Psychonomic Bulletin & Review
https://doi.org/10.3758/s13423-024-02483-7

THEORETICAL/REVIEW

EZ-CDM: Fast, simple, robust, and accurate estimation of circular diffusion model parameters

Hasan Qarehdaghi[1] · Jamal Amani Rad[1]

Accepted: 17 February 2024

Abstract
The investigation of cognitive processes that form the basis of decision-making in paradigms involving continuous outcomes has gained the interest of modeling researchers who aim to develop a dynamic decision theory that accounts for both speed and accuracy. One of the most important of these continuous models is the circular diffusion model (CDM, Smith. *Psychological Review*, *123*(4), 425. 2016), which posits a noisy accumulation process mathematically described as a stochastic two-dimensional Wiener process inside a disk. Despite the considerable benefits of this model, its mathematical intricacy has limited its utilization among scholars. Here, we propose a straightforward and user-friendly method for estimating the CDM parameters and fitting the model to continuous-scale data using simple formulas that can be readily computed and do not require theoretical knowledge of model fitting or extensive programming. Notwithstanding its simplicity, we demonstrate that the aforementioned method performs with a level of accuracy that is comparable to that of the maximum likelihood estimation method. Furthermore, a robust version of the method is presented, which maintains its simplicity while exhibiting a high degree of resistance to contaminant responses. Additionally, we show that the approach is capable of reliably measuring the key parameters of the CDM, even when these values are subject to across-trial variability. Finally, we demonstrate the practical application of the method on experimental data. Specifically, an illustrative example is presented wherein the method is employed along with estimating the probability of guessing. It is hoped that the straightforward methodology presented here will, on the one hand, help narrow the divide between theoretical constructs and empirical observations on continuous response tasks and, on the other hand, inspire cognitive psychology researchers to shift their laboratory investigations towards continuous response paradigms.

Keywords Decision-making · Continuous response · Cognitive modeling · Circular diffusion model · Response time · Method of moments

Introduction

Time-course of simple decision-making

Typically, experimental data in cognitive psychology are derived from tasks requiring subjects to choose between a small number of alternatives (usually sets of two or three), between a large number of alternatives, or along a continuum. Occasionally, the bare choice data of the subjects in these experiments may not suffice to draw inferences about the covert underlying phenomenon of interest in the study. Examples of such scenarios may occur in research pertaining to a particular temporal stage of decision-making, such as the duration of a specific cognitive process. In addition, these situations may arise when the data are influenced by the inherent criterion setting that regulates the trade-off between speed and accuracy of performance. A dynamic cognitive model of the decision-making time course could be quite beneficial in these circumstances.

The class of speeded decision-making models known as *evidence accumulation models*, or *sequential sampling models*, involve the accumulation of information over time for possible choice options. These models can assess both

The early idea of this work was presented as a talk at the 44th Annual Conference of the Cognitive Science Society (CogSci 2022) as well as the MathPsych/ICCM 2022 in July 2022 and published as a short paper in the CogSci Proceedings.

✉ Jamal Amani Rad
j_amanirad@sbu.ac.ir; j.amanirad@gmail.com

1 Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran

Published online: 08 April 2024

response time and accuracy simultaneously (Stone, 1960; Ratcliff, 1978). The most successful models in this context view the process of information accumulation as a stochastic process that is continuous in both time and evidence space. In these models, the relevant information is continually sampled, and each sample contributes an amount of evidence in favor of or against one or more possible response options. The evidence from consecutive samples is aggregated over time until a criterion level of evidence supporting a particular response option is reached, which triggers a corresponding behavioral response. The influential parameters in these models are the *drift rate*, which refers to the relative speed of information accumulation in favor of each alternative; the *decision criterion*, which represents the amount of information required to choose each alternative; and the *non-decision time*, which accounts for the time taken for processes other than evidence accumulation, such as encoding stimulus and executing response.

Many models in the framework of evidence accumulation have been developed for the two alternative decision tasks (Ratcliff, 1978; Heath, 2000; Usher & McClelland, 2001; Brown & Heathcote, 2005, 2008; Drugowitsch et al., 2012; Hawkins et al., 2015; Voss et al., 2019; Tillman et al., 2020; Shinn et al., 2020; van Ravenzwaaij et al., 2020; Hawkins & Heathcote, 2021; Miletić et al., 2021) which are extensively used in vast areas of clinical and experimental psychology such as evaluating the sustained attention in children with attention deficit hyperactivity disorder (ADHD) (Huang-Pollock et al., 2020), the role of speed in ADHD-related working memory deficits (Weigard & Huang-Pollock, 2017), the pattern of information processing in children with ADHD (Nejati et al., 2022), perceptual decision-making in autistic children (Manning et al., 2022a), visual motion and decision-making in dyslexia (Manning et al., 2022b), the optimality of decision policies (Bogacz et al., 2006; Drugowitsch et al., 2012; Evans et al., 2020; Evans & Brown, 2017; Evans et al., 2018; Starns & Ratcliff, 2012), stop signal paradigms (Matzke et al., 2013, 2017a, b), Go/No-Go paradigms (Gomez et al., 2007; Ratcliff et al., 2018), learning strategies (Pedersen et al., 2017; Fontanesi et al., 2019; Sewell et al., 2019), attentional choice (Krajbich et al., 2010, 2012), neural processes (Gold & Shadlen, 2007) and so on (interested readers can also refer to these two reviews, Ratcliff et al. (2016) and Forstmann et al. (2016)).

Dynamic decision models for continuous report

While two-alternative decision tasks (*two-choice tasks*) have been extensively used in the history of cognitive psychological laboratory research, there has been a growing interest in the use of tasks that entail deciding between many alternatives or along a continuum (*continuous report tasks*). Although two-choice tasks have undergone frequent validation and are straightforward to implement, with participants finding them easy to comprehend and having sophisticated cognitive models, at times these tasks may not accurately reflect the natural cognitive processes being studied, or they may fail to capture the desired information due to their restriction to a binary or small number of choices. Recent applications of continuous report tasks show intriguing findings that are very difficult, if not impossible, to derive by using two-choice tasks (Zhang & Luck, 2008; Bays et al., 2009; Fougnie & Alvarez, 2011; Bays et al., 2011; Fougnie et al., 2012; van den Berg et al., 2012; Marshall & Bays, 2013; Fan & Turk-Brown, 2013; Kool et al., 2014; Swan & Wyble, 2014; van den Berg et al., 2014; Ma et al., 2014; Gunseli et al., 2015; Green & Pratte, 2022).

Different task conditions that manipulate instruction, reward, cue, and difficulty can alter performance in continuous-report tasks (Ratcliff (2018), and Kvam (2019b)). Variations in performance exist among individuals in addition to conditions. Differences in the underlying processes of such tasks are anticipated to be the source of these conditional and individual differences. This highlights the necessity for a dynamic theoretical framework of decision-making processes that can effectively model continuous-report paradigms.

In the past few years, there has been a development of theories pertaining to the accumulation of evidence in continuous decision tasks. The *circular diffusion model* (CDM) is a primary theory that is employed for a commonly used subset of continuous decision tasks where the alternatives are naturally arranged around a circle (Smith, 2016; Smith & Corbett, 2019; Smith, 2019; Kvam, 2019b; Smith et al., 2020; Zhou et al., 2021). The *spatially continuous diffusion model* is a more general model that involves a greater number of parameters and incorporates additional cognitive factors into the model (Ratcliff, 2018; Ratcliff & McKoon, 2020). Additionally, alternative models have been proposed, including the *geometric similarity representation model* (Kvam & Turner, 2021), which synthesizes the preceding two models, and the *multiple anchored accumulation theory* (Kvam et al., 2021), which accounts for the influence of reference points in the stimulus spectrum.

Parameter estimation

Although powerful continuous models have significant advantages in explaining experimental data for continuous-report tasks based on underlying psychological processes, their mathematical complexity has limited their practical use among researchers. Consequently, cognitive researchers continue to rely on experiments with a limited number of options. A significant limitation in the widespread adoption of these models within the field of cognitive psychology may be attributed to the technical challenges associated with model fitting and parameter estimation, which are necessary for capturing the effects of experimental manipulations and individual differences.

Different methods can be employed to fit a model to the experimental data. *maximum likelihood estimation* (MLE) is a standard approach that usually yields a low disparity between the estimated values and the actual data-generating values of model parameters, i.e., has high *accuracy* in estimation (read Stigler (2007) for the historical story behind the MLE). The approach involves identifying the parameter values that maximize the probability of producing the observed data. The implementation of the MLE may pose certain challenges in practice, as the computation of the likelihood function with an adequate level of precision or the utilization of an efficient optimization algorithm for the maximization of the likelihood can turn out to be difficult tasks. Yet, in some instances, the maximum likelihood estimates could be derived from simpler equations (see section 2.4 in Bishop (2006)). For example, the maximum likelihood estimate of a normal model, parameterized by its mean and variance, is simply the mean and variance statistics of data points (see section 2.3.4 in Bishop (2006)). In such cases, there is no need to calculate the likelihood or use an optimization algorithm.

However, when the model is not normal, the mean and variance of the data are not the maximum likelihood estimates. Despite this, it can still give roughly good estimates, which is why the *method of moments* has reasonable accuracy in estimating parameters. This method is a traditionally used estimation technique that has its roots in the work of Pearson (1894). In the method of moments, the k parameters of a model are estimated in such a way that the first k moments of the data and the model become equal (the mean of X^i is the i^{th} moment of a random variable X). The proposed methodology, while straightforward, may exhibit suboptimal precision in its estimations due to its reliance on a small number of statistics that may not contain sufficient information about the whole dataset. Wagenmakers et al. (2007) used this method to overcome the challenge of fitting the *diffusion decision model (DDM)* (Ratcliff, 1978) in two-choice tasks by introducing a simple approach named the "EZ diffusion model" or "EZ-DDM". In the EZ diffusion model, the parameters were calculated from summary statistics of the data using quite simple formulas (see Wagenmakers et al. (2008); Wagenmakers (2009); van Ravenzwaaij & Oberauer (2009); van Ravenzwaaij et al. (2017) for more details).

Across-trial variability

In the framework of evidence accumulation, most models come in versions that incorporate the variability of the primary parameters (Laming, 1968; Ratcliff, 1978; Ratcliff & Tuerlinckx, 2002; Ratcliff & Smith, 2004; Brown & Heathcote, 2008; Ratcliff, 2013). In this situation, the researcher is faced with the dilemma of determining the appropriate model to utilize. We first briefly discuss the advantages and disadvantages of each version of the model using the celebrated diffusion decision model for two-choice decision-making as an example.

The core subprocesses of decision-making may experience fluctuations due to various potential factors. So, it is plausible to contemplate the variability in the main parameters of the evidence accumulation model. Besides that, considering the across-trial variability in the model, the model predicts specific patterns that are actually present in the empirical data. For example, the inclusion of starting point variability (or decision criterion) enables the diffusion decision model to produce fast errors (Laming, 1968). Also, variability in drift generates a slow error pattern (Ratcliff, 1978) and variability in non-decision time could improve the fit to the leading edge of the response time distribution Ratcliff & Tuerlinckx (2002). However, incorporating across-trial variability into the model may not be advantageous in all respects. For example, in the case of the diffusion decision model, the reliability of parameter values across experimental sessions is limited (Lerche & Voss, 2017; Yap et al., 2012). Furthermore, the estimation of variability parameters necessitates a greater degree of technical expertise (Boehm et al., 2018). On the other hand, incorporating variability into a model increases its complexity, potentially leading to accuracy issues in recovery when compared to more parsimonious models. For example, in datasets with small to moderate trial numbers, the estimation of variability parameters are imprecise, leading to inaccuracies in the estimation of primary parameters within the diffusion decision model (Ratcliff & Tuerlinckx, 2002; Ratcliff & Childers, 2015; Lerche & Voss, 2016; Lerche et al., 2017). Therefore, it can probably be said that a model with reduced complexity, assuming no variability by fixing variability parameters to zero, exhibits superior performance in discerning variations in parameter values, despite the presence of non-zero variability in the true underlying data-generating model (van Ravenzwaaij & Oberauer, 2009; van Ravenzwaaij et al., 2017). In summary, the complex model, including variability parameters, is a more realistic representation and exhibits superior performance in fitting data. However, it is not suitable for measuring the primary parameter values and conducting group or condition comparisons based on these measurements.

The circular diffusion model also benefits from incorporating variability when fitted to the data (Smith et al., 2020; Zhou et al., 2021). For example, variability in drift length and decision criterion could predict slow and fast errors, respectively (Smith, 2016). Also, it is expected that there will be fluctuations in psychological processes related to each of the primary parameters of the CDM. So the inquiry pertains to the inclusion or exclusion of variability parameters. Does the inclusion of variability impede the measurement of parameter values in the circular diffusion model? The relation between this topic and EZ-CDM is that the EZ-CDM formulation is based on the simple CDM, considering no variability. So it is not known if this method could perform satisfactorily when the data are generated by the model with variability, as is expected to happen in empirical applications. We will examine this issue thoroughly in this article.

The nature of EZ-CDM

There may be confusion about the genuine nature of the EZ-CDM, which originates from the unsettled view of the essence of the EZ-DDM. Wagenmakers et al. (2007) introduced the EZ-DDM in the article titled “An EZ-diffusion **model** for response time and accuracy”. This article perceives EZ-DDM **as a model** based on the DDM with some simplifying assumptions. There is an alternate perspective that views the EZ-DDM **as an estimation method** for the DDM. For example, see the article by Ratcliff (2008) titled “The EZ diffusion **method**: Too EZ?”. We argue that the latter view is more consistent with the EZ-DDM’s inherent nature.

When EZ-DDM is viewed as a model, two important issues arise. The first issue concerns the theoretical importance of the EZ-DDM. However, it has no theoretical significance in the psychology of decision-making, independent of the DDM. Another issue is that the EZ-DDM, as a simple model, fails to predict some qualities of empirical data, like fast or slow errors. This indicates that EZ-DDM is a false model. These problems cease to exist when the EZ-DDM is viewed as an estimation method. No one expects an estimation method to have theoretical significance. Furthermore, it is meaningless to judge whether an estimation method is true or false.

In this article, our perspective on EZ-CDM is as an estimation method for the CDM. So our only concern is its accuracy in estimating the parameters of this model.

Our work

In this study, we present a novel method, EZ-CDM, for parameter estimation of the circular diffusion model. This method utilizes both maximum likelihood estimation and method of moments techniques, resulting in a straightforward, efficient, quite precise, and easily executable method. Maximum likelihood estimation (MLE) is employed in cases where an analytical expression can be derived for the estimation. In situations where such an expression is not feasible, the method of moments is used. This approach preserves both precision and simplicity in parameter estimation. However, the method may suffer from the lack of accuracy induced by the method of moments. We conducted an analysis to assess the extent of this issue. Another problem that the method could suffer from is the inclusion of corrupted data. We investigated the problem and proposed a resilient version of the method, named robust EZ-CDM, that could handle the contamination. Also, the EZ-CDM faces a challenge by disregarding an important aspect of the model which is the inclusion of across-trial variability in main parameters. This matter has been thoroughly investigated in the present article. As the last part of our analysis, we offer an illustration of the employment of the estimation technique for the mixed model involving the CDM and guess.

The subsequent sections of the article are organized in the following manner: “Circular diffusion model” section provides a brief overview of the circular diffusion model. In “EZ-CDM” section, the EZ-CDM is presented as a method for estimating the CDM parameters. In “Parameter recovery” section, a comparison is made between the parameter recovery of the method and that of the maximum likelihood estimation method. The fifth section comprises an examination of the influence of contaminants on the EZ-CDM and the introduction of the robust EZ-CDM. In “EZ-CDM performance in presence of across-trial variability” section, the EZ-CDM’s performance is demonstrated when the data comes from the CDM with across-trial variability in the main parameters. “Experimental data containing guesses” section presents the implementation of the method in conjunction with the approach for estimating guest proportion on experimental data. “Discussion” and “Conclusion” sections of the article comprise a discussion and conclusion, respectively.

Circular diffusion model

The circular diffusion model (Smith, 2016) belongs to the evidence accumulation category, which describes the dynamic process of decision-making in a circular decision task. In this task, participants should select one point from a circle of decision alternatives. The responses in these tasks exhibit a continuous nature wherein extreme options are naturally bound together (i.e., there is no extreme). Direction, orientation, or color (when represented in a wheel) are of this type (for example see tasks in Unsworth et al. (2014)). The quantification of participants’ performance in each trial can be represented by the response time and the angle of the response on the circle (or the angular difference between the response and the correct option if the data are aggregated for different correct options). An example of a color task popular in visual working memory research (Zhang & Luck, 2008; Bays et al., 2009), along with the model representation for a single trial, is illustrated in Fig. 1.

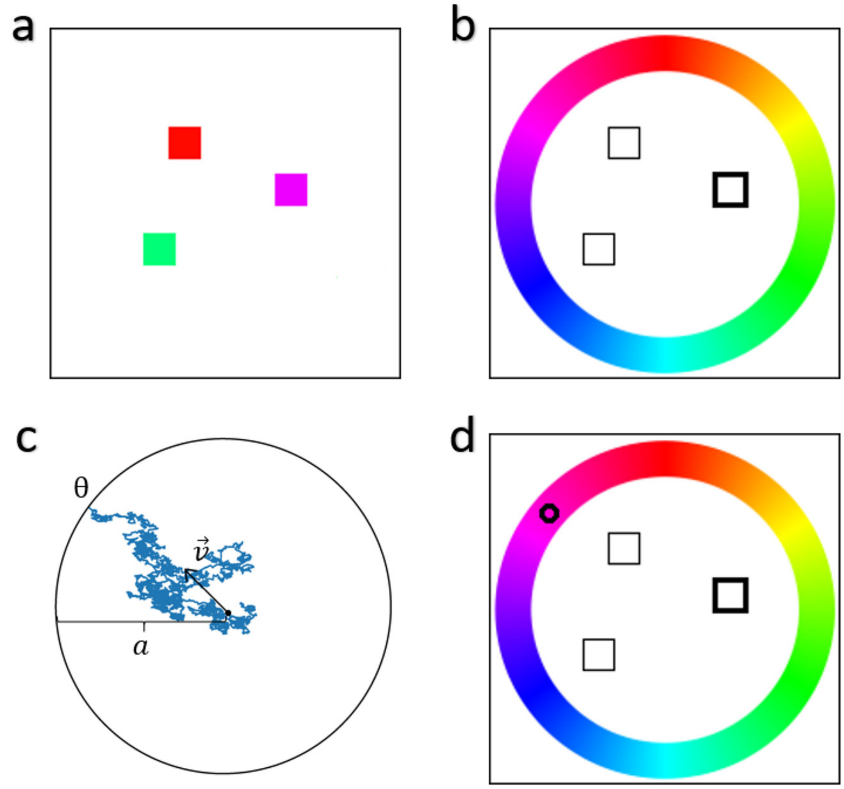


Fig. 1 The CDM model in one trial of a circular decision task. **a** The participant is presented with three squares of different colors. **b** The participant is asked to choose the color of the probed square by moving the cursor to a location on the circle. **c** The CDM model assumes that the process of decision-making consists of a Wiener diffusion process on the plane with a drift vector v, representing the encoded stimulus in memory. The process runs until it hits the boundary circle of radius a. **d** The hitting point determines the decided color, and the response time is determined by the time it takes for the process to hit the boundary plus non-decision time T_{er}

This model depicts the belief or evidence space as a two-dimensional plane where each point represents a potential state that the accumulated evidence may be in during the decision-making process. In this space, each direction corresponds to a distinct response option, such that the angle of the state point can be interpreted as a manifestation of the option with the highest accumulated evidence so far. Also, the magnitude of the point serves as a measure of the quantity of evidence collected in favor of this option (for more description, see Kvam (2019a); Kvam & Turner (2021)).

