



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/148435/>

Version: Accepted Version

Article:

Hickish, R., Fletcher, D. and Harrison, R. (2019) A rail network performance metric to capture passenger experience. *Journal of Rail Transport Planning and Management*, 11. 100138. ISSN: 2210-9706

<https://doi.org/10.1016/j.jrtpm.2019.06.002>

Article available under the terms of the CC-BY-NC-ND licence
(<https://creativecommons.org/licenses/by-nc-nd/4.0/>).

Reuse

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) licence. This licence only allows you to download this work and share it with others as long as you credit the authors, but you can't change the article in any way or use it commercially. More information and the full terms of the licence here: <https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

1 A rail network performance 2 metric to capture passenger 3 experience

4 Bob Hickish¹, David I Fletcher^{1*} and Robert F
5 Harrison²

6 ¹Department of Mechanical Engineering, The University of
7 Sheffield, UK

8 ²Department of Automatic Control and Systems Engineering,
9 The University of Sheffield, UK

10 *Corresponding author: d.i.fletcher@sheffield.ac.uk,
11 Department of Mechanical Engineering, The University of
12 Sheffield, Mappin Street, Sheffield, S1 3JD, UK

13 Declarations of interest: none

14 Abstract

15 For passenger rail operators worldwide a common concern is to
16 better understand and improve passenger experience. Based on
17 factors including train movement times and crowding, the
18 Journey Time Metric and Disutility Metric can be used to
19 quantitatively assess the journey experience of individual
20 passengers. However an assessment of overall network
21 performance is also desirable. This paper presents a whole-
22 network assessment metric that captures passenger experience
23 by aggregating and normalizing individual journey
24 assessments. The newly developed metric is validated against
25 customer satisfaction data measured in passenger surveys of the
26 London Underground Limited Victoria Line with a statistically
27 significant correlation ($P < 0.005$) between the predictions and
28 the measurements. It is found that there is a high degree of
29 correlation ($\rho = 1.00$, $P < 0.005$) between the network scores
30 calculated using the new whole-network assessment metric
31 with either the Journey Time Metric or Disutility Metric despite
32 their different formulations and countries of origin. Through
33 development of the new metric it is identified that many
34 commonly used network assessment metrics (e.g. Public
35 Performance Measure and the end-to-end journey time of
36 passengers) are insensitive to crucial aspects of passenger
37 experience. The newly developed metric could be used by rail
38 operators to better select strategies for improving passenger
39 experience.

40

41

42 **Keywords**

43 Journey Time Metric, Disutility Metric, rail, network,
44 passenger, assessment

45 **Highlights**

- 46 • A new whole-network assessment metric is developed
47 to capture passenger experience by aggregating and
48 normalizing individual journey assessments.
- 49 • Two different passenger journey assessment metrics of
50 different international origin are compared.
- 51 • The new whole-network assessment metric is validated
52 against measured data from the London Underground
53 Limited Victoria Line.

Nomenclature

ψ – individual passenger journey score*
 Ψ – distribution of passenger journey scores
 ϕ – network score*
 I – number of states in a passenger journey
 i – counter for enumerating sequence of states
 t_i – time passenger spends in their i^{th} state (seconds)
 Ω – Value of Time weighting function*
 α_i – passenger journey stage of i^{th} state*
 β_i – vector describing conditions of passenger's i^{th} state*
 ε – number of passenger train changes
 ω – crowding penalty function*
 δ – number of passengers on train
 δ_{\max} – train maximum capacity
 γ – train crush capacity
 μ – number of seats
 η – crowding factor
 c_1 to c_3 – constants
 k_1 to k_7 – constants
 R – number of passengers
 d_r – distance travelled by r^{th} passenger
 τ_B – Kendall Rank Correlation Coefficient B*

* Values specific to a metric are indicated with the superscript text: *JTM* or *DM*

54 **1 Introduction**

55 With demand for rail travel having doubled in the last 20 years
56 (Davis, 2018) and 40% more passengers predicted by 2040
57 (Carne, 2018), rail travel has an increasingly important role to
58 play in meeting the passenger journey needs of Great Britain
59 (GB). To fulfil this role the rail industry Technical Leadership
60 Group (2017) set targets for the GB network that included

61 “improving [the] customer experience” of passengers. The
62 experience of passengers and their satisfaction is also a network
63 performance indicator for other rail networks internationally,
64 e.g. across Europe (TNS Political & Social, 2013) and in Japan
65 (Kunimatsu et al., 2012). Traditionally, however, rail networks
66 have been assessed with train-focussed metrics. For example,
67 the GB industry standard Public Performance Measure (PPM)
68 describes the percentage of services that arrive at their *final*
69 destination within five minutes (ten for long distance trains) of
70 the timetabled time, this metric having no sensitivity to the
71 effect on passengers if the train arrives late at intermediate
72 stations, or to the comfort of their journey. In this paper a new
73 method is developed which combines assessments of individual
74 passenger journeys, i.e. *journey scores*, for all passengers in a
75 network to give a *network score* that quantifies the experience
76 of passengers. In a case study relating to the Victoria Line of
77 the London Underground Limited (LUL) network, the whole-
78 network assessment metric is validated against measured data
79 from passenger surveys surmised by LUL (2018a).
80 Furthermore, international comparison is made when the
81 whole-network assessment metric is used with individual
82 passenger journey assessment metrics from different countries
83 of origin. The developed whole-network assessment metric will
84 allow operators to provide a parameter summarising overall
85 network performance from the passenger perspective, enabling
86 this to be effectively optimised.

87 2 Metrics to assess networks

88 The aggregate of passenger end-to-end journey time has been
89 used as a metric to assess network performance, for example by
90 Vuchic and Newell (1968), Chang et al. (2000) and Cacchiani
91 and Toth (2012). However, there is evidence that end-to-end
92 journey time does not fully capture the passenger experience.
93 For example, Susilo and Cats (2014) show that, for public
94 transport travellers, factors such as station environment, ease of
95 transfer, service frequency and safety are significant
96 determinants of passenger satisfaction. Because Chen and Chen
97 (2010) describe customer satisfaction as being affected by
98 customer experience, in the current paper it is assumed that the
99 satisfaction of a passenger is an indicator of their experience,
100 and the effect of other factors such as ticket pricing is
101 disregarded. Consequently, in the current paper, decreasing
102 passenger dissatisfaction or disutility and increasing passenger
103 satisfaction are considered to be equivalent to “improving
104 passenger experience”. The disconnect between passenger
105 journey time and passenger satisfaction is evident in the results
106 of a rail passenger survey by Transport Focus (2016) which
107 showed that journey time has a smaller influence upon
108 passenger satisfaction than punctuality of the service or
109 cleanliness. Therefore, to better capture passenger satisfaction

110 it is necessary to quantify a passenger journey in greater depth
111 than journey time or punctuality alone.

112

113 2.1 Describing a passenger journey with stages

114 A passenger journey can be modelled as the combination and
115 repetition of specific activities, i.e. *stages*. For example, Wang
116 et al. (2015) state that a passenger journey can be well
117 represented with the stages: walking into and out of a station,
118 waiting on the platform, riding on a train and transferring
119 between platforms. However, they do not take into account the
120 relative impact of time in each stage upon the whole passenger
121 experience. Vansteenwegen and Van Oudheusden (2007) and
122 Sels et al. (2016) describe a passenger journey using two stages
123 (“In Station” and “On Train”) and capture the varying impact
124 of time in different stages by weighting these times with a
125 different Value of Time (VoT). The VoT concept has been
126 developed in Transport Economics and describes, in monetary
127 terms, the disutility experienced by a passenger over a time
128 period. It can be thought of as the price a passenger would pay
129 to reduce their travel time by one unit, hence a greater VoT
130 indicates a worse experience for passengers. As well as being
131 sensitive to the journey stage of a passenger, a VoT can be
132 sensitive to the mode of transport, journey purpose and
133 distance, for example having different values for travel by car,
134 bus, train or other public transport (ARUP et al., 2015).
135 Wardman (2004) showed that the VoT is sensitive to the
136 activity of the passenger, and Vansteenwegen and Van
137 Oudheusden provide values showing that passengers rate 1
138 minute of waiting in a station to be equivalent to 2.5 minutes on
139 a moving train. By modelling the amount of time passengers
140 spend in both of these stages and weighting it by the VoT for
141 each stage, Vansteenwegen and Van Oudheusden create a
142 network assessment metric which can capture the relative effect
143 on passengers of time savings in either stage. However, their
144 metric does not capture the effect of *crowding* (i.e. the number
145 of passengers on a train relative to the number of seats and
146 standing space) which can reduce the personal space and
147 comfort of passengers, causing additional disutility and hence
148 increasing the VoT.

