for a specific type of commodity can be estimated using the ONS supply and use tables to estimate, for each production or service sector in a zone, how much input it requires of the specified commodity. A simple example is that the intermediate consumption of metal ores (NST:3) will largely be a function of the production level in the zone of metal products (NST:10).

5.4.3 The zonal estimates of both final and intermediate demand should additionally distinguish between:

- consumption of domestically produced goods
- consumption of imported goods.

5.4.4 The allocation of final consumption among zones can be determined as a function of the zonal population and zonal employment, combined with a suitable weighting.

5.4.5 There may be a case to take account of differences in regional consumption rates per capita by commodity type, through making use of the ONS regional household expenditure estimates[22]. However, it may be argued that the regional differences in consumption rates for goods are much less pronounced than the differences in some other non-commodity rates, such as housing expenditure.

5.5 DISTRIBUTION CENTRE LOCATIONS

5.5.1 The two main potential sources for zonal annual estimates of potential distribution centre activity are:

- employment data for those within the industry sector from the Business Register and Employment Survey (BRES), which is collected by ONS for detailed SIC codes.
- Non-domestic rating and floorspace data from the Valuation Office Agency for England and Wales[23] and from the Scottish Assessors[24] in Scotland.

5.5.2 Because of the need to distinguish between national, regional and local distribution centres, the VOA data is better suited to providing this distinction, as it includes a measure of the size of the hereditament, as well as of its rateable value. It also includes a SCat (Special Category) code that describes in detail the type of usage of each hereditament.

5.6 ASSEMBLING DATA TO SUPPORT A SUPPLY CHAIN BASED MODEL

5.6.1 As explained above, there are reasonable data sources available to estimate the various zonal totals required for input to the model. The more challenging task is to assemble a suitable database to calibrate the initial demand model in the base year. The required data on flows segmented by logistic legs is not currently assembled via the CSRGT nor in any other DfT statistical sources.

5.6.2 There are a number of ways of assembling the empirical foundations for the representation of supply chains within the freight demand model. These are listed here, in order of their increasing resource requirements and corresponding improvements in accuracy and coverage, together with cross-references to where each data assembly procedure is discussed in greater detail in this report.

a) Process the existing CSRGT journey records to allocate a probable land use type at each loading and unloading stop, based on their respective post codes. This allocation could be

---

[22] https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/articles/developmentofregionalhouseholdexpendituremeasures/2025-03-12

[23] https://www.gov.uk/government/statistics/non-domestic-rating-stock-of-properties-2024

[24] https://www.saa.gov.uk/

based on cross-referencing to VOA databases on non-domestic property rates[25] and on Council Tax records on dwellings. Based on the combination of: the pair of land use types; the commodity type; the trip length; and the characteristics of the load and vehicle type, allocate a probable logistic leg type to the trip (Annex B, 11.1).

b) Adjust future CSRGT interview questionnaires to collect added information for each journey on the land use type at each loading and unloading stop. Use this information as in the previous option (Annex B, 11.1).

c) Develop a new automated shipment-based truck survey or perhaps even a full survey of supply chains (Annex B, 11.2).

6 ESTIMATING THE MATRICES

6.1 INTRODUCTION

6.1.1 Although we have stressed the importance at the outset of retaining a procedure based on the concept of PC **matrices** (i.e. forming the connections between the zonal Ps and Cs described in the previous section), and then moving from that to a set of modal OD matrices, it must be made clear that the available data cannot support an independent estimate of the PC matrices. We therefore describe a process whereby we use as much of the data as possible to construct both sets of matrices simultaneously. However, the methodology is still anchored to the zonal Ps and Cs.

6.1.2 In line with what was set out in Section 4.2, we require an underlying model which can be partially calibrated to various sources of data. This model is sketched out in the following Section.

6.2 A MODEL OF FREIGHT DISTRIBUTION

6.2.1 A persistent feature of freight modelling is the need to recognise the different ways in which freight can be transported between production and consumption zones, and in particular the possibilities for transshipment. The model we propose is based on that developed by one of the authors for the European Commission – TRIMODE (Noekel et al., 2018), though this itself is an extension of an earlier Spatial Input-Output model formulation similar to that used in the BYFM model[26] (WSP, 2011) which in turn was largely based on the original formulation in Williams (1979). The full set of equations used in TRIMODE to represent the formulation of the distribution channel choice model is presented within the TRIMODE (Noekel et al., 2018) paper, and the main subset of them is explained below in the rest of this section.

6.2.2 The key concept is that of the "distribution channel" *h*, which is generally similar to the "transport chain" proposed by Jensen *et al* (2019). The channel represents the P-C movement

[25] VOA rating lists, including post-code, for all individual non-domestic properties for England and Wales are available for download for 2023 at https://voaratinglists.blob.core.windows.net/html/rlidata.htm, though some discussion between the Department and the VOA over the conditions of use of the dataset may be required.

[26] All the equations used within the BYFM model to represent distribution chain choices are presented and explained within Appendix I of the BYFM Final report (WSP, 2011). In para. I.6.4 it explains the difference between the *distribution chain choice* procedure that was adopted in BYFM and the *distribution channel choice* procedure that was later adopted in TRIMODE. It is this latter procedure that is recommended for use for the NFTM, as presented within this section.

using different kinds of transport with the possibilities of unloading and loading at warehouse and other distribution centres *en route*.

6.2.3 The model is represented as a hierarchical logit choice model where the upper model of channel choice is supported by a lower level which, for each type of channel, considers the actual "routes" in terms of logistic "legs" – the key stages in the channel. An important element is the **type** of logistic leg, which we describe as primary, secondary or tertiary, since these relate to consignment size and hence the type of transport that is likely to be used.

6.2.4 A **leg** is defined as a stage between when a good is loaded onto the vehicle and when it is unloaded. This leg may sometimes be direct from where the good is originally produced to where it is ultimately consumed. More typically, however, there will be a succession of logistic legs accessing National (NDC), Regional (RDC) or Local (NDC) Distribution Centres (warehouse facilities), that when combined comprise the **distribution channel** that delivers the goods from the producer through eventually to the consumer.

6.2.5 Each possible channel type h is defined by the number of legs (K_h), the sequence of facilities $\{k\}$ pertaining to each leg, and the types of each leg. A simplified example of possible channels is provided in Annex C.

6.2.6 The lower level model works back from zones of final consumption $\{c\}$ to choose the possibilities for sourcing the demand. The whole model is run independently for each of the 20 commodity groups r discussed in Section 3.9. The number of possible channels and the form of the legs will vary substantially with the commodity type.

6.2.7 The structure is illustrated in the diagram below:

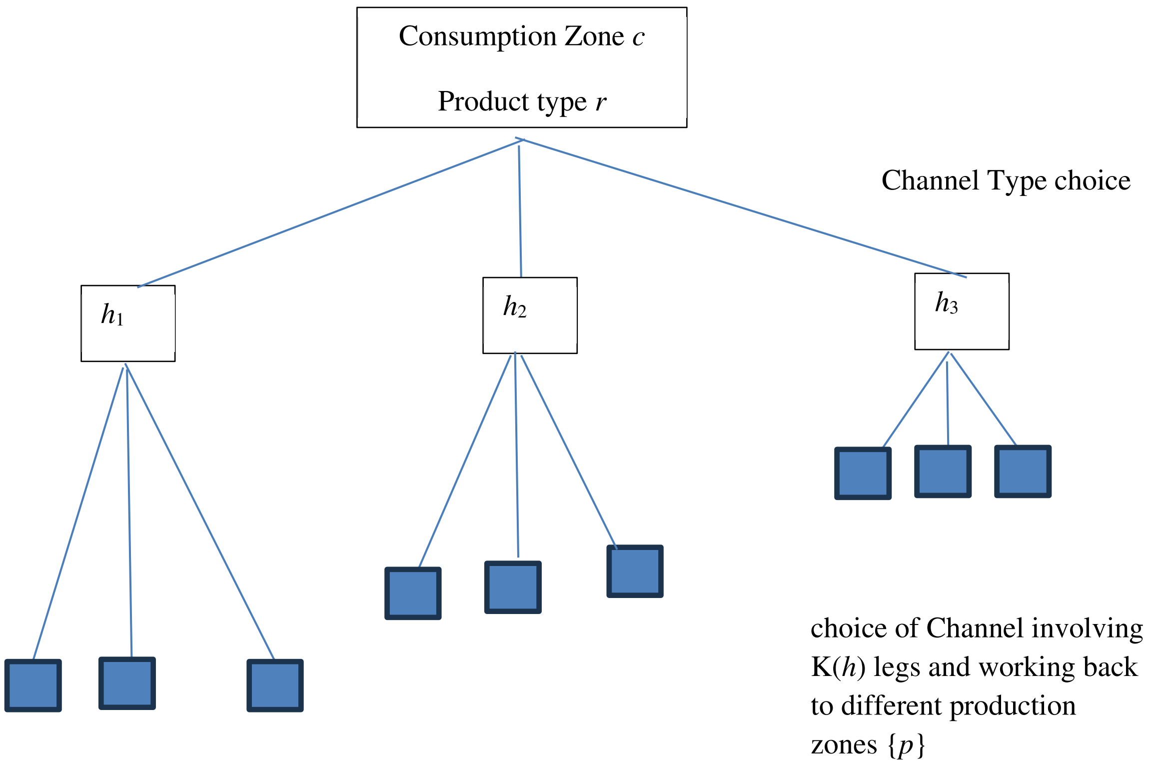


FIGURE 1 CHANNEL TYPE CHOICE ILLUSTRATION

6.2.8 The choice model is based on the generalised cost per unit of freight u, which includes both the transport costs (based on distances and times provided from a network) as well as the handling and storage costs at intermediate points.

