Immediate effects of EVA midsole resilience and upper shoe structure on running biomechanics: a machine learning approach

Andrea N Onodera [1,2], **Wilson PG Neto** [3], **Maria Isabel Roveri** [1], **Wagner R Oliveira** [2], **Isabel CN Sacco** [Corresp. 1]

[1] Physical Therapy, Speech and Occupational Therapy department, University of São Paulo, School of Medicine, Sao Paulo, Sao Paulo, Brazil

[2] Dass Nordeste Calçados e Artigos Esportivos Inc, Ivoti, Rio Grande do Sul, Brazil

[3] School of Engeneering & IT, Centro Universitário Ritter dos Reis, Porto Alegre, Rio Grande do Sul, Brazil

Corresponding Author: Isabel CN Sacco
Email address: icnsacco@usp.br

Background. Resilience of midsole material and the upper structure of the shoe are conceptual characteristics that can interfere in running biomechanics patterns. Artificial intelligence techniques can capture features from the entire waveform, adding new perspective for biomechanical analysis. This study tested the influence of shoe midsole resilience and upper structure on running kinematics and kinetics of non-professional runners by using feature selection, information gain, and artificial neural network analysis.

Methods. Twenty-seven experienced male runners (63±44 km/week run) ran in four-shoe design that combined two resilience-cushioning materials (low and high) and two uppers (minimalist and structured). Kinematic data was acquired by 6 infrared cameras at 300 Hz, and ground reaction forces were acquired by 2 force plates at 1200 Hz. We conducted a Machine Learning analysis to identify features from the complete kinematic and kinetic time series and from 42 discrete variables that had better discriminate the four shoes studied. And for that, we built an input data matrix of dimensions 1080 (10 trials x 4 shoes x 27 subjects) x 1254 (3 joints x 3 planes of movement x 101 data points + 3 vectors forces x 101 data points + 42 discrete calculated kinetic and kinematic features).

Results. The applied feature selection by information gain and artificial neural networks successfully differentiated the two resilience materials using 200(16%) biomechanical variables with an accuracy of 84.8% by detecting alterations of running biomechanics, and the two upper structures with an accuracy of 93.9%.

Discussion. The discrimination of midsole resilience resulted in lower accuracy levels than did the discrimination of the shoe uppers. In both cases, the ground reaction forces were among the 25 most relevant features.The resilience of the cushioning material caused significant effects on initial heel impact, while the effects of different uppers were distributed along the stance phase of running. Biomechanical changes due to shoe midsole resilience seemed to be subject-dependent, while those due to upper structure seemed to be subject-independent.

1   **Immediate effects of EVA midsole resilience and upper shoe structure on running**

2   **biomechanics: a machine learning approach**

3   Andrea N Onodera[1,2], Wilson P Gavião Neto[3], Maria Isabel Roveri[1], Wagner R Oliveira[2], Isabel

4   CN Sacco[1]

6   [1] University of São Paulo, School of Medicine, Physical Therapy, Speech and Occupational

7   Therapy dept., São Paulo, SP, Brazil.

8   [2] Dass Nordeste Calçados e Artigos Esportivos Inc, Ivoti, Rio Grande do Sul, Brazil.

9   [3] Ritter dos Reis University Center – UNIRITTER - Laureate International Universities, School of

10   Engeneering & IT and Master of Design, Porto Alegre, RS, Brazil.

12   **Isabel CN Sacco (Corresponding Author)**

13   E-mail: icnsacco@usp.br

15   **Abstract**

16   **Background.** Resilience of midsole material and the upper structure of the shoe are conceptual

17   characteristics that can interfere in running biomechanics patterns. Artificial intelligence

18   techniques can capture features from the entire waveform, adding new perspective for

19   biomechanical analysis. This study tested the influence of shoe midsole resilience and upper

20   structure on running kinematics and kinetics of non-professional runners by using feature

21   selection, information gain, and artificial neural network analysis.

22   **Methods.** Twenty-seven experienced male runners (63±44 km/week run) ran in four-shoe

23   design that combined two resilience-cushioning materials (low and high) and two uppers

24   (minimalist and structured). Kinematic data was acquired by 6 infrared cameras at 300 Hz, and

25   ground reaction forces were acquired by 2 force plates at 1200 Hz. We conducted a Machine

26   Learning analysis to identify features from the complete kinematic and kinetic time series and

27   from 42 discrete variables that had better discriminate the four shoes studied. And for that, we

28   built an input data matrix of dimensions 1080 (10 trials x 4 shoes x 27 subjects) x 1254 (3 joints

29   x 3 planes of movement x 101 data points + 3 vectors forces x 101 data points + 42 discrete

30   calculated kinetic and kinematic features).

31   **Results.** The applied feature selection by information gain and artificial neural networks

32   successfully differentiated the two resilience materials using 200(16%) biomechanical variables

33   with an accuracy of 84.8% by detecting alterations of running biomechanics, and the two upper

34   structures with an accuracy of 93.9%.

35   **Discussion.** The discrimination of midsole resilience resulted in lower accuracy levels than did

36   the discrimination of the shoe uppers. In both cases, the ground reaction forces were among

37 the 25 most relevant features. The resilience of the cushioning material caused significant

38 effects on initial heel impact, while the effects of different uppers were distributed along the

39 stance phase of running. Biomechanical changes due to shoe midsole resilience seemed to be

40 subject-dependent, while those due to upper structure seemed to be subject-independent.

41   **Introduction**

42   Sports shoes have many roles in running; among them, providing adequate impact-force

43   absorption [1, 2], stability for foot/ankle movements [3] and comfort [4]. These roles have been

44   the most studied in running and shoe biomechanics so far. Running shoes are basically

45   constituted by upper, midsole and sole. Among different possible combinations of these three

46   elements in the shoe construction, the upper is definitely the one more subject to variations in

47   its construction, such as color, model design, added elements and materials, and certainly, the

48   last two factors will have a particular influence on running biomechanics. Runners select a

49   comfortable running shoe using their own comfort criteria [5] and, because the shoe upper

50   maintains a large contact area with the foot, it would have a stronger influence over fit and

51   comfort, which in turn would impact in runner's kinematic and kinetic strategies during practice

52   and competitions. It has been demonstrated that a firmer foot contact with a shoe resulted in

53   lower loading rates due to a better coupling of foot-footwear, which optimizes the use of the

54   midsole impact absorption technology by favoring a better foot positioning inside the shoe [6].

55   Investigate the isolated influence of upper types in running biomechanics would contribute to a

56   more comprehensive and efficient approach in the shoe construction process.

57   Nevertheless, the most manipulated and studied shoe part in biomechanics is still the

58   midsole [7-10]. The majority of shoe companies invests a large amount of time, effort and

59   money development of damping materials technologies, such as gels, air, and springs for

60   supposedly improving sports performance. The midsole hardness is the most explored physical

61   characteristic of the midsole in biomechanical studies [1, 7, 10-13]. Running with hard shoes

62   resulted in same peak magnitude vertical GRF as running with soft ones [1, 12, 14] and faster

63   time to achieve the first peak [1]. Therefore, midsole hardness affected the loading rate but not

64   in a proportional rate [7]. Most of running kinematic changes due to the midsole hardness

65   occurs at the ankle joint [1, 10, 15] and some authors state that these different midsole

66   hardness lead to different impact perception by runners [7, 11], which in turn causes distinct

67   alteration on running mechanics [12, 13] and may mislead the real impact damping by shoe

68   midsoles [7, 11].

