Journal of Biomechanics 175 (2024) 112302

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech

Inverse kinematics in cervical spine models: Effects of scaling and model degrees of freedom for extension and flexion movements

Hamidreza Barnamehei [a], Yu Zhou [b], Xudong Zhang [b], Anita N. Vasavada [a,*]

[a] Voiland School of Chemical Engineering and Bioengineering; and Department of Integrative Physiology and Neuroscience, Washington State University, Pullman, WA, USA

[b] Department of Industrial & Systems Engineering, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Keywords:
Neck model
Intervertebral kinematics
Inverse kinematics
Musculoskeletal modeling
Model scaling
Model complexity

ABSTRACT

Intervertebral kinematics can affect model-predicted loads and strains in the spine; therefore knowledge of expected vertebral kinematics error is important for understanding the limitations of model predictions. This study addressed how different kinematic models of the neck affect the prediction of intervertebral kinematics from markers on the head and trunk. Eight subjects executed head and neck extension-flexion motion with simultaneous motion capture and biplanar dynamic stereo-radiography (DSX) of vertebrae C1-C7. A generic head and neck model in OpenSim was scaled by marker data, and three versions of the model were used with an inverse kinematics solver. The models differed according to the number of independent degrees of freedom (DOF) between the head and trunk: 3 rotational DOF with constraints defining intervertebral kinematics as a function of overall head-trunk motion; 24DOF with 3 independent rotational DOF at each level, skull-T1; 48DOF with 3 rotational and 3 translational DOF at each level. Marker tracking error was lower for scaled models compared to generic models and decreased as model DOF increased. The largest mean absolute error (MAE) was found in extension-flexion angle and anterior-posterior translation at C1-C2, and superior-inferior translation at C2-C3. Model scaling and complexity did not have a statistically significant effect on most error metrics when corrected for multiple comparisons, but ranges of motion were significantly different from DSX in some cases. This study evaluated model kinematics in comparison to gold standard radiographic data and provides important information about intervertebral kinematics error that are foundational to model validity.

1. Introduction

Neck musculoskeletal models have the potential to improve the diagnosis and treatment of neck musculoskeletal disorders by estimating the loads and displacements of tissue components. Accurate model prediction depends, among other factors, on the definition of vertebral body positions (Vasavada et al., 2018), which are important to calculating muscle and ligament lengths. Vertebral positions cannot be known explicitly from external measurements (Johnson, 1998), requiring modeling assumptions about the relation between external landmarks and intervertebral kinematics.

Inverse kinematics (IK) algorithms, which minimize error in external marker trajectories, are often used to define joint kinematics in musculoskeletal models. These algorithms have been used extensively in the upper and lower limbs (reviewed in (Begon et al., 2018)). IK algorithms are more challenging for the spine, however, because external markers cannot be placed on every segment, and thus multiple joint kinematics must be estimated with the same marker trajectories. Constraints on model joint definitions, which decrease the degrees of freedom (DOF) are often used to make the problem tractable. In thoracolumbar spine models, it was found that increasing the number of DOF always decreases marker error and decreases intervertebral kinematics error to some extent, but an excessive number of DOF results in unrealistic intervertebral kinematics (Alemi et al., 2021, Wang et al., 2021). In addition, vertebral size is related to subject stature (DeSantis Klinich et al., 2004); therefore linear measurements such as intervertebral translations that relate to vertebral size are likely affected by model size.

We have previously found that the predictions of biomechanical models are affected by the assumptions used to define vertebral positions in models (Nevins et al., 2014, Vasavada et al., 2018). However,

* Corresponding author at: Voiland School of Chemical Engineering and Bioengineering, Department of Integrative Physiology and Neuroscience, Washington State University, Pullman, WA 99164-1505, USA.
E-mail address: vasavada@wsu.edu (A.N. Vasavada).

https://doi.org/10.1016/j.jbiomech.2024.112302
Accepted 30 August 2024
Available online 1 September 2024

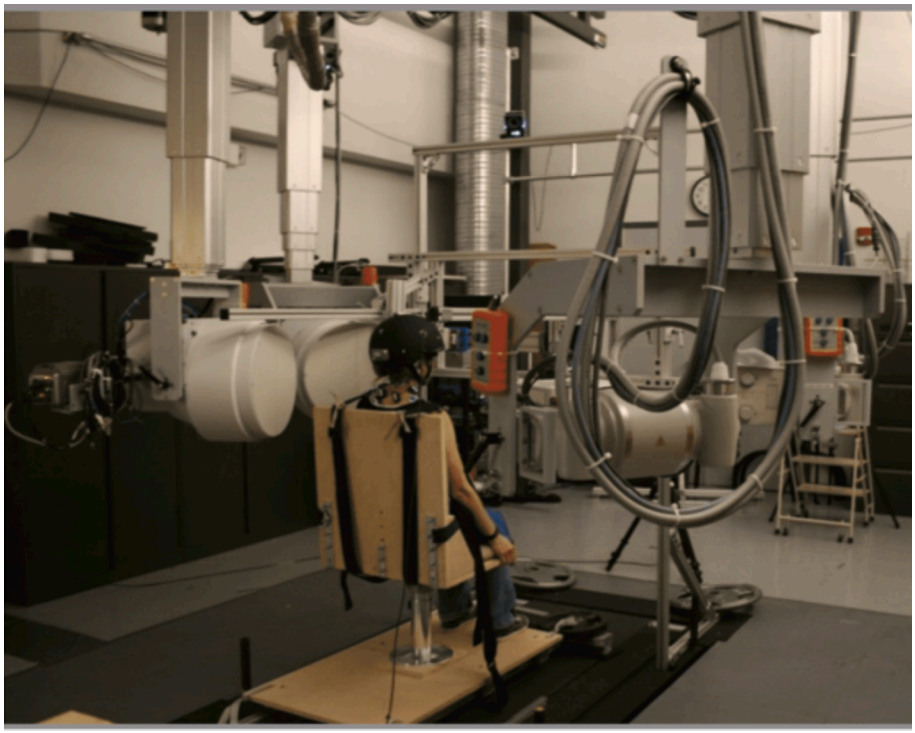

Fig. 1. Experimental setup consisting of height-adjustable chair, dynamic radiographic imaging system, Vicon surface-markers, and Vicon motion analysis system. Adapted from Zhou et al., 2020.

Table 1
Representation of model complexity in 3, 24, and 48 DOF models.

	3DOF	24DOF	48DOF
Allows independent translations (*i.e.*, center of rotation is not fixed)			X
Allows independent rotational motion at all joints between skull and trunk (*i.e.*, each vertebral motion is not a fixed percentage of head-trunk angle)		X	X
Allows non-sagittal motion	X	X	X
Number of rotational DOFs (each level)	3 (Constrained)	3	3
Number of translational DOFs (each level)	0	0	3

cervical spine intervertebral kinematics have not been estimated with an IK solver and compared to the "gold standard"—dynamic stereo-radiography (DSX) data. Moreover, it is not known how model joint kinematic definitions (specifically, the number of rotational and translational DOF) affect model predictions and error.

The purpose of this study was to evaluate how scaling and model complexity (number and type of DOF) affect inverse kinematics predictions in a neck model. We hypothesized that both scaling model size and increasing the DOF will decrease marker tracking errors and joint kinematics errors. The results of this study also provide an estimate of expected kinematics error when radiographic data are not available.