The process of evidence accumulation typically begins at the origin, where no prior evidence or tendency towards any option is assumed. The state of evidence evolves as new information is received or processed over time. This is depicted in the model by the change in the location of the state point in the plan. Based on CDM principles, the diffusion process governs this updating of the state point as:

$$\boldsymbol{X}(t) = \overrightarrow{\boldsymbol{v}}\, t + \sigma \boldsymbol{W}(t), \qquad (1)$$

where $\boldsymbol{X}(t)$ is the state point in the plane at time t, $\boldsymbol{W}(t)$ is a two-dimensional Wiener diffusion process, $\overrightarrow{\boldsymbol{v}}$ and σ are respectively drift vector and diffusion coefficient. See Chapter 10 in Ross (2007) for an introduction to the diffusion process.

The above equation shows that the mean vector and variance-covariance matrix of location change distribution in time Δt will be $\overrightarrow{\boldsymbol{v}}\,\Delta t$ and $\sigma^2 \Delta t \boldsymbol{I}$ where $\boldsymbol{I}$ is the identity matrix. Therefore, σ^2 indicates the rate of variability that is accumulated over time, independently for two dimensions. The drift vector controls the systematic tendency of state change over time. The mean directional change of the state point aligns with the drift rate direction, leading to the interpretation of the average response identity for the drift angle (represented as θ_v). Also, the state point moves in rate of drift length (represented as v) giving it the meaning of the intensity of the average response. The relationship between the polar and Cartesian coordinates for the drift vector is as follows:

$$\begin{cases} v_1 = v \cos\theta_v, \\ v_2 = v \sin\theta_v, \end{cases} \qquad \begin{cases} \theta_v = \text{atan2}(v_2, v_1), \\ v = \sqrt{v_1^2 + v_2^2}, \end{cases}$$

where atan2 function is an arctangent function with two arguments that give the angle whose cosine and sine ratio and signs are identical to the ratio and signs of the arguments in respective order.

Considering the interpretation of the state point magnitude, it is reasonable to adopt a threshold (denoted as a) as the decision process termination criterion. This criterion signifies the achievement of a satisfactory level of evidence accumulated for selecting an option. So, the termination occurs when the state point reaches a circle of radius a around the origin. The angle of the hitting point determines the chosen alternative, and the response time is the sum of the time taken to hit the circle plus the non-decision time (represented as t_0). Please see Fig. 1 which represents the accumulating evidence on a single experimental trial.

As a final point, note that the model in its current form lacks parameter identifiability. Since multiplying v, a, and σ by a constant, would not change the predictions of the model. Typically, in order to constrain the model, a constant diffusion coefficient is employed. Throughout this paper, we consider $\sigma = 1$. For alternative constraints, the equations should be reconsidered.

In sum, the circular diffusion model has the following parameters:

- Drift angle denoted by θ_v, represents the average response identity.
- Drift rate length denoted by v, represents the average response intensity.
- Decision criterion denoted by a, represents the satisfactory level of evidence for choosing.
- Non-decision time denoted by t_0, represents the time taken by mechanisms other than the evidence accumulation.

EZ-CDM

Before going to the technical material, we would like to first give a clear top-down picture of how the parts relate to one another or where the argument is going in the form of a road map here. The key point here is that EZ-CDM is based on predicting four simple statistics rather than using the entire dataset: two statistics from the choice angle data (circular mean MCA and circular variance VCA), and two statistics from the response time data (mean MRT and variance VRT). The unique characteristics of the choice angle and response time distributions in the CDM allow for the formulation of equations that can determine the mean and variance of these distributions based on the values of the CDM parameters. The idea is that the mean and variance of the choice angle and response time distributions of CDM should match the corresponding statistics derived from the data. Based on this idea, a system of four equations can be formulated for the four statistics, and then the system can be solved with respect to our four parameters (drift rate angle θ_v, drift rate magnitude v, threshold a, and non-decision time t_0).

The possibility of an EZ-CDM relies on three results established in Smith (2016). First, the predicted decision times and decision outcomes in the simple CDM (where there is no across-trial variability in drift rate and decision criterion) are independent of each other, so we could fit the model separately on the decision outcome and decision time data. This means, the presence of statistics composed of the mixture of choice angle and response time data is not necessary, and we are allowed to use separate marginal statistics of the mean and variance of choice angle and response time. Second, the marginal distribution of the choice angle is von Mises in form and has straightforward formulas for its mean and variance, enabling us to relate the parameter values of the CDM to the predictions of the circular mean and variance of the choice angle, ignoring the decision times. Third, the marginal distribution of the decision time is the first passage time of a Bessel process with drift, which has a closed-form moment-generating function. This enables us to relate the parameter values of CDM to the mean and variance of the response time distribution. In particular, the precision parameter of the von Mises distribution is equal to the product of the drift rate norm and the decision criterion, expressed in diffusion coefficient units. If the latter is treated as the scaling constant for the process and set to unity, precision is simply the drift rate norm times the criterion.

To summarize all these mentioned features and have a top-down view, the road map can be defined as follows: the estimation of the parameters of the decision-time distribution can be handled independently of the outcomes using the expression for the moment generating function for the decision times in Smith (2016). The mean and variance of the decision times also depend on the drift rate norm and the decision criterion. However, unlike the expression for the precision of the outcome distribution, the moments of the decision-time distribution depend both on the product of drift rate and criterion and their ratio. The trick is that this property provides sufficient structure to allow the equations to be solved for each of these parameters independently (It is analogous to solving simultaneous equations for the logs of two unknowns). The mean non-decision time can be then estimated from the mean RT and the mean decision time by subtraction in the usual way.

The arrangement of content in the rest of this section is as follows: Initially, we illustrate the independence of the choice angle and response time distribution in the CDM. Subsequently, we show the derivation of the formulas relating the mean and variance of the choice angle and response time to the model parameters. In the next step, we build a system of four equations utilizing the aforementioned formulas, wherein the mean and variance of the model's predictions are substituted with the corresponding statistics of the data. Ultimately, we solve this equation with respect to the parameters of the CDM.

Response time and response outcome are independent

The likelihood function for the CDM is given in Smith (2016) as an infinite series that, when expressed in polar coordinates, takes the following form:

$$
\begin{aligned}
Pr(\theta, t) =& \frac{1}{2\pi a^2} \exp\Big(va\cos(\theta - \theta_v) - \frac{v^2(t-t_0)}{2}\Big) \\
&\times \sum_{i=1}^{\infty} \frac{j_{0,i}}{J_1(j_{0,i})} \exp\big(-\frac{j_{0,i}^2}{2a^2}(t-t_0)\big),
\end{aligned} \tag{2}
$$

where θ and t are choice angle and response time, J_1 is the first-order Bessel function of the first kind and $j_{0,i}$ is the i^{th} zero of the zero-order Bessel function of the first kind. The equation above can be rewritten as:

$$
\begin{aligned}
Pr(\theta, t) =& \frac{1}{2\pi I_0(av)} \exp\Big(va\cos(\theta - \theta_v)\Big) \\
&\times \frac{I_0(av)}{a^2} \sum_{i=1}^{\infty} \frac{j_{0,i}}{J_1(j_{0,i})} \exp\big(-\frac{(t-t_0)}{2}(v^2 + \frac{j_{0,i}^2}{a^2})\big),
\end{aligned}
$$

where I_0 is the first kind of modified Bessel function of order zero. The first factor is a probability function (integrates to one) and depends only on the θ (not t), and the second factor depends only on t (not θ). Therefore, the first and second factors are, respectively, the marginal probability of θ and t. Hence, we can write:

$$Pr(\theta, t) = Pr(\theta) \times Pr(t),$$

which yields that the choice angle and response time are independent. So, the coincidence of them is not important and we treat the choice and response time data separately.

Marginal distribution of response outcome

The choice angle has a von Mises probability distribution (see Section 3.5.4 in Mardia & Jupp (1999) for more information about this distribution). This distribution is defined on the circle (its support could be considered $[-\pi, \pi]$) and parametrized by its mean μ and concentration κ (acts as inverse of dispersion):

$$Pr(\theta) = e^{\kappa\cos(\theta-\mu)}/2\pi I_0(\kappa).$$

So, the relationship between the parameters of the CDM and von Mises is as follows:

$$
\begin{cases}
\mu = \theta_v, \\
\kappa = av.
\end{cases}
$$

This indicates that the drift angle is the only parameter controlling the mean choice angle. The most probable outcome of response is aligned with the drift angle. Also, the multiplication of drift length and decision criterion controls the width of the choice angle distribution. The higher drift length and decision criterion result in closer responses to the mean.

It is known that the maximum likelihood estimate of von Mises parameters should satisfy the equations (see Section 5.3.1 in Mardia & Jupp (1999)):

$$
\begin{cases}
\mu = \operatorname{atan2}\Big(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n, \frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\Big), \\
1 - \frac{I_1(\kappa)}{I_0(\kappa)} = 1 - \sqrt{\Big(\frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\Big)^2 + \Big(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n\Big)^2},
\end{cases}
$$

where I_1 is the first kind of modified Bessel function of order one, N is the number of data points, and θ_n is the choice angle in the n^{th} trial.

The right-hand side of equality in the first equation above is known as the circular mean of data. The right-hand side of the second equation is called the circular variance of data (see Chapter 2 in Mardia & Jupp (1999)). The left-hand side of equalities shows the circular mean and variance of the von Mises distribution (see Section 3.5.4 in Mardia & Jupp (1999)). So, the equations that relate the circular mean and variance of the data to the parameters of von Mises distribution are actually the formulation of the maximum likelihood estimation.

Replacing the von Mises parameters with the CDM parameters will yield our first two equations:

$$
\begin{cases}
\theta_v = \operatorname{atan2}\Big(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n, \frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\Big), \\
1 - \frac{I_1(av)}{I_0(av)} = 1 - \sqrt{\Big(\frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\Big)^2 + \Big(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n\Big)^2}.
\end{cases}
$$

Now we need two more equations to be able to solve for our four parameters. For this, we will turn into the response time data and build equations by using the first two moments of the model and data.

Marginal distribution of response time

The marginal probability of response time is the first passage time of a Bessel process of order two with drift (plus the non-decision time t_0) (see Section 4.6 in Borodin & Salminen (2002) for information about the Bessel process). This distribution has the moment generating function (Laplace transform) that if we add t_0, will be:

$$E[e^{-\lambda t}] = \frac{e^{-\lambda t_0} I_0(av)}{I_0\left(a\sqrt{2\lambda + v^2}\right)}.$$

This function can help to derive formulas for moments of response time (read Section 2.3 in Casella & Berger (2001)). The first moment is:

$$E[T] = -\frac{d}{d\lambda} E[e^{-\lambda t}]\Big|_{\lambda=0} = t_0 + \frac{a}{v}\frac{I_1(av)}{I_0(av)},$$

where we have used the fact that $(d/dx) I_0(x) = I_1(x)$. Also, the second moment is:

$$\begin{aligned} E[T^2] &= \frac{d^2}{d\lambda^2} E[e^{-\lambda t}]\Big|_{\lambda=0} \\ &= \frac{1}{I_0^4}\left(I_0^4 t_0^2 + 2\frac{a}{v} I_0^3 I_1 t_0 + 2\frac{a}{v^3} I_0^3 I_1 - \frac{a^2}{v^2} I_0^4 + 2\frac{a^2}{v^2} I_0^2 I_1^2\right), \end{aligned}$$

where all Bessel functions are calculated at av (superscripts are exponents). Also, it should be noted that we used the fact that $(d/dx) I_1(x) = I_0(x) - I_1(x)/x$ in the above calculations.

We convert the formulas of moments to mean and variance. The mean equals the first moment and the variance can be calculated from the first two moments as:

$$E[T^2] - E[T]^2 = \frac{a^2}{v^2}\frac{I_1^2(av)}{I_0^2(av)} + \frac{2a}{v^3}\frac{I_1(av)}{I_0(av)} - \frac{a^2}{v^2}.$$

Upon examining the aforementioned formula and mean response time equation, it is evident that the non-decision time solely impacts the mean response time by causing a shift. The impact of decision criterion and drift length on the mean and variance of response time is nonlinear.

System of equations

The estimation procedure begins by calculating the following statistics of data (mean of choice angle (MCA), the variance of choice angle (VCA), mean of response time (MRT), and variance of response time (VRT)):

$$\begin{cases} MCA = atan2\left(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n, \frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\right), \\ VCA = 1 - \sqrt{\left(\frac{1}{N}\sum_{n=1}^{N}\cos\theta_n\right)^2 + \left(\frac{1}{N}\sum_{n=1}^{N}\sin\theta_n\right)^2}, \\ MRT = \frac{1}{N}\sum_{n=1}^{N} t_n, \\ VRT = \frac{1}{N}\sum_{n=1}^{N}\left(t_n - \frac{1}{N}\sum_{n=1}^{N} t_n\right)^2, \end{cases} \tag{3}$$

where t_n is the n^{th} response time data point. Then, these values will be passed to the system of equations:

$$\begin{cases} \theta_v = MCA, \\ 1 - \frac{I_1(\kappa)}{I_0(\kappa)} = VCA, \\ t_0 + \frac{a}{v}\frac{I_1(av)}{I_0(av)} = MRT, \\ \frac{a^2}{v^2}\frac{I_1^2(av)}{I_0^2(av)} + \frac{2a}{v^3}\frac{I_1(av)}{I_0(av)} - \frac{a^2}{v^2} = VRT. \end{cases}$$

Note that $\kappa = av$. Now we try to solve this system of equations for the parameters of the CDM.

The first equation directly gives the θ_v. We try to solve the second equation for κ. First, It is more convenient and easier to work with a statistic called mean resultant length instead of circular variance. The mean resultant length (represented by R) is a concentration parameter that equals one minus the circular variance (see section 2.3.1 in Mardia & Jupp (1999)). So the second equation becomes:

$$\frac{I_1(\kappa)}{I_0(\kappa)} = 1 - VCA = R.$$

An approximate solution for this equation is given in Banerjee et al. (2005) as:

$$\kappa_0 = \frac{R(2 - R^2)}{1 - R^2}.$$

This approximation is not accurate enough for our purpose here, so we use the Newton–Raphson iteration as suggested in Banerjee et al. (2005):

$$\kappa_1 = \kappa_0 - \frac{I_1(\kappa_0)/I_0(\kappa_0) - R}{1 - I_1^2(\kappa_0)/I_0^2(\kappa_0) - I_1(\kappa_0)/\kappa_0 I_0(\kappa_0)}.$$

To clarify the point of view here, it may be appropriate to briefly mention that the κ_0 is an approximate solution to the equation, deriving the concentration parameter from the mean resultant length. The solution is given in the Banerjee et al. (2005) and has a heuristic nature. This solution considers the fact that when R is zero and one, κ approaches zero and infinity, respectively. Also, the convergence rate is taken into account. κ_1 is one step Newton-Raphson iteration which moves the κ_0 one step towards the accurate solution of the equation.

One iteration is satisfactory for our use here as the ratio of the estimate to the exact solution of the equation is between 0.995 and $1 + 10^{-15}$ and we have checked that this magnitude of error has no considerable effect on the parameter recovery of the method. Now we rewrite the fourth equation with respect to v, by replacing the a with κ_1/v, and $I_1(av)/I_0(av)$ with R:

$$v^4 = \frac{1}{VRT}\left(\kappa_1^2 R^2 + 2\kappa_1 R - \kappa_1^2\right),$$

which will give v, and a could be computed as:

$$a = \kappa_1/v.$$

Then the t_0 can be calculated from the third equation:

$$t_0 = MRT - \frac{a}{v}R.$$

To summarize, the estimation starts with calculating statistics by Eq. 3 and doing the following calculations in the given order:

$$\begin{cases} \theta_v = MCA, \\ R = 1 - VCA, \\ \kappa_0 = \dfrac{R(2 - R^2)}{1 - R^2}, \\ \kappa_1 = \kappa_0 - \dfrac{I_1(\kappa_0)/I_0(\kappa_0) - R}{1 - I_1^2(\kappa_0)/I_0^2(\kappa_0) - I_1(\kappa_0)/\kappa_0 I_0(\kappa_0)}, \\ v = \sqrt[4]{\dfrac{1}{VRT}\left(\kappa_1^2 R^2 + 2\kappa_1 R - \kappa_1^2\right)}, \\ a = \kappa_1/v, \\ t_0 = MRT - \dfrac{a}{v}R. \end{cases} \tag{4}$$

The tutorial for implementing the method is available at https://github.com/HasanQD/EZ-CDM.

Parameter recovery

In this section, we will examine the accuracy of the proposed method in recovering the parameters that generate data. To do so, we choose a range of parameter values, simulate data from them, estimate the parameters from that data, and look at the degree of difference between the actual data-generating parameter values and the estimated values. Because of the stochastic nature of the model, it is impossible for this disparity to equal zero, as different parameter values can produce the same data. So, to have a reference for comparing the performance, the parameters are also recovered by the maximum likelihood estimation method.

A wide range for drift length and the decision criterion is considered that covers the previous estimates of the CDM on empirical data (Kvam, 2019b; Smith et al., 2020; Zhou et al., 2021) and also produces a reasonable joint choice angle and response time distributions. Variation for drift angle is not considered, as the bias and precision of the estimation will be identical for different values of drift angle in both the MLE and EZ-CDM methods. Changing the drift angle by $\Delta\theta_v$, will only rotate all choice data by this amount. By adding this quantity to θ_v in the likelihood function (Eq. (2)) the effect of this change will be canceled. Consequently, the maximum likelihood estimation for all parameters will remain unchanged, except for the drift angle, which will shift by $\Delta\theta_v$. Also, for the EZ-CDM, only the value of MCA will change by $\Delta\theta_v$ in Eq. 3 which will solely impact the calculation of the θ_v in Eq. 4, shifting it by the same amount of $\Delta\theta_v$. Similar phenomena occur with the non-decision time. The change in data-generating non-decision time value will just displace all the response time data by this amount of change, and adding this to t_0 in likelihood (Eq. (2)) will cancel the change. Also, this modification will solely impact the mean response time (MRT) in Eq. 3. This, in turn, will only affect t_0 in Eq. 4, causing it to shift proportionally to the alteration amount of data generating nondecision time.

A total of 10,000 parameter sets are produced by sampling from a uniform distribution with a range of [.5, 4.5] for drift length. Another independent sample is taken from the same distribution for the decision criterion. Assuming the value of zero for both the drift angle and non-decision time. Two data sets with 100 and 1000 trials have been simulated for each parameter set, representing the standard and maximum number of trials in experimental settings, respectively. The process of simulation involves executing a random walk approximation of the stochastic process in Eq. 1:

$$\begin{cases} \Delta X_1(t) = v\Delta t + \xi_1\sqrt{\Delta t}, \\ \Delta X_2(t) = \xi_2\sqrt{\Delta t}, \end{cases}$$

where subscripts indicate coordinated components and ξ is a sample of standard normal distribution. We used $\Delta t = 0.001$ seconds. The second equation lacks a drift component due to the assumption that the drift vector is aligned with the horizontal axis, resulting in a zero vertical component.

The random walk terminates upon intersecting the criterion circle, whereby the resulting angle and time of the event (non-decision time is zero) constitute the simulated trial.

The EZ-CDM is fitted to each data set by the procedure outlined in the preceding section. Maximum likelihood estimation is carried out using Eq. (2). One problem with this formula is that when $t - t_0$ becomes small, the exponential term decreases at a slow rate by increasing i. Consequently, the summands exhibit a slow convergence towards zero. In this situation, a precise calculation of the likelihood requires the computation of many terms. The severity of this problem increases as the decision criterion value increases. Another issue is that when drift magnitude and decision criterion are large, the exponential term in the first summands increases, which requires the calculations in the formula to be done with higher precision. We performed a computation of the summation up to 200 terms and forced the resulting values to zero when they were negative or when the last summand was nonzero. The calculations are carried out in double precision. We checked that in the range of parameters we use here, this calculation procedure addresses the problems described above (although problems start to emerge for a slightly bigger value of v and a). The Nelder–Mead optimization algorithm is used for finding the maximum of the likelihood function (actually the minimum of negative log-likelihood) (Nelder & Mead, 1965). For each data set, the algorithm is executed 20 times with randomized initial simplexes. Results show that in most of the cases, the algorithm converges to the same optimal point (on average, 16 out of 20), demonstrating the robustness of the optimization procedure.