6.2.9 In outline, the lower level model proceeds as follows. We denote the start and end point of each leg by i and j respectively. Since we work backwards from the consumption zone, for the first leg (k = 1) within any channel type, we will have $j = c$, and for the final leg ($k = K_h$) we will have $i = p$.

6.2.10 From zone $j = c$, we calculate the likelihood of receiving the consignment from each possible zone i, based on $u_{ij} = \hat{u}_{ij} + w_i$ where $\hat{u}_{ij}$ is the transport cost and w_i is the warehouse gate cost[27] at the start of the leg, to cover the cumulative cost of delivery to the unloading, storage and reloading at the initiating distribution centre, or else in the case of the final leg, the original production cost in zone p.

6.2.11 The probability of receiving the consignment from zone i is then given by Eq [1]:

[1] $$p_{i|jhk} = \frac{S_i\,.\exp(-\lambda.u_{ijk}^h)}{\sum_{i'} S_{i'}.\exp(-\lambda.u_{i'jk}^h)}$$

where S_i is a zonal size term that denotes the capacity of the origin zone i for the delivery of the freight type, and λ is a parameter which governs the level of variability in the choice process.

6.2.12 In a similar way we work back over the legs within the distribution channel h until the final leg is reached, where we get the probability of choice of the production zone.

6.2.13 The composite utility (logsum) $\tilde{u}_c^h$ from this choice process for each channel h for the consumption zone c is then calculated and used to provide the choice of distribution channel as in Eq [2]

[2] $$p_{h|c} = \frac{\exp\left(-\theta.\lambda\left(\tilde{u}_c^h\right)\right)}{\sum_{h'}\exp\left(-\theta.\lambda\left(\tilde{u}_c^{h'}\right)\right)}$$

where θ is the structural parameter in the hierarchy lying in the range [0 $\leq \theta \leq$ 1]

6.2.14 The resulting distribution channel probabilities are applied to the total consumption demand for the freight type at zone c.

6.2.15 The parameters λ and θ need to be calibrated to available data (separately for each freight type and, in the case of λ, for each logistic leg type) , and the utility terms adjusted to meet a set of constraints defined on the basis of the surveys, counts etc. as set out in Chapter 4.

---

[27] The terminology *warehouse gate cost* for the zone i is introduced because this is the cost term that is relevant to the decision being made by the shipping agent with respect to choosing which type of distribution channel h from which origin zone i to select for the delivery of the consignment. How the consignment has actually been delivered to that specific distribution centre i is not of direct interest to the shipping agent, but only the cost of purchase that is charged there.

In addition, summing over all consumption zones c, the implied volumes at the production zones $\{p\}$ will need to be constrained to the separately estimated P volumes. Clearly there will be considerable elements of judgment in this process, given the manifold uncertainties.

6.2.16 It may be noted that the proposed approach deliberately avoids the concept of distribution channel routes/paths from P to C, which would potentially generate exponential choice sets. While it would be possible to derive the P-C matrices by following through all the distribution channels, our view is that it is not strictly necessary, as the main emphasis is on providing the set of O-D matrices which retain the implied connection to the zonal Ps and Cs.

6.3 O-D MODAL MATRICES

6.3.1 Given the construction of the supply chains between P and C, we can derive the O-D modal matrices between the locations relating to each successive leg, and then sum over all consumption zones c. Separately for each commodity type, this will provide the (annual) freight volumes between all zone pairs ij distinguished by the type of leg and the associated modes/ vehicle types.

6.3.2 In relation to the secondary aim of providing LGV and HGV[28] matrices to the NTM, there are a number of further tasks required. Firstly, the pure freight delivery matrices need to be supplemented for empty running. The level of empty running should probably be commodity-specific, since – for example - heavy bulk flows have a higher empty return rate than some other NSTs and leg types. Appropriate values can be derived from analysis of the CSRGT loaded and empty vehicle movement patterns and, on a more limited basis, from van survey results. In BYFM empty running was applied as a probability to the reverse of the estimated loaded movements, and a similar process was incorporated in the Swedish SAMGODS model[29] (where it was noted that 24-25% of all domestic vehicle kilometres of lorries registered in Sweden is driven empty).

6.3.3 An alternative simpler approach might be to apply it to O-Ds as a reverse factor segmented by HGV type and trip length, rather than at the commodity tonnes level. It is likely to need some CSRGT analysis and experimentation with approaches to finalise the most suitable approach.

6.3.4 The next operation is to aggregate over all commodity types, since the NTM is only interested in vehicles, regardless of their consignments. For the LGVs, it will be important to avoid any double counting, as noted in Section 3.3.

6.3.5 The matrices are still at an annual level, and at a coarser zonal level than is required for the NTM. In terms of the NTM, we require average weekday movements, and suggestions regarding the average day are set out in Section 3.7. The disaggregation to the finer NTM zone system can be achieved by the mechanism set out in paragraph 3.6.4.

[28] Note that we think there may be some value in terms of assignment from using different types of HGV, as discussed in Section 3.8
[29] Significance (2015)

7 CONNECTION TO FORECASTING MODEL

7.1.1 Here we provide summary recommendations on the manner in which future economic change is fed through to provide suitable inputs to scenario runs of the model for future years. We are aware that there is a separately commissioned thinkpiece on this topic.

7.1.2 The expectation is that much of the use of the NFTM will be to examine longer term transport growth patterns within an era of rapid technological and economic change and associated scenarios on the potential impacts of future policy initiatives.

7.1.3 It would seem beneficial to provide a formal linkage between the change in production and consumption levels for individual commodity types that are input to the freight demand model for forecast years and a UK economic and trade forecasting model that maintains a detailed representation of individual industry types which can be mapped across to the commodity type segmentation in the transport model, ideally including a regional or more detailed spatial disaggregation across the UK. The use of this type of national economic model to guide the future growth trends within the NFTM's inputs of zonal tonnages produced and consumed would provide a coherent mechanism to represent scenarios on future zonal changes in patterns of trade and of economic growth for each commodity type segment.

7.1.4 We agree with the view presented on FFR p.12 about taking freight transport tonnage data, rather than economic monetary value data, as the starting point for freight demand model development. This differs from the BYFM approach which applied commodity type-specific volume-to-value ratios to zonal economic estimates measured in monetary value units to then estimate zonal freight production and consumption tonnages for each year and commodity type. Accordingly, when forecasting, it appears preferable to apply commodity type-specific, dimensionless, relative economic growth factors to the base year zonal production and consumption tonnages by commodity type. It is likely to be simpler and at least as realistic to make use of these dimensionless growth factors constructed from the outputs of the economic model, rather than using volume-to-value ratios, when bridging the gap between monetary value units and physical volume units within the freight demand growth estimation procedure in the freight model.

7.1.5 The ideal way ahead would be to interface to some existing national economic model that is already in use across Government Departments, so as to maximise in-house consistency in policy assessments between Transport and these other Departments. It would be helpful to investigate whether some such candidate in-house model might already be available or whether it would be preferable instead to interface to one of the existing external models developed by economic consultancies in the UK?

8 SUMMARY

8.1.1 Time has been short, and there are many aspects which will require much more discussion before a viable blueprint for the construction of the base year matrices which are an essential part of a National Freight Transport Model [NFTM]. However, we have endeavoured to make useful recommendations on a number of key issues.

8.1.2 At the outset, we set out our commitment to the P-C approach Even though in practice we do not consider it necessary to develop the matrices themselves, they provide the basis for the construction of the essential logistic supply chains, as discussed in Chapter 2, while maintaining consistency with zonal production and consumption.

8.1.3 In Chapter 3 we make recommendations for the dimension of the matrices, largely in line with the recommendations in the FAME Feasibility Report. This leads to a maximum of 80 P-C matrices in conception. However, while we agree that ultimately the NFTM should include the whole of the island of Ireland, we are concerned at the incompatibility of data definitions, as discussed in Section 3.5, and incline towards the idea of a staged approach beginning with a model for Great Britain only (while including freight movements to and from Northern Ireland.

8.1.4 We have given considerable attention to data issues, which remain challenging! As well as setting out the main sources, we have also provided recommendations in Chapter 4 as to how some existing surveys, such as CSRGT, could be expanded and improved in coverage, while Chapter 5 discusses the sources required to construct the zonal production and consumption estimates.

8.1.5 Finally, in Chapter 6, we have briefly sketched out the modelling procedures which will deliver both the matrices required for NRTF and those for the (passenger) NTM. Our recommendations are to stay with the general approach developed for the earlier Base Year Freight Matrices [BYFM] study (WSP, 2011), which has had practical demonstration, with some further developments for the European Commission TRIMODE model.

8.1.6 Further attention will need to be given to the detailed calibration of the model, where, in addition to the general approach used in BYFM and TRIMODE, we recommend making use of other sources of data (such as vehicle counts). We have expressed a more sceptical view of validation, with a general preference for incorporating as much data as possible in the calibration process.

8.1.7 In Chapter 7 we have set out some preliminary ideas as to how the base year matrices could be used in the forecasting procedure within NFTM. In the four annexes, we have provided further discussion on inconsistencies of data definition, improving access to UK logistics data, supply chains and logistic legs, and traffic mix patterns across HGV types.