69   Apart from midsole hardness, resilience is also an important mechanical property of

70   midsole that has been seldom studied [8, 16, 17]. It represents the energy restored by the

71   cushioning material after an applied force ceases. Managing resilience while maintaining the

72   hardness of a polymeric foam is possible by adding different kinds of compounds to its formula.

73   Ethylene-vinyl acetate (EVA) added to elastomers could have ideal softness and high resilience

74   characteristics, would have a full-recovery capacity for the next foot step after a heel strike,

75   while a less resilient (more viscous) material would have the capacity of attenuating more

76   energies at initial loading cycles, easily achieving compression flattening after some cycles. It is

77   expected that different resilience materials would mainly reflect different initial impact forces,

78   because more resilient materials will quickly restore the cushioning property while less resilient

79   materials will take a little longer to restore the cushioning property [18]. Sinclair et al. [16] have

80   shown that running with shoes with energy return component resulted in greater tibial

81   acceleration peak, calcaneous eversion and internal tibia rotation compared to conventional

82   running shoes. In a later study, they have shown lower oxygen consumption and respiratory

83   exchange ratio with more resilient model of shoes [17]. Worobets et al. [8] manipulated only

84   the midsole materials, maintaining the upper structure, and also reported lower oxygen

85   consumption when running with a more compliant/resilient midsole condition. However, the

86   isolated effect of resilience changing in shoe midsole is still unknown in running biomechanics.

87   The majority of biomechanics studies vary the shoe model as a whole to investigate the

88   effects of various structural shoe properties and elements while running [19-22]. Such an

89   approach deeply interferes with an appropriate differentiation and interpretation of which

90   shoe characteristics most influence the kinetic and kinematic changes during running.

91   Identifying more precisely which shoe characteristics really matter for impact attenuation and

92   kinematic adaptation would facilitate the development of more effective products for athletes.

93   Nonetheless, individual's mechanical and neuromuscular adaptations to changes in

94   shoes are influenced by mechanical, neurophysiological, anatomical and even psychological

95   factors and, therefore, is likely to observed different individuals using different strategies in

96   response to changes in running shoes [9, 12]. Thereby, one may conclude that regardless the

97   type of shoe modification, the biomechanical responses observed may be subject-dependent.

98   We proposed to identify the most relevant biomechanical features to differentiate two

99   different midsole resiliencies and upper structures of shoes during running using an entirely

100   supervised approach based on machine learning (ML). This approach has been used to identify

101   crucial features and relevant patterns for classifying and predicting locomotor patterns in a

102   given condition [23-25]. Our strategy consisted of using Information Gain (IG) to select

103   important features and Artificial Neural Networks (ANN) to classify and predict the different

104   shoe resiliencies and upper types. Our assumptions were: (h1) low versus high resilient

105   cushioning effects on running kinematics and ground reaction forces are classifiable by using a

106   ML approach, (h2) structured versus minimalist upper effects on running kinematics and

107 ground reaction forces are classifiable by using a ML approach, and (h3) there are

108 biomechanical changes due to shoe midsole resilience or upper structure that are subject-

109 independent.

110

111 **Methods**

112 ***Subjects***

113   Twenty-seven experienced non-professional male runners (36.0±7.3 years old,

114 1.72±0.05 m, 73.9±6.2 kg, 62.9±43.8 km/week run, 7.5±7.1 years of practice) with a rearfoot

115 strike pattern and with no experience in minimalist shoe participated in this study. All runners

116 were free of injury or musculoskeletal pain according to the definition of Macera et al. [26]; had

117 no major foot or ankle postural alterations or deformities, excessive static pronation or

118 supination of the foot and ankle complex according to the Foot Posture Index [27]; and did not

119 present leg length discrepancy greater than 1cm. All subjects agreed to participate in the study

120 approved by the Ethics Committee of the School of Medicine of the University of Sao Paulo

121 (Ethical Application CEP-FMUSP: Protocol #054/14) and signed a written consent form.

122

123 ***Tested running shoes***

124   Four running shoes were especially developed by a local sportive shoe manufacturer.

125 The final masses of the 4 constructions were equivalent to avoid negative effects due to mass

126 differences [28]. All shoes were constructed using the same last, the same design of upper

127 pieces, midsole geometry, and outsole. The hardness of cushioning materials was fixed at 40

128   Asker C, measured by a durometer (GoTechAskerC, Taichung, Taiwan). The two shoe uppers

129   had the same design and shape, but the different structure and materials:

130   (1) **SU - structured upper**: 15 mm of soft foam in the heel collar and tongue, hard heel cup

131   involving the medial, lateral and posterior parts of the heel, synthetic pieces sewed in the vamp

132   and doubled fabric over the whole shoe (Figure 1A).

133   (2) **MU - minimalist upper**: light-weight mesh, tongue without foam, without heel cup, and

134   almost all pieces of the upper were connected by means of heat fusion (Figure 1B).

135   Both cushioning materials were made of ethylene-vinyl acetate (EVA); were inserted in

136   the same rearfoot area within the midsole; and had an oval shape of 10mm thickness, 50mm

137   width, and 70mm length. Resilience was assessed by vertical resiliometer (GoTech GT7042-V1,

138   Taichung, Taiwan). The midsoles were: (1) **LR - low resilience -** 5% of resilience (+/- 3%) (Figure

139   1C), and (2) **HR - high resilience** - 55% of resilience (+/- 3%) (Figure 1D).

140   **Figure 1**

141   The first tested condition (condition 1 - upper SU and cushioning material LR) was the

142   same for all runners; the other three testing conditions were randomized for each subject using

143   simple draw. The other three conditions were: condition 2 - upper SU and cushioning material

144   HR, condition 3 - upper MU and cushioning material LR, and condition 4 - upper MU and

145   cushioning material HR. The subjects were asked to lace their shoes tightly and comfortably, in

146   the same way they typically lace during their running practice.

147

148   ***Experimental protocol***

149   Kinematic data were acquired by six infrared cameras (VICON T-40, Oxford, UK). Sixteen

150   passive-reflexive markers (14 mm diameter) were fixed on both lower limbs (two anterior

151   superior iliac spines; two posterior iliac spines; two lateral epicondyles of the knees; two

152   markers over the lower lateral 1/3 surface of the thighs; two lateral malleolus; two markers

153   over the lower 1/3 of the shank; two second metatarsal heads; two posterior surface of

154   calcaneous at the same height above the plantar surface of the foot as the toes markers)

155   according to Plug´n Gait marker set, [29]. The two foot markers were fixed on the shoes

156   (second metatarsal heads and calcaneous) after deep palpation of bone prominences. The

157   laboratory coordinate system was established at one corner of one force plate, and all initial

158   calculations were based on this global coordinate system. In Nexus software (Vicon Nexus 1.7,

159   Oxford, UK), each data sample from each lower limb segment (foot, shank, and thigh) was

160   modeled as a rigid body with a local coordinate system that coincided with anatomical axes.

161   Translations and rotations of each segment were reported relative to neutral positions defined

162   during the static standing trial.

163   The program calculates the joint angles by means of a decomposition matrix based on

164   Cardan sequences and six degrees of freedom model. The decomposition matrix describes the

165   relationship between two local coordinate systems, one for each segment between which the

166   relative angle is determined. The joint kinematics was considered as the movement of the distal

167   segment in relation to the proximal; e.g. for determining the knee angle, the thigh was the

168   proximal segment and the shank the distal one. The movements occur around 3 different axes

169   which describe two definition of movement each: flexion/extension, abduction/adduction, and

170   internal/external rotation [30].