149 Horowitz (1978) showed that, as well as the journey stage, the
150 “environmental conditions” that a passenger experiences during
151 a stage (referred to as *conditions* in the current paper) affect the
152 VoT. Horowitz considered weather conditions, that are not
153 considered here, but also standing vs seated travel and
154 crowding levels. Models to quantify the effect that crowding
155 has upon the VoT have been developed for example by

156 Wardman and Murphy (2015) and Qin (2014). Two metrics
157 developed in different international systems to assess individual
158 passenger journeys across journey stages and crowding levels
159 are the Journey Time Metric (JTM) and the Disutility Metric
160 (DM).

161
162 The JTM has been developed by LUL and shared with the
163 authors by private communication, the most informative
164 accessible documentation being the investigations of Chan
165 (2007) and Hickey (2011). It describes passenger journeys
166 using five stages “Buying Ticket”, “Moving Through Station”,
167 “On Platform”, “On Platform (Left Behind)” (where a
168 passenger has not been able to board a suitable train because it
169 is overly occupied) and “On Train”. The effect of crowding
170 conditions are considered in the “On Train” stage by modifying
171 the VoT with a *crowding penalty* that is dependent on the
172 number of passengers, train capacity and seats. The DM has
173 been developed in Japan and is documented in English by
174 Kunimatsu et al. (2009, 2012). It takes a similar approach to the
175 JTM, but resolves a journey using two stages (“On Train” and
176 “In Station”) with weightings different to those used by the
177 JTM. Similar to the JTM, the DM applies a crowding penalty
178 for passengers in the “On Train” stage that is sensitive to the
179 same factors as the JTM crowding penalty, however a different
180 formula is used. The DM is used again by Kanai et al. (2011) to
181 assess individual journeys as part of a network assessment
182 metric used in a decision support tool for delay management.
183 They discuss different methods of combining journey scores
184 into a network score, however none of their methods normalize
185 for the distance travelled by passengers, meaning that networks
186 providing shorter journeys could compare favourably against
187 networks providing longer journeys.

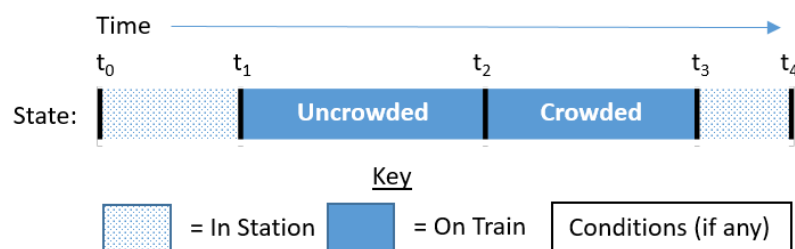
188
189 Moving from individual journey to network metrics, Ali et al.
190 (2017) predict network performance by combining journey
191 scores calculated using an individual journey metric with
192 similarities to the JTM and DM. The network metric is
193 demonstrated to predict observed simple qualitative
194 relationships between timetable features and network
195 performance, e.g. fewer train services result in worse network
196 performance as determined by their metric.

197
198 The JTM, DM and the metric described by Ali et al. are the
199 only metrics, found for this review, to capture the multi-stage
200 nature of passenger journeys and weight the time spent in each
201 stage *including* the effect of crowding. They therefore capture
202 individual passenger journeys in more detail than the other
203 metrics identified here which consider journey stages *or*
204 crowding only. However, the parameter values used within the
205 metric of Ali et al. could not be retrieved so this is excluded

206 from further analysis. To the best of the authors' knowledge, no
 207 publicly available documents describe the validation or
 208 comparison of the JTM and DM, or network assessment
 209 metrics based upon them. This gap defines the targets of this
 210 paper, to make a comparison of the JTM and DM methods, and
 211 to develop a validated network metric based upon them.

212 3 Network assessment metrics that capture 213 the passenger perspective

214 To assess a rail network we evaluate individual passenger
 215 journeys and examine the distribution of experiences. To
 216 evaluate modelled passenger journeys, we introduce the term
 217 *state* to describe a specific combination of journey stage and
 218 conditions. A passenger journey is decomposed into a sequence
 219 of states as shown in Figure 1, which illustrates an example
 220 four-state passenger journey. Shading is used to indicate which
 221 journey stage the passenger is in (“On Train” or “In Station”).
 222 Crowding is only considered in the “On Train” stage and text is
 223 used to indicate this. The markers t_0 to t_4 indicate the times at
 224 which the passenger changed state. At t_0 , the passenger enters
 225 the origin station, at t_1 the passenger boards their train. At t_2 the
 226 train stops at an intermediate station where more passengers
 227 board making it crowded. The passenger journey *stage* does not
 228 change, but the *state* does. At t_3 the passenger reaches their
 229 destination station and exits at t_4 . The number of states in a
 230 passenger journey, I , is variable dependant on the journey and
 231 we use the counter, i , to enumerate the sequence of states, $i = 1,$
 232 $2, \dots I$.



233
 234
 235
 236
 237
 238

Figure 1 – An example passenger journey decomposed into four states. The journey is described with two stages: On Train and In Station. The shading of the state indicates the stage. Text is used to describe the conditions of the state. The markers t_0 to t_4 relate to the times when the passenger changed state.

239 The sum of VoT weightings across all states of a passenger
 240 journey can be used as an individual journey score. The
 241 following section describes how this is calculated when either
 242 the JTM or DM is used. The following section also compares
 243 how the JTM and DM calculate the crowding penalty. Section
 244 3.2 then describes how the distribution of journey scores is
 245 evaluated to give a network score.
 246

247 3.1 Calculating an individual journey score
 248 A journey score calculated using the JTM is computed from the
 249 formula:

$$250 \quad \psi^{JTM} = \sum_{i=1}^{i=I} t_i \Omega^{JTM}(\alpha_i^{JTM}, \beta_i^{JTM}, \omega^{JTM})$$

251 (1)

252 Where ψ denotes the journey score, t_i , the time (in seconds)
 253 spent in the i^{th} state, Ω , the VoT weighting function, α_i and β_i ,
 254 respectively the journey stage and conditions of the passenger's
 255 i^{th} state and ω the crowding penalty function. ψ^{DM} (given by
 256 (2)) is calculated similarly to ψ^{JTM} , but has an additional term
 257 to capture the relative disutility experienced by passengers
 258 changing train with a parameter for the number of times a
 259 passenger must change trains, ε , and a weighting factor, k_1 . A
 260 value of 600 is used by Kunimatsu et al. for k_1 , meaning that
 261 each train change has an associated disutility equivalent to 10
 262 minutes (600 seconds) travelling on an otherwise unoccupied
 263 train. Table 1 provides the other parameter values for each
 264 metric.

$$265 \quad \psi^{DM} = \sum_{i=1}^{i=I} t_i \Omega^{DM}(\alpha_i^{DM}, \beta_i^{DM}, \omega^{DM}) + k_1 \varepsilon$$

266 (2)

$\alpha_s^{JTM} =$	1	2	3	4	5
Description	On Train	On Platform	On Platform (Left Behind)	Moving Through Station	Buying Ticket
$\Omega^{JTM} =$	$1 + \omega^{JTM}(\beta_i^{JTM})$	2.5	3	2.7	2.5
$\alpha_i^{DM} =$	1	2			
Description	On Train	In Station			
$\Omega^{DM} =$	$1 + \omega^{DM}(\beta_i^{DM})$	3			

267 *Table 1 – The VoT weighting, Ω , for both metrics dependent on the journey stage, α ,*
 268 *of a passenger's i^{th} state. A description of the journey stage relating to α is also*
 269 *shown. The VoT weighting for the On Train state is dependent on a crowding penalty*
 270 *function, ω , calculated using the conditions of the state, β . For the JTM, these*
 271 *values have been shared with the authors by personal communication and for the*
 272 *DM they are taken from Kunimatsu et al. (2012).*

273 Table 1 shows the relative weighting both metrics put on each
 274 state (a lower value of Ω indicates a better passenger
 275 experience) and that the JTM describes a journey using five
 276 journey stages whereas the DM uses two. Both methods

277 consider crowding only when passengers are in the “On Train”
 278 journey stage. The JTM crowding penalty, ω^{JTM} , is determined
 279 with the formula given by (3) using values given in Table 2.

$$280 \quad \omega^{JTM} = \begin{cases} 0, & \delta \leq \mu \\ c_1 + c_2 \frac{\delta - \mu}{\gamma} - c_3 \frac{\delta\mu - \mu^2}{\gamma^2}, & \mu < \delta \leq \delta_{\max} \end{cases} \quad (3)$$