2. Methods

2.1. Subjects

Eight healthy subjects (4 male, 4 female) participated in the study, with average (± standard deviation) mass 67.2 (± 10.7) kg, height 170.2 (± 10.1) cm and age 29.4 (± 7.7) years. All participants were asymptomatic, free from any injuries or musculoskeletal disorders, with no history of neck surgery or chronic neck pain. The study was approved by the Institutional Review Board of University of Pittsburgh (where the human subject data were collected), and all subjects provided informed consent.

2.2. Instrumentation and tasks

Subjects were seated within a custom experimental setup (Fig. 1) with a dynamic stereo-radiography (DSX) system and Vicon motion capture system (Zhou et al., 2020). The DSX system included two cardiac-cine angiography generators (EMD Technologies, CPX-3100CV), two X-ray tubes with 0.3/0.6 mm focal spot, two 16-in Thalus image intensifiers, and two high-speed digital cameras (4-megapixels Phantom v10, Vision Research). Radiographic images were recorded with a sampling frequency of 30 frames per second for 3.2 s. The DSX system parameters were pulsed exposure time=2.5 ms, excitation voltage=70 kV, and current=160 mA. Intervertebral kinematics were calculated from a previously-validated model-based tracking process (Anderst et al., 2011). 6DOF intervertebral kinematics (3 translations and 3 rotations) were calculated for each intervertebral joint from C1 to C7.

Twelve Vicon motion capture cameras (Vantage-Series, Vicon Motion Labs, USA) with sampling frequency of 60 Hz were used to record marker trajectories. Skin-mounted markers were attached to ten anatomical landmarks: the left and right tragi (LTRAG/RTRAG), inferior border of each orbit (LORBIT/RORBIT), glabella (GLAB), acromion processes (LSHO/RSHO), suprasternal notch (SNOTCH), a lower point on the sternum (STERN), and C7 spinous process (C7). Vicon Nexus 2.2.5 software was used to create time series of three-dimensional marker positions. Marker and DSX data were time-synched and transformed to a common coordinate system.

Subjects executed two kinds of tasks: (1) neutral posture (static), looking directly forward, and (2) continuous motion from a starting position through a full range of extension and flexion and back to starting position. Starting position was neutral in 2 subjects, flexed in 5 subjects and extended in one subject.

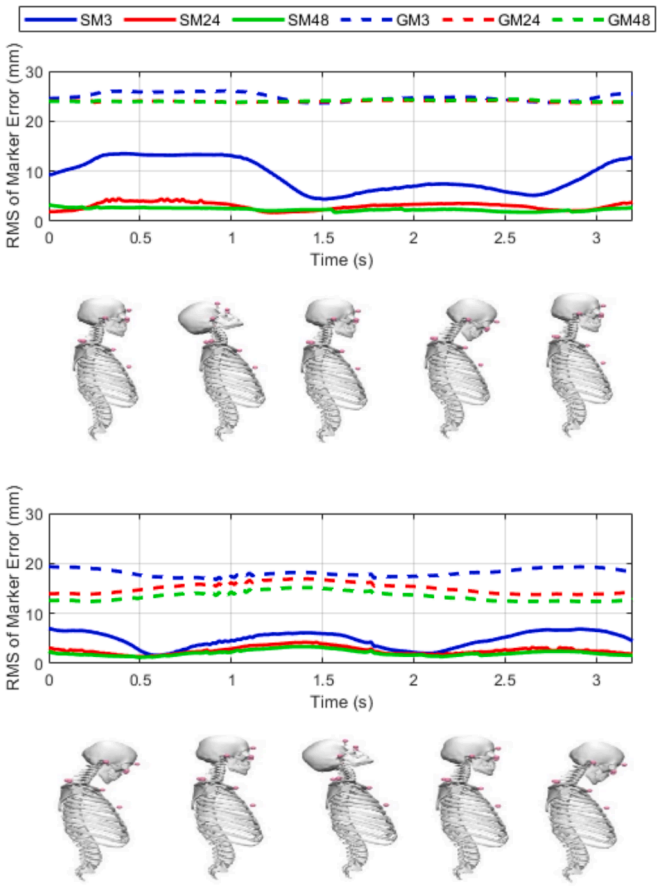

Fig. 2. The RMS error (mm) during the extension-flexion tasks for 10 markers from OpenSim inverse kinematics analysis for 6 head-neck models (generic and scaled) with respect to time for two subjects. One subject started in a neutral position and the other started in flexion, but both went through a complete cycle of full extension and flexion before returning to the starting position.

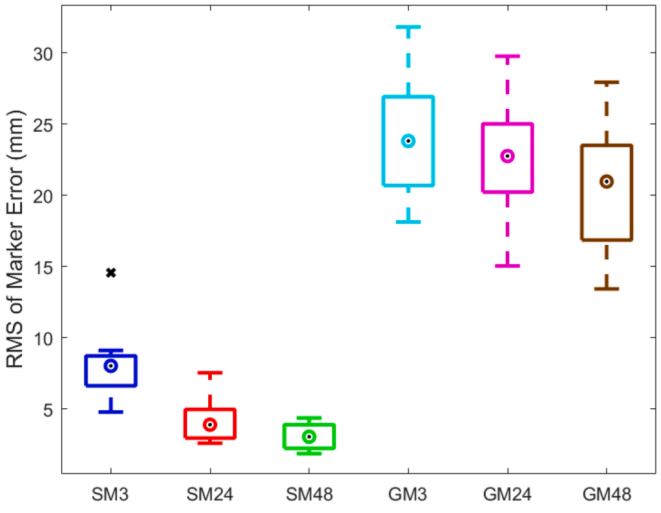

Fig. 3. Boxplots of average RMSE (mm) over 10 markers during the extension-flexion tasks for 8 subjects and 6 models (scaled and generic, with differing degrees of freedom). The central marker indicates the median, the box indicates the interquartile range (25th–75th percentile), the whiskers are 1.5 times the interquartile range, and x's are considered outliers.

2.3. Musculoskeletal modeling

A head and neck model (Vasavada et al., 1998) was implemented in OpenSim (Delp et al., 2007). Three versions of the base model differed according to the number of independent degrees of freedom (DOF) between the head and trunk (Table 1).

3DOF: The 3 DOF are Flexion-Extension, Lateral Bending, and Axial Rotation angles of the skull with respect to T1. The amount of rotation at each intervertebral joint is assumed to be a fixed percentage of the total angular motion between the skull and T1 (Supplemental Table S1; (White and Panjabi, 1990)). The center of rotation is fixed according to the literature (Penning, 1978, Amevo et al., 1991).

24DOF: These models have one rotational DOF each for Flexion-Extension, Lateral Bending and Axial Rotation at each level from skull-T1 (3*8=24DOF). There is no constraint on the relative amount of rotation at each level. The centers of rotation are fixed as in the 3DOF models.

48DOF: Each intervertebral joint between skull-T1 includes six DOFs: Flexion-Extension, Lateral Bending, and Axial Rotation rotational DOFs; and Superior-Inferior, Anterior-Posterior, and Medial-Lateral translational DOFs (6*8=48DOF). There is no constraint on the amount of rotation or translation at any level.

2.4. Model scaling and inverse kinematics

The generic (base) model represented a 50th percentile male. Subject-specific (scaled) models were created by using the scaling tool of OpenSim 4.1 software (Delp et al., 2007). The generic models were first scaled to match the size of each participant using neutral posture marker data. Next, marker registration was accomplished by modifying the position of virtual markers to match the position of experimental markers on the subject using photographs and the "Adjust Model Marker" part of OpenSim's scaling tool. These two processes (scaling and registration) were iterated several times based on marker error to determine the best scaled model. The trajectories of the head and trunk markers were input to the Inverse Kinematics Solver of OpenSim, and the intervertebral kinematics were calculated at each intervertebral level for each model.