171   Ground reaction forces were acquired at 1200 Hz by two force plates (AMTI BP600600,

172 Watertown, USA) embedded in the center of a 25 m walkway. Acquisitions of kinematic and

173 force data were synchronized by a 64 multichannel Vicon MX Giganet Lab and A/D converter.

174   Running velocity was kept between 9.5 to 10.5 km/h (mean 10.1±0.5 km/h), monitored

175 by 2 photoelectrical sensors (Tecsistel Speed View, Novo Hamburgo, Brazil). Ten trials per

176 subject for each shoe condition were collected, resulting in 40 trials on the dominant limb. The

177 limb dominance was defined as the leg used to kick a soccer ball [31].

178

179 ***Biomechanical data analysis***

180   The marker coordinates were filtered using a 12 Hz zero-lag fourth-order low-pass

181 Butterworth filter. Force data was filtered with a 300 Hz zero-lag fourth-order low-pass

182 Butterworth filter. The angular and force data from initial contact to take-off were normalized

183 in stance time (interpolated 0-100%) and in magnitude by the body weight.

184   The 30 discrete kinematic features analyzed were: peak angles (degrees), angles at the

185 beginning of stance phase (degrees), instant of peak angle (seconds), range of motion from the

186 beginning of stance phase to peak angle (degrees), and final angle of stance phase (degrees);

187 for ankle, knee and hip joints; for sagittal and frontal planes of movement (5x3x2).

188   For vertical and antero-posterior forces, the magnitude of first vertical force peak (1VFP)

189 (body weight), time of 1VFP (milliseconds), magnitude of second vertical force peak

190 (2VFP)(body weight), time of 2VFP (milliseconds), loading rate (slope of 20% to 80% of 1VFP)

191 (body weight/ second), time of minimal vertical force in midstance (milliseconds), propulsion

192 rate (slope of curve between minimal vertical force in midstance and the 2VFP) (body weight/

193   second), minimum breaking antero-posterior force (body weight), breaking antero-posterior

194   impulse, median frequency of 1VFP (Hz), time of stance phase (milliseconds), and decay rate

195   (slope of curve from 2VFP to the end of stance phase) (body weight/ second) were calculated.

196   The whole interpolated time-series of all three planes of motion (sagittal, frontal and

197   transversal) and forces (vertical, antero-posterior and medio-lateral) were also analyzed.

198   Usually, there are high-dimensional and redundant datasets in cross-sectional studies that

199   investigated shoe effects in running biomechanics [13, 32], and therefore a careful feature

200   selection must be conducted to minimize bias and help identifying the biomechanical

201   parameters that is most influenced by shoe characteristics [13, 23, 32]. This was the main

202   reason why we chose to include both discrete and whole time-series points in the analysis. As

203   explained in the next sections, we adopted an approach based on ANN, which can receive large

204   numbers of data simultaneously and the pieces of data do not have to be isolated from each

205   other [33].

206

207   ***Machine learning approach***

208   IG and Artificial Neural Networks (ANN) are ML techniques that have been successfully

209   used for feature selection and classification in many areas. Whereas ANN have been used to

210   classify patterns in biomechanics studies [24, 34-36], IG has not been explored in biomechanics.

211   The IG is a supervised method that ranks variables individually without applying data

212   transformations, and it has the potential to facilitate the interpretation of the influence of a

213   single variable on the underlying discrimination process [37, 38]. IG is a methodology to select

214   the most relevant variables in a big and redundant set of variables.

215

216   Input variables and feature selection by IG

217   The 3D joint angular displacement time series was vectorized to a 1254 dimensional

218   vector (3 lower limb joints x 3 planes of angular displacement x 101 interpolated data points + 3

219   vectors of ground reaction forces x 101 interpolated data points + 42 discrete calculated kinetic

220   and kinematic features). An input data matrix M was then created (10 trials x 4 shoe condition x

221   27 subjects), resulting in a matrix dimension of 1080 x 1254. The lines of M represented each

222   subject trial in terms of 1254 variables, some of which may be more affected by specific

223   characteristics of shoe design. In this context, finding a small subset of variables is a desired

224   result [36], because that may indicate a more discriminative and less redundant subset of

225   features that would improve the results of the classification.

226   To assess the effects of shoe interventions, many studies have generated high-

227   dimensional and redundant datasets [13, 32], therefore, a variable selection stage must be

228   carefully conducted because it plays a critical role in minimizing bias and the influence of

229   intrinsic factors such as individual characteristics [13, 23, 32]. Frequently, redundant data

230   imposes challenges for understanding the phenomena of interest. To overcome these

231   challenges, approaches that capture features of the entire waveform instead of isolated

232   parameters add new perspectives to the analysis of the complex effects that a shoe may have

233   on a movement pattern [39]. The feature selection was conducted by analyzing the accuracy

234   levels of variables subsets. It was not practical to test all subsets of the 1254 variables/columns

235   available in M matrixes. We used IG to rank the variables in decreasing order of relevance. As a

236   supervised method, IG ranks the variables according to their discriminative power in separating

237   subjects' trials in terms of the two target variables, i.e., low vs high resilience and minimalist

238   versus structured upper. Diverse variables subsets were tested; they contained an increasing

239   number of variables (25, 50, 100, 150 and 200) and bigger subsets were systematically formed

240   by aggregating less relevant variables according to the IG criteria. This step was crucial to

241   determine the smallest number of variables that achieved accurate discrimination of the

242   resiliencies and upper structures. In this context, the greatest subset of variables involved in our

243   experiments included 200 variables, since our analysis indicated that 200 variables were

244   enough to evaluate the hypotheses of this study.

245

246   Classification Procedure

247   As a supervised classification approach, we adopted the standard k-fold-cross-validation, which

248   divides the subject trials into folds: $k-1$ folds for training and the remaining fold for testing. In

249   the training set, a subset of variables was selected by using IG, their values were scaled from -1

250   to 1, and then an ANN learning model was trained. The classification accuracy of the resulting

251   ANN model was then computed on the test fold. We reported results when the ANN method

252   achieved the best classification accuracy. We used a traditional feed-forward network with a

253   single hidden layer [40]. As it requires a parameter setting, which is still a research issue, we

254   perform an exhaustive searching through a subset of parameters values: the number of

255   neurons in the hidden layer was selected from the set $\{10,\ldots,75\}$; the learning rate $\in$

256     $\{0.05,\ldots,0.25\}$ and the number of training cycles $\in \{300,\ldots,1500\}$. To reduce the risk of

257   overfitting, we adopted the decay procedure [40], as implemented in the RapidMiner software.

258   We conducted the classification in two contexts:

259   **I.** a 4-fold-cross-validation for *each subject* (40 trials) to assess the existence of effects

260   from shoes conditions, finding one accuracy value to discriminate the shoe condition for each

261   subject and for each subset of features;

262   **II.** a standard 10-fold-cross-validation involving *all subjects*' trials to assess the existence

263   of subject-independent changes induced by the shoes interventions, finding one accuracy value

264   for each subset of features.

