Supplementary Material: Machine Learning Workflows for Motion Capture-driven Biomechanical Modelling

May 4, 2025

Contents

1 Hyperparameters explored for various machine learning techniques and optimal hyperparameters

Here, we report the optimal hyperparameters obtained for each machine learning (ML) technique using cross-validation. The full set of hyperparameters explored is provided on https://github.com/rahul2512/MoCap_ML_workflow/tree/main/hyperparameters. Note that the output variable normalisation (norm out) was also treated as a hyperparameter. For Convolutional and RNN models, a sliding window approach with a window size of 5 was used to process the input.

Table 1: Hyperparameters choices explored for Linear model (LM) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	p	regularizer val	NN variant	norm out
Joint angles (SE)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1
Joint moments (SE)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1
Joint reaction forces (SE)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1
Joint angles (SN)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1
Joint moments (SN)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1
Joint reaction forces (SN)	Adam	random normal	256	100	linear	200	0	rmse	rmse	0.001	0	0	LM	1

Table 2: Hyperparameters choices explored for Feedforward Neural Networks (FFNN) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	p	regularizer val	NN variant	norm out
Joint angles (SE)	Adam	glorot normal	64	100	relu	400	4	rmse	mse	0.001	0.100	0	NN	1
Joint moments (SE)	Adam	glorot normal	64	100	relu	400	4	rmse	mse	0.001	0.100	0	NN	1
Joint reaction forces (SE)	Adam	glorot normal	64	50	relu	400	4	rmse	mse	0.001	0.100	0	NN	1
Joint angles (SN)	Adam	glorot normal	64	100	relu	400	4	rmse	mse	0.001	0.100	0	NN	1
Joint moments (SN)	Adam	glorot normal	64	100	relu	600	2	rmse	mse	0.001	0.100	0	NN	1
Joint reaction forces (SN)	Adam	glorot normal	64	100	relu	400	4	rmse	mse	0.005	0.100	0	NN	0

Table 3: Hyperparameters choices explored for Random Forests (RF) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	n estimators	max features	max depth	min samples split	min samples leaf	bootstrap	criterion	norm out
Joint angles (SE)	50	log2	10	5	1	True	absolute error	0
Joint moments (SE)	50	log2	20	2	2	True	absolute error	0
Joint reaction forces (SE)	50	log2	10	5	1	True	absolute error	0
Joint angles (SN)	50	log2	10	2	1	True	absolute error	0
Joint moments (SN)	50	log2	20	2	2	True	absolute error	0
Joint reaction forces (SN)	50	log2	30	2	1	True	absolute error	0

Table 4: Hyperparameters choices explored for Extreme Gradient Boosting (XGBR) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	n estimators	learning rate	max depth	objective	alpha	lambda1	norm out
Joint angles (SE)	200	0.010	50	reg:squarederror	0.100	0.100	0
Joint moments (SE)	200	0.010	50	reg:squarederror	0.100	0.100	0
Joint reaction forces (SE)	200	0.010	50	reg:squarederror	0.100	0.100	0
Joint angles (SN)	200	0.010	50	reg:squarederror	0.100	0.100	0
Joint moments (SN)	100	0.010	20	reg:squarederror	0.200	0.100	0
Joint reaction forces (SN)	500	0.010	50	reg:squarederror	0.200	0.100	0

Table 5: Hyperparameters choices explored for Convolutional Neural Networks (CNN) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	pool size	regularizer val	NN variant	filt size	stride	norm out
Joint angles (SE)	Adam	glorot normal	64	50	relu	32	2	rmse	rmse	0.005	2	0	CNN	3	1	0
Joint moments (SE)	Adam	glorot normal	64	100	relu	64	2	rmse	mse	0.001	2	0	CNN	3	3	1
Joint reaction forces (SE)	Adam	glorot normal	64	50	relu	32	2	rmse	rmse	0.005	2	0	CNN	3	3	1
Joint angles (SN)	Adam	glorot normal	64	50	relu	64	1	rmse	mse	0.005	2	0	CNN	3	1	0
Joint moments (SN)	Adam	glorot normal	64	100	relu	32	2	rmse	mse	0.001	2	0	CNN	3	3	1
Joint reaction forces (SN)	Adam	glorot normal	64	100	relu	32	2	rmse	rmse	0.001	2	0	CNN	3	1	1

Table 6: Hyperparameters choices explored for CNNLSTM (CNNLSTM) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	pool size	regularizer val	NN variant	filt size	stride	LSTM units	norm out
Joint angles (SE)	Adam	glorot normal	64	100	relu	32	1	rmse	mse	0.001	2	0	CNNLSTM	3	1	256	1
Joint moments (SE)	Adam	glorot normal	64	100	relu	32	1	rmse	mse	0.001	2	0	CNNLSTM	3	1	256	1
Joint reaction forces (SE)	Adam	glorot normal	64	50	relu	32	1	rmse	rmse	0.001	2	0	CNNLSTM	3	3	256	1
Joint angles (SN)	Adam	glorot normal	64	100	relu	32	1	rmse	mse	0.001	2	0	CNNLSTM	3	1	256	1
Joint moments (SN)	Adam	glorot normal	64	100	relu	32	1	rmse	mse	0.001	2	0	CNNLSTM	3	1	256	1
Joint reaction forces (SN)	Adam	glorot normal	64	100	relu	32	1	rmse	mse	0.001	2	0	CNNLSTM	3	3	256	1

Table 7: Hyperparameters choices explored for Convolutional LSTM (convLSTM) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	pool size	regularizer val	NN variant	filt size	stride	norm out
Joint angles (SE)	Adam	glorot normal	64	100	relu	32	2	rmse	rmse	0.001	2	0	CNN	3	3	0
Joint moments (SE)	Adam	glorot normal	64	100	relu	64	2	rmse	mse	0.001	2	0	CNN	3	3	1
Joint reaction forces (SE)	Adam	glorot normal	64	200	relu	32	2	rmse	mse	0.005	2	0	CNN	3	3	0
Joint angles (SN)	Adam	glorot normal	64	100	relu	64	1	rmse	mse	0.005	2	0	CNN	3	1	0
Joint moments (SN)	Adam	glorot normal	64	100	relu	64	1	rmse	mse	0.001	2	0	CNN	3	1	1
Joint reaction forces (SN)	Adam	glorot normal	64	50	relu	32	2	rmse	mse	0.001	2	0	CNN	3	3	1

Table 8: Hyperparameters choices explored for Vanilla Recurrent Neural Networks (VRNN) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	NN variant	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	p	regularizer val	norm out
Joint angles (SE)	SimpleRNN	RMSprop	glorot normal	128	100	sigmoid	256	1	rmse	rmse	0.001	0.200	0	1
Joint moments (SE)	SimpleRNN	RMSprop	glorot normal	128	200	sigmoid	256	2	rmse	mse	0.001	0.200	0	1
Joint reaction forces (SE)	SimpleRNN	RMSprop	glorot normal	128	100	sigmoid	256	2	rmse	rmse	0.001	0.100	0	1
Joint angles (SN)	SimpleRNN	RMSprop	glorot normal	128	200	relu	128	0	rmse	mse	0.001	0.100	0	1
Joint moments (SN)	SimpleRNN	RMSprop	glorot normal	64	100	relu	256	0	rmse	mse	0.001	0.200	0	1
Joint reaction forces (SN)	SimpleRNN	RMSprop	glorot normal	128	200	sigmoid	256	2	rmse	mse	0.001	0.200	0	1

Table 9: Hyperparameters choices explored for Long Short-Term Memory (LSTM) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	NN variant	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	p	regularizer val	norm out
Joint angles (SE)	LSTM	Adam	glorot normal	64	50	relu	128	0	rmse	mse	0.001	0.200	0	0
Joint moments (SE)	LSTM	Adam	glorot normal	64	50	sigmoid	128	2	rmse	rmse	0.001	0.100	0	1
Joint reaction forces (SE)	LSTM	Adam	glorot normal	128	50	relu	512	1	rmse	rmse	0.001	0.100	0	0
Joint angles (SN)	LSTM	Adam	glorot normal	64	50	relu	128	0	rmse	mse	0.001	0.200	0	0
Joint moments (SN)	LSTM	Adam	glorot normal	64	50	sigmoid	128	2	rmse	rmse	0.001	0.100	0	1
Joint reaction forces (SN)	LSTM	Adam	glorot normal	128	50	relu	512	1	rmse	rmse	0.001	0.100	0	0

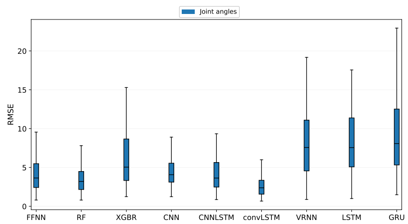
Table 10: Hyperparameters choices explored for Gated Recurrent Unit (GRU) in subject-exposed (SE) and subject-naive (SN) settings and the optimal hyperparameters found for each MSK output category.

	NN variant	optim	kinit	batch size	epoch	act	num nodes	H layer	metric	loss	lr	p	regularizer val	norm out
Joint angles (SE)	GRU	RMSprop	glorot normal	64	100	sigmoid	256	0	rmse	mse	0.001	0.100	0	1
Joint moments (SE)	GRU	RMSprop	glorot normal	128	100	sigmoid	256	1	rmse	rmse	0.001	0.100	0	1
Joint reaction forces (SE)	GRU	RMSprop	glorot normal	128	200	sigmoid	128	3	rmse	rmse	0.001	0.100	0	1
Joint angles (SN)	GRU	RMSprop	glorot normal	128	100	relu	256	2	rmse	rmse	0.001	0.100	0	1
Joint moments (SN)	GRU	Adam	glorot normal	128	100	sigmoid	256	2	rmse	rmse	0.001	0.100	0	1
Joint reaction forces (SN)	GRU	Adam	glorot normal	64	50	relu	128	0	rmse	rmse	0.001	0.100	0	0

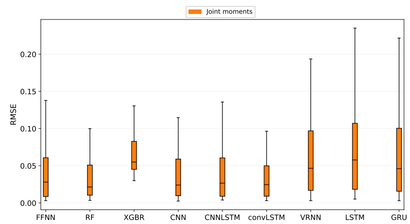
2 Statistics for various ML techniques

Here, in this section, we have reported comprehensive statistics for all model classes considered in this work, including: a) Subject-exposed model: In the subject-exposed model, at least one trial from test subjects is used in training the model, b) Subject-exposed model for unseen data: We further examined the generalisability of subject-exposed models on unseen subjects. Here, we observed that the accuracy of subject-exposed models decreases considerably for unseen subjects; however, this decrease is different for different classes of ML models. c) Subject-naive model: in which no test subject trials are part of training data, this significantly improves the accuracy for unseen subjects over subject-exposed models, and d) finally, we examined the generalisability of the subject-naive models to unseen subjects with wrist braces (mimicking transradial prosthesis usage). In general, we observed that the accuracy of subject-naive models decreases; however, this decrease is considerably different for different classes of ML models, highlighting their ability to generalize.

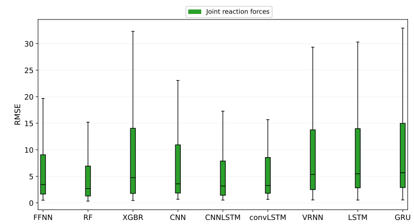
We found convLSTM to be the overall best technique, determined based on the average RMSE or NRMSE. convLSTM is the best model for three out of six different techniques (subject-exposed and subject-naive models for joint angles, joint moments, and joint reaction forces) and is very close to the best-performing techniques for the remaining categories.



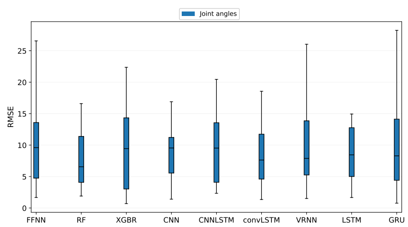
(a) Subject-exposed



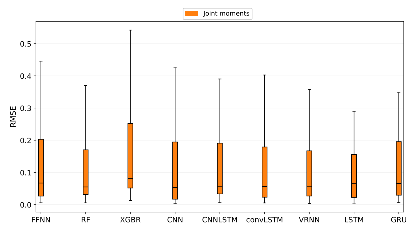
(b) Subject-exposed



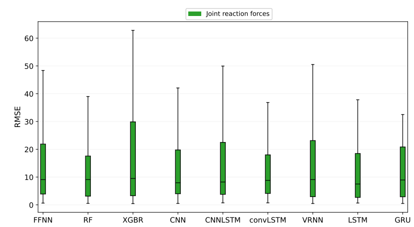
(c) Subject-exposed



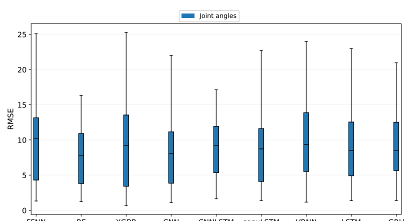
(d) Subject-exposed for unseen subjects



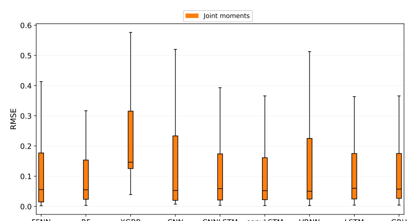
(e) Subject-exposed for unseen subjects



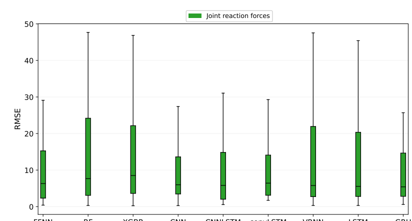
(f) Subject-exposed for unseen subjects



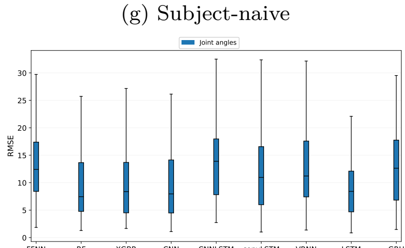
(g) Subject-naive



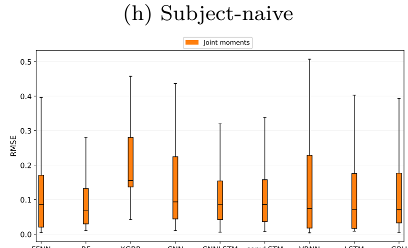
(h) Subject-naive



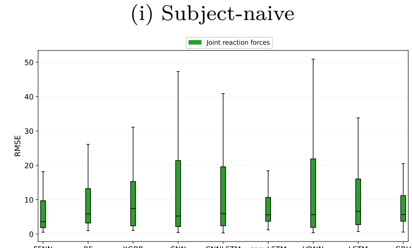
(i) Subject-naive



(j) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)

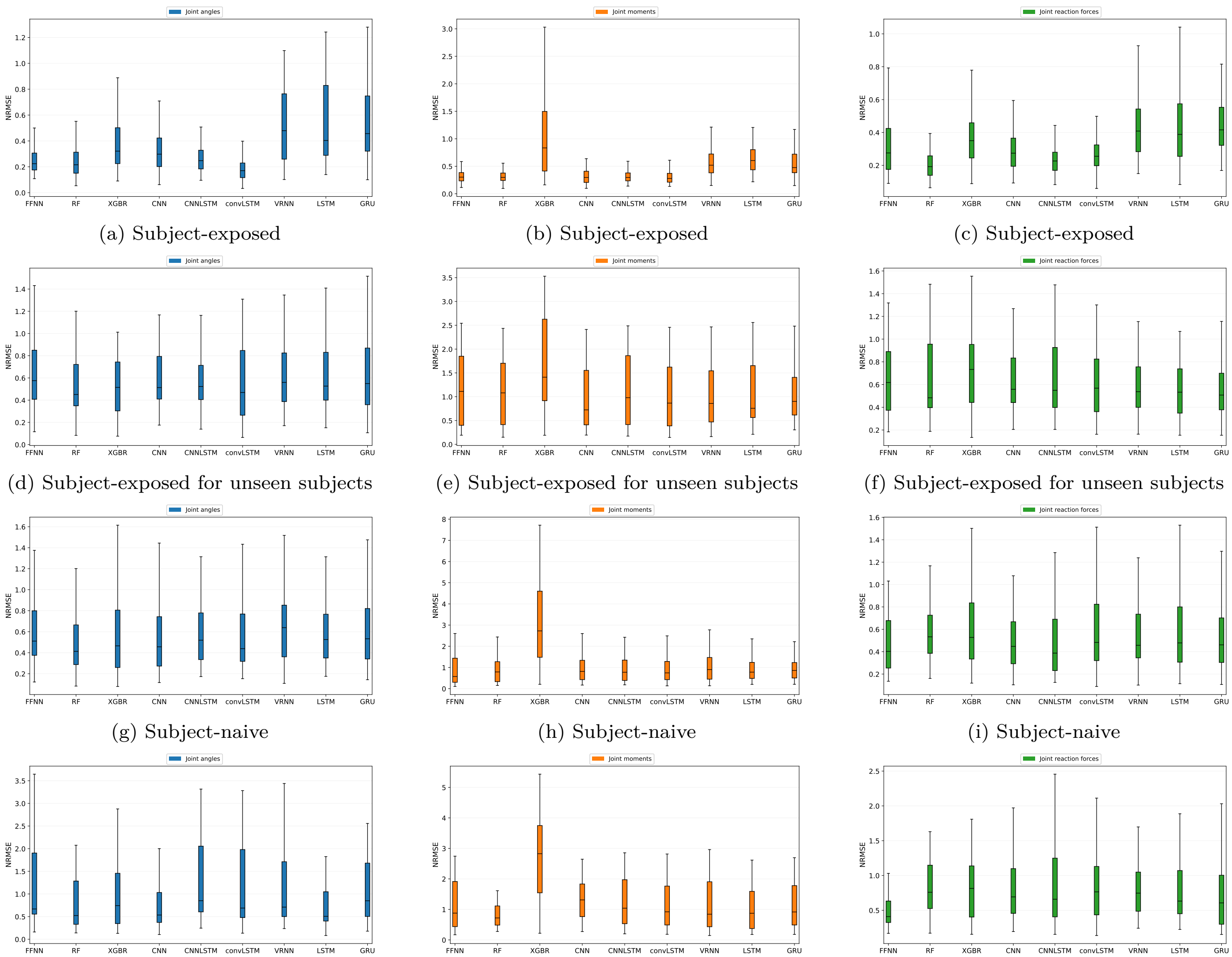


(k) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)



(l) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)

Figure 1: This figure reports the performance of widely used supervised ML techniques, including Random Forests (RF), Extreme Gradient Boosting (XGBR), Feedforward Neural Networks (FFNN), Vanilla Recurrent Neural Networks (VRNN) and their advanced variants (Long Short-Term Memory [LSTM] and Gated Recurrent Unit [GRU]), Convolutional Neural Networks (CNN), and hybrid architectures (viz. CNNLSTM and Convolutional LSTM [convLSTM]) in mapping OMC inputs to MSK-based biomechanical outputs. Accuracy is reported in terms of average RMSE (root-mean-square error) for held-out test data (consisting of multiple trials and/or subjects), with the error bar showing the standard deviation across trials and/or subjects. Note that the subject-naive models were not retrained for wrist-braced subjects and, therefore, are a measure of the transferability of the model.



(a) Subject-exposed

(b) Subject-exposed

(c) Subject-exposed

(d) Subject-exposed for unseen subjects

(e) Subject-exposed for unseen subjects

(f) Subject-exposed for unseen subjects

(g) Subject-naive

(h) Subject-naive

(i) Subject-naive

(j) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)

(k) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)

(l) Subject-naive models transferability to braced subjects (i.e., mimicked transradial prosthesis usage)

Figure 2: This figure reports the performance of widely used supervised ML techniques, including Random Forests (RF), Extreme Gradient Boosting (XGBR), Feedforward Neural Networks (FFNN), Vanilla Recurrent Neural Networks (VRNN) and their advanced variants (Long Short-Term Memory [LSTM] and Gated Recurrent Unit [GRU]), Convolutional Neural Networks (CNN), and hybrid architectures (viz. CNNLSTM and Convolutional LSTM [convLSTM]) in mapping OMC inputs to MSK-based biomechanical outputs. Accuracy is reported in terms of average NRMSE (normalised root-mean-square error) for held-out test data (consisting of multiple trials and/or subjects), with the error bar showing the standard deviation across trials and/or subjects. Note that the subject-naive models were not retrained for wrist-braced subjects and, therefore, are a measure of the transferability of the model.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.15	0.58	0.95	-0.88	0.98	4.44	2.19	12.38	1.03	2.52	73.73	51.15	279.93	5.94	49.64
Joint moments	0.11	0.51	0.92	-0.88	0.91	4.08	1.94	10.10	1.25	2.95	0.45	0.48	2.21	0.04	0.44
Joint reaction forces	0.15	0.50	0.94	-0.88	0.81	4.88	2.58	12.90	1.21	3.28	138.20	194.70	1030.79	4.43	114.61
	Subject-naive model														
Joint angles	-0.07	0.62	0.98	-0.94	1.11	8.38	5.97	22.67	1.52	9.35	140.95	134.28	561.37	12.01	122.99
Joint moments	0.26	0.50	0.95	-0.93	0.80	7.49	5.27	22.10	1.10	8.25	0.86	0.95	3.68	0.03	1.16
Joint reaction forces	0.13	0.53	0.95	-0.89	0.93	7.90	6.48	27.28	1.40	7.78	182.19	260.13	1351.60	3.12	162.37
	Subject-exposed model for unseen data														
Joint angles	0.26	0.61	0.96	-0.95	1.01	6.91	5.27	26.16	1.70	4.92	117.38	103.06	412.80	12.58	122.40
Joint moments	0.20	0.44	0.90	-0.62	0.79	5.69	3.46	19.15	1.52	3.90	0.57	0.50	1.74	0.07	0.54
Joint reaction forces	0.18	0.52	0.93	-0.94	0.72	5.27	3.17	17.34	1.48	3.54	114.22	122.97	572.44	4.76	119.62
	Subject-naive model for braced data														
Joint angles	0.09	0.64	0.98	-0.94	1.23	10.16	6.60	28.17	1.49	10.12	139.98	122.62	456.52	13.04	143.98
Joint moments	0.37	0.50	0.92	-0.74	0.79	12.07	7.66	27.64	2.18	13.50	1.20	1.14	4.52	0.07	0.84
Joint reaction forces	-0.01	0.63	0.92	-0.97	1.29	10.51	6.46	26.89	1.49	8.76	150.03	181.60	730.97	3.60	160.70

Table 11: Average Pearson's correlation coefficient and average NRMSE/RMSE values for LM predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.87	0.24	1.00	-0.41	0.10	0.25	0.11	0.61	0.11	0.13	4.00	2.15	11.02	0.81	3.07
Joint moments	0.80	0.22	0.99	-0.05	0.21	0.32	0.12	0.76	0.12	0.15	0.04	0.04	0.17	0.00	0.05
Joint reaction forces	0.80	0.31	0.99	-0.85	0.16	0.32	0.18	0.89	0.09	0.25	7.92	10.42	50.56	0.50	7.40
	Subject-naive model														
Joint angles	0.74	0.39	1.00	-0.67	0.28	0.62	0.37	1.83	0.12	0.42	9.64	5.80	25.08	1.32	8.85
Joint moments	0.73	0.29	0.99	-0.03	0.32	0.90	0.70	2.60	0.10	1.14	0.13	0.17	0.58	0.00	0.16
Joint reaction forces	0.60	0.54	0.99	-0.93	0.48	0.54	0.46	2.85	0.14	0.42	13.58	20.13	123.26	0.40	12.96
	Subject-exposed model for unseen data														
Joint angles	0.62	0.50	0.99	-0.81	0.44	0.68	0.35	1.43	0.12	0.44	11.59	9.41	44.03	1.68	8.84
Joint moments	0.55	0.48	0.98	-0.81	0.56	1.20	0.82	2.54	0.19	1.45	0.15	0.18	0.62	0.01	0.18
Joint reaction forces	0.49	0.62	0.98	-0.90	0.76	0.75	0.56	2.67	0.18	0.52	18.73	24.30	115.47	0.63	17.98
	Subject-naive model for braced data														
Joint angles	0.52	0.56	0.99	-0.91	0.86	1.36	1.32	5.55	0.16	1.35	14.20	9.09	36.98	1.86	8.98
Joint moments	0.41	0.59	0.99	-0.94	0.85	1.14	0.84	2.74	0.17	1.48	0.14	0.15	0.53	0.00	0.15
Joint reaction forces	0.74	0.34	0.97	-0.87	0.17	0.62	0.54	2.44	0.17	0.31	7.72	8.51	37.05	0.53	7.85

Table 12: Average Pearson's correlation coefficient and average NRMSE/RMSE values for NN predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range. The statistics highlighted in blue represent the best performance (in terms of RMSE/NRMSE) among all the models tested.