281
 282 Where δ denotes the number of passengers, μ , the number of
 283 seats on the train, δ_{\max} , the maximum passenger capacity, γ ,
 284 the crush capacity and c_1 to c_3 constants. The crowding penalty
 285 formula given by (3) has been shared with the authors by
 286 personal communication from the Transport Planning
 287 department of LUL (Kelt, 2015). The second term of (3)
 288 captures the number of standing passengers relative to the crush
 289 capacity of the train and the third term captures the effect of
 290 seated passengers also. The value of γ describes the theoretical
 291 maximum number of people that can fit into the train assuming
 292 seven passengers per square meter of standing floor space.
 293 However, LUL have determined that the practical maximum
 294 capacity of a train is less than γ and under “normal operating
 295 conditions” the value of δ_{\max} is defined as 71% of γ . The DM
 296 crowding penalty, ω^{DM} , is determined with the formula given
 297 by (4) and requires computing the crowding factor, η , given by
 298 (5). The constants k_2 to k_7 and c_1 to c_3 are shown by Table 2.

$$299 \quad \omega^{DM} = \begin{cases} k_2\eta, & \eta < 1 \\ k_3\eta - k_4, & 1 \leq \eta < 1.5 \\ k_5\eta - k_6, & 1.5 \leq \eta \leq 2 \end{cases} \quad (4)$$

$$300 \quad \eta = \frac{k_7\delta}{\delta_{\max}} \quad (5)$$

303

Name	c_1	c_2	c_3	k_2	k_3	k_4	k_5	k_6	k_7
Value	0.85	1.915	1.03	0.027	0.0828	0.0558	0.179	0.2	2

304 *Table 2 - Constant values used to calculate the crowding penalty, ω^{JTM} and ω^{DM} , in*
 305 *(3) and (4). For the JTM, these values have been shared with the authors by*
 306 *personal communication and the DM constants k_2 to k_6 are taken from Kunimatsu et*
 307 *al. (2012). The value of k_7 is informed by Nippon (2018).*

308 The values of c_1 to c_3 have been derived by LUL and shared
 309 with the authors by personal communication (Kelt, 2015). The
 310 values of k_2 to k_6 are listed by Kunimatsu et al. (2012).
 311 Although Kunimatsu et al. do not explicitly define η , they
 312 describe it as the “congestion rate of the train”, therefore it can
 313 be inferred as being proportional to δ/δ_{\max} . However because
 314 Nippon (2018) report the largest crowding factor (η) observed
 315 in Japan during 2017 as 2 (relating to when “bodies come into
 316 contact with each other and one feels considerable pressure”),
 317 the scaling factor k_7 is introduced into (5) and given a value of

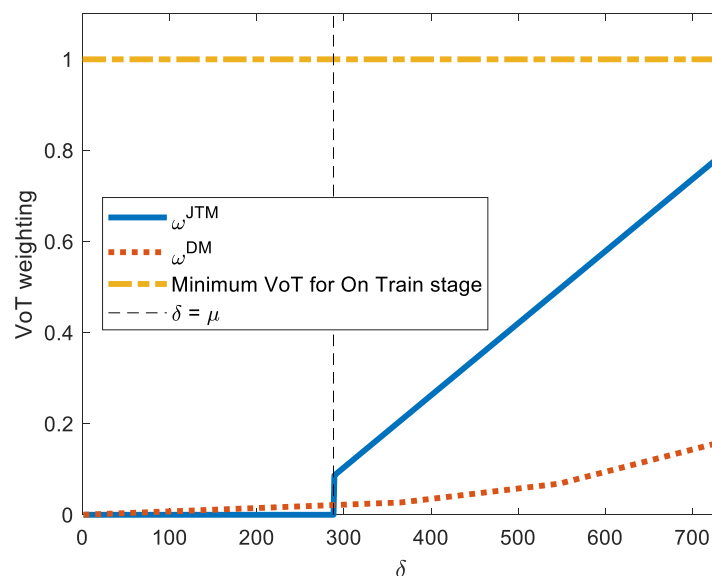
318 2. The values of μ , δ_{\max} and γ are rolling stock specific and are
 319 defined by LUL for each fleet. For the LUL 2009 rolling stock
 320 (used on the Victoria Line and the subject of this investigation)
 321 their values are 288, 730 and 1028 respectively (Kelt, 2015)

322

323 Figure 2 compares ω^{JTM} and ω^{DM} on the y-axis for varying
 324 number of passengers (δ). The number of seats on the train is
 325 shown by a vertical dashed line and reflects that when $\delta \leq \mu$,
 326 the JTM does not apply a crowding penalty. A crowding
 327 penalty is applied by the DM even at this level of occupancy,
 328 but it is small in comparison to the minimum VoT weighting
 329 for passengers in the “On Train” journey stage (the dash-dot
 330 horizontal line). When $\delta > \mu$, the JTM applies a crowding
 331 penalty that is 4 to 8 times greater than the DM crowding
 332 penalty. For both metrics, the crowding penalty is always less
 333 than the minimum VoT weighting for the “On Train” stage.
 334 Both the JTM and DM models of crowding assume that
 335 passengers are homogenously distributed throughout the train
 336 and that passengers will always find and occupy a seat if one is
 337 available. Although this may not be realistic, it is the same for
 338 both models so the comparison is like-for-like.

339

340 The VoT weightings (in Table 1) and crowding penalty
 341 function for the JTM and the DM have been derived for the
 342 LUL network and Japanese railway respectively. It is therefore
 343 expected for these values to capture local preferences and
 344 expectations.



345

346 *Figure 2 - The crowding penalty, ω , applied by the JTM and the DM for different*
 347 *numbers of passengers, δ , in LUL 2009 rolling stock up to its maximum capacity.*
 348 *The number of seats, μ , is shown by a vertical dash line. The minimum VoT*
 349 *weighting applied by both metrics to passengers that are in the “On Train” stage is*
 350 *shown by a horizontal dash-dot line.*

351 3.2 Calculating a network score from journey scores
 352 Networks provide journeys for multiple passengers so there is a
 353 distribution of journey scores. To improve the network
 354 assessment metric and ensure that journey scores only capture
 355 the quality of the service provided to the passenger by the
 356 network (and not the distance of the passenger journey which is
 357 a passenger choice), we normalize journey scores by the
 358 distance travelled. This allows like-for-like comparison of
 359 journey scores within the distance-normalized journey score
 360 distribution, Ψ , given by:

$$361 \quad \Psi = \left[\frac{\psi_1}{d_1}, \frac{\psi_2}{d_2}, \dots, \frac{\psi_R}{d_R} \right] \quad (6)$$

362 Where ψ_r and d_r respectively denote the journey score and
 363 distance travelled relating to the r^{th} passenger and R the number
 364 of passengers. Different features of Ψ can be used to provide
 365 the network score, ϕ , for all R passengers conveyed. Although
 366 we wish to capture the effect of passenger numbers upon
 367 crowding, we also wish the network score to be independent of
 368 the number of journey scores within Ψ . Consequently, an
 369 additional passenger-number normalization step is included so
 370 ϕ^{JTM} and ϕ^{DM} are defined by:

$$371 \quad \phi = \frac{1}{R} \sum_{r=1}^{r=R} \frac{\psi_r}{d_r} \quad (7)$$

372 Beyond this network score the characteristics of the distribution
 373 of Ψ can offer additional insight. For example, an operator
 374 wishing to examine the consistency of their service to
 375 passengers taking different journeys may evaluate the range of
 376 Ψ in addition to ϕ . In the current paper we focus primarily on
 377 ϕ to study quality of service provided to all passengers within
 378 the network.

381 4 Validation and comparison

382 To validate the network assessment metric, ϕ values are
 383 calculated using either the JTM or DM (ϕ^{JTM} or ϕ^{DM}) for the
 384 Victoria Line of the LUL network. For the same network, a
 385 network score is determined from measured Customer
 386 Satisfaction Survey (CSS) data, ϕ^{CSS} . The predictive values of
 387 ϕ^{JTM} and ϕ^{DM} are compared against the measured ϕ^{CSS} values
 388 and the correlation between their changes relative to a baseline
 389 year is quantified. The predictive values are then compared to
 390 each other to determine a relationship between the network
 391 assessment metric when either journey score metric is used. To
 392 calculate ϕ^{JTM} and ϕ^{DM} data describing the network operation
 393 was combined with data describing the passenger load and
 394 captures the effect of varying timetables and passenger loads

395 over ten years. For the Victoria Line in the period investigated,
396 the formation, length and interior layout of rolling stock remain
397 constant, therefore the frequency of trains (determined by the
398 timetable) has the greatest effect upon the passenger carrying
399 capacity of the network. Decreasing the speed of trains on a
400 line slows travel but also reduces headway with potential to
401 decrease intervals between trains, so typically there is a trade-
402 off between journey times and frequency. To meet increasing
403 demand for travel, minimise crowding and generate more
404 revenue, whilst maintaining competitive journey times against
405 other transport modes, there is a pressure on LUL to balance
406 this trade-off when updating their timetable.