265   By comparing classification accuracies between the contexts I and II, it was possible to

266   analyze the subject-dependency of the results and provisionally evaluate a pattern induced by

267   different resilience and upper conditions. Classification accuracy of higher than 80% was

268   considered good [23]. The machine learning procedures were conducted in the software

269   RapidMiner (v.5.3.015, Dortmund, Germany).

270

271   **Results**

272   All 1254 variables were involved in the experiments of resiliencies and uppers. In both

273   cases, the accuracy of discriminating the effect of shoe conditions indicated that context I

274   outperformed context II.

275   The composition of the most relevant features subsets according to IG are detailed in

276   Tables 1 and 2 for the midsole resilience and upper structures comparisons, respectively.

277   Although forces have been the most discriminative variables for both midsole and upper, the

278   top 5 features for upper and resilience came from different components, respectively vertical

279   and medio-lateral forces.

280   **Table 1 and 2**

281

282   *Resilience effect*

283   In context II, 200 variables were sufficient to distinguish midsole resiliencies with an

284   accuracy of 84.8% (Figure 2, red line). The accuracies indicated that context I, which considers

285   only trials of a single subject, outperformed context II for all subsets of features (Figure 2, blue

286   line). A mean accuracy of 89.4% (±8.3%) to classify the two resiliencies was reached by

287   considering only the 25 most relevant features, while the best accuracy of 93.9% (±5.0) to

288   classify the two resiliencies was reached with the 150 most relevant features.

289   **Figure 2**

290   Among the 200 most relevant variables to discriminate between low and high resilience

291   cushioning materials, six of them were discrete biomechanical variables (4 ground-reaction

292   force variables and 2 kinematic variables) (Figure 3). The 5 most relevant variables came from

293   medio-lateral force, between 6% and 10% of stance phase.

294   **Figure 3**

295

296   *Upper structure effect*

297   The results indicated that the upper structures effects were less complex than the

298   cushioning materials ones. In the context II, accuracy higher than 85% was achieved by

299   considering only 25 variables to differentiate upper structures (Figure 4, red line). As in the case

300   of resiliencies, results on uppers shown that context I outperformed context II; it was possible

301   to obtain a mean accuracy of 93.4% (±4.8%) with 25 variables, and 95.6% (±3.8) with 150

302   variables (Figure 4, blue line).

303   **Figure 4**

304   Among the 200 most relevant variables to discriminate structured and minimalist

305   uppers 11 of them came from the discrete biomechanical variables (9 force variables and 2

306   kinematic variables) (Figure 5). The 5 most relevant variables to discriminate between uppers

307   were vertical forces from 11% to 14% of stance phase and the first peak.

308   **Figure 5**

309   **Discussion**

310   We proposed an entirely supervised approach based on ANN to distinguish the effects

311   of different midsole resiliencies and upper structures of shoes on running biomechanics. The

312   results confirmed our first and second hypotheses because the effects on running kinematics

313   and kinetics caused by low and high resilient cushioning midsoles and structured and minimalist

314   uppers were separable and classifiable by the adopted ML approach. IG was efficient in

315   selecting important features, as was confirmed by the proportionally slower increase in

316   classification accuracy with respect to increasing numbers of input features (Figure 4). The top

317   features in the IG rank may be considered the most responsible for distinguishing the effects of

318   the upper structures and midsole resiliencies.

319   When analyzing all subjects together, the methodology successfully differentiated the

320   two resiliencies with 84.8% accuracy using 200 variables and the two shoe uppers with 85.3%

321 accuracy using 25 variables, which is higher than the classification rate threshold of 80% chosen

322 by Hoerzer et al. [23]. Intra-subject analysis increased the classification accuracy for resiliencies

323 to a mean of 89.4% and for uppers using just 25 variables to a mean of 93.4%. This indicated

324 that the adopted ML analysis achieved more accuracy to identify different conditions

325 (cushioning materials and shoe uppers) within a given subject than within the set of all subjects,

326 which is consistent with the higher inter-subject variability.

327 Among the variables in the interpolated time series, the 25 most relevant features for

328 discriminating midsole resilience were mainly forces, ankle flexion-extension, and hip rotation

329 variables. Considering the discrete variables that differentiated the resiliencies, 4 were related

330 to vertical force (first peak and minimal force) and 2 were related to ankle kinematics

331 (dorsiflexion and eversion). Resilience causes significant effects on the initial impact of the heel

332 with the ground while running. The cushioning materials were inserted only under the heel part

333 of the shoe; considering we focused our study on rearfoot strikers, it was acceptable that the

334 resilience especially influenced the variables related to this first part of the stance phase.

335 However, resilience probably did not affect only this part of the stance for all runners.

336 The 200 most relevant features that discriminated the two resiliencies were distributed

337 in short time windows spread over all cycle periods of the kinematic and kinetic time series.

338 This does not mean, however, that these 200 features were equally relevant for all 27 subjects.

339 These short and sparse windows corresponded to individual patterns of biomechanical

340 responses, or a group of individuals with the same biomechanical responses, which leads us to

341 refute part of the third hypothesis and conclude that changes due to shoe midsole resilience

342 seemed to be subject-dependent. This means that is possible to discriminate the two

343   resiliencies with a fewer number of variables and with a higher accuracy if we analyze

344   individualized biomechanical data. In summary, each runner responded differently to resilience

345   in the cushioning materials, changing a different pool of biomechanical variables that represent

346   distinct motor strategies. Additionally, the most discriminative kinematic features were not

347   found at the same time periods as the most discriminative kinetic features in the stance phase,

348   suggesting that kinematics adjustments in the lower limbs caused by shoe changes might not

349   be influencing impact attenuation. This differs from what Hennig et al. [11] and Milani et al. [7]

350   suggested in their studies.

351   The upper structures classification had higher discriminative power than midsole

352   resiliencies, as was reflected in the higher accuracy levels of upper structure classification. The

353   25 most relevant features for discriminating upper structures were composed mainly of force

354   variables (both discrete variables and time series features) and sagittal ankle variables. From

355   the 200 features, we found that the most relevant biomechanical variables for the classification

356   of uppers were concentrated in the first and last third of the stance phase for all three-force

357   components, sagittal plane of ankle, and for all planes of knee. The exceptions were for the

358   frontal and transverse planes of the ankle, which had the central third as the most relevant part

359   for classification. Therefore, changes in shoe uppers seem to be more subject-independent.

360   Eleven discrete biomechanical features were among the most relevant for classifying

361   upper structures. Nine were force variables related to first and second vertical peaks, minimal

362   vertical force, and breaking antero-posterior force; two were from ankle kinematics. According

363   to the relevant discrete parameters, it seemed that force features had highest discriminative

364   power, so were more relevant for differentiating upper structures than kinematic variables and

365   they were temporally distributed across the stance phase. The flexibility of the upper (due to

366   different applied materials) could affect the flexibility of the foot inside the shoe, its own

367   capability to absorb impact by stretching the foot arches and generate the propulsion force, so

368   it is reasonable to assume that the force features across the entire stance phase are relevant

369   variables for distinguishing upper differences. The differences in structure probably affect

370   sensitivity to the ground, and consequently how runners modulate applied force to the ground.

371   As seen by Hagen and Hennig [6], firmer foot contact with a shoe could result in lower loading

372   rates due to a better coupling between the foot and shoe, which in turn facilitates the use of

373   the impact-absorbing technology of the midsole.

374   In a training context, it would be ideal if each runner performed a biomechanical

375   assessment of his or her running shoes. Then it would be possible to understand how this

376   external factor (running shoe) changes (or not) each runner's particular running mechanics.

