Appendix

Figure A.1. Sleep stage distribution of all patients based on manual annotation.
(A) "All data" includes the manual sleep staging from all human experts. "Dataset A" comprises only 30-s epochs where all raters were in agreement, representing 67.84% of all epochs. (B) "Dataset B" additionally includes epochs with mixed labels, most commonly involving W and AS. A total of 1.32% of epochs (black segment) were excluded due to complete disagreement among all raters. W = Wake, QS = Quiet Sleep, AS = Active Sleep.
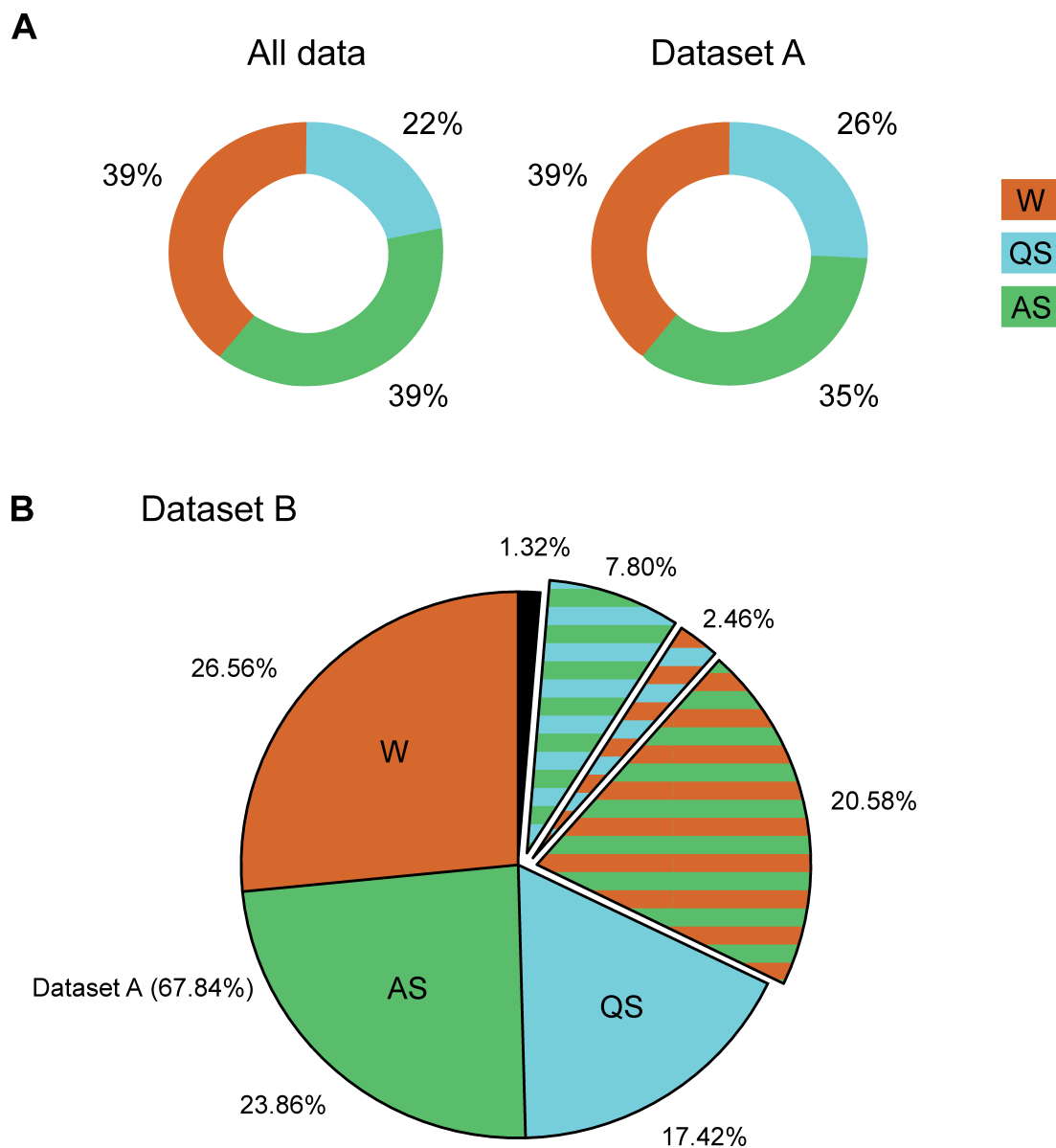
Figure A.2. Examples illustrating the removal of artifact-prone periods and falsely detected R-peaks (for details see section 2.4.2). (A) Automatically identified false-positive R-peaks (red circles) were removed to refine heartbeat