407 4.1 Data sources

408 The data sources used in this investigation are: Victoria Line
409 Working Timetable (WTT) numbers 31 to 41 (London
410 Underground Limited, 2007, 2009, 2011, 2012a, 2012b, 2014,
411 2015a, 2015b, 2016b, 2016c, 2017), Access, Egress and
412 Interchange (AEI) data provided by LUL (2016a), the
413 Performance Data Almanac (PDA) (London Underground
414 Limited, 2018a) and the Rolling Origin Destination Survey
415 database (RODS) (London Underground Limited, 2018b). In
416 the following section, the data is described in more detail.
417

418 4.2 Input data

419 The network operation data is taken from the WTTs and the
420 AEI data. For each day, the WTTs provide the average train
421 frequency and interstation run times for the three weekday
422 operational periods on which our investigation concentrates:
423 Morning Peak, Midday Off Peak and Evening Peak. Later
424 operational periods are excluded because their timings are not
425 consistent between the WTTs. The effect of this exclusion is
426 unlikely to be significant because observing the RODS
427 database indicates that this period is when the fewest
428 passengers travel and so it has the least weighting on the
429 network score. Weekends and holidays are not considered
430 because they are more likely to be affected by events (e.g.
431 sporting events or planned line closures for maintenance works)
432 that affect passenger experience but are not captured in all the
433 input data sources. The operational pattern described in the
434 WTT is applied for every day the timetable was in effect (LUL
435 update their timetable irregularly, but the date of introduction is
436 provided by each WTT). The WTTs also provide the distance
437 between adjacent station pairs. The AEI data describes the
438 passenger travel time from station door to platform and vice
439 versa, and platform to platform. The AEI data available relates
440 to every four week period of the year beginning 2011 (the LUL
441 reporting year begins on 1st April), over which the year mean is
442 2.23 minutes. Because data is only available for one year, this

443 is applied for all years of the investigation, implicitly assuming
444 that personal mobility within the station remains constant over
445 this period.

446

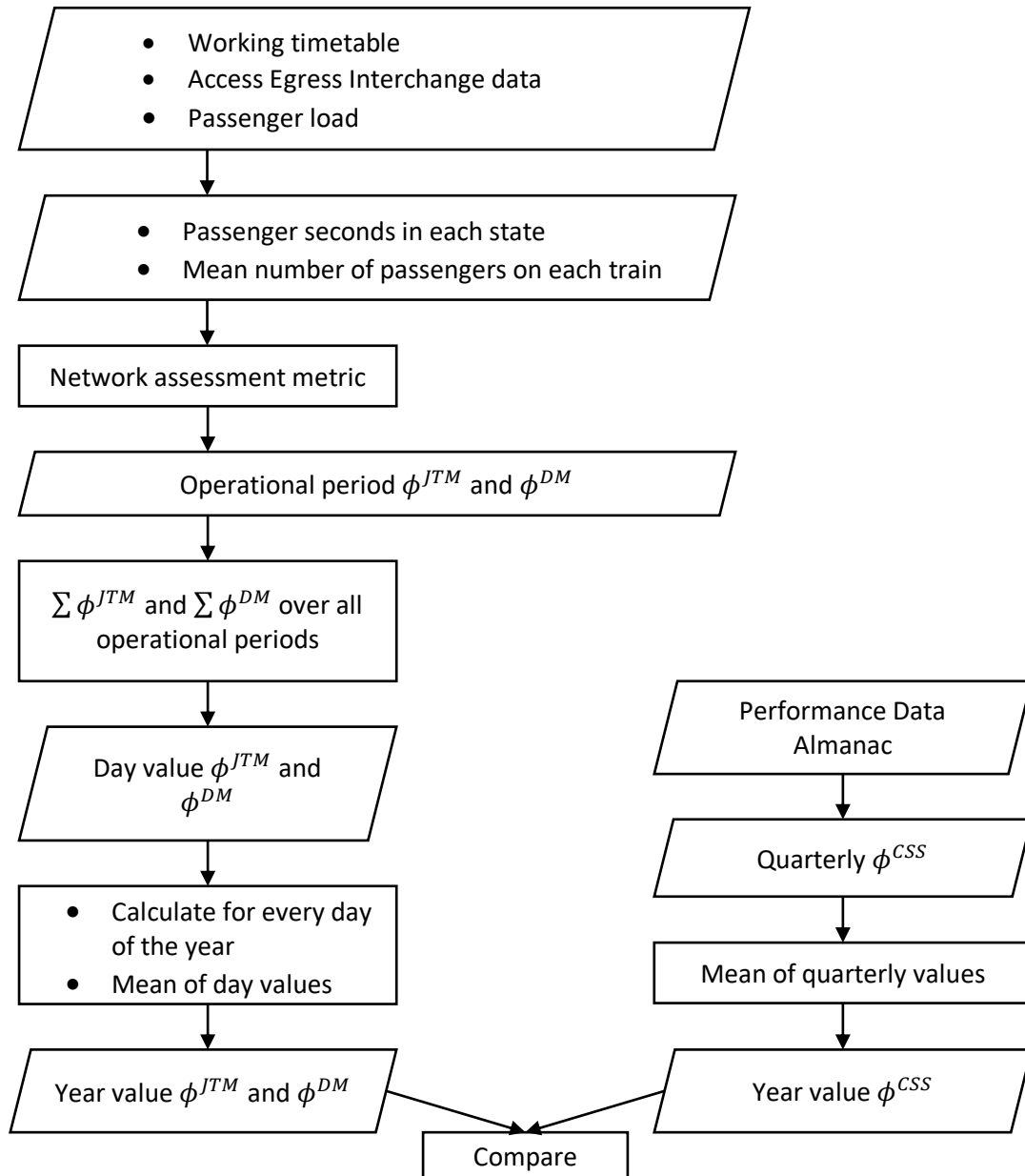
447 The passenger load data is a combination of two data sources:
448 the PDA and RODS. RODS provides the proportion of
449 passengers included within the database that travel between
450 adjacent station pairs in an operational period, i.e. *line section*
451 *loadings*. However, this data does not describe whole
452 passenger journeys (i.e. an origin and destination with any
453 transfer stations). The PDA provides the total number of
454 passengers travelling on the Victoria Line each year, and the
455 quarterly CSS data. To collect the CSS data, LUL use
456 questionnaires to ask approximately 2,500 passengers per
457 quarter to rate, on a scale of 1 to 10, their satisfaction with their
458 travel on the line of the last leg of their journey. The mean of
459 the ratings is then multiplied by 10 and reported for each line
460 by LUL.

461

462 4.3 Methodology

463 To calculate ϕ^{JTM} and ϕ^{DM} , the line section loading data was
464 scaled by the yearly passenger numbers data and used to
465 disaggregate the journeys of passengers who travelled further
466 than the station adjacent to their origin, into a series of journeys
467 between adjacent station pairs. For each operational period
468 (Morning Peak, Midday Off Peak and Evening Peak) and line
469 section, the number of passengers per train was calculated by
470 dividing the number of passenger journeys in that period by the
471 number of trains. Where demand for travel exceeded provision,
472 the excess passengers were modelled as being “left behind” by
473 one train before catching the next. The frequency of trains was
474 used to determine the total passenger time spent in the “On
475 Train”, “On Platform” and “On Platform (Left Behind)” stages.
476 The journey score metrics were used to calculate the VoT
477 weighting for these states. To avoid over-counting, the AEI
478 time and weighting was only applied twice for each whole
479 passenger journey defined by the PDA data rather than the
480 RODS data. The “Buying Ticket” journey stage was
481 disregarded because the use of pre-paid travel cards (“Oyster”
482 cards) and contactless payment at ticket gates is common for
483 this network. For example, in 2012 Oyster cards were used for
484 over 80% of public transport travel in London (Transport for
485 London, 2012). The inter-station distances were multiplied by
486 the line section loadings so that the aggregate of the VoT
487 weightings could be normalized by the total passenger distance
488 travelled. This analysis was conducted for the Morning Peak,
489 Midday Off Peak and Evening Peak operational periods of
490 every weekday and was dependent on the daily timetable and
491 yearly number of passenger journeys. To calculate the network
492 score for that day, the values from the three operational periods

493 of the day were summed. The year value was calculated as the
 494 mean of the year's day values. This process is illustrated by
 495 Figure 3 which shows ϕ^{JTM} , ϕ^{DM} and ϕ^{CSS} being calculated for
 496 corresponding years so that comparison is like-for-like.
 497 Because the CSS data is already normalized for passenger
 498 numbers and distance travelled, it is not relevant to normalize
 499 ϕ^{CSS} using (7).
 500



