Dear Reviewers,

Thank you for your feedback on my manuscript.

Following the suggestions of Reviewer 2, I have made two major modifications to the manuscript.

First, I have included a second dataset with 8 subjects, for which the Hips model could be calculated. This confirmed the main result of the paper: large improvements in accuracy can be obtained even with limited kinematic information. Unfortunately, I could not find another public dataset with full body kinematics.

Second, to demonstrate the practical applications of the method, I have added the calculation of a commonly used balance indicator: the control of foot placement using the CoM state. I show that, whereas different kinematic models provide very different measures of balance (r-square of the foot placement prediction ranging from 0.13 to 0.38), the measure derived from the novel method is much more reliable (r-square of 0.58 for all three kinematic methods). The novel method may therefore be less biased when evaluating balance indicators in populations for which no accurate anthropometric data is available.

Please find below in blue the responses to your comments.

Best regards,

Charlotte Le Mouel

I. Reviewer 1

The paper present the use of Kalman filter to estimate the COM. The method is compared to model based methods and validated against the COM acceleration during the flight phase of running.
Overall the paper is clear and complete.
I only have small comments:

- The use of kalman filter method should be added to the title and abstract. Optimal merging, for me, recall more optimization methods than kalman filter.
  The title has been changed accordingly.
- In the methods section, it's not clear how the model were personalized.
  This section has been expanded (lines 201 to 206): "The coordinates of 37 markers are used to calculate three-dimensional segment coordinate systems. An anthropometric table (Dumas et al., 2007; Dumas & Wojtusch, 2018) is then used to place the segment CoM in the longitudinal, antero-posterior and lateral directions relative to the segment origin, and to determine the fraction of the segment's weight relative to the person's weight. The whole body CoM is then obtained as a weighted sum of the segment CoMs."
- In table 3, it's not clear if the data is in meters or mm. It's written in meters but then the percentage is with respect to something in mm.
  The data is in mm, thank you for pointing this out.
- Figure 3 and 4 have the same figure.
  Thank you for pointing this out! Figure 3 is now the correct figure.

II. Reviewer 2

The manuscript presents a novel approach for determining the position of the whole-body center of mass (CoM) by combining kinematic and kinetic data through the use of a Kalman filter. The authors effectively address the inherent limitations in both kinematic and kinetic methods when used in isolation, proposing a solution that reduces errors in CoM estimation, especially during dynamic activities such as running. The study shows that the

combination of kinematic and kinetic information can significantly reduce the error in estimating CoM positions, with the proposed method being applicable across different kinematic models. The paper is well-written and structured, with the problem statement clearly outlined and the methodology appropriately justified. The manuscript's integration of statistical methods, such as bootstrapping and the Mann-Whitney U test, shows a thoughtful approach to handling the small sample size.
Below you can find my detailed review of each section of the manuscript:

Title and Abstract:
- The title of the manuscript seems to clearly reflect the topic of the manuscript.
- The abstract summarizes the key findings, including the error reduction achieved with the Kalman filter method, and clearly communicates the study's contributions.

Introduction:
- The introduction outlines the challenges in CoM estimation and proposes combining kinematic and kinetic data to improve accuracy. The hypotheses are logically presented.
- The literature review presented in the introduction, addresses the different types of CoM estimation methods properly, however, it comes to my concern that among the "Kinetic" methods the "LowPass filtering of the Center of Pressure" method (Caron et al., 1997), was left unmentioned. Moreover, a relatively recent and popular CoM estimation technique named "Statically Equivalent Serial Chain" (Cotton et al., 2009) was also left unmentioned in the literature section. This technique also utilizes both kinetic and kinematic information for CoM estimation. I would recommend this section to be modified accordingly.
These are now mentioned in the introduction:
Lines 64 to 67: “An alternative kinetic method is to low-pass filter the CoP (Caron et al., 1997). During quiet standing, this should reflect the low-frequency movements of the CoM. This method is however less accurate than either kinetic methods using the GRF or the kinematic methods described below (Lafond et al., 2004).”
Lines 124 to 137: “Finally, kinetics can be used to obtain individualised body segment inertial parameters for a personalised kinematic model (Cotton et al., 2008; Jovic et al., 2016; Bonnet et al., 2016; Chebel & Tunc, 2023). This requires an identification phase, during which the subject performs a series of poses. The inertial parameters providing the best fit between the kinematic CoM and the kinetics are determined. This can reduce the mismatch between kinematics and kinetics by 10 to 45 %, but at the cost of the subject performing an elaborate seven minute “dance” during the identification phase (Bonnet et al., 2016). In robotics, the statically equivalent serial chain technique is used to obtain personalised kinematic models (Cotton et al., 2008). This allows the CoM to be obtained from a more limited set of parameters than the whole set of body inertial parameters. The identification these parameters nevertheless requires the subject to perform 75 static poses, and the mismatch between the horizontal CoM and CoP locations during these static poses is still 25.5 mm with the optimal parameters (Chebel & Tunc, 2023). Identification techniques therefore require the subject to perform a long series of poses (which may not be feasible for pathological populations) for only a moderate improvement in accuracy.”