Output	r					NRMSE					RMSE				
	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.88	0.19	1.00	0.11	0.16	0.24	0.12	0.67	0.05	0.16	3.65	2.06	9.85	0.80	2.32
Joint moments	0.80	0.21	0.99	-0.06	0.19	0.32	0.12	0.62	0.10	0.14	0.03	0.03	0.14	0.00	0.04
Joint reaction forces	0.85	0.26	0.99	-0.83	0.10	0.23	0.15	0.84	0.06	0.12	4.86	5.08	21.66	0.34	5.63
	Subject-naive model														
Joint angles	0.75	0.47	1.00	-0.77	0.16	0.55	0.42	2.18	0.08	0.38	8.08	5.28	29.08	1.25	7.11
Joint moments	0.66	0.43	1.00	-0.61	0.37	0.91	0.64	2.44	0.15	0.94	0.12	0.14	0.53	0.00	0.13
Joint reaction forces	0.52	0.61	0.99	-0.88	0.75	0.64	0.40	2.28	0.16	0.34	16.62	20.13	98.73	0.33	21.12
	Subject-exposed model for unseen data														
Joint angles	0.82	0.39	1.00	-0.83	0.13	0.56	0.34	1.54	0.08	0.37	8.54	5.89	29.80	1.90	7.31
Joint moments	0.61	0.52	1.00	-0.91	0.39	1.12	0.73	2.44	0.15	1.29	0.14	0.17	0.60	0.01	0.14
Joint reaction forces	0.55	0.61	0.99	-0.89	0.65	0.70	0.51	2.63	0.19	0.56	16.56	22.16	114.07	0.55	14.45
	Subject-naive model for braced data														
Joint angles	0.57	0.57	0.99	-0.80	0.72	0.84	0.70	3.15	0.14	0.96	9.41	6.61	25.72	1.29	8.89
Joint moments	0.23	0.71	0.99	-0.97	1.35	0.81	0.39	1.61	0.28	0.63	0.09	0.07	0.28	0.01	0.10
Joint reaction forces	0.51	0.59	0.98	-0.75	0.83	0.99	0.82	3.77	0.17	0.62	11.50	12.15	54.50	0.97	10.00

Table 13: Average Pearson's correlation coefficient and average NRMSE/RMSE values for RF predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range. The statistics highlighted in blue represent the best performance (in terms of RMSE/NRMSE) among all the models tested.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.83	0.27	1.00	-0.86	0.22	0.47	0.47	3.34	0.09	0.28	6.17	3.93	24.01	1.24	5.35
Joint moments	0.76	0.27	0.99	-0.36	0.20	1.07	0.78	3.21	0.16	1.08	0.07	0.04	0.19	0.03	0.04
Joint reaction forces	0.82	0.27	0.99	-0.76	0.18	0.38	0.20	1.10	0.09	0.21	10.05	13.12	71.50	0.45	12.27
	Subject-naive model														
Joint angles	0.69	0.47	1.00	-0.87	0.52	0.65	0.54	2.14	0.08	0.55	9.61	7.39	41.34	0.65	10.13
Joint moments	0.62	0.34	1.00	-0.50	0.41	3.06	2.02	7.72	0.21	3.12	0.23	0.18	0.82	0.04	0.19
Joint reaction forces	0.49	0.59	0.99	-0.88	0.86	0.66	0.43	2.43	0.12	0.50	16.66	20.84	105.23	0.24	18.54
	Subject-exposed model for unseen data														
Joint angles	0.73	0.47	1.00	-0.89	0.27	0.62	0.50	2.04	0.08	0.44	9.83	7.67	36.23	0.69	11.31
Joint moments	0.56	0.49	1.00	-0.74	0.56	1.67	0.99	3.53	0.19	1.71	0.17	0.19	0.64	0.01	0.20
Joint reaction forces	0.44	0.63	0.99	-0.88	0.93	0.79	0.55	3.01	0.13	0.51	22.22	29.27	130.39	0.43	26.60
	Subject-naive model for braced data														
Joint angles	0.56	0.51	0.98	-0.85	0.79	1.02	0.86	3.33	0.13	1.11	11.12	8.74	38.13	1.66	9.19
Joint moments	0.27	0.59	0.97	-0.90	0.82	2.69	1.50	5.43	0.22	2.21	0.20	0.10	0.46	0.04	0.14
Joint reaction forces	0.54	0.49	0.98	-0.71	0.79	1.04	1.00	4.79	0.16	0.73	11.64	11.69	55.42	1.04	12.90

Table 14: Average Pearson's correlation coefficient and average NRMSE/RMSE values for xgbr predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

Output	r					NRMSE					RMSE				
	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.78	0.38	1.00	-0.80	0.21	0.33	0.18	1.13	0.06	0.22	4.66	2.34	12.55	1.23	2.44
Joint moments	0.81	0.19	0.99	0.18	0.20	0.32	0.14	0.64	0.10	0.20	0.04	0.03	0.14	0.00	0.05
Joint reaction forces	0.77	0.37	0.99	-0.66	0.23	0.32	0.19	1.14	0.09	0.17	6.87	7.04	33.83	0.68	9.08
	Subject-naive model														
Joint angles	0.77	0.42	1.00	-0.91	0.23	0.59	0.45	2.26	0.12	0.47	8.89	6.64	32.43	1.08	7.30
Joint moments	0.74	0.27	0.99	-0.08	0.35	0.99	0.68	2.60	0.17	0.91	0.14	0.16	0.62	0.01	0.21
Joint reaction forces	0.72	0.47	0.99	-0.90	0.24	0.55	0.41	2.26	0.10	0.38	13.16	18.14	97.93	0.28	10.17
	Subject-exposed model for unseen data														
Joint angles	0.72	0.46	1.00	-0.56	0.25	0.66	0.39	1.78	0.18	0.38	9.73	6.28	34.45	1.40	5.67
Joint moments	0.58	0.44	0.99	-0.63	0.44	1.00	0.68	2.41	0.19	1.14	0.14	0.17	0.56	0.00	0.18
Joint reaction forces	0.48	0.60	0.99	-0.86	0.95	0.69	0.42	2.43	0.21	0.39	17.35	21.62	105.16	0.53	15.78
	Subject-naive model for braced data														
Joint angles	0.69	0.48	1.00	-0.55	0.23	0.82	0.64	2.83	0.11	0.66	9.75	6.60	26.13	1.09	9.66
Joint moments	0.35	0.61	0.99	-0.90	0.81	1.31	0.69	2.65	0.27	1.07	0.14	0.14	0.51	0.01	0.18
Joint reaction forces	0.75	0.40	0.98	-0.63	0.14	0.84	0.51	2.32	0.19	0.64	14.13	17.30	67.10	0.47	19.22

Table 15: Average Pearson's correlation coefficient and average NRMSE/RMSE values for CNN predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.82	0.30	1.00	-0.57	0.18	0.26	0.10	0.51	0.10	0.14	4.27	2.55	13.07	0.87	3.17
Joint moments	0.79	0.21	0.99	-0.22	0.22	0.32	0.13	0.70	0.14	0.14	0.04	0.04	0.15	0.00	0.05
Joint reaction forces	0.81	0.29	0.99	-0.83	0.20	0.26	0.14	0.85	0.08	0.11	5.76	6.25	26.51	0.52	6.46
	Subject-naive model														
Joint angles	0.73	0.43	0.99	-0.72	0.23	0.63	0.40	1.99	0.17	0.44	9.23	6.03	38.51	1.64	6.59
Joint moments	0.71	0.25	0.99	0.03	0.36	0.96	0.64	2.43	0.18	0.96	0.13	0.16	0.53	0.00	0.15
Joint reaction forces	0.67	0.49	0.98	-0.88	0.33	0.55	0.47	2.39	0.13	0.46	13.52	18.68	103.69	0.56	12.82
	Subject-exposed model for unseen data														
Joint angles	0.66	0.52	1.00	-0.83	0.28	0.61	0.35	1.86	0.14	0.31	9.90	6.91	35.92	2.34	9.47
Joint moments	0.58	0.38	0.98	-0.61	0.64	1.16	0.76	2.49	0.17	1.45	0.15	0.18	0.61	0.01	0.16
Joint reaction forces	0.52	0.61	0.97	-0.91	0.86	0.74	0.53	2.64	0.21	0.53	17.53	23.01	114.42	0.73	18.71
	Subject-naive model for braced data														
Joint angles	0.61	0.41	0.98	-0.51	0.70	1.36	1.19	5.06	0.25	1.45	14.32	7.94	33.68	2.73	10.16
Joint moments	0.38	0.54	0.98	-0.95	0.41	1.26	0.85	2.85	0.20	1.44	0.14	0.14	0.51	0.01	0.11
Joint reaction forces	0.71	0.32	0.98	-0.29	0.22	0.84	0.54	2.46	0.15	0.84	14.57	18.01	71.12	0.35	17.16

Table 16: Average Pearson's correlation coefficient and average NRMSE/RMSE values for CNNLSTM predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.87	0.23	1.00	-0.07	0.17	0.19	0.12	0.63	0.03	0.11	2.69	1.51	8.13	0.67	1.79
Joint moments	0.80	0.22	0.99	-0.18	0.24	0.31	0.13	0.62	0.13	0.16	0.03	0.03	0.14	0.00	0.04
Joint reaction forces	0.81	0.29	0.99	-0.82	0.17	0.28	0.15	0.98	0.06	0.13	5.79	5.64	26.25	0.66	6.75
	Subject-naive model														
Joint angles	0.76	0.38	1.00	-0.78	0.26	0.59	0.39	1.85	0.15	0.45	8.93	5.92	30.62	1.41	7.52
Joint moments	0.71	0.30	0.99	-0.22	0.27	0.95	0.71	2.86	0.13	0.86	0.13	0.16	0.63	0.00	0.14
Joint reaction forces	0.66	0.48	0.99	-0.91	0.48	0.62	0.46	2.29	0.09	0.50	12.70	17.40	99.11	1.73	11.00
	Subject-exposed model for unseen data														
Joint angles	0.82	0.39	1.00	-0.67	0.14	0.60	0.45	2.10	0.06	0.58	8.48	5.43	25.10	1.35	7.14
Joint moments	0.52	0.51	0.99	-0.93	0.54	1.05	0.74	2.46	0.14	1.23	0.14	0.18	0.60	0.01	0.16
Joint reaction forces	0.53	0.59	0.98	-0.86	0.65	0.70	0.48	2.53	0.16	0.46	16.36	21.57	109.58	0.73	13.90
	Subject-naive model for braced data														
Joint angles	0.67	0.48	0.99	-0.54	0.58	1.27	1.25	4.90	0.13	1.50	12.84	8.34	32.67	1.02	10.58
Joint moments	0.60	0.44	0.98	-0.81	0.40	1.18	0.78	2.82	0.18	1.28	0.13	0.14	0.46	0.01	0.12
Joint reaction forces	0.70	0.39	0.99	-0.70	0.33	0.94	0.75	3.63	0.14	0.70	9.10	8.81	38.71	1.23	6.95

Table 17: Average Pearson's correlation coefficient and average NRMSE/RMSE values for convLSTM predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range. The statistics highlighted in blue represent the best performance (in terms of RMSE/NRMSE) among all the models tested.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.77	0.37	1.00	-0.66	0.20	0.57	0.40	2.26	0.10	0.51	8.26	4.76	24.07	0.88	6.54
Joint moments	0.65	0.35	0.98	-0.56	0.39	0.62	0.36	1.95	0.15	0.34	0.06	0.06	0.26	0.00	0.08
Joint reaction forces	0.70	0.41	0.99	-0.61	0.34	0.48	0.33	2.36	0.15	0.26	10.41	11.22	48.21	0.56	11.26
	Subject-naive model														
Joint angles	0.73	0.46	1.00	-0.93	0.24	0.67	0.39	1.96	0.11	0.49	10.32	6.38	29.35	1.16	8.36
Joint moments	0.63	0.33	0.98	-0.29	0.50	1.05	0.73	3.08	0.14	1.02	0.15	0.18	0.69	0.00	0.20
Joint reaction forces	0.63	0.49	0.97	-0.77	0.47	0.59	0.39	2.42	0.10	0.39	17.12	23.33	104.68	0.33	19.21
	Subject-exposed model for unseen data														
Joint angles	0.72	0.47	0.99	-0.95	0.31	0.65	0.34	1.35	0.17	0.44	10.75	7.73	31.18	1.52	8.61
Joint moments	0.59	0.38	0.98	-0.57	0.53	1.01	0.65	2.47	0.16	1.07	0.14	0.17	0.57	0.00	0.14
Joint reaction forces	0.56	0.57	0.98	-0.79	0.60	0.63	0.38	2.23	0.16	0.36	18.78	24.59	96.70	0.50	20.21
	Subject-naive model for braced data														
Joint angles	0.56	0.53	0.99	-0.81	0.69	1.30	1.29	5.24	0.24	1.21	13.54	8.72	34.91	1.38	10.18
Joint moments	0.40	0.62	0.99	-0.95	0.80	1.17	0.89	2.96	0.14	1.47	0.15	0.16	0.51	0.00	0.21
Joint reaction forces	0.71	0.33	0.96	-0.68	0.19	0.87	0.55	2.62	0.24	0.56	15.93	20.15	75.99	0.37	19.93

Table 18: Average Pearson's correlation coefficient and average NRMSE/RMSE values for VRNN predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.79	0.37	1.00	-0.83	0.11	0.59	0.42	2.12	0.14	0.54	8.40	4.86	25.27	1.00	6.30
Joint moments	0.60	0.37	0.98	-0.52	0.48	0.69	0.44	2.39	0.22	0.37	0.07	0.07	0.31	0.01	0.09
Joint reaction forces	0.74	0.39	0.99	-0.82	0.30	0.46	0.29	2.11	0.08	0.32	10.10	10.74	44.57	0.53	11.15
	Subject-naive model														
Joint angles	0.74	0.45	0.99	-0.93	0.19	0.63	0.41	2.04	0.18	0.42	9.75	6.80	32.53	1.38	7.64
Joint moments	0.59	0.39	0.97	-0.72	0.58	0.94	0.61	2.59	0.20	0.76	0.13	0.15	0.60	0.00	0.15
Joint reaction forces	0.65	0.47	0.99	-0.85	0.47	0.64	0.49	2.97	0.11	0.49	18.65	26.30	128.63	0.33	17.57
	Subject-exposed model for unseen data														
Joint angles	0.75	0.49	0.99	-0.92	0.15	0.64	0.35	1.41	0.15	0.43	10.68	8.08	33.78	1.67	7.76
Joint moments	0.46	0.50	0.97	-0.77	0.85	1.07	0.69	2.56	0.21	1.09	0.14	0.17	0.62	0.00	0.13
Joint reaction forces	0.56	0.58	0.99	-0.70	0.74	0.70	0.59	3.03	0.15	0.39	18.84	27.40	131.23	0.66	15.82
	Subject-naive model for braced data														
Joint angles	0.69	0.44	0.98	-0.49	0.18	0.78	0.58	2.28	0.08	0.65	9.19	5.90	26.31	0.86	7.42
Joint moments	0.35	0.63	0.97	-0.97	0.77	1.09	0.81	2.62	0.18	1.22	0.13	0.14	0.43	0.01	0.16
Joint reaction forces	0.71	0.37	0.97	-0.72	0.28	0.90	0.69	3.07	0.23	0.62	12.64	14.55	58.74	0.75	13.30

Table 19: Average Pearson's correlation coefficient and average NRMSE/RMSE values for LSTM predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

	r					NRMSE					RMSE				
Output	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR	Mean	SD	Max	Min	IQR
	Subject-exposed model														
Joint angles	0.78	0.30	0.98	-0.32	0.19	0.60	0.41	2.31	0.10	0.43	8.96	5.31	26.16	1.48	7.20
Joint moments	0.63	0.35	0.99	-0.52	0.37	0.59	0.30	1.65	0.15	0.34	0.06	0.06	0.23	0.00	0.08
Joint reaction forces	0.71	0.39	0.98	-0.72	0.33	0.48	0.30	2.39	0.17	0.23	10.85	11.64	49.36	0.56	12.09
	Subject-naive model														
Joint angles	0.68	0.45	1.00	-0.88	0.40	0.63	0.35	1.57	0.14	0.48	9.71	6.16	30.42	1.40	6.87
Joint moments	0.56	0.39	0.97	-0.43	0.59	0.96	0.59	2.79	0.20	0.73	0.13	0.15	0.61	0.00	0.15
Joint reaction forces	0.67	0.49	0.99	-0.88	0.31	0.58	0.44	2.73	0.11	0.40	15.03	22.40	118.03	0.61	11.84
	Subject-exposed model for unseen data														
Joint angles	0.73	0.45	0.97	-0.90	0.22	0.62	0.36	1.52	0.11	0.51	10.78	8.36	32.53	0.77	9.72
Joint moments	0.48	0.49	0.97	-0.79	0.69	1.04	0.56	2.48	0.30	0.79	0.13	0.15	0.54	0.01	0.17
Joint reaction forces	0.58	0.57	0.98	-0.88	0.58	0.61	0.38	2.26	0.15	0.32	17.74	23.38	98.00	0.51	17.94
	Subject-naive model for braced data														
Joint angles	0.54	0.56	0.99	-0.87	0.63	1.31	1.30	5.39	0.18	1.18	13.71	8.60	35.92	1.48	10.94
Joint moments	0.30	0.59	0.96	-0.93	0.81	1.15	0.78	2.69	0.18	1.29	0.13	0.14	0.44	0.01	0.14
Joint reaction forces	0.71	0.41	0.98	-0.74	0.25	0.84	0.73	2.84	0.15	0.70	9.84	12.14	58.82	0.56	7.48

Table 20: Average Pearson's correlation coefficient and average NRMSE/RMSE values for GRU predictions compared with Musculoskeletal (MSK) model outputs. The subject-exposed model is tested on held-out (unseen) test data, highlighting the need for subject-naive models. For braced data sets, the accuracy of subject-naive models is reported, highlighting the transferability of the models. Note: The average is taken over all output features and over all test trials (for a given output category). IQR refers to inter-quartile range.

3 Robustness of models against the inherent noise in Optical Motion Capture systems

Here we have shown the comparison of musculoskeletal (MSK) model outputs with the model predictions for a trial in the held-out test data in Subject-exposed and Subject-naive settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the models is illustrated by the variability in the model predictions and also quantified in terms of signal-to-noise ratio (SNR).

In general, all ML models are highly robust to the added noise, except for a few output variables where the range of motion is negligible. Ideally, such variables can be removed from the analysis and model training, but are often kept for completion and also to measure the noise impact in this case. Notably, the SNR ratio for the linear model is poor.