9 REFERENCES

Arup (2016) Provision of technical advice and support for matrix building guidance. Technical Report for Department for Transport.
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/663319/matrix-building-guidance.pdf

de Bok M, L Tavasszy and S Thoen (2022) Application of an empirical multi-agent model for urban goods transport to analyze impacts of zero emission zones in the Netherlands. *Transport Policy,* vol. 124, August, 119-127.
https://pure.tudelft.nl/ws/portalfiles/portal/126562058/1_s2.0_S0967070X19307383_main.pdf

Department for Transport (2004) Survey of Privately Owned Vans Results of survey October 2002 - September 2003. Transport Statistics Bulletin SB (04) 21.
http://webarchive.nationalarchives.gov.uk/20060820162202/http://dft.gov.uk/stellent/groups/dft_transstats/documents/page/dft_transstats_026991.hcsp

Department for Transport (2008) Road Freight Statistics 2006, Revised August 2008. Transport Statistics Bulletin SB (06) 23. http://webarchive.nationalarchives.gov.uk/20110503144935/http://www.dft.gov.uk/pgr/statistics/datatablespublications/freight/goodsbyroad/

Department for Transport, UK (2011) Discrepancies between Road Freight and Road Traffic HGV traffic estimates: Methodology Note.

Department for Transport (2021) Van Statistics 2019-2020: Technical Report. DfT Release 15 April 2021. https://www.gov.uk/guidance/van-statistics-information

Eurostat (2015) Rail Transport Statistics Methodology. Version 8.01. http://ec.europa.eu/eurostat/web/transport/publications

Eurostat (2016) Road freight transport methodology. Reference Manual for the implementation of the Council Regulation No 1172/98 on statistics on the carriage of goods by road. https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-16-005

Guilbault M, E Gouvernal & M Soppé (2008) Demand for transport and production systems? Shipper surveys in France in 1988 and 2004. Freight & Logistics stream, ETC Conference papers 2008. https://aetransport.org/past-etc-papers/conference-papers-pre-2012/conference-papers-2008?abstractId=2851&state=b

Jensen A F, Thorhauge M, de Jong G, Rich J, Dekker Th, Johnson D, Cabral M O, Bates J J and Nielsen O A (2019) A disaggregate freight transport chain choice model for Europe. Transportation Research Part E: Logistics and Transportation Review, Volume 121, January 2019, pp 43-62, https://doi.org/10.1016/j.tre.2018.10.004

Le Vine S, J Luan and J Polak (2013) Van travel in Great Britain: What do we know from the National Travel Survey? Report for the Independent Transport Commission. http://www.theitc.org.uk/our-research/research-reports-2/

MDS Transmodal (2024) DfT FAME: Feasibility Report. Produced for Department for Transport. https://www.gov.uk/government/publications/freight-analysis-and-modelling-environment

Noekel K., I. Williams & D Fiorello (2018) Long-Distance, Multimodal Freight in a Continental Transportation Model. *Transportation Research Record*, vol. 2672, (44), 103-113. https://journals.sagepub.com/doi/pdf/10.1177/0361198118791910.

ORR (2024) Freight rail usage and performance: Quality and methodology report. Office of Rail and Road. https://dataportal.orr.gov.uk/media/1233/freight-quality-report.pdf

Petterson H (2021) Scope for automation in the Swedish Commodity Flow survey. UNECE Working Party on Transport Statistics, 2021-06-11.

Significance (2015) Method Report - Logistics Model in the Swedish National Freight Model System, Deliverable for Trafikverket, Report D6B, Project 15017, Significance, The Hague.

Williams I N (1979) An approach to solving spatial-allocation models with constraints. Environment and Planning A, vol. 11, pp. 3-22.

Williams I (2024) Peer Review of Freight Analysis and Modelling Environment (FAME) – Stage 1. Produced for Department for Transport. https://www.gov.uk/government/publications/freight-analysis-and-modelling-environment

WSP (2011) Base Year Freight Matrices: Final Report and Appendices. For contract no. PPRO 04/03/14 for UK Department for Transport. https://www.data.gov.uk/dataset/ad3c0fd4-17e9-4a8a-af5a-0d03954907a0/base-year-freight-matrices

10 ANNEX A: EXAMPLES OF DATA DEFINITION INCONSISTENCIES

10.1.1 To ensure that analysts constructing the base OD matrices will be able to assemble and combine the required datasets in a consistent and informative fashion, it is crucial that they understand the types of underlying data definition consistencies that may need to be addressed. The following three examples for: LGVs; HGV traffic levels; and weight of goods transported, highlight just a few of the many meta-data inconsistencies that arise when interpreting combinations of standard freight data sources.

10.1.2 The meta-data supplied within the "Van Statistics 2019-2020: Technical Report" (DfT, 2021) indicates that its analysis and its comparison with earlier surveys will need to be carried out with great care. This 2020 van survey defines a van as:

> "(Light commercial vehicles / light vans) 4-wheel vehicles constructed for transporting goods. Must have a gross weight of 3.5 tonnes or less.
>
> The definition of a van used in this research means that the population covered by this study is smaller than number of licensed vans in Great Britain." (page 1, DfT 2021)

10.1.3 This difference in definitions needs to be accounted for when expanding the survey results to the national fleet. More generally, the definitions of road vehicle types adopted within the statistics on the new and used national licensed vehicle fleets or in the odometer data on vehicle mileage available through the MOT, do not necessarily correspond to those used in traffic count statistics.

10.1.4 Road goods vehicle type categorisations are not consistent between the DfT traffic counts used for the AADF estimates and the traffic counts collected by National Highways (NH) for the trunk road network. Differences between visual based vehicle type classifications and automated counts with vehicle type classifications based on vehicle lengths could potentially lead to significant differences in estimated vehicle type splits. Accordingly for the reasons discussed below, **all matrix adjustments of vehicle ODs, segmented by HGV type, aimed to better match to observed individual link flows should be based on the DfT road traffic open dataset[30] and not on NH vehicle type counts.**

10.1.5 More specifically, the DfT road traffic open data provides classified vehicle data for all of the motorway and "A" road network of Great Britain and for some of its minor roads. For HGVs this is further classified into 3 categories of rigids plus four of artics based on the vehicle's number of axles and whether with trailer. This classification can be matched reasonably well to the HGV type classification proposed earlier in Section 3.8. Buses and

[30] https://roadtraffic.dft.gov.uk/about

coaches are in a separate further category, and LGVs under 3.5 tonnes gross vehicle weight are also in a separate vehicle category. DfT's 200 permanent automatic traffic counter sites (ATCs) record information about vehicles passing over them, including vehicle length and wheelbase, to classify into the required vehicle types.

10.1.6 National Highways have automatic traffic counters across their whole trunk road network, but these ATCs automatically classify vehicles into just 4 types which are based only on vehicle length: (a) 0 – 5.2 metres; (b) 5.21 – 6.6 metres; (c) 6.61 – 11.6 metres; (d) 11.6+ metres. The Traffic Statistics Methodology Review[31] carried out by DfT in 2018 concluded that it would be inadvisable to directly use the vehicle typology mix from the NH ATCs. Instead it backed combining: the percentage split of traffic by vehicle type using DfT counter vehicle type splits; together with Highways England ATC traffic totals. This combined approach has been used for the road traffic estimates from 2017 onwards for Great Britain.

10.1.7 DfT statisticians have in the past carried out an investigation into why an increasing gap had arisen between their two estimates of HGV traffic that are based on different data sources.

> "Both the Road Freight statistics series and the Road Traffic (RT) series obtain and publish estimates in vehicle-kilometres of Heavy Goods Vehicle (HGV) traffic. The Road Traffic HGV traffic estimates are consistently higher than those from the Road Freight series and **the gap has increased by nearly 11% to 29%** between 2000 and 2010." (p.1, DfT, 2011, emphasis added)

[31] https://assets.publishing.service.gov.uk/media/633337988fa8f51d2a86317f/traffic-statistics-methodology-review-alternative-data-sources.pdf

TABLE 3 SUMMARY OF THE MAIN DISCREPANCY FACTORS BETWEEN THE ROAD TRAFFIC ESTIMATES AND THE CSRGT HGV TRAFFIC ESTIMATES AND OF THEIR IMPACTS

Factor	Impact
ATC misclassification of 2-axle rigid HGVs, LGVs and buses	Over-estimation of the RT estimate (both total estimate and 2-axle rigid estimate)
Miscellaneous large non-goods-carrying vehicles	Over-estimation of the RT estimate (both total estimate and some sub-classification estimates)
Underreporting in the CSRGT	Under-estimation of the CSRGT estimate (both total estimate and all sub-classification estimates)
Magnification of underreporting due to methodology of grossing up CSRGT to annual estimates	Under-estimation of the CSRGT estimate (both total and all sub-classifications)
Foreign and NI-registered HGVs	Coverage difference which should be included in the RT estimate but excluded from the CSRGT estimate
Differing treatment of drawbar trailers as rigids and artics	Definitional differences causing differences in estimates at a sub-classification level only
Differing treatment of axles – by total number, or number on the ground	Definitional differences causing differences in estimates at a sub-classification level only
Increasing numbers of LGVs over time	An increasing impact on ATC misclassification (and thus increasing over-estimation of the RT estimate) contributing to the continuing divergence over time
Impact of EU expansion	Affects the proportion of foreign HGV traffic attributed to each country but unlikely to be a significant cause of continuing divergence over time
Increasing foreign HGV traffic over time	Only a slight increase in foreign HGV traffic so only a marginal contribution to the continuing divergence over time

Source: DfT (2012?) font colouring added.

10.1.8 In particular, CSRGT only includes vehicles used for the carriage of goods and so excludes other heavy vehicles such as fire engines, mobile cranes, etc. According to the DVLA database 79% of licensed HGVs are commercial (goods carrying) with the residual 21% having an HGV body type but not used for commercial freight. However, the daily mileage of many such vehicles might be expected to be much lower than that of the goods carrying HGVs.

