A COMPARISON OF THE SUITABILITY OF ACCUMULATION MODELS IN TRAVEL BEHAVIOUR MODELLING

Thomas Hancock (Corresponding Author)
Institute for Transport Studies
University of Leeds
tstoh@leeds.ac.uk

Stephane Hess
Institute for Transport Studies
University of Leeds
S.Hess@its.leeds.ac.uk

Charisma Choudhury
Institute for Transport Studies
University of Leeds
C.F.Choudhury@leeds.ac.uk

Word Count: 5713 words + 0 figures + 4 tables = 6713

ABSTRACT

The importance of increasing the behavioural realism of activity and travel choice models is increasingly being realised by the modellers. Despite the fact that choice modellers have attempted to add behavioural insights into their models, few attempts as of yet have looked instead at directly applying models developed in mathematical psychology for choice modelling. This is perhaps surprising given that a key aim of both areas is to understand behaviour. The present paper looks at two such models from mathematical psychology in which the preference for alternatives dynamically accumulates over time. Whilst decision field theory (DFT) has been shown to capture various context effects due to its stochastic nature, the multi-attribute linear ballistic accumulator (MLBA) has been designed to be more mathematically tractable. This paper considers how best to operationalise these models for typical route choice datasets, finding that both can incorporate underlying preferences for alternatives, meaning that both can be used for labelled route choice tasks. Whilst DFT outperforms MLBA in model fit for both our stated preference route choice datasets, MLBA performs comparably to typical random utility models in both datasets and has good model fit for three simulated datasets, suggesting that both models provide a good account of behaviour and work well with typical choice data, especially in capturing non-linear sensitivities, which are often found in transport choice modelling. Though the models are yet to be tested with revealed preference data, they hold significant promise in leading to more robust choice models applicable to transport as well as wider choice contexts.

INTRODUCTION

Whilst choice modelling has been grounded in firm economic foundations (*1*), attempts at understanding decision-making behaviour in other fields has been implemented with very different aims and objectives. Since the work of Kahneman and Tversky in the 1970s (*2–4*), the field of behavioural economics has considered choice from an economic viewpoint whilst simultaneously demonstrating that decision-makers are subject to biases and heuristics that result in non-optimal choices. Meanwhile, mathematical psychologists have tended to try and build models that can explain these effects (*5–7*) as well as decision-making under time pressure (*8*). Whilst there have been some recent efforts to incorporate behavioural ideas into choice models such as the minimisation of regret (*9, 10*) and the incorporation of heuristics (*11*), few papers as of yet have tested whether models developed in mathematical psychology can be used for decision-making in general. Recently, one such popular model from mathematical psychology, decision field theory, has been tested against models traditionally used in choice modelling (*12, 13*). This current paper further aims to test the suitability of using decision field theory in route choice modelling as well as comparing it against the multi-attribute linear ballistic accumulator (MLBA), a carefully designed alternative model from mathematical psychology (*14*).

Under a decision field theory (DFT) model, the preference values of the alternatives update stochastically over time. The idea of this is that the underlying cognitive processes when making a decision are captured. At each moment an attibute is compared across alternatives and a valence is added to the preference value for each alternative. At some moment, the decision-maker comes to a conclusion. Either, one of the alternatives reaches some threshold (similar to satisficing (*15, 16*)) or an external cue forces the decision-maker to make a choice, in which case the decision-maker chooses the alternative with the highest preference value at that moment. MLBA models have a similar accumulation process for the preference of alternatives, but does not update stochastically. Instead, decision-makers start with some random amount of initial ’evidence’ for each alternative, that then ’drifts’ linearly until one of the alternatives reaches a threshold.

This means that the mathematics underlying the two models is vastly different. Whilst we require expected values and covariances to calculate the probability of alternatives for a decision field theory model, the multiattribute linear ballistic accumulator has been specifically designed such that it is ’simple’ (*17*) and is mathematically tractable such that the probability of alternatives can easily be calculated. Whilst recent work has tested decision field theory against traditional choice models for route choice (*13*), other models from mathematical psychology are yet to have been compared.

There have been few applications of DFT (that are mostly theoretical) in the transport literature and none for MLBA as far as the authors are aware. For example, DFT conceptually should be an appropriate model for dealing with a variety of travel situation effects including situational dynamics, type of travel, cultural habits and societal norms (*18*). Additionally, DFT has been combined with the Queuing Network-Model Human Processer to model a driver’s speed control (*19*). It has also been demonstrated that DFT accurately predicts the share of participants who choose park and ride, car, bus or subway (*20*). The large and rich datasets typically found in transport choice datasets have meant that computational limitations have until now limited the use of DFT in transport applications (*21*).