Methods:
- The methods are detailed enough, including the use of a publicly available database and clearly defined protocols for data processing and analysis.
- Regarding the sample size, this can be considered a major limitation for this study (as mentioned in the limitation section), as only using 2 subjects, cannot reflect the generalizability of the study, I strongly recommend increasing the number of the participants (could be through incorporating data from another dataset).
Unfortunately, I could not find another public dataset with full body kinematics. Instead, I have added a second dataset with 8 subjects from which the Hips model could be calculated. This confirmed the main result of the paper: large improvements in accuracy can be obtained even with limited kinematic information.
- Regarding the statistical analysis, the use of Mann-Whitney U test and bootstrapping for confidence intervals, are appropriate and described in detail. Although the study could benefit from a larger sample size for more robust conclusions, there is no mention of a formal power analysis, which is typically used to determine whether a study has sufficient sample size to detect meaningful effects.

The number of samples yielded by such a power analysis correspond to the summed duration of the flight phases (for the Acceleration Bias) and of the gait cycle durations (for the Position Error). The narrow confidence intervals (Tables 1 and 2) indicate that there are enough samples (i.e. the trials are long enough).
The reviewer may however be refering to the low number of subjects (2 in the first submission). I have therefore included data from a second dataset (8 additional subjects) which confirms the initial conclusions.

Results:
- The results are correctly described and interpreted. They support the hypothesis that combining kinematic and kinetic data reduces error in CoM estimation.

Discussion:
- The conclusions are supported by the results and do not overstate the findings. The author does acknowledge the need for further research to validate the method with larger datasets. However, the discussion could have been more focused on the practical implications of the proposed method. For instance, how could this Kalman filter-based approach be integrated into existing motion capture systems? How would it improve upon current methods in clinical gait analysis or athletic performance monitoring? The study would benefit from more concrete examples or applications.
To demonstrate the practical applications of the method, I have added the calculation of a commonly used balance indicator: the control of foot placement using the CoM state. I show that, whereas different kinematic models provide very different measures of balance (r-square of the foot placement prediction ranging from 0.13 to 0.38), the measure derived from the novel method is much more reliable (r-square of 0.58 for all three kinematic methods). The novel method may therefore be less biased when evaluating balance indicators in populations for which no accurate anthropometric data is available.
- Another concern which should be pointed out, is that the study focuses heavily on the use of a Kalman filter to combine kinematic and kinetic data, without exploring alternative filtering or data fusion techniques. The assumption of linearity and Gaussian noise makes the Kalman filter suboptimal for handling non-linear dynamics or complex noise models, which may be more prevalent in biomechanical data. The author could have explored or at least discussed other potential methods.
The most common approach for handling non-linear dynamics in biomechanics is inverse dynamics. This approach (and its limited accuracy) are mentioned in the introduction (lines 119 to 122). The more complex structure of the noise is discussed in the Limitations (lines 589 to 596).
- Moreover, the lack of direct ground truth for horizontal CoM Coordinates should also be emphasized. Although the GRF are regarded as almost ground truth, it should be mentioned that it is still not optimal.
I believe this is a misunderstanding. I use as ground truth the fact the CoM acceleration during flight is equal to gravity. This provides a ground truth for the horizontal acceleration: it is 0 m/s^2. This ground truth is used to calculate the Acceleration Bias in all directions.
The GRF is used only for the Position Error. I have added emphasis in the discussion that this cannot be considered as ground truth (lines 514 to 517): “The accuracy of the kinematic models was also evaluated using a classical method: comparing the kinematic position to the kinetic position. As explained in the introduction, this method relies on pre-processing and double integration of the GRF, and therefore cannot be considered as ‘ground truth’.”

Summary of the review:
While this study provides an innovative solution to the problem of CoM estimation by combining kinematic and kinetic data through a Kalman filter, there are several areas where the work could be significantly improved. The small sample size, assumptions about Gaussian noise, and lack of exploration of alternative methods weaken the overall rigor of the study. Additionally, more discussion around computational complexity and practical applications would greatly enhance the manuscript’s relevance. Addressing these points would result in a more robust, generalizable, and practically useful contribution to the field, otherwise this manuscript wouldn't be ready for publication.

Minor comments:
- Line 44: Please mention a reference for this statement "with an accuracy typically within 1 % of the person’s weight".

This sentence was changed to: “Good quality force-plates have a measurement noise of a few Newton, and can provide a very accurate measure of the CoM acceleration”.

- Line 156: "A few short gaps ...", I would advise to be more specific about the number/frequency of gaps available in the dataset.

This has been specified: “Over all subjects and trials, 8 markers were invisible for a duration of up to 3 seconds. These gaps were filled through linear interpolation.”

- The manuscript would benefit from a thorough language revision to improve clarity and readability.