The parameter recovery results are graphically displayed together with a numerical assessment of the RMSE root mean squared error (RMSE). The mean squared error is a measure of an estimator's accuracy. This indicator incorporates both the bias of estimation (the distance between the true value of the parameter and the average estimated value) and the variance (the spread of the estimated parameter value):

$$\begin{aligned}\mathrm{MSE} &= E\big[(\hat{\theta}-\theta)^2\big] = E\big[(\hat{\theta}-E[\hat{\theta}]+E[\hat{\theta}]-\theta)^2\big]\\ &= E\big[(\hat{\theta}-E[\hat{\theta}])^2\big] + 2E\big[\hat{\theta}-E[\hat{\theta}]\big](E[\hat{\theta}]-\theta) + (E[\hat{\theta}]-\theta)^2\\ &= \mathrm{Var}(\hat{\theta}) + \mathrm{Bias}^2(\hat{\theta}).\end{aligned}$$

In the above equation, θ is the data-generating parameter value, and $\hat{\theta}$ is the estimated parameter value. Note that

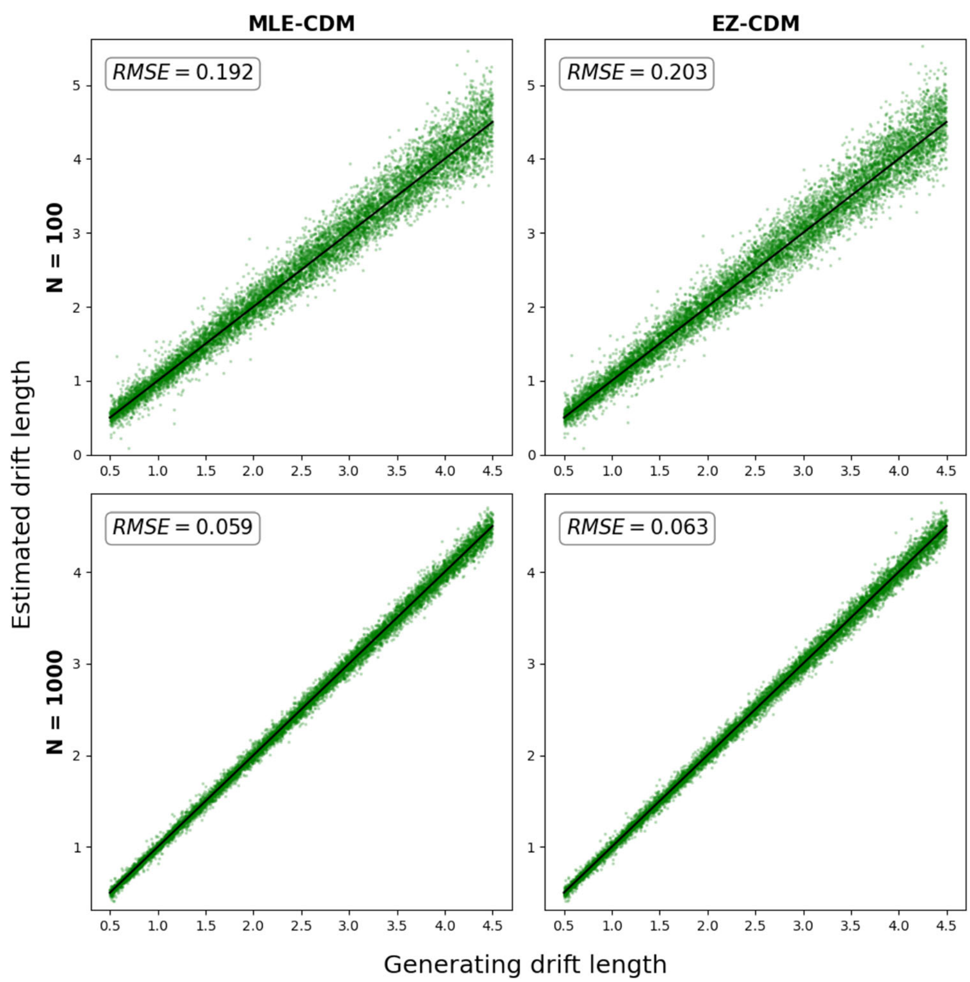


Fig. 2 Parameter recovery results for drift length. The plots in the first row are for data sets of 100 trials, and the second row is for 1000 trial data sets. Each *point* represents the recovered value against the data-generating values of the parameter for a single data set. The first column shows the results for the maximum likelihood estimation method, and the second column is for the EZ-CDM. The *black line* is the identity line $y = x$. The root mean squared error is represented in each plot

when there are varying generating parameters (drift length and decision criterion) the calculated MSE could be thought of as the average mean squared error over the whole range of generating parameter values (see Chapter 20 in Dekking et al. (2005) for an introduction to estimation efficiency and mean squared error). The square root of the mean squared error will be more illuminating, indicating the standard deviation of the estimated parameter values around the true data-generating parameter value.

Generating simulated datasets and calculating likelihoods are carried out using the NumPy 1.20.1 library (Harris et al., 2020). 'optimize.minimize' in the SciPy 1.7.1 library (Virtanen et al., 2020) is used for optimization. All calculations are implemented in Python 3.8.12 (vanRossum, 1995).

The results for parameter recovery of drift length for the MLE and EZ are plotted in Fig. 2. For the MLE, results show no considerable bias across the entire range. Variability in recovery is lower when there are more data points. However, variability increases with an increase in the generating value of v. The EZ is similar in performance as it has no evident bias and no great increase in variability compared to the MLE across the whole range and in two possible data set sizes.

Besides, Fig. 3 contains recovery results for drift angle. The recovered values for drift angle by the MLE and EZ are identical. This is because the only effect of drift angle is on choice angle data, for which we used the maximum likelihood estimation formula in the EZ-CDM. Apparently, estimations seem unbiased, and the variability is lower for larger dataset sizes.

Moreover, Fig. 4 illustrates the recovery results for the decision criterion. The EZ performs similarly to the MLE, both showing no considerable bias. An increase in variability is evident for larger values of the data-generating decision criterion and for smaller data set sizes.

The recovery results for non-decision time are given in Fig. 5. The MLE shows no bias and lower variability for the larger data set size. The same holds for the EZ, albeit with greater recovery variability relative to the MLE.

Another aspect of a reliable estimation method is its ability to remain unaffected by variations in other parameters while estimating a specific parameter. For instance, if two groups of subjects differ only in their drift lengths, the method should solely identify the difference in drift length and not in any

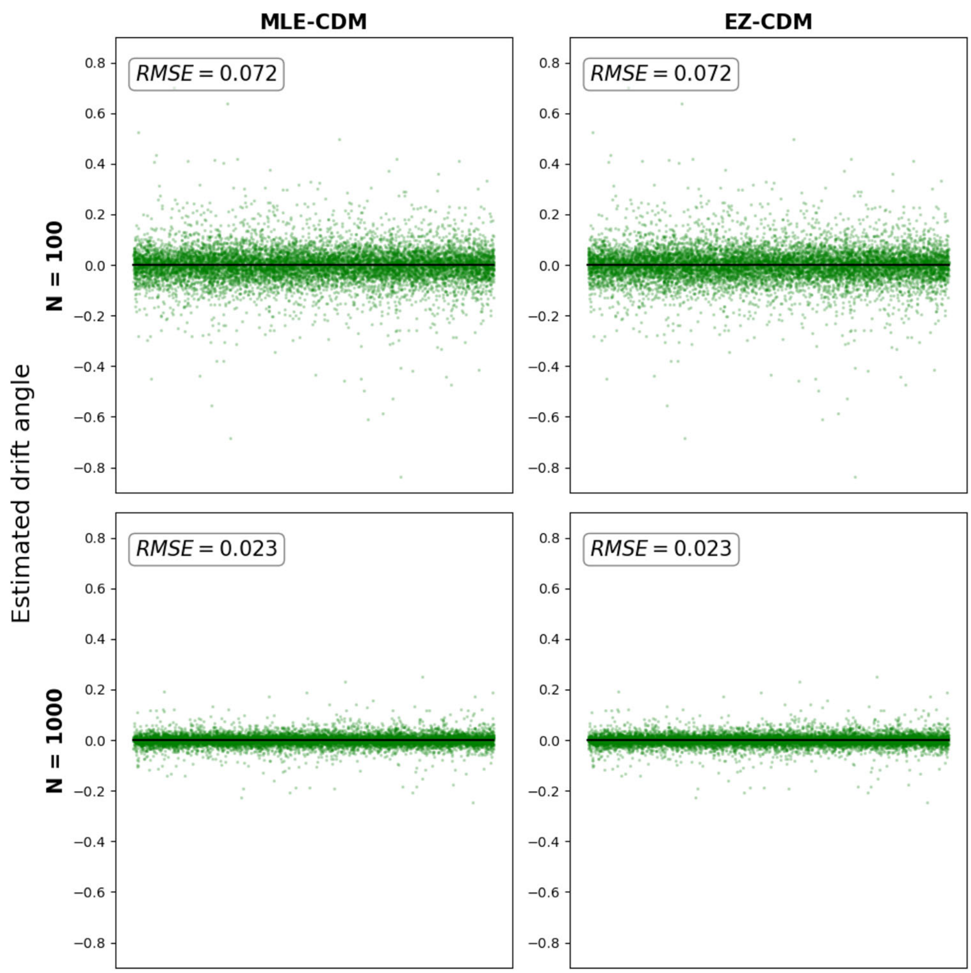


Fig. 3 Parameter recovery results for drift angle. The plots in the first row are for data sets of 100 trials, and the second row is for 1000 trial data sets. Each point's vertical coordinate represents the recovered value of the parameter for a single data set. The values are in radians. The points are jittered on the horizontal axis for better illustration, so the horizontal variation has no meaning. The first column shows the results for the maximum likelihood estimation method, and the second column is for the EZ-CDM. The *black line* is the line $y = 0$ representing the data-generating parameter value. The root mean squared error is represented in each plot

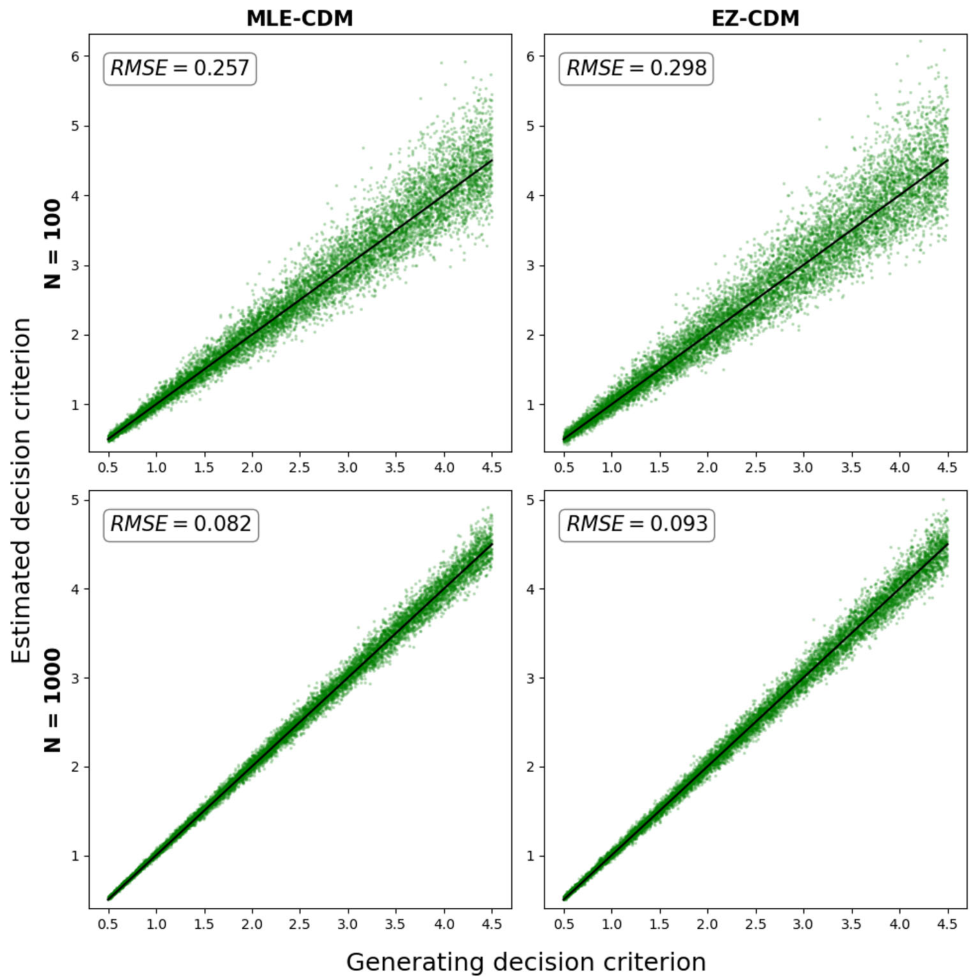


Fig. 4 Parameter recovery results for the decision criterion. The description of the plot is similar to Fig. 2

other parameters. To give a better understanding of the performance of the proposed approach, this issue is examined here.

As described earlier, the change in drift angle and non-decision time will only change the estimated values of the equivalent parameter. So, we just need to investigate the behavior of the EZ-CDM in recovering the parameters under varying drift lengths and decision criteria. Figures 6 and 7 show the recovered parameter values for different parameters with respect to variations of the drift length and decision criterion. The plots show no evident bias in recovery, which yields the desired property for the EZ-CDM. Additionally, variability in drift angle decreases with an increase in drift length and decision criterion. Also, variability in non-decision time decreases with an increase in drift length and a decrease in the decision criterion. These patterns are also present in maximum likelihood estimates. The increase in variability tells us that it will require more data to be able to detect differences in the parameter, but still, it is unlikely to incorrectly attribute changes in one parameter to a difference in other parameters as the recoveries show no bias.

Contaminant responses

A common issue encountered while dealing with experimental data is the presence of contamination from influential processes other than the modeled process. For instance, in decision-making tasks, factors such as fast guessing or a momentary lapse of attention can contaminate the data. Some of these contaminants can be detected and removed from the data by cutoffs because their response pattern is disjoint from the main uncontaminated data. However, the problem arises when the contaminants are intermixed with the rest of the data (Ratcliff & Tuerlinckx, 2002).

If the contaminated response time data significantly impacts the statistical measures utilized in the EZ-CDM (mean and variance), it may pose a challenge for the method. In this section, we analyze the potential impact of contamination on the EZ-CDM.

Following Ratcliff & Tuerlinckx (2002); Ratcliff (2008) on analyzing contaminants effects, we took the data sets in "Parameter recovery" section, and added a uniformly selected number from [0, 2] seconds to the response time for 5% of trials (five trials in 100 trial data sets and 50 in 1000 trial data

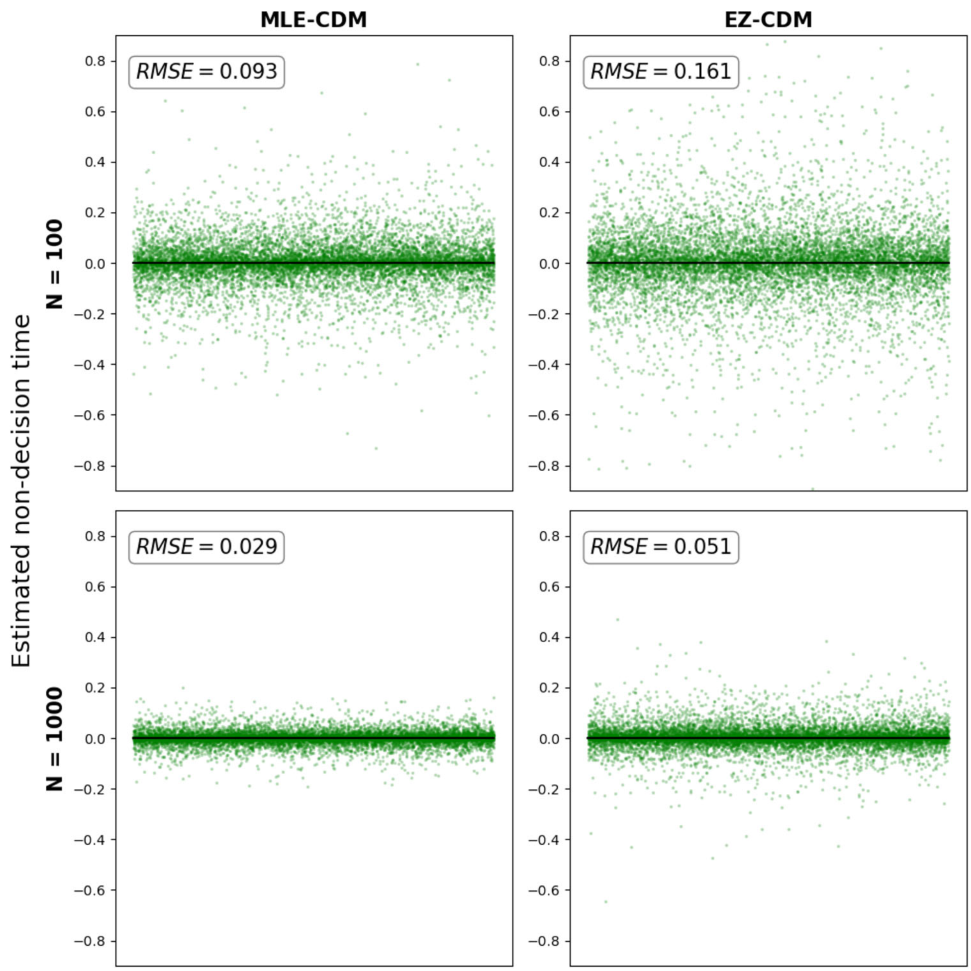


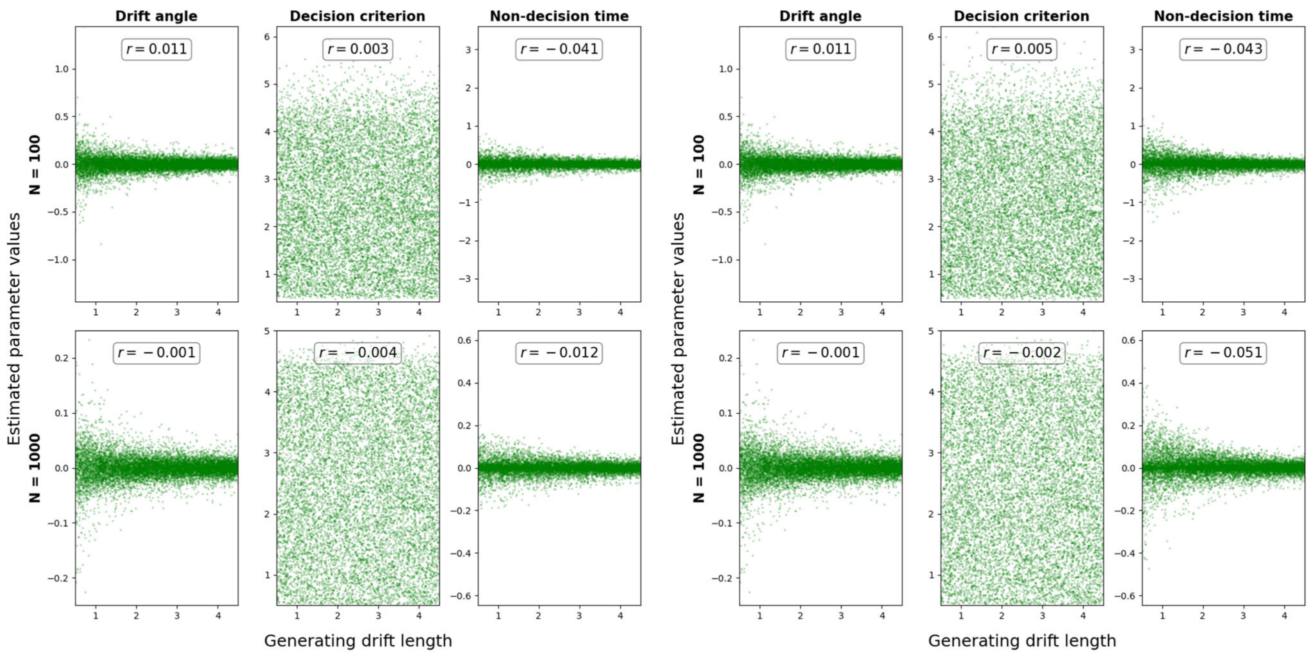
Fig. 5 Parameter recovery results for the non-decision time. The values are in seconds. The description of the plot is similar to Fig. 3

sets). The recovery of the EZ-CDM is compared to the recovery of the MLE for original uncontaminated data in Fig. 8a. Also, Fig. 8b shows the result for the more severe case of 10% contamination (ten trials in 100 trial data sets and 100 in 1000 trial data sets).