Research in mainstream mathematical psychology has demonstrated that MLBA outperforms DFT (*14*) and the linear ballistic accumulator - a simpler form of the model, where each alternative has a drift rate which is estimated without incorporating the values of the alternatives’ attributes

(*22*) , performs well for best-worst choice datasets. MLBA thus holds significant promise in making a successful transition into choice modelling.

However, there has not been any previous research that investigates the applicability of these models to mainstream choice modelling data. This motivates our research where we operationalise these models for typical route choice datasets and investigate the strengths and weaknesses over traditional models.

The remainder of this paper is organised as follows. First, we present the frameworks behind both decision field theory and the multi-attribute linear ballistic accumulator. Next, we present our empirical applications comparing these models against multinomial logit and random regret minimisation, considering the impact of scaling methods as well as looking at how to add complexities to the models. The final section summarises the findings and presents the directions of future research.

MODEL FRAMEWORKS

In this section we describe the equations underlying decision field theory and the multi-attribute linear ballistic accumulator and demonstrate how these can be used to calculate the probability of alternatives in such models.

Decision field theory

Decision field theory, as an accumulator model, has preference values for each alternative that update over time.

$$P_t = S \cdot P_{t-1} + V_t \tag{1}$$

The previous values, P_{t-1}, are multiplied by a feedback matrix, S, and a valence vector V_t is added. The feedback matrix has two parameters that control for the level of occurence of the attraction, similarity and compromise effects (5, 7, 23) and is defined:

$$S = I - \phi_2 \times \exp(-\phi_1 \times D^2) \tag{2}$$

where ϕ_1 is a memory parameter and ϕ_2 is a sensitivity parameter. D is the distance between the attributes of the alternatives. Whilst the relative importance of the different attributes can be taken into account with a psychological distance function (*23*) and new distance functions have been described (*12*), the Euclidean distance between the attributes can also be used for simplicity (*20*). The memory parameter, if less than one, results in preference values deteriorating over time (if no valence vectors were added) and results in the preference values stabilising for a large number of timesteps (*5, 12*). The sensitivity parameter effects how much the alternatives compete with each other. At each timestep, a decision field theory model assumes that the decision-maker compares a single attribute across all of the alternatives. This results in a random valence vector at time t, V_t, which can be calculated:

$$V_t = C \cdot M \cdot W_t + \varepsilon_t \tag{3}$$

where C is a contrast matrix used to rescale the values such that they total zero, M is the matrix of attribute values and $W_t = [0..1..0]'$ with entry $j = 1$ if and only if attribute j is the attribute being attended to by the decision-maker at timestep t. A DFT model thus estimates a weight, w_j, for the likelihood of attending to attribute j. The model assumes that at all timesteps the decision-maker is attending to one of the attributes. This implies that $\sum_j w_j = 1$, therefore a standard uniform distribution $X \sim U(0,1)$ can be used to select which attribute a decision-maker attends to at each time step. There is also a random error vector, $\varepsilon_t = [\varepsilon..\varepsilon]'$, added on to allow for flexibility in the variation of probability values that DFT predicts. This is in essence an error or noise parameter (*5*), for which higher values would be expected for more complex decision-making tasks (*23*).

Estimation of decision field theory

It has been demonstrated that the probabilities of alternatives at time t can be calculated with the expected value and the covariance of the preference values (ξ_t and Ω_t) (*5*).

To calculate the expected value of the preference values, we must first expand equation 1, which results in:

$$P_t = \sum_{k=0}^{t-1} S^k \cdot V_{t-k} + S^t \cdot P_0 \tag{4}$$

where P_0 is the initial preference vector. This is often assumed to be a vector of zeros (24) but can also be used to capture underlying preferences for different alternatives (13).

The attribute weights w_j are stationary, therefore W_t can be considered a stationary stochastic process. This means that V_t is also a stationary stochastic process with mean $E[V_t]$ and a variance covariance matrix given by $Cov[V_t]$. We let ε_t vary according to a normal distribution with mean zero and variance ε. Thus if we have $\mu = E[V_t]$, it can be calculated as $\mu = C \cdot M \cdot w_m$, where w_m is a vector containing the probabilities of each of the attributes being considered. We also have $Cov[V_t] = \Phi = C \cdot M \cdot \Psi \cdot M' \cdot C' + \varepsilon$, where $\Psi = Cov[W_t]$ (C and M are matrices of constants). We can then calculate the expected value and the expected covariance of P_t. With S being a constant, $E[P_t]$ reduces to:

$$E[P_t] = \xi_t = \sum_{k=0}^{t-1} S^k \cdot \mu + S^t \cdot P_0 \tag{5a}$$

$$= (I-S)^{-1}(I-S^t) \cdot \mu + S^t \cdot P_0 \tag{5b}$$

We can also now calculate the covariance of the preference values:

$$Cov[P_t] = \Omega_t = Cov\left[\sum_{k=0}^{t-1} S^k \cdot V_{t-k} + S^t \cdot P_0\right] \tag{6a}$$

$$= \sum_{k=0}^{t-1} \left[S^k \cdot \Phi \cdot S^{k'}\right] \tag{6b}$$

We can further simplify $Cov[P_t]$ such that we can avoid requiring a summation (13). We replace the feedback matrices with a matrix Z of size $n^2 \times n^2$ and reshape Φ (with entries p_{ij}) into a column matrix:

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n^2} \\ z_{21} & z_{22} & \cdots & z_{2n^2} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n^21} & z_{n^22} & \cdots & z_{n^2n^2} \end{bmatrix} \overline{\Phi} = \begin{bmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \\ p_{21} \\ \vdots \\ p_{nn} \end{bmatrix} \tag{7}$$

By setting $z_{(j-1)n+i,(k-1)n+l} = S_{il}\,S'_{kj}$, $Cov[P_t]$ reduces:

$$Cov[P_t] = \Omega_t = \sum_{k=0}^{t-1}\left[S^k \cdot \Phi \cdot S^{k'}\right] \tag{8a}$$

$$= \sum_{k=0}^{t-1}\left[Z^k \cdot \overline{\Phi}\right] \tag{8b}$$

$$= (I-Z)^{-1}\,(I-Z^t)\,\overline{\Phi} \tag{8c}$$

These succinct forms for ξ_t and Ω_t mean that we can now calculate the probabilities of the alternatives. On the basis of the multivariate central limit theorem, P_t converges to the multivariate normal distribution (5). Under a decision field theory model, the alternative chosen is the one with the greatest preference value at the conclusion of the deliberation process. Thus the probability of choosing alternative A from a set of n alternatives at time t is:

$$Prob\left[\max_{i\in n} P_t\,[i] = P_t\,[A]\right] = \int_{X>0} exp\left[-(X-\Gamma)'\Lambda^{-1}(X-\Gamma)/2\right]/(2\pi|\Lambda|^{0.5})dX \tag{9}$$

with $X = [P_t\,[A] - P_t\,[B]\,,\ldots,P_t\,[A] - P_t\,[n]]'$, $\Gamma = L\xi_t$, $\Lambda = L\Omega_t L'$ where

$$L = \begin{bmatrix} 1 & -1 & 0 & \ldots & \ldots & 0 \\ 1 & 0 & -1 & \ddots & & \vdots \\ 1 & \vdots & \ddots & \ddots & \ddots & \vdots \\ 1 & \vdots & & \ddots & -1 & 0 \\ 1 & 0 & \ldots & \ldots & 0 & -1 \end{bmatrix} \tag{10}$$

L is a matrix comprised of a column vector of 1s and a negative identity matrix of size $n-1$ where n is the number of attributes. The column vector of 1s is placed in the i^{th} column where i is the chosen option.

Thus to estimate the probability of alternatives under a decision field theory model, we require estimates for $n-1$ weight parameters (where n is the number of attributes) and estimates for four psychological parameters (ϕ_1 and ϕ_2, the sensitivity and memory parameters respectively, the number of timesteps, t, and the variance of the error term, ε).

The multi-attribute linear ballistic accumulator

Under the multi-attribute linear ballistic accumulator, the key component is the drift rates of the alternatives, d_j, which are drawn from normal distributions:

$$d_j \sim N\left(\frac{10}{1+exp(-\gamma \cdot v_j)}, s\right) \tag{11}$$

where v is a valence vector with a value v_j for each alternative. γ is a logistic parameter and s is a parameter for the standard deviation of the drift rates. This is typically set to the same value for all alternatives, but a different value could be estimated for each drift rate (6). Small values of the logistic parameter γ would result in $exp(-\gamma \cdot v_j) \to 1$, meaning that the valences are less influential and the probabilities of the alternatives become less deterministic. The chosen alternative in a multi-attribute linear ballistic accumulator model is the first alternative to pass a threshold value, χ. Start points for each of the alternatives are drawn from a uniform distribution $U[0,A]$ where A is estimated.

