

Time-frequency deep learning model for sleep classification of dairy cows using polysomnography and RumiWatch data.

Chuanyi Guo^{1,2}, Emma Ternman³, Xinjie Zhao¹, Kai Liu², Mutian Niu¹

¹*Animal Nutrition Group, ETH Zurich, Zürich, Switzerland.*

²*Department of Infectious Diseases and Public Health, City University of Hong Kong, Hong Kong SAR, China.*

³*Faculty of Biosciences and Aquaculture, Nord University, Steinkjer, Norway.*

What is sleep?

- Sleep is a recurring, reversible neuro-behavioral state of relative perceptual **disengagement from** and **unresponsiveness to** the environment (Mary and William, 1989).

For humans:

- Neural restoration
- Immune function
- Memory consolidation
- Maintaining overall health
- ...

For cows:



Sleep is essential for mammals, yet remains poorly understood in cows.

Current gaps in cow sleep

Ground truth measurement:

- Visual Observation
 - Labor-intensive
 - Overestimation
- Electrophysiological Recording
 - Time-intensive
 - Annotator bias

Usability for on-farm applications

- PSG is impractical
- Alternative solution is needed

Comparison of Sleep Scoring Methods

(Ternman et al, 2014)

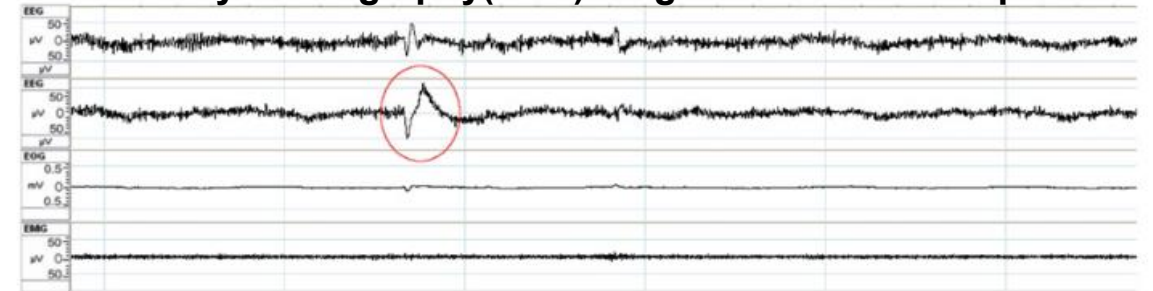
Sleep stages	Visual	Electrophysiological
NREM*	124 ± 17	20 ± 5
REM**	14 ± 4	10 ± 3

* NREM - Non-rapid eye movement

** REM – Rapid eye movement

*** Duration unit: Min

Polysomnography(PSG) Diagram of NREM Sleep



Objective

- 1. Propose an automated sleep classification method**
- 2. Evaluate accessible sensors for sleep detection**

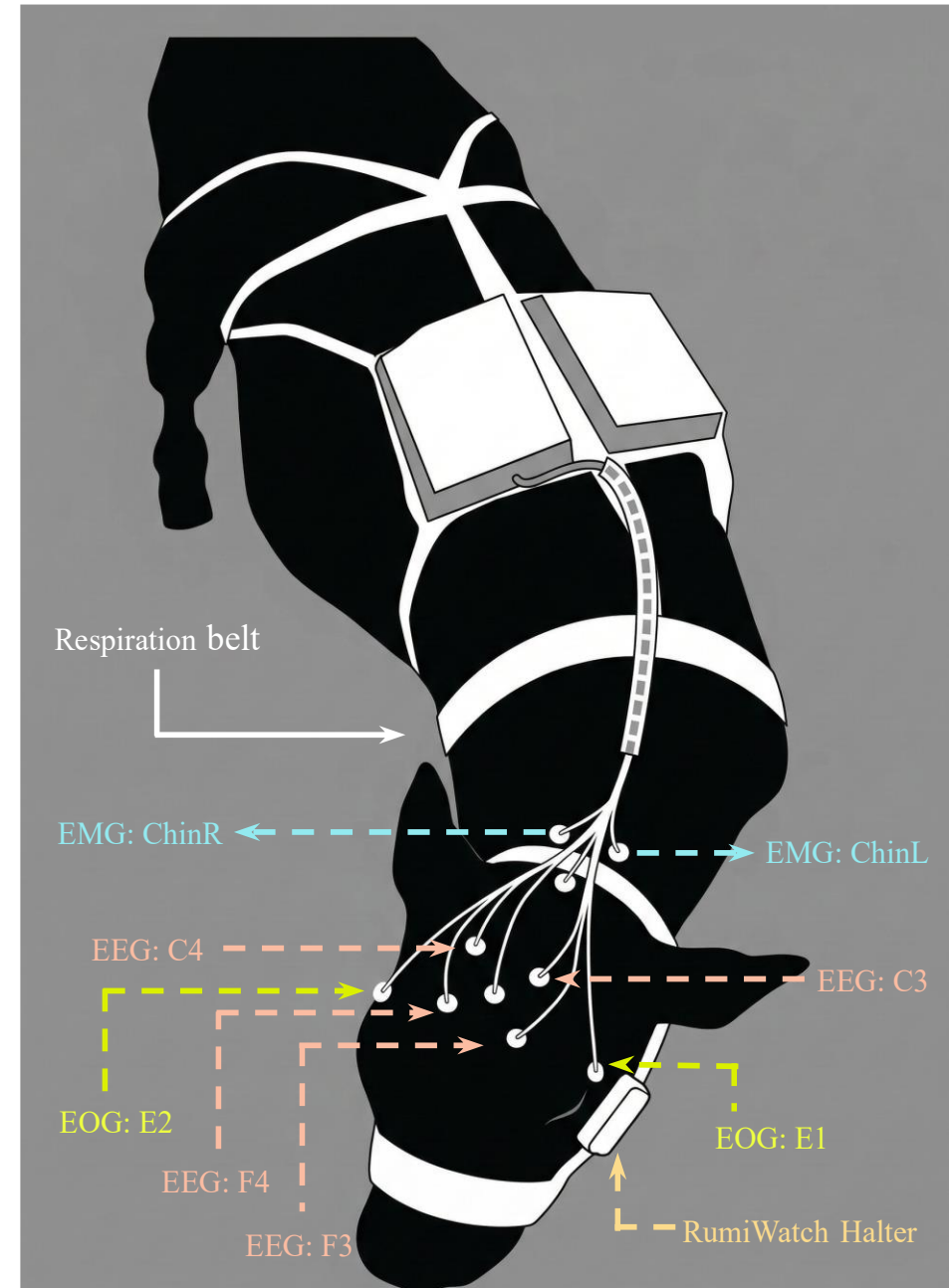
Data collection

9-Channel of Polysomnography (PSG)