377   Moreover, only few variables among the 1254 available (about 12% for different cushioning

378   materials and 4% for different shoe uppers) are evidently necessary to detect alterations of

379   running biomechanics with high accuracy (93.9% ± 5.0% and 94.3% ± 4.5%, respectively). This

380   result is plausible with the theory of "functional groups" [9, 23] and it is appropriate to

381   understand particular cases. Our results showed that runners have different responses to the

382   materials used in the shoes and a general conclusion arising from a heterogeneous sample may

383   lead to wrong outcomes. Further studies could analyze the influence of smaller sub-groups with

384   similar biomechanical responses to shoe characteristics. This procedure would be useful for

385   understanding how different resilience of cushioning material affects the running mechanics for

386   specific "types" of athletes. It was also possible to demonstrate that a biomechanical study of

387   sports shoes isolating characteristics to be tested provided more specificity in the

388   comprehension of the influence each part on running biomechanics.

389

390   **Conclusion**

391   The applied methodology based on feature selection by IG and classification by ANN

392   successfully differentiated the high and low resilience materials and the structured and

393   minimalist uppers with accuracies of higher than 85% using 200 features (16% of 1254 available

394   features). The classification of upper structures presented higher accuracy levels than that of

395   midsole resilience probably due to higher inter-individual variability, but in both cases, the

396   forces are among the 25 most relevant features subset. The ground reaction forces are the

397   most important features to differentiate midsole resilience and the resilience caused valuable

398   effects on initial heel impact while running. The different patterns of biomechanical response

399   chosen by runners to adapt to different resilience probably led to lower accuracy levels for this

400   classification. We can therefore conclude that biomechanical changes due to shoe midsole

401   resilience seem to be subject-dependent and changes due to upper structure seem to be

402   subject-independent.

403

404   **Acknowledgements**

405   We would like to thank Dass Nordeste Calçados e Artigos Esportivos Inc. for supplying

406   the shoes prototypes and the laboratory used in this study, and Alexandro Rodrigo da Rosa and

407   Jonas Schneider for designing and developing the shoes used in this study.

408

409   **References**

410   1.   Clarke TE, Frederick EC, Cooper LB. Effects of shoe cushioning upon ground reaction forces in
411   running. Int J Sports Med. 1983;4(4):247-51. doi: 10.1055/s-2008-1026043. PubMed PMID: 6654550.
412   2.   Hennig EM. Eighteen years of running shoe testing in Germany – a series of
413   biomechanical studies Footwear Science. 2011;3(2):71-81. Epub Oct 2011.
414   3.   Cheung RT, Wong MY, Ng GY. Effects of motion control footwear on running: a systematic
415   review. J Sports Sci. 2011;29(12):1311-9. doi: 10.1080/02640414.2011.591420. PubMed PMID:
416   21751855.
417   4.   Nigg BM. Biomechanics of sport shoes. 1st edition ed. Alberta, Canada: Topline Printing Inc.;
418   2010. 300 p.
419   5.   Nigg BM, Baltich J, Hoerzer S, Enders H. Running shoes and running injuries: mythbusting and a
420   proposal for two new paradigms: 'preferred movement path' and 'comfort filter'. Br J Sports Med. 2015.
421   doi: 10.1136/bjsports-2015-095054. PubMed PMID: 26221015.
422   6.   Hagen M, Hennig EM. Effects of different shoe-lacing patterns on the biomechanics of running
423   shoes. J Sports Sci. 2009;27(3):267-75. doi: 10.1080/02640410802482425. PubMed PMID: 19156560.
424   7.   Milani TL, Hennig EM, Lafortune MA. Perceptual and biomechanical variables for running in
425   identical shoe constructions with varying midsole hardness. Clin Biomech (Bristol, Avon).
426   1997;12(5):294-300. PubMed PMID: 11415737.
427   8.   Worobets J, Wannop JW, Tomaras E, Stefanyshyn D. Softer and more resilient running shoe
428   cushioning properties enhance running economy. Footwear Science. 2014;6(3):147-53. Epub 17 Jun
429   2014.
430   9.   Nigg BM, Stefanyshyn D, Cole G, Stergiou P, Miller J. The effect of material characteristics of
431   shoe soles on muscle activation and energy aspects during running. J Biomech. 2003;36(4):569-75.
432   PubMed PMID: 12600347.
433   10.   Maclean CL, Davis IS, Hamill J. Influence of running shoe midsole composition and custom foot
434   orthotic intervention on lower extremity dynamics during running. J Appl Biomech. 2009;25(1):54-63.
435   PubMed PMID: 19299830.
436   11.   Hennig EM, Valiant GA, Liu Q. Biomechanical variables and the perception of cushioning for
437   running in various types of footwear. Journal of Applied Biomechanics. 1996;12:143 - 50.
438   12.   Kersting UG, Brüggemann GP. Midsole material-related force control during heel-toe running.
439   Res Sports Med. 2006;14(1):1-17. doi: 10.1080/15438620500528158. PubMed PMID: 16700401.
440   13.   Nigg BM, Baltich J, Maurer C, Federolf P. Shoe midsole hardness, sex and age effects on lower
441   extremity kinematics during running. J Biomech. 2012;45(9):1692-7. doi:
442   10.1016/j.jbiomech.2012.03.027. PubMed PMID: 22507350.
443   14.   Nigg BM, Bahlsen HA, Luethi SM, Stokes S. The influence of running velocity and midsole
444   hardness on external impact forces in heel-toe running. J Biomech. 1987;20(10):951-9. PubMed PMID:
445   3693376.
446   15.   Hardin EC, van den Bogert AJ, Hamill J. Kinematic adaptations during running: effects of
447   footwear, surface, and duration. Med Sci Sports Exerc. 2004;36(5):838-44. PubMed PMID: 15126719.
448   16.   Sinclair J, Franks C, Fau-Goodwin J, Naemi R, Chockalingam N. Influence of footwear designed to
449   boost energy return on the kinetics and kinematics of running compared to conventional running shoes.
450   Comparative Exercise Physiology. 2014;10:199-206.
451   17.   Sinclair J, Mcgrath R, Brook O, Taylor PJ, Dillon S. Influence of footwear designed to boost
452   energy return on running economy in comparison to a conventional running shoe. J Sports Sci.
453   2016;34(11):1094-8. doi: 10.1080/02640414.2015.1088961. PubMed PMID: 26367197.