The valences are similar to a decision field theory model's valences with the exception that they attempt to additionally capture the comparison process achieved by DFT's feedback matrix. Thus we have

$$V = C \cdot M \cdot W \tag{12}$$

where W is a vector comprising of a set of attribute weights that sum to 1, M is the attribute matrix and C is a $n \times n$ comparison matrix (n being the number of alternatives) with diagonal entries of 1 and off-diagonal elements:

$$C_{i,j\neq i} = \frac{exp(-\phi \cdot D_{i,j}) - 1}{n-1}. \tag{13}$$

ϕ is a sensitivity parameter such that high values mean that the distance between alternatives, D, become insignificant. Low values allow for more similar alternatives to compete more with each other relative to less similar alternatives. Whilst we can use the Euclidean distance for the distance between the alternatives, we could also use a psychological distance function (*14, 23*).

Estimation of the multi-attribute linear ballistic accumulator

Given the drift rates of the alternatives and known parameter values for the start and end points A and χ respectively, we can calculate the probability of each alternatives' accumulator being the first to finish (*17*). The amount of evidence that needs to be accumulated for an alternative to reach the threshold is U[χ-A,χ] (assuming $\chi > A$). Given an alternative's mean drift rate, D_j, the cumulative distribution function for the time taken for the accumulator associated with alternative j is given by

$$F_j(t) = Prob\left(\frac{U[\chi - A, \chi]}{D_j} < t\right) \tag{14}$$

Brown and Heathcote (2008) demonstrate that this reduces to:

$$\begin{aligned} F_j(t) = 1 + \frac{\chi - b - t \cdot D_j}{A} \cdot \Phi\left(\frac{\chi - A - t \cdot D_j}{t \cdot s}\right) - \frac{b - t \cdot D_j}{A} \cdot \Phi\left(\frac{b - t \cdot D_j}{t \cdot s}\right) \\ + \frac{t \cdot D_j}{A} \cdot \phi\left(\frac{\chi - A - t \cdot D_j}{t \cdot s}\right) - \frac{t \cdot D_j}{A} \cdot \phi\left(\frac{b - t \cdot D_j}{t \cdot s}\right) \end{aligned} \tag{15}$$

where ϕ and Φ are the standardised normal distibution's density and cumulative density functions respectively. The associated probability density function is then:

$$\begin{aligned} f_j(t) = \frac{1}{A}\Bigg[-D_j \cdot \Phi\left(\frac{\chi - A - t \cdot D_j}{t \cdot s}\right) + D_j \cdot \Phi\left(\frac{b - t \cdot D_j}{t \cdot s}\right) \\ + s \cdot \phi\left(\frac{\chi - A - t \cdot D_j}{t \cdot s}\right) - s \cdot \phi\left(\frac{b - t \cdot D_j}{t \cdot s}\right)\Bigg] \end{aligned} \tag{16}$$

To then calculate the probability of alternatives we need to calculate the probability density function of alternative j being quicker than all other alternatives i [1].

[1] For full derivations of equations 15, 16 and 17, refer to appendix A of Brown and Heathcote (2008).

$$PDF_j(t) = f_j(t) \prod_{i \neq j} (1 - F_i(t)) \tag{17}$$

Thus we have

$$Prob(Alternative\ j) = \int_0^\infty PDF_j(t)dt \tag{18}$$

EMPIRICAL APPLICATION

In this section, we first look at how best to apply the multi-attribute linear ballistic accumulator using a Danish route choice dataset where participants face route choice tasks with two alternatives each with 2 attributes. We then compare MLBA's performance against both decision field theory and multinomial logit and random regret minimisation on a UK route choice dataset, looking at the effects of incorporating additional parameters to capture underlying preferences for alternatives. We conclude by looking at some simulated datasets and the impacts of considering choices generated by different models.

Danish Data

Our first dataset is a subset from the Danish value of time dataset (*25*). There is a total of 545 participants and 4,214 choices, with each choice decision comprising of two alternative routes described by travel cost and travel time.

Scaling of attributes

As both MLBA and DFT are scale-variant (*14, 24*), we first compare the effects of using several different scaling methods for the attributes. The first scaling method follows unity-based normalisation (*12*), where we rescale the attributes to values between 0 and -1. For the second method, no scaling method is used. The third method uses standard score normalisation. The fourth and fifth scaling methods use minimum and maximum rescaling respectively (by choice set), as previously shown to be effective for MLBA (*14*).