- 4-channel – **Brain Activity (EEG)**
- 2-channel – **Muscle Activity (EMG)**
- 2-channel – **Eye Movement (EOG)**
- 1-channel – **Respiration (RS)**

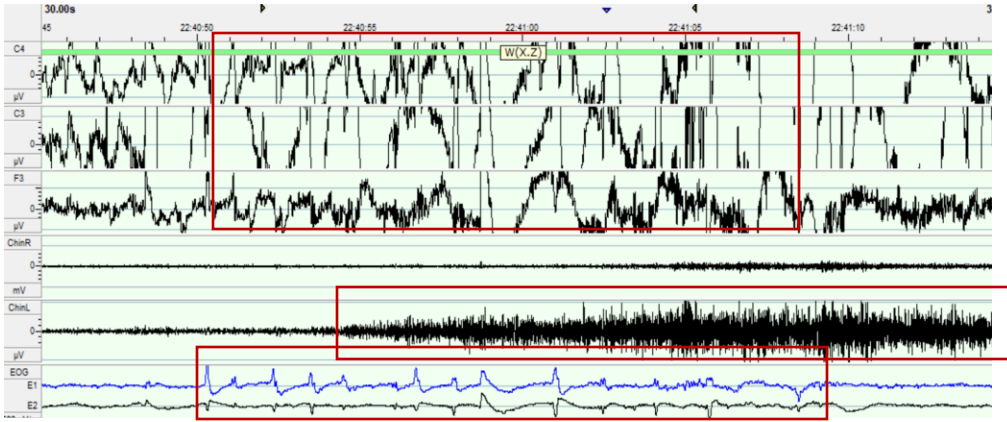
4-Channel of RumiWatch System (RWS)

- 1-channel – **Pressure (Jaw Movement)**
- 3-channel – **Accelerometers (X, Y, Z)**