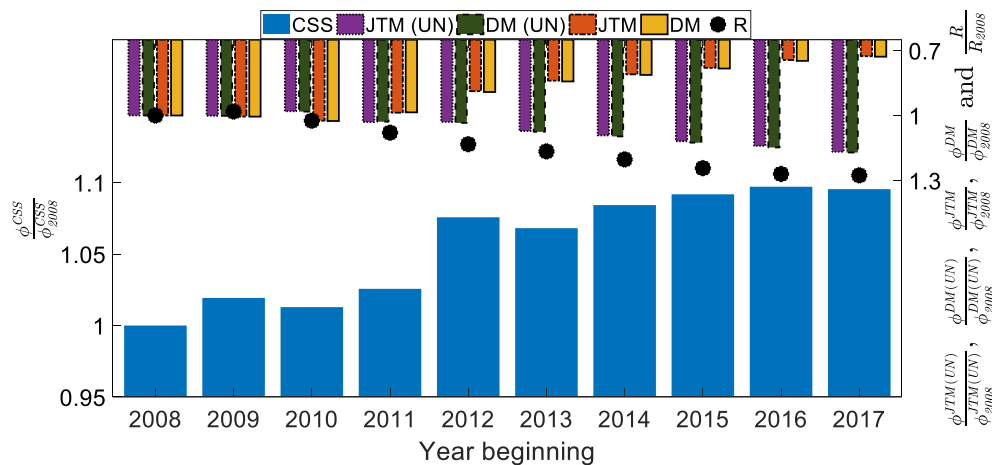
501
 502 *Figure 3 - The method for calculating the measured network score, ϕ^{CSS} , and*
 503 *predicted network score using the Journey Time Metric or Disutility Metric, ϕ^{JTM}*
 504 *and ϕ^{DM} respectively, from the Working Timetable (WTT), Access Egress and*
 505 *Interchange (AEI) data, passenger load data and Customer Satisfaction Survey*
 506 *(CSS) data.*

507 4.4 Results

508 Figure 4 enables comparison of ϕ^{CSS} with ϕ^{JTM} and ϕ^{DM} , and
 509 also presents data where no distance or passenger normalization

510 is applied, $\phi^{JTM(UN)}$ and $\phi^{DM(UN)}$, for the years 2008 to 2017.
 511 The number of passengers, R , is also included in the plot.
 512 Upward-pointing bars with values displayed on the left ordinate
 513 are used for ϕ^{CSS} , while $\phi^{JTM(UN)}$, $\phi^{DM(UN)}$, ϕ^{JTM} and ϕ^{DM}
 514 are represented by downward-pointing bars with values
 515 displayed on the right ordinate. Because the prediction metrics
 516 measure dissatisfaction and ϕ^{CSS} measures satisfaction, the
 517 right ordinate is inverted. A positive change in the vertical
 518 position of a bar-top for ϕ^{CSS} indicates a “better” performing
 519 network. R is also represented by markers with values
 520 displayed on the right ordinate. To allow comparison of relative
 521 changes on different scales and using different units, all series
 522 have been normalized against their 2008 value.

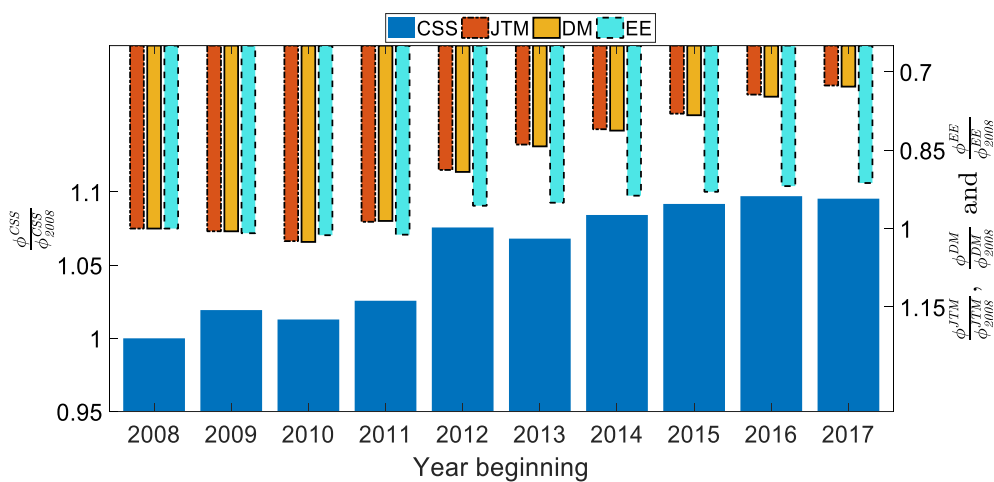
523
 524 It can be seen that over time, in general, the measured network
 525 scores (ϕ^{CSS}) indicate improving network performance, with
 526 rising values relative to 2008. In general, this behaviour is
 527 successfully predicted by ϕ^{JTM} and ϕ^{DM} . However, $\phi^{JTM(UN)}$
 528 and $\phi^{DM(UN)}$ predict deteriorating network performance and
 529 correlate with the increasing passenger numbers. It should be
 530 noted that, whilst the prediction metrics appear to give equal
 531 scores in 2008, this is because of the series normalization
 532 process. The importance of normalizing the predictive values
 533 by passenger numbers and distance travelled is clear if the
 534 metrics are to be compared over time.



535 *Figure 4 - Bar chart to compare predicted and measured network scores for*
 536 *different years and different prediction methods. Measured customer satisfaction*
 537 *scores, ϕ^{CSS} , are shown by the left ordinate. Predictions using the Journey Time*
 538 *Metric, ϕ^{JTM} , Journey Time Metric with no distance or passenger normalization,*
 539 *$\phi^{JTM(UN)}$, Disutility Metric, ϕ^{DM} , and Disutility Metric with no distance or*
 540 *passenger normalization, $\phi^{DM(UN)}$, are shown by the right ordinate which has been*
 541 *inverted. The right ordinate also displays the number of passengers, R . All values*
 542 *have been normalized against the corresponding 2008 value.*

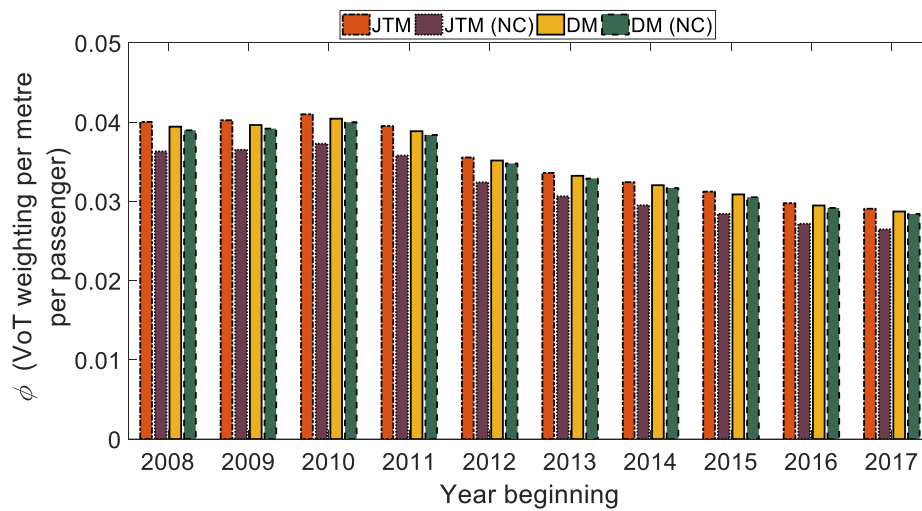
543 To investigate the importance of applying VoT weightings to
 544 different passenger states, Figure 5 enables comparison of
 545 ϕ^{CSS} , ϕ^{JTM} , ϕ^{DM} and a simple end-to-end journey time, ϕ^{EE} .
 546 To ensure like-for-like comparison, ϕ^{EE} has been normalized

547 for passenger numbers and distance. The ordinates are similar
 548 to Figure 4 with the right ordinate now displaying ϕ^{EE}
 549 normalized against the 2008 value. To quantify the level of
 550 agreement between predicted and measured performance,
 551 Kendall's rank correlation coefficient B , τ_B , is calculated
 552 between the series of ϕ^{CSS} with each series of: ϕ^{JTM} , ϕ^{DM} and
 553 ϕ^{EE} . For the series of ϕ^{CSS} with ϕ^{JTM} and ϕ^{CSS} with ϕ^{DM} a
 554 value of -0.82 ($P < 0.005$) is found (-1.0 indicates perfect
 555 (negative) correlation between prediction and measurement and
 556 0 indicates no correlation). For the series of ϕ^{CSS} with ϕ^{EE} a
 557 value of -0.73 ($P < 0.005$) is found, indicating worse correlation
 558 and that network assessment metric is improved by
 559 representing a passenger journey as a series of states and
 560 applying weighting to these.