2.5. Data analysis

Six model types were compared for marker errors and kinematics errors: GM3, GM24 and GM48 (generic models) and SM3, SM24, SM48 (scaled models), with 3, 24 and 48 DOF respectively. The distance between experimental marker and virtual marker position in the model was calculated as the root mean square error (RMSE) in OpenSim.

Several different metrics were used to evaluate intervertebral kinematics. The mean absolute error (MAE) over the extension-flexion motion was used to compare the model prediction and experimental radiographic (DSX) kinematics. Pearson correlation coefficients (r) were used to quantify the similarity of shape between model predictions and experimental data for each level and model. The range of motion (ROM) was compared among the DSX and scaled model data as a gross measure of magnitude of movement. Translational kinematics were only evaluated for the 48DOF models because translation was constrained by the fixed center of rotation in 3DOF and 24DOF models.

2.6. Statistical analysis

Data were checked for normality using the Shapiro-Wilk test. 16% of the variables had a non-normal distribution. Rather than mixing parametric and nonparametric tests and because the sample size was relatively small, all data were analyzed using nonparametric tests. Results are reported as median and interquartile range (IQR). The Friedman test

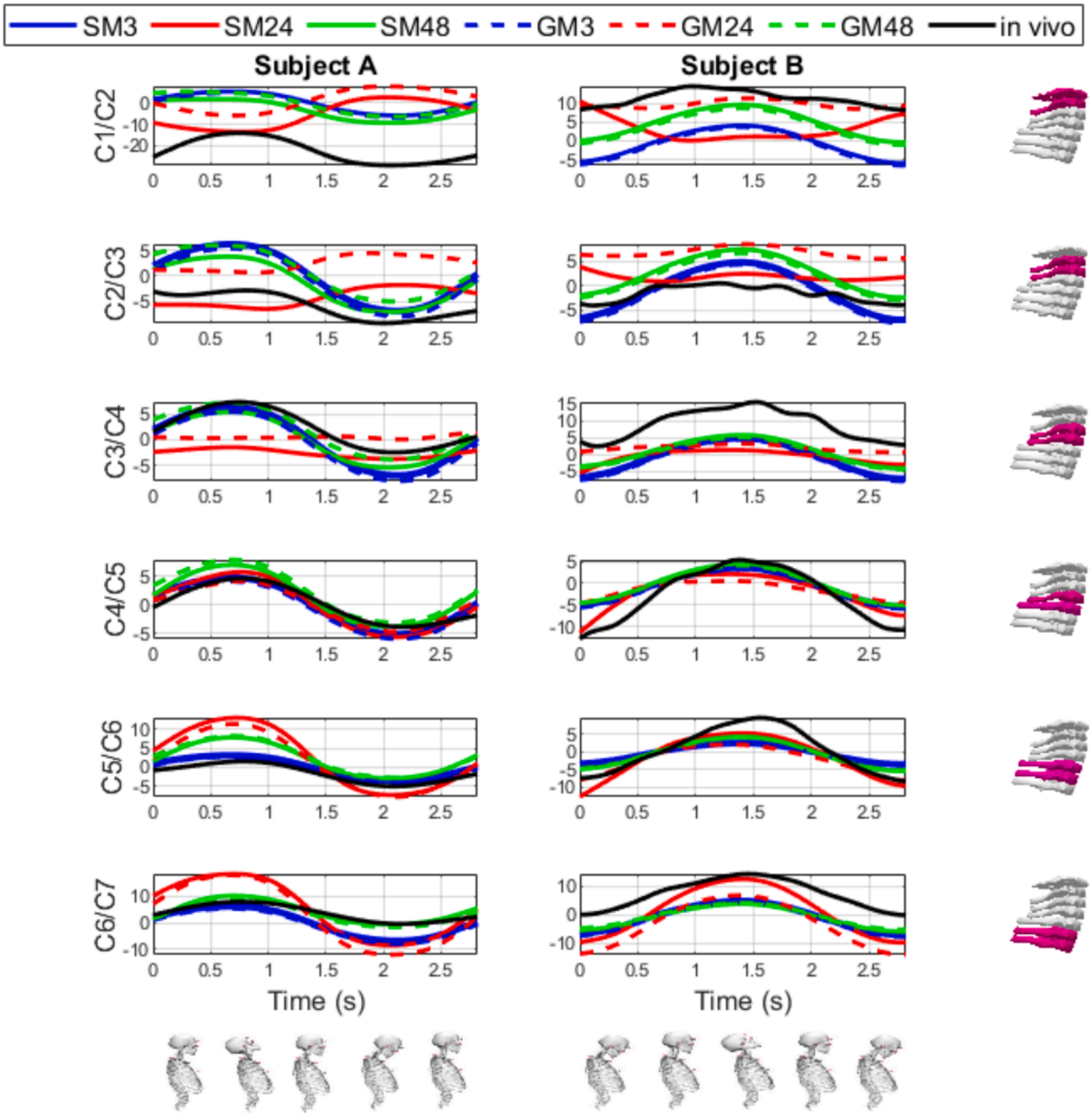

Fig. 4. The extension-flexion angles (deg) for 3 head-neck models (generic and scaled) and in vivo (DSX) results for six cervical intervertebral levels (C1/C2, C2/C3, C3/C4, C4/C5, C5/C6, and C6/C7) with respect to time for the same two subjects as in Fig. 2.

(nonparametric version of one-way repeated measures ANOVA) was used to compare more than two means. If the result indicated significant difference ($p<0.05$), post-hoc pairwise comparisons were performed using the Wilcoxon Signed Ranks Test (nonparametric version of paired *t*-test). A Bonferroni correction was applied to adjust the significance level as $\alpha_c=0.05/c$, where c is the number of planned comparisons. The 6 intervertebral levels (C1/C2, C2/C3, C3/C4, C4/C5, C5/C6, and C6/C7) were analyzed separately.

To test the hypotheses about the effects of model scaling (generic vs. scaled) and complexity (3, 24 or 48 DOF), the variables evaluated were (1) overall marker RMSE, (2) MAE of model rotational kinematics for each level, and (3) Pearson correlation coefficient *r* for each level. If Friedman test indicated significant differences, 9 pairwise comparisons by Wilcoxon Signed Ranks test were planned: 3 to evaluate the effect of scaling (SM3 vs. GM3, SM24 vs. GM24, and SM48 vs. GM48) and 6 to evaluate the effect of model complexity (SM3 vs. SM24, SM3 vs. SM48, SM24 vs. SM48, GM3 vs. GM24, GM3 vs. GM48, and GM24 vs. GM48). For translational MAE, SM48 and GM48 were compared by Wilcoxon Signed Ranks test.

The Friedman test was also employed to compare rotational ROM between DSX and scaled models (SM3, SM24, and SM48). Six pairwise comparisons were planned to compare rotational ROM: DSX vs. SM3, DSX vs. SM24, DSX vs. SM48, SM3 vs. SM24, SM3 vs. SM48, and SM24 vs SM48. For translational ROM, DSX and SM48 were compared by Wilcoxon Signed Ranks test.