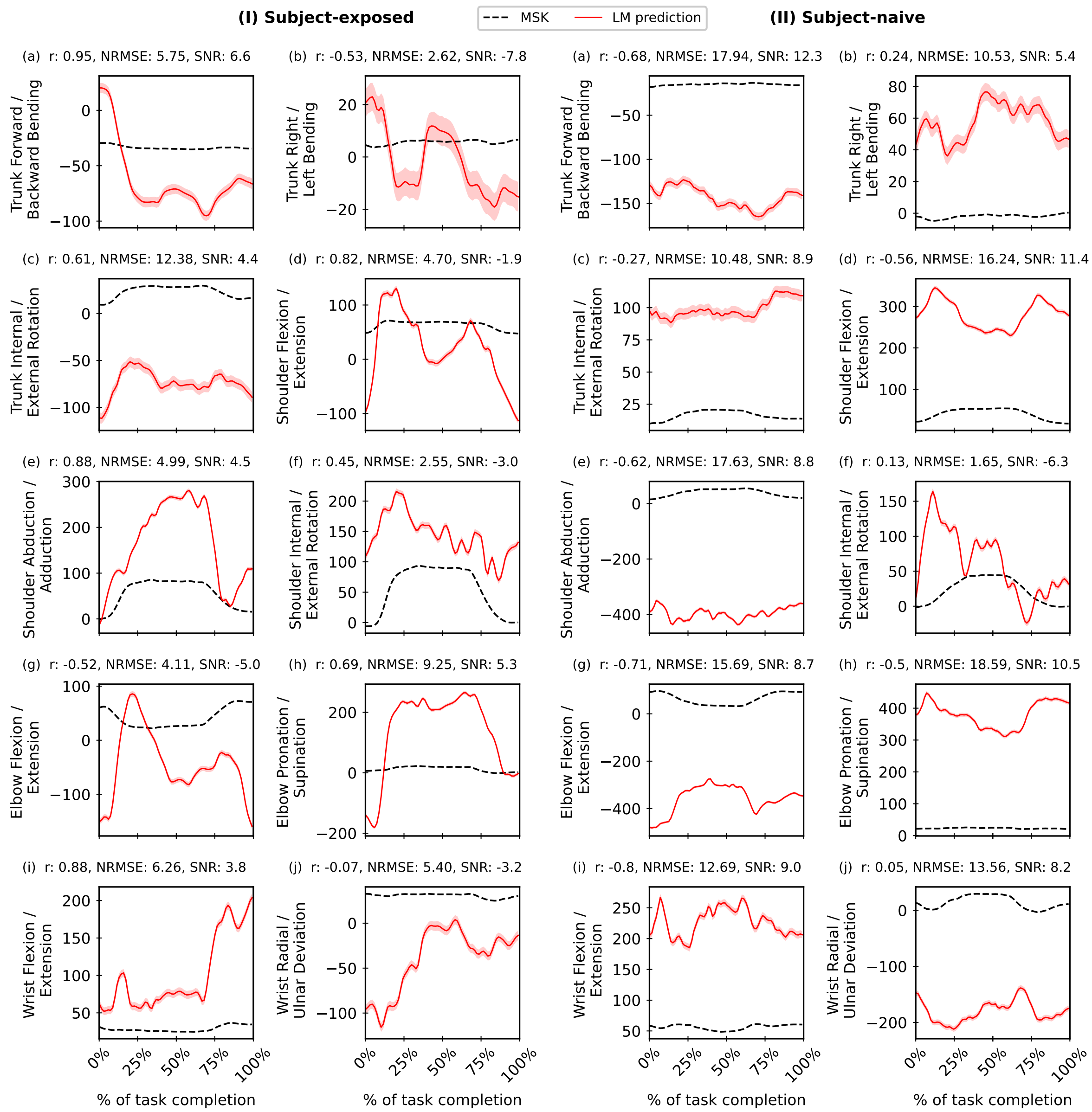


Figure 3: Comparing musculoskeletal (MSK) model outputs with the Linear model (LM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 0.4±4.75 and 7.55±5.35 decibels.

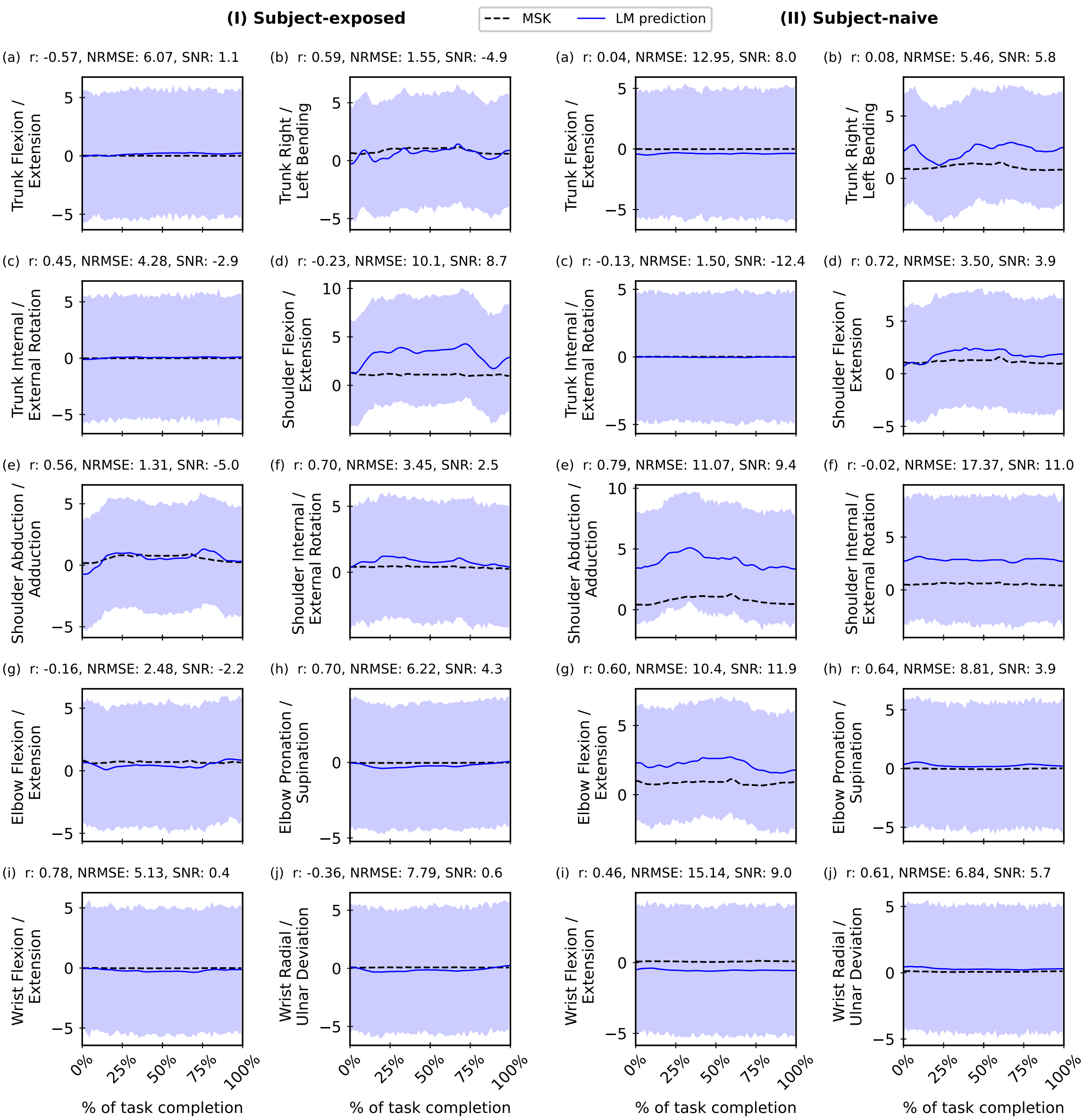


Figure 4: Comparing musculoskeletal (MSK) model outputs with the Linear model (LM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 0.4±4.1 and 5.65±6.56 decibels.

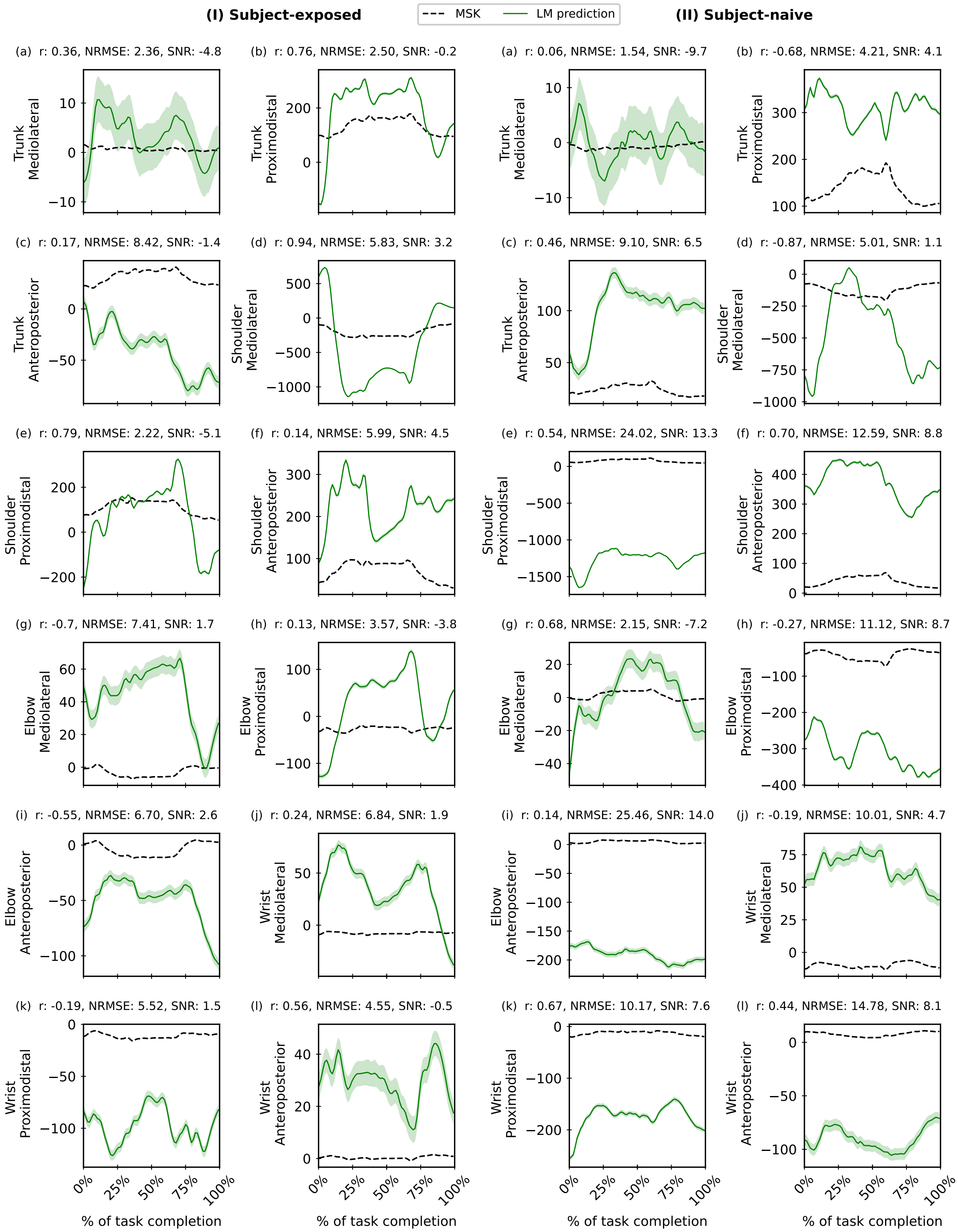


Figure 5: Comparing musculoskeletal (MSK) model outputs with the Linear model (LM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 0.0±3.12 and 5.11±6.91 decibels.

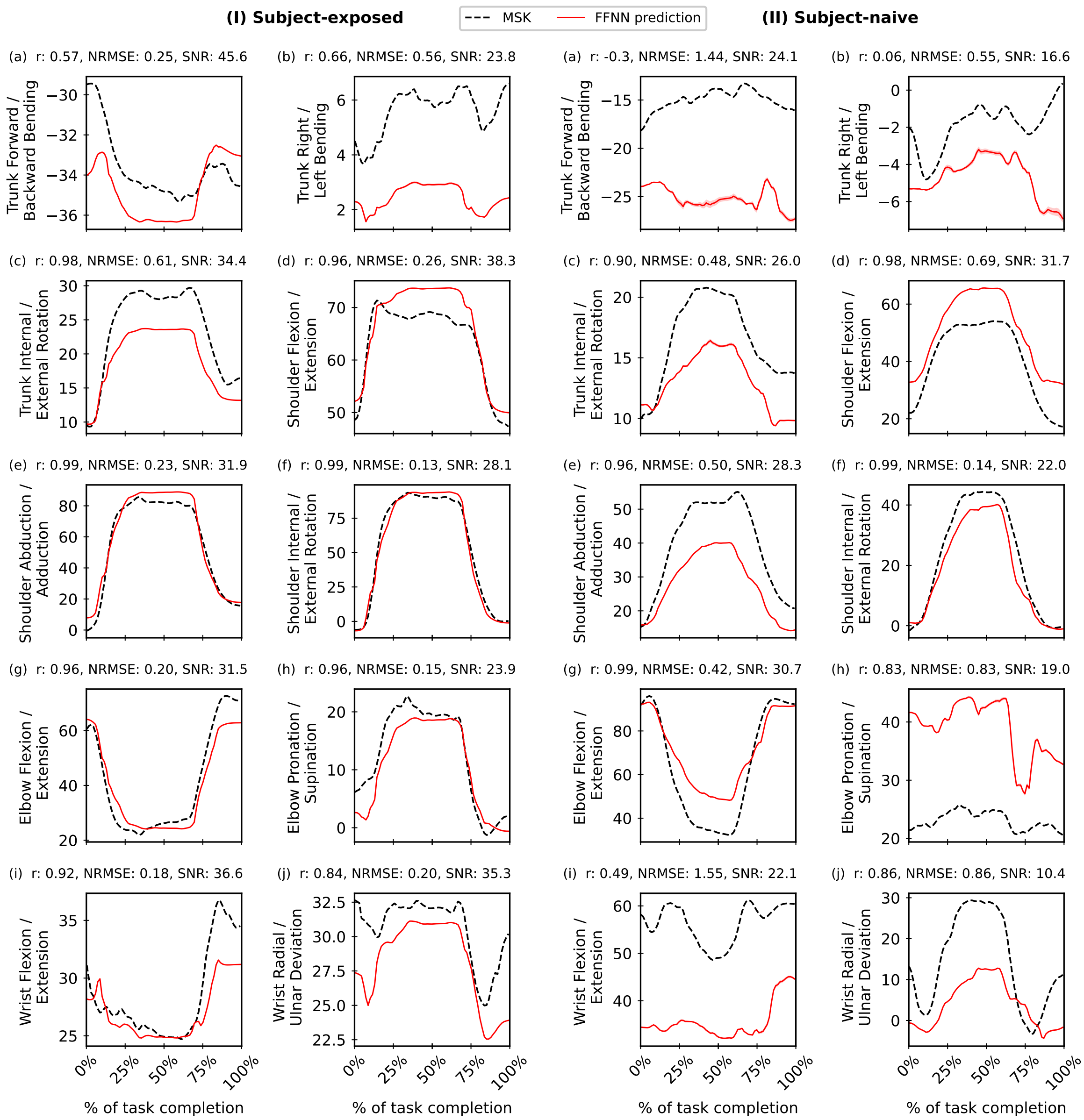


Figure 6: Comparing musculoskeletal (MSK) model outputs with the Feedforward Neural Networks (FFNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the FFNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 33.71±6.62 and 23.7±6.19 decibels.

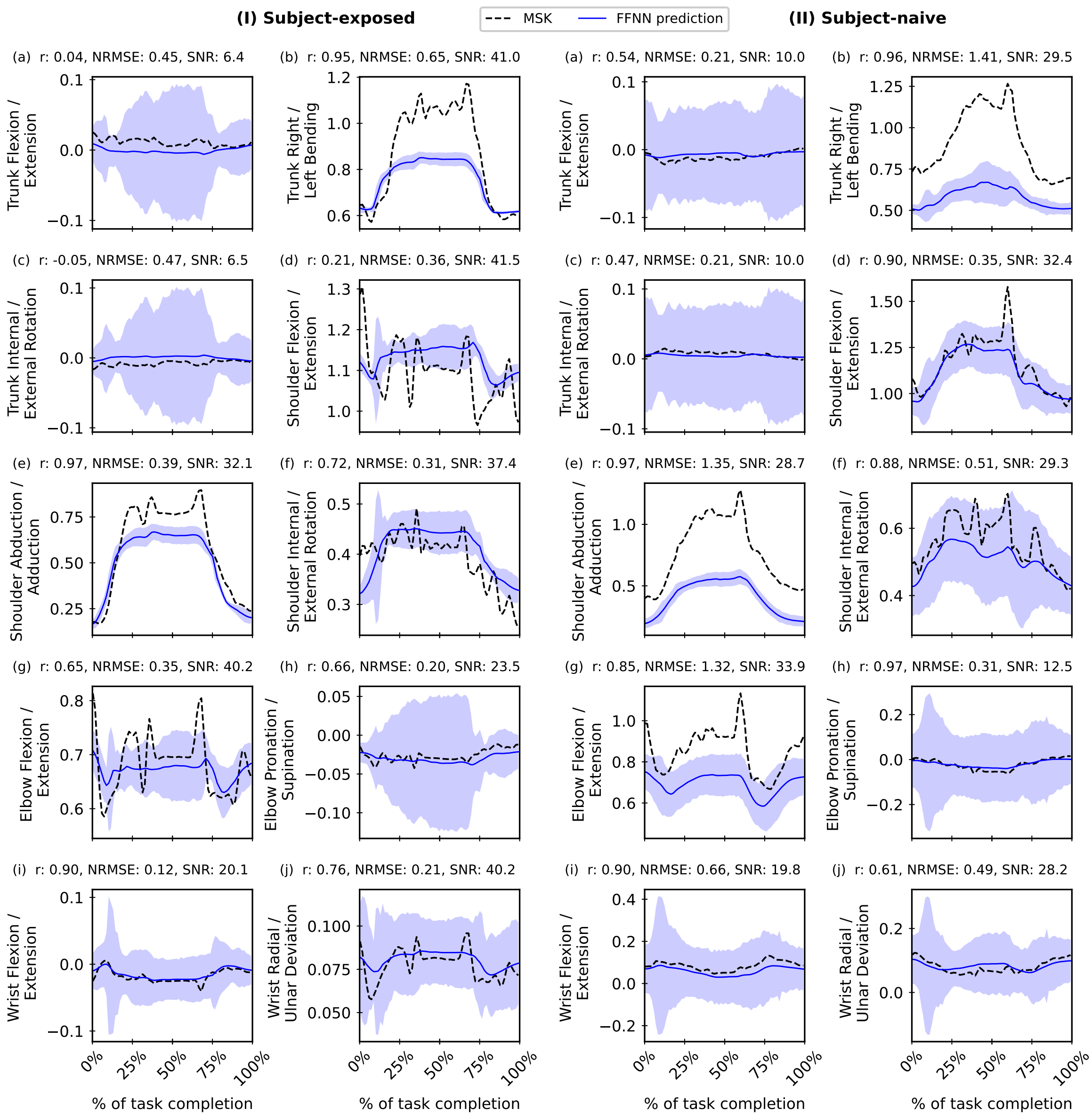


Figure 7: Comparing musculoskeletal (MSK) model outputs with the Feedforward Neural Networks (FFNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the FFNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 30.8±13.25 and 23.93±9.18 decibels.

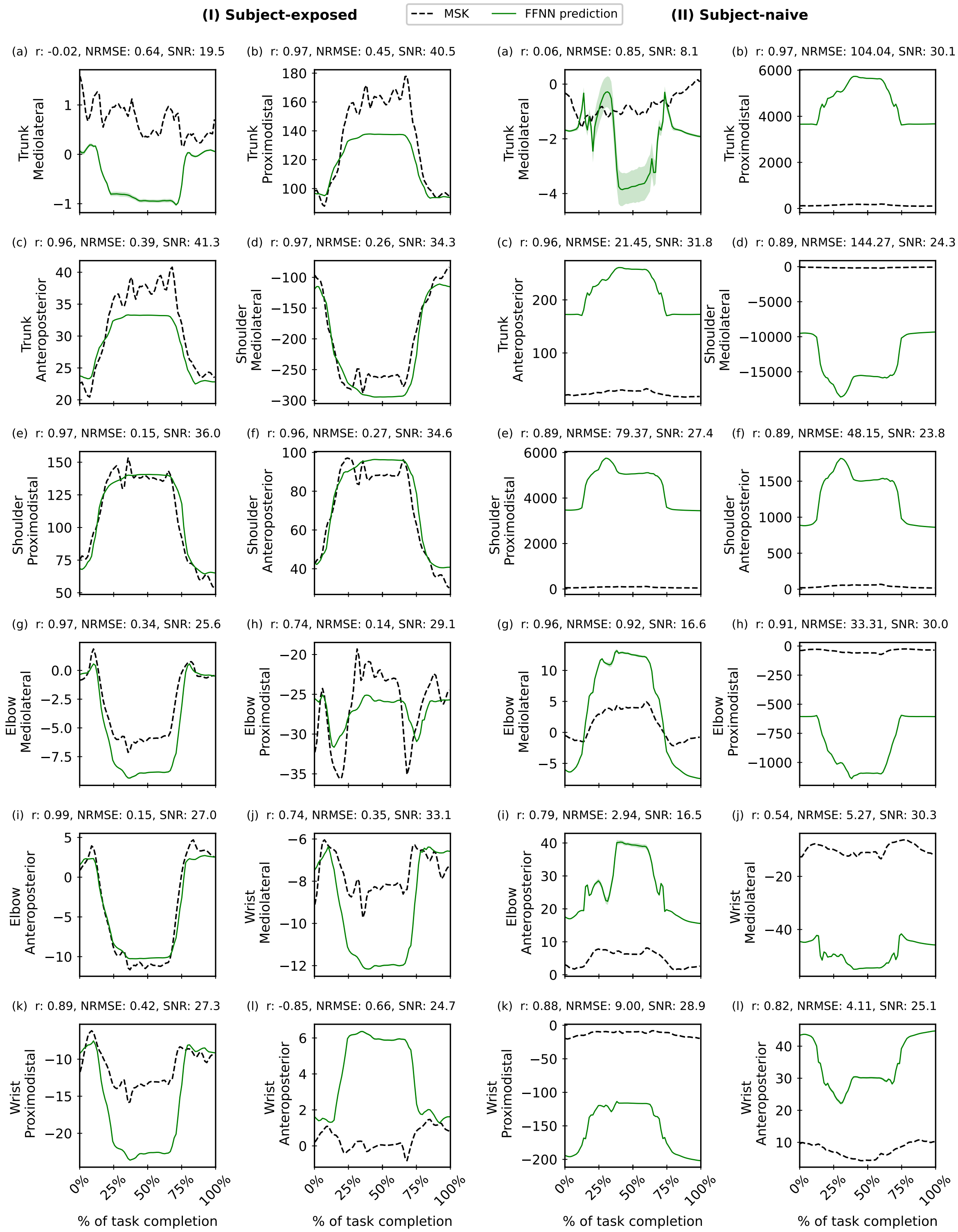


Figure 8: Comparing musculoskeletal (MSK) model outputs with the Feedforward Neural Networks (FFNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the FFNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 32.27±6.45 and 24.79±6.84 decibels.