detections for subsequent feature extraction. (B) First row: Filtered ECG with noisy segments. Second row: spectrogram of the ECG. Third row: normalized bandpower (BP) in the frequency range 1–30 Hz, baseline is defined as 25th percentile of BP, threshold is calculated using Equation 1, with $C = 15$. Fourth row: RR intervals and median RR interval. Periods exceeding the threshold were extended by 1 second on either side. During these noisy periods, no features, like the median RR interval, were calculated. As a result of this artifact removal process, 2.12% of all 30-second epochs were discarded.
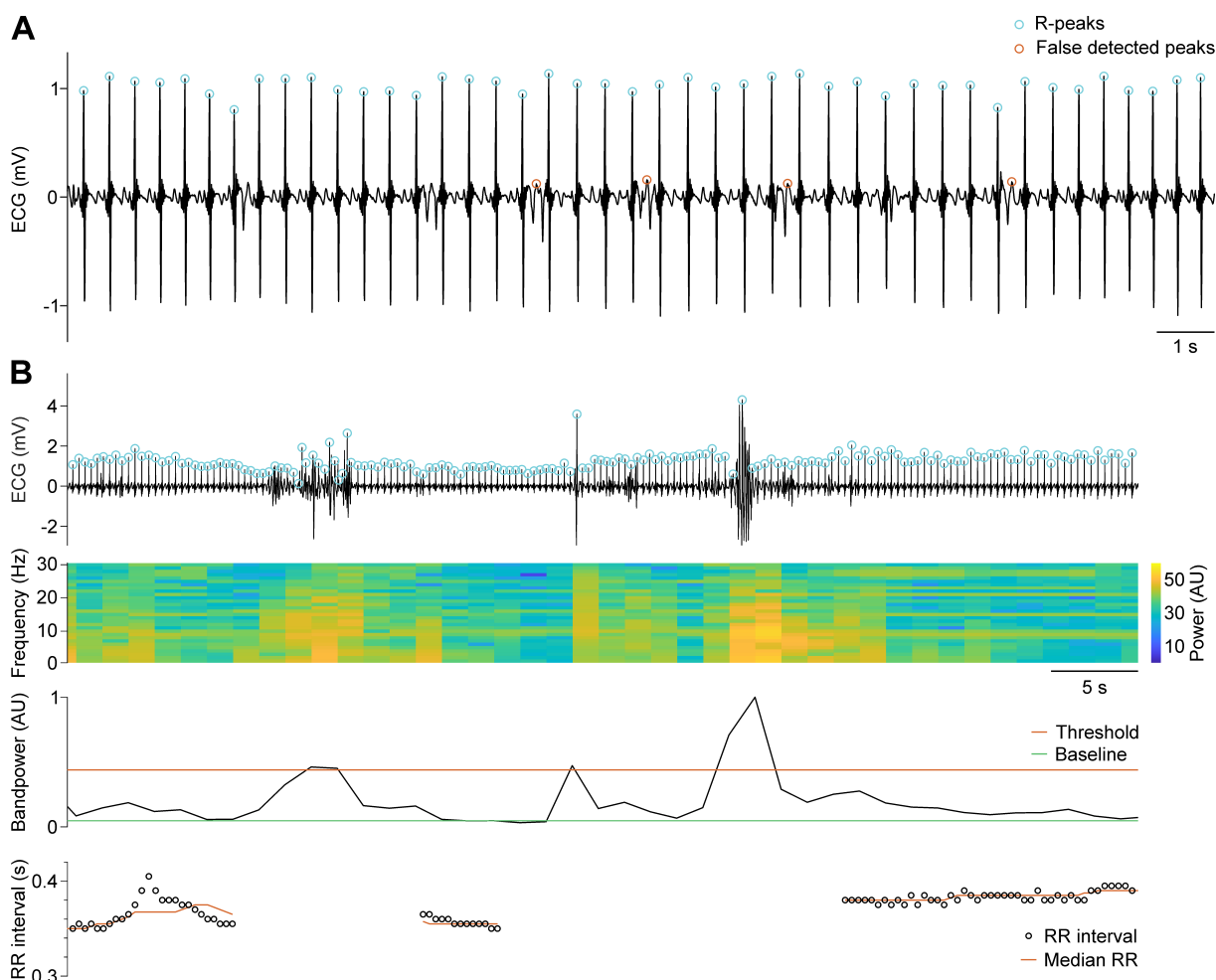
Figure A.3. Feature importance analysis across models and sleep stages (for details see section 2.4.4). Average absolute SHAP (Shapley Additive