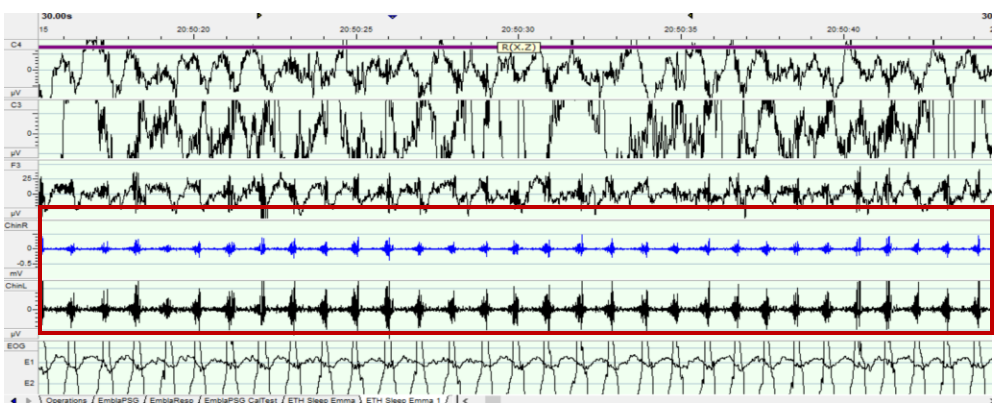


Data annotation

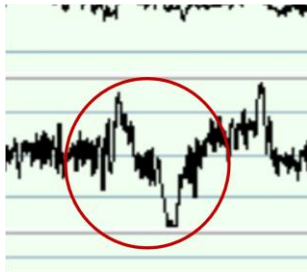
Wake 41.62%



Rumination 45.59%



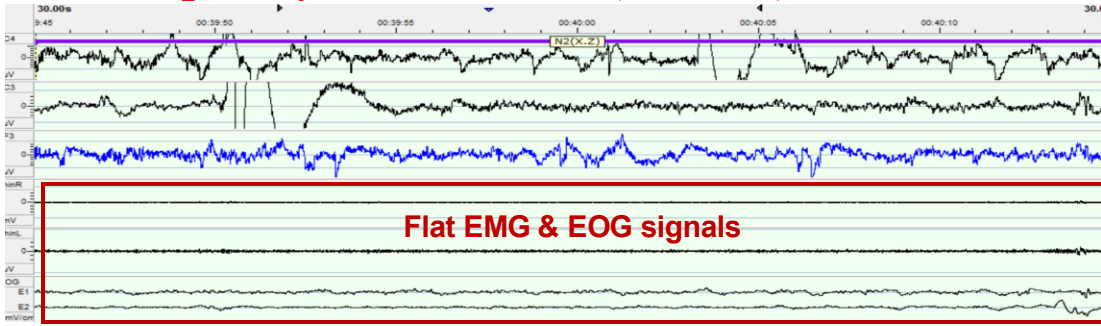
Drowsing



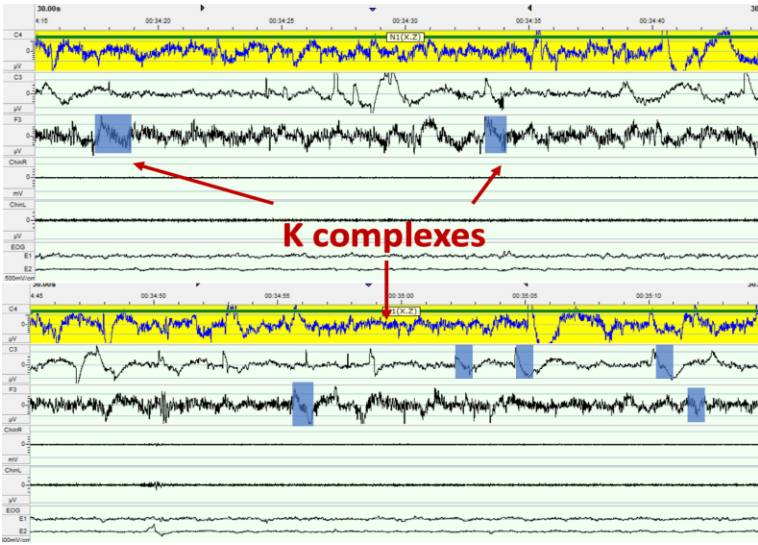
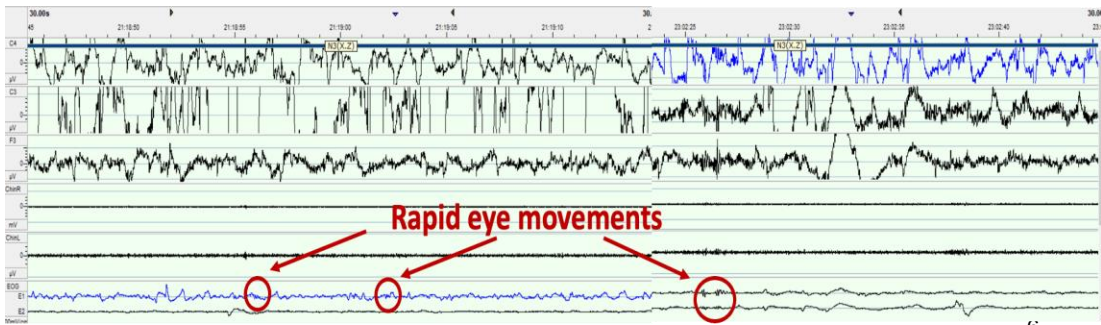
Sleep

12.79%

Non-Rapid Eye Movement (NREM)



Rapid Eye Movement (REM)