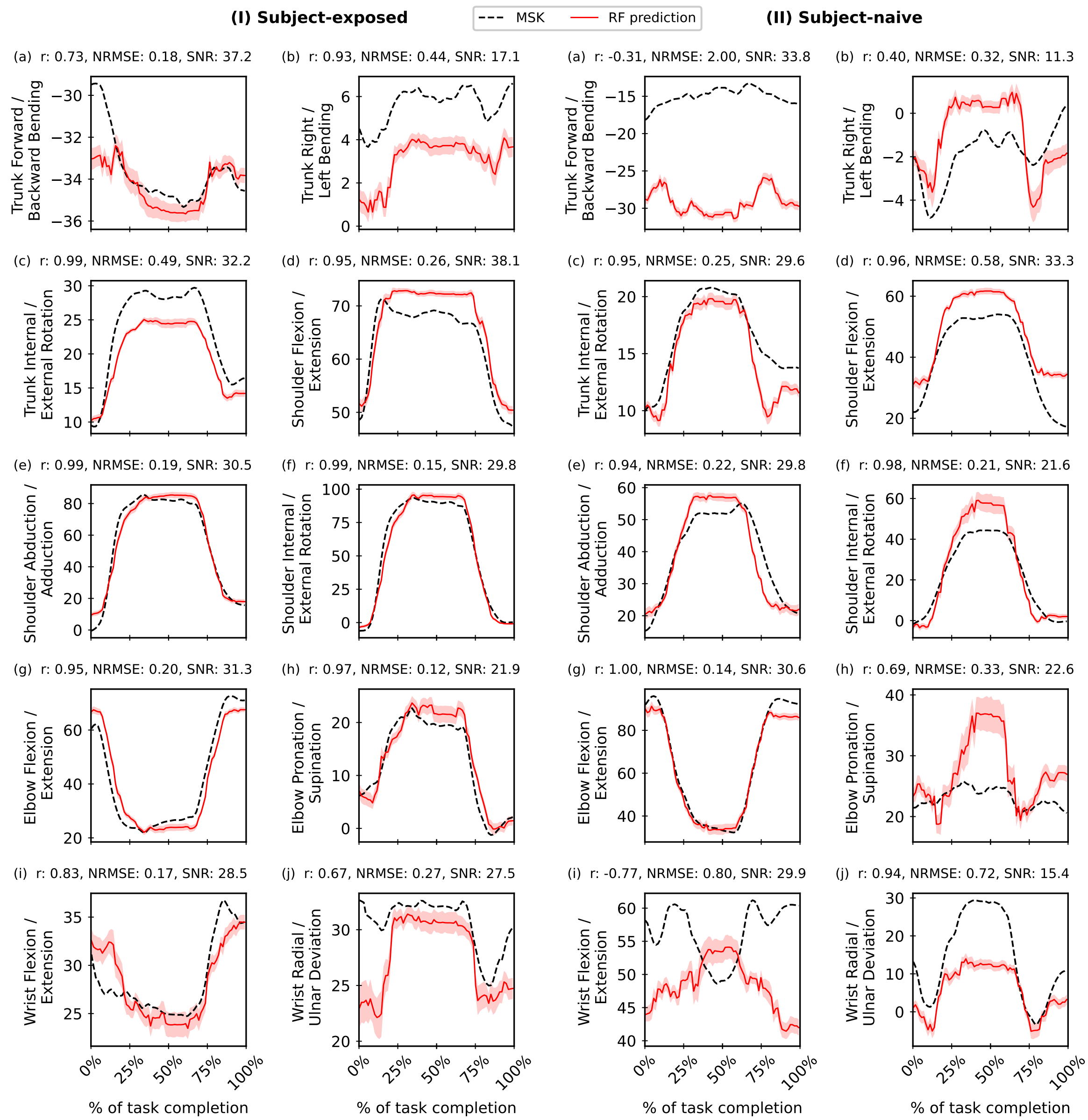


Figure 9: Comparing musculoskeletal (MSK) model outputs with the Random Forests (RF) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the RF predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 29.4±6.01 and 25.79±7.33 decibels.

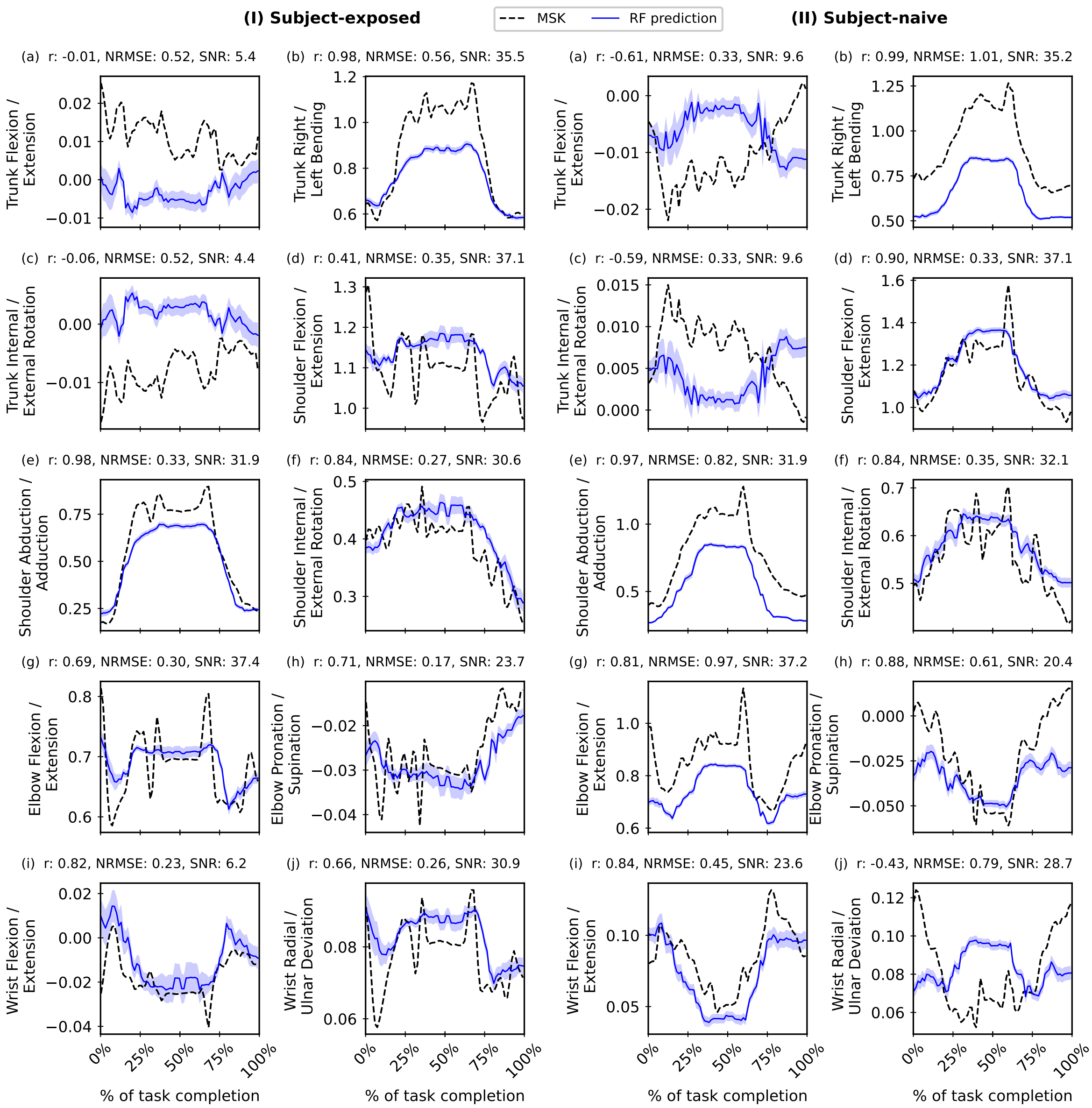


Figure 10: Comparing musculoskeletal (MSK) model outputs with the Random Forests (RF) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the RF predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 24.32±12.97 and 26.54±9.92 decibels.

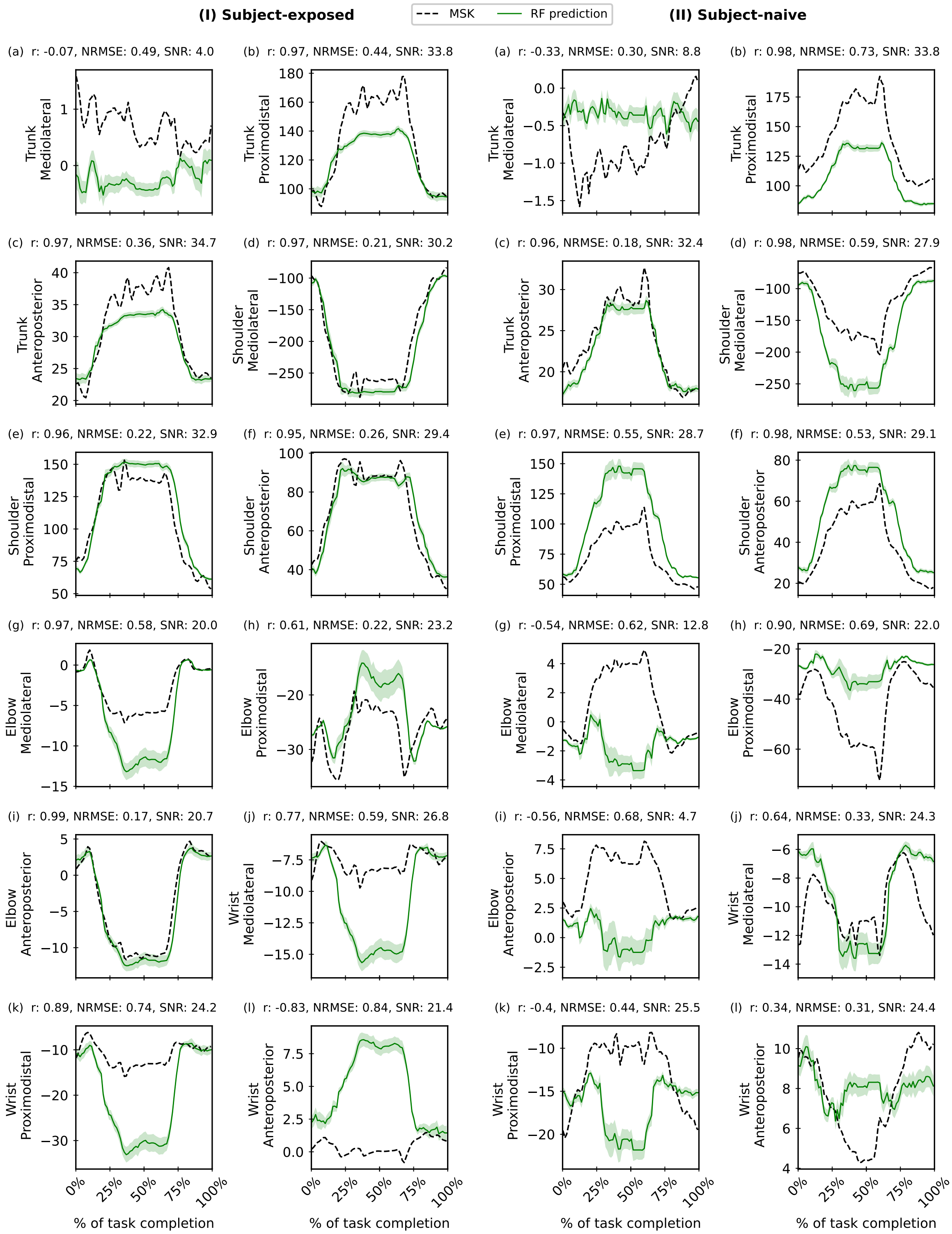


Figure 11: Comparing musculoskeletal (MSK) model outputs with the Random Forests (RF) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the RF predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 25.12±8.09 and 22.87±8.89 decibels.

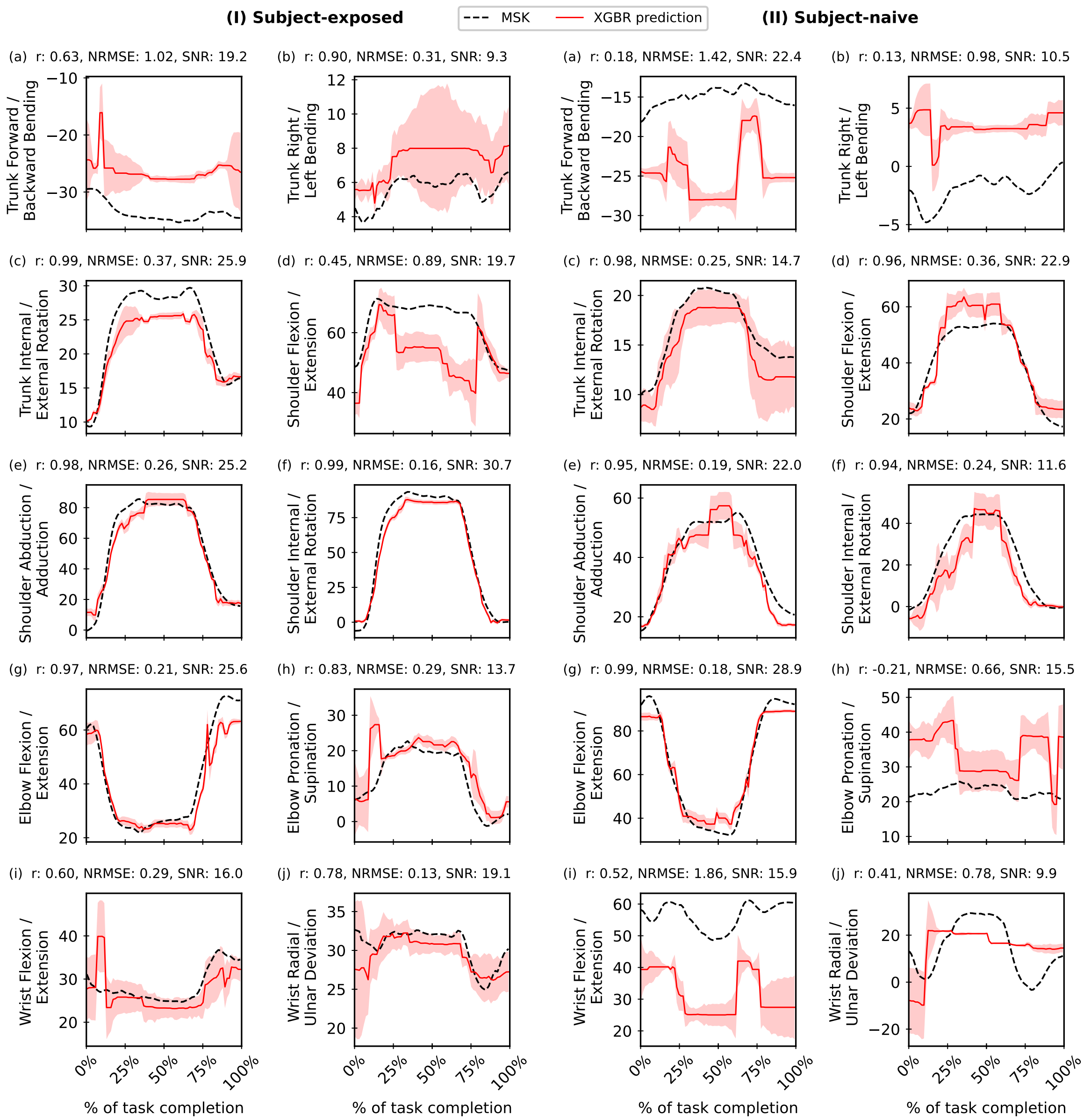


Figure 12: Comparing musculoskeletal (MSK) model outputs with the Extreme Gradient Boosting (XGBR) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the XGBR predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 20.44±6.14 and 17.42±6.01 decibels.

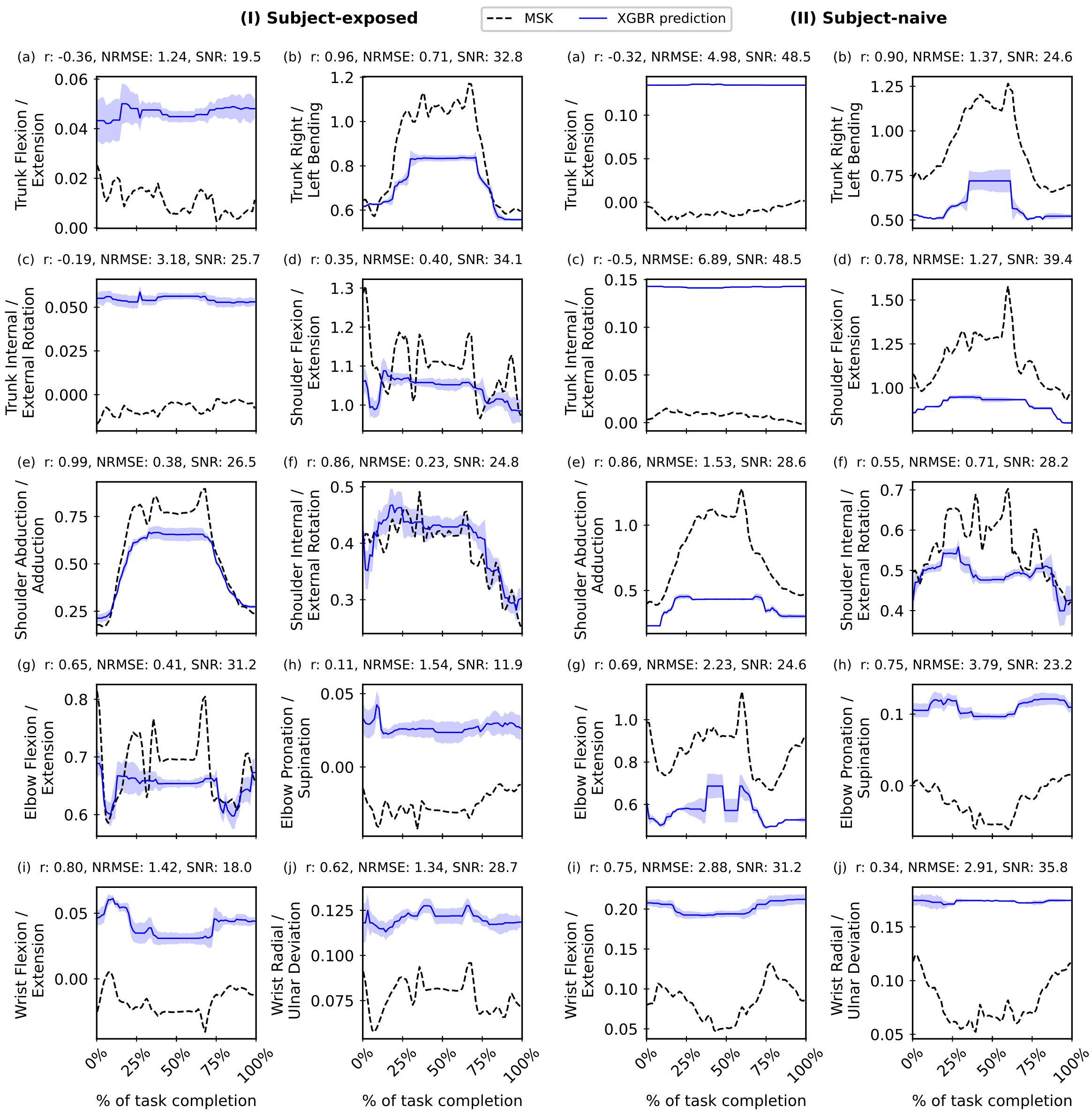


Figure 13: Comparing musculoskeletal (MSK) model outputs with the Extreme Gradient Boosting (XGBR) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the XGBR predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 25.32±6.69 and 33.27±9.01 decibels.

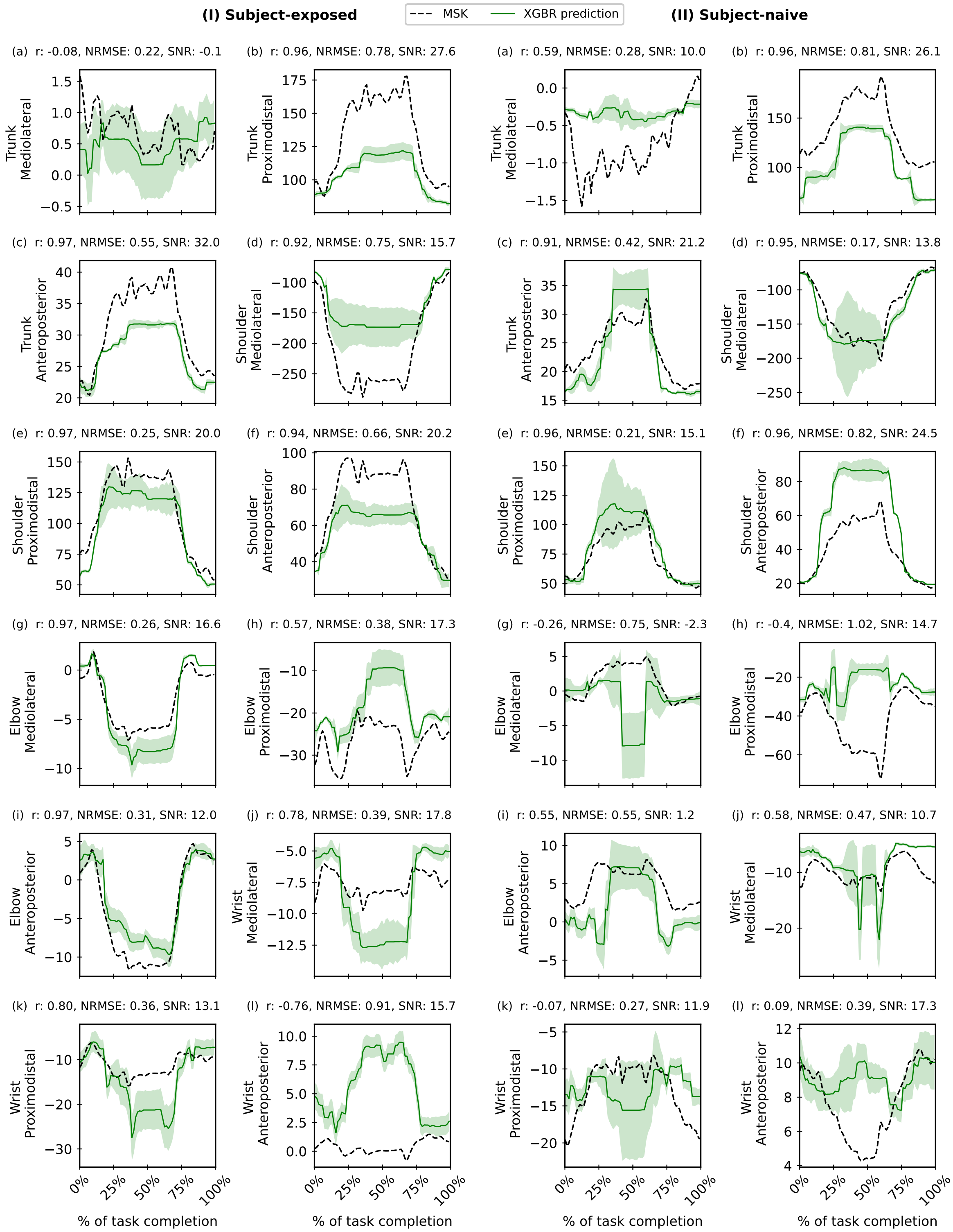


Figure 14: Comparing musculoskeletal (MSK) model outputs with the Extreme Gradient Boosting (XGBR) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the XGBR predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 17.32±7.58 and 13.68±8.06 decibels.