Upon analyzing the data, it is evident that the EZ-CDM exhibits a tendency to underestimate and display increased variability in the recovery of drift length, particularly in cases where the generating drift length is higher. Furthermore, we observe that the recovery of drift angle is identical to the MLE on uncontaminated data since the contamination only affected the response time, not the choice angle data, on which the estimation of drift angle is based. Furthermore, we observe an overestimation and increased variability in recovering decision criteria throughout the entire range. Also, it is apparent that the recovery of non-decision time exhibits both underestimation and increased variability. All problematic patterns are more severely evident when the ratio of contaminant data rises from 5 to 10%.

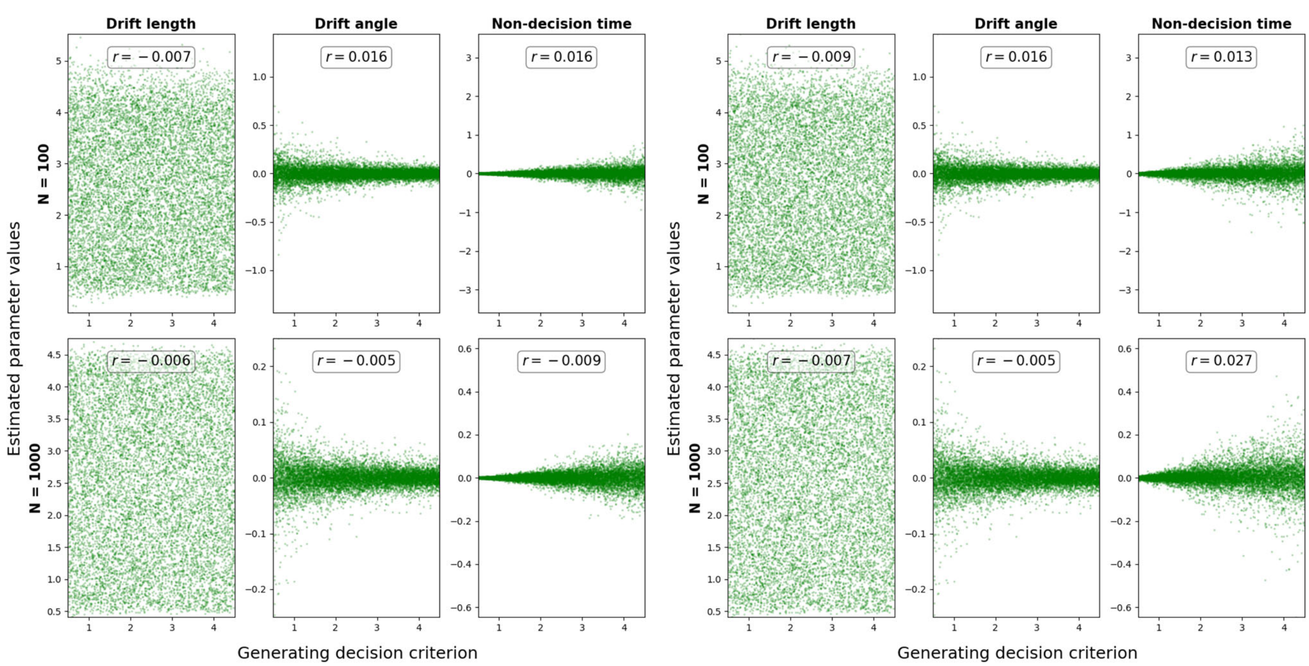
The origin of this poor performance can be traced to the sensitivity of the values of statistics in Eq. 3 to contaminants. The additive nature of contamination in response time data leads to an increase in both the mean and variance of response time. Referring to the estimation formulas for EZ-CDM in Eq. 4, the enlarged response time variance causes a reduction in the estimation of drift length according to the fifth equation. This, in turn, leads to an increase in the estimated value of the threshold by the sixth equation. These patterns decrease the estimation of non-decision indicated by the seventh equation, although this effect is alleviated by the increased mean response time.

The median and interquartile range are more robust options for measuring central tendency and dispersion. In order to incorporate these novel statistics into the EZ-CDM framework, we need to establish the relationship between the primary statistics and their substitutes. Assuming a normal distribution of response time, the mode and median coincide as identical points. Furthermore, in a normal distribution, the interquartile range equals $iqr \times \sigma$, where σ is the standard deviation of the distribution and iqr is the interquartile range of the standard normal distribution ($iqr \approx 1.349$). Despite the non-normality of the response time distribution, these transformation functions between the statistics could be approximately good. So we can replace the statistics of



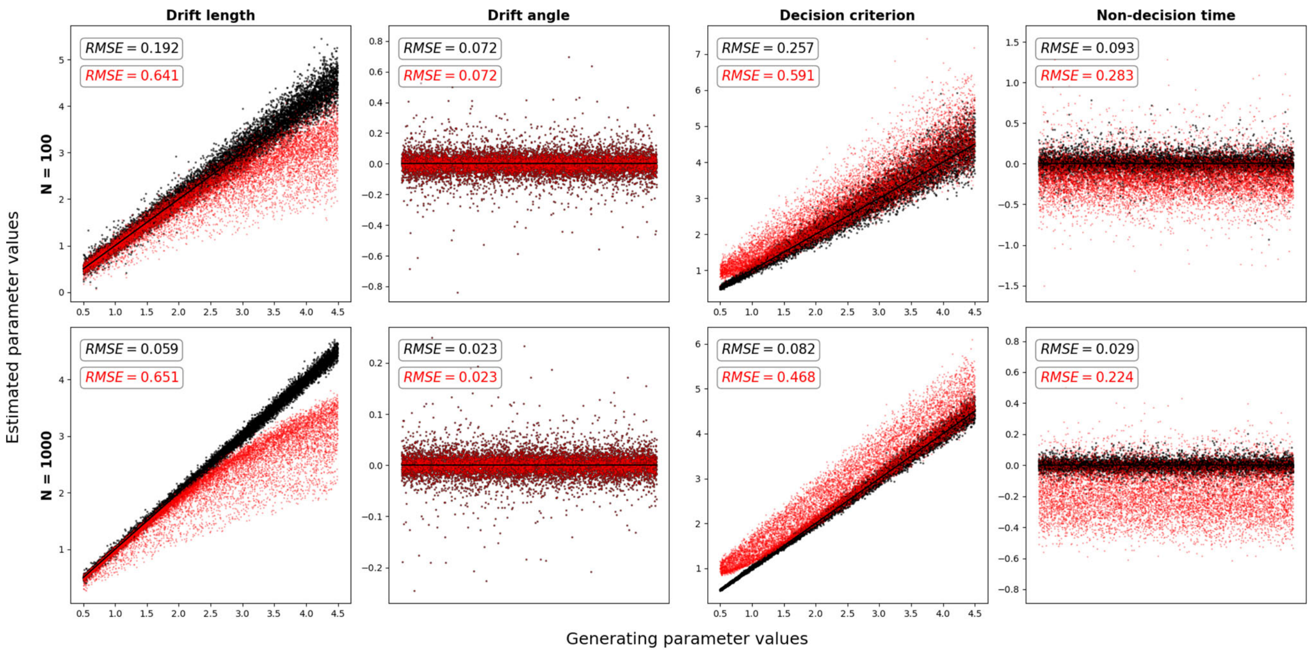
(a) MLE-CDM (b) EZ-CDM

Fig. 6 The pattern of estimated parameters for different values of drift length. The estimated values of drift angle, decision criterion, and non-decision time are plotted against the generating values of drift length. The vertical axis range is different for the two rows, but it is the same for the two plots. The correlation values between the estimated parameters and data-generating drift lengths are given in each plot

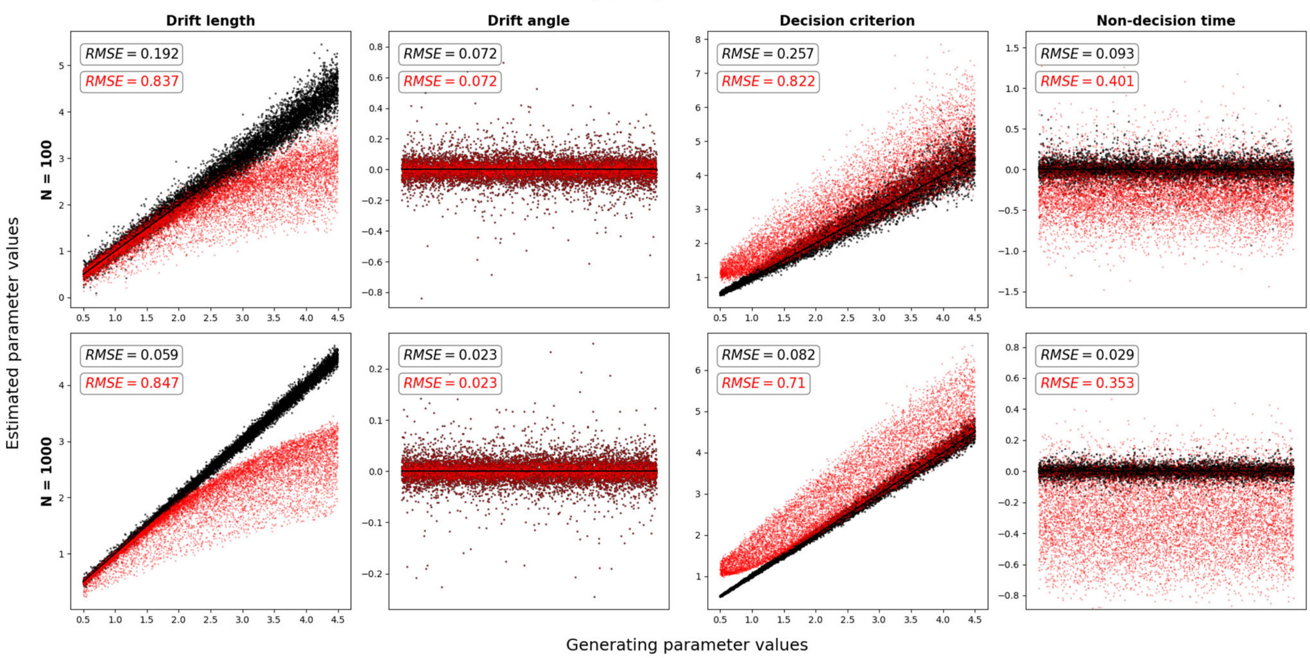


(a) MLE-CDM (b) EZ-CDM

Fig. 7 The pattern of estimated parameters for different values of decision criterion. The estimated values of drift length, drift angle, and non-decision time are plotted against the generating values of the decision criterion. The vertical axis range is different for the two rows, but it is the same for the two plots. The correlation values between the estimated parameters and data-generating decision criterion values are given in each plot

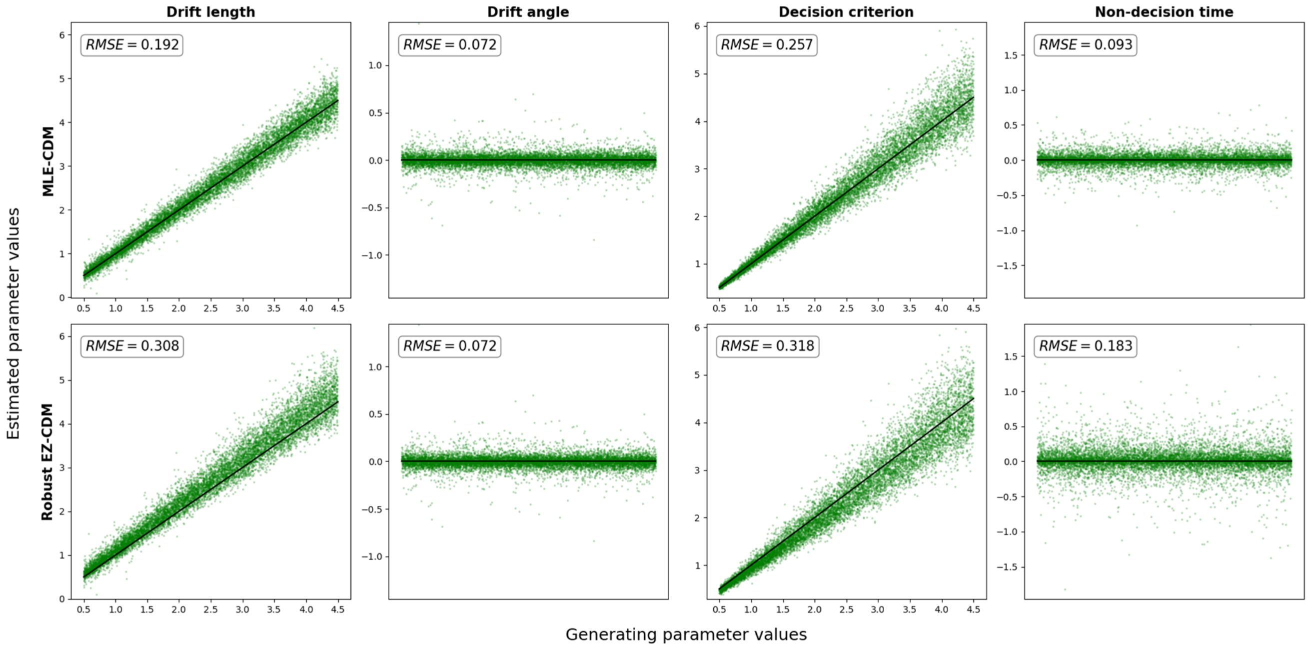


(a) 5 percent

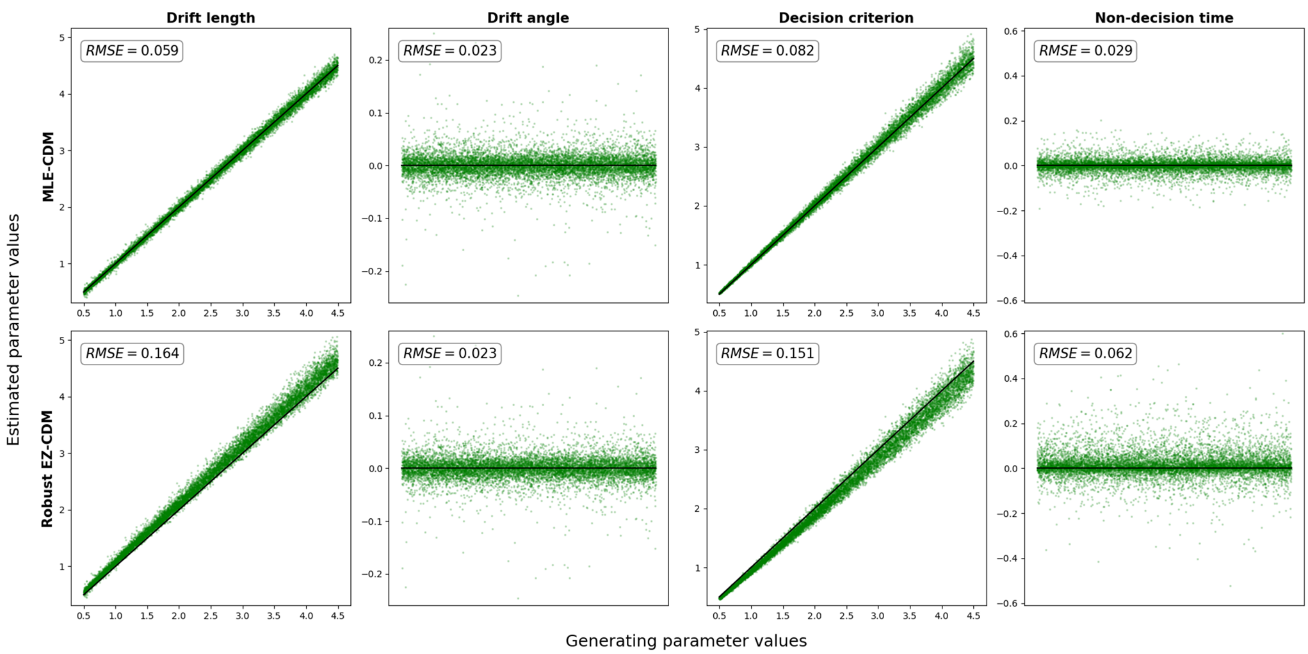


(b) 10 percent

Fig. 8 Performance of the EZ-CDM when confronted with contaminant data. The *black dots* are the maximum likelihood estimates on the uncontaminated data. The *red dots* are the estimates of the EZ-CDM on contaminated data. The root mean squared error for the MLE (*in black*) and EZ (*in red*) is reported in each plot



(a) $N = 100$



(b) $N = 1000$

Fig. 9 Parameter recovery results for the robust EZ-CDM compared to the MLE for two data set sizes

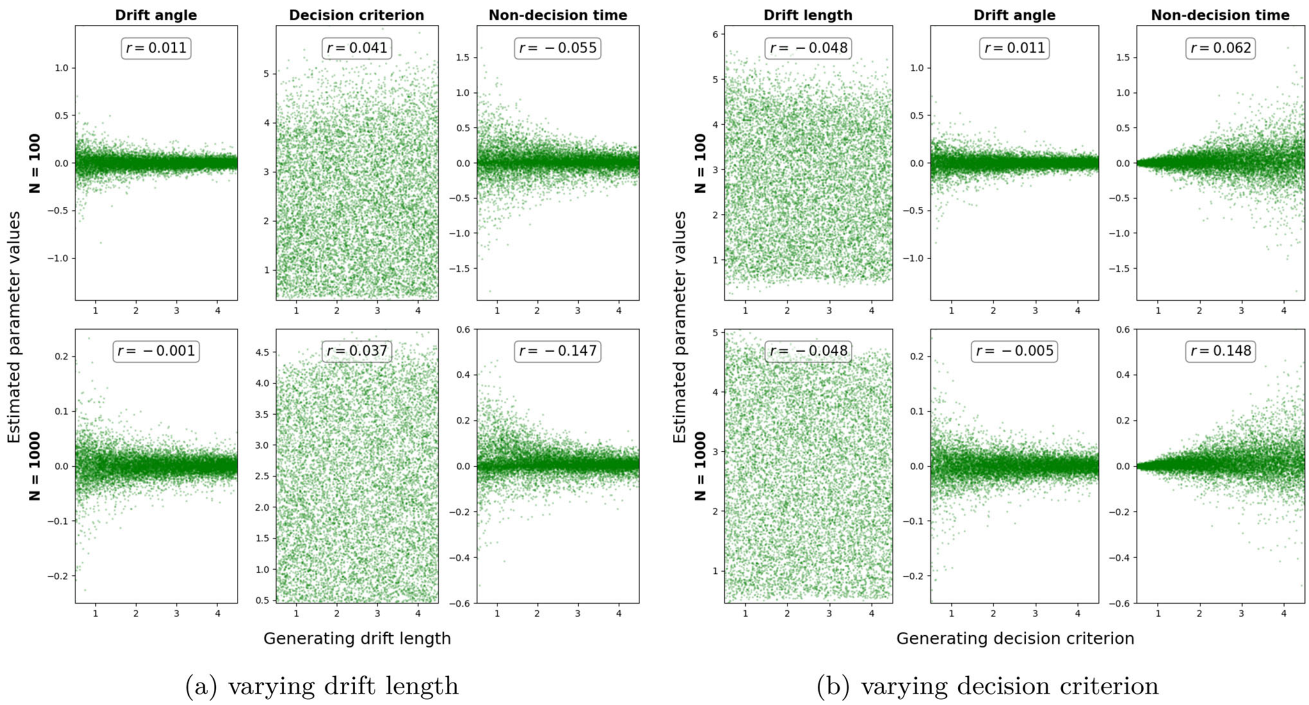


Fig. 10 The pattern of estimated parameters across varying drift lengths and decision criteria. **a** The estimated values of drift angle, decision criterion, and non-decision time are plotted against the generating values of drift length. **b** The estimated values of drift length, drift angle, and non-decision time are plotted against the generating values of the decision criterion

response time data in Eqs. 3 and 4, i.e., MRT and VRT, with the subsequent statistics:

$$\begin{cases} \text{median}\left[\{t_n\}_{n=1}^{N}\right], \\ \\ \left(\text{IQR}\left[\{t_n\}_{n=1}^{N}\right]/iqr\right)^2, \end{cases} \tag{5}$$