561 *Figure 5 - Bar chart to compare predicted and measured network scores for*
 562 *different years and different prediction methods. Measured customer satisfaction*
 563 *scores, ϕ^{CSS} , are shown by the left ordinate. Predictions using the Journey Time*
 564 *Metric, ϕ^{JTM} , Disutility Metric, ϕ^{DM} , and end-to-end journey time, ϕ^{EE} , are shown*
 565 *by the right ordinate which has been inverted. All year scores have normalized*
 566 *against the 2008 value for the corresponding metric.*

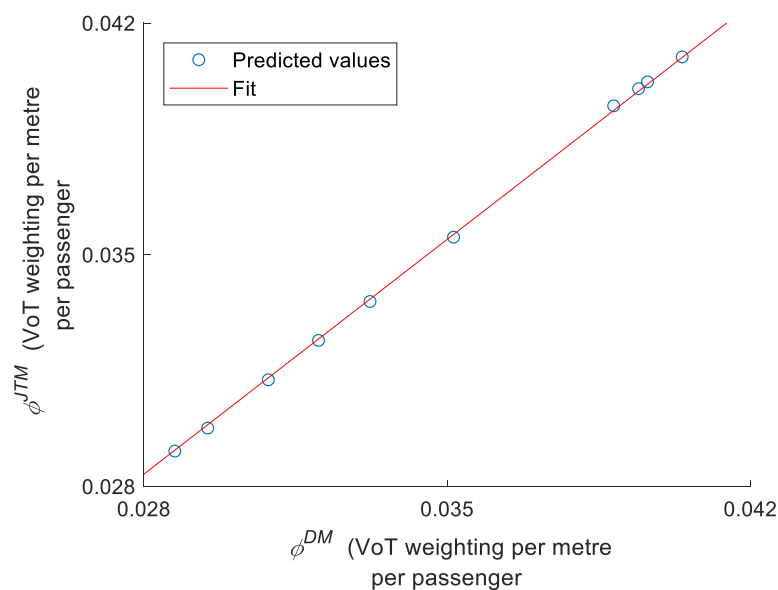
567 To explore the importance of the crowding penalty Figure 6
 568 enables comparison of ϕ^{JTM} and ϕ^{DM} against the case where
 569 no crowding penalty has been applied in the calculation,
 570 $\phi^{JTM(NC)}$ and $\phi^{DM(NC)}$. The y-axis displays the raw values of
 571 ϕ , i.e. they are not normalized against the 2008 value. To
 572 determine what proportion of the network score is contributed
 573 by factors other than the crowding penalty, the value of
 574 $\phi^{(NC)}/\phi$ is calculated. For the JTM and DM series
 575 respectively, a mean value of 0.91 and 0.99 is found both with a
 576 standard deviation less than or equal to 0.002. This behaviour is
 577 discussed in Section 5.
 578
 579



580 Figure 6 – Bar chart to compare the predicted network scores, ϕ , for different years
 581 and different prediction methods. Predictions using the Journey Time Metric, ϕ^{JTM} ,
 582 and the Disutility Metric, ϕ^{DM} , are compared against the case where no crowding
 583 penalty is applied, $\phi^{JTM(NC)}$ and $\phi^{DM(NC)}$ respectively.

584 Figure 7 plots ϕ^{DM} against ϕ^{JTM} for the data from the years
 585 2008 to 2017. The strong linear relationship of the data
 586 ($\rho=1.00$, $P<0.005$) suggests that, in general, similar changes in
 587 network performance are predicted by the JTM and the DM. A
 588 linear fit to this data shows a gradient of 1.013 (95%
 589 confidence bounds of 1.012 and 1.015). The intercept has been
 590 forced to the origin because both metrics are zero under the
 591 same condition: when no passenger time is spent in the
 592 network. The gradient implies that ϕ^{JTM} is consistently
 593 approximately 1.3% greater than ϕ^{DM} , but both are reacting
 594 consistently to external change over the period investigated.

595



596

597 Figure 7- The relationship between the ten network score predictions for the
 598 Victoria Line from 2008 to 2017. The fit has an intercept forced to the origin and a
 599 gradient of 1.013.

600 5 Discussion

601 The results in Figure 4 indicate that, to successfully predict
 602 behaviour of ϕ^{CSS} , it is necessary to normalize the network
 603 assessment metric by the number of passengers and the
 604 distance they travel. In this investigation, the ratio between
 605 different line section loadings remains constant for all years
 606 therefore the value of R plotted in Figure 4 represents changes
 607 to passenger numbers and distance travelled. Consequently, the
 608 results in Figure 4 show that without passenger numbers and
 609 distance normalization, the predicted network scores become
 610 sensitive to both. This effect is unwanted therefore including
 611 passenger number and distance normalization within our
 612 network assessment metric is supported.

613
 614 Choosing a typical significance level of 0.005, the results
 615 shown in Figure 5 are statistically significant evidence that the
 616 null hypothesis (that predicted and measured data are
 617 uncorrelated) can be rejected. Although the choice of
 618 significance level is arbitrary (Wasserstein and Lazar, 2016),
 619 considering the JTM and DM have been developed from
 620 empirical studies of passenger preferences and there is evidence
 621 that end-to-end journey time influences passenger experience
 622 (Transport Focus, 2016), we choose to accept the alternate
 623 hypothesis that there is correlation between CSS data and
 624 predictions with our network assessment metric when using the
 625 JTM, DM or end-to-end journey time. Because τ_B^{JTM} and τ_B^{DM}
 626 are closer to -1 than τ_B^{EE} , these results suggest that using our
 627 network performance metric with the JTM or DM better
 628 predicts relative changes to the CSS data than using end-to-end
 629 journey time. However, observing tables calculated by Walker
 630 (2016) indicate that even the 80% confidence intervals of
 631 τ_B^{JTM} , τ_B^{DM} and τ_B^{EE} are too large to determine a statistically
 632 significant difference between the values of τ_B^{JTM} , τ_B^{DM} and τ_B^{EE} .
 633 To determine a statistically significant difference by reducing
 634 the confidence interval without altering the significance level,
 635 more years of data for comparison are needed in the series of ϕ .
 636 It is unsurprising that τ_B^{JTM} and τ_B^{DM} do not equal -1.0 because,
 637 in this study, ϕ^{JTM} and ϕ^{DM} do not capture the effect of some
 638 factors, beyond the timetable and passenger load, which may
 639 affect ϕ^{CSS} , e.g. delayed trains. Our network assessment metric
 640 using the JTM or DM can capture the effect of some of these
 641 other factors, but the limitation of data available to this study
 642 means that they are not well captured by the model of network
 643 operation used. Similarly, because of factors such as survey
 644 design and implementation, the CSS data may not fully capture
 645 influencers to passenger experience that distinguish ϕ^{JTM} , ϕ^{DM}
 646 and ϕ^{EE} , e.g. if the surveys were not conducted during times of
 647 high travel demand the effect of crowding will not be well
 648 captured. Consequently, not being able to determine a

649 statistically significant difference in the accuracy of ϕ^{JTM} , ϕ^{DM}
650 and ϕ^{EE} might also be a limitation of the measured CSS data.

651

652 Section 3.1 describes that for low passenger numbers, ϕ^{JTM} is
653 insensitive to crowding (because no crowding penalty is
654 applied), whereas ϕ^{DM} is. However when some passengers are
655 standing (the normal operating regime for many GB services,
656 e.g. 70% of services into London St. Pancras during the
657 morning peak (Peluffo, 2018)), ϕ^{JTM} will be more sensitive to
658 crowding than ϕ^{DM} because it applies a crowding penalty four
659 to eight times greater. This is confirmed by the results of Figure
660 6 which demonstrate that the contribution of the crowding
661 penalty to the network score is on average 9% and 1% for the
662 ϕ^{JTM} and ϕ^{DM} respectively. Section 3.1 also describes that the
663 DM applies a greater VoT weighting than the JTM to
664 passengers who are “In Station”. Because the VoT weightings
665 of the JTM and DM have been derived from surveying
666 passengers, this may reflect local differences in passenger
667 expectations where the metric was developed. For example,
668 when used in our network assessment metric the JTM
669 (developed in London) penalises crowding more and delay on
670 the platform, less, than the DM (developed in Japan). This
671 suggests that when considering a specific network, it is
672 important to ensure the use of VoT weightings relevant to the
673 passengers of that network. However, the similarity of the
674 ϕ^{JTM} and ϕ^{DM} values in the results indicate that the difference
675 in weightings placed on different passenger journey states
676 approximately cancel out (for the study network in the years
677 investigated). The results in Figure 7 show a high degree of
678 correlation ($\rho=1.00$, $P<0.005$) between network scores
679 calculated using the JTM and network scores calculated using
680 the DM, despite their different formulations and countries of
681 origin.