10.1.9 The DfT conclusions on impacts of data inconsistencies are summarised in TABLE 3. Those converted above to red font relate to differences in data definitions or of data coverage between distinct data sources. These impacts have been quantified as part of an adjustment procedure that the Department carried out to generate an improved match between the two data sources, to enable the Department to compare their trends in HGV traffic growth on the basis of a mutually consistent set of data definitions.

10.1.10 Due to the different weight measures adopted within the standard modal statistical sources, namely: **gross-gross** weight on rail; but just **gross** weight on road, sea and air, the actual weight registered for a shipment may appear to change whenever there are transfers of shipments between rail and other modes. Nor will this shipment weight match to the trade statistics definition which measures only the **net weight** of the goods, **excluding any packaging or carriage equipment.**

10.1.11 More specifically, when measured on road the "gross" weight of goods term is defined as:

> "… total weight of the goods and all packaging, but **excluding the tareweight** of any container, swap-body and pallets containing goods". (P10-B-62, Eurostat, 2016, bold added)

When measured on rail, the "gross-gross" weight measurement that is used is defined as:

> "The weight to be taken into consideration **includes**, in addition to the weight of the goods transported, the weight of packaging and **the tare weight of containers, swap bodies, pallets as well as road vehicles transported by rail** in the course of combined transport operations". (p.20, Eurostat, 2015, bold added).

10.1.12 Paradoxically, within the ORR rail freight statistics publications this "gross-gross" weight measurement definition for rail freight is actually termed “net” tonnage to denote

> “Freight lifted - The mass of goods (tonnes) carried on the rail network, excluding the weight of the locomotives and wagons.” (ORR, 2024, p.6)

More generally for rail the term “gross” tonnage includes the weight of the locomotives and wagons.

10.1.13 For bulk and semi-bulk goods this distinction between modes in the form of weight measurement is of limited significance. However, for high value finished products, particularly light products such as textiles and equipment that generally will be shipped in unitised format, the added weight of the units used for carriage may be comparable with the weight of the goods carried themselves. Accordingly, it will be necessary to discount the tare weight of these units whenever the rail tonnages are being merged with other modes or with trade statistic's based measurements of goods flows.

10.1.14 This example of major naming inconsistencies between modes highlights the care that is needed with interpreting terminology across modal datasets, plus the great need to introduce a common unambiguous set of data-definitions throughout all datasets in use.

11 ANNEX B: IMPROVING ACCESS TO UK LOGISTICS DATA

11.1 CSRGT REPROCESSING AND REFINEMENTS

11.1.1 To realistically represent the economic logic underpinning shippers' decisions within a freight demand model, it is important that realistic information on the observed current patterns of goods movements along supply chains is available for use during its calibration. However, explicit information on the stage in the distribution chain for road goods movements has never been collected within the CSRGT. Nor is it available for the rail or maritime modes, though this is less of an issue because most (though not all) of their flows will in practice be primary logistic legs, because of the large size of the shipments typically carried by these non-road modes.

11.1.2 The simplest and fastest option would be to explore the effectiveness of **reprocessing the existing CSRGT** data in order to allocate a land use type to the unloading and loading points of every stop.

11.1.3 It would be most resource effective to initiate this option through a small scale **feasibility pilot study** to identify the likely resulting quality of the allocation procedure and to explore the most suitable methodology and background datasets to use to support this allocation of land use type. The first step of this pilot study would extract a random sample of the completed CSRGT questionnaires over the last few years to examine the patterns of trips for which valid postcode details had, and had not, been provided for both ends of the trip. This step would identify the maximum potential for the allocation procedure to be successful. It would also identify whether certain types of commodities, vehicle types, area types or trip types have lower than average compliance rates in providing postcodes. The second step for each sampled trip with compliant postcodes, would allocate based on its postcodes a land use type to each end of this trip. The allocation is based on matching this post code to a property entry in either a suitable domestic or a non-domestic dataset of properties and of their usage. Potential datasets for examination include: the VOA non-domestic rating files; the Council Tax dwelling files; and the Ordnance Survey NGD Land Use files.

11.1.4 If the land use type at the start of a trip is a quarry, farm, factory, international port, etc. then the trip would be expected to be a primary logistics leg. If the land use type at the end of a trip is a house, office or retail premises, then the trip is likely to be a tertiary logistics leg. Because the VOA differentiates warehouses by size, the differentiation between national, regional and local distribution centres should be feasible. The CSRGT also now enables terminals to be explicitly identified by mode, which would help to build up intermodal chain patterns. Other similar rules can be built up to take account of the pair of land use types for the trip to determine its likely place within the supply chain for that commodity type. It will be helpful when designing the rules that identify the location within the supply chain of a trip to also make use of information on: the mode of carriage (e.g. container, pallet, bulk, packaged, etc.); the number of pickups and drops on the journey; and the type of HGV.

11.1.5 Provided that the pilot study has successfully demonstrated that the allocation procedure generates realistic land use types for a sufficiently representative cross-section of HGV trips within the sample, then the **main CSRGT reprocessing task** could be initiated. Ideally, an automated matching procedure could be developed, perhaps based on recent AI capability improvements. This allocation should be applied to all compliant trips within the CSRGT dataset covering a few recent years. To expand to the national annual set of movements, expansion factors would then need to be applied to each compliant trip. These expansion factors need to be estimated in a form that takes appropriate account of the observed differences in postcode compliance rates across the different classes of trips.

11.1.6 In this manner, there is potential to generate an improved database of HGV flows that is suited to supporting a well-designed integrated freight and logistics demand model. Clearly, the experimentation and checking of the processing methodology proposed above would be a significant task, but its outputs would provide valuable information for use more widely within FAME.

11.1.7 On the other hand, if the pilot study results are not encouraging, then better and more comprehensive information could be obtained by **extending the CSRGT** so that it explicitly collects information on the type of economic activity at the point of pick-up and of delivery of each load. This approach had been pioneered in the 2003-05 van surveys carried out by the Department (DfT, 2004 and 2008) and the information that it generated played an important part in facilitating the inclusion of LGV as a freight mode within the construction of the BYFM model. Recent improvements have been achieved within the CSRGT in explicitly identifying feeder legs to/from other modes (rail terminals, sea and airports). The Department could extend this approach from just terminals to also include on the questionnaire a list of land use types for selection at each pickup or drop stop.

11.1.8 While any such change to the CSRGT design would be of long-term major use for improving the understanding of the responses to policies of the road freight industry, it would take some years to provide concrete results and so it might not be available in time to support the initial development of this national freight model.

11.2 INFORMATIVE SURVEYS OF SHIPMENTS

11.2.1 The ideal data source to underpin the development of a supply chain based freight demand model would be a large scale shippers' survey that traces shipments through the sequence of stages from the original producer through to the final consumer. Unfortunately, such surveys are complex to design and implement so they are not commonplace. The ECHO 2004 survey (Guilbault et al., 2008) of around 3,000 shippers and 10,000 shipments in France is one of the few surveys that examined the sequence of agents and their activities along the transport chain of a shipment from its producer through to the final consignee, while covering all freight modes. Its focus was more on the movements of finished products, than on the large flows of bulk intermediate goods. It provides good insights into French supply chain structures, despite its relatively limited sample size. While the successful introduction of this type of comprehensive shippers survey to the UK would be very informative, it is likely that it would be relatively expensive and complex to implement it effectively.

11.2.2 More recently, the HARMONY Horizon 2020 research project describes the development of

> "an exceptionally large dataset of truck trip travel diaries for The Netherlands, collected by Statistics Netherlands (CBS) using an innovative automated procedure to collect the truck trip diary data. The database includes information on the vehicle, the route, and shipments that were carried, and has a high data density. First of all, the survey is mandatory: carriers are obliged to report truck trip travel diaries for the trucks in their fleet that were included in the sample of CBS. On top of that, the data was collected using an automated XML-interface with the transportation management systems that reduces the administrative burden for carriers to complete the survey." (de Bok, et al., 2022).

11.2.3 The survey data collected by CBS was for more than 2 million trips, both domestic and international. It assembled the data within a three level hierarchy:

- **Truck** data included the vehicle ownership type, fuel consumption rate, carrying capacity, vehicle type and home base address;
- **Tour** data included the distance from origin to destination of the tour, date and time of start and end; origin and destination location and zip code, haulier or own-account, vehicle capacity utilisation in %m^2 and %m^3, border crossing location;
- **Shipment** data included the distance from loading to unloading point, gross weight, cargo/loading type (solid bulk, liquid bulk, container, pallet, etc.), Loading and unloading location and zip code, loading and unloading location type (production, consumption/processing, retail, seaport, distribution centre, etc.), goods type (NST), invoice value, volume (l or m^3).

11.2.4 The truck and tour level data that was collected is broadly similar to that already collected within the CSRGT. However, the specific focus on individual shipments is not adopted within the CSRGT, which instead collects information on the total tonnes loaded and unloaded at each stop. This alternative individual shipment orientation enables key information to be assembled that provides insights into supply chain structures. The identification of the location type of the loading and unloading points of individual shipments for each commodity type, enables the differences to be quantified for each of the different types of logistic leg within a supply chain: in its spatial patterns of movements; in cargo type / form of loading; in shipment size; in vehicle size; in vehicle loading efficiency; etc.

11.2.5 If a similar breakthrough on data availability at the shipment level, could be achieved in the UK, then this would lessen the costs and risks involved in the development of a future freight demand model. Access to such data in the UK would support the identification of the relative proportions of different types of logistic legs within supply chains and the typical characteristics of each of these types of logistic legs. In turn, this data would provide the empirical foundation for the full explicit representation of supply chain activity and of its impact on transport demand patterns, which would enable major improvements in the functionality of the freight demand model component of a national freight model.