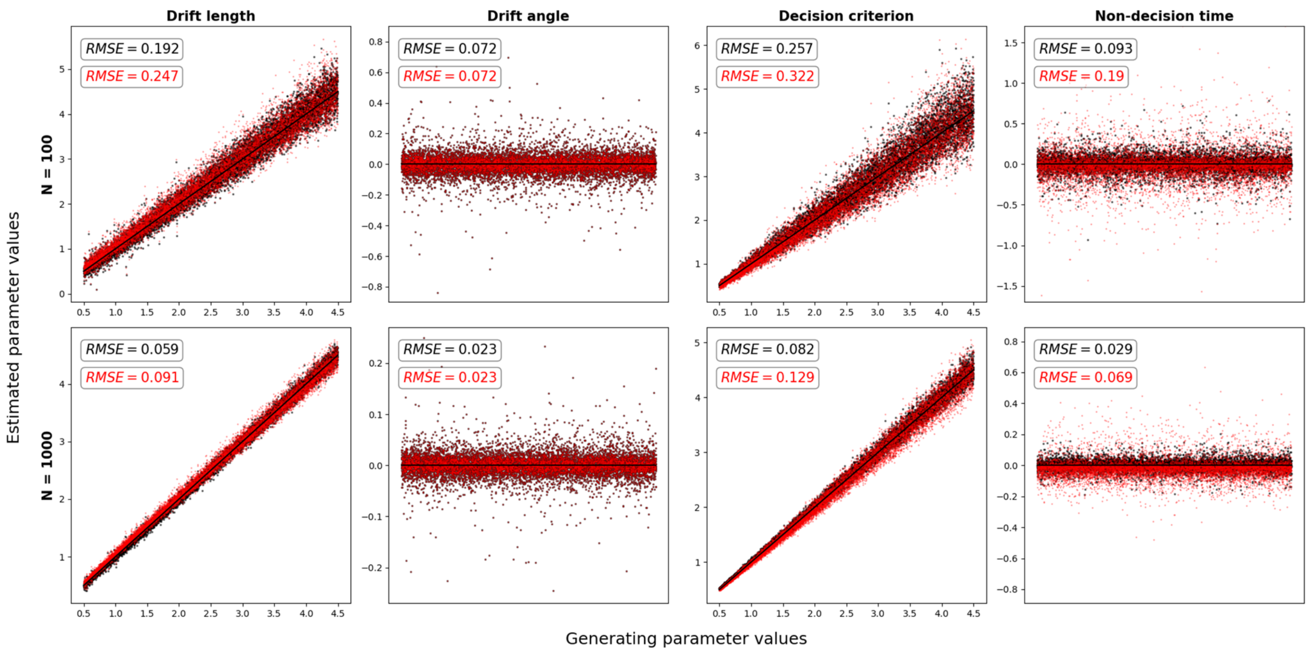
where IQR is the interquartile range.

Before analyzing the performance of this new method on contaminated data, it is imperative to conduct an assessment of its ability to recover parameters for uncontaminated data (analogous to “Parameter recovery” section). Figure 9 shows the recovery results of the robust EZ-CDM on datasets from “Parameter recovery” section.

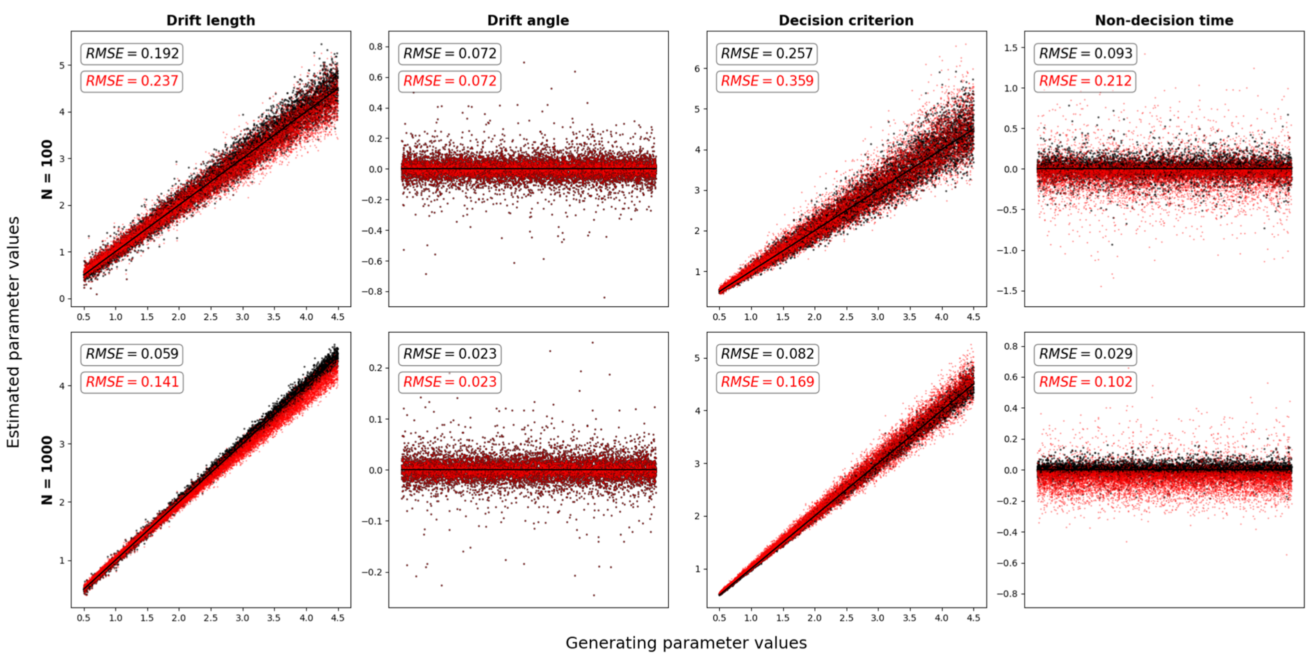
The robust EZ-CDM exhibited comparable performance to the MLE in recovery variability, with the exception of heightened variability in non-decision time and slight increases in variability of drift length and decision criterion. The estimates for drift angle are the same for the two methods since we only changed the statistics for response time and not the choice angle in the robust version of the EZ-CDM. In terms of bias, the method overestimates the drift length and underestimates the decision criterion. This means that the exact value of the parameters will be estimated by error, but more importantly, the differences will still be detected with nearly the accuracy of the MLE due to consistent bias across the entire range. The observed rise in the root mean squared error is expected to be primarily attributable to bias rather than a significant consequence of increased estimation variance.

The sensitivity of the robust EZ-CDM estimations to changes in the drift length and decision criterion is illustrated in Fig. 10. Note that the varying drift angle and non-decision time do not pose a problem. A change in the data-generating drift angle will only shift the estimated drift angle by the amount of the change. Similarly, modifying the data-generating non-decision time will only shift the median of the response time, and this will only shift the estimation of the non-decision time to the extent of the change in generating non-decision time. The findings indicate that it is unlikely to wrongly attribute the change in one parameter to other parameters, although a small correlation is detected between non-decision time with both drift length and decision criterion.

Now, we investigate the performance of the robust EZ-CDM on recovering parameters for the same contaminated data. Figure 11 represents the results. The observed performance is very satisfactory. This good performance can be attributed to the resilience of the response time statistics against contamination. The nature of the median and interquartile range is such that for a considerable change in them, a substantial portion of the dataset needs to be impacted by contamination.



(a) 5 percent



(b) 10 percent

Fig. 11 Performance of the robust EZ-CDM when confronted with contaminant data. *Black dots* are the maximum likelihood estimates on the uncontaminated data. *Red dots* are the estimates of the robust EZ-CDM on contaminated data. The root mean squared error for the MLE (*in black*) and EZ (*in red*) is reported in each plot

EZ-CDM performance in presence of across-trial variability

This section aims to evaluate the performance of the EZ-CDM in recovering the primary parameters of the CDM from data that is generated by the complex CDM model, which encompasses across-trial variability. We conduct a comparative analysis of its performance against the maximum likelihood estimation method for the true data-generating model. Two models are analyzed. One, only considering drift variability (with the advantage of having a closed-form likelihood formula). And the full model, in which all primary parameters are subject to variability.

The results are illustrated in plots showing recovered values against true data-generating values of parameters. Two indices, namely bias, and correlation, are utilized to report quantitative performance. Bias is calculated as the average deviation between estimated and data-generating parameter values. The Spearman correlation between estimated and data-generating values of parameters is reported as an indication of how much the estimation method can preserve order between the parameter values. A methodology exhibiting a strong correlation would be more successful in detecting systematic differences in parameter values between groups and conditions.

The likelihood of the model with a normally distributed drift vector is as follows Smith & Corbett (2019):

$$Pr(\theta,t) = \frac{1}{\sqrt{[\eta_1^2(t-t_0)+1][\eta_2^2(t-t_0)+1]}} \exp\Big(-\frac{v^2}{2\eta_1^2} + \frac{[a\eta_1^2\cos(\theta-\theta_v)+v]^2}{2\eta_1^2[\eta_1^2(t-t_0)+1]} + \frac{[a\eta_2^2\sin(\theta-\theta_v)]^2}{2\eta_2^2[\eta_2^2(t-t_0)+1]}\Big) \times \frac{1}{2\pi a^2}\sum_{i=1}^{\infty}\frac{j_{0,i}}{J_1(j_{0,i})}\exp\Big(-\frac{j_{0,i}^2}{2a^2}(t-t_0)\Big). \tag{6}$$

The drift vector is assumed to come from a bivariate normal distribution, where the mean vector has length v and angle θ_v. Also, its covariance matrix has two eigenvectors with angles θ_v and $\theta_v+\pi/2$ and eigenvalues η_1^2 and η_2^2 respectively. This means that drift variability has two independent components. One is in alignment with the mean vector, and the other is perpendicular to it. This is why η_1 and η_2 are named the radial and tangential components of drift variability (Smith, 2019).

Variable drift

We investigated the performance over a range of trial numbers, including 50, 100, 150, 200, 300, 500, and 1000. At each of these levels, 500 datasets were generated. The parameters for each dataset were selected from the following distributions:

$$\begin{aligned}
&a \sim U[0.5, 3],\\
&v \sim U[0, 3],\\
&\theta_v = 0,\\
&\eta_1 \sim v \times U[0, 0.7],\\
&\eta_2 \sim v \times U[0, 0.7],\\
&t_0 = 0,
\end{aligned}$$

where U denotes a uniform distribution, which takes the endpoints of the range as its inputs. Drift angle and non-decision time are taken to be constants, as the discussion in "Parameter recovery" section remains applicable in this context. The drift variability is taken to be a factor of the drift length as it results in more reasonable simulated data.

The simple and robust versions of EZ-CDM were fitted to each dataset according to the procedure outlined in "EZ-CDM" and "Contaminant responses" sections. For maximum likelihood estimation, we used Eq. (6) but as described in "Parameter recovery" section, when t is close to t_0, the accurate determination of the likelihood value requires a high-precision computation of numerous terms in the infinite series. Recently, Smith et al. (2023) proposed a method to address this issue. The method involves utilizing an approximate formula for small values of $t - t_0$, while adopting a truncated series with low precision calculation for large values of $t - t_0$. However, for the range of parameter values searched by the optimization algorithm in our work, this approach seems not viable. For example, when $a = 5$, $v = 5$, $\eta_1 = 1.3 \times v$ for some values of t, we think that both of the formulas fail to give accurate values of likelihood. It is important to note that this issue only occurs within the specified range of parameter values used here. It is possible that the parameter space is more constrained and the likelihood approximation performs well in experimental settings.

In order to tackle this matter, we calculated the series up to 2000 terms whenever needed, with a precision of 50 decimal points. Due to the amount of computation involved, the aforementioned calculations were performed only once, and the resultant values were stored for subsequent use. The series was computed for paired values of a and $t - t_0$ that were equally spaced within a range with step sizes of $da = 0.0005$ and $dt = 0.0005$. The mpmath module in Python was utilized for performing calculations with high precision (mpmath development team, 2023). Implementations for calculating the likelihood function with this method are available in https://github.com/HasanQD/CDM.

To optimize the likelihood function, we performed 1000 Nelder–Mead (Nelder & Mead, 1965) runs with random starting points, and we chose the best one. When the differential evolution (Storn & Price, 1997) was run 1000 times, the outcomes were almost identical.

The bias of recoveries is represented in Fig. 12. In general, as the size of the data set increases, the maximum likelihood method's bias decreases. The method's unbiasedness for larger datasets is evident in the convergence towards zero. However, for some parameters, the convergence appears to be slightly deviating from zero, which might be a sign of a small issue with the MLE implementation. The EZ methods do not experience this reduction in bias with increasing the trial number. Also, convergence appears to be to a non-zero value, which is expected as the data-generating model is different than the method's expected model. The EZ methods underestimate the values for drift length, whereas MLE overestimates them. The robust EZ-CDM (REZ) performs better or is comparable to the MLE in terms of bias magnitude for trial numbers up to 200. The performances on the drift angle are similar among the methods. For the decision criterion, MLE overestimates, and the EZ methods underestimate the values. At least up to 200 trial numbers, the EZ outperforms or matches up to the MLE. The MLE performs best for non-decision time across the whole range, although the performance of REZ is also notable, tending to overestimate about 20 ms.

The recoveries' variability is seen in Fig. 13. Overall, as the size of the data set rises, so does the performance of each method. While the performance of the EZ methods is not poor across the whole range, the drift length performance of the REZ is comparable to the MLE only for datasets under 100 trial numbers. The drift angle performance is similar among the methods. When it comes to decision criteria, the EZ outperforms the MLE up to a dataset of 100 trial numbers, but the MLE surpasses for larger datasets, even if the performance of the EZ methods is still quite good. The MLE wins in non-decision time, while REZ's performance is also noticeable. The recovery of the variability parameters is not very good, even for large data sets.

As an indication of type I statistical error when the difference is on another parameter, a correlation analysis between the data producing values of one parameter and the estimated values for another parameter is also performed. Significant correlations have been found, mostly with the EZ approaches; nevertheless, further research is needed to determine the full impact on statistical analysis. These results are available in the supplementary material.

The estimated parameter values against the true data-generating values are depicted in Fig. 14 for the 100 trial number case. The distribution of the estimated values in the range of true parameter values is better represented by this figure. For all methods, the bias and variability in the recoveries rise with higher values of drift length and decision criterion. This pattern is preserved for the other set sizes. The supplemental material contains the analogous graphs for the other set sizes.

Full-model

Here, we considered variability in all of the main parameters of the CDM. Similar to the preceding subsection, the drift vector is postulated to be derived from a bivariate normal distribution. The decision criterion is considered to come from a uniform distribution with a mean a and a range r_a. The non-decision time variability follows the uniform distribution with a mean t_0 and a range r_t.

A total of 250 datasets, for each trial number in 50, 100, 200, 300, 500, and 1000, were simulated using randomly selected parameter values from the following distributions:

$$
\begin{aligned}
&a \sim U[0.5, 3],\\
&r_a \sim a \times U[0, 1],\\
&v \sim U[0, 3],\\
&\theta_v = 0,\\
&\eta_1 \sim v \times U[0, 0.7],\\
&\eta_2 \sim v \times U[0, 0.7],\\
&t_0 = 0,\\
&r_t \sim U[0, 0.25].
\end{aligned}
$$

The variability parameters controlling the dispersion of the drift and decision criterion are considered to be factors of their respective mean values of the distribution.

For calculating the likelihood, we used numerical integration on a and t_0 in Eq. 6. The rectangle rule was utilized by calculating the values on the points for which we had the stored values of the series ($da = 0.0005$ and $dt = 0.0005$) for the case of 100 trial numbers. Using the Simpson rule did not improve the performance of the MLE. Also, Increasing the step size resulted in a slight degradation of the outcomes, indicating that the integration was performed with nearly precise accuracy. However, we are not sure if the numerical integration is performed with the highest possible precision.

The maximum of likelihood function was determined through the use of the differential evolution optimization method (Storn & Price, 1997). We used the implementation of this method in the SciPy 1.9.3 library with its default values for parameters of differential evolution (Virtanen et al., 2020). The reliability of the optimization method is evidenced by the superior optimal values obtained compared to the optimal likelihood values achieved by running the Nelder–Mead method with an initial simplex containing the true parameter values as one of its nodes. In lower trial numbers, most of the datasets had better optimums by differential evolution than the informed Nelder–Mead.