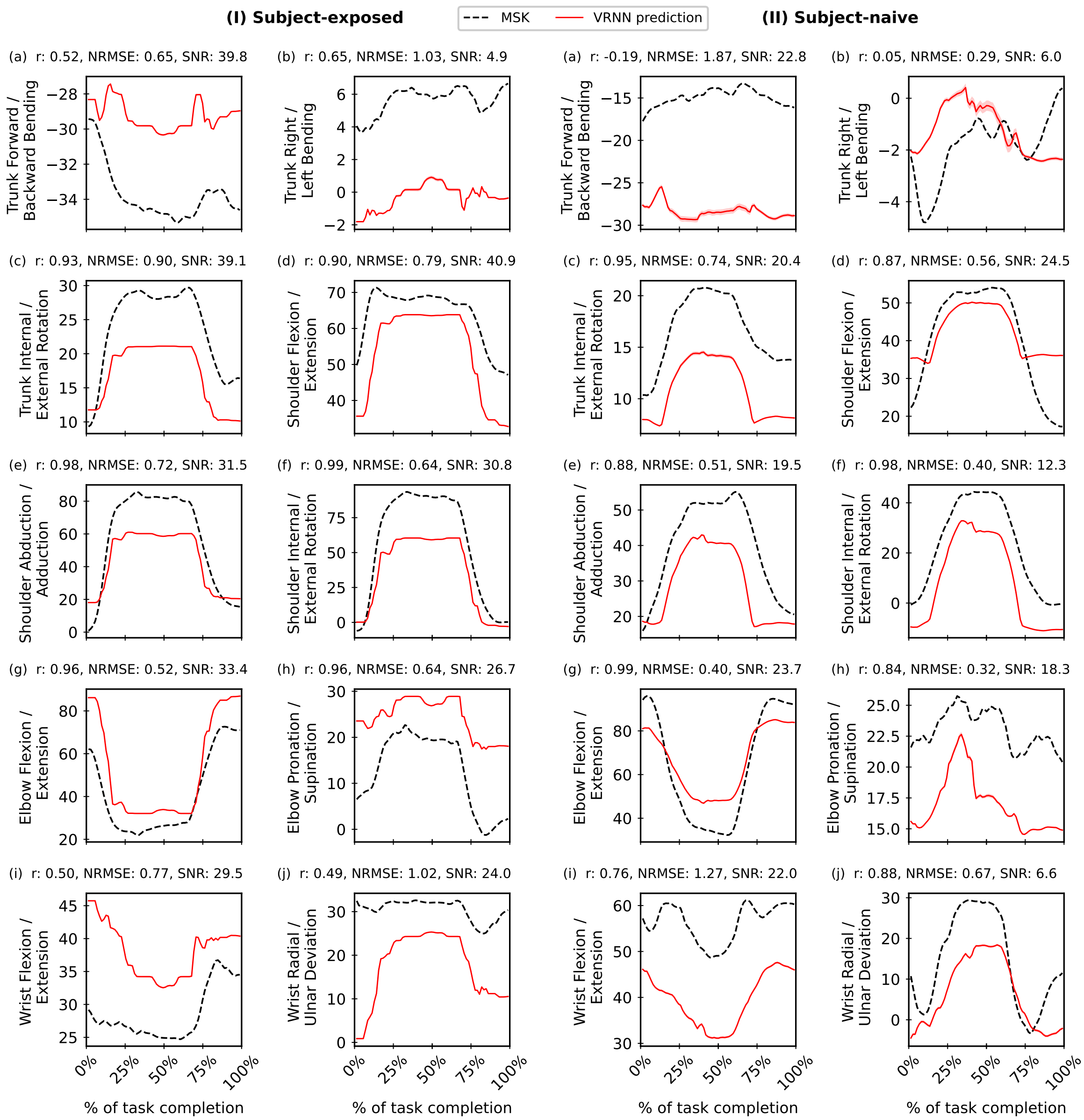


Figure 15: Comparing musculoskeletal (MSK) model outputs with the Vanilla Recurrent Neural Networks (VRNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the VRNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 30.07±9.96 and 17.7±6.56 decibels.

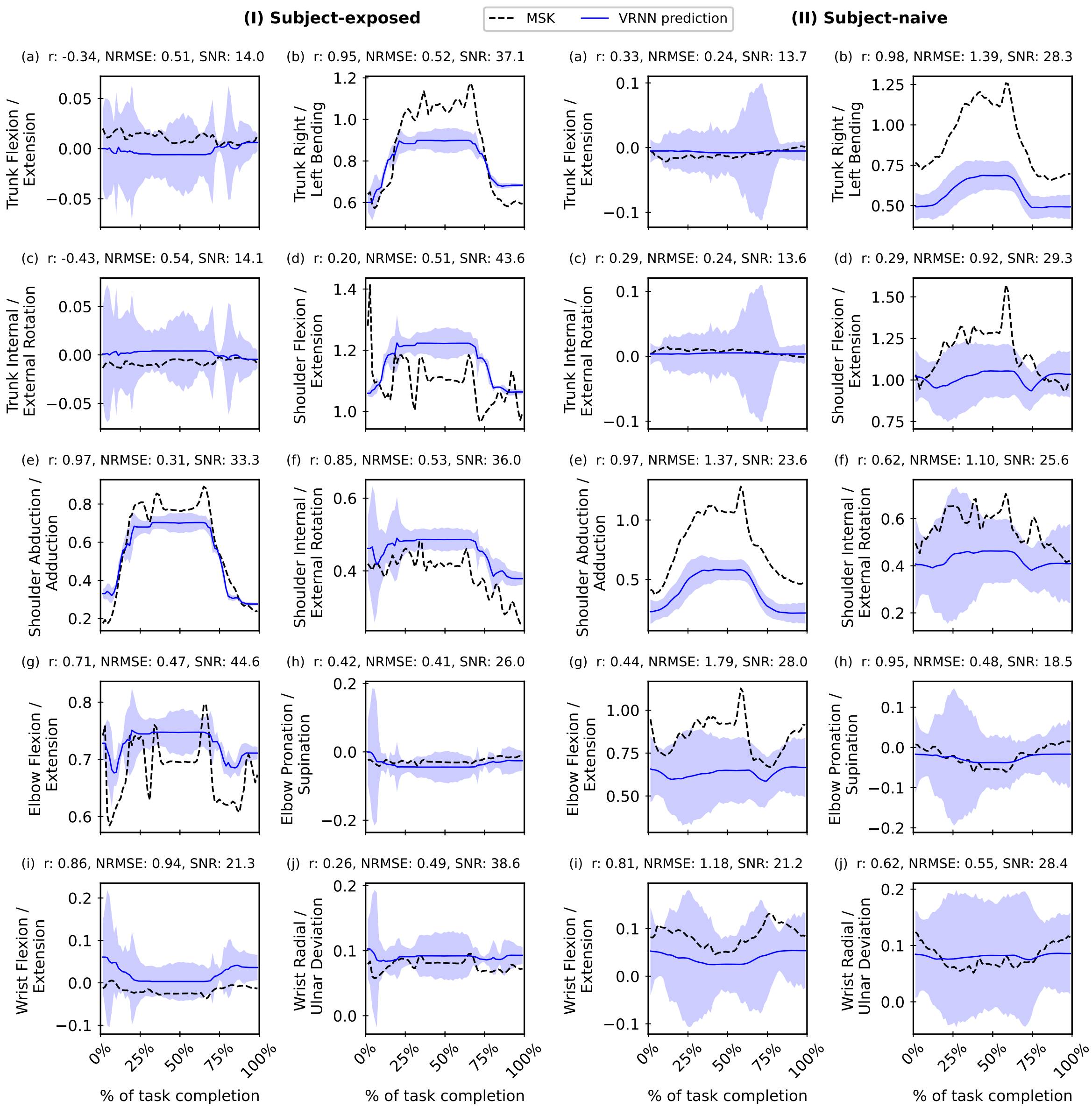


Figure 16: Comparing musculoskeletal (MSK) model outputs with the Vanilla Recurrent Neural Networks (VRNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the VRNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 30.85±10.78 and 23.12±5.72 decibels.

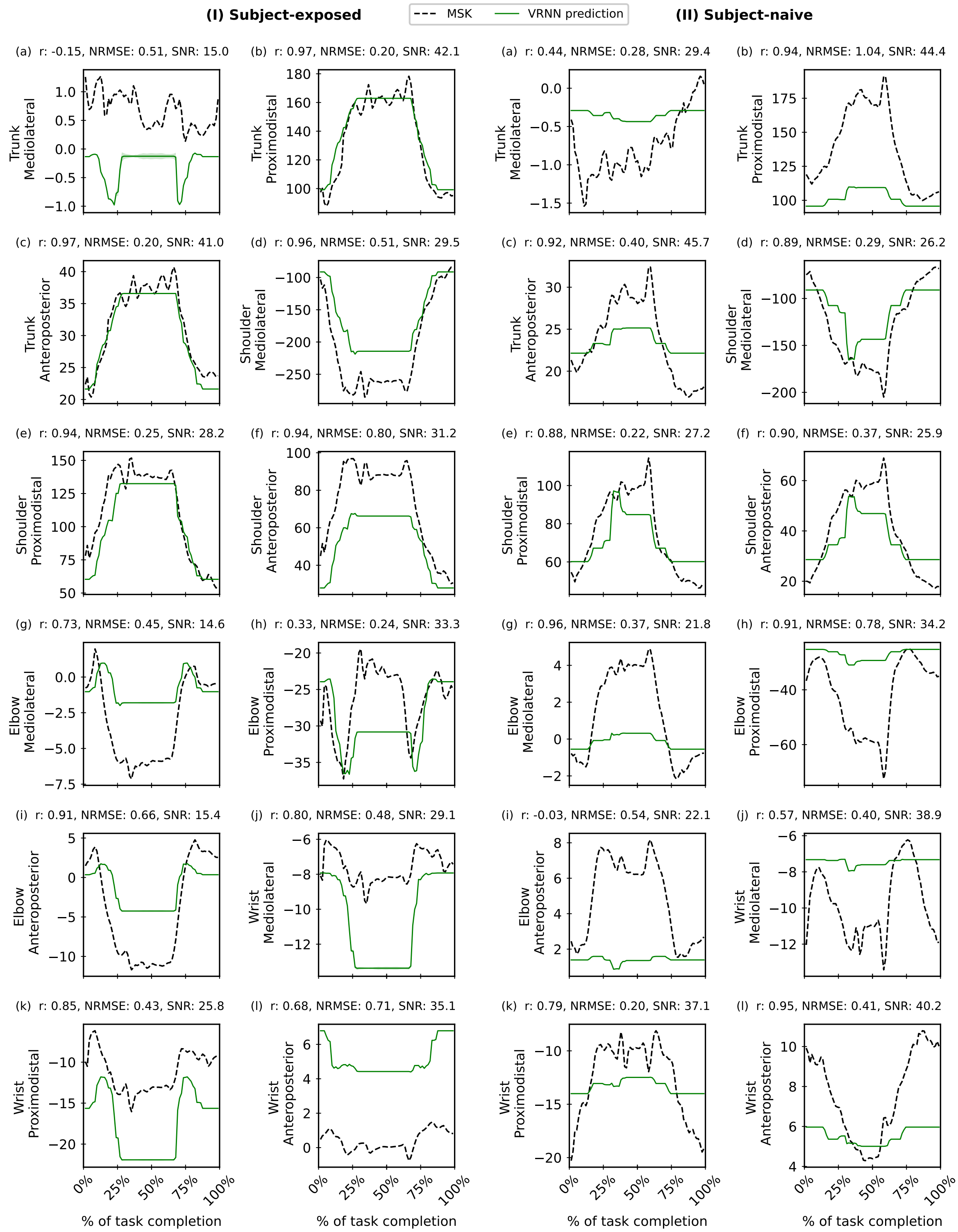


Figure 17: Comparing musculoskeletal (MSK) model outputs with the Vanilla Recurrent Neural Networks (VRNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the VRNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 28.4±8.95 and 32.78±8.08 decibels.

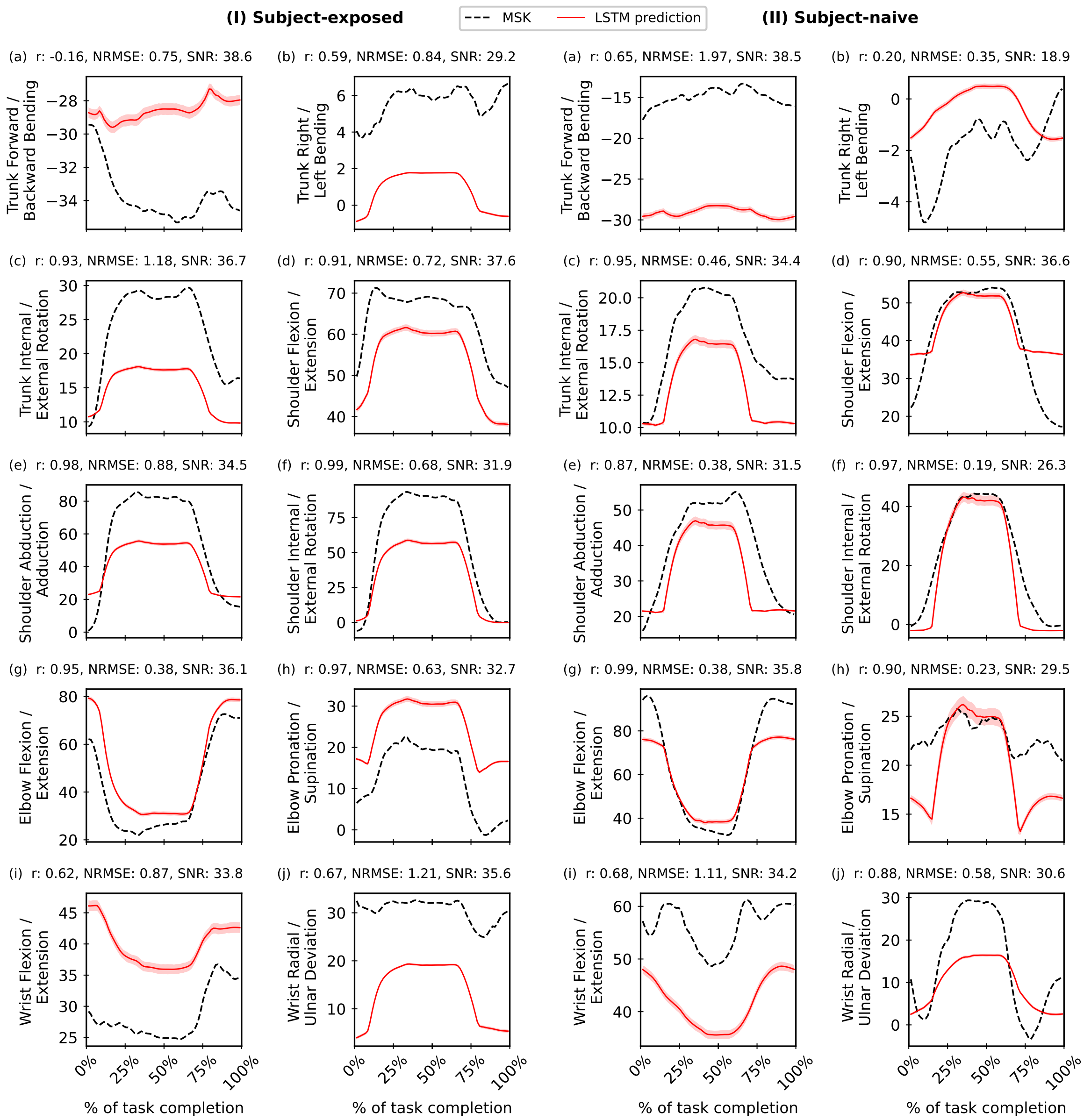


Figure 18: Comparing musculoskeletal (MSK) model outputs with the Long Short-Term Memory (LSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 34.68±2.71 and 31.61±5.46 decibels.

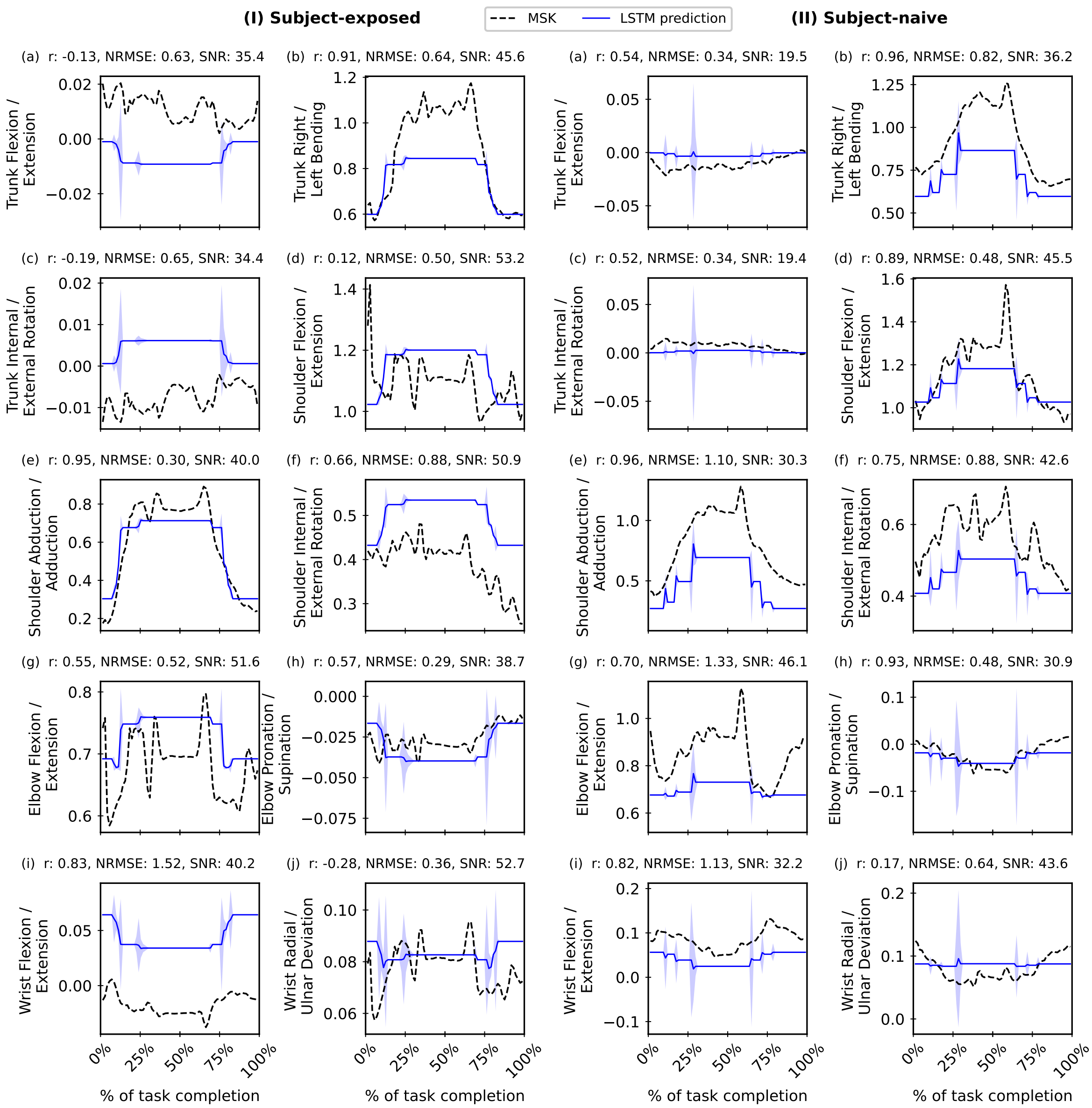


Figure 19: Comparing musculoskeletal (MSK) model outputs with the Long Short-Term Memory (LSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 44.15±7.04 and 34.68±9.45 decibels.