3. Results

3.1. Model scaling

The male participant percentile ranges were 1st to 64th percentile mass and 12th to 99th percentile height, and the female participant percentiles ranged from 34th to 62nd percentile mass and 10th to 87th percentile height (Gordon et al., 2014). Correspondingly, scale factors

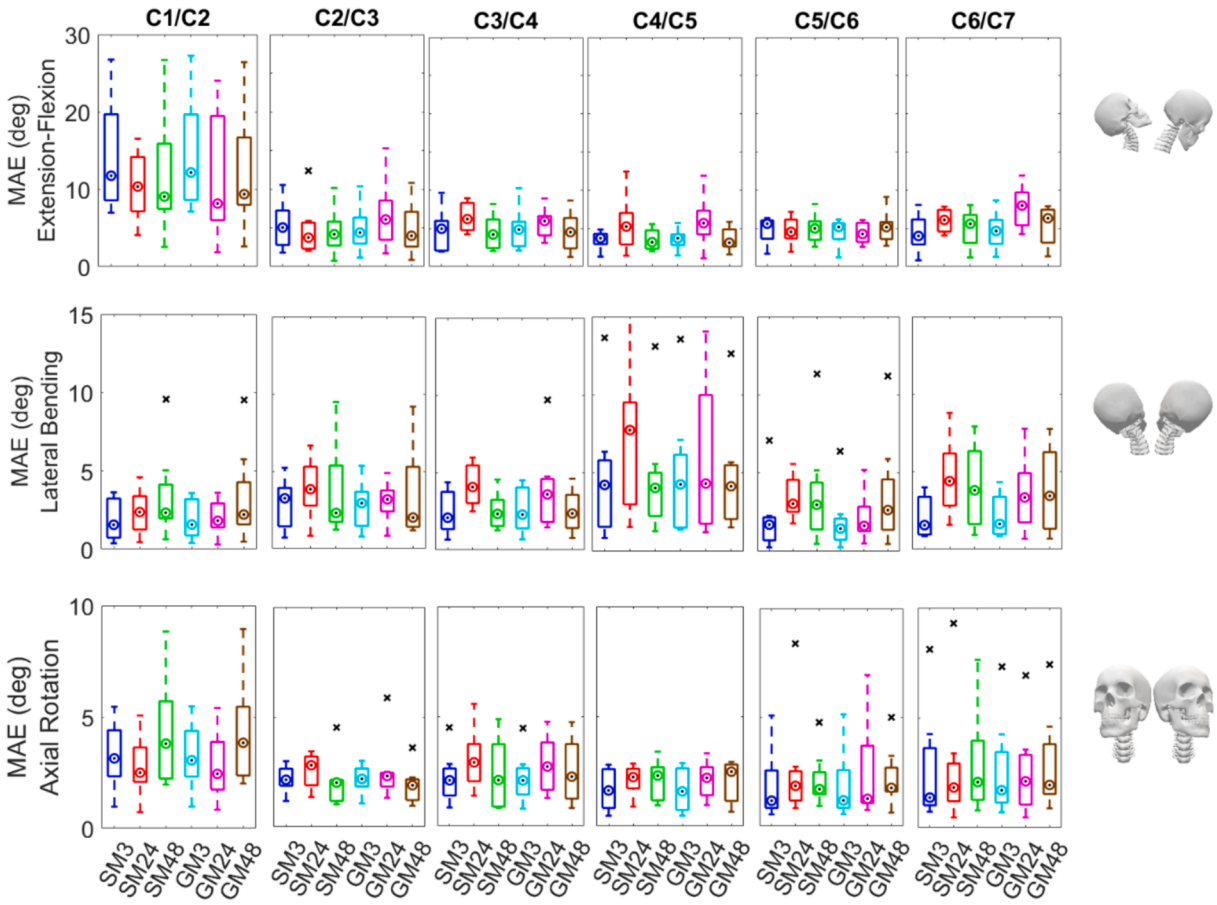

Fig. 5. Boxplots of rotational MAE (deg) during the extension-flexion tasks for 8 subjects and 6 models (scaled and generic, with differing degrees of freedom). The central marker indicates the median, the box indicates the interquartile range (25th-75th percentile), the whiskers are 1.5 times the interquartile range, and x's are considered outliers. Rows are from top to bottom: extension-flexion, lateral bending, axial rotation. Columns are from left to right: C1/C2 to C6/C7.

ranged from 0.88 to 1.13 for males and 0.86 to 1.05 for females.

3.2. Marker error

Fig. 2 shows the marker RMSE for two subjects over the range of motion. Over all subjects, the average marker RMSE of generic models over the ROM ranged from 13 mm to 32 mm, and the RMSE of scaled models ranged from 2 mm to 15 mm (Fig. 3).

Friedman test indicated significant differences among the models ($p<0.001$). Median [IQR] RMSE was lower for scaled vs. generic models for 3, 24 and 48 DOF, and marker error decreased with model complexity (Fig. 3). Among scaled models, the largest median [IQR] marker error was in the SM3 model (8.0 [6.6–8.7] mm), followed by SM24 (3.9 [2.9–5.0] mm), and lowest in the SM48 model (3.0 [2.2–3.9] mm). Among generic models, GM3 had the highest median marker RMS error (23.8 [20.7–26.9] mm), followed by GM24 (22.8 [20.2–25.0] mm), and GM48 (21.0 [16.9–23.5] mm). Marker error decreased more with scaling than with model complexity, but none of the decreases were significant ($0.008<p<0.039$) at the corrected $\alpha_c=0.0056$ for significance for 9 paired comparisons. All p-values are in Supplemental Table S2.

Markers on the trunk had greater error than markers on the head, especially for generic models (Supplemental Table S3). For generic models, trunk marker errors ranged from 6 mm to 75 mm, whereas head marker errors ranged from 1 mm to 24 mm over all subjects. For scaled models, trunk marker errors ranged from 1 mm to 21 mm, and head marker errors ranged from 1 mm to 15 mm.

3.3. Kinematics errors

Mean Absolute Error. Fig. 4 shows the extension-flexion angles for all 6 models with DSX data for each intervertebral level for two subjects. Across all subjects, the largest extension-flexion errors occurred at the C1/C2 level, with MAE as high as 30°, but the median values at C1/C2 were less than 12°. At other levels MAE did not exceed 15° in any subject, and the median was between 3° and 8° for all six types of models (Fig. 5, top row). Friedman test showed no differences for model scaling or complexity at any level ($0.115<p<0.787$; Supplemental Table S2).

Median MAE for lateral bending was 2° to 8° for all six types of models over all subjects, with maximum MAE of 15° (Fig. 5, middle row). Friedman test indicated significant differences at C3/C4 ($p=0.035$) and C4/C5 ($p=0.025$). However, post hoc tests did not find significant differences among any models ($0.008<p<1$) at the critical $\alpha_c=0.0056$ for 9 comparisons (Supplemental Table S2).

Median MAE for axial rotation was 1° to 4° for all six types of models over all subjects, with maximum errors up to 9° (Fig. 5, bottom row). Friedman test showed no differences for model scaling or complexity at any level ($0.084<p<0.744$; Supplemental Table S2).