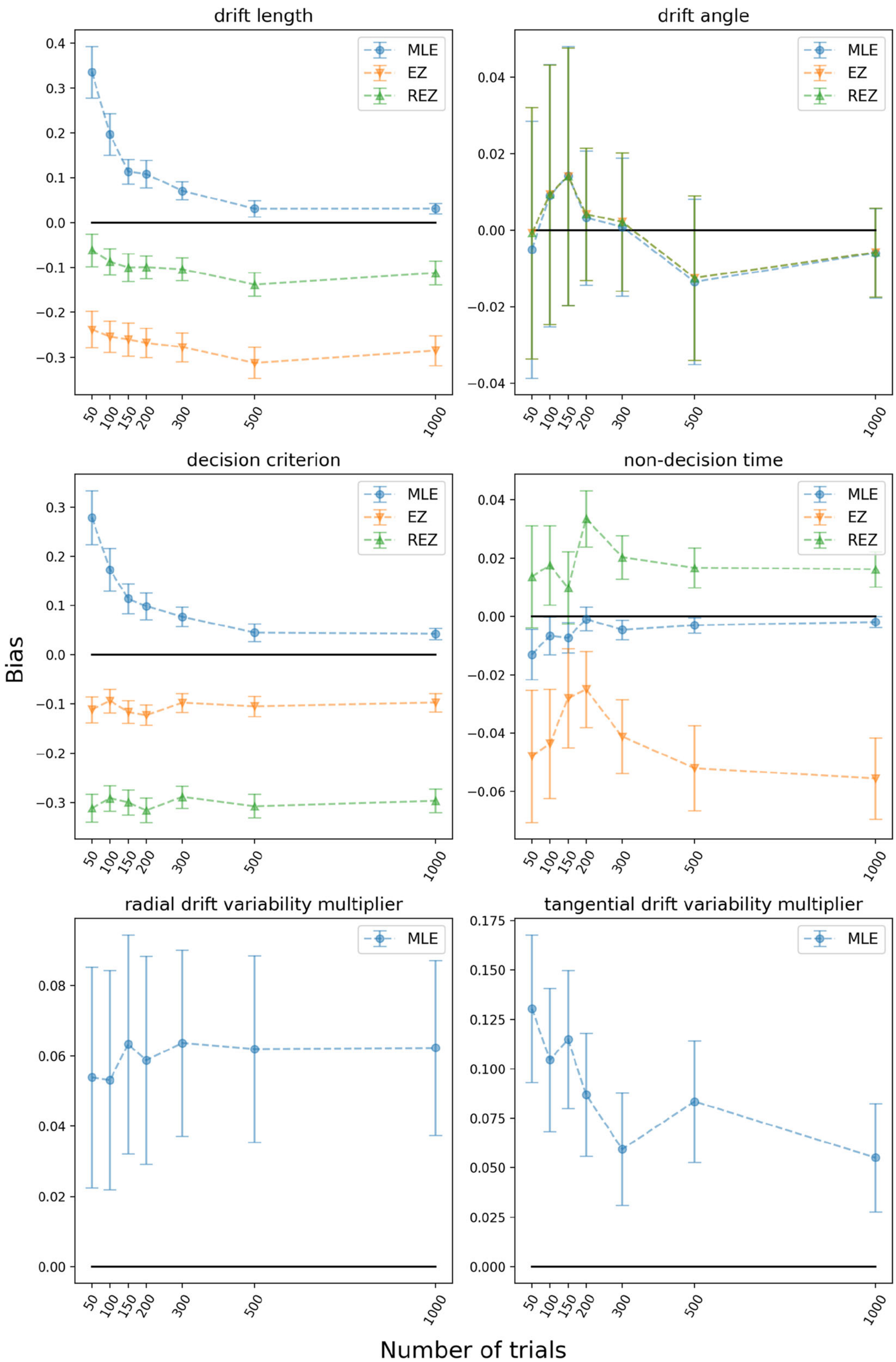


Fig. 12 Biases of estimation methods across the number of trials for all parameters of the CDM with drift variability. Each *point* shows the mean of the difference between the estimated and true parameter values. The *error bars* represent the 95% confidence interval evaluated by considering the normal distribution with the standard error as the standard deviation. This will be very close to the confidence interval derived from the *t*-distribution because of the large sample size (see Test 2, Section VI, Part 2, in Sheskin (2003))

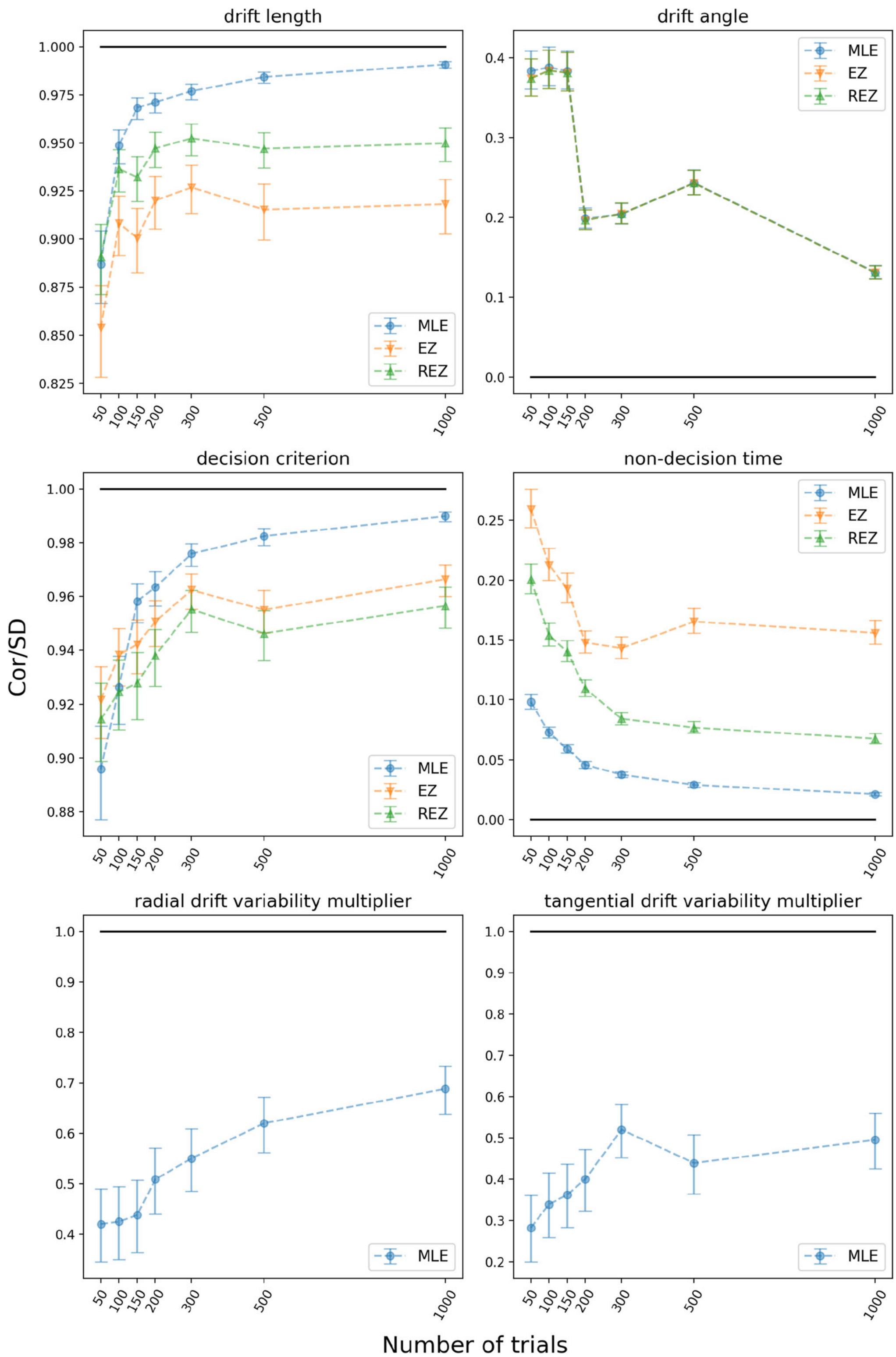


Fig. 13 Variabilities of estimation methods across the number of trials for all parameters of the CDM with drift variability. For the drift angle and non-decision time where the true parameter is fixed, the standard deviation is reported, and for the remaining parameters, Spearman's correlation coefficient – an indication of maintaining the order of the true parameters – is reported. *Error bars* are 95% confidence intervals. Confidence intervals for the standard deviation are calculated using the chi-squared distribution (see Test 3, Section VI, Part 2, in Sheskin (2003)). Also, confidence intervals of correlations are calculated using the Fisher transform (see Test 29, Section VI, Part 5, and Test 28, Section VI, Part 4, in Sheskin (2003))

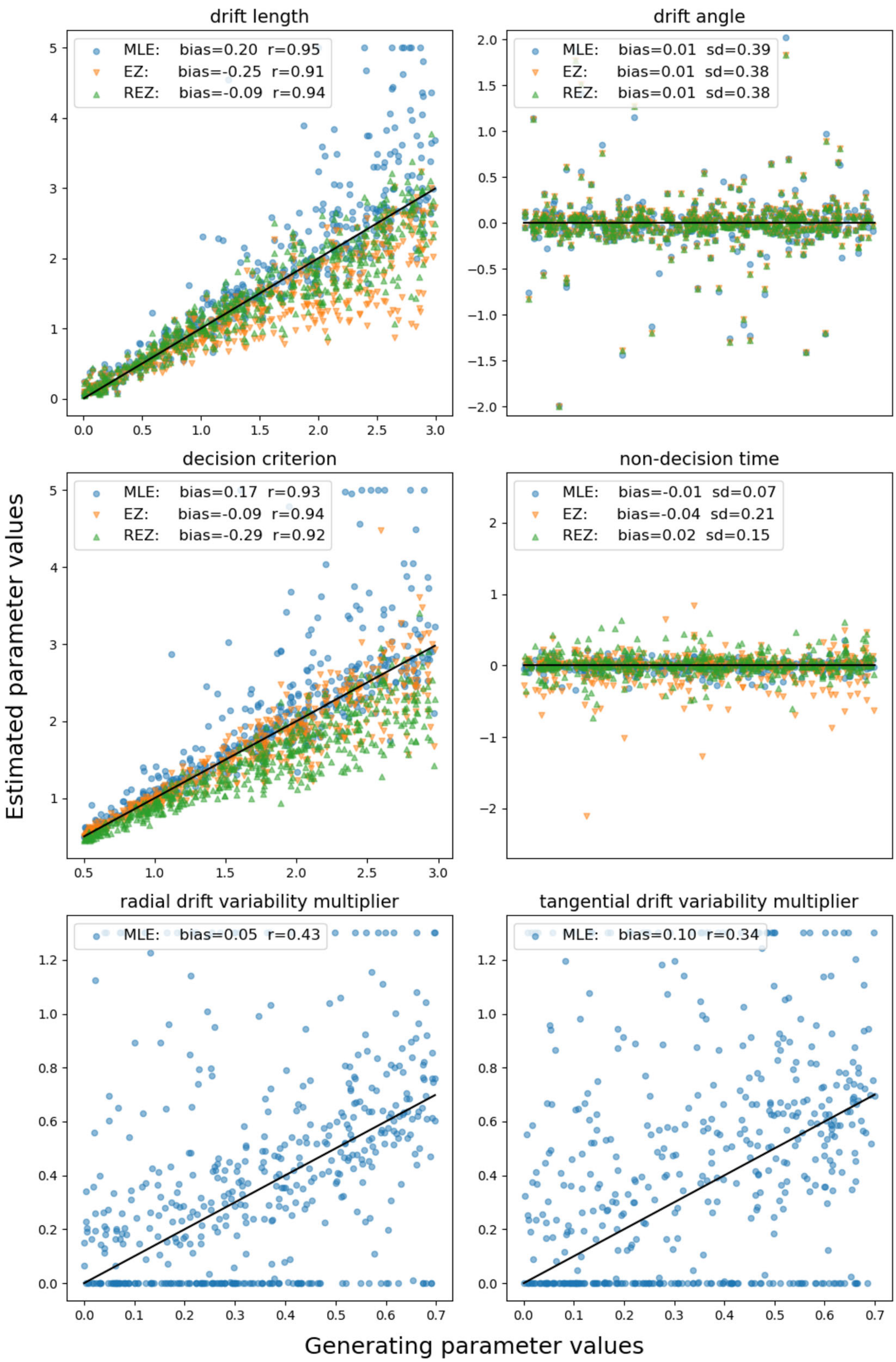


Fig. 14 Parameter recovery results where datasets are generated by CDM with drift variability in the case of 100 trial number. The performances of three methods of maximum likelihood estimation (MLE), simple EZ-CDM (EZ), and robust EZ-CDM (REZ) are reported. Each *point* represents the estimated value of the parameter against the true data-generating value for a single dataset. The bias, Spearman correlation coefficient, or standard deviation values are displayed in each plot. The *black line* in each plot indicates perfect recovery

However, for high trial number, differential evolution has an advantage for around half of the datasets (although the relative performance of three estimation methods was intact for both halves.). However, it remains possible that the outcomes may be enhanced through refinement of the optimization methodology. Specifically, note that we did not check the repeated optimization of Nelder–Mead and differential evolution due to high computational and time demands.

In summary, the setup of the maximum likelihood estimation may exhibit slight deviations from its optimal implementation, but we have an upper bound on the performance of the MLE. We do not expect that the recovery of the MLE in the full-CDM will be better than its recovery in the nested model, which has only variability in drift. So we can consider this in the comparison of performance between EZ and MLE.

Figure 15 represents the recoveries' bias. Generally, the bias of the maximum likelihood method reduces with increasing data set size. This reduction does not happen for EZ methods. The MLE overestimates the drift length value, whereas the EZ approaches underestimate it. For dataset sizes under 500, the REZ outperforms the MLE in terms of drift length bias magnitude. The performances on the drift angle are similar among the methods. For the decision criterion, MLE overestimates, and the EZ methods underestimate the values. The EZ performs superiorly to the MLE for trial numbers under 500. For non-decision time, the performance of REZ is comparable to MLE.

The variability of recoveries is represented in Fig. 16. Overall, all three methods benefit in performance from increasing the trial number. For drift length, REZ performs superiorly to MLE for small set sizes, but MLE retains superiority above 500 trial numbers. Yet the performance is not bad for EZ methods. The performance on the drift angle is similar between methods. When it comes to decision criteria, the EZ performs similarly to the MLE for datasets smaller than 500, but the MLE outperforms the EZ for larger datasets, even if the EZ methods' performance is still decent. The MLE performs best for non-decision time across the whole range of trial numbers, although the performance of REZ is also notable. The recoveries of variability parameters are not satisfactory even for 1000 trials.

To determine the mistake of methods for incorrectly associating a difference in one parameter with another, we examined the correlation between true and estimated values of non-identical parameters. While the EZ methods occasionally demonstrated significant correlations, most correlations were not significant for the important situation of drift length and the decision criterion. However, more investigation is required to fully understand the influence on statistical analysis. The supplemental material contains these results.

The estimated parameter values against the true data-generating values are depicted in Fig. 17 for the 100 trial number case. The distribution of the estimated values in the range of true parameter values is better represented by this figure. For all methods, the bias and variability in the recoveries rise with higher values of drift length and decision criterion. This pattern is preserved for the other set sizes. The supplemental material contains the analogous graphs for the other set sizes.

Experimental data containing guesses

Some important applications of circular experimental tasks are in situations where the paradigms are designed in such a way that a considerable amount of data is the result of guessing alongside the main cognitive process (Zhang & Luck, 2008; Green & Pratte, 2022). For example, in the realm of visual working memory, it has been posited that such guess data arises in instances where the stimulus is not retained in memory, providing insight into the limits of working memory capacity. In such scenarios, it is critical to calculate the ratio of data originating from guessing. It would be advantageous to estimate the CDM parameters pertaining to the primary cognitive process in addition to the guess proportion. In what follows, we aim to provide a comprehensive approach to accomplishing this task.

The uniform distribution of choices over angles is a clear indicator of guess data. This property enables the capture of the guess proportion by fitting a mixed distribution of uniform and a modal distribution (often considered von Mises distribution) over choices.

To use the CDM in this situation, it is necessary to identify the guess data and exclude them so that only the data relevant to the main decision process takes part in the CDM fitting. However, the problem is that guess trials are intermixed with main trials, especially in regions of high accuracy. In such a situation, the CDM can be extended to include guessing. However, the difficulty lies in the fact that, to the best of our knowledge, there is no coherent and accepted model for guessing in continuous report tasks that takes into account response time. So we want to estimate the CDM parameter values and guess proportion, without incorporating any specific model for the response time of guesses.

Here, we will try to achieve the above-mentioned goal by introducing a method that is compatible with the EZ-CDM. We illustrate the application of this method on the dataset in Smith et al. (2020) (the data for the condition with high chromatic noise is used as it contains a heavy-tailed choice angle distribution). The experimental data pertains to a circular color task conducted on a sample of four individuals (further information regarding the experiment can be found in Smith et al. (2020)). It should be mentioned that in the

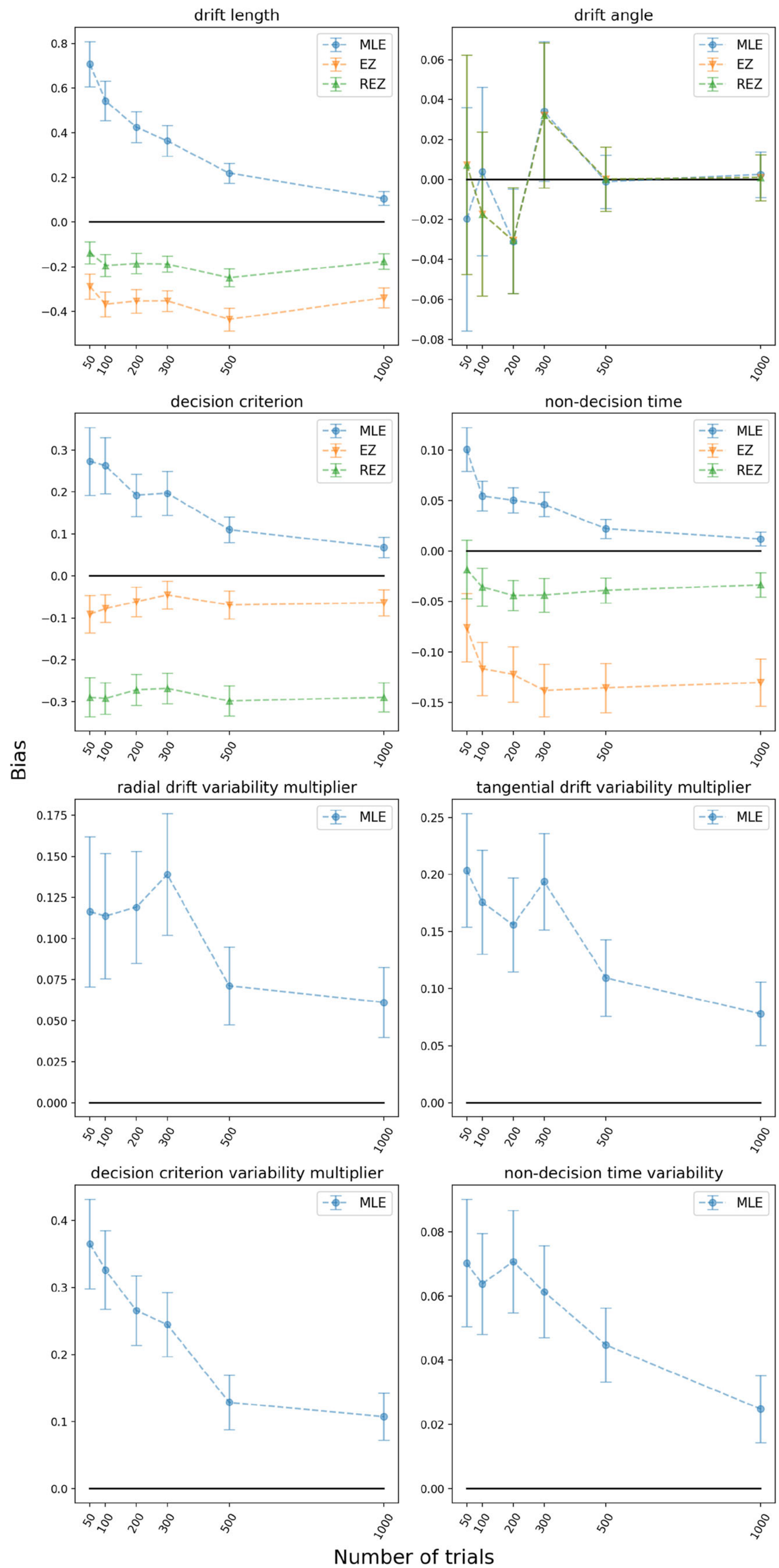


Fig. 15 Biases of estimation methods across the number of trials for all parameters of the CDM with full variability. The description of the plot is similar to Fig. 12

Fig. 16 Variabilities of estimation methods across the number of trials for all parameters of the CDM with full variability. The description of the plot is similar to Fig. 13