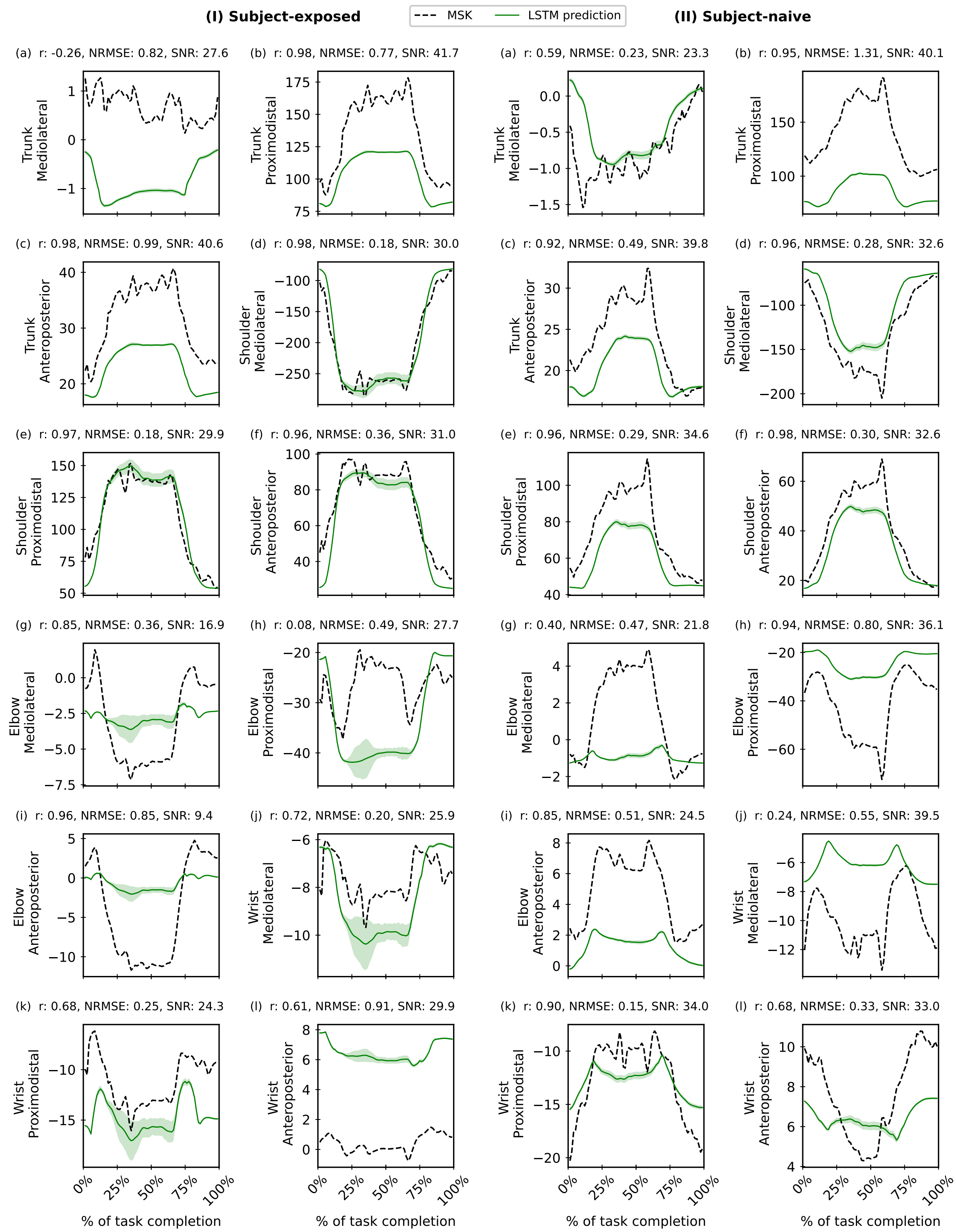


Figure 20: Comparing musculoskeletal (MSK) model outputs with the Long Short-Term Memory (LSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the LSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 27.9±8.45 and 32.65±6.06 decibels.

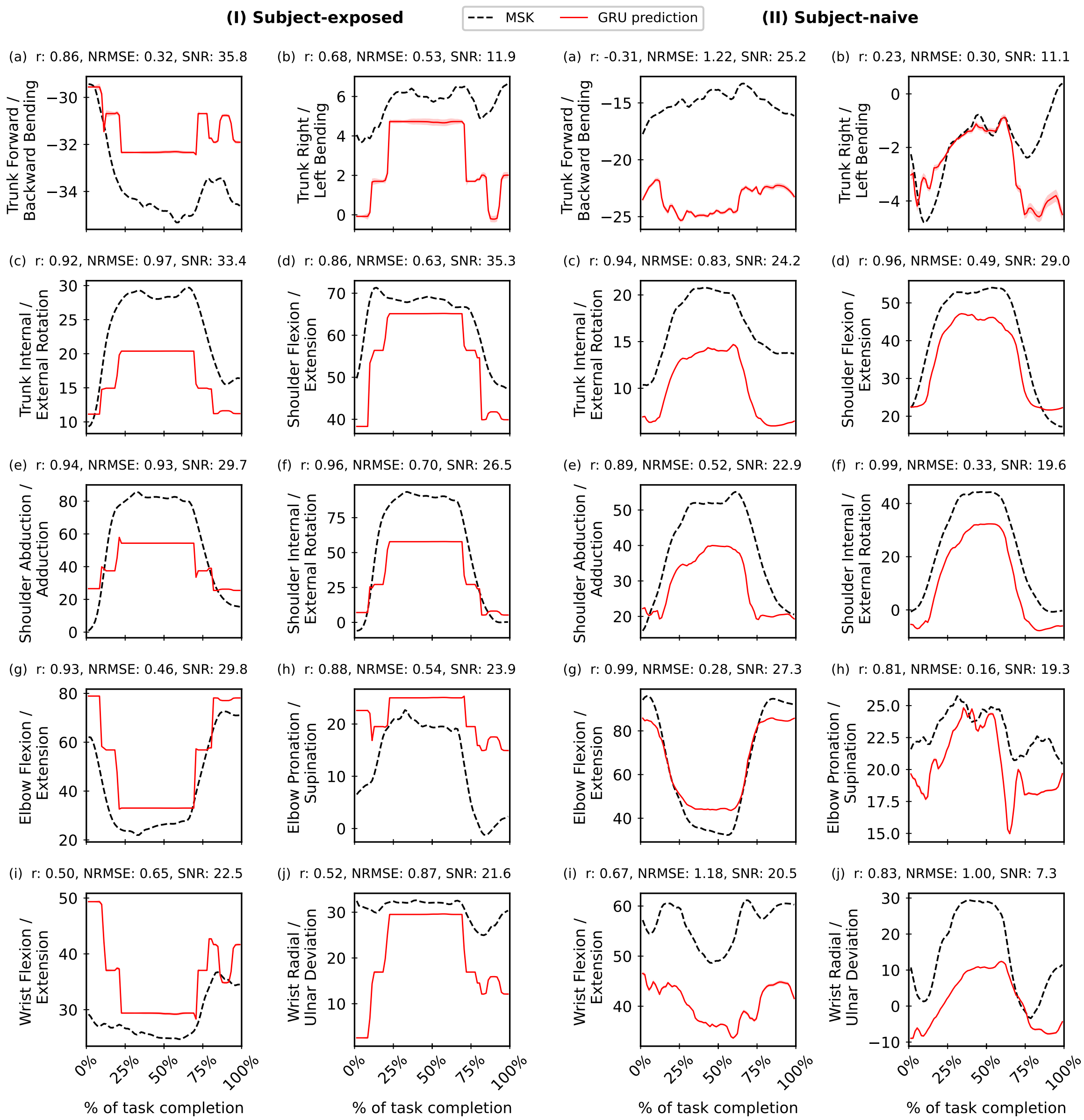


Figure 21: Comparing musculoskeletal (MSK) model outputs with the Gated Recurrent Unit (GRU) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the GRU predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 27.03±6.98 and 20.58±6.5 decibels.

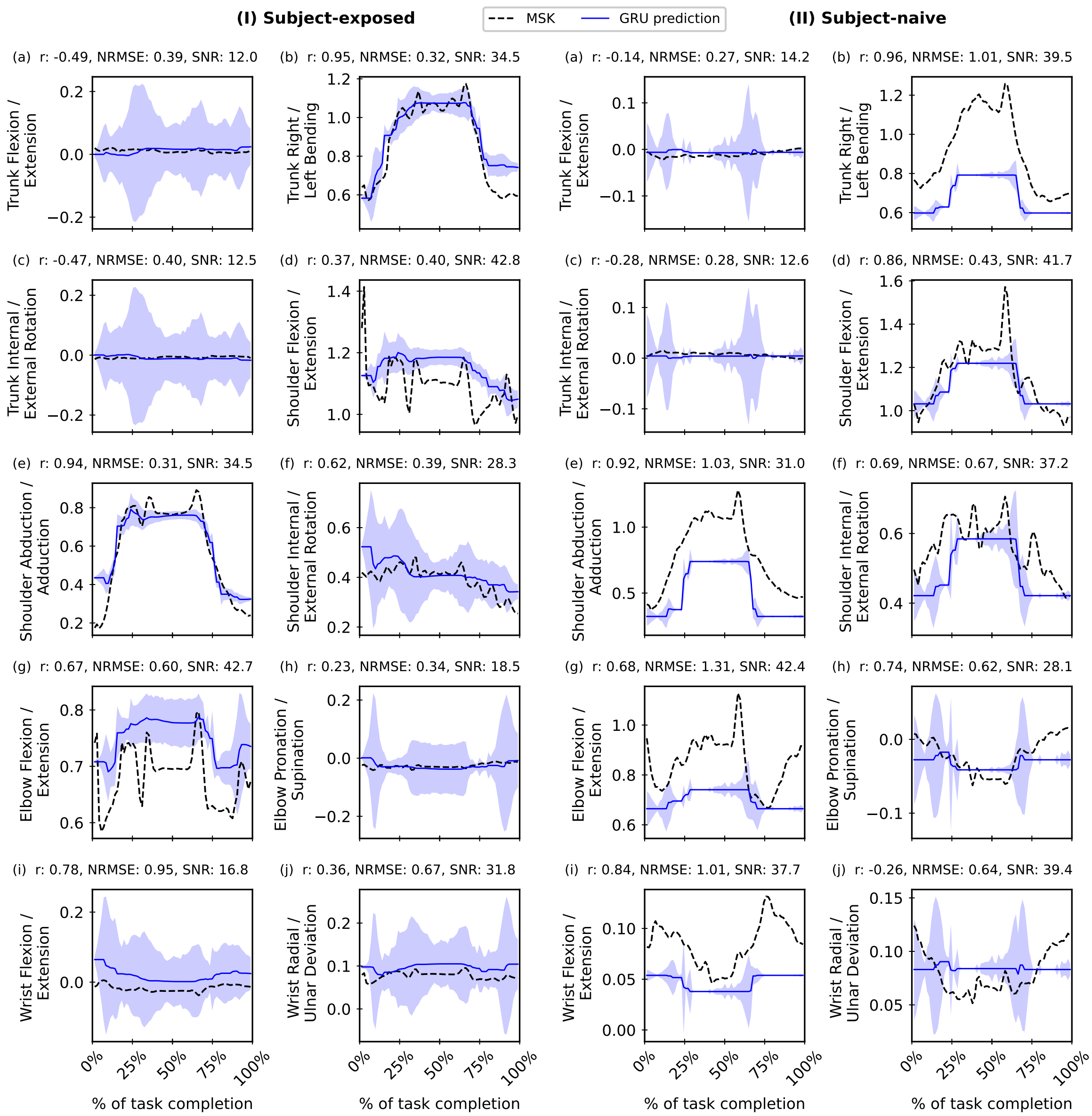


Figure 22: Comparing musculoskeletal (MSK) model outputs with the Gated Recurrent Unit (GRU) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the GRU predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 27.47±11.16 and 32.34±10.38 decibels.

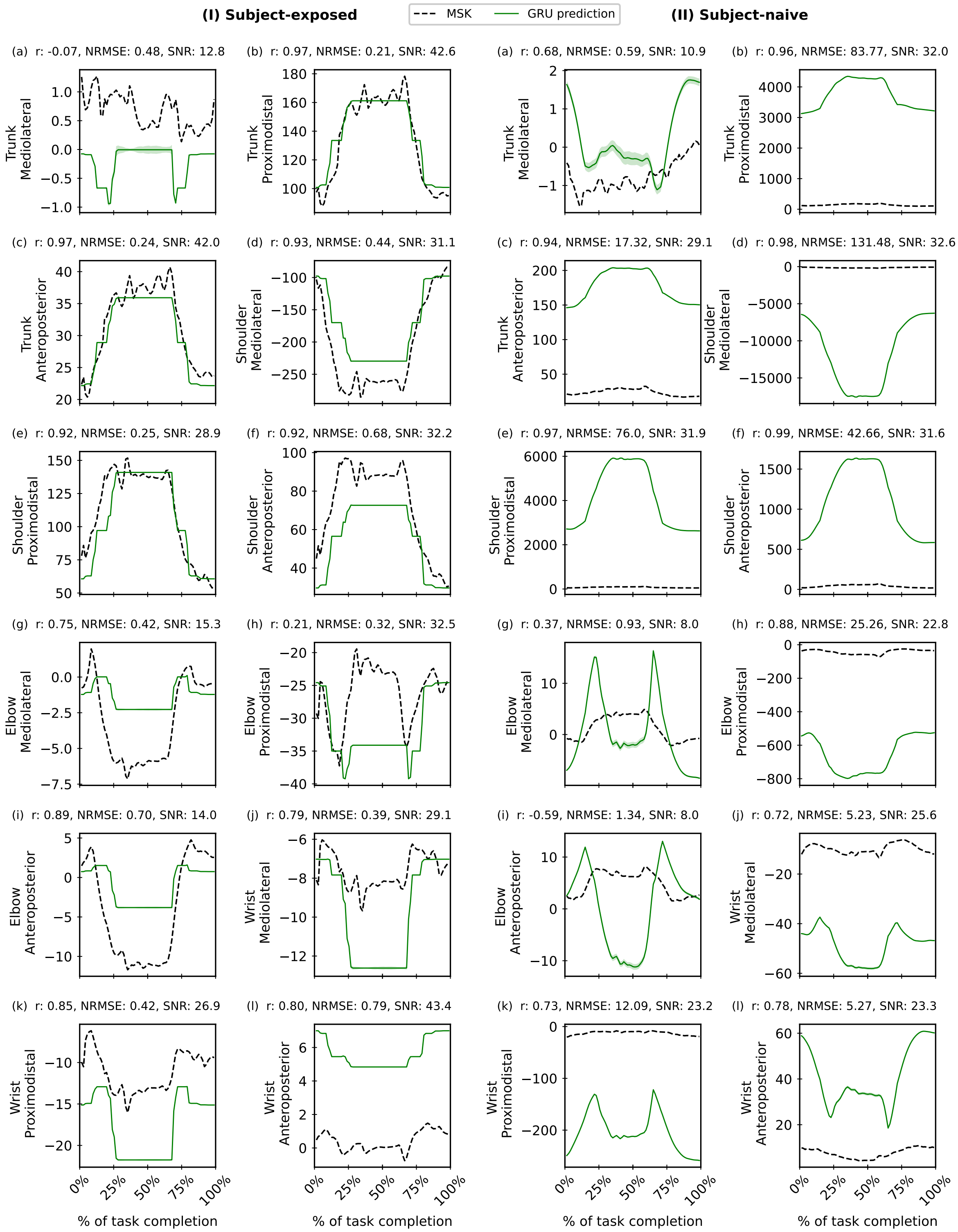


Figure 23: Comparing musculoskeletal (MSK) model outputs with the Gated Recurrent Unit (GRU) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the GRU predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 29.17±10.27 and 23.19±8.98 decibels.

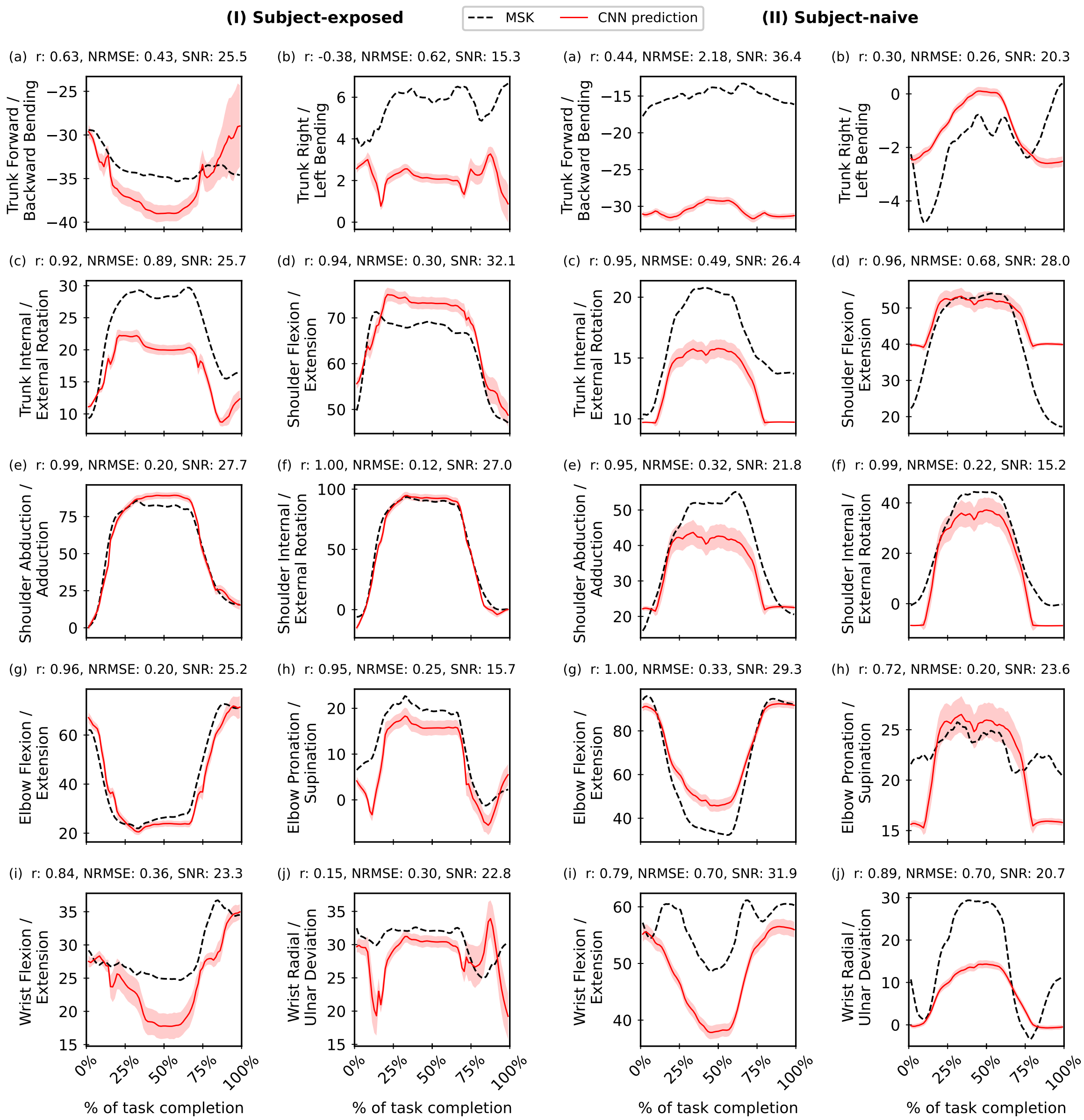


Figure 24: Comparing musculoskeletal (MSK) model outputs with the Convolutional Neural Networks (CNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 24.04±4.93 and 25.34±5.94 decibels.

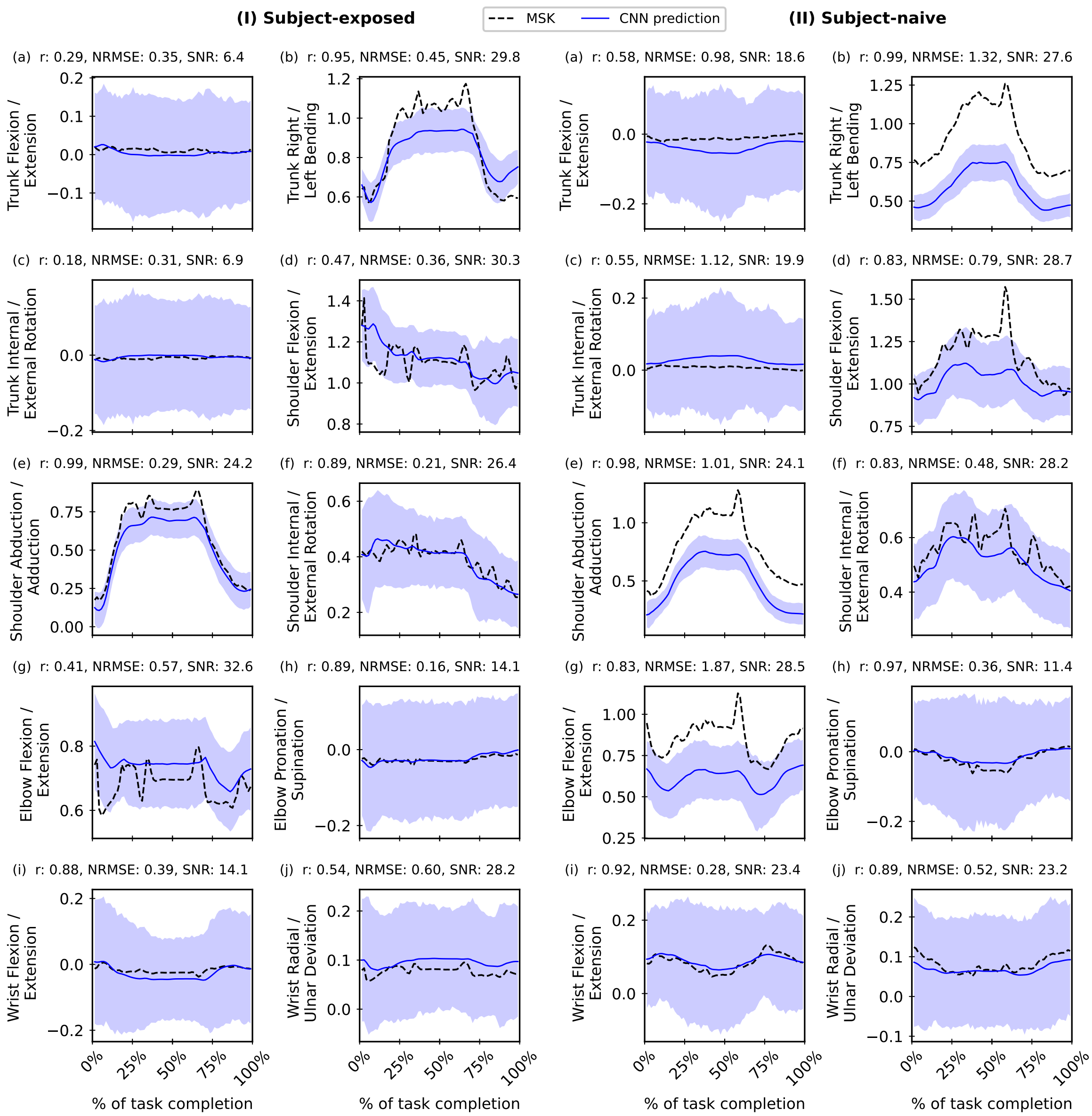


Figure 25: Comparing musculoskeletal (MSK) model outputs with the Convolutional Neural Networks (CNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 21.27±9.45 and 23.38±5.28 decibels.