454   18.   Sun PC, Wei HW, Chen CH, Wu CH, Kao HC, Cheng CK. Effects of varying material properties on
455   the load deformation characteristics of heel cushions. Med Eng Phys. 2008;30(6):687-92. doi:
456   10.1016/j.medengphy.2007.07.010. PubMed PMID: 17888713.
457   19.   Azevedo AP, Brandina K, Bianco R, Oliveira VH, Souza JR, Mezencio B, et al. Effects of replica
458   running shoes upon external forces and muscle activity during running. J Sports Sci. 2012;30(9):929-35.
459   doi: 10.1080/02640414.2012.682080. PubMed PMID: 22530777.
460   20.   Braunstein B, Arampatzis A, Eysel P, Brüggemann GP. Footwear affects the gearing at the ankle
461   and knee joints during running. J Biomech. 2010;43(11):2120-5. doi: 10.1016/j.jbiomech.2010.04.001.
462   PubMed PMID: 20462583.
463   21.   McNair PJ, Marshall RN. Kinematic and kinetic parameters associated with running in different
464   shoes. Br J Sports Med. 1994;28(4):256-60. PubMed PMID: 7894957; PubMed Central PMCID:
465   PMCPMC1332086.
466   22.   Dixon SJ. Use of pressure insoles to compare in-shoe loading for modern running shoes.
467   Ergonomics. 2008;51(10):1503-14. doi: 10.1080/00140130802239562. PubMed PMID: 18803091.
468   23.   Hoerzer S, von Tscharner V, Jacob C, Nigg BM. Defining functional groups based on running
469   kinematics using Self-Organizing Maps and Support Vector Machines. J Biomech. 2015;48(10):2072-9.
470   doi: 10.1016/j.jbiomech.2015.03.017. PubMed PMID: 25869722.
471   24.   Muniz AM, Liu H, Lyons KE, Pahwa R, Liu W, Nobre FF, et al. Comparison among probabilistic
472   neural network, support vector machine and logistic regression for evaluating the effect of subthalamic
473   stimulation in Parkinson disease on ground reaction force during gait. J Biomech. 2010;43(4):720-6. doi:
474   10.1016/j.jbiomech.2009.10.018. PubMed PMID: 19914622.
475   25.   Schöllhorn WI, Nigg BM, Stefanyshyn DJ, Liu W. Identification of individual walking patterns
476   using time discrete and time continuous data sets. Gait Posture. 2002;15(2):180-6. PubMed PMID:
477   11869912.
478   26.   Macera CA, Pate RR, Powell KE, Jackson KL, Kendrick JS, Craven TE. Predicting lower-extremity
479   injuries among habitual runners. Arch Intern Med. 1989;149(11):2565-8. PubMed PMID: 2818115.
480   27.   Redmond AC, Crosbie J, Ouvrier RA. Development and validation of a novel rating system for
481   scoring standing foot posture: the Foot Posture Index. Clin Biomech (Bristol, Avon). 2006;21(1):89-98.
482   doi: 10.1016/j.clinbiomech.2005.08.002. PubMed PMID: 16182419.
483   28.   Frederick EC. Physiological and ergonomics factors in running shoe design. Appl Ergon.
484   1984;15(4):281-7. PubMed PMID: 15676526.
485   29.   Kadaba MP, Ramakrishnan HK, Wootten ME. Measurement of lower extremity kinematics
486   during level walking. J Orthop Res. 1990;8(3):383-92. doi: 10.1002/jor.1100080310. PubMed PMID:
487   2324857.
488   30.   Hamill J, Selbie WS, Kepple TM. Three-Dimensional Kinematics, Research Methods in
489   Biomechanics. Human Kinetics. 2nd ed ed2014.
490   31.   Greenberger HB, Paterno MV. Relationship of knee extensor strength and hopping test
491   performance in the assessment of lower extremity function. J Orthop Sports Phys Ther. 1995;22(5):202-
492   6. doi: 10.2519/jospt.1995.22.5.202. PubMed PMID: 8580947.
493   32.   Maurer C, Federolf P, von Tscharner V, Stirling L, Nigg BM. Discrimination of gender-, speed-,
494   and shoe-dependent movement patterns in runners using full-body kinematics. Gait Posture.
495   2012;36(1):40-5. doi: 10.1016/j.gaitpost.2011.12.023. PubMed PMID: 22304784.
496   33.   Barton JG, Lees A. An application of neural networks for distinguishing gait patterns on the basis
497   of hip-knee joint angle diagrams. Gait & Posture. 1997;5(1):28-33. doi: http://dx.doi.org/10.1016/S0966-
498   6362(96)01070-3.
499   34.   Begg R, Kamruzzaman J. A machine learning approach for automated recognition of movement
500   patterns using basic, kinetic and kinematic gait data. J Biomech. 2005;38(3):401-8. doi:
501   10.1016/j.jbiomech.2004.05.002. PubMed PMID: 15652537.

502   35.   Fischer A, Do M, Stein T, Asfour T, Dillmann R, Schwameder H. Recognition of individual
503   kinematic patterns during walking and running - A comparison of artificial neural networks and support
504   vector machines. International Journal of Computer Science in Sport. 2011;10(1):63-7.
505   36.   Rupérez MJ, Martín-Guerrero JD, Monserrat C, Alcañiz M. Artificial neural networks for
506   predicting dorsal pressures on the foot surface while walking. Expert Systems with Applications.
507   2012;39:5349 - 57. doi: 10.1016/j.eswa.2011.11.050.
508   37.   Guyon I, Gunn S, Nikravesh M, Zadeh LA. Feature extraction: Foundations and applications. New
509   York: Springer-Verlag Berlin Heidelberg; 2006. 778 p.
510   38.   Witten IH, Frank E, Hall MA. Data Mining: Practical machine learning tools and techniques. 3rd.
511   ed. San Diego, CA, USA: Morgan Kaufmann; 2011. 664 p.
512   39.   Muniz AM, Nadal J. Application of principal component analysis in vertical ground reaction force
513   to discriminate normal and abnormal gait. Gait Posture. 2009;29(1):31-5. doi:
514   10.1016/j.gaitpost.2008.05.015. PubMed PMID: 18640040.
515   40.   Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York, NY, USA:
516   Springer New York Inc.; 2001.

517

IG rank subset for midsole resilience

Table 1 - Number of variables of each feature subset according to IG rank for midsole resilience comparison.

1   Table 1 - Number of variables of each feature subset according to IG rank for midsole resilience

2   comparison.

RESILIENCE COMPARISON	25 most relevant variables	50 most relevant variables	100 most relevant variables	150 most relevant variables	200 most relevant variables
Medio-lateral GRF	9	11	14	17	18
Sagittal Ankle	6	9	16	18	25
Transversal Hip	3	12	17	33	39
Vertical GRF	2	4	6	12	16
Antero-posterior GRF	2	2	7	11	17
Discrete Force Variables	2	3	4	4	4
Discrete Kinematic Variables	1	1	1	2	2
Frontal Knee		7	12	12	14
Sagittal Hip		1	19	35	38
Transversal Knee			4	5	12
Frontal Ankle				1	5
Frontal Hip					6
Transversal Ankle					4
% of total features (1254)	2%	4%	8%	12%	16%

IG rank subset for upper structure

Table 2 - Number of variables of each feature subset according to IG rank for upper

structures comparison.

1   Table 2 - Number of variables of each feature subset according to IG rank for upper structures

2   comparison.

UPPER STRUCTURE COMPARISON	25 most relevant variables	50 most relevant variables	100 most relevant variables	150 most relevant variables	200 most relevant variables
Antero-posterior GRF	10	19	33	38	40
Vertical GRF	8	12	16	19	21
Discrete Force Variables	4	4	4	6	9
Sagittal Ankle	2	10	20	22	28
Medio-lateral GRF	1	4	12	18	22
Transversal Knee		1	7	19	25
Frontal Ankle			5	13	26
Transversal Ankle			2	9	10
Discrete Kinematic Variables			1	2	2
Sagittal Knee				4	7
Transversal Hip					6
Frontal Knee					4
% of total features (1254)	2%	4%	8%	12%	16%

Figure 1

Image of shoes prototypes used in the experiment

Figure 1 – Illustration of testing shoes. (A) Structured Upper condition (SU). (B) Minimalist Upper condition (MU). (C) Low Resilience cushioning material condition (LR). (D) High Resilience cushioning material condition (HR).