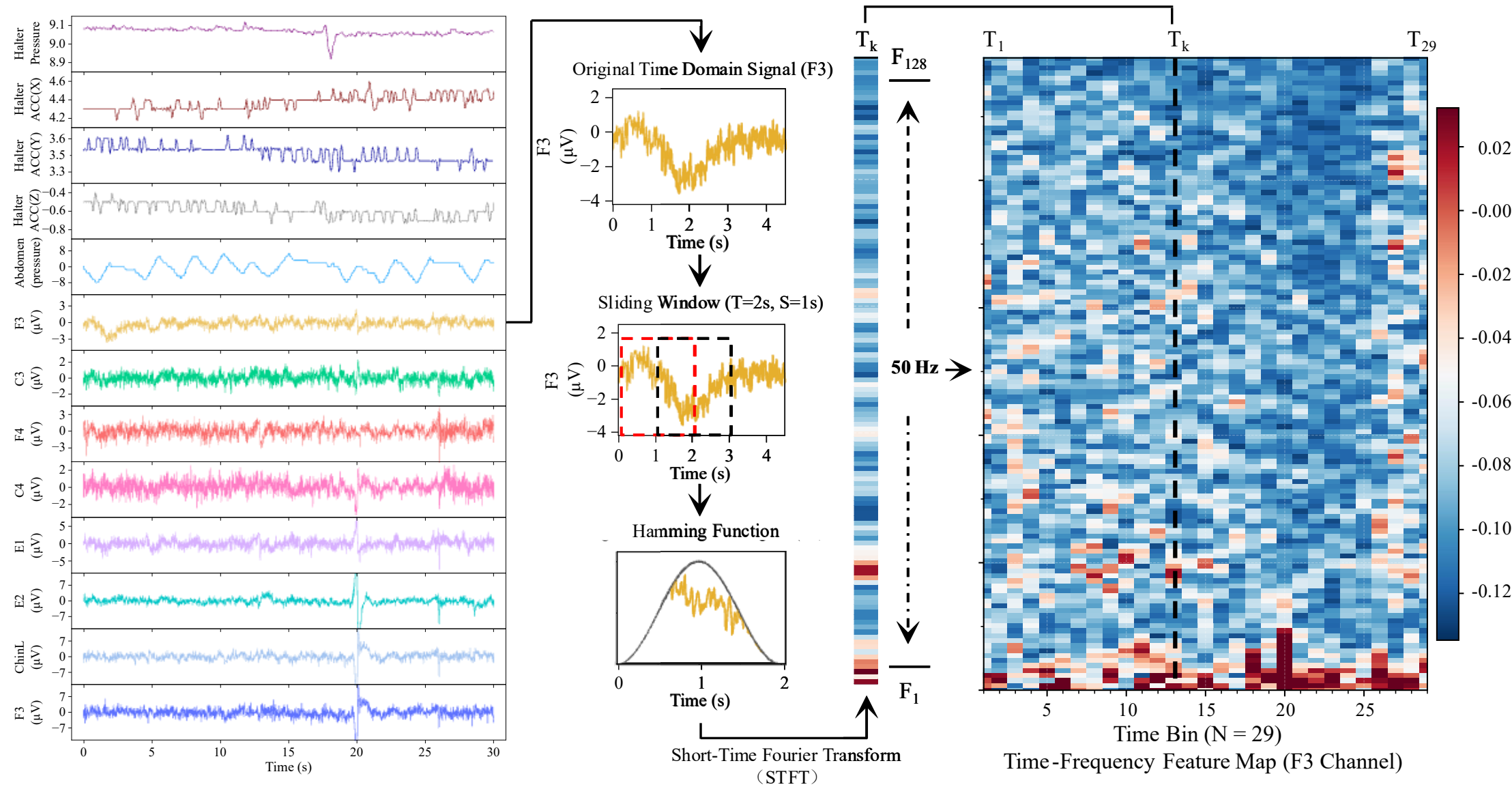
Data annotation

Cow ID	Behaviour Classes			Total Epochs	Duration (hour)
	Wake	Sleep	Rumination		
Cow 1	766	158	234	1,158	9.65
Cow 2	473	139	780	1,392	11.60
Cow 3	346	150	672	1,168	9.73
Cow 4	853	222	315	1,390	11.58
Cow 5	756	201	451	1,408	11.73
Cow 6	449	142	750	1,341	11.18
Cow 7	577	172	650	1,399	11.66

➤ 7 Cows, 77 hours data

➤ 1 epoch = a 30s data segment

Data preprocessing – Time-frequency-based representations



(a) Overall Architecture of DairySleepNet

Time-Frequency Input (C, T, F)
 C = # of channels | T = 29 time frames | F = 128 frequency bins (52 hz)

Single-Channel Feature Extraction

Single Channel Encoder (Mamba Block) N_s

Concat
 $(T, C \times F)$

RMSNorm + Dropout

Multi-Channel Feature Extraction

Multiple Channel Encoder (Mamba Block) N_m

$\oplus$ $(T, C \times F)$

RMSNorm
 $(T, C \times F)$

Flatten
 $(T \times C \times F)$

Classification Head

Linear (1024) | GELU | Dropout | Linear (3)

Wake / Sleep / Rumination

```

graph TD
    H[H] --> RMSNorm[RMSNorm]
    RMSNorm --> Linear1[Linear]
    Linear1 --> Z[Z]
    Linear1 --> X[X]
    Z --> SiLU1[SiLU]
    X --> Conv1D[Conv1D]
    Conv1D --> SiLU2[SiLU]
    SiLU2 --> Linear2[Linear]
    Linear2 --> Delta[Δ]
    Linear2 --> B[B Matrix]
    Linear2 --> C[C Matrix]
    Delta --> Linear3[Linear]
    Linear3 --> Softplus[Softplus]
    B --> Discretize[Discretize ZOH]
    C --> Discretize
    Discretize --> SelectiveScan[Selective Scan]
    SiLU1 --> Add1((+))
    Softplus --> Add1
    SelectiveScan --> Add1
    Add1 --> Linear4[Linear + Dropout]
    Linear4 --> Add2((+))
    H --> Add2
    Add2 --> Output[Output H_out]

```

The diagram illustrates the proposed architecture for the multi-view multi-label classification task. It shows the flow of information from input features x_t and hidden states h_{t-1} through various processing blocks to produce the output y_t and the next hidden state h_t .