Translational MAE was only compared between SM48 and GM48. MAE for anterior-posterior (AP) translational motion was greatest at C1/C2, with median [IQR] values of 13.9 [8.3–17.1] mm for SM48 and 15.1 [9.6–17.5] mm for GM48, and maximum values up to 19 mm (Fig. 6, top row). At all other levels, the median AP translational MAE was 3 to 4 mm, with maximum of 7 mm. MAE for superior-inferior (SI) translational motion was greatest at C2/C3, with median [IQR] values of 26.3 [24.6–29.2] mm for SM48 and 27.6 [25.4–29.3] mm for GM48, and maximum values up to 32.5 mm (Fig. 6, middle row). At all other levels,

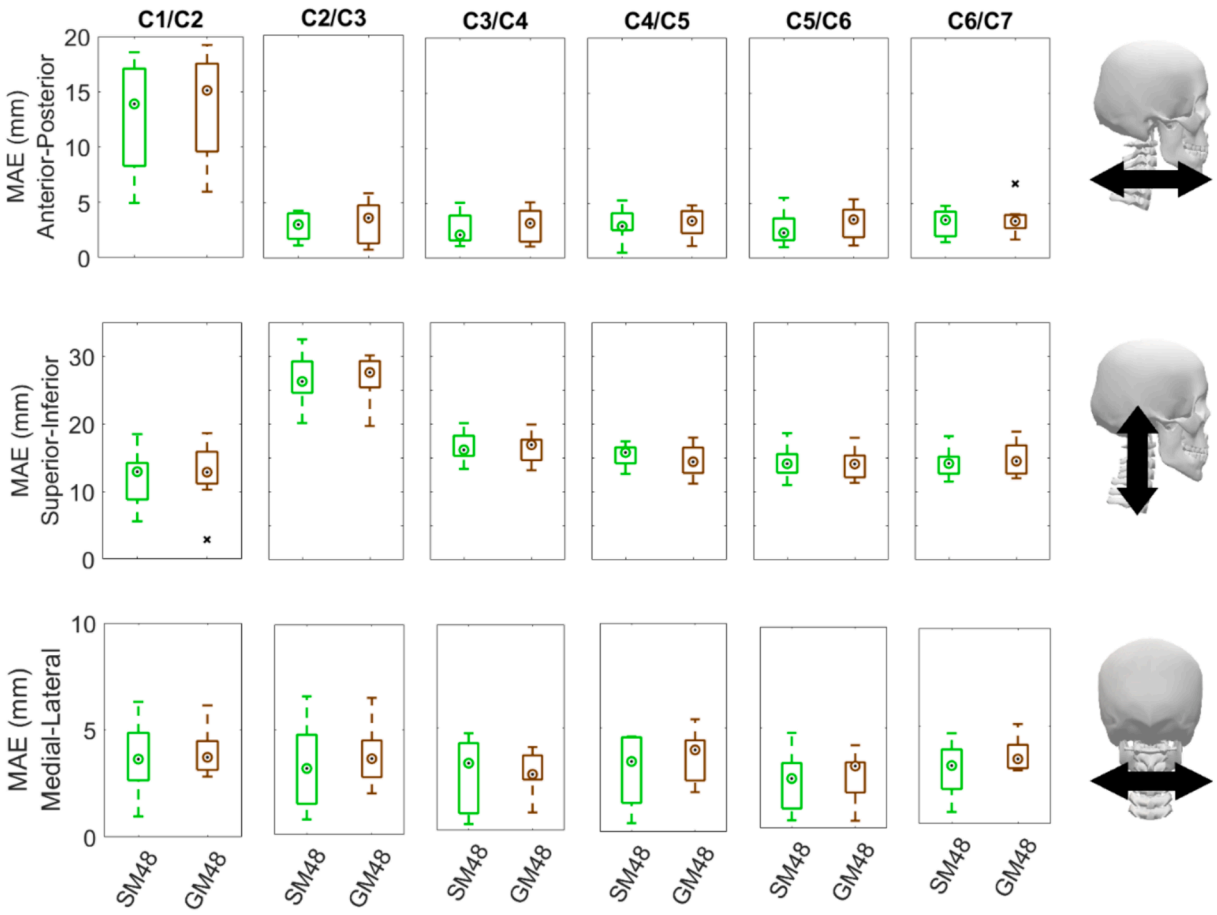

Fig. 6. Boxplots of translational MAE (mm) during the extension-flexion tasks for 8 subjects and 2 models (scaled and generic with 48 DOF). The central marker indicates the median, the box indicates the interquartile range (25th-75th percentile), the whiskers are 1.5 times the interquartile range, and x's are considered outliers. Rows are from top to bottom: anterior-posterior (AP), superior-inferior (SI) and medial–lateral (ML). Columns are from left to right: C1/C2 to C6/C7.

Table 2
Median [Interquartile Range] over 8 subjects for Pearson Correlation Coefficient of extension-flexion kinematics in degrees for 6 model types (scaled and generic with different levels of complexity), compared to DSX data. Although the Friedman Test found significant differences among models at all levels, pairwise comparisons did not find significance at the critical $\alpha_c = 0.0056$ for 9 comparisons (Supplemental Table S2).

	SM3	SM24	SM48	GM3	GM24	GM48
C1/C2	0.93 [0.89–0.95]	−0.79 [-0.91-(−0.24)]	0.92 [0.89–0.94]	0.93 [0.90–0.95]	−0.28 [-0.78–0.28]	0.91 [0.89–0.94]
C2/C3	0.96 [0.95–0.98]	0.62 [-0.30–0.78]	0.96 [0.94–0.98]	0.96 [0.95–0.98]	0.74 [0.23–0.90]	0.96 [0.95–0.98]
C3/C4	0.96 [0.95–0.98]	0.85 [0.83–0.91]	0.97 [0.95–0.98]	0.97 [0.95–0.98]	0.80 [0.25–0.89]	0.97 [0.95–0.98]
C4/C5	0.95 [0.92–0.97]	0.89 [0.80–0.93]	0.95 [0.90–0.97]	0.95 [0.92–0.97]	0.90 [0.24–0.92]	0.95 [0.91–0.97]
C5/C6	0.96 [0.95–0.98]	0.95 [0.87–0.97]	0.96 [0.94–0.97]	0.96 [0.95–0.98]	0.95 [0.83–0.97]	0.96 [0.94–0.98]
C6/C7	0.95 [0.92–0.98]	0.92 [0.86–0.96]	0.95 [0.90–0.98]	0.95 [0.92–0.98]	0.93 [0.86–0.96]	0.94 [0.92–0.99]

the median SI translational MAE was 13 to 17 mm, with maximum of 20 mm. Median MAE for ML translational motion was 3–4 mm, with maximum of 7 mm (Fig. 6, bottom row). There were no significant differences between SM48 and GM48 at any level for any translational direction ($p>0.157$; Supplemental Table S1).

Pearson Correlation Coefficient. 24DOF models had qualitatively different traces (Fig. 4). For extension-flexion motion, models with 3DOF or 48DOF had median r values over 0.9 at all levels (Table 2). However, models with 24DOF had negative r values at the C1/C2 level for most subjects, and some subjects had r values that were either negative or below 0.9 at all other levels. Friedman test indicated significant differences in the Pearson Correlation Coefficient at all levels ($p<0.018$; Supplemental Table S2). However, post hoc tests did not find significant differences among any models ($0.008<p<0.945$) at the critical α_c=0.0056 for 9 comparisons (Supplemental Table S2).

Range of Motion. There was considerable variability in intervertebral ROM among subjects. Because marker tracking was better for scaled models compared to generic models, and MAE and Pearson Correlation Coefficient were not affected by scaling, Friedman test was performed to compare DSX and scaled model (SM3/SM24/SM48) ROM. For six pairwise comparisons, the critical α_c=0.0083.