11.2.6 Commodity flow surveys have been carried out 5 times in Sweden[32], approximately every 5 years since 2001, with the most recent survey being in 2021. The focus in commodity flow surveys is on the movements of freight more than on the vehicles and vessels that transport the goods. This Swedish shippers survey measures the value and the tonnage of goods transported domestically and abroad and the sequence of transport modes that were used along the transport chain for individual shipments. The scope for automating the implementation of this survey has been explored by the Swedish Transport Administration (Petterson, 2021). Commodity flow Surveys are also carried out in the US[33] every 5 years starting in 2012, with a switch to online, rather than paper, information collection within the 2022 survey.

12 ANNEX C: OVERVIEW OF SUPPLY CHAINS AND LOGISTIC LEGS

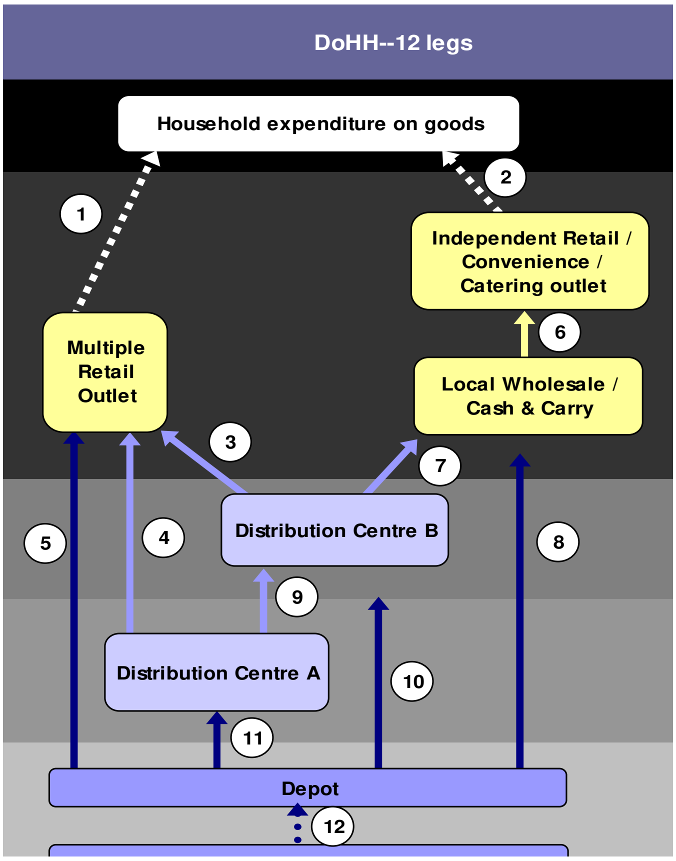
12.1 CONTEXT

12.1.1 This provides a brief overview of logistic structures within UK supply chains to provide context for the recommended introduction of an explicit representation of supply chains within the freight demand model.

12.1.2 Although some bulk primary products, such as ores or grains, may be transported directly without intermediate storage from where they are produced to where they are consumed or processed, this direct transfer is not the norm for most goods. Instead, foodstuffs and consumer goods usually pass through many intermediate distribution centres within the supply chain that connects the initial producer of the goods through to their final consumer.

[32] https://www.trafa.se/en/travel-survey/commodity-flows/#:~:text=The%20survey%20constitutes%20official%20statistics,and%20between%20Sweden%20and%20abroad

[33] https://www.census.gov/programs-surveys/cfs.html



Logistic leg	Typical vehicle types
1	Car, LGV
2	Car, LGV, bike, robot
3, 4, 5	Artic, Medium rigid
6	LGV, Small rigid
7,8	Small, Medium rigid
9, 10, 11	Artic, rail
12	Artic, Medium, Large rigid

DC type	Dwell times
NDC - A	Weeks
RDC - B	Days
LDC, FfC	Hours

FIGURE 2 SUPPLY CHAIN AND LOGISTIC LEGS FOR CONSUMER GOODS

12.1.3 An example of some of the supply chain structures that are used to ship goods to households is illustrated in Figure 2. This chart illustrates

- Some of the different *supply chain* options that may be in use
- the intermediate *distribution centres* (warehouses) at which goods are stored
- the individual *logistic legs* travelled between each intermediate storage stage within a specific *distribution channel* that connects a *producer* through to a *consumer* in a supply chain.

12.1.4 In general, when consumer goods leave the producer's depot they are likely to travel to a large central National Distribution Centre (NDC - A), being shipped as a large consignment within a large artic. Typically many of the items within this consignment may be stored at the NDC for some weeks before they have gradually all been shipped out to different regional locations, again mainly within large artics. The goods in a consignment will spend a few days at a Regional Distribution Centre (RDC -B) before gradually being shipped within an artic or a medium rigid HGV as smaller consignments to either a large supermarket or retail outlet or else to a Local Distribution Centre (LDC), a Local Fulfilment Centre (LFfC) or a local parcel delivery hub. From the LDC, goods may be delivered to or collected by smaller local retailers or caterers within LGVs or small rigids. Alternatively, items may be delivered from a local parcel delivery hub directly to households or firms as parcel deliveries using vans or cargo bikes doing a local round. In recent years via online ordering procedures, groceries in large supermarkets are increasingly being delivered from local fulfilment centres or from grocery shops in LGVs or even using autonomous delivery robots[34] direct to consumers, rather than being collected by retail customers in their own cars.

12.1.5 The aim underpinning the evolving design of the supply chain structures within firms is to minimise the overall cost of their goods distribution system (the combined costs of transport, warehousing, capital, etc.), while meeting the consumption needs of their customers. The most cost-effective way of avoiding temporary stock-outs is through carrying substantial stock levels in a central NDC from where items can then be distributed when and where they are needed, while minimising the requirement to carry large stocks at a regional and especially at a local level. This is why goods will typically be stored: for weeks in NDCs; for a few days in RDCs; and often for just a few hours in local parcel delivery hubs. The average consignment size required at the delivery end of a logistic leg typically decreases when progressing through the individual stages of the supply chain from the producer through to the consumer. This is why the typical vehicle size used gradually decreases through the sequence of logistic legs from the producer to the consumer.

12.1.6 A key feature of distribution centre activity that should be recognised within the formulation of a freight demand model is that the annual tonnage of goods leaving any distribution centre should be similar to that entering that distribution centre.

12.2 LOGISTIC LEG TYPES AND DISTRIBUTION CHANNEL CHOICE

12.2.1 In the logistics industry, it has been convenient to classify the different types of logistic legs into three generic categories that have their own distinct transport and logistic characteristics.

1. **Primary distribution** legs (e.g. from producers to major distribution centres or to major individual consumers) – these will typically be single drop, regular, large consignments moved on rail or on large road vehicles that are fully and efficiently loaded (pallets), often over long distances and often operating on a 24-hour basis.
2. **Secondary distribution** legs (e.g. from distribution centres to major retailer outlets or to local distributors / cash-and-carries) - these will typically be relatively large consignments with few drops per vehicle with some delivering using roll-cages rather than just pallets, typically over medium distances.
3. **Tertiary distribution** legs (e.g. from local distributors to dispersed small-scale consumers) - these will typically be spread as multi-drops of smaller consignments, packages and parcels often using smaller vehicle sizes that may perhaps be only partially loaded, operating only during the daytime when the receiving shops and offices are open for business, and typically over shorter distances.

12.2.2 The leg sequence within a specific distribution channel typically will comprise a mixture of types of logistic legs that is designed to produce the most cost-effective channel structure for the delivery of specific types of commodities to particular customer markets. The driving force behind the design of a distribution channel is to minimise the overall costs (transport, storage, administration, wastage, depreciation, etc.) of the distribution system for the delivery of goods to particular types of consumer markets.

---

[34] Starship robot delivery vehicles are currently in use for local deliveries from corner shops to households in Milton Keynes (since 2018), Manchester, Cambridge, Leeds and other cities.
https://www.starship.xyz/company/

12.2.3 As an example, we assume that there are H=3 types of distribution channel *h* that are available to supply a particular product. A shipping agent determines how much it would cost to obtain a unit of the consignment via each of these H distribution channel types, so that a decision can be made on which type of channel to choose. The cost choice facing the agent is based on the following characteristics of each of the individual types of distribution channel *h* and of their respective K_h logistic legs.

- $h=1$, $K_h=1$ - **direct from the producer**, in zone *i*. For this type of logistic leg *l*, the transport cost may be high even in the case where the producer is located in the same region as the factory because the consignment size currently required from that specific producer may often be well below a full lorry load. The total delivery cost from a producer is the factory gate cost $\hat{c}_i^r$ in zone *i*, plus the shipment transport cost $\hat{u}_{ic}^{rl}$ on leg $k=1$ of logistic leg type *l*, from zone *i* to *c*.

- $h=2$, $K_h=3$ - **indirectly via a *local* distribution centre** for stage $k=1$ in zone *i* that in turn is stocked for stage $k=2$ **from regional distribution centres**. Here, because other types of goods are also required to be delivered to the consuming factory from this local distribution centre, a full lorry load can be used for its delivery of the consignment. Because of the relatively small scale of the local distribution centre and its limited storage space, it is restocked regularly from a regional distribution centre, which in turn in stage $k=3$ is stocked directly **from the producers**. Hence this distribution channel has a total of $K=3$ logistic legs. Each logistic leg of type *l* will have its expected transport cost, plus just for the stages $k=1$ *or* 2, the warehousing cost to cover the unloading, storage and reloading at the initiating distribution centre, or else in the case of stage $k=3$, the original production cost in zone *p*.