Figure 2

Accuracy levels to discriminate midsole resilience materials in various contexts considering different subsets of variables

Figure 2 – Mean accuracy and standard deviation for each subset of input variables with the highest IG values to discriminate resilience materials. Red line represents the context II and considers all subjects together. Blue line represents the context I and considers each subject in isolation.
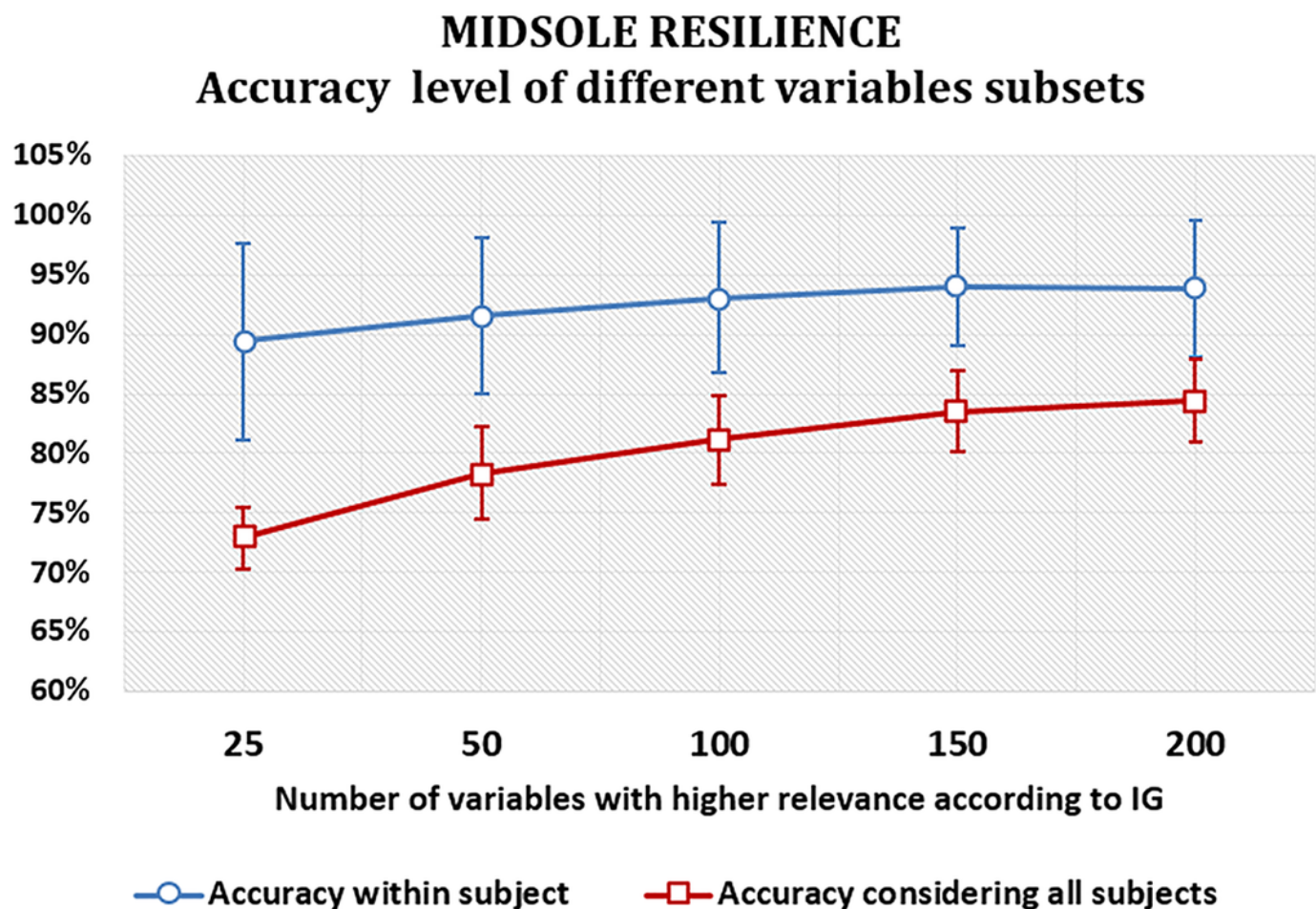
Figure 3

Ground reaction force and Kinematics time-series during running with different resilience midsoles.

Figure 3 - (A) Mean time series of ground reaction force for different resilience of cushioning materials. (B) Mean time series of joints kinematics in all planes of motion for different resilience materials. Blue lines represent the low resilience cushioning condition and red dotted lines represent the high resilience cushioning condition. The 200 highest IG variables are highlighted in the yellow boxes.
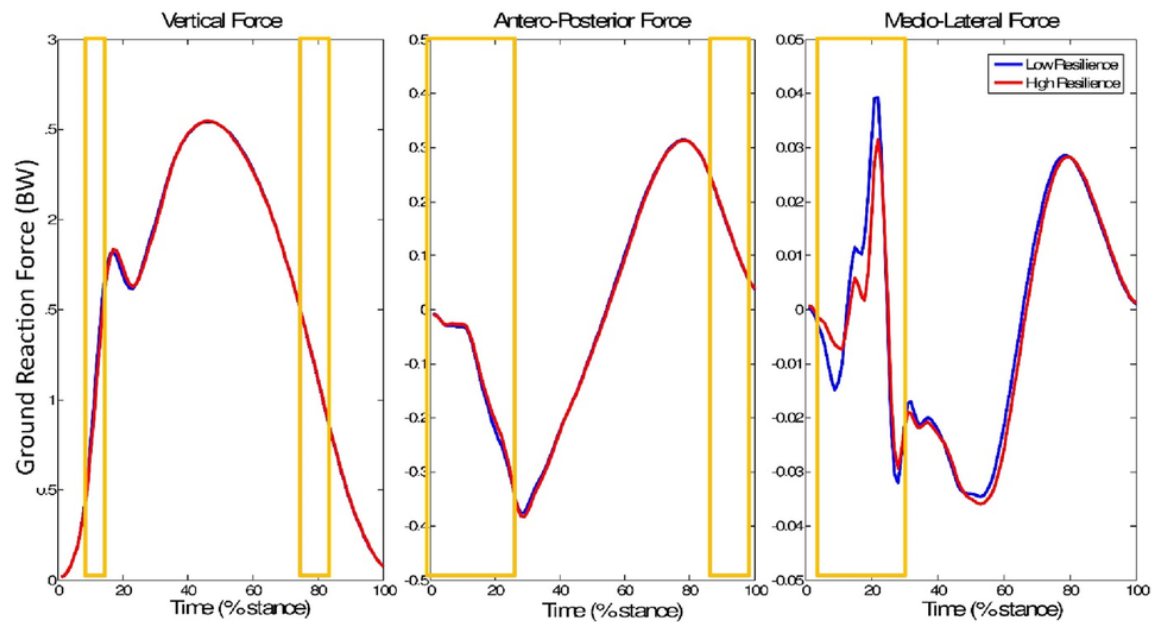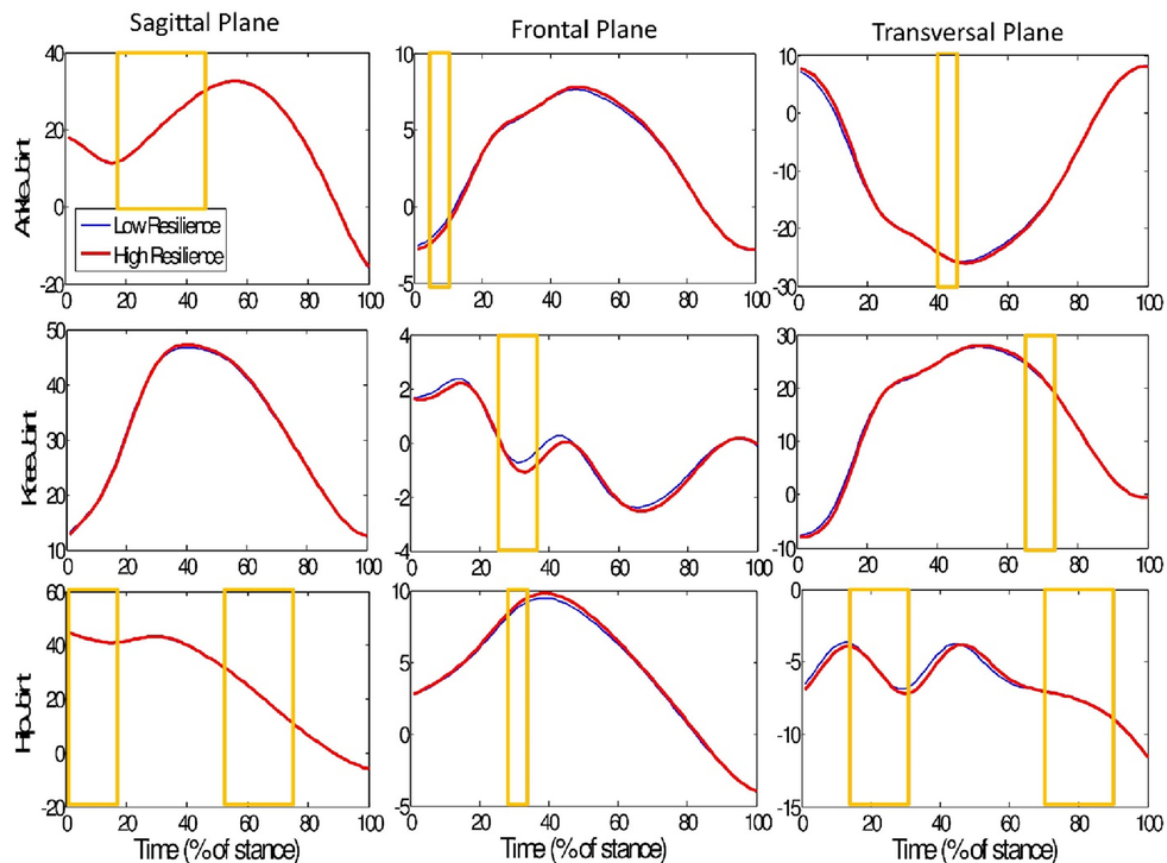
Figure 4