TABLE 1 : The log-likelihood values obtained from models for the Danish dataset

Danish Data	par.	scale 1	scale 2	scale 3	scale 4	scale 5
MLBA	5	-2,287.68	-2,192.75	-2,192.51	-2,254.32	-2,280.73
MLBA + psych.	6	-2,273.08	-2,192.75	-2,192.51	-2,246.96	-2,280.73
MLBA + χ	6	-2,231.50	-2,191.02	-2,192.51	-2,242.58	-2,280.63
DFT	5	-2,039.16	-2,032.74	-2,015.57	-2,110.59	-2,121.18
RUM	3	-2,301.25	-	-	-	-
RUM+log terms	5	-2,211.69	-	-	-	-

The results of using the different scaling methods are given in table 1. We also compare MLBA to DFT and multinomial logit (RUM). For MLBA, we fix the upper threshold, χ to 20 and use Euclidean distance for the distance between attributes, D. The random utility multinomial logit model estimates parameters for cost, time, log cost, log time and an alternative specific constant for the first option. This results in us being able to compare the models directly as they each use five parameters. It appears that no matter which scaling method is used, the DFT model achieves a higher log-likelihood value than both the RUM and MLBA models. Table 1 also suggests that neither adding a psychological distance parameter (*14*) nor setting χ as a free parameter result in MLBA making a significant enough improvement to have a better fit than DFT. MLBA only has a higher log-likelihood value than RUM if scales 2 or 3 are used. As there is no significant impact of including extra parameters for MLBA for either of these scaling methods, we use Euclidean distance and set χ to 20 for the remaining MLBA models in this paper.

UK data

The second stated preference dataset we consider has a total of 368 participants, each completing 10 choice tasks resulting in 3,680 choices. The participants are all public transport commuters living in the UK. Each task involves an invariant reference trip and two hypothetical alternatives. Each alternative is described by travel time, cost, rate of crowded trips, rate of delays (both out of 10 trips), the average length of delays (across delayed trips) and cost of a provision of a delay information service (*26*). The final of these attributes was found to be insignificant in all models, therefore was omitted. We use this dataset to look at the impact of including alternative specific constants in the different models. We add on alternative specific constants to the utility and regret functions (*10*), respectively, for the random utility model (RUM) and random regret minimisation (RRM) models. For DFT, we add parameters to estimate the starting value of the initial preference matrix, P_0. Finally, the multi-attribute linear ballistic accumulator model incorporates alternative specific constants by adjusting the drift rates such that:

$$d_j \sim N\left(\frac{10}{1+exp(-\gamma \cdot v_j)}, s\right) + ASC_j \tag{19}$$

The results of these models are shown in table 2.

TABLE 2 : The log-likelihood values obtained from models for the UK dataset

UK data	Fare			relative	Log Fare		
	LL	par.	BIC	runtime	LL	par.	BIC
MLBA	-3549.41	9	7172.718	213.68	-3435.82	9	6945.532
MLBA+p	-3530.82	11	7151.949	292.47	-3424.40	11	6939.123
DFT	-3598.87	8	7263.425	174.78	-3424.49	8	6914.671
DFT+p	-3577.97	10	7238.039	181.89	-3380.58	10	6843.267
RUM	-3732.75	5	7506.555	1.00	-3440.43	5	6921.917
RUM+p	-3721.67	7	7500.815	1.07	-3401.58	7	6860.635
RRM	-3707.79	5	7456.623	1.58	-3423.34	5	6887.733
RRM+p	-3699.49	7	7456.455	1.82	-3402.59	7	6862.655

Whilst it initially appears that the best fitting model is MLBA, considering the log of the fares rather than the fare prices for the alternatives results in much less difference in log-likelihoods between the models. This suggests that the logistic parameter, γ, significantly improves the MLBA model before log fare prices are considered, but that MLBA loses this advantage after the other models can also take into account this non-linear sensitivity to costs. This also appears to be the case for RRM, which loses its advantage over RUM once log fares are considered. We also see a fairly consistent improvement in all of our different models by including parameters for underlying preferences.

Simulated data: Setup

For our simulated datasets, we use the choice datasets available from the choice modelling centre website (*27*) and then simulate choices three times using a random utility model, a DFT model and an MLBA model. For each dataset, we use the model to calculate the probability of an alternative being chosen and then use the same set of randomly drawn numbers to assign a choice based on the calculated probabilities for each choice task.