Extension-flexion ROM ranged from 5° to 28° in DSX measurements across all intervertebral levels (Fig. 7, top row). Friedman test indicated significant differences at all levels ($p<0.001$; Supplemental Table S2), except for C1/C2 ($p=0.175$) and C4/C5 ($p=0.058$). Median DSX ROM was significantly less than SM3 at C2/C3 (by 5°; $p=0.008$), significantly greater than SM24 at C3/C4 (by 6°; $p=0.008$) and significantly less than SM24 at C6/C7 (by 10°; $p=0.008$). DSX extension-flexion ROM was not

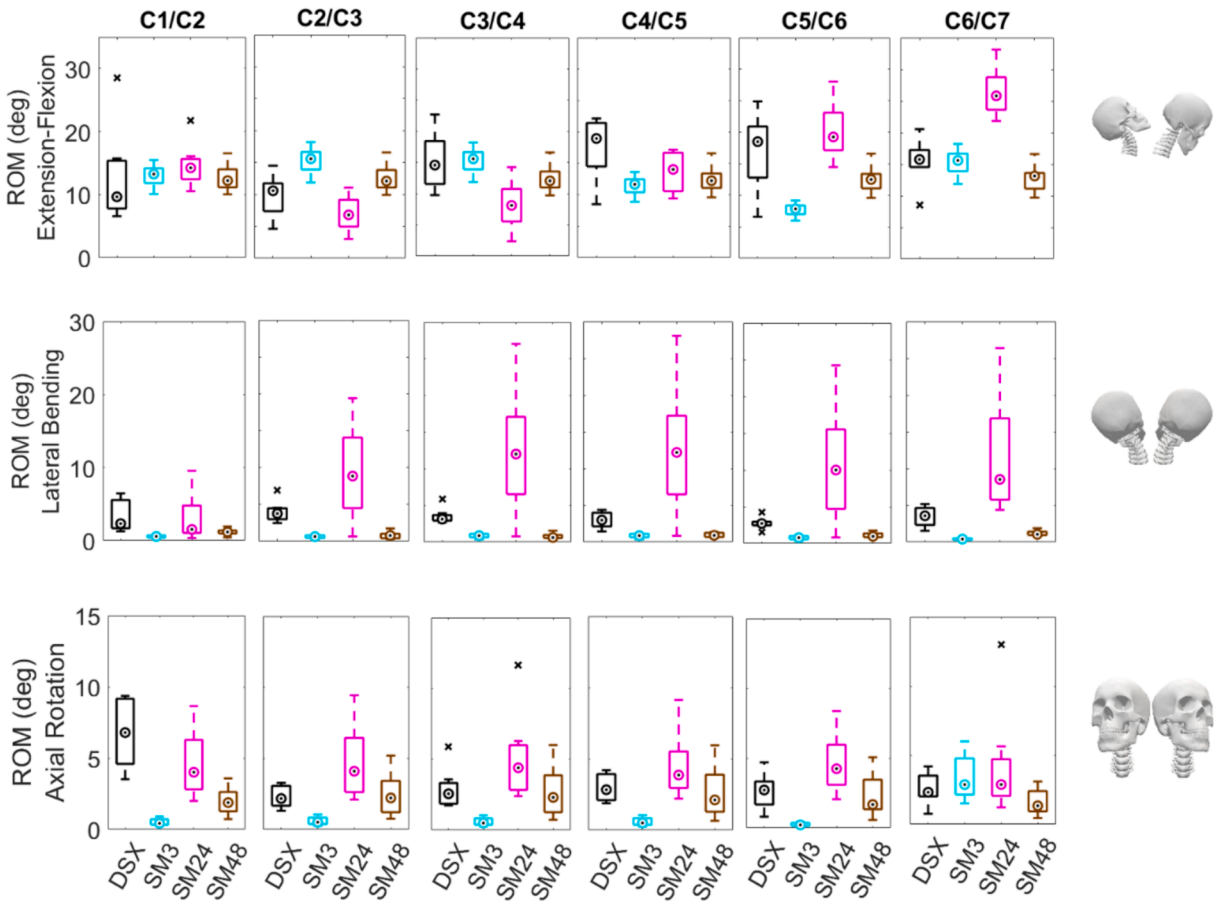

Fig. 7. Boxplots of rotational ROM (deg) during the extension-flexion tasks for 8 subjects, DSX data and 3 models (scaled, with 3, 24 or 48 degrees of freedom). The central marker indicates the median, the box indicates the interquartile range (25th-75th percentile), the whiskers are 1.5 times the interquartile range, and x's are considered outliers. Rows are from top to bottom: extension-flexion, lateral bending, axial rotation. Columns are from left to right: C1/C2 to C6/C7.

significantly different from SM48 at any levels at the critical α_c=0.0083 (0.039<p<0.109). There were significant differences among the 3 models at C2/C3, C3/C4, C5/C6 and C6/C7 (p=0.008). The largest differences were at C6/C7, where median ROM of SM24 (26°) was twice that of median ROM of SM48 by 13°, while the DSX ROM was 16°.

Lateral bending ROM ranged from 1° to 7° in DSX measurements across all intervertebral levels (Fig. 7, middle row). Friedman test indicated significant differences at all levels (p<0.005; Supplemental Table S2). Median DSX ROM was significantly greater than SM3 (p=0.008) and SM48 (p=0.008) by 2°-3° at all levels except C1/C2 in SM48 (p=0.023). Although median SM24 ROM was approximately 5°-9° larger than DSX ROM, this was only significant at C6/C7 (p=0.008) and not significantly different at any level between C1/C2 to C5/C6 (0.016<p<0.742). SM24 ROM was significantly greater (by 7°-11°) than SM3 at C6/C7 (p=0.008) and SM48 at C2/C3, C3/C4 and C6/C7 (p=0.008). SM3 ROM was significantly less than SM48 at C5/C6 and C6/C7 (p=0.008), but the difference was only 1°.

Axial rotation ROM ranged from 1° to 9° in DSX measurements across all intervertebral levels (Fig. 7, bottom row). Friedman test indicated significant differences at all levels (p<0.001; Supplemental Table S2) except C6/C7 (p=0.348). Median DSX ROM was significantly greater than SM3 at all levels C1/C2 through C5/C6 (p=0.008) by 2°-6°. DSX ROM was significantly less (by 2°) than SM24 at C2/C3 (p=0.008). DSX ROM was significantly greater (by 5°) than SM48 at C1/C2 (p=0.008). There were significant differences between SM3 and SM24 (by 4°) and between SM3 and SM48 (by 1°-2°) at all levels C1/C2 through C5/C6 (p=0.008). SM24 ROM was not significantly different from SM48 at any level (0.023<p<0.148).

Translational ROM was only compared for the 48DOF model relative to DSX with the Wilcoxon Signed Ranks test. Anterior-posterior (AP) translation ranged from 1 mm to 9 mm in DSX measurements across all intervertebral levels (Fig. 8, top row). The DSX AP translations were significantly less (by 2 to 5 mm) than SM48 AP translations at C1/C2, C3/C4 and C6/C7 levels (p<0.023; Supplemental Table S2). Superior-inferior (SI) translation ranged from 1 mm to 3 mm in DSX measurements across all intervertebral levels (Fig. 8, middle row), and the DSX SI translations were significantly less (by 2 to 3 mm) than SM48 SI translations at all levels (p<0.016). Medial-lateral (ML) translations ranged from 1 mm to 2 mm in DSX measurements across all intervertebral levels (Fig. 8, bottom row), and there were no significant differences between DSX and SM48 ML translations (p>0.313).