- $h=3$, $K_h=2$ **- indirectly via a *regional* distribution centre** in zone *i* that in turn is stocked for stage $k=2$ **directly from the producers**. The total widget delivery cost from here is the warehouse gate cost w_{ik}^{rh} in zone *i*, plus the shipment transport cost $\hat{u}_{ic}^{rl}$ on leg $k=1$ of logistic leg type *l* from zone *i* to *c*.

12.2.4 The terminology *warehouse gate cost* for the zone *i* has been introduced because this is the cost term that is relevant to the decision being made by the shipping agent with respect to choosing which type of distribution channel *h* from which origin zone *i* to select for the delivery of the consignment. How the widgets have actually been delivered to that specific distribution centre *i* is not of direct interest to the shipping agent, but only the cost of purchase that is charged there.

12.3 POTENTIAL DISTRIBUTION CHANNEL STRUCTURES

12.3.1 This Section first describes examples of the composition of the set of distribution channels that could be considered for use within the model. Then for each channel it provides some examples of the real world context within which it may occur.

12.3.2 A set of the main types of distribution channels to potentially be distinguished within the model is listed in Table 4. This specifies for each distribution channel: its code number; name; number of distinct legs; and the type of each of the sequence of logistic legs traversed in moving from the producer to the consumption end of the movement.

TABLE 4 DISTRIBUTION CHANNEL TYPOLOGY

Distrib. Channel	Name	#Legs	Logistic leg type sequence from P to C				
11	Direct	1	1				Direct P to C
12	Direct local	1	3				Direct P to local C
21	NDC	2	1	2			Via NDC to large consumer
31	NRDC	3	1	2	3		Via NDC & RDC
41	NRLDC	4	1	2	2	3	Via NDC, RDC & LDC
22	RLDC	2	2	3			Via LDC to consumer
23	SHub	2	3	1			Spoke & Hub
24	HubS	2	1	3			Hub & Spoke
25	SS	2	3	3			Spoke & Spoke
32	SHHS	3	3	1	3		Spoke/Hub/Hub/Spoke

Source: Own analysis

12.3.3 This list of distribution channel types has been devised to try to represent the more common sequences of logistic leg types that occur when moving goods from where they are produced to where they are consumed. Further explanations are now outlined that provide the context for the occurrence in practice for these channels.

12.3.4 **Channel 12 – direct local** - has been introduced to represent single leg distribution channels that involve short distance movements within a locality. These may be road movements from a farm of milk to a local creamery or of animals to a local abattoir. They might also arise within local industrial clusters, such as: movements of metal products from steel mills to adjacent steel product manufacturers; or of automotive components to adjacent car plants; etc. The outbound leg is likely to be fully and efficiently loaded, whereas the return leg may often be empty. It is because of the need to calibrate these movements to be over relatively short distances that they are separated from the next channel category.

12.3.5 **Channel 11 – direct** –represents the more typical medium to longer distance single leg channel of a primary logistic leg from a producer to a consumer. Because the majority of production industries will have specialisms that cater for markets over a wide spatial area, the journey length distribution for their products is likely to be spatially dispersed, particularly for those commodity NSTs that represent the production of a complex range of finished products, **r**ather than just some bulk intermediate outputs for local consumption in related industrial sectors. For low value high density goods, it is unlikely that road would be a competitive mode to rail or water except over shorter distances, so that the journey length on road for such goods is likely to be short. For high value or perishable goods, road may have strong security, reliability or speed advantages over other modes, so that the journey lengths by road may be long, even when distributing import/export finished goods from/to ports.

12.3.6 To transport direct deliveries cost-effectively from a producer to a consumer, the consignment size needs to be large enough to fill the road vehicle. Whereas this may be achievable between pairs of large manufacturers, it is uncommon for a group of households to require a full lorryload of goods from some producer at once. Accordingly, direct deliveries are unlikely for final demand but do arise for intermediate demand, though more for bulk goods, rather than for finished products.

12.3.7 More generally, the design of the operation of the distribution system aims to provide cost-effective deliveries that can cover all consignment sizes for a full range of delivery urgency and security conditions. The non-direct distribution channel options that are discussed below can between them cater for this spectrum of requirements.

12.3.8 **Channel 21 – NDC** –represents a two stage movement, starting with a primary leg from the producer to a National Distribution Centre (NDC) that will warehouse a large stock of goods in a centralised accessible location. Subsequently, a full vehicle load consignment or perhaps a two or three drop secondary leg will be used to deliver goods from the NDC to the consumer. This distribution channel 21 option is not normally used for final demand because to be effective it would require large consignment sizes. It is more typical of intermediate demand requirements that often do require large consignments.

12.3.9 One reason for including an NDC within a distribution channel is that the consumer may not require a full vehicle load (large consignment) from any one individual producer, whereas the wide range of goods within the NDC may be able to satisfy a sufficient range of consumption requirements of the consumer to successfully fill the vehicle. Another advantage from the inclusion of the NDC arises with imported goods that have a long lead time from their order to their delivery. The use of an NDC truncates the time incurred from initially making the order through to eventually receiving the delivery on site.

12.3.10 The historic approach to inventory management was for each firm to carry its own large safety stock. However, this was wasteful in terms of space, wastage and obsolescence of goods, stock capital tied up, warehouse staff, risks of stockout, etc. Nowadays, with lean just-in-time systems, the goods are stocked efficiently instead at a single NDC, from which they could then be delivered reliably and quickly to consumers, by using the improved road network and vehicle stock. The large savings in local warehousing costs more than outweighs the added cost of accessing the NDC, rather than using direct deliveries and carrying high local warehousing costs and stock-out risks. Within the calibration of the distribution channel choice model, large option-specific penalties are introduced to take account of the large indirect operational costs that would arise from direct deliveries and associated local warehousing, versus the use instead of NDCs and perhaps also of RDCs that are now discussed.

12.3.11 **Channels 31 – NRDC**–represents adding in a further intermediate Regional Distribution Centre (RDC) and a subsequent tertiary logistic leg to deliver to the local consumption locations around that RDC. For goods that have a large turnover, or for perishable foodstuffs, or urgent goods that require a rapid delivery once ordered, operating just with an NDC is unsatisfactory because it may be located a long distance from a consumer. Within RDCs the goods stored have short turnaround cycles and are restocked efficiently from the NDC, often running on a 24-hour basis. These RDCs are introduced to the supply chain to provide a more effective and flexible delivery service.

12.3.12 **Channels 41 –NRLDC** –represents adding in a further Local Distribution Centre (LDC - e.g. cash and carry, parcel depot, cold-store) from which small or urgent consignments are delivered on tertiary logistic legs to local consumers. Because these final deliveries are in relatively small consignments that may not even fill a small vehicle and may require a considerable number of (un)loading stops, it would be less cost-effective to try to deliver them from a relatively remote RDC.

12.3.13 **Channel 22 – RLDC** –represents a regional movement to an LDC from which local deliveries are then made. This will typically arise for bulk goods such as petroleum products, where the goods may first be consigned by pipeline to a large storage depot, or for cement factories from which large fully loaded vehicles move the product through to local smaller depots spread around the surrounding region, from which the tertiary delivery legs are delivered using smaller vehicles that carry smaller consignments to consumers, such as intermediate demand of cement to construction sites or final demand of heating oil to offices or households.

12.3.14 **Channel 23 – SHub** –represents a spoke tertiary leg to a local consolidation hub from which a primary leg then transports the goods to the final destination. The NST 14.1 - household waste collection is one widespread example.

12.3.15 **Channel 24 – HubS** –represents a primary leg to a local distribution hub from which small consignments are distributed on tertiary leg spokes. Prepared meals or sandwiches (perishables) for the local catering trade are an example.

12.3.16 **Channel 25 – SS** –represents tertiary spoke to hub to tertiary spoke. Local post is an example.

12.3.17 **Channel 32 – SHHS** – represents tertiary spoke to hub to hub to tertiary spoke. This is how the national post and national courier and parcel delivery services operate. Items are moved in the evening from a wide set of local depots to a regional hub. From there they are moved in large vehicles on a primary leg to one of a small number of other regional hubs from which they are then dispersed on a tertiary spoke across the local depots of that region early in the morning, ready for local delivery to houses and to firms during that day. Many of these final parcel delivery stages from the local depot to households or offices may be in vans or in London on cargo bikes.

12.3.18 Although other distribution channel options exist, the set described above here should be sufficient to cover the main patterns of distribution in operation currently within the UK.

13 ANNEX D: TRAFFIC MIX PATTERNS ACROSS HGV TYPES

13.1.1 This annex compiles a selection of DfT data that identifies distinct differences between HGV vehicle types in the nature of their traffic patterns. Significant differences in the mix of traffic across HGV types are illustrated:

- across road classes
- across days of the week, for a given class of road
- across times of day, for a given class of road and day of the week combination.

13.1.2 Accordingly, when scaling annual HGV traffic flows down to specific periods during weekdays or weekends, distinct factors should be applied for different types of HGVs. This is why it is necessary to segment by individual HGV types within the base year HGV OD matrices.