For our random utility model, we define the utility a respondent n obtains from alternative j from choice task t as:

$$U_{jnt} = ASC_{M_j} + ASC_{F_j} + \beta_{TT_j} * TT + \beta_{TC} * TC + \beta_{AC} * AC + \varepsilon_{jnt} \tag{20}$$

where ASC_{M_j} and ASC_{F_j} are alternative specific constants for male and female participants, respectively. TT is the travel time, TC is the travel cost and AC is the access time. There are also coefficients for travel cost and access time and also alternative specific travel time coefficients, β_{TT_j}. For the decision field theory simulated dataset, we incorporate underlying preferences by setting $P_{0_j} = ASC_{M_j} + ASC_{F_j}$. After using standard score normalisation to rescale the attributes ($a_j = -\frac{a_j - mean(a)}{sd(a)}$), the alternative specific travel time coefficients are added to the travel times: $TT_j = TT_j + \beta_{TT_j}$ for both DFT and MLBA. We incorporate the underlying preference parameters for MLBA in the drift rate, as in equation 19. All of the values used for the parameters to generate probabilities for each alternative are given in table 4.

Simulated datasets: Results

The specification of our simulated datasets results in us being able to test various different complexities of each of our models. The most basic MNL model uses only three parameters: β_{AC}, β_{TC} and a single coefficient for travel times across all alternatives, β_{TT}. The second model incorporates three additional parameters for the alternative specific travel time coefficients. The third also has 6 parameters and instead adds on three alternative specific constants. The fourth model includes six alternative specific constants, such that male and female participants' preferences can be captured. The final 'full' model allows all twelve parameters to be estimated. Similar restrictions are placed on MLBA and DFT such that we have five levels of complexity for each of these models too.

The log-likelihoods obtained from these models are displayed in table 3. For the dataset with choices generated by RUM, it appears that the basic DFT and MLBA models fit significantly better than the basic RUM model. Whilst the difference between the different models decreases quickly as the model complexity increases, the best fit achieved across the dataset is found by a full specification of the DFT model. Comparisons across 12 parameters find that the MLBA model achieves the best fit, but it does not improve by incorporating the final cost parameters. Similar results for the basic models are found for the datasets with choices generated by MLBA and DFT. It also appears that, as expected, MLBA fits the MLBA generated choices with a higher log-likelihood and DFT fits the DFT generated choices best. It appears that the random utility model does not come close to competing with MLBA and DFT until full model specifications are used, demonstrating that further complexities added to the models and generation of the choices may result in RUM further closing the gap to DFT and MLBA.

TABLE 3 : The log-likelihood values obtained from models for the simulated datasets

Simulated Datasets				
RUM		dataset		
	par.	RUM	DFT	MLBA
basic	3	-6538.56	-6366.14	-6253.36
+ costs	6	-6216.54	-6125.14	-5975.15
+ASC	6	-6210.02	-6088.14	-5905.16
+gender	9	-6139.94	-5916.24	-5828.92
full	12	-6076.57	-5768.44	-5665.93
DFT		dataset		
	par.	RUM	DFT	MLBA
basic	6	-6222.67	-5941.57	-5771.53
+ costs	9	-6113.13	-5902.77	-5757.53
+ASC	9	-6114.06	-5904.76	-5740.43
+gender	12	-6051.34	-5720.31	-5672.07
full	15	-6035.67	-5716.03	-5652.74
MLBA		dataset		
	par.	RUM	DFT	MLBA
basic	6	-6233.93	-5986.93	-5769.17
+ costs	9	-6116.45	-5937.23	-5732.36
+ASC	9	-6113.89	-5936.34	-5731.16
+gender	12	-6042.95	-5758.16	-5652.51
full	15	-6042.92	-5745.77	-5648.98

Recovery of parameters
Table 4 gives parameter estimates for each of the three full models for their own dataset, thus demonstrating the stability of the parameters. We see, as expected, that the parameters used in the basic versions of each model are more recoverable than the preference parameters. Most crucially for DFT, the weight parameters are almost perfectly recovered whilst the sensitivity and memory parameters, ϕ_1 and ϕ_2 are not, further demonstrating that a DFT model without these parameters may be worth investigating (*13*).

TABLE 4 : Parameter values used to generate datasets and estimates for full models for their respective datasets