4. Discussion

The goals of this study were to reveal how scaling and model degrees of freedom affect the estimated intervertebral kinematics during extension-flexion motion. We hypothesized that scaling would decrease model error. Scaling did decrease the marker error in all cases (by 54% up to 92%), but the hypothesis was not supported by the rigorous statistical correction for 9 comparisons. Scaling had less effect on rotational kinematics (MAE or Pearson Correlation Coefficient). However, we believe that scaling models is important for accurately tracking external markers and should also improve linear measures such as muscle length or translational kinematics.

We also hypothesized that increasing model complexity would decrease model error. Although increasing model complexity decreased marker error in all subjects except one, the effect of model complexity was less than the effect of scaling, and the hypothesis was not supported

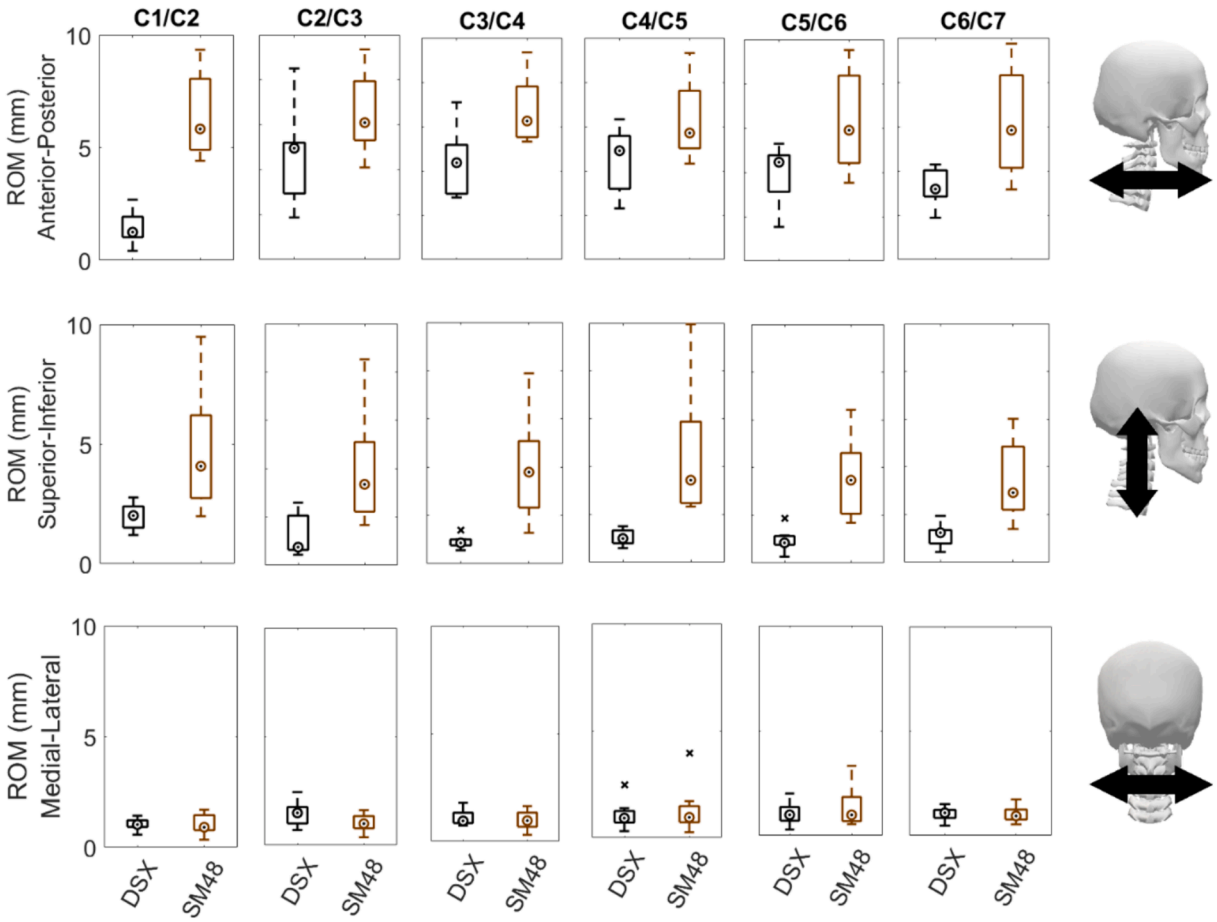

Fig. 8. Boxplots of translational ROM (mm) during the extension-flexion tasks for 8 subjects, DSX data and 2 models (scaled and generic, with 48 degrees of freedom). The central marker indicates the median, the box indicates the interquartile range (25th-75th percentile), the whiskers are 1.5 times the interquartile range, and x's are considered outliers. Rows are from top to bottom: anterior-posterior (AP), superior-inferior (SI) and medial–lateral (ML). Columns are from left to right: C1/C2 to C6/C7.

statistically for marker error. For MAE, significant differences were found at two levels in lateral bending but were not significant with correction for multiple comparisons. The effect of model complexity was evident when evaluating the shape of the curves, especially at C1/C2, where the 24DOF models were often out of phase and had low correlation to DSX kinematics. It is possible that constraining the relative amount of rotation at intervertebral joints resulted in more physiological motions for this simple extension-flexion task, whereas when the 24DOF model was allowed freedom of magnitude, errors propagated to the upper cervical spine, culminating in large differences in the C1/C2 kinematics. In general, 24DOF models did not track *in vivo* kinematics as well, but we did not find that 48DOF models were superior to models with 3DOF. ROM results showed that in some cases, the *difference* between model ROM and DSX ROM was larger than the actual DSX ROM, especially for 24DOF models. Because of the shape of the kinematics curves and ROM, we particularly urge care in using 24DOF models. For ROM, there were more significant differences amongst DSX data and models, in part because there were fewer comparisons and thus the critical p-value for ROM was higher than for MAE and Pearson Correlation Coefficient. For this reason, we have presented all the p-values in the Supplementary Materials so that readers can evaluate the graphical data in context.

The results of this study are generally consistent with other studies examining inverse kinematics, most of which have been done in the lower limb. For example, marker uncertainty of 2 cm could lead to uncertainty of 16 degrees in peak ankle angle (Uchida and Seth, 2022). Sensitivity analyses have also shown that soft tissue artifact had a greater effect on gait kinematics compared to bony landmark placement (Myers et al., 2015). In our study, soft tissue artifact is likely a large reason for the high marker error in trunk markers compared to skull markers, and thus the weights were lower for trunk markers (Supplemental Table S3). It is important to recognize, however, that limb kinematics can usually be measured with three markers per segment, but this is not the case for vertebral kinematics.

The challenge of spine inverse kinematics has been examined in at least two thoracolumbar models. Wang and colleagues used dynamic optimization for solving thoracolumbar spine kinematics (Wang et al., 2021). They had biplanar x-rays in the initial standing position and examined flexion–extension and lateral bending motions while seated with external markers. While standing upright, MAE of vertebral centers was 15–24 mm, and MAE of joint orientations was 4-6° compared to X-ray data. During motion, marker RMSE were 5–8 mm. Inverse kinematics solutions with 6DOF at each joint led to unrealistic vertebral positions, but even models with 3DOF at each joint had abnormally large ROM. Alemi and colleagues evaluated 7 kinematic constraint conditions, ranging from 3DOF to 51DOF (Alemi et al., 2021). Marker RMS error averaged 3–7 mm, with lowest error in the 51DOF models. However, intervertebral ROM of 51DOF was larger than constrained models or literature values (White and Panjabi, 1990), while the 3DOF models' ROM was smaller than literature. They concluded that adding additional DOF up to 6DOF improved marker error and segmental motion smoothness, but adding more DOF did not improve kinematic predictions. The findings of these two studies are generally in line with our results.