Accuracy levels to discriminate upper strustures in various contexts considering different subsets of variables

Figure 4 – Mean accuracy and standard deviation for each subset of input variables with the highest IG values to classify upper structures. Red line represents the context II and considers all subjects together. Blue line represents the context I and considers each subject in isolation.
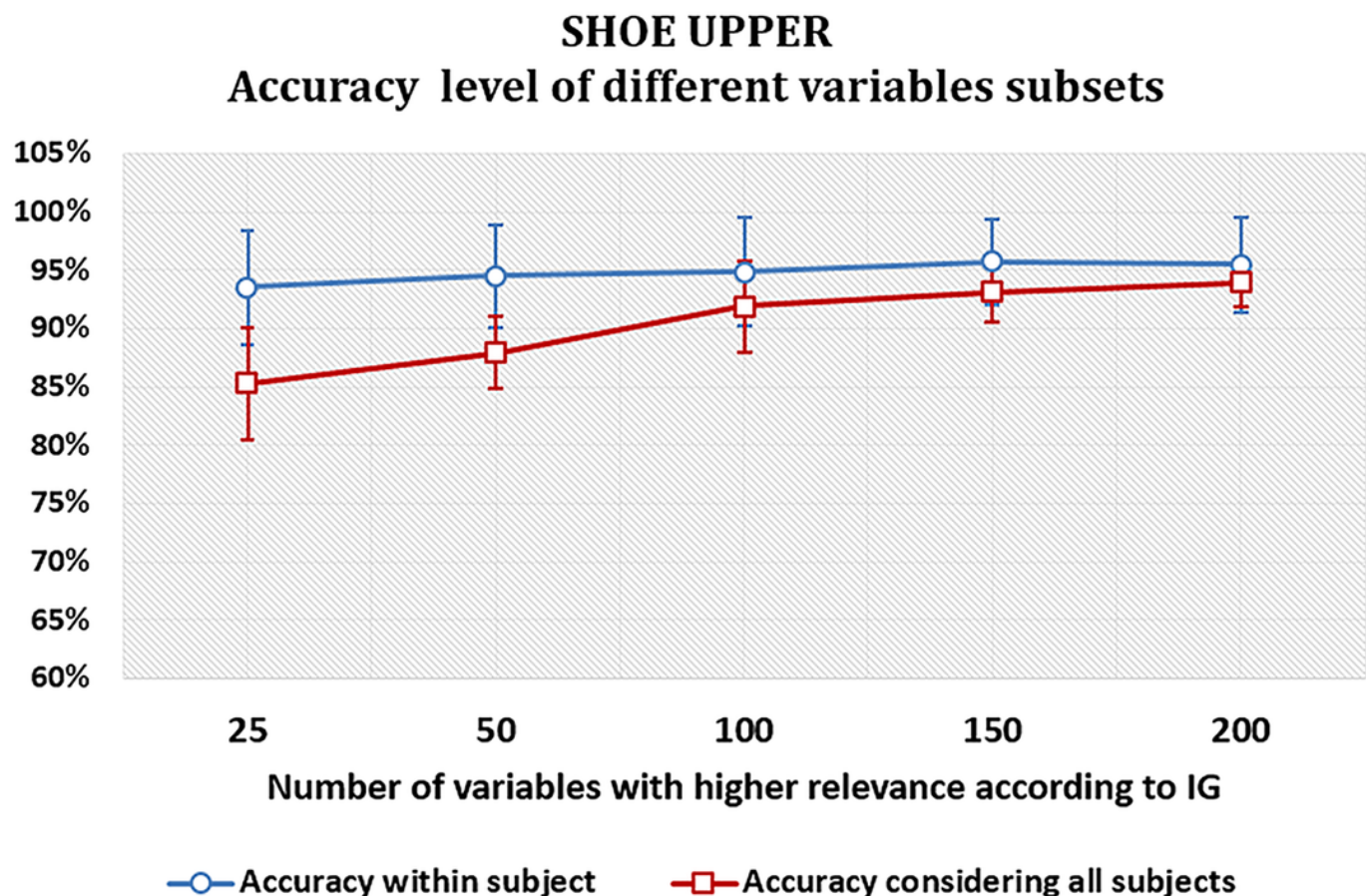
Figure 5

Ground reaction force and kinematics time-series during running with different shoe upper structures.

Figure 5 - (A) Mean time series of ground reaction force for different shoe upper structures. (B) Mean time series of joints kinematics in all planes of motion for different shoe upper structures. Black lines represent the structured upper condition and Pink dotted lines represent the minimalist upper condition. The 200 highest IG variables are highlighted in the yellow boxes.