TABLE 5 PERCENTAGE OF VKMS BY HGV TYPE FOR EACH ROAD CLASS, GB 2023

	Motorway	Trunk Rural 'A' Roads	Trunk Urban 'A' Roads	Principal Rural 'A' Roads	Principal Urban 'A' Roads	Minor Roads: Rural	Minor Roads: Urban
Articulated	**71**	**65**	**60**	**44**	**29**	**17**	**12**
3 or 4 axles	4	4	3	5	4	6	2
5 axles	21	19	19	14	9	4	4
6+ axles	46	42	37	24	16	8	7
Rigid	**29**	**35**	**40**	**56**	**71**	**83**	**88**
2 axles	20	22	24	34	49	51	61
3 axles	4	5	6	8	9	17	22
4+ axles	5	8	10	13	13	14	5
Total	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: DfT Traffic Table TRA3105

13.1.3 Looking first at the spread of traffic by HGV type across different road classes, Table 5 illustrates various patterns. Moving across road classes from motorways through to minor urban roads,

- the percentage of articulated HGVs within overall HGV usage reduces greatly from 71% down to 12%
- within the articulated HGV type, the usage of the largest 6+ axle class reduces much more than that of the smallest 3 or 4 axle class
- within the rigids, it is the smallest 2 axle HGV type that increases most strongly, up from 20% to 61% of all HGV traffic, whereas the increase in the share of the largest rigids is less substantial.

13.1.4 This illustrates the differences in the mix of HGV types and sizes that is typically observed on each type of road. The differences in mix are likely to be much greater on individual road links, than in the above averages across all links of a specific type within GB.

13.1.5 Table 6 illustrates, for each of articulated and rigid HGVs, an index of how their traffic intensity on motorways varies across the days of the week and across the hours in each day in the week. This index is set to 100 based on the average hourly traffic for the specific road type and HGV type combination. Although the colour patterns might initially appear broadly similar overall, important differences between vehicle types emerge from closer inspection of the variations of this traffic index.

a) On Monday to Friday on motorways, significant artic traffic continues through the night between 10PM and 5AM, at around a third of the artic traffic level of the weekday peak. For rigids on motorways, night time traffic is around a ninth of the weekday peak.
b) Rigid weekday motorway traffic is very concentrated to between 6AM and 4PM.
c) Artic weekday motorway traffic drops significantly between 7 to 9AM, presumably artic drivers avoid AM peak congestion by going for a breakfast break. In contrast, weekday motorway traffic of rigids peaks to its highest AM rate from 7 to 8 AM, until exceeded in the early afternoon.

13.1.6 Although not illustrated in Table 6, the DfT data also indicates that distinct differences between rigids and artics in their time of day and of week traffic patterns can be observed for each of the other classes of roads. For example, Table 7 illustrates the ratio of total artic to total rigid traffic kms for each hour of the week for road class motorway and for urban ‘A’ roads.

13.1.7 **On motorways**, artic traffic is at least 75% **greater** than the level of rigids’ traffic during almost all the week, except Friday PM and Saturday during the day. In the very early morning, artic motorway traffic is around 5 times that of rigids.

13.1.8 **On urban ‘A’ roads,** artic traffic runs **below** 50% of the level of rigids’ traffic between 7AM and 6PM for most of the week. It is only overnight between 10PM and 5AM that artic traffic exceeds rigid’s traffic.

TABLE 6 TIME PROFILE TRAFFIC INDICES FOR EACH OF ARTIC AND RIGID, BY DAY OF WEEK ON MOTORWAYS, 2018, GB

	Articulated HGVs on Motorways						
Time of day	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
00:00-01:00	33.77	61.39	64.01	65.50	65.19	53.42	20.55
01:00-02:00	34.47	61.72	64.65	66.02	66.09	53.91	19.93
02:00-03:00	38.21	65.79	68.16	69.47	69.74	55.56	19.93
03:00-04:00	47.88	74.60	77.23	78.38	78.64	59.95	21.60
04:00-05:00	73.92	99.20	101.19	102.08	101.70	65.76	23.96
05:00-06:00	119.52	141.29	142.49	143.85	141.38	70.70	28.01
06:00-07:00	152.19	170.47	172.74	175.16	169.32	71.20	31.00
07:00-08:00	146.28	158.98	163.15	164.84	161.04	69.82	35.21
08:00-09:00	144.23	153.41	156.21	158.58	155.70	65.96	37.32
09:00-10:00	158.50	169.19	170.82	171.99	163.67	62.86	39.36
10:00-11:00	164.37	176.72	177.61	178.33	168.43	59.77	40.91
11:00-12:00	167.01	179.30	179.58	180.28	167.15	56.40	43.27
12:00-13:00	168.63	181.48	182.10	181.20	163.05	54.10	45.28
13:00-14:00	172.46	182.78	183.45	181.28	157.72	51.05	47.03
14:00-15:00	172.20	180.54	180.76	180.12	146.58	47.85	48.26
15:00-16:00	165.50	171.62	170.28	168.68	132.13	43.91	48.41
16:00-17:00	150.05	151.13	149.43	146.70	110.16	39.32	48.63
17:00-18:00	129.48	130.58	128.72	126.06	95.60	35.89	48.30
18:00-19:00	117.14	120.43	120.25	119.17	88.72	33.40	47.76
19:00-20:00	102.96	107.55	108.93	108.50	84.85	30.48	46.46
20:00-21:00	92.77	96.24	97.96	98.23	80.93	27.51	43.53
21:00-22:00	81.05	83.77	85.12	85.55	72.12	24.60	38.96
22:00-23:00	70.65	72.96	73.94	74.45	63.30	22.40	34.49
23:00-00:00	63.63	65.86	66.98	67.04	55.69	20.92	32.92

	Rigid HGVs on Motorways						
Time of day	**Monday**	**Tuesday**	**Wednesday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
00:00-01:00	17.06	23.73	25.18	26.24	27.66	23.76	14.06
01:00-02:00	16.21	23.63	24.74	25.79	27.02	22.21	12.11
02:00-03:00	18.77	26.83	27.58	28.84	29.39	22.07	10.44
03:00-04:00	27.80	33.83	34.58	36.08	35.46	23.12	11.09
04:00-05:00	56.74	58.33	59.79	61.37	58.48	29.90	13.86
05:00-06:00	116.44	112.82	114.73	116.77	109.98	43.32	17.72
06:00-07:00	198.02	200.32	203.99	206.21	196.69	64.82	26.48
07:00-08:00	207.60	211.17	218.83	220.79	217.21	77.81	34.83
08:00-09:00	199.64	200.09	203.81	205.67	201.37	76.30	38.84
09:00-10:00	205.25	204.16	207.91	209.36	202.13	77.81	43.61
10:00-11:00	202.86	204.79	207.55	208.93	205.11	78.83	47.42
11:00-12:00	203.10	207.00	208.81	211.42	210.07	77.27	51.55
12:00-13:00	201.86	207.28	210.77	212.60	208.88	73.57	52.32
13:00-14:00	208.06	216.01	219.44	222.18	211.94	69.25	52.65
14:00-15:00	210.75	218.95	222.03	224.27	206.08	64.94	53.21
15:00-16:00	199.56	204.71	207.04	209.14	184.69	61.25	54.33
16:00-17:00	165.98	168.33	169.07	170.10	146.03	57.27	54.50
17:00-18:00	118.49	121.13	123.29	124.45	110.88	52.57	52.67
18:00-19:00	86.81	91.88	94.83	98.31	86.24	45.91	49.89
19:00-20:00	64.91	69.23	72.74	78.03	68.71	36.88	44.88
20:00-21:00	49.81	52.67	55.58	60.11	53.65	29.31	38.00
21:00-22:00	38.42	40.36	42.77	45.54	40.90	23.29	30.70
22:00-23:00	30.92	32.87	34.11	35.81	32.58	19.52	23.43
23:00-00:00	25.86	27.72	28.80	30.18	26.70	16.64	19.65

Source: DfT Traffic distribution data, colouring added.

TABLE 7 RATIO OF TRAFFIC OF ARTIC TO RIGID, BY TIME OF DAY BY DAY OF WEEK ON MOTORWAYS AND URBAN 'A' ROAD TYPES, 2018, GB

Ratio of Artic to Rigid

	Motorways						
Time of day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
00:00-01:00	4.33	5.66	5.56	5.46	5.16	4.92	3.20
01:00-02:00	4.65	5.72	5.72	5.60	5.35	5.31	3.60
02:00-03:00	4.46	5.37	5.41	5.27	5.19	5.51	4.18
03:00-04:00	3.77	4.83	4.89	4.75	4.85	5.67	4.26
04:00-05:00	2.85	3.72	3.70	3.64	3.81	4.81	3.78
05:00-06:00	2.25	2.74	2.72	2.70	2.81	3.57	3.46
06:00-07:00	1.68	1.86	1.85	1.86	1.88	2.40	2.56
07:00-08:00	1.54	1.65	1.63	1.63	1.62	1.96	2.21
08:00-09:00	1.58	1.68	1.68	1.69	1.69	1.89	2.10
09:00-10:00	1.69	1.81	1.80	1.80	1.77	1.77	1.98
10:00-11:00	1.77	1.89	1.87	1.87	1.80	1.66	1.89
11:00-12:00	1.80	1.90	1.88	1.87	1.74	1.60	1.84
12:00-13:00	1.83	1.92	1.89	1.87	1.71	1.61	1.89
13:00-14:00	1.81	1.85	1.83	1.79	1.63	1.61	1.96
14:00-15:00	1.79	1.80	1.78	1.76	1.56	1.61	1.98
15:00-16:00	1.81	1.83	1.80	1.77	1.57	1.57	1.95
16:00-17:00	1.98	1.96	1.93	1.89	1.65	1.50	1.95
17:00-18:00	2.39	2.36	2.28	2.22	1.89	1.49	2.01
18:00-19:00	2.95	2.87	2.78	2.65	2.25	1.59	2.10
19:00-20:00	3.47	3.40	3.28	3.04	2.70	1.81	2.27
20:00-21:00	4.08	4.00	3.86	3.58	3.30	2.05	2.51
21:00-22:00	4.62	4.54	4.36	4.11	3.86	2.31	2.78
22:00-23:00	5.00	4.86	4.74	4.55	4.25	2.51	3.22
23:00-00:00	5.39	5.20	5.09	4.86	4.56	2.75	3.67