In this study, the base model used was a generic 50th percentile male model, although 4 of 8 subjects were female. The participant pool spanned a broad range of height and weight percentiles, and scaling factors, which were based on markers on the head and trunk, ranged from 0.86 to 1.13. Different marker placements or weights could result in slightly different scaling factors for the same subjects. These differences could affect muscle attachments and lengths, as well as the position of the joint centers, but a formal sensitivity analysis on marker placement and weights has not been conducted for spine inverse kinematics, as has been done in the lower limb (Myers et al., 2015, Uchida and Seth, 2022). We used the same kinematic constraints and (scaled) centers of rotation for male and female model kinematics because previous studies have not found sex differences in cervical spine posture or kinematics (DeSantis Klinich et al., 2004, Zheng, 2011). Although there can be large variation in cervical curvature (usually lordosis) in the "neutral" posture (among subjects, but not sex-specific), it was not possible to adjust subject-specific lordosis using solely external markers.

Although several studies have examined inverse kinematics, few studies have compared model-predicted kinematics to "gold standard" radiographic data as done here. We found that increasing model complexity improved marker error but not necessarily kinematics error. A limitation of our study is that DSX data were not available for skull-C1 or C7-T1 motion. Further, centers of rotation and motion distribution were consistent over the range of motion in these models; future studies should address the sensitivity of model predictions to centers of rotation and motion distribution that vary over the range of motion (Anderst et al., 2013, Anderst et al., 2015). In some cases, the MAE was similar to or larger than the DSX ROM, especially for lateral bending, axial rotation ROM (which were small), and for AP and SI translation ROM in the upper cervical spine. Intervertebral kinematics can be very important for calculating strains (and consequently stresses) in muscles, ligaments or discs, and the large superior-inferior MAE in particular can have a large effect on estimated strain. This study can provide input into future analyses of the effect of the kinematics error on other model-predicted values such as loads.

CRediT authorship contribution statement

Hamidreza Barnamehei: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yu Zhou:** Writing – review & editing, Formal analysis, Data curation. **Xudong Zhang:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Data curation, Conceptualization. **Anita N. Vasavada:** Writing – review & editing, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by a research grant from the Centers for Disease Control and Prevention/National Institute for Occupational Safety and Health (Grant No. R01OH010587).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiomech.2024.112302.

References

Alemi, M.M., Burkhart, K.A., Lynch, A.C., Allaire, B.T., Mousavi, S.J., Zhang, C., Bouxsein, M.L., Anderson, D.E., 2021. The Influence of Kinematic Constraints on Model Performance During Inverse Kinematics Analysis of the Thoracolumbar Spine. Front. Bioeng. Biotechnol. 9, 688041.

Amevo, B., Worth, D., Bogduk, N., 1991. Instantaneous axes of rotation of the typical cervical motion segments: a study in normal volunteers. Clin. Biomech. (Bristol, Avon) 6 (2), 111–117.

Anderst, W.J., Donaldson 3rd, W.F., Lee, J.Y., Kang, J.D., 2015. Cervical motion segment contributions to head motion during flexion\extension, lateral bending, and axial rotation. Spine J. 15 (12), 2538–2543.

Anderst, W. J., Baillargeon, E., Donaldson, W. F., 3rd, Lee, J. Y. and Kang, J. D. (2011). Validation of a noninvasive technique to precisely measure in vivo three-dimensional cervical spine movement. *Spine (Phila Pa 1976)* **36**(6): E393-400.

Anderst, W., Baillargeon, E., Donaldson, W., Lee, J. and Kang, J. (2013). Motion path of the instant center of rotation in the cervical spine during in vivo dynamic flexion-extension: implications for artificial disc design and evaluation of motion quality after arthrodesis. *Spine (Phila Pa 1976)* **38**(10): E594-601.

Begon, M., Andersen, M.S., Dumas, R., 2018. Multibody Kinematics Optimization for the Estimation of Upper and Lower Limb Human Joint Kinematics: A Systematized Methodological Review. J. Biomech. Eng. 140 (3).

Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G., 2007. OpenSim: open-source software to create and analyze dynamic simulations of movement. I.E.E.E. Trans. Biomed. Eng. 54 (11), 1940–1950.

DeSantis Klinich, K., Ebert, S.M., Van Ee, C.A., Flannagan, C.A.C., Prasad, M., Reed, M.P., Schneider, L.W., 2004. Cervical Spine Geometry in the Automotive Seated Posture: Variations with Age, Stature, Gender. Stapp Car Crash J. 48, 301–330.

Gordon, C.C., Blackwell, C.L., Bradtmiller, B., Parham, J.L., Barrientos, P., Paquette, S.P., Corner, B.D., Carson, J.M., Venezia, J.C., Rockwell, B.M., Mucher, M., Kristensen, S., 2014. 2012 anthropometric survey of U.S. Army Personnel: Methods and Summary Statistics. Development and Engineering Center, Natick, MA, U.S. Army Natick Soldier Research.

Johnson, G. M. (1998). The correlation between surface measurement of head and neck posture and the anatomic position of the upper cervical vertebrae. *Spine (Phila Pa 1976)* **23**(8): 921-927.

Myers, C.A., Laz, P.J., Shelburne, K.B., Davidson, B.S., 2015. A probabilistic approach to quantify the impact of uncertainty propagation in musculoskeletal simulations. Ann Biomed Eng 43 (5), 1098–1111.

Nevins, D.D., Zheng, L., Vasavada, A.N., 2014. Inter-individual variation in vertebral kinematics affects predictions of neck musculoskeletal models. J. Biomech. 47 (13), 3288–3294.

Penning, L., 1978. Normal movements of the cervical spine. AJR Am. J. Roentgenol. 130 (2), 317–326.

Uchida, T.K., Seth, A., 2022. Conclusion or Illusion: Quantifying Uncertainty in Inverse Analyses From Marker-Based Motion Capture due to Errors in Marker Registration and Model Scaling. Front. Bioeng. Biotechnol. 10, 874725.

Vasavada, A. N., Li, S. and Delp, S. L. (1998). Influence of muscle morphometry and moment arms on the moment-generating capacity of human neck muscles. *Spine (Phila Pa 1976)* **23**(4): 412-422.

Vasavada, A.N., Hughes, E., Nevins, D.D., Monda, S.M., Lin, D.C., 2018. Effect of Subject-Specific Vertebral Position and Head and Neck Size on Calculation of Spine Musculoskeletal Moments. Ann. Biomed. Eng. 46 (11), 1844–1856.

Wang, W., Wang, D., Falisse, A., Severijns, P., Overbergh, T., Moke, L., Scheys, L., De Groote, F., Jonkers, I., 2021. A Dynamic Optimization Approach for Solving Spine Kinematics While Calibrating Subject-Specific Mechanical Properties. Ann. Biomed. Eng. 49 (9), 2311–2322.

White, A.A., Panjabi, M.M., 1990. Clinical Biomechanics of the Spine. Lippincott Williams & Wilkins.

Zheng, L., 2011. Sex Differences in Human Neck Musculoskeletal Biomechanics and Modeling. Washington State University, Ph.D.

Zhou, Y., Chowdhury, S., Reddy, C., Wan, B., Byrne, R., Yin, W., Zhang, X., 2020. A state-of-the-art integrative approach to studying neck biomechanics in vivo. Sci. China Technol. Sci. 63, 1235–1246.