682

683 Considering all the results together suggests that using our
684 newly developed network performance metric with the JTM or
685 DM can be used to predict network performance from the
686 passenger perspective, and successfully aggregates across
687 passenger states to capture effects such as crowding and
688 different journey stages. There is evidence that the network
689 assessment metric, using either the JTM or DM, better predicts
690 changes to customer satisfaction than end-to-end journey time.
691 Because the JTM, CSS data and network operation data are all
692 related to LUL, this result might be considered special to this
693 case where there is a “closed-loop” between metric and
694 validation. However, the DM has no connection to the LUL
695 data but is demonstrated here to achieve similar outcomes. This
696 indicates the result is not special to the “closed-loop” case.

697 6 Conclusions

698 Passenger journeys are multi-stage and the conditions of a
699 journey stage, e.g. crowding when on a train, can vary. We
700 have introduced the term “state” to describe a specific
701 combination of stage and conditions. A passenger journey can
702 be described as a series of states and the literature has shown
703 that the relative time spent in each of these will have different
704 effect on the overall experience of the passenger. Measuring the
705 passenger end-to-end journey time alone, or the train
706 punctuality at final destination (as used in the common UK
707 performance measure, PPM) will not capture this. The JTM and
708 DM are journey assessment metrics that can capture individual
709 journey experience by applying a VoT weighting to time spent
710 in each state. Both metrics sum the weighted time spent in each
711 state, but they use different weightings, journey stages and the
712 DM applies an additional penalty for train changes. Both apply
713 a crowding penalty to capture the additional disutility caused to
714 a passenger when traveling on a train with other passengers.
715 For networks operating in the regime where some passengers
716 cannot find a seat, the crowding penalty applied by the JTM is
717 four to eight times greater than the DM. In this regime, the
718 assessment of network performance using the JTM is more
719 sensitive to crowding than when using the DM. Both the JTM
720 and the DM can be used as part of a network assessment metric
721 we introduce where the network score is taken to be the
722 aggregate of journey scores normalized by the distance
723 travelled and the number of passengers. It is found that, for the
724 Victoria Line of the LUL network from 2008 to 2017, there is a
725 high degree of correlation ($\rho=1.00$, $P<0.005$) between the
726 network scores calculated with the JTM and network scores
727 calculated with the DM, despite their different formulations and
728 countries of origin. Extending the number of different networks
729 in this comparison is an area for future work, to determine if
730 this result is network-specific or general.

731
732 When comparing network scores against measured values of
733 customer satisfaction for the same network (obtained from
734 surveys) there is statistically significant evidence ($P<0.005$) to
735 reject the null hypothesis that predicted and measured changes
736 do not correlate. Considering other evidence from the literature,
737 we therefore accept the hypothesis that predicted and measured
738 changes are correlated which means our network assessment
739 metric can be applied to predict the relative performance of
740 different networks from the passenger perspective. For the data
741 available, our network assessment metric using the JTM or the
742 DM better predicted relative changes to customer satisfaction
743 than end-to-end journey time. However, to determine a
744 statistically significant difference more data for comparison is
745 required. Therefore future work is to investigate networks
746 where more than ten measurements of network performance

747 can be collected and corresponding predictions computed (in
 748 the case of our experiment each measurement corresponds to a
 749 year over which passenger satisfaction data is available
 750 corresponding to the timetable operated that year, but any
 751 timescale in which a system change and its effect can be
 752 measured may be considered in future experiments). This might
 753 be achieved by re-investigating the Victoria Line in the future
 754 as additional years of customer satisfaction data become
 755 available. Further future work is to investigate networks where
 756 a more detailed description of the passenger route is available
 757 so that the effect of train transfer on passenger experience can
 758 be captured. The network assessment metric could then be
 759 validated for journeys which include this activity and might
 760 also allow a statistically significant difference with end-to-end
 761 journey time to be discerned. Updating the network assessment
 762 metric with new VoT weightings to capture other factors which
 763 influence passenger experience (e.g. cleanliness and journey
 764 purpose) is also an area for future work.

765 7 Acknowledgements

766 Funding was received from Network Rail and EPSRC grant
 767 number EP/M508135/1. The authors would like to thank Nigel
 768 Kelt and David Winslett at London Underground for discussion
 769 and access to data, and Dr Chikara Hirai and Dr Taketoshi
 770 Kunimatsu at Railway Technical Research Institute in Tokyo
 771 for discussions and documentation regarding use of network
 772 metrics in Japan.

773 8 References

- 774 ALI, A. A., ELIASSON, J. & WARG, J. 2017. Measuring the
 775 Socio-economic Benefits of Train Timetables
 776 Application to Commuter Train Services in Stockholm.
 777 *Transportation Research Procedia*, 27, 849-856.
- 778 ARUP, INSTITUTE FOR TRANSPORT STUDIES LEEDS &
 779 ACCENT. 2015. *Provision of market reserch for value*
 780 *of travel time savings and reliability: Non-Technical*
 781 *Summary Report* [Online]. London, UK: Department for
 782 Transport. Available:
 783 [https://assets.publishing.service.gov.uk/government/upl](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/470229/vtts-phase-2-report-non-technical-summary-issue-august-2015.pdf)
 784 [oads/system/uploads/attachment_data/file/470229/vtts-](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/470229/vtts-phase-2-report-non-technical-summary-issue-august-2015.pdf)
 785 [phase-2-report-non-technical-summary-issue-august-](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/470229/vtts-phase-2-report-non-technical-summary-issue-august-2015.pdf)
 786 [2015.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/470229/vtts-phase-2-report-non-technical-summary-issue-august-2015.pdf) [Accessed 20/03/2019].
- 787 CACCHIANI, V. & TOTH, P. 2012. Nominal and robust train
 788 timetabling problems. *European Journal of Operational*
 789 *Research*, 219, 727-737.
- 790 CARNE, M. 2018. *Strategic Buisness Plan 2019 - 2024*
 791 [Online]. London, UK: Network Rail,. Available:
 792 [https://cdn.networkrail.co.uk/wp-](https://cdn.networkrail.co.uk/wp-content/uploads/2018/02/Strategic-business-plan-high-level-summary.pdf)
 793 [content/uploads/2018/02/Strategic-business-plan-high-](https://cdn.networkrail.co.uk/wp-content/uploads/2018/02/Strategic-business-plan-high-level-summary.pdf)
 794 [level-summary.pdf](https://cdn.networkrail.co.uk/wp-content/uploads/2018/02/Strategic-business-plan-high-level-summary.pdf) [Accessed 15/11/2018].

- 795 CHAN, J. 2007. *Rail transit OD matrix estimation and journey*
796 *time reliability metrics using automated fare data.*
797 Master of Science in Transportation, Massachusetts
798 Institute of Technology.
- 799 CHANG, Y. H., YEH, C. H. & SHEN, C. C. 2000. A
800 multiobjective model for passenger train services
801 planning: application to Taiwan's high-speed rail line.
802 *Transportation Research Part B-Methodological*, 34,
803 91-106.
- 804 CHEN, C. & CHEN, F. 2010. Experience quality, perceived
805 value, satisfaction and behavioral intentions for heritage
806 tourists. *Tourism Management*, 31, 29-35.
- 807 DAVIS, R. 2018. *Rail Factsheet* [Online]. London, UK:
808 Department for Transport. Available:
809 [https://assets.publishing.service.gov.uk/government/upl](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/761352/rail-factsheet-2018.pdf)
810 [oads/system/uploads/attachment_data/file/761352/rail-](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/761352/rail-factsheet-2018.pdf)
811 [factsheet-2018.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/761352/rail-factsheet-2018.pdf) [Accessed 13/12/2018].
- 812 HICKEY, S. 2011. *Improving the estimation of platform wait*
813 *time at the London Underground.* Master of Science in
814 Transportation, Massachusetts Institute of Technology.
- 815 HOROWITZ, A. J. 1978. The subjective value of the time
816 spent in travel. *Transportation Research*, 12, 385-393.
- 817 KANAI, S., SHIINA, K., HARADA, S. & TOMII, N. 2011. An
818 optimal delay management algorithm from passengers'
819 viewpoints considering the whole railway network.
820 *Journal of Rail Transport Planning & Management*, 1,
821 25-37.
- 822 KELT, N. Personal communication. "RE: LU Crowding and
823 Left Behind". 2015, 3rd Decemeber.
- 824 KUNIMATSU, T., HIRAI, C. & TOMII, N. 2012. Train
825 timetable evaluation from the viewpoint of passengers
826 by microsimulation of train operation and passenger
827 flow. *Electrical Engineering in Japan*, 181, 51-62.
- 828 KUNIMATSU, T., HIRAI, C., TOMII, N. & TAKABA, M.
829 2009. Evaluation of timetables by estimating
830 passengers' personal disutility using micro-simulation.
831 *In: HANSEN, I. (ed.) RailZuirch2009 - Third*
832 *International Seminar on Railway Operations*
833 *Modelling and Analysis.* Zürich, Switzerland: ETH
834 Zürich.
- 835 LONDON UNDERGROUND LIMITED 2007. Victoria Line
836 Working Timetable No. 31. London.
- 837 LONDON UNDERGROUND LIMITED 2009. Victoria Line
838 Working Timetable No. 32. London.
- 839 LONDON UNDERGROUND LIMITED 2011. Victoria Line
840 Working Timetable No. 33. London.
- 841 LONDON UNDERGROUND LIMITED 2012a. Victoria Line
842 Working Timetable No. 34. London.
- 843 LONDON UNDERGROUND LIMITED 2012b. Victoria Line
844 Working Timetable No. 35. London.