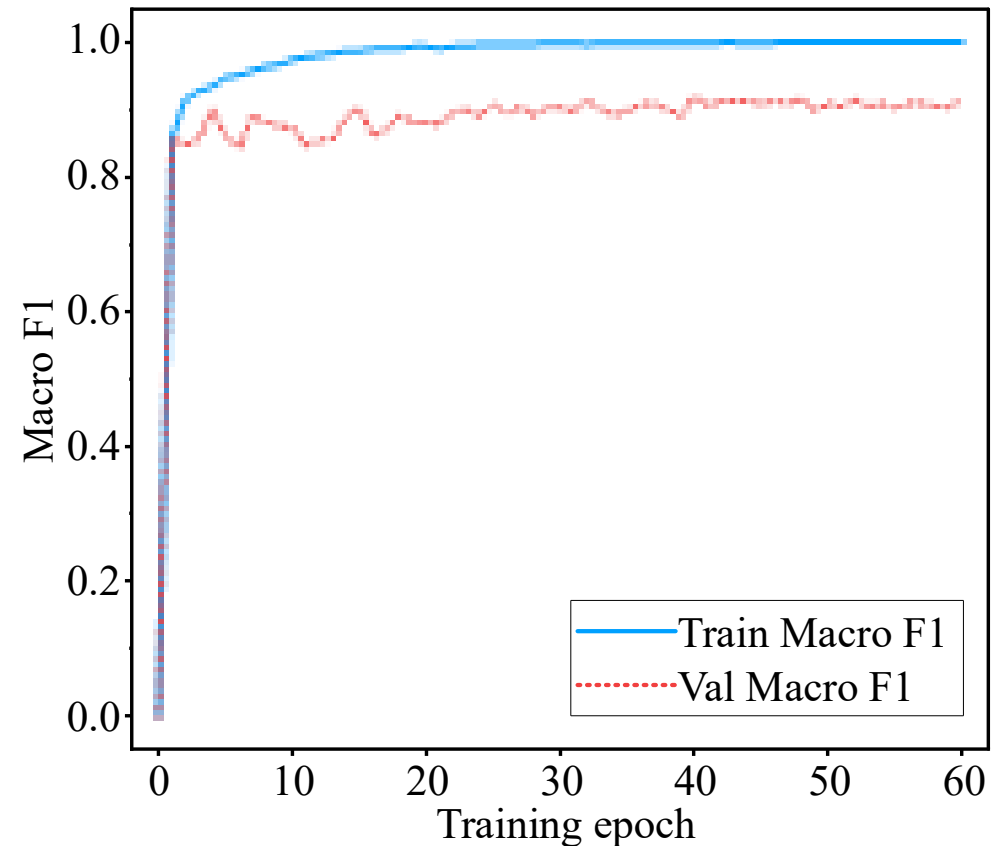
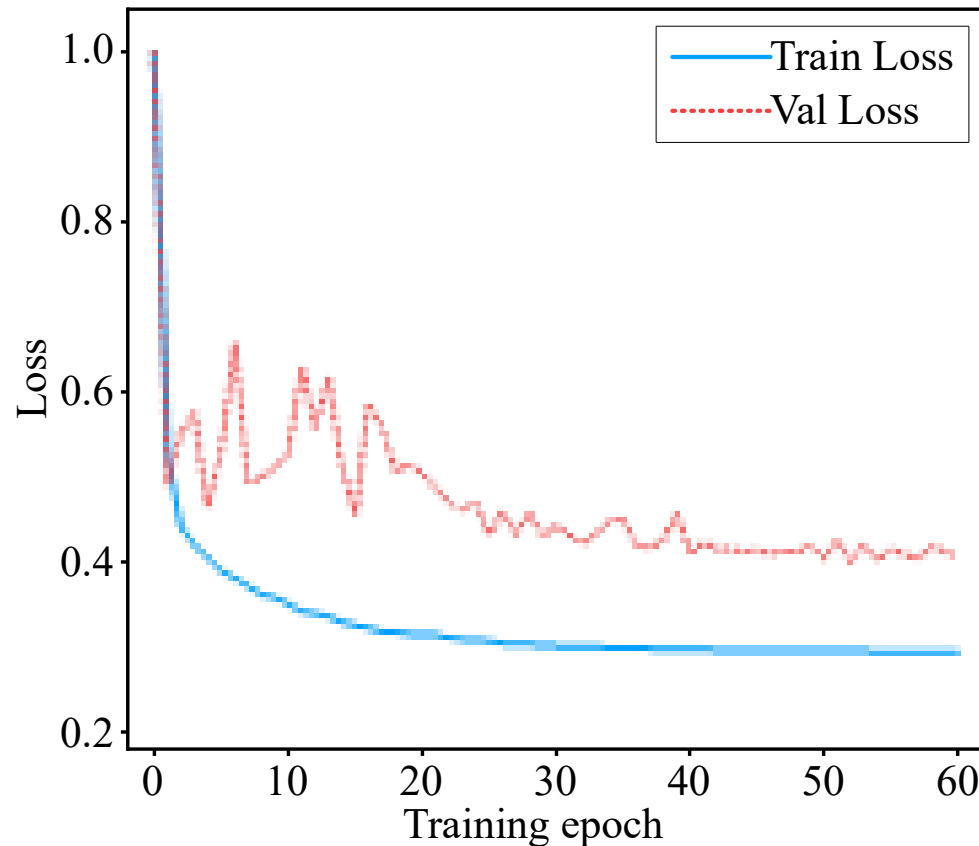
- Input x_t :** A green vector of size 4 is input to block B_t .
- Block B_t :** A gray block that receives input from x_t and the **Project** block. It outputs four parallel features to block A .
- Block A :** A gray block that receives inputs from B_t and the **Discretize** block. It outputs four parallel features to block C_t and the next hidden state h_t .
- Block C_t :** A gray block that receives inputs from A and the **Project** block. It outputs a green vector of size 4, which is the final output y_t .
- Block Δ_t :** A gray block that receives input from the **Project** block. It is connected to block A via a dashed arrow labeled **Discretize**.
- Project Block:** A purple block that receives input from x_t and outputs to B_t , C_t , and Δ_t .
- Hidden States:** The previous hidden state h_{t-1} (blue box) is input to block A via four parallel dashed lines. The current hidden state h_t (blue box) is output from block A via four parallel dashed lines.

$$\Delta = \text{softplus}(\text{Linear}_{\Delta}(x_t))$$

Model training

Animal-level 4-fold cross-validation, 10% of training set was held out as validation set

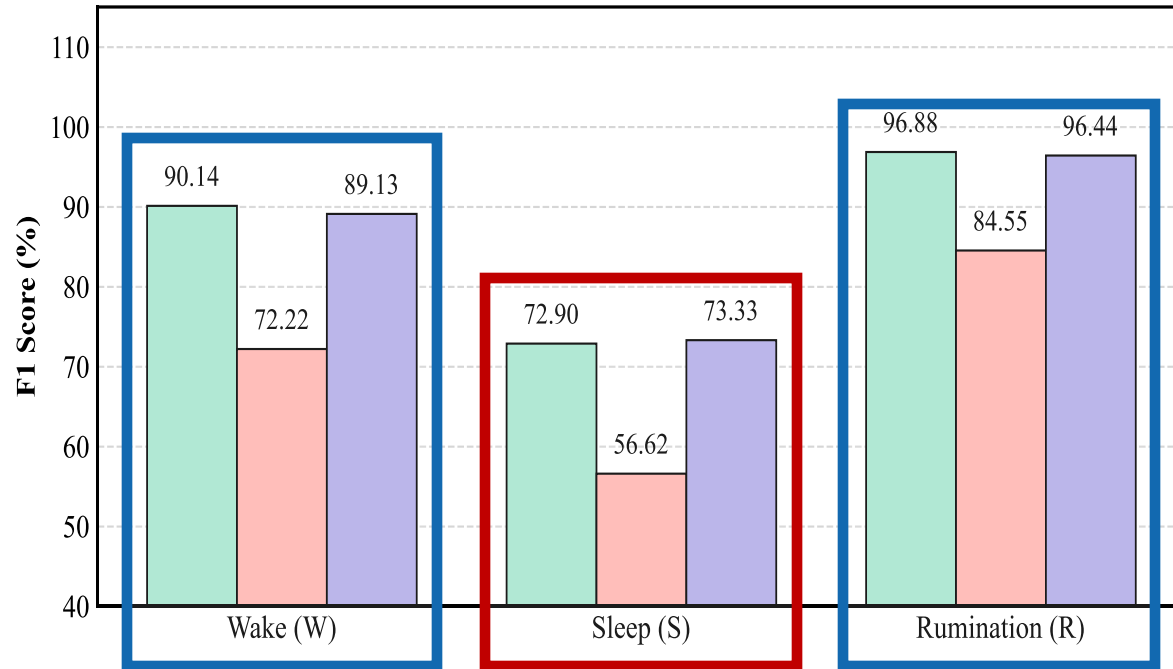
Training epochs was set to 200 with early stopping patience 20 epochs.