Ratio of Artic to Rigid

	Urban 'A' roads						
Time of day	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
00:00-01:00	0.92	1.50	1.57	1.54	1.54	1.32	0.65
01:00-02:00	1.13	1.77	1.77	1.81	1.85	1.55	0.82
02:00-03:00	1.25	1.74	1.87	1.92	1.86	1.76	1.06
03:00-04:00	1.20	1.61	1.62	1.70	1.72	1.64	1.19
04:00-05:00	1.07	1.13	1.14	1.14	1.17	1.24	1.19
05:00-06:00	0.80	0.79	0.80	0.80	0.82	0.89	1.01
06:00-07:00	0.51	0.52	0.53	0.54	0.54	0.62	0.78
07:00-08:00	0.44	0.45	0.46	0.46	0.46	0.47	0.69
08:00-09:00	0.46	0.47	0.47	0.47	0.47	0.41	0.64
09:00-10:00	0.45	0.46	0.46	0.46	0.46	0.36	0.53
10:00-11:00	0.44	0.45	0.46	0.46	0.44	0.36	0.47
11:00-12:00	0.45	0.45	0.46	0.46	0.44	0.34	0.45
12:00-13:00	0.46	0.47	0.48	0.47	0.45	0.35	0.45
13:00-14:00	0.45	0.47	0.47	0.47	0.45	0.35	0.43
14:00-15:00	0.44	0.45	0.46	0.45	0.42	0.34	0.42
15:00-16:00	0.44	0.45	0.45	0.45	0.42	0.35	0.43
16:00-17:00	0.43	0.43	0.43	0.44	0.40	0.35	0.45
17:00-18:00	0.46	0.47	0.46	0.48	0.45	0.35	0.48
18:00-19:00	0.55	0.53	0.54	0.54	0.49	0.36	0.51
19:00-20:00	0.64	0.64	0.65	0.64	0.61	0.39	0.53
20:00-21:00	0.80	0.81	0.80	0.79	0.76	0.46	0.58
21:00-22:00	0.91	0.95	0.97	0.95	0.93	0.49	0.66
22:00-23:00	1.11	1.16	1.15	1.17	1.05	0.57	0.73
23:00-00:00	1.26	1.26	1.27	1.27	1.11	0.54	0.77

Source: Own calculation, based on DfT Traffic distribution data and table TRA3105.

13.1.9 Table 8 presents the number of HGV trips in 2024, segmented by both vehicle type and length of haul. The proportion of HGV trips that are in large artics (> 33 tonnes gross vehicle weight) varies from a minimum of 18% for distances up to 25kms, increasing to 80% for those over 300kms, with an overall average of 37% of HGV trips being in large artics.

TABLE 8 TRIPS BY LENGTH OF HAUL, PERCENTAGE BY HGV TYPE: 2024, GB

	Length of haul (kilometres)							
	Up to 25km	Over 25km to 50km	Over 50km to 100km	Over 100km to 150km	Over 150km to 200km	Over 200km to 300km	Over 300km	All lengths
Rigid								
> 3.5t to 7.5t	21	16	12	13	7	7	-	**14**
>7.5t to 17t	11	7	6	6	4	-	-	**7**
> 17t to 25t	17	17	12	11	11	8	13	**13**
>25t	32	39	32	15	9	7	5	**27**
All rigids	**81**	**79**	**62**	**44**	**31**	**22**	**18**	**61**
Articulated								
> 3.5t to 33t	1	1	1	2	3	3	2	**2**
> 33t	18	20	36	53	66	75	80	**37**
All artics	**19**	**21**	**38**	**56**	**69**	**78**	**82**	**39**
All vehicles	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Source: Own estimates based on DfT Road Freight Statistics (2024) Tables: RFS0113, 0109 and 0110.

13.1.10 Table 9 identifies, for vehicles of any given size, the major differences between commodity types in their average loads per loaded vehicle. For example, an artic carrying coal (NST:2) may have a load that is almost three times the weight of an artic carrying textiles (NST:5). This implies that each 1000 tonnes of textiles in an output OD matrix may generate three times as many artics as would 1000 tonnes of coal.

13.1.11 Within each vehicle type specific column in Table 9, the darkest green cells denote the heaviest loads, whereas the darkest red cells denote the lightest loads per loaded vehicle. These colour patterns indicate that across the set of HGV types, their relative load size patterns are not always the same[35] across the set of NSTs.

[35] However, in a few cases, such as the low 2.3 tonnes average load for NST:2 coal on rigids from 7.5 to 17 tonne gvw, the atypical value is associated with a small sample size and corresponding potentially large sampling errors.

TABLE 9 AVERAGE LOAD PER LOADED VEHICLE BY NST AND VEHICLE TYPE

	Rigids					tonnes
NST	Over 3.5t to 7.5t	Over 7.5t to 17t	Over 17t to 25t	Over 25t	Artics	All
1	1.1	3.2	8.2	13.2	18.7	**13.5**
2	1.3	2.3	13.1	18.3	26.0	**18.9**
3	1.7	6.2	12.3	18.2	25.4	**17.4**
4	0.9	2.4	5.1	9.0	14.2	**9.5**
5	1.2	3.2	5.1	4.4	9.6	**6.5**
6	1.2	3.1	6.9	8.7	20.0	**13.1**
7	1.2	3.7	6.6	11.5	23.2	**13.7**
8	1.2	2.9	6.4	12.8	17.1	**11.8**
9	1.1	3.2	8.2	12.7	19.5	**10.8**
10	1.3	3.6	5.9	8.0	16.8	**8.8**
11	1.2	4.9	7.2	9.4	10.3	**6.3**
12	1.2	4.9	7.2	9.4	10.3	**6.3**
13	1.1	2.9	5.2	9.4	11.8	**6.4**
14	1.0	3.4	4.7	10.6	17.4	**6.9**
Total	**0.9**	**2.7**	**5.2**	**8.4**	**10.8**	**7.0**

Source: CSRGT data from 2004-07 for GB, processed for BYFM model.

14 ANNEX E: SUMMARY OF DATA IMPROVEMENT TASKS

14.1.1 The set of requirements for the data improvements needed to support the implementation of the NFTM has been discussed throughout the think piece and is summarised here.

ROAD FREIGHT

14.1.2 The top priority is to improve access to logistic data that can quantify for each commodity type (i) its pattern of distribution chains and (ii) the characteristics of their sequence of successive logistic legs. The required steps are.

a) DfT to carry out the feasibility pilot study (Annex B, 11.1) of post processing a sub-set of the CSRGT data to allocate a land use type to each trip end – i.e. for each consignment pick-up and drop location.
b) If this allocation pilot test is successful, then apply the same approach in an automated fashion to a few successive recent years for the complete CSRGT.
c) If this feasibility test is unsuccessful, enhance the CSRGT to start to collect information on land use type at each trip end

14.1.3 Although remotely sensed HGV movement data that is based on GPS or other technologies may be useful, the way in which the data is collected needs to be fully understood. Unless there is a guaranteed unbiased method to expand from the sample collected up to the full population of vehicle movements of interest, it could be misleading. Moreover, as such remote data may lack information on the type and weight of goods carried and on the vehicle type, its information content may be limited. If these issues can be addressed, then remotely sensed data could be excellent for achieving improved spatial detail in the OD matrices.

14.1.4 A longer term option would be to develop a new automated shipment-based truck survey or perhaps even a full survey of supply chains (Annex B, 11.2).

RAIL FREIGHT

14.1.5 Obtain from Network Rail an OD matrix of annual actual[36] tonnes of goods moved during the base year, segmented by type of commodity (ideally matching to the NST commodity type category used in CSRGT) while providing the maximum feasible level of zonal detail (based on the model's zoning system), including distinguishing ports and intermodal transfer terminals.

14.1.6 Obtain guidance on the broad rail tariff structure for each type of movement.

MARITIME AND PORTS

14.1.7 The availability of data for the maritime side of ports generally is good. In contrast, port statistics are almost non-existent relating to transport activity on the inland side of the port. Because ports are the locations where much of the UK competition between road and rail (and waterborne) modes occurs, they could provide an effective means for identifying those contexts where each of the inland modes is more successful. Section 4.4 outlines a medium priority task to collect key data items about inland transport usage for a suitable sample of import and of export consignments passing through the UK ports (plus Eurotunnel).

WAREHOUSING LOCATIONS AND COSTS

14.1.8 Because of the importance of different types of distribution centres within the intermediate stages of PC movements, good data on their locations, size and characteristics is needed. Access to this data at a level of detail suitable to meet the model's requirements needs to be confirmed across the UK, via the VOA and Scottish Assessors (Section 5.5).

14.1.9 It will also be necessary to locate suitable data sources on typical unit warehousing costs of each commodity type within NDCs, RDCs and LDCs, to support the calculation of warehouse gate costs (para. 6.2.10).

ENSURING CONSISTENCY IN FREIGHT DATA DEFINITIONS

14.1.10 A key message throughout this think piece is the need to ensure full consistency in freight data definitions whenever data from different data sources are combined. Annex A provides examples of some of the types of issues that need to be addressed. Unless there is appropriate resolution of the large number of potential inconsistencies arising between the many data sources required to support a national freight model, the quality of that model may be compromised.

[36] See the discussion from para. 10.1.11 above onwards, on definitional issues with "net" tonnage on rail.