Explanations) values over all 30-s epochs are shown for all used features, separated by sleep stages (W = Wake, QS = Quiet Sleep, AS = Active Sleep) across the three models: Piezo+ECG Model, Piezo Model, and ECG Model. The importance of each feature is ranked based on its contribution to the model's predictions. Please note the different x-axis in C for visualization purposes. The feature definitions can be found in Table 3.
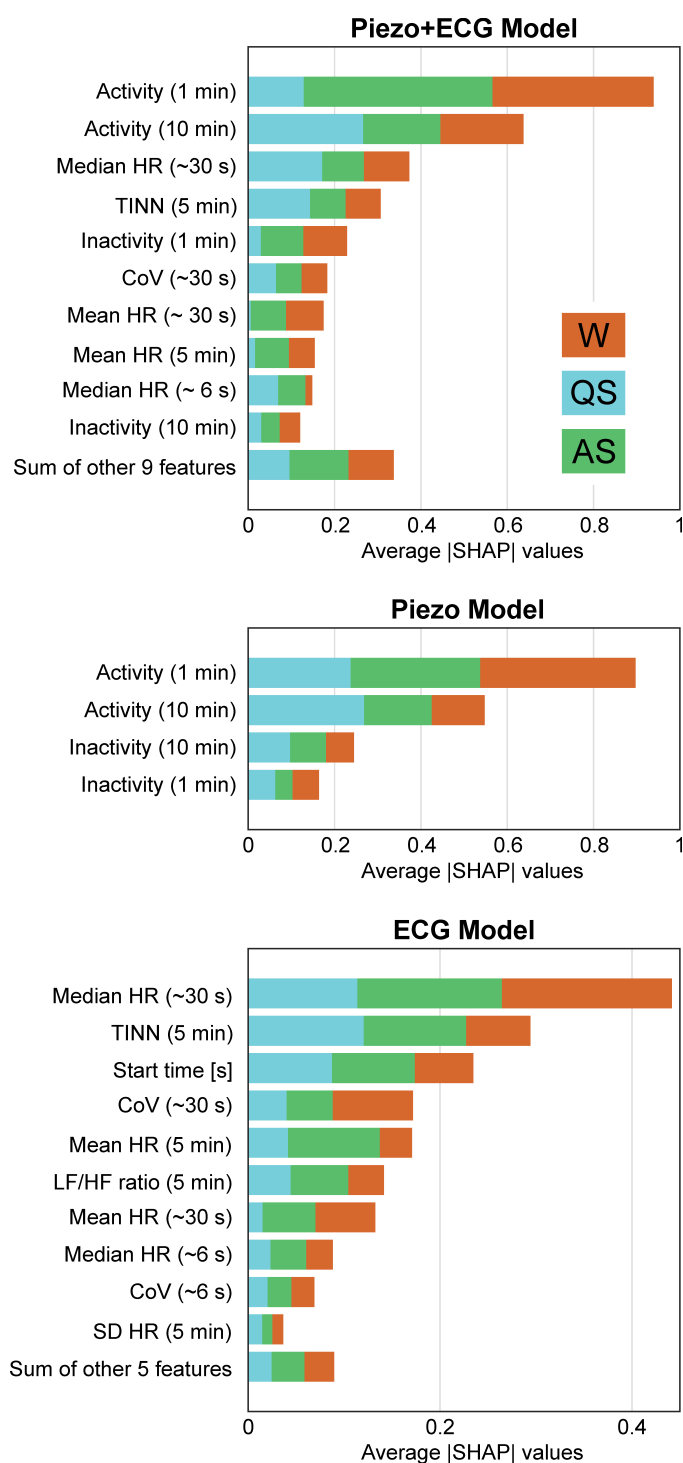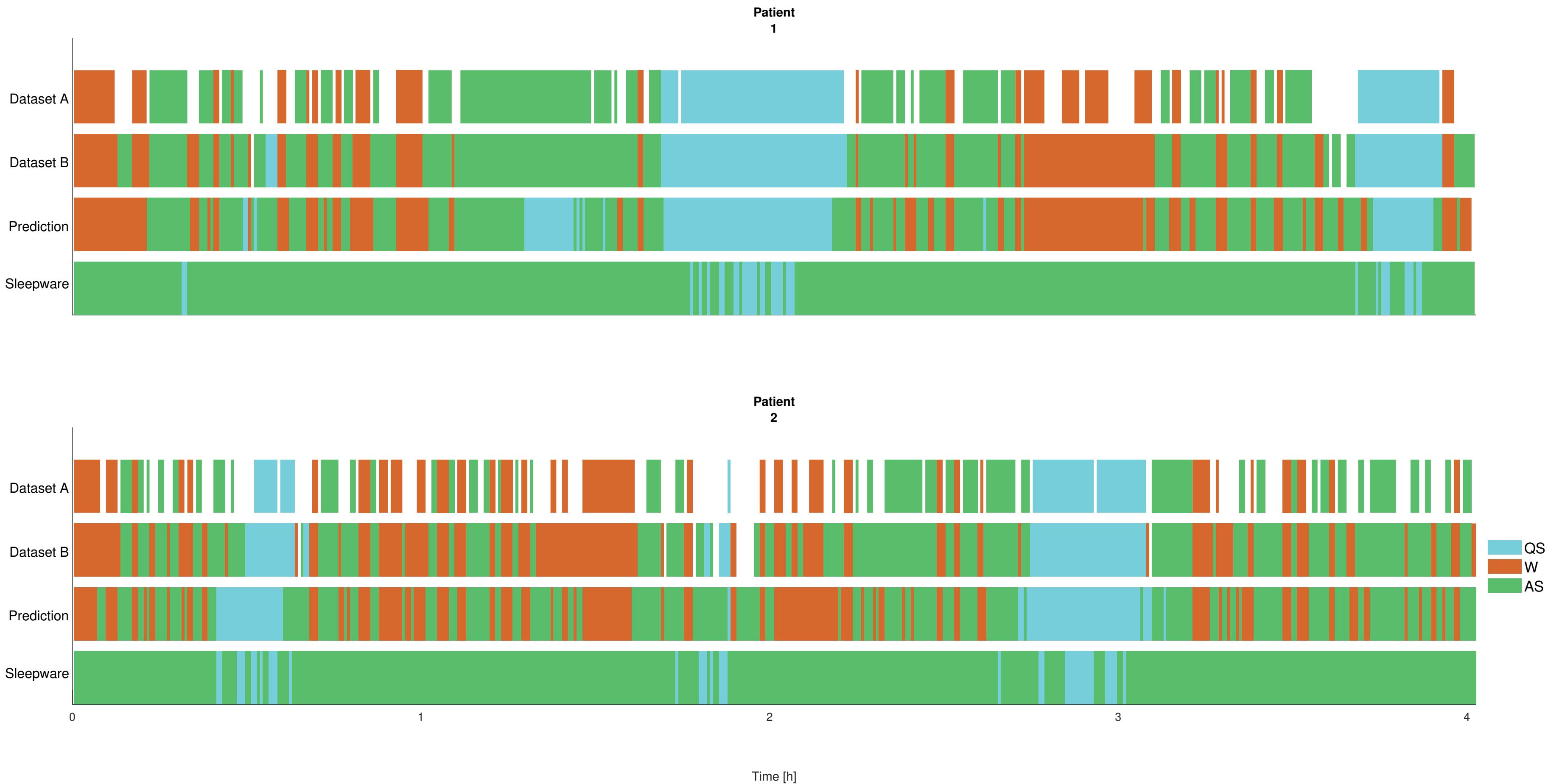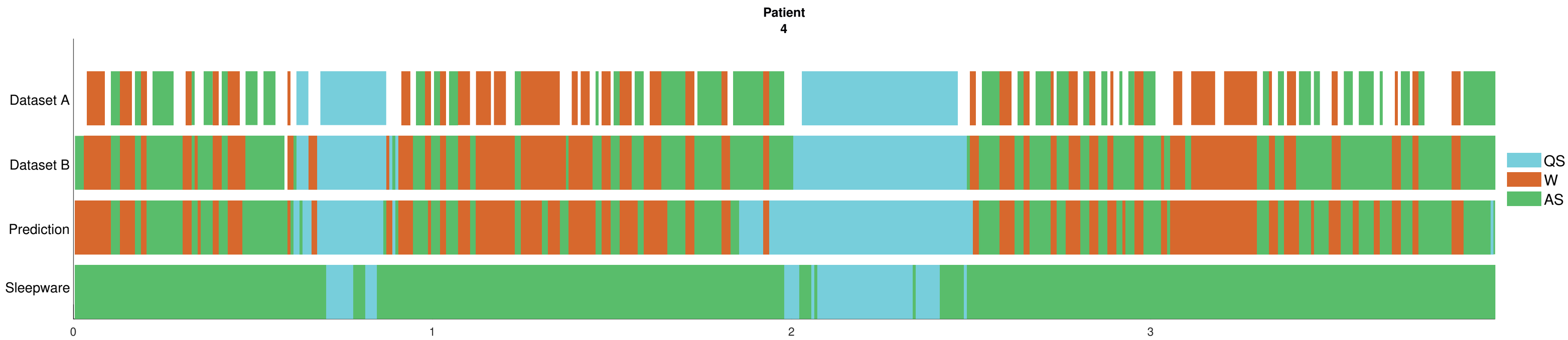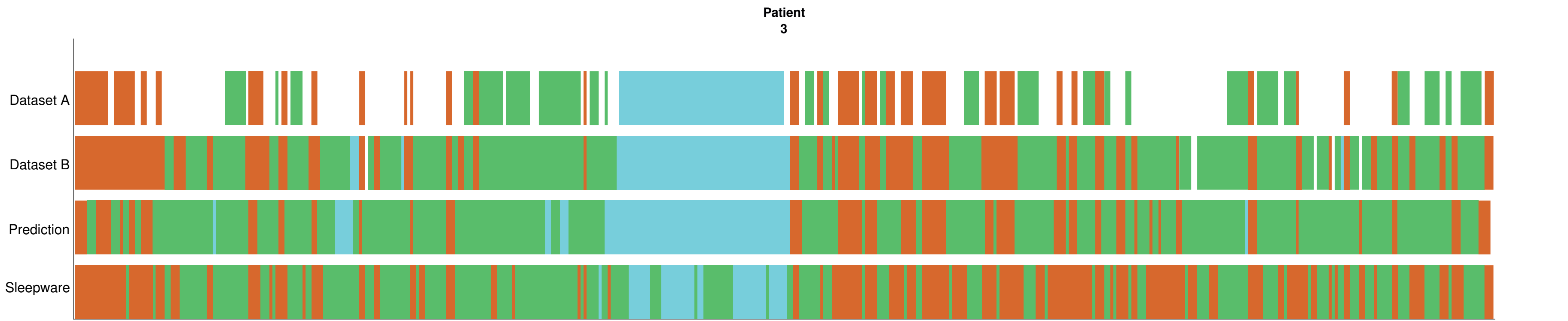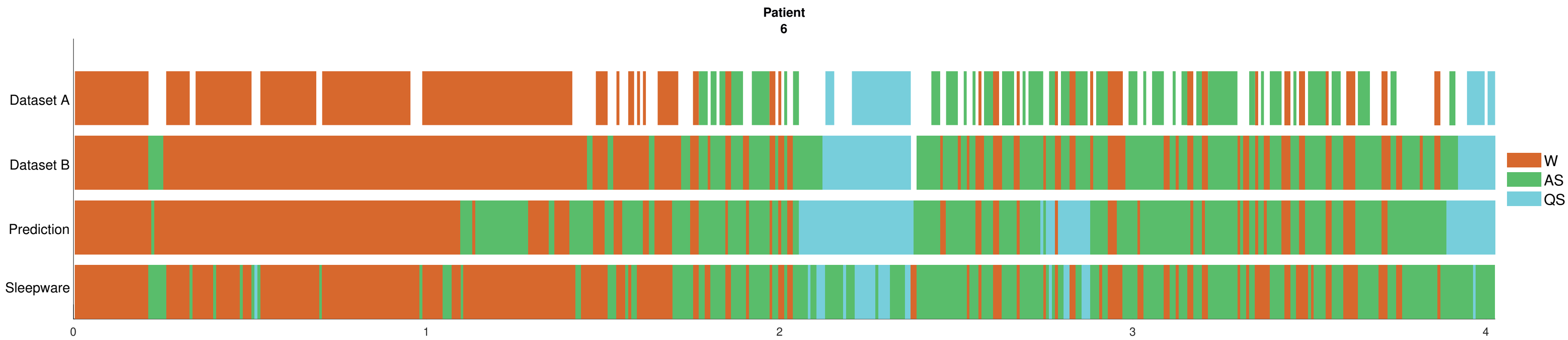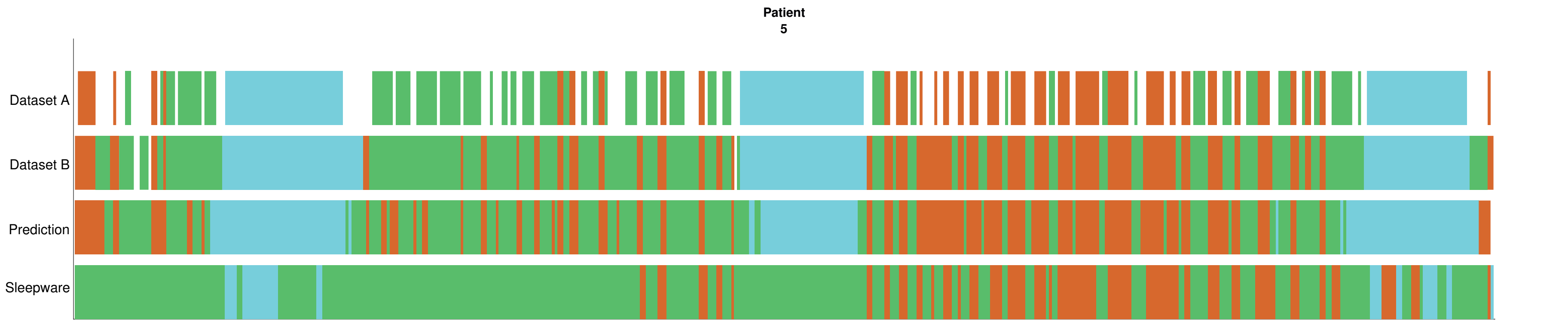
Table A.1. All performance measures for all models trained on Dataset A and B and for Sleepware generated sleep stages.