Model performance with different data inputs

Class-wise F1 Scores

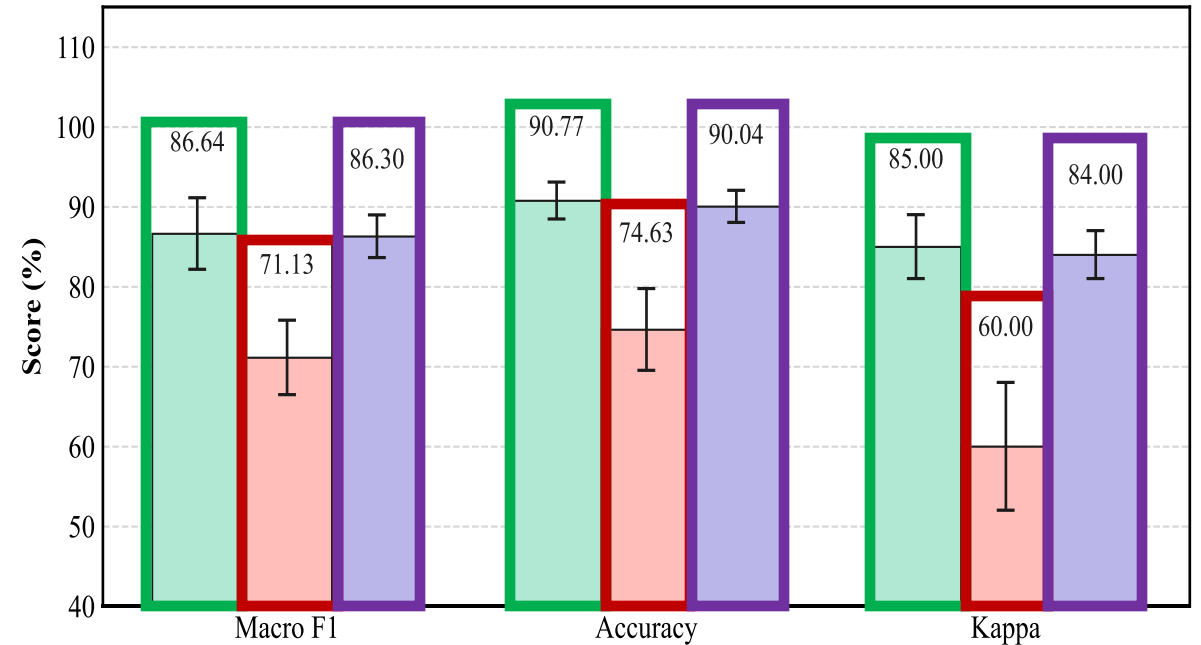
PSG RWS PSG+RWS



- Accurate classification for Wake / Rumination
- Challenge remains for Sleep

Overall Metrics Comparison

PSG RWS PSG+RWS



- PSG is the most informative data source
- RWS is insufficient for sleep classification
- PSG + RWS has less variances among folds

Model performance with different data inputs

PSG Modality

True Stage	W	S	R
W	3892	267	61
S	350	834	0
R	173	3	3676
		Predicted Stage	

RWS Modality

True Stage	W	S	R
W	2863	929	428
S	259	870	55
R	587	90	3175
		Predicted Stage	

PSG+RWS Modality

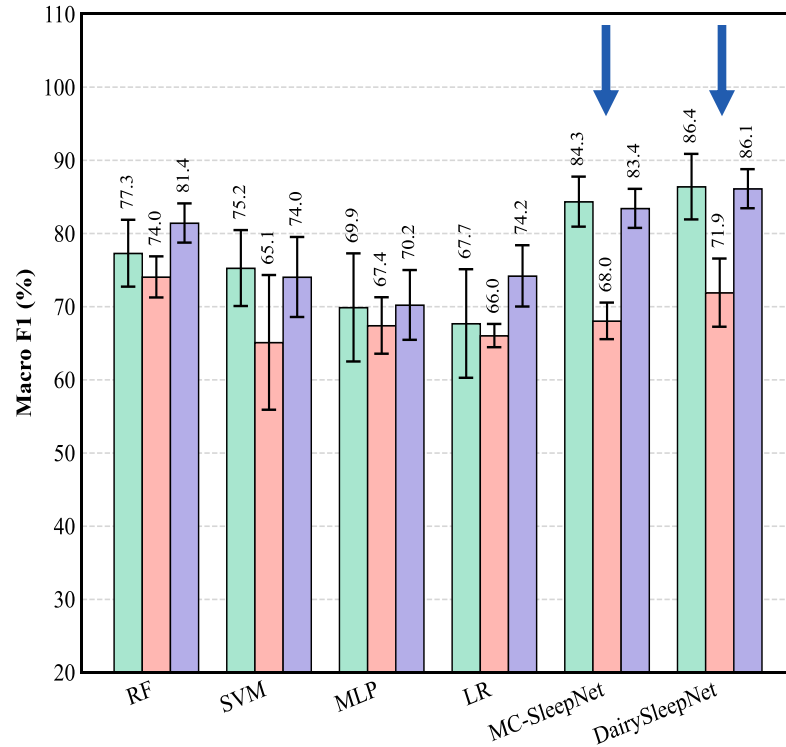
True Stage	W	S	R
W	3773	367	80
S	285	899	0
R	188	2	3662
		Predicted Stage	

- Most of confusion arises between Wake & Sleep
- RWS has more mistakes classifying Wake & Rumination