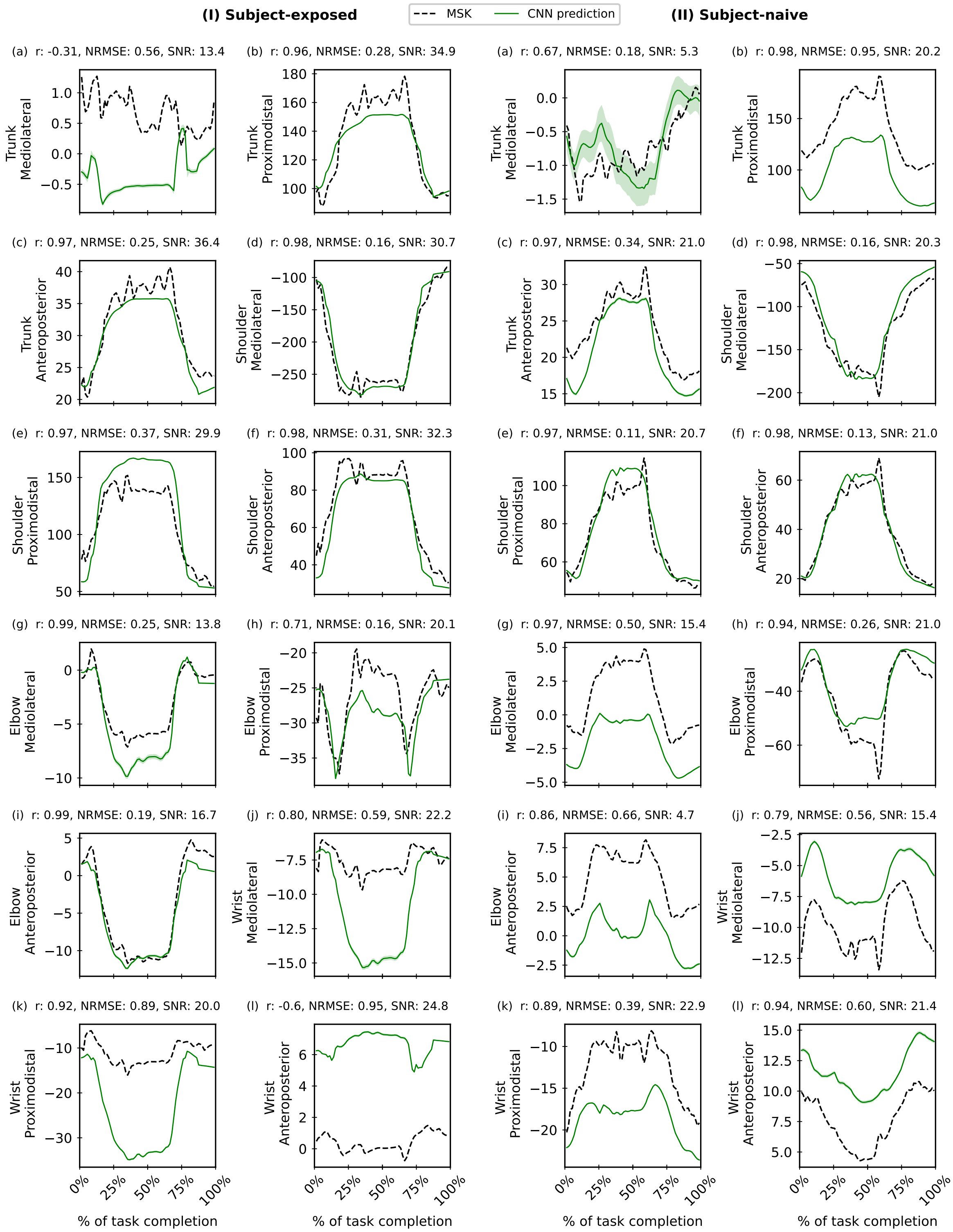


Figure 26: Comparing musculoskeletal (MSK) model outputs with the Convolutional Neural Networks (CNN) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNN predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 24.58±7.75 and 17.43±5.97 decibels.

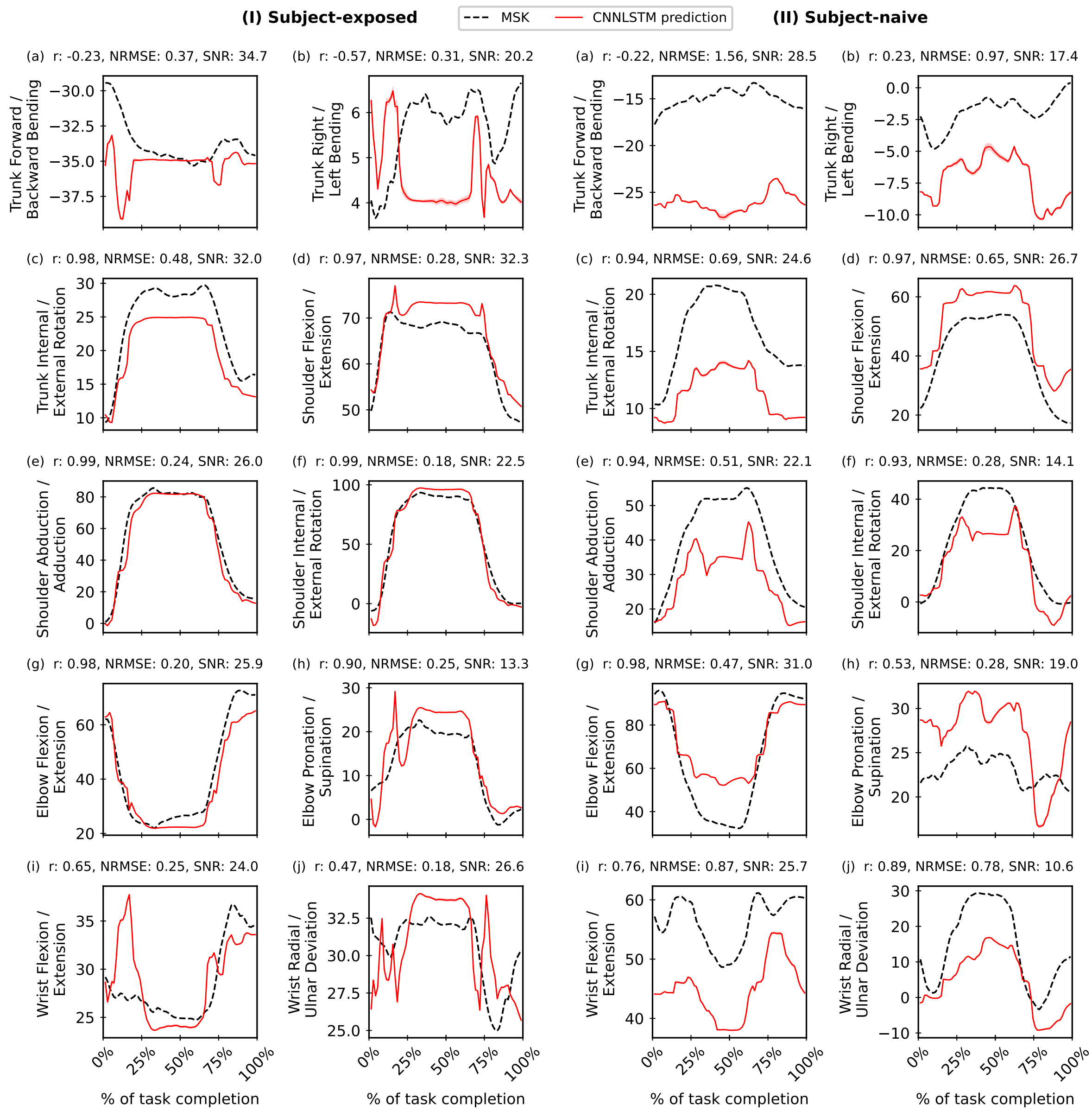


Figure 27: Comparing musculoskeletal (MSK) model outputs with the CNNLSTM (CNNLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNNLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 25.86±5.99 and 21.97±6.24 decibels.

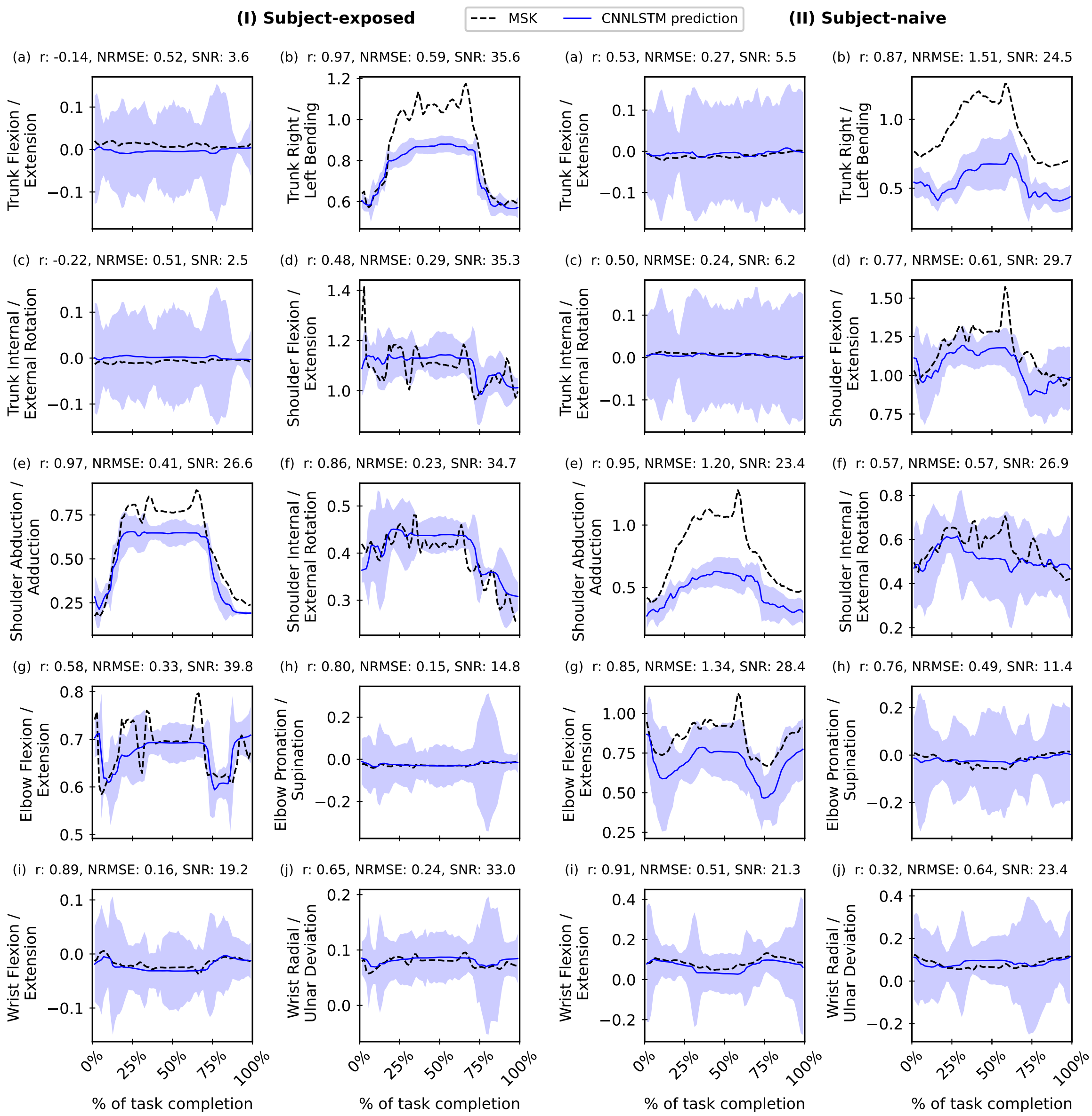


Figure 28: Comparing musculoskeletal (MSK) model outputs with the CNNLSTM (CNNLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNNLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 24.49±13.0 and 20.09±8.54 decibels.

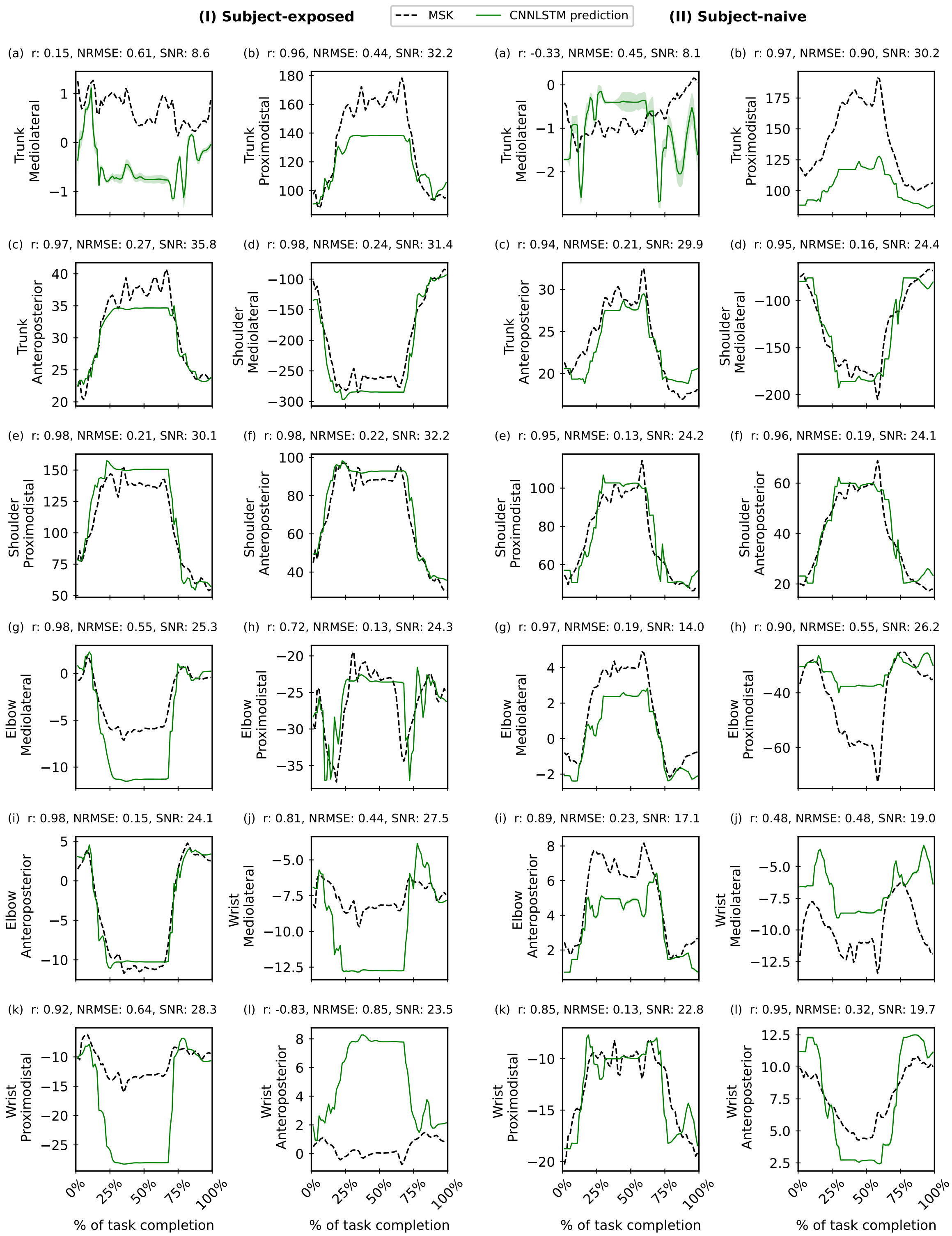


Figure 29: Comparing musculoskeletal (MSK) model outputs with the CNNLSTM (CNNLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the CNNLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 26.91 ± 6.66 and 21.64 ± 6.18 decibels.

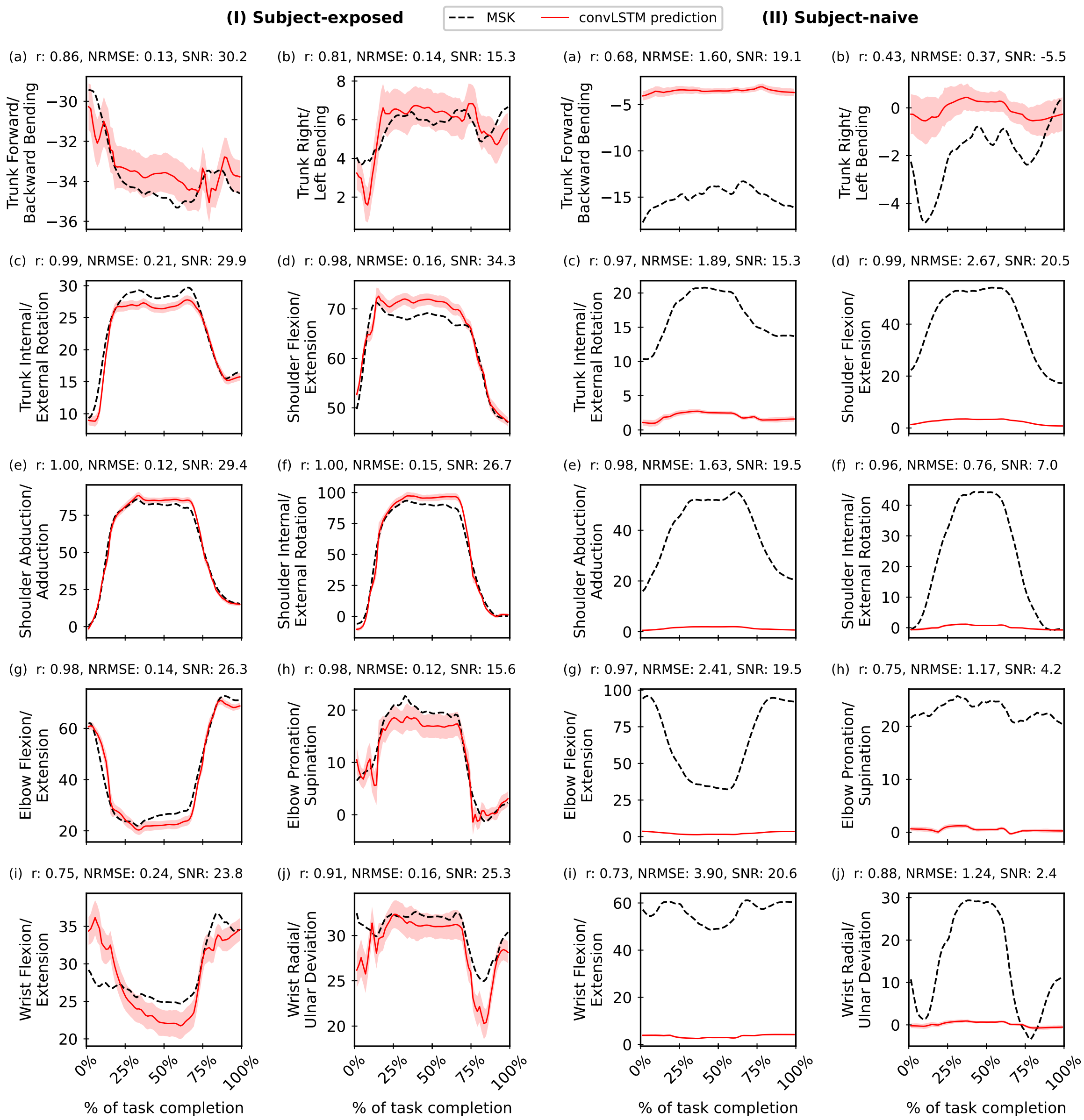


Figure 30: Comparing musculoskeletal (MSK) model outputs with the Convolutional LSTM (convLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the convLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 25.64±5.82 and 12.23±8.9 decibels.

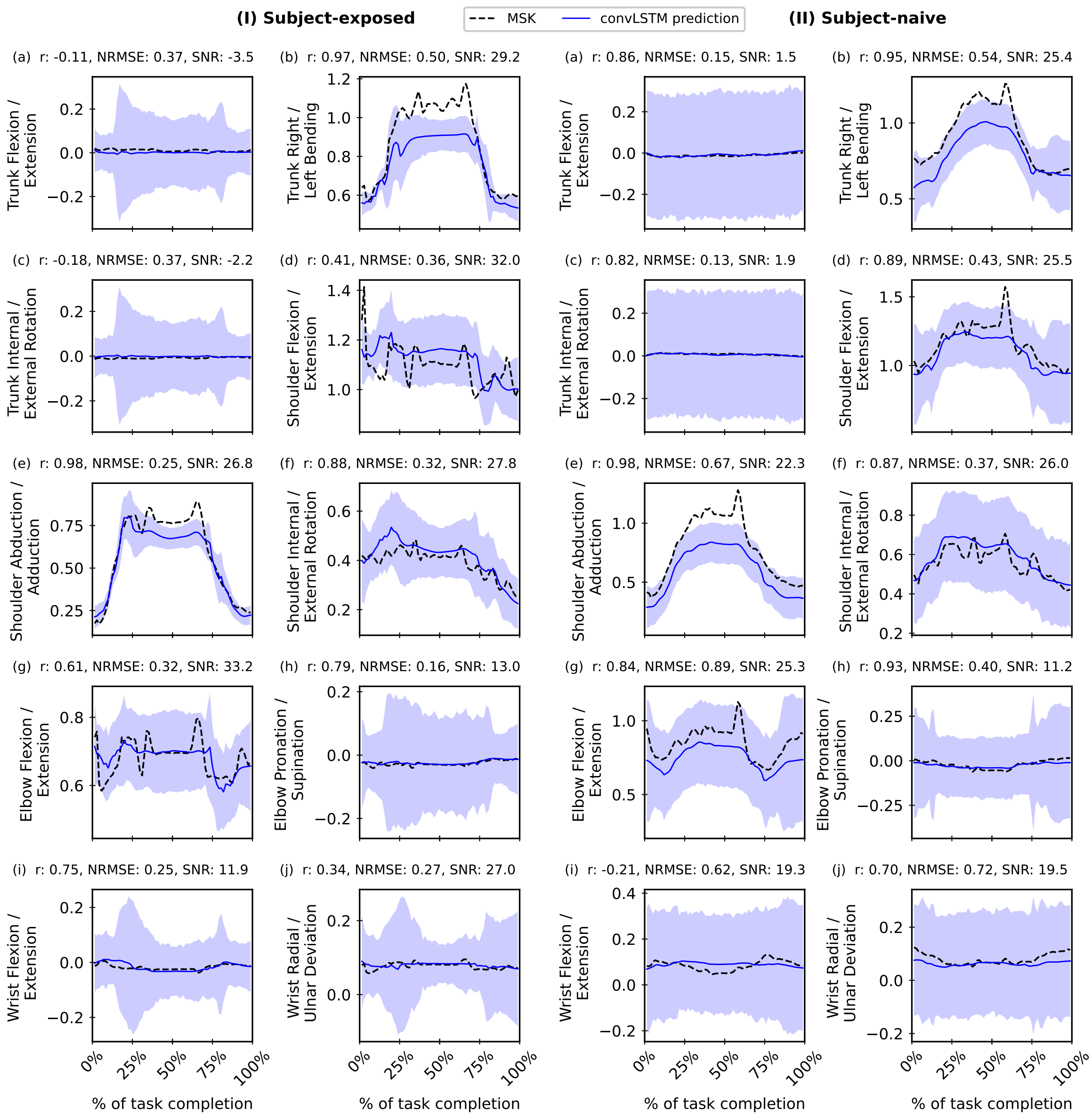


Figure 31: Comparing musculoskeletal (MSK) model outputs with the Convolutional LSTM (convLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the convLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 19.52±13.09 and 17.78±9.08 decibels.