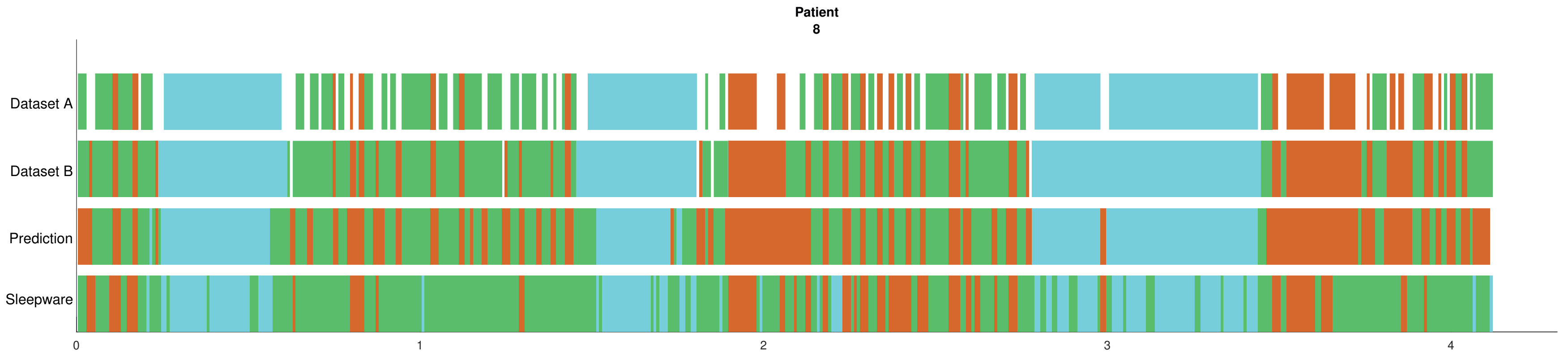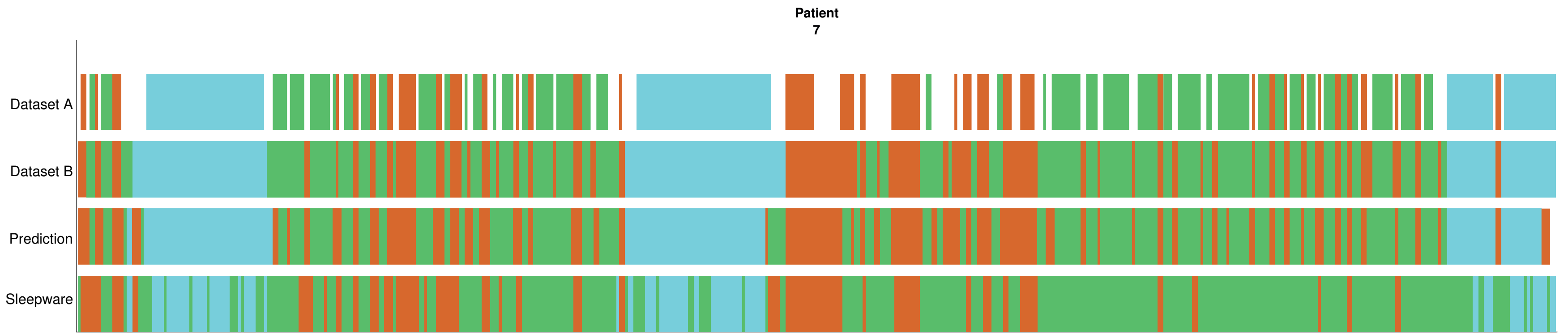| Dataset A	Piezo+ECG Model				Piezo Model				ECG Model				Sleepware			
	mean±SD	W	AS	QS	mean±SD	W	AS	QS	mean±SD	W	AS	QS	mean±SD	W	AS	QS
Accuracy	0.92 (±0.01)	0.93	0.89	0.95	0.89 (±0.004)	0.91	0.84	0.91	0.88 (±0.002)	0.86	0.84	0.93	0.68 (±0.16)	0.67	0.53	0.86
Kappa	0.83 (±0.002)	0.84	0.77	0.87	0.75 (±0.01)	0.81	0.66	0.76	0.73 (±0.01)	0.71	0.66	0.82	0.31(±0.21)	0.19	0.19	0.55
Sensitivity	0.88 (±0.01)	0.9	0.87	0.88	0.83 (±0.01)	0.86	0.8	0.83	0.82 (±0.004)	0.82	0.8	0.86	0.53 (±0.38)	0.19	0.95	0.46
Specificity	0.94 (±0.01)	0.95	0.9	0.97	0.92 (±0.003)	0.94	0.87	0.94	0.91 (±0.002)	0.89	0.87	0.96	0.76 (±0.4)	0.97	0.30	0.99
Precision	0.89 (±0.01)	0.91	0.83	0.92	0.8 (±0.01)	0.89	0.73	0.78	0.82 (±0.003)	0.82	0.78	0.88	0.73 (±0.28)	0.82	0.42	0.96
Dataset B	Piezo+ECG Model				Piezo Model				ECG Model				Sleepware			
	mean	W	AS	QS	mean	W	AS	QS	mean	W	AS	QS	mean±SD	W	AS	QS
Accuracy	0.87	0.87	0.83	0.93	0.82	0.84	0.77	0.87	0.83	0.80	0.78	0.90	0.68 (±0.17)	0.66	0.53	0.86
Kappa	0.71	0.72	0.65	0.78	0.59	0.65	0.52	0.63	0.61	0.59	0.54	0.69	0.27(±0.19)	0.16	0.16	0.48
Sensitivity	0.81	0.80	0.83	0.78	0.73	0.75	0.74	0.71	0.74	0.74	0.74	0.72	0.50 (±0.39)	0.17	0.93	0.40
Specificity	0.90	0.91	0.82	0.97	0.86	0.89	0.78	0.92	0.86	0.84	0.80	0.95	0.74 (±0.41)	0.96	0.26	0.99
Precision	0.82	0.87	0.75	0.87	0.73	0.81	0.69	0.71	0.75	0.75	0.71	0.79	0.70 (±0.23)	0.75	0.45	0.90
Accuracy Ratio Dataset A/B	0.95	0.93	0.93	0.98	0.93	0.92	0.91	0.96	0.94	0.93	0.93	0.96				