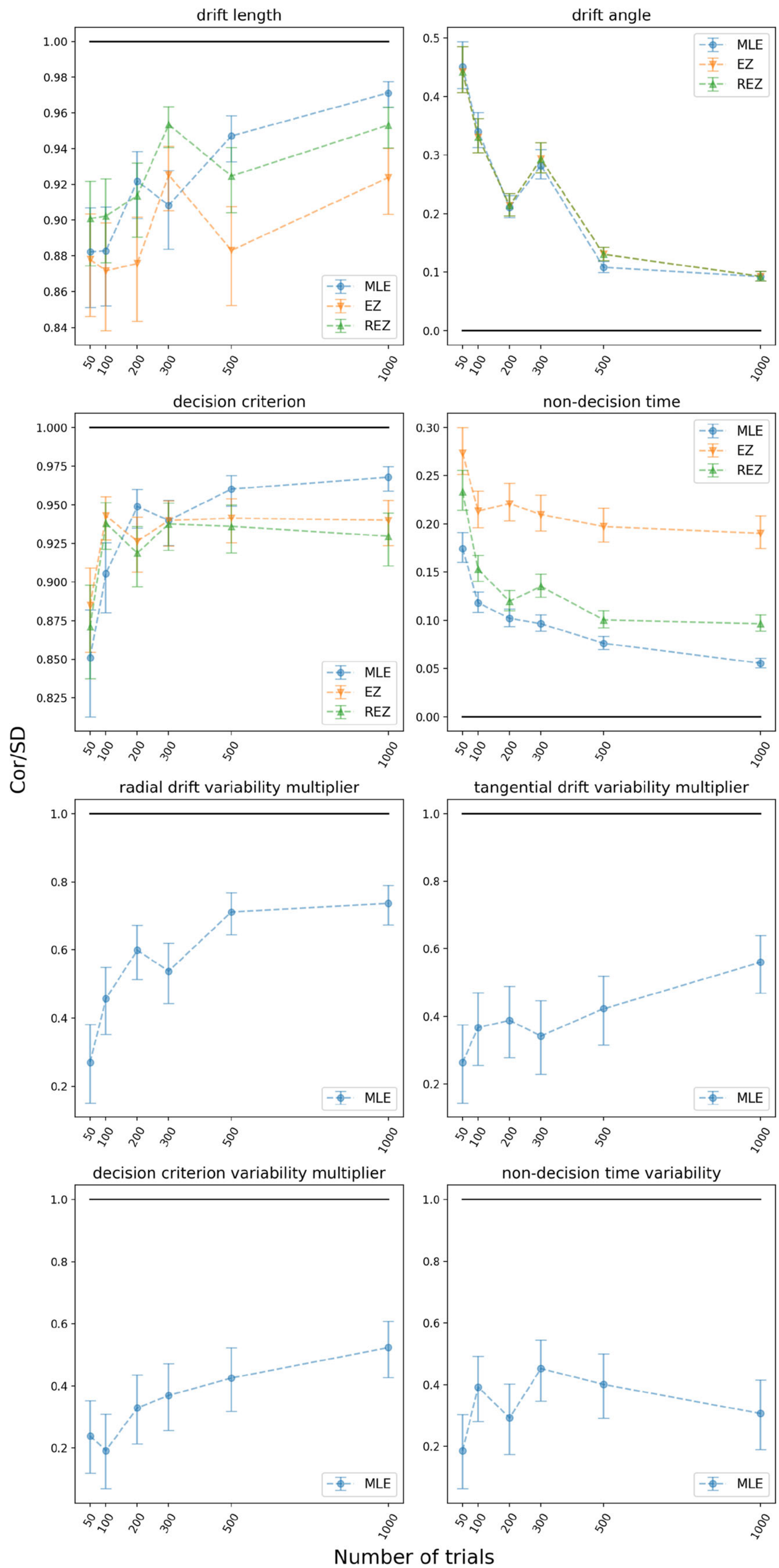
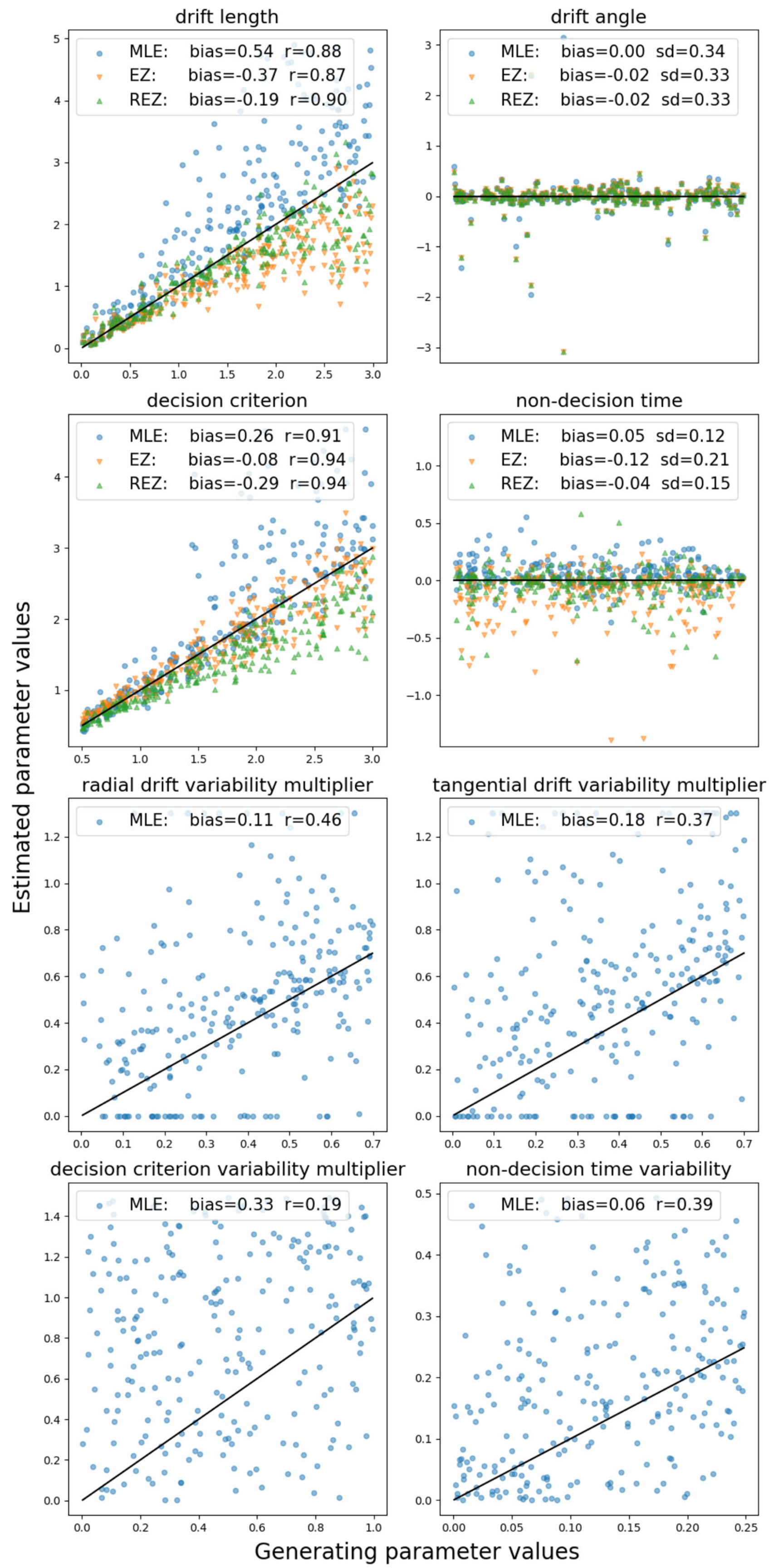


Fig. 17 Parameter recovery results where datasets are generated by full-CDM. The description of the plot is similar to Fig. 14



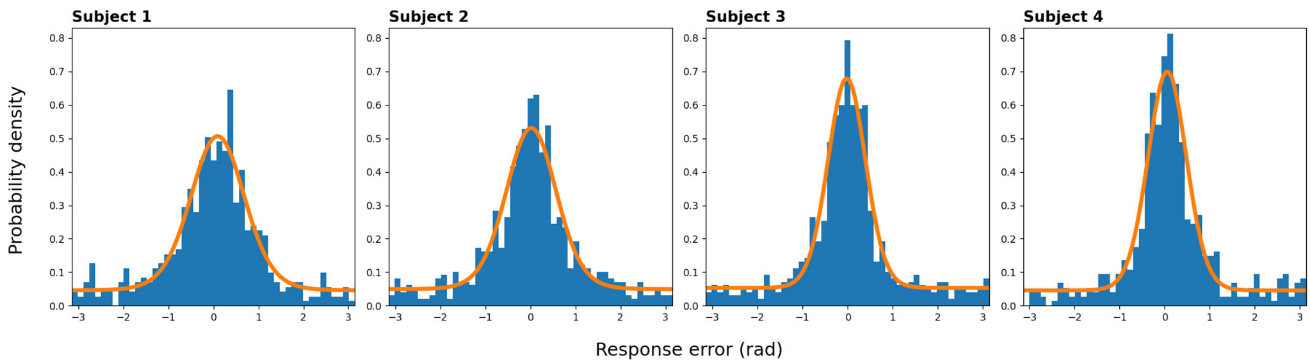


Fig. 18 Histogram of choice angle data and fits of the mixed model for the four participants

original paper, the data was considered to come from two sources: the primary decision-making process and the other process, which they called 'encoding failure'. The second process also exhibits a uniform distribution of response outcomes, but they are not the result of guessing. We keep calling the origin of these data a guess, but we should keep in mind that the referred work does not view the phenomenon this way. Here, the utilization of this data serves solely as an example to illustrate our methodology, rather than theoretical derivations pertaining to the original work.

Given the lack of information regarding response time, we seek to infer the proportion of guesses by analyzing the data on choice angles. As frequently used, a mixed model comprising uniform and von Mises distributions can be fitted to the given data to determine the guess proportion and the von Mises distribution parameters. The maximum likelihood estimation method can be employed to fit the mixed model. Worthy of being explained here, the likelihood of the mixed model is:

$$Pr(\theta) = \frac{p}{2\pi} + (1-p)\frac{e^{\kappa\cos(\theta-\mu)}}{2\pi I_0(\kappa)}, \tag{7}$$

where p is the probability of guessing, μ is the angular mean, and κ is the concentration parameter. The Nelder–Mead algorithm is utilized to determine the optimal parameter values that maximize the likelihood of the observed data. The fits of the mixed model with these parameters are given in Fig. 18. Also, Table 1 contains the estimated parameter values of the mixed model. The fits are reasonable, and parameter values are consistent with differences in the response patterns of subjects. The data for subject 3 shows a little heavier tails compared to other subjects and has a larger value of the estimated guess proportion (p). There is not much evident angular bias in the data for subjects, and the low values of μ are consistent with this. Also, the data for subjects 3 and 4 seems more compact and concentrated, and the values of κ agree with this pattern.

Now that we have μ and κ, we only need the statistics of the response time data to be able to use EZ-CDM. We aim to derive the mean and variance of the response time data related to the primary decision process. Let T_1 be this distribution with a respective mean and variance of MRT_1 and VRT_1. Additionally, consider MRT_2 and VRT_2 as the mean and variance of the guess response time distribution T_2. If the two distributions are mixed, with the probability of the guess being q, then the mean and variance of the mixed distribution T will be:

$$\begin{cases} E(T) = \int t P(T=t) \\ \quad = \int t\big[(1-q)P(T_1=t) + qP(T_2=t)\big] \\ \quad = (1-q)MRT_1 + qMRT_2, \\ V(T) = E[T^2] - E[T]^2 \\ \quad = \int t^2 P(T=t) - E[T]^2 \\ \quad = \int t^2\big[(1-q)P(T_1=t) + qP(T_2=t)\big] - E[T]^2 \\ \quad = (1-q)(VRT_1 + MRT_1^2) + q(VRT_2 + MRT_2^2) \\ \quad - [(1-q)MRT_1 + qMRT_2]^2 \\ \quad = (1-q)VRT_1 + qVRT_2 + q(1-q)(MRT_1 - MRT_2)^2. \end{cases}$$

By partitioning the response time data into two segments, each characterized by its mean E and variance V and asso-

Table 1 Parameter values of the mixed model fitted on choice angle data of four subjects

	p	μ	κ
Subject 1	0.281	0.095	2.902
Subject 2	0.308	0.018	3.345
Subject 3	0.334	-0.017	5.835
Subject 4	0.287	0.057	5.555

ciated guess proportion q, we could solve the resulting equations for the means and variances of the primary decision and guess response time as follows (superscripts indicate each of the two segments):

$$\begin{cases} MRT_1 = \dfrac{q^{(2)}E^{(1)} - q^{(1)}E^{(2)}}{q^{(2)} - q^{(1)}}, \\ MRT_2 = \dfrac{(1-q^{(2)})E^{(1)} - (1-q^{(1)})E^{(2)}}{q^{(1)} - q^{(2)}}, \\ VRT_1 = \dfrac{q^{(2)}V^{(1)} - q^{(1)}V^{(2)}}{q^{(2)} - q^{(1)}} - q^{(1)}q^{(2)}(MRT_1 - MRT_2)^2, \\ VRT_2 = \dfrac{(1-q^{(2)})V^{(1)} - (1-q^{(1)})V^{(2)}}{q^{(1)} - q^{(2)}} - (1-q^{(1)})(1-q^{(2)})(MRT_1 - MRT_2)^2. \end{cases} \tag{8}$$

In order to achieve a good separation of statistics for two processes, it is better to partition the data into sections that exhibit the maximal difference in guess proportion. We split the experimental data into two parts of high and low accuracy that contain an equal number of trials based on choice angle data (look at Fig. 19). This is done by finding the range centered around the circular mean that contains half of the data. The length of this range is twice the median of the absolute values of the choice angle data, subtracted by the drift angle. Given the range of length r and a total guess proportion of p, the proportion of guesses within this range can be determined as pr/π. This is due to the fact that the probability of a guess falling within this region is $pr/2\pi$, and the total probability of the entire region is $1/2$. A similar argument will demonstrate that the guess proportion in the low accuracy region is $p(2\pi - r)/\pi$. By substituting the given values as q into the aforementioned equations and incorporating the means and variances of response time data for each region, the resulting mean and variance of the primary decision process response time distribution can be obtained.

Following the formulation in "EZ-CDM" section, we can use the available values of μ, κ, MRT_1 and VRT_1 to derive the parameter values of CDM:

$$\begin{cases} \theta_v = \mu, \\ R = \dfrac{I_1(\kappa)}{I_0(\kappa)}, \\ v = \sqrt[4]{\dfrac{1}{VRT}\left(\kappa_1^2 R^2 + 2\kappa_1 R - \kappa_1^2\right)}, \\ a = \kappa_1/v, \\ t_0 = MRT - \dfrac{a}{v}R. \end{cases} \tag{9}$$

The resulting values of parameters are represented in Table 2.

We cannot provide a plot for fitting results on the response time data as we have no model for the response time of the guess. However, the method is based on the exact prediction

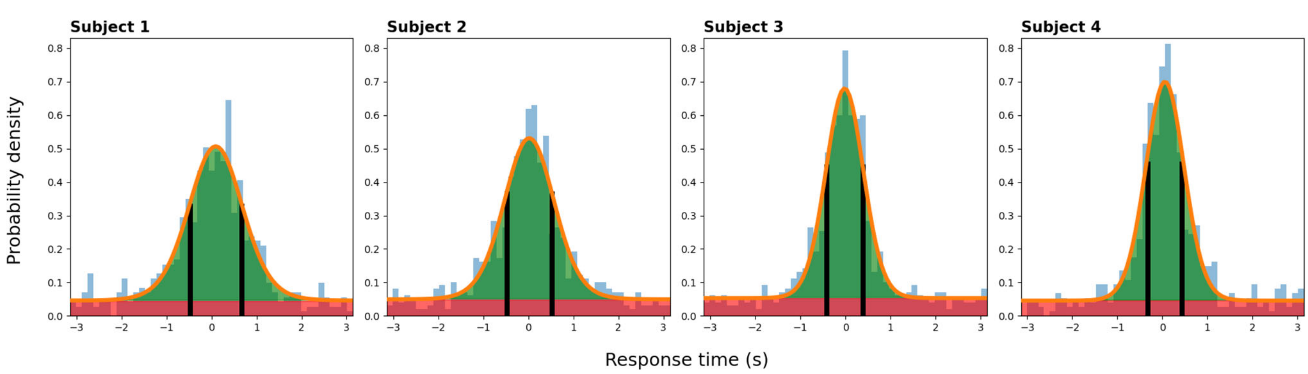


Fig. 19 The response region is split into two categories of high and low accuracy. The region between two vertical black lines contains half of the data points considered as high accuracy. The distribution of response in each region is an intermix of guessing (in red) and the primary decision process (in green)

Table 2 Parameter values of the CDM for 4 subjects

	v	θ_v	a	t_0
Subject 1	1.53	0.095	1.90	0.43
Subject 2	1.99	0.018	1.68	0.67
Subject 3	2.57	-0.017	2.27	0.29
Subject 4	2.01	0.057	2.76	0.58

of the mean and variance of response time for both parts and, consequently, for the entire dataset.

As a final point, note that we calculated the mean and variance of two high and low accuracy regions on response time data with a cutoff of 4 s to alleviate the effect of outliers. Also, we could employ the robust statistics of mean and interquartile range instead of the mean and variances and use Eq. (5) to convert them to appropriate values to insert into Eq. (9) to derive CDM parameter values.

Discussion

In the introduction, a discussion revolved around the advantages and indispensability of employing evidence accumulation models in order to achieve a more precise delineation of the fundamental psychological mechanism in continuous report paradigms. The reader may possess additional motives. Many of these motivations are those that make researchers extensively use the evidence accumulation models for two alternative decisions. Despite the motivational drive, the application of these models to paradigms of continuous report has been limited to a handful of works. We think that a major reason for this could be the hardship of the implementation of these models, which entails fitting the model onto the empirical data and approximating the interpretable parameters. Unlike models for binary choice models, dependable and user-friendly techniques and software tools for parameter estimation have yet to be devised for continuous decision models.

To alleviate this hardship, we proposed a straightforward procedure for estimating the parameters of the circular diffusion model and showed that, besides being simple, it exhibits comparable performance to maximum likelihood estimation in terms of estimation accuracy. The only exception was the non-decision time, where the variability in the EZ-CDM was higher than the MLE. So if the non-decision time is important for the research purpose and the estimates of the EZ-CDM do not give significant information, there is a possibility that using other methods of estimation will be helpful. The results demonstrate that the EZ-CDM is also able to detect differences in conditions and groups.

The small deficiency of the EZ-CDM compared to the MLE can be attributed to two sources. One is that the statistics we used for response time data do not contain sufficient information about the whole dataset. The other source is that we fitted the choice angle data with maximum likelihood, but the application of maximum likelihood to both choice angle and response time data may not exhibit identical performance as compared to the utilization of maximum likelihood solely on choice angle data. This means that the EZ-CDM exhibits a slightly improved fit with the choice angle data in comparison to the MLE, albeit at the cost of a slight misfit in the response time data.

We showed that the robust EZ-CDM remains relatively unaffected by a particular category of contaminant response time. However, based on the nature of the statistics used in the approach, it is expected that the method will exhibit robustness against other forms of contamination. Also, there is no need to exclude extreme data points when using the robust EZ-CDM. But still, the presence of some kind of contaminants in choice angle data may pose a challenge for the EZ-CDM. Some techniques have been devised to estimate the parameters of the von Mises distribution with enhanced robustness (Kato & Eguchi, 2016). These techniques can serve as a substitute for the maximum likelihood method employed in the initial phase of the EZ-CDM if required.

The key to the improved performance of the robust EZ-CDM is the utilization of statistics with more resilience against contamination in the method of moments. This insight has potential implications in the field of model-based cognitive science, specifically in the development of a robust approach to fitting models to data. For instance, this could be incorporated into the EZ-DDM methodology to enhance the accuracy of estimates for the parameters of the diffusion decision model. The integration of this technique into the EZ-DDM can be achieved by using Eq. (5). A potential benefit of this approach for EZ-DDM is the resolution of its issue with estimating non-decision times that are not within the natural range, as it produces negative non-decision times in various instances.

One point worth mentioning is that, in practice, the MLE is not suitable for use on experimental data as it is extremely sensitive to contaminant and outlier data. Alternative modifications to the approach have been implemented, like using the MLE on partitions of the response space instead of single responses (Ratcliff, 2018), assigning a minimal value for the probability of data points that have a very small probability (Smith et al., 2020), or using the MLE on mixed models with contaminants (Ratcliff & Tuerlinckx, 2002). These alternative methods are expected to have poorer performance compared to the MLE. The real rivals of the simple and robust versions of the EZ-CDM are these alternative and lower-performance methods, but we compared the simple EZ-CDM with the MLE and the robust EZ-CDM on contaminated data with the MLE on uncontaminated data.