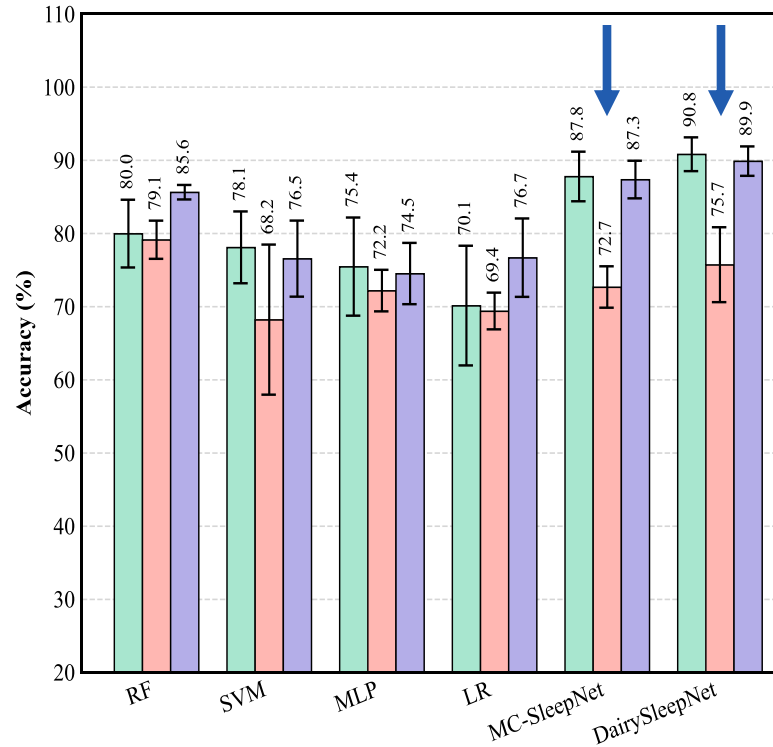
Performance comparison with baselines

PSG RWS PSG+RWS

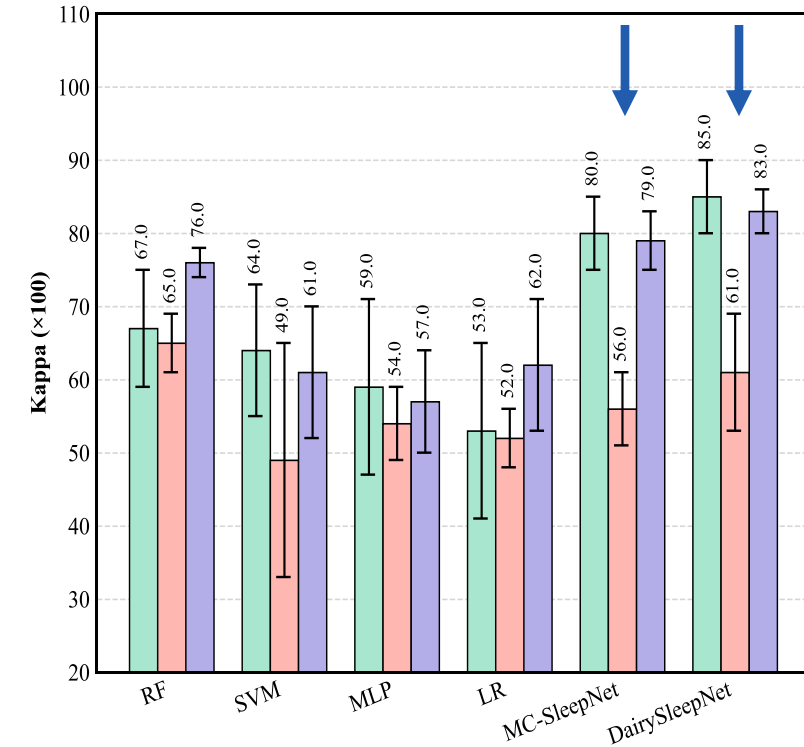
Macro F1 Comparison



Accuracy Comparison

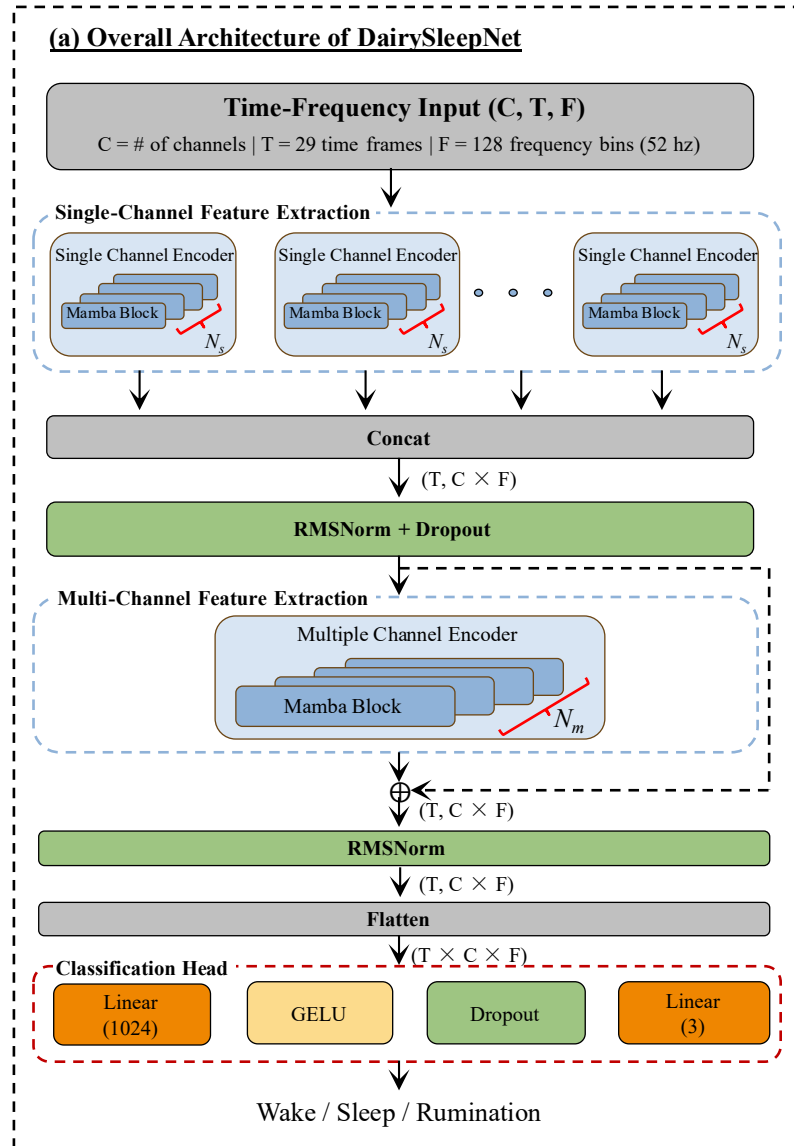


Kappa Comparison



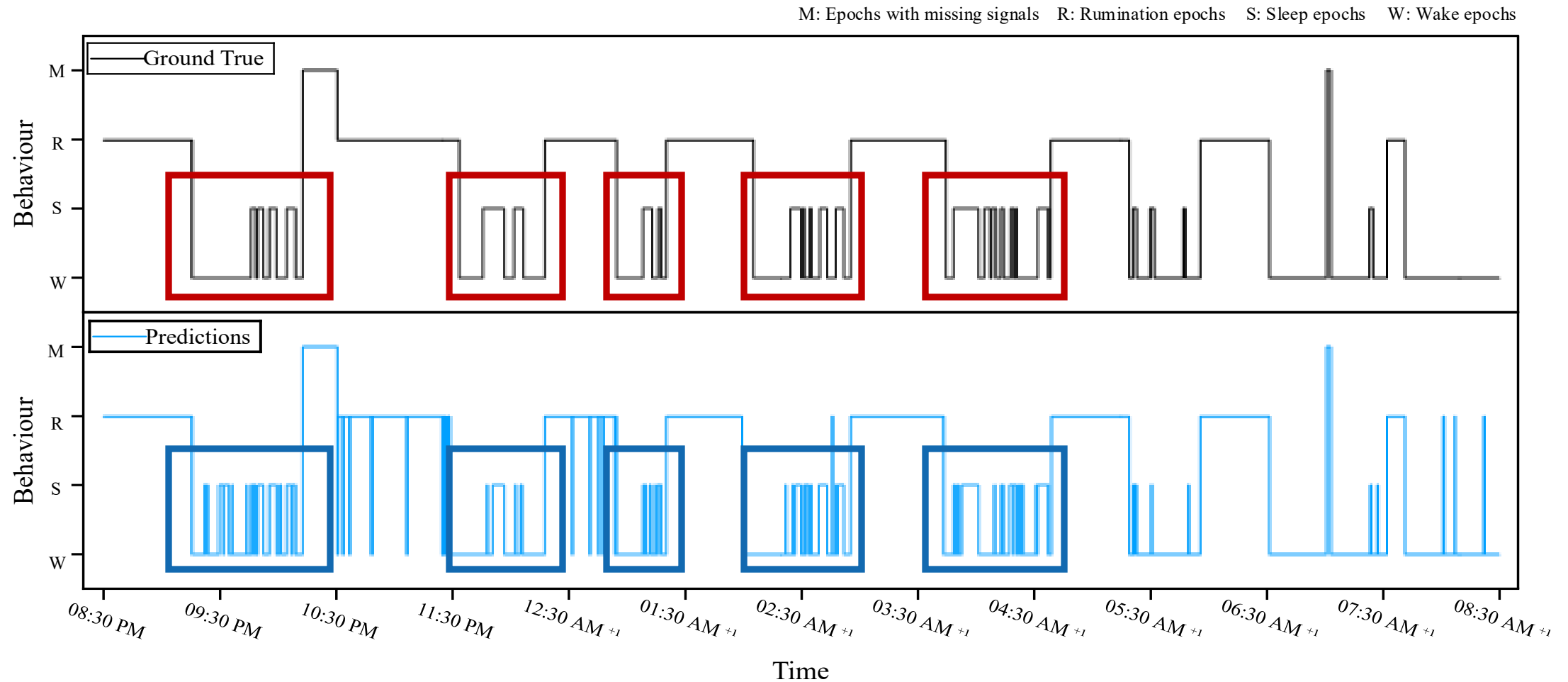
- ML method (Mean, Standard deviation, Root mean square, Range, Skewness, Kurtosis)
- Time-frequency representations are more informative than time-based feature representations

Sensitivity analysis



N_s	N_m	ACC	Macro F1	κ
4	2	88.64 ± 2.81	84.43 ± 4.12	0.81 ± 0.05
4	4	90.80 ± 2.31	86.37 ± 4.48	0.85 ± 0.05
4	8	89.58 ± 3.06	85.28 ± 4.81	0.83 ± 0.05
8	2	90.34 ± 2.23	86.09 ± 3.96	0.84 ± 0.04
8	4	90.23 ± 2.39	85.35 ± 4.24	0.84 ± 0.04
8	8	89.94 ± 2.72	85.42 ± 4.29	0.83 ± 0.05
16	2	89.65 ± 2.21	85.22 ± 3.53	0.83 ± 0.04
16	4	88.81 ± 3.03	84.68 ± 3.99	0.81 ± 0.05
16	8	90.29 ± 2.69	86.3 ± 4.06	0.84 ± 0.05

Hypnogram



- Overall pattern is well captured
- Mistake happens in transitions between **Wake & Sleep**

Takeaway message

- ***Selective state space model*** shows promising potential in classifying cow's sleep behavior.
- Time-frequency representations are ***more informative*** for sleep classification.
- Cow sleep is highly ***polyphasic***, making sleep-wake transitions much harder to capture than rumination.
- Next-generation farm sensors fusing ***multiple vital signs*** might be a more feasible alternative to PSG in routine farm settings

Thank you!

chuguo@ethz.ch