	MNL		DFT		MLBA	
	generation	estimated	generation	estimated	generation	estimated
$ASC_{M_{car}}$	0.0000		0.0000		0.0000	
$ASC_{F_{car}}$	0.0000		0.0000		0.0000	
$ASC_{M_{rail}}$	-0.4000	1.1354	-2.0000	-4.2785	-1.0000	-3.1019
$ASC_{F_{rail}}$	0.2000	0.8344	1.0000	5.4253	0.5000	2.2545
$ASC_{M_{air}}$	-0.6000	1.1774	-3.0000	-3.0371	-1.5000	-2.7179
$ASC_{F_{air}}$	-0.2000	0.4847	-1.0000	3.3436	-0.5000	1.4751
$ASC_{M_{hsr}}$	-0.2000	1.2178	-1.0000	0.2815	-0.5000	-2.2905
$ASC_{F_{hsr}}$	-0.4000	-0.1860	-2.0000	-1.5017	-1.0000	-0.6221
$\beta_{TT_{car}}$	-0.0040	0.0001	0.1000	0.5522	0.1000	-0.2858
$\beta_{TT_{rail}}$	-0.0055	-0.0082	-0.0500	0.5204	-0.1000	0.0140
$\beta_{TT_{air}}$	-0.0050	-0.0054	-0.1000	-0.1888	-0.0500	-0.5323
$\beta_{TT_{hsr}}$	-0.0050	-0.0035				
$\beta_{TC}/\ w_{TC}$	-0.0200	-0.0185	0.5000	0.4997	0.5000	0.4861
β_{AC}	-0.0040	-0.0039				
w_{TT}			0.4500	0.4510	0.4500	0.4655
ϕ_1			0.0500	6.2422		
ϕ_2			0.0100	0.2454		
ε			2.0000	1.5848		
t			5.0000	3.9121		
ϕ					0.0500	0.0344
y					1.0000	1.5639
A					10.0000	14.4977
s					1.0000	1.0656

CONCLUSIONS

This paper looks at the comparison of two accumulator choice models from mathematical psychology against models typically used in choice modelling. We find that the multi-attribute linear ballistic accumulator does not provide as good a fit as decision field theory for our two stated preference route choice datasets but that both provide better fits than RUM for one dataset and both do for the other depending on how attributes are considered. This is because it appears that the logistic parameter within MLBA provides a natural mechanism for capturing non-linear sensitivities in the data. Simple versions of the model work well but gains in fit are less substantial for MLBA than random utility models when parameters are added to capture complexities in the data. This is particularly evident in our simulated datasets which demonstrate that both MLBA and DFT fit the data substantially better than RUM when 6-9 parameters are used, but that RUM gains back some of this deficit back when 12-15 parameters are used.

We also consider the effects of different scaling methods on the attributes. As a contrast to previous results (*14*), we find that DFT outperforms MLBA regardless of which type of scaling is used and regardless of whether we use a Euclidean or a psychological distance function in our MLBA model.

This paper also considered the impacts of including parameters to capture underlying preferences in MLBA and DFT. Results from our UK dataset suggest that DFT makes more substantial gains than MLBA as well as our RUM and RRM models. We have, however, only considered one method for incorporating preferences in both of these models. Whilst we added parameters to the drift rate in MLBA, alternative specifications would allow for the adjustment of the starting point A or the threshold χ. It is easily possible that some alternatives may not require as much evidence to be chosen (for example, a commuter's usual route to work), meaning that an MLBA model including alternative specific thresholds may work well. This will be investigated in future research.

Additionally, the linear ballistic accumulator (*17*), a simplified version of MLBA for two alternatives, has been demonstrated to work well with dynamic datasets where attributes change over time (*28*). A similar concept could be applied to both DFT and MLBA, for which changing attributes could easily be incorporated. Thus DFT and MLBA may work well with dynamic revealed preference datasets such as the lane merging decisions made by drivers.

This research thus serves as a 'proof of concept' that accumulator models such as DFT and MLBA are an attractive alternative approaches to RUM and hold significant promise in improving the behavioural realism in choice models, both in transport and beyond.

ACKNOWLEDGEMENTS

The authors acknowledge the financial support by the European Research Council through the consolidator grant 615596-DECISIONS.

REFERENCES

[1] K. Train, "Discrete choice with simulation," 2003.

[2] A. Tversky, "Elimination by aspects: A theory of choice." *Psychological review*, vol. 79, no. 4, p. 281, 1972.

[3] A. Tversky and D. Kahneman, "Availability: A heuristic for judging frequency and probability," *Cognitive psychology*, vol. 5, no. 2, pp. 207–232, 1973.

[4] ——, "Judgment under uncertainty: Heuristics and biases," in *Utility, probability, and human decision making*. Springer, 1975, pp. 141–162.

[5] R. M. Roe, J. R. Busemeyer, and J. T. Townsend, "Multialternative decision field theory: A dynamic connectionst model of decision making." *Psychological review*, vol. 108, no. 2, p. 370, 2001.

[6] J. S. Trueblood, S. D. Brown, A. Heathcote, and J. R. Busemeyer, "Not just for consumers context effects are fundamental to decision making," *Psychological science*, vol. 24, no. 6, pp. 901–908, 2013.

[7] T. Noguchi and N. Stewart, "In the attraction, compromise, and similarity effects, alternatives are repeatedly compared in pairs on single dimensions," *Cognition*, vol. 132, no. 1, pp. 44–56, 2014.