Patient 1

| | Dataset A | Dataset B | Prediction | Sleepware |

Patient 2

| | Dataset A | Dataset B | Prediction | Sleepware |

Legend: QS, W, AS

Time [h]

Patient 3

Dataset A	
Dataset B	
Prediction	
Sleepware	

Patient 4

Dataset A	
Dataset B	
Prediction	
Sleepware	

Legend: QS, W, AS

Time [h]: 0, 1, 2, 3

Patient 7

Dataset A
Dataset B
Prediction
Sleepware

Patient 8

Dataset A
Dataset B
Prediction
Sleepware

W
AS
QS

Time [h]

0   1   2   3   4

Patient 12

	Dataset A
	Dataset B
	Prediction
	Sleepware

Patient 14

	Dataset A
	Dataset B
	Prediction
	Sleepware

Legend: W, AS, QS

Time [h]
0   1   2   3

Patient 21

Dataset A	
Dataset B	
Prediction	
Sleepware	

Patient 22

Dataset A	
Dataset B	
Prediction	
Sleepware	

QS
W
AS

Time [h]

Patient 27

Dataset A
Dataset B
Prediction
Sleepware

Patient 28

Dataset A
Dataset B
Prediction
Sleepware

0   1   2   3   4

Time [h]

QS
W
AS