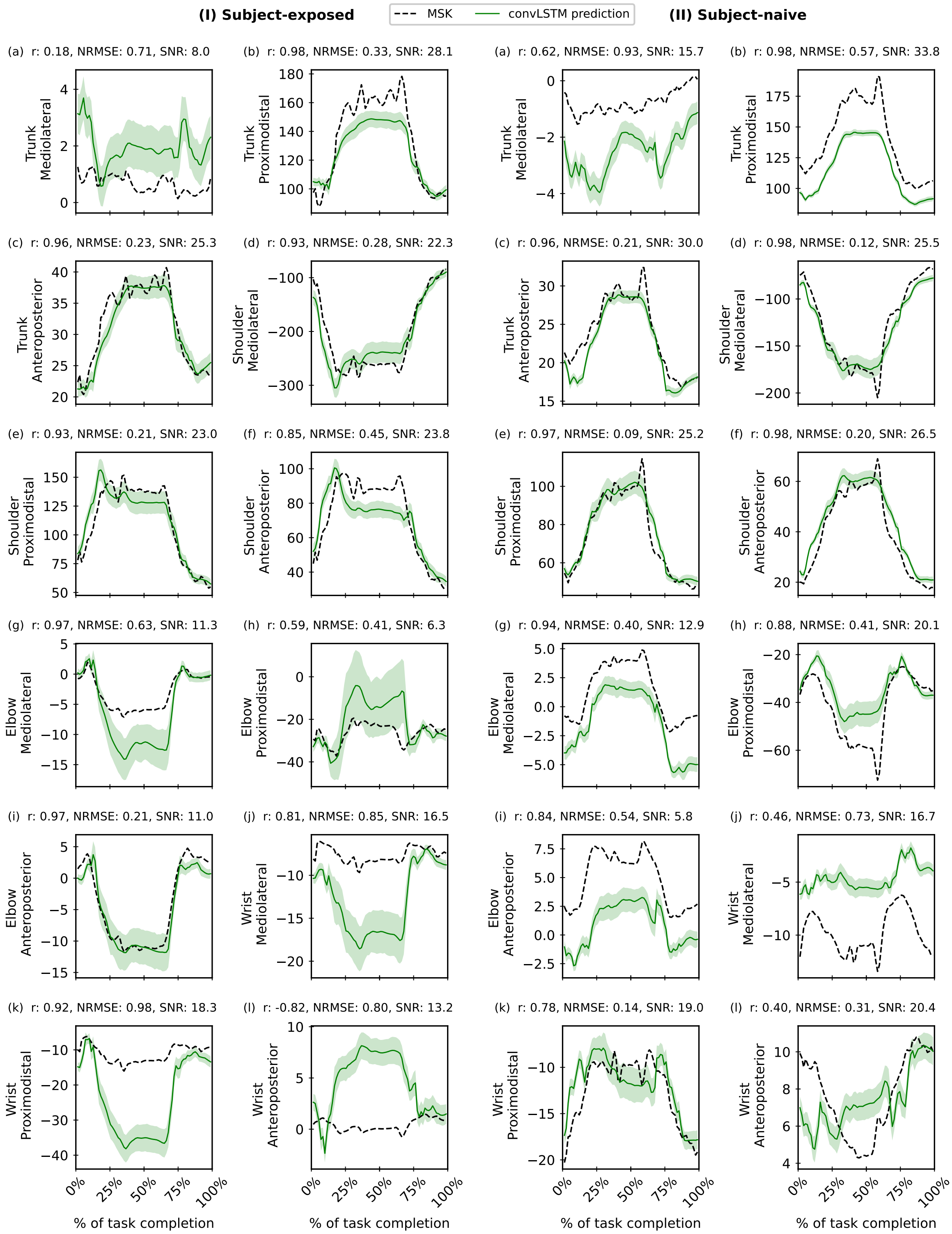
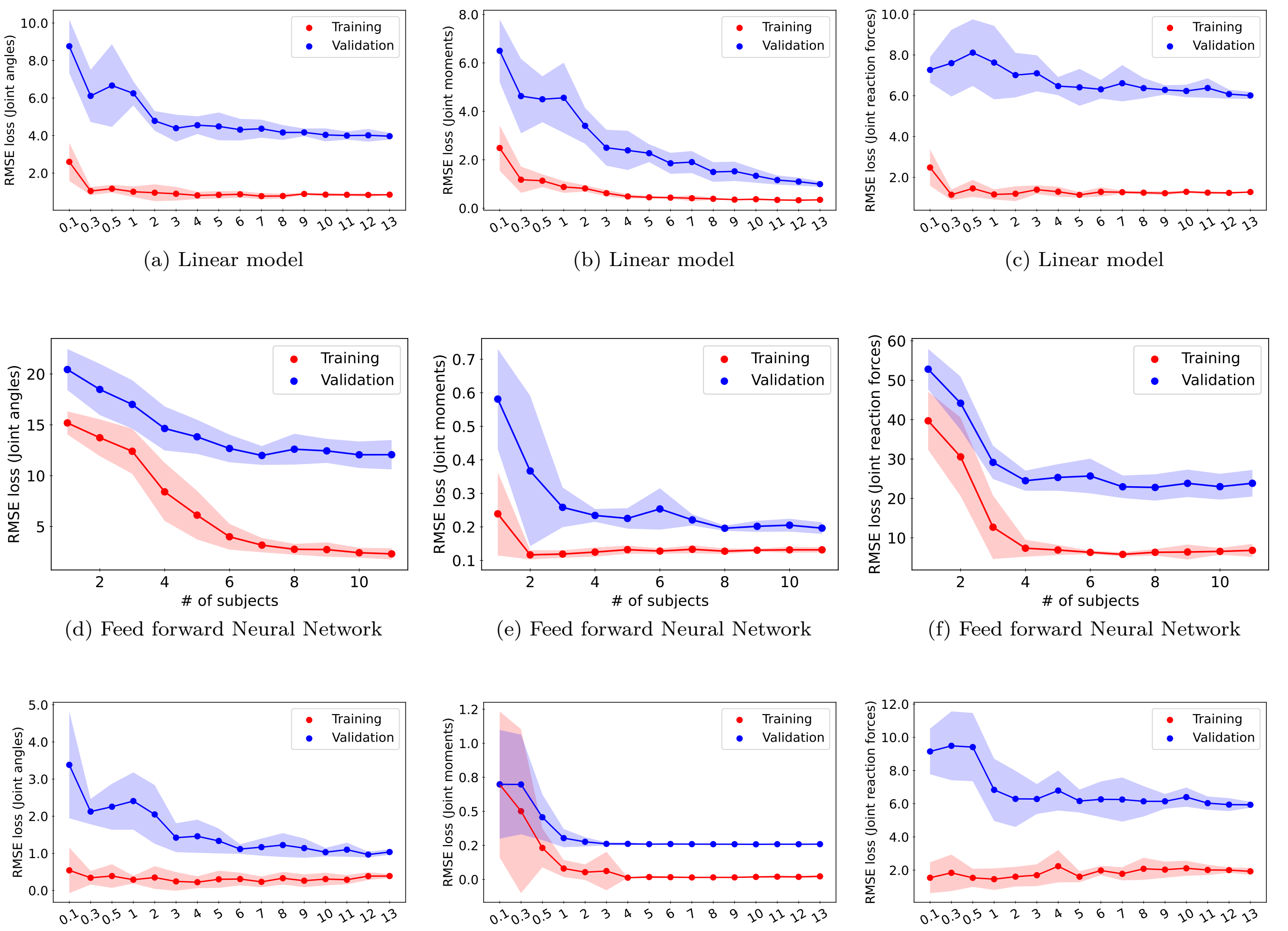


Figure 32: Comparing musculoskeletal (MSK) model outputs with the Convolutional LSTM (convLSTM) predictions for a trial in the held-out test data in Subject-exposed (left) and Subject-naive (right) settings. Additionally, the OMC input is subjected to simulated noises/inherent errors in soft tissue artefacts, marker misplacement, and OMC motion capture system instrumental error, and the robustness of the ML models is illustrated by the variability in the convLSTM predictions. Robustness to added noise is quantified in terms of signal-to-noise ratio (SNR) and is defined in Methods. The mean SNRs across all output categories for subject-exposed and subject-naive models are 17.25±6.96 and 20.98±7.43 decibels.

4 Learning curves



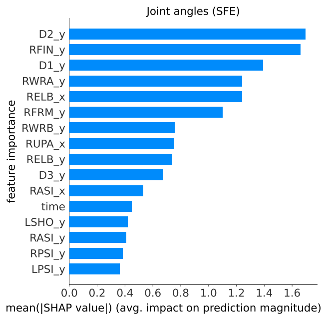
(a) Linear model (b) Linear model (c) Linear model

(d) Feed forward Neural Network (e) Feed forward Neural Network (f) Feed forward Neural Network

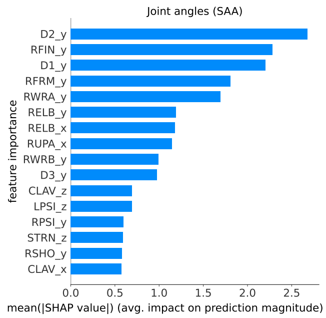
(g) convLSTM (h) convLSTM (i) convLSTM

Figure 33: Learning curve, i.e., training and validation error vs training data (# of subjects) for different models using optimal hyperparameters (SI section 1). 0.1, 0.3, and 0.5 represent 10, 30, and 50% data of a subject. RMSE loss is the average loss (i.e., the average of losses in all the output variables), and the spread around the mean is the standard deviation of the RMSE loss for ten different combinations of training subjects or multiple random model initialization.

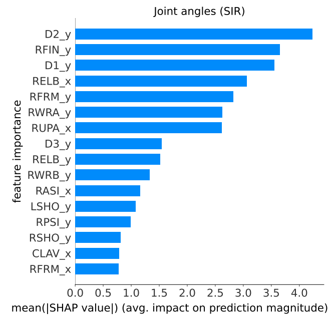
5 Explainable AI



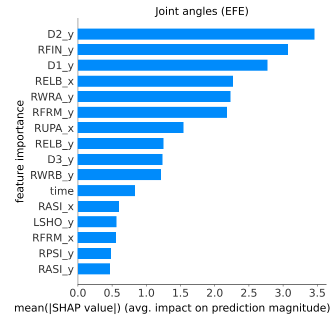
(a)



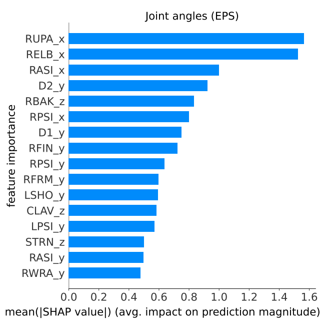
(b)



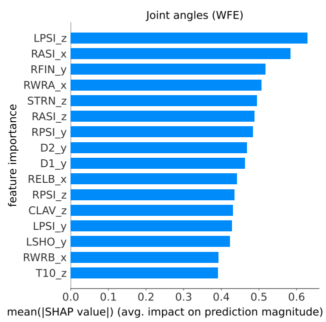
(c)



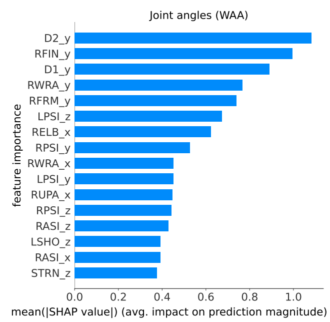
(d)



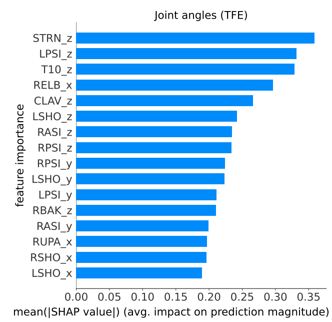
(e)



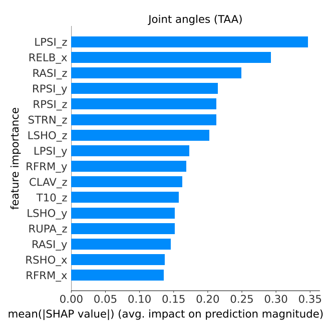
(f)



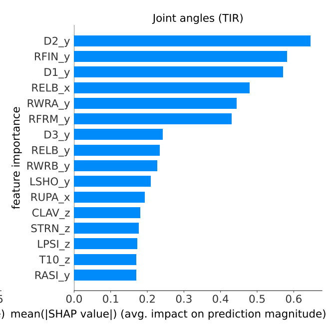
(g)



(h)



(i)



(j)

Figure 34: SHAP results for random forest (joint angles). SHAP values tell the importance of a given feature (marker location, x, y, and z) in explaining the output. Joint angles have ten outputs (details in methods).

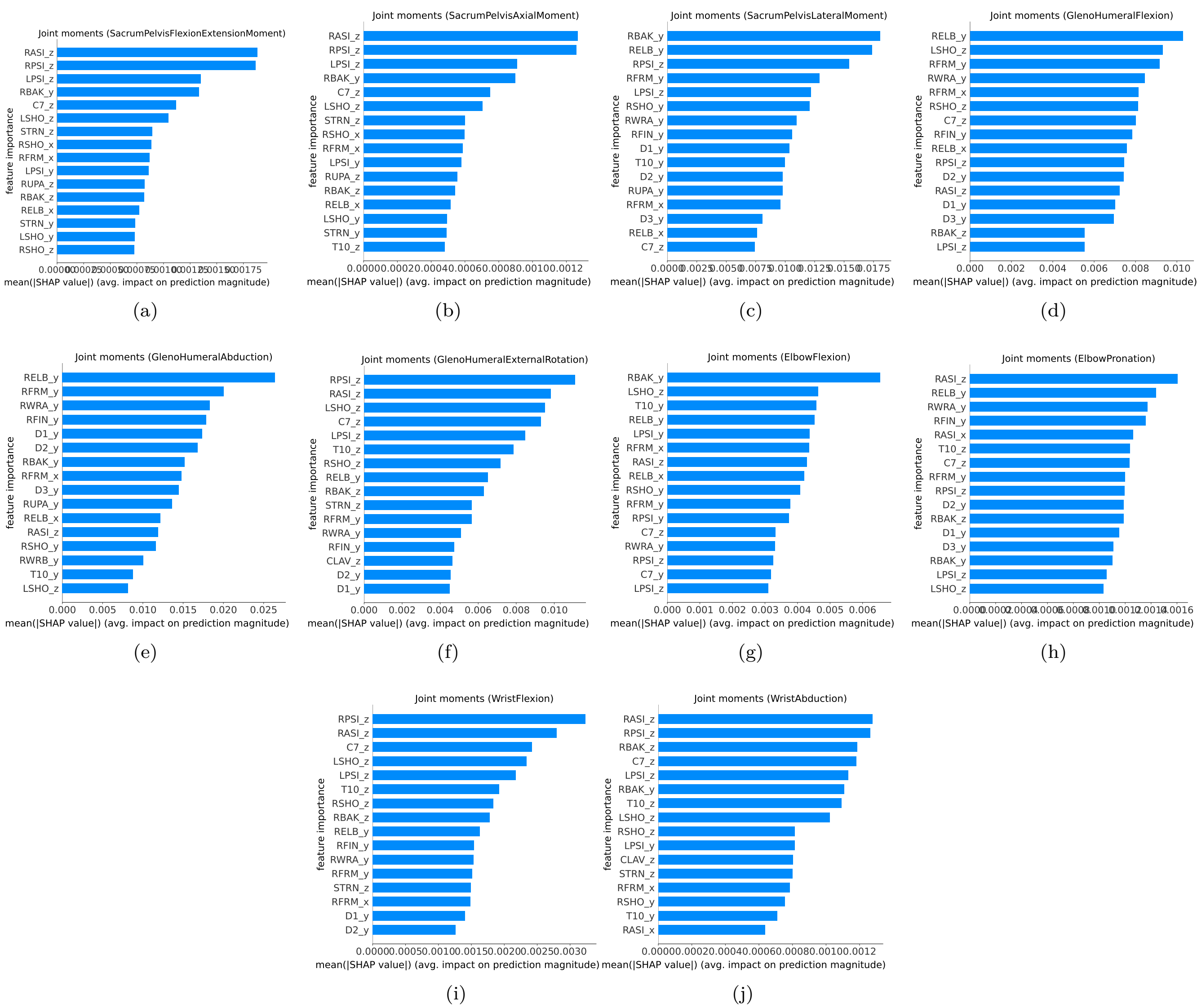


Figure 35: SHAP results for random forest (joint moment). SHAP values tell the importance of a given feature (marker location, x, y, and z) in explaining the output. Joint moments have ten outputs (details in methods).

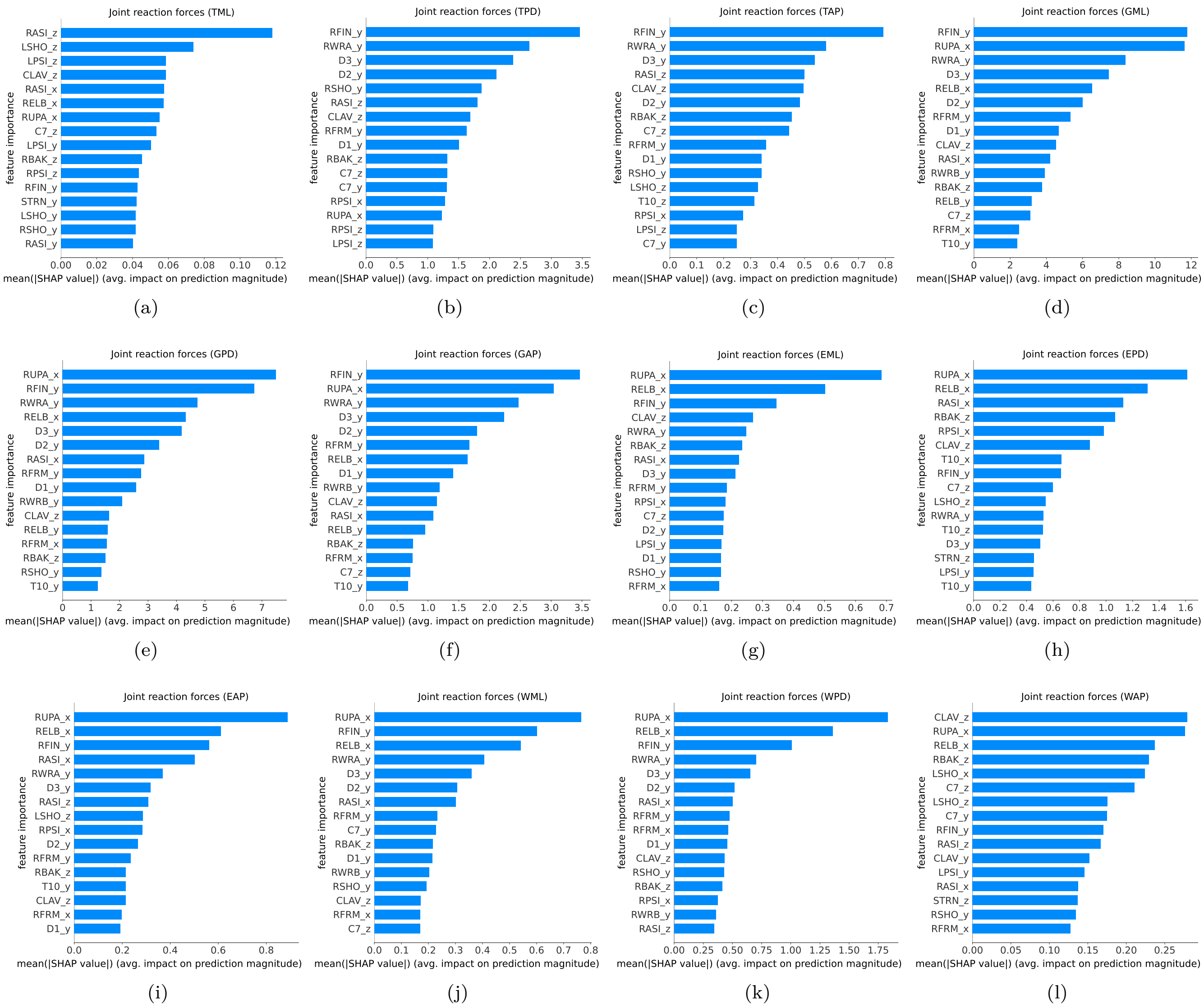


Figure 36: SHAP results for random forest (Joint reaction force). SHAP values tell the importance of a given feature (marker location, x, y, and z) in explaining the output. Joint reaction force has twelve outputs (details in methods).

6 Model complexity

(a) Subject-exposed (Carbon footprint per epoch)

(b) Subject-naive (Carbon footprint per epoch)

(c) Subject-exposed (Carbon footprint)

(d) Subject-naive (Carbon footprint)

(e) Subject-exposed (Prediction time)

(f) Subject-naive (Prediction time)

(g) Subject-exposed (# parameters)

(h) Subject-naive (# parameters)

Figure 37: The figure reports computational costs of training various ML models in terms of CO_2 gram equivalent and prediction time for best-fit models. (a, b) plot CO_2 footprint for training the best-fit model per epoch, while (c, d) plot the total cost. Note as reported in SI section 1, optimal number of epochs is different for different model (100 in most cases). For RF and XGBR, number of estimators were considered equivalent of epochs in deep learning models. (e, f) plot the average prediction time per frame (estimated from the test data), which is significantly faster than the frequency at which OMC data is captured (100 Hz). The best-fit convLSTM and RF models are $\sim$30 and $\sim$63 times faster than the sampling frequency of OMC data, and is suitable for real-time predictions. (g, h) plot number of parameters in various models. For RF and XGBR, the number of parameters is quantified as the total number of nodes in all decision trees. Note that a direct comparison of the number of parameters in RF and XGBR with NN models might not be appropriate due to the fundamentally different nature of these models. This plot reflects that certain architectures require more computational resources than others, despite having fewer parameters.