[8] J. R. Busemeyer and J. T. Townsend, "Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment." *Psychological review*, vol. 100, no. 3, p. 432, 1993.

[9] C. G. Chorus, T. A. Arentze, and H. J. Timmermans, "A random regret-minimization model of travel choice," *Transportation Research Part B: Methodological*, vol. 42, no. 1, pp. 1–18, 2008.

[10] C. G. Chorus, "A new model of random regret minimization," *EJTIR, 10 (2), 2010*, 2010.

[11] J. Swait, "A non-compensatory choice model incorporating attribute cutoffs," *Transportation Research Part B: Methodological*, vol. 35, no. 10, pp. 903–928, 2001.

[12] N. A. Berkowitsch, B. Scheibehenne, and J. Rieskamp, "Rigorously testing multialternative decision field theory against random utility models." *Journal of Experimental Psychology: General*, vol. 143, no. 3, p. 1331, 2014.

[13] T. O. Hancock, S. Hess, and C. F. Choudhury, "Decision field theory: improvements to current methodology and comparisons with standard choice modelling techniques," 2017, the 5th International Choice Modelling Conference.

[14] J. S. Trueblood, S. D. Brown, and A. Heathcote, "The multiattribute linear ballistic accumulator model of context effects in multialternative choice." *Psychological review*, vol. 121, no. 2, p. 179, 2014.

[15] B. E. Kaufman, "A new theory of satisficing," *Journal of Behavioral Economics*, vol. 19, no. 1, pp. 35–51, 1990.

[16] B. Schwartz, A. Ward, J. Monterosso, S. Lyubomirsky, K. White, and D. R. Lehman, "Maximizing versus satisficing: happiness is a matter of choice." *Journal of personality and social psychology*, vol. 83, no. 5, p. 1178, 2002.

[17] S. D. Brown and A. Heathcote, "The simplest complete model of choice response time: Linear ballistic accumulation," *Cognitive psychology*, vol. 57, no. 3, pp. 153–178, 2008.

[18] E. Stern and H. W. Richardson, "Behavioural modelling of road users: current research and future needs," *Transport Reviews*, vol. 25, no. 2, pp. 159–180, 2005.

[19] G. Zhao, C. Wu, and B. Ou, "Mathematical modeling of average driver speed control with the integration of queuing network-model human processor and rule-based decision field theory," in *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55, no. 1. Sage Publications, 2011, pp. 856–860.

[20] H. Qin, H. Guan, and Y.-J. Wu, "Analysis of park-and-ride decision behavior based on decision field theory," *Transportation research part F: traffic psychology and behaviour*, vol. 18, pp. 199–212, 2013.

[21] T. Otter, J. Johnson, J. Rieskamp, G. M. Allenby, J. D. Brazell, A. Diederich, J. W. Hutchinson, S. MacEachern, S. Ruan, and J. Townsend, "Sequential sampling models of choice: Some recent advances," *Marketing letters*, vol. 19, no. 3-4, pp. 255–267, 2008.

[22] G. E. Hawkins, A. Marley, A. Heathcote, T. N. Flynn, J. J. Louviere, and S. D. Brown, "Integrating cognitive process and descriptive models of attitudes and preferences," *Cognitive science*, vol. 38, no. 4, pp. 701–735, 2014.

[23] J. M. Hotaling, J. R. Busemeyer, and J. Li, "Theoretical developments in decision field theory: comment on tsetsos, usher, and chater (2010)." *Psychological review*, vol. 117, no. 4, pp. 1294–1298, 2010.

[24] J. R. Busemeyer and A. Diederich, "Survey of decision field theory," *Mathematical Social Sciences*, vol. 43, no. 3, pp. 345–370, 2002.

[25] M. Fosgerau, "Investigating the distribution of the value of travel time savings," *Transportation Research Part B: Methodological*, vol. 40, no. 8, pp. 688–707, 2006.

[26] S. Hess and A. Stathopoulos, "A mixed random utility - random regret model linking the choice of decision rule to latent character traits," *Journal of choice modelling*, vol. 9, pp. 27–38, 2013.

[27] CMC, "CMC choice modelling code for R, Choice Modelling Centre, University of Leeds, UK," 2017. [Online]. Available: http://www.cmc.leeds.ac.uk

[28] W. R. Holmes, J. S. Trueblood, and A. Heathcote, "A new framework for modeling decisions about changing information: The piecewise linear ballistic accumulator model," *Cognitive psychology*, vol. 85, pp. 1–29, 2016.