A limitation encountered by the EZ-CDM is its reliance on the assumption that there is no across-trial variability in the data-generating model. Nevertheless, there is compelling theoretical and empirical evidence contradicting this assumption (Smith, 2016, 2019; Smith et al., 2020; Zhou et al., 2021). The impact of this assumption on the performance of EZ-CDM in parameter value measurement was analyzed in the presence of across-trial variability on the main parameters of the model generating the data. Our findings demonstrate that this assumption actually works as an advantage for EZ-CDM for data sets with a small to medium number of trials. In such cases, the robust EZ performed superior to or comparable to MLE in recovering the drift length in terms of bias magnitude and variability. For the decision criterion, the EZ was better. The only exception was the non-decision time, where the MLE outperformed. The high values of Spearman correlations indicated that the methods are suitable for detecting the order of parameter values among individuals, groups, and conditions.

The MLE's poor performance in small data set size can be attributed to the model's inherent complexity. The recovery of the variability parameters exhibited undesirable performance, which may have consequential impacts on the recovery of other primary parameters. Utilizing a parsimonious model and disregarding parameters with low recoverability may result in enhanced performance in the retrieval of primary parameter values. Particularly in situations where there is uncertainty regarding the precise specifications of the exact details of the data-generating model.

Incorporating variability into the model raises the issue of selecting an appropriate distribution. The inherent variability could follow other distributions than those considered in the model. In these contexts, it is necessary to assess the impact of the model's misspecification (Ratcliff, 2013). So the methods based on the models that consider across-trial variability should deal with this problem. However, the EZ-CDM approach offers a potential solution by disregarding variability.

In addition to the issue of misspecification, it is necessary to verify the robustness of the methods against contamination. The adjusted versions of the maximum likelihood methods, while robust, are expected to have lower performance than the MLE on uncontaminated, unmisspecified data. The robust EZ-CDM is expected to preserve its performance to a considerable extent when faced with these challenges. The case of misspecification and contaminants was not analyzed, as the EZ methods yielded promising results in comparison to the MLE already. However, alternative approaches require validation to address these issues before they can be selected over EZ techniques for parameter estimation. For example, in the case of non-decision time, the EZ performance was inferior to the MLE. Nevertheless, it may be a viable option among estimators who exhibit resilience towards misrepresentation and impurities.

The MLE procedure was attempted to be implemented with optimal precision. Our approach involved the precise calculation of probability and the utilization of a reliable optimization procedure. But there is still potential for improvement in the MLE results; however, this will require more time, processing capacity, and using more sophisticated methods. Our implementation of the MLE fitting procedure relies on the technique we used, which utilizes both memory and computational power. The technique involves the calculation and storage of the computationally demanding portion of the likelihood function. This allows for retrieval from memory as needed, rather than repeated calculations. Lucky for us, this computationally demanding portion only depends on the a and $t - t_0$, so the stored file is compact and manageable. The feasibility of employing this methodology can be examined in other instances of fitting stochastic models.

We also proposed a method for fitting the mixed model of the CDM and guess. This model does not incorporate any assumptions regarding the duration of time taken to provide a guess response. In the statistical context, models of this nature are referred to as "semiparametric". The estimates of the parameters by the method are expected to converge to the data-generating values of the parameters with enough data, no matter what the guess distribution is. The method uses the coupling information of the choice angle and response time. This coupling is important and contains the necessary information for separating guesses from main decision process data, but the method only uses the coupling associated with two regions of choice angle (high and low accuracy) and not the coupling of single data points. There is potential for improving the technique by employing coupling at the individual data point level. However, the application of the current method for experimental analysis, particularly in the realm of visual working memory, holds potential promise.

Conclusion

We proposed a simple method for estimating the parameters of the circular diffusion model. The procedure involves the computation of particular data statistics (namely, the circular mean and variance of the choice angle as well as the mean and variance of the response time) through the utilization of Eq. 3. Subsequently, these computed values are employed in equations to make estimations of the parameter values via

Eq. (4). Furthermore, a robust version of the method was presented, which incorporates statistics in Eq. 5 and exhibits a lower susceptibility to contaminants.

Our endeavor in this article involved establishing trust in the utilization of these techniques for measuring the CDM parameters in practical scenarios. Extensive analysis has been conducted on the performance of these methods in situations where the data contains contaminants or is generated by the model incorporating across-trial variability. The performance of the methods exhibited a high level of satisfaction, which fulfilled the promise of the title by introducing a method that is fast, simple, robust, and accurate.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-024-02483-7.

Acknowledgements The authors express gratitude towards P.L. Smith and co-authors in Smith et al. (2020) for providing open-access data, which was utilized in this study. We also thank P.L. Smith for carefully reading the manuscript and providing his insightful remarks. Also, HQ thanks all friends at the CMPLab, with special recognition to Mehdi Ebrahimi Mehr, for their support.

Funding JAR was supported by the IRAN Cognitive Sciences & Technologies Council (Grant No. 9311), and Iran National Science Foundation, INSF (Grant No. 99010447 and 4001703).

Code Availability The codes accompanied with a tutorial for using the simple and robust EZ-CDM, along with the mixed model with guess are available at https://github.com/HasanQD/EZ-CDM. Also, implementations for calculating the likelihood function with this method and the simulation of the data are available in https://github.com/HasanQD/CDM.

References

Banerjee, A., Dhillon, I. S., Ghosh, J., Sra, S., & Ridgeway, G. (2005). Clustering on the unit hypersphere using von Mises-Fisher distributions. *Journal of Machine Learning Research* 6(9)

Bays, P., Gorgoraptis, N., Wee, N., Marshall, L., & Husain, M. (2011). Temporal dynamics of encoding, storage, and reallocation of visual working memory. *Journal of Vision, 11*, 6.

Bays, P. M., Catalao, R. F., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision, 9*(10), 7.

Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.

Boehm, U., Annis, J., Frank, M. J., Hawkins, G. E., Heathcote, A., Kellen, D., Krypotos, A. M., Lerche, V., Logan, G. D., Palmeri, T. J., et al. (2018). Estimating across-trial variability parameters of the diffusion decision model: Expert advice and recommendations. *Journal of Mathematical Psychology, 87*, 46–75.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision-making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review, 113*(4), 700–765.

Borodin, A. N., & Salminen, P. (2002). *Handbook of Brownian motion-facts and formulae*, 2nd edn. Springer Science & Business Media

Brown, S., & Heathcote, A. (2005). A ballistic model of choice response time. *Psychological Review, 112*(1), 117.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153–178.

Casella, G., & Berger, R. L. (2001). *Statistical inference*. Cengage Learning.

Dekking, F. M., Kraaikamp, C., Lopuhaä, H. P., & Meester, L. E. (2005). *A Modern Introduction to Probability and Statistics: Understanding why and how,* (vol. 488). Springer.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience, 32*(11), 3612–3628. https://doi.org/10.1523/JNEUROSCI.4010-11.2012

Evans, N. J., & Brown, S. D. (2017). People adopt optimal policies in simple decision-making, after practice and guidance. *Psychonomic Bulletin & Review, 24*(2), 597–606.

Evans, N. J., Bennett, A. J., & Brown, S. D. (2018). Optimal or not; depends on the task. *Psychonomic Bulletin & Review, 26*, 1027–1034.

Evans, N. J., Hawkins, G. E., & Brown, S. D. (2020). The role of passing time in decision-making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(2), 316.

Fan, J. E., & Turk-Brown, N. B. (2013). Internal attention to features in visual short-term memory guides object learning. *Cognition, 129*, 292–308.

Fontanesi, L., Gluth, S., Spektor, M. S., & Rieskamp, J. (2019). A reinforcement learning diffusion decision model for value-based decisions. *Psychonomic Bulletin & Review, 26*, 1099–1121.

Forstmann, B. U., Ratcliff, R., & Wagenmakers, E. J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology, 67*, 641–666.

Fougnie, D., & Alvarez, G. A. (2011). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *Journal of Vision, 11*, 3.

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications, 3*, 1229.

Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision-making. *Annual Review of Neuroscience, 30*, 535–574.

Gomez, P., Ratcliff, R., & Perea, M. (2007). A model of the go/no-go task. *Journal of Experimental Psychology: General, 136*(3), 389.

Green, M. L., & Pratte, M. S. (2022). Local motion pooling is continuous, global motion perception is discrete. *Journal of Experimental Psychology: Human Perception and Performance, 48*(1), 52.

Gunseli, E., van Moorselaar, D., Meeter, M., & Olivers, C. N. L. (2015). The reliability of retro-cues determine working memory representations. *Psychonomic Bulletin & Review, 22*, 1334–1341.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ..., & Oliphant, T. E. (2020). Array programming with numpy. *Nature, 585*(7825), 357–362.

Hawkins, G. E., & Heathcote, A. (2021). Racing against the clock: Evidence-based versus time-based decisions. *Psychological Review, 128*(2), 222. https://doi.org/10.1037/rev0000259

Hawkins, G. E., Forstmann, B. U., Wagenmakers, E. J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience, 35*(6), 2476–2484. https://doi.org/10.1523/JNEUROSCI.2410-14.2015

Heath, R. A. (2000). The Ornstein-Uhlenbeck model for decision time in cognitive tasks: An example of control of nonlinear network dynamics. *Psychological Research, 63*(2), 183–191. https://doi.org/10.1007/PL00008177

Huang-Pollock, C., Ratcliff, R., McKoon, G., Roule, A., Warner, T., Feldman, J., & Wise, S. (2020). A diffusion model analysis of sustained attention in children with attention deficit hyperactivity disorder. *Neuropsychology, 34*, 641–653. https://doi.org/10.1037/neu0000636

Kato, S., & Eguchi, S. (2016). Robust estimation of location and concentration parameters for the von Mises-Fisher distribution. *Statistical Papers, 57*(1), 205–234.

Kool, W., Conway, A. R. A., & Turk-Browne, N. B. (2014). Sequential dynamics in visual short-term memory. *Attention, Perception, & Psychophysics, 76*, 1885–1901.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298.

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology, 3*, 193.

Kvam, P. D. (2019). A geometric framework for modeling dynamic decisions among arbitrarily many alternatives. *Journal of Mathematical Psychology, 91*, 14–37.

Kvam, P. D. (2019). Modeling accuracy, response time, and bias in continuous orientation judgments. *Journal of Experimental Psychology: Human Perception and Performance, 45*(3), 301.

Kvam, P. D., & Turner, B. M. (2021). Reconciling similarity across models of continuous selections. *Psychological Review*.

Kvam, P. D., Marley, A., & Heathcote, A. (2021). A unified theory of discrete and continuous responding. *Psychological Review*. https://doi.org/10.1037/rev0000378

Laming, D. R. J. (1968). *Information theory of choice-reaction times*. Academic Press.

Lerche, V., & Voss, A. (2016). Model complexity in diffusion modeling: Benefits of making the model more parsimonious. *Frontiers in Psychology*, 1324

Lerche, V., & Voss, A. (2017). Retest reliability of the parameters of the ratcliff diffusion model. *Psychological Research, 81*, 629–652.

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? a comparison of different optimization criteria. *Behavior Research Methods, 49*, 513–537.

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17*, 347–356.

Manning, C., Hassall, C. D., Hunt, L. T., Norcia, A. M., Wagenmakers, E. J., Evans, N. J., & Scerif, G. (2022). Behavioural and neural indices of perceptual decision-making in autistic children during visual motion tasks. *Scientific Reports, 12*(1), 1–19. https://doi.org/10.1038/s41598-022-09885-4

Manning, C., Hassall, C. D., Hunt, L. T., Norcia, A. M., Wagenmakers, E. J., Margaret, Snowling G., & J adn Scerif, Evans NJ,. (2022). Visual motion and decision-making in dyslexia: Reduced accumulation of sensory evidence and related neural dynamics. *Journal of Neuroscience, 42*(1), 121–134. https://doi.org/10.1523/JNEUROSCI.1232-21.2021

Mardia, K. V., & Jupp, P. E. (1999). *Directional statistics*. Wiley Online Library

Marshall, L., & Bays, P. M. (2013). Obligatory encoding of task-irrelevant features depletes working memory resources. *Journal of Vision, 13*, 21.

Matzke, D., Dolan, C. V., Logan, G. D., Brown, S. D., & Wagenmakers, E. J. (2013). Bayesian parametric estimation of stop-signal reaction time distributions. *Journal of Experimental Psychology General, 142*(4), 1047.

Matzke, D., Hughes, M., Badcock, J. C., Michie, P., & Heathcote, A. (2017). Failures of cognitive control or attention? the case of stop-signal deficits in Schizophrenia. *Attention, Perception, & Psychophysics, 79*, 1078–1086.

Matzke, D., Love, J., & Heathcote, A. (2017). A bayesian approach for estimating the probability of trigger failures in the stop-signal paradigm. *Behavior Research Methods, 49*, 267–281.

Miletić, S., Boag, R. J., Trutti, A. C., Stevenson, N., Forstmann, B. U., & Heathcote, A. (2021). A new model of decision processing in instrumental learning tasks. *Elife, 10*, e63055. https://doi.org/10.7554/eLife.63055

mpmath development team, T. (2023). mpmath: A Python library for arbitrary-precision floating-point arithmetic (version 1.3.0). http://mpmath.org/

Nejati, V., Hadian-Rasanan, A. H., Rad, J. A., Alavi, M. M., Haghie, S., & Nitsche, M. A. (2022). Transcranial direct current stimulation (tDCS) alters the pattern of information processing in children with ADHD: Evidence from drift diffusion modeling. *Neurophysiologie Clinique, 52*, 17–21. https://doi.org/10.1016/j.neucli.2021.11.005

Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *The Computer Journal, 7*(4), 308–313.

Pearson, K. (1894). Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London A, 185*, 71–110.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review, 24*, 1234–1251.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59.

Ratcliff, R. (2008). The EZ diffusion method: Too EZ? *Psychonomic Bulletin & Review, 15*(6), 1218–1228.

Ratcliff, R. (2013). Parameter variability and distributional assumptions in the diffusion model. *Psychological Review, 120*(1), 281.

Ratcliff, R. (2018). Decision making on spatially continuous scales. *Psychological Review, 125*(6), 888.

Ratcliff, R., & Childers, R. (2015). Individual differences and fitting methods for the two-choice diffusion model of decision making. *Decision, 2*(4), 237.

Ratcliff, R., & McKoon, G. (2020). Decision making in numeracy tasks with spatially continuous scales. *Cognitive Psychology, 116*, 101259.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review, 111*(2), 333.

Ratcliff, R., & Tuerlinckx, F. (2002). Estimating parameters of the diffusion model: Approaches to dealing with contaminant reaction times and parameter variability. *Psychonomic Bulletin & Review, 9*(3), 438–481.

Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences, 20*(4), 260–281.

Ratcliff, R., Huang-Pollock, C., & McKoon, G. (2018). Modeling individual differences in the go/no-go task with a diffusion model. *Decision, 5*(1), 42–62.

Ross SM (2007) Introduction to probability models, 9th edn. Academic press

Sewell, D. K., Jach, H. K., Boag, R. J., & Heer, C. A. V. (2019). Combining error-driven models of associative learning with evidence accumulation models of decision-making. *Psychonomic Bulletin & Review, 26*, 868–893.

Sheskin, D. J. (2003). *Handbook of parametric and nonparametric statistical procedures*. Chapman and hall/CRC

Shinn, M., Lam, N. H., & Murray, J. D. (2020). A flexible framework for simulating and fitting generalized drift-diffusion models. *ELife, 9*, e56938. https://doi.org/10.7554/eLife.56938

Smith, P., Garrett, P., & Zhou, J. (2023). Obtaining stable predicted distributions of response times and decision outcomes for the circular diffusion model

Smith, P. L. (2016). Diffusion theory of decision making in continuous report. *Psychological Review, 123*(4), 425.

Smith, P. L. (2019). Linking the diffusion model and general recognition theory: Circular diffusion with bivariate-normally distributed drift rates. *Journal of Mathematical Psychology, 91*, 145–158.

Smith, P. L., & Corbett, E. A. (2019). Speeded multielement decision-making as diffusion in a hypersphere: Theory and application to double-target detection. *Psychonomic Bulletin & Review, 26*(1), 127–162.

Smith, P. L., Saber, S., Corbett, E. A., & Lilburn, S. D. (2020). Modeling continuous outcome color decisions with the circular diffusion model: Metric and categorical properties. *Psychological Review, 127*(4), 562.

Starns, J. J., & Ratcliff, R. (2012). Age-related differences in diffusion model boundary optimality with both trial-limited and time-limited tasks. *Psychonomic Bulletin & Review, 19*(1), 139–145.

Stigler, S. M. (2007). The epic story of maximum likelihood. *Statistical Science*, 598–620.

Stone, M. (1960). Models for choice-reaction time. *Psychometrika, 25*(3), 251–260. https://doi.org/10.1007/BF02289729

Storn, R., & Price, K. (1997). Differential evolution-a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization, 11*(4), 341.

Swan, G., & Wyble, B. (2014). The binding pool: A model of shared neural resources for distinct items in visual working memory. *Attention, Perception, & Psychophysics, 76*, 2136–2157.

Tillman, G., Van Zandt, T., & Logan, G. D. (2020). Sequential sampling models without random between-trial variability: The racing diffusion model of speeded decision making. *Psychonomic Bulletin & Review, 27*(5), 911–936.

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71*, 1–26.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review, 108*(3), 550.

van den Berg, R., Shin, H., Chou, W. C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences, 109*, 8780–8785.

van den Berg, R., Awh, E., & Ma, W. J. (2014). Factorial comparison of working memory models. *Psychological Review, 121*, 124–149.

vanRossum, G. (1995) Python reference manual. Department of Computer Science [CS] (R 9525)

van Ravenzwaaij, D., & Oberauer, K. (2009). How to use the diffusion model: Parameter recovery of three methods: EZ, fast-dm, and DMAT. *Journal of Mathematical Psychology, 53*(6), 463–473.

van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review, 24*(2), 547–556.

van Ravenzwaaij, D., Brown, S. D., Marley, A., & Heathcote, A. (2020). Accumulating advantages: A new conceptualization of rapid multiple choice. *Psychological Review, 127*(2), 186–215. https://doi.org/10.1037/rev0000166

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, I., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., SciPy 10 Contributors. (2020). Scipy 1.0: Fundamental algorithms for scientific computing in python. *Nature Methods*, *17*(3), 261–272

Voss, A., Lerche, V., Mertens, U., & Voss, J. (2019). Sequential sampling models with variable boundaries and non-normal noise: A comparison of six models. *Psychonomic Bulletin & Review, 26*(3), 813–832. https://doi.org/10.3758/s13423-018-1560-4

Wagenmakers, E. J. (2009). Methodological and empirical developments for the Ratcliff diffusion model of response times and accuracy. *European Journal of Cognitive Psychology, 21*(5), 641–671.

Wagenmakers, E. J., Van Der Maas, H. L., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review, 14*(1), 3–22.

Wagenmakers, E. J., van der Maas, H. L. J., Dolan, C. V., & Grasman, R. P. P. P. (2008). Ez does it! extensions of the EZ-diffusion model. *Psychonomic Bulletin & Review, 15*(6), 1229–1235.

Weigard, A., & Huang-Pollock, C. (2017). The role of speed in ADHD-related working memory deficits: A time-based resource-sharing and diffusion model account. *Clinical Psychological Science, 5*, 195–211. https://doi.org/10.1177/2167702616668320

Yap, M. J., Balota, D. A., Sibley, D. E., & Ratcliff, R. (2012). Individual differences in visual word recognition: Insights from the english lexicon project. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 53.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235.

Zhou, J., Osth, A. F., Lilburn, S. D., & Smith, P. L. (2021). A circular diffusion model of continuous-outcome source memory retrieval: Contrasting continuous and threshold accounts. *Psychonomic Bulletin & Review, 28*(4), 1112–1130.