- 845 LONDON UNDERGROUND LIMITED 2014. Victoria Line
846 Working Timetable No. 36. London.
- 847 LONDON UNDERGROUND LIMITED 2015a. Victoria Line
848 Working Timetable No. 37. London.
- 849 LONDON UNDERGROUND LIMITED 2015b. Victoria Line
850 Working Timetable No. 38. London.
- 851 LONDON UNDERGROUND LIMITED. 2016a. *FOI 0121*
852 *1617.xlsx* [Online]. London, UK: What Do They Know.
853 Available:
854 [https://www.whatdotheyknow.com/request/i_need_to_u](https://www.whatdotheyknow.com/request/i_need_to_understand_the_definit)
855 [nderstand_the_definit](https://www.whatdotheyknow.com/request/i_need_to_understand_the_definit) [Accessed 20/12/2018].
- 856 LONDON UNDERGROUND LIMITED 2016b. Victoria Line
857 Working Timetable No. 39. London.
- 858 LONDON UNDERGROUND LIMITED 2016c. Victoria Line
859 Working Timetable No. 40. London.
- 860 LONDON UNDERGROUND LIMITED 2017. Victoria Line
861 Working Timetable No. 41. London.
- 862 LONDON UNDERGROUND LIMITED. 2018a. *Performance*
863 *Data Almanac* [Online]. London, UK: Transport for
864 London. Available:
865 [https://tfl.gov.uk/corporate/publications-and-](https://tfl.gov.uk/corporate/publications-and-reports/underground-services-performance)
866 [reports/underground-services-performance](https://tfl.gov.uk/corporate/publications-and-reports/underground-services-performance) [Accessed
867 29/11/2018].
- 868 LONDON UNDERGROUND LIMITED. 2018b. *Rolling*
869 *Origin Destination Survey* [Online]. London, UK:
870 Transport for London. Available: [https://tfl.gov.uk/info-](https://tfl.gov.uk/info-for/open-data-users/our-open-data)
871 [for/open-data-users/our-open-data](https://tfl.gov.uk/info-for/open-data-users/our-open-data) [Accessed
872 20/12/2018].
- 873 NIPPON. 2018. *Japan's Most Congested Train Lines* [Online].
874 Tokyo, Japan: The Nippon Communications
875 Foundation. Available:
876 <https://www.nippon.com/en/features/h00255> [Accessed
877 23/10/2018].
- 878 PELUFFO, N. 2018. *Rail passenger numbers and crowding on*
879 *weekdays in major cities in England and Wales: 2017*
880 [Online]. London, UK: Department for Transport.
881 Available:
882 [https://assets.publishing.service.gov.uk/government/upl](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/728526/rail-passengers-crowding-2017.pdf)
883 [oads/system/uploads/attachment_data/file/728526/rail-](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/728526/rail-passengers-crowding-2017.pdf)
884 [passengers-crowding-2017.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/728526/rail-passengers-crowding-2017.pdf) [Accessed 29/12/2018].
- 885 QIN, F. 2014. Investigating the In-Vehicle Crowding Cost
886 Functions for Public Transit Modes. *Mathematical*
887 *Problems in Engineering*, 2014, 1-13.
- 888 SELS, P., DEWILDE, T., CATTRYSSSE, D. &
889 VANSTEENWEGEN, P. 2016. Reducing the passenger
890 travel time in practice by the automated construction of
891 a robust railway timetable. *Transportation Research*
892 *Part B-Methodological*, 84, 124-156.
- 893 SUSILO, Y. O. & CATS, O. 2014. Exploring key determinants
894 of travel satisfaction for multi-modal trips by different

- 895 traveler groups. *Transportation Research Part a-Policy*
896 *and Practice*, 67, 366-380.
- 897 TECHNICAL LEADERSHIP GROUP. 2017. *Rail Technical*
898 *Strategy Capability Delivery Plan* [Online]. London,
899 UK: Rail Safety and Standards Board. Available:
900 [https://www.rssb.co.uk/rts/Documents/2017-01-27-rail-](https://www.rssb.co.uk/rts/Documents/2017-01-27-rail-technical-strategy-capability-delivery-plan-brochure.pdf)
901 [technical-strategy-capability-delivery-plan-brochure.pdf](https://www.rssb.co.uk/rts/Documents/2017-01-27-rail-technical-strategy-capability-delivery-plan-brochure.pdf)
902 [Accessed 7/11/2018].
- 903 TNS POLITICAL & SOCIAL. 2013. *Europeans' satisfaction*
904 *with rail services* [Online]. The European Commission.,
905 Available:
906 [http://ec.europa.eu/commfrontoffice/publicopinion/flash](http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_382a_en.pdf)
907 [/fl_382a_en.pdf](http://ec.europa.eu/commfrontoffice/publicopinion/flash/fl_382a_en.pdf) [Accessed 08/02/2019].
- 908 TRANSPORT FOCUS 2016. National Rail Passenger Survey -
909 Spring 2016 Main Report. London, UK.
- 910 TRANSPORT FOR LONDON. 2012. *Join in the celebrations*
911 *across the capital this summer with a limited edition*
912 *Summer Oyster card* [Online]. London, UK: Transport
913 for London. Available: [https://tfl.gov.uk/info-](https://tfl.gov.uk/info-for/media/press-releases/2012/june/join-in-the-celebrations-across-the-capital-this-summer-with-a-limited-edition-summer-oyster-card)
914 [for/media/press-releases/2012/june/join-in-the-](https://tfl.gov.uk/info-for/media/press-releases/2012/june/join-in-the-celebrations-across-the-capital-this-summer-with-a-limited-edition-summer-oyster-card)
915 [celebrations-across-the-capital-this-summer-with-a-](https://tfl.gov.uk/info-for/media/press-releases/2012/june/join-in-the-celebrations-across-the-capital-this-summer-with-a-limited-edition-summer-oyster-card)
916 [limited-edition-summer-oyster-card](https://tfl.gov.uk/info-for/media/press-releases/2012/june/join-in-the-celebrations-across-the-capital-this-summer-with-a-limited-edition-summer-oyster-card) [Accessed
917 4/2/2019].
- 918 VANSTEENWEGEN, P. & VAN OUDHEUSDEN, D. 2007.
919 Decreasing the passenger waiting time for an intercity
920 rail network. *Transportation Research Part B-*
921 *Methodological*, 41, 478-492.
- 922 VUCHIC, V. R. & NEWELL, G. F. 1968. Rapid Transit
923 Interstation Spacings for Minimum Travel Time.
924 *Transportation Science*, 2, 303-339.
- 925 WALKER, D. A. 2016. JMASM38: Confidence Intervals for
926 Kendall's Tau with Small Samples (SPSS). *Journal of*
927 *Modern Applied Statistical Methods*, 15, 868-883.
- 928 WANG, Y., TANG, T., NING, B., VAN DEN BOOM, T. J. J.
929 & DE SCHUTTER, B. 2015. Passenger-demands-
930 oriented train scheduling for an urban rail transit
931 network. *Transportation Research Part C: Emerging*
932 *Technologies*, 60, 1-23.
- 933 WARDMAN, M. 2004. Public transport values of time.
934 *Transport Policy*, 11, 363-377.
- 935 WARDMAN, M. & MURPHY, P. 2015. Passengers' valuations
936 of train seating layout, position and occupancy.
937 *Transportation Research Part a-Policy and Practice*,
938 74, 222-238.
- 939 WASSERSTEIN, R. L. & LAZAR, N. A. 2016. The ASA's
940 Statement on p-Values: Context, Process, and Purpose.
941 *The American Statistician*, 70, 129-133.
942