SLEEP 2017
BOSTON • JUNE 3-7

Multimodal Ambulatory Sleep Detection Using Recurrent Neural Networks

Chen, W.[1]*, Sano, A.[1]*, Lopez, D.[1,2], Taylor, S.[1], McHIll, A. W.[3,4], Phillips, A. J.[3,4], Barger, L.K.[3,4], Czeisler, C. A.[3,4], Picard, R. W.[1] (*equal contribution)

[1]Media Lab, Massachusetts Institute of Technology, Cambridge, MA,
[2]Harvard-MIT Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA,
[3]Sleep Health Institute and Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, MA,
[4]Division of Sleep Medicine, Harvard Medical School, Boston, MA.

Conflict of Interest Disclosures for Speakers

X 1. I do not have any relationships with any entities **producing**, **marketing, re-selling, or distributing** health care goods or services consumed by, or used on, patients, **OR**

2. I have the following relationships with entities **producing**, **marketing, re-selling, or distributing** health care goods or services consumed by, or used on, patients.

Type of Potential Conflict	Details of Potential Conflict
Grant/Research Support	
Consultant	
Speakers' Bureaus	
Financial support	
Other	

3. The material presented in this lecture has no relationship with any of these potential conflicts, **OR**

4. This talk presents material that is related to one or more of these potential conflicts, and the following objective references are provided as support for this lecture:

1.

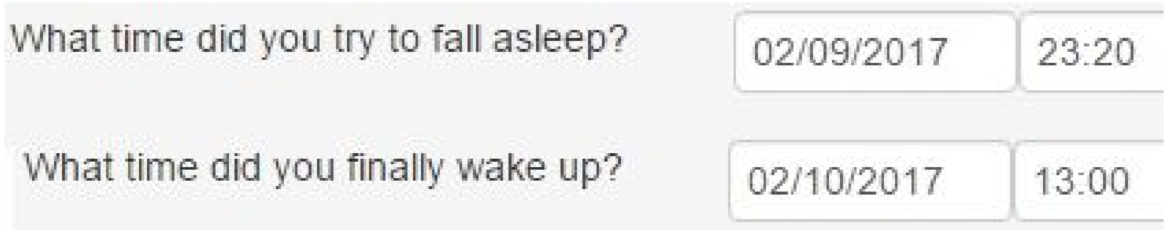
2.

3.

Motivation

Polysomnography (PSG)

Impractical for long-term home use



What time did you try to fall asleep? 02/09/2017 23:20

What time did you finally wake up? 02/10/2017 13:00

Actigraphy + Sleep Diary

Requires significant effort of users to maintain accurate diaries, and of researchers to check the diary entries for anomalies

There is a need for tools to enable accurate long-term evaluation of sleep timing and duration in daily life with less burden on users and researchers.

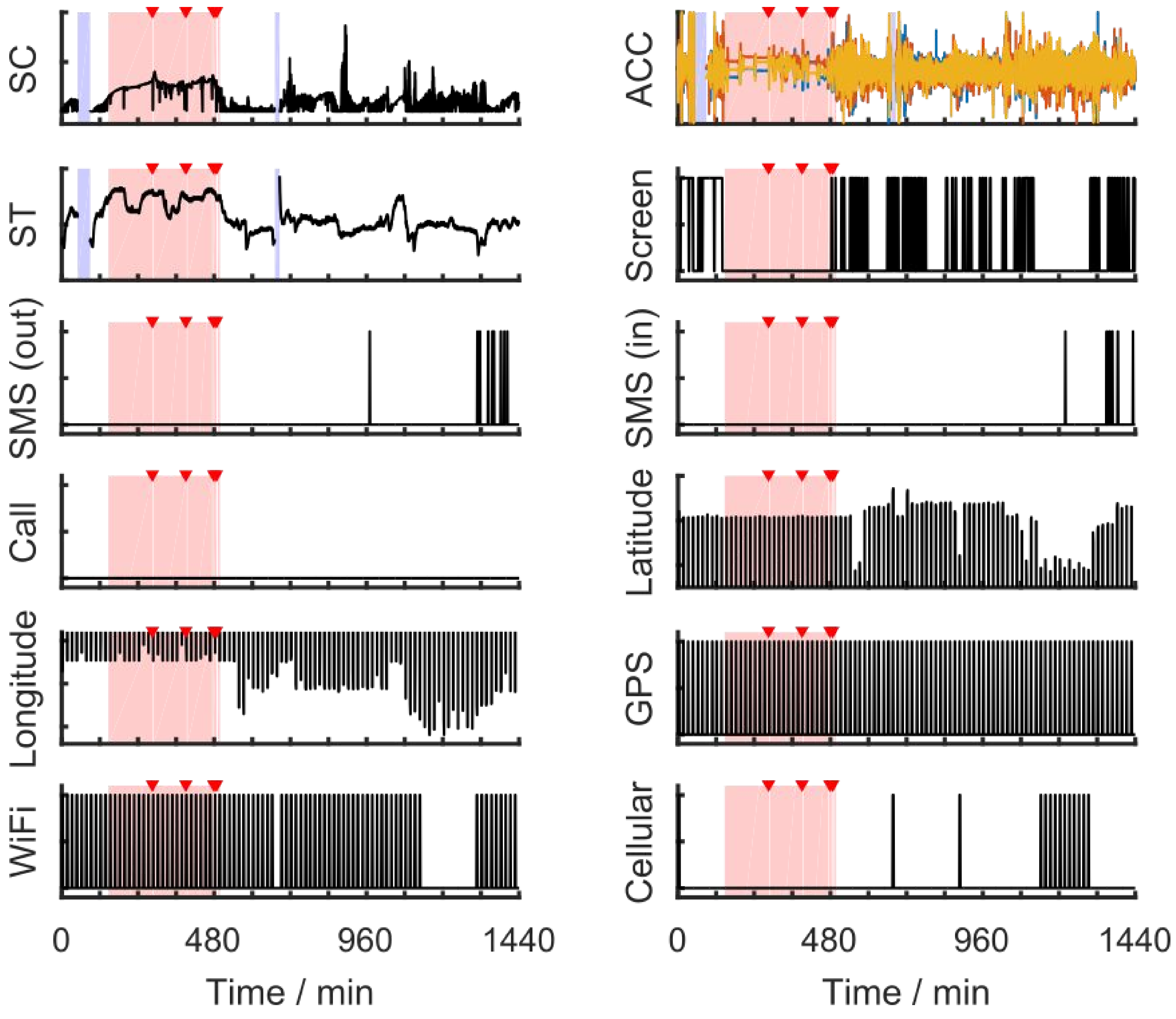
Data

- 5580 days of **multimodal data** from **a wrist sensor and an Android phone**
- 186 undergraduate students, 30 days each

- Wrist Sensor
  Skin conductance (SC)
  Acceleration (ACC)
  Skin temperature (ST)
- Phone
  Call
  SMS
  Location
  Screen

- Time

- Labels of sleep/wake:
  Human scored actigraphy with sleep diaries
  based on a previously established method (Barger et al., 2014)
  Resolution: 1 min -> 1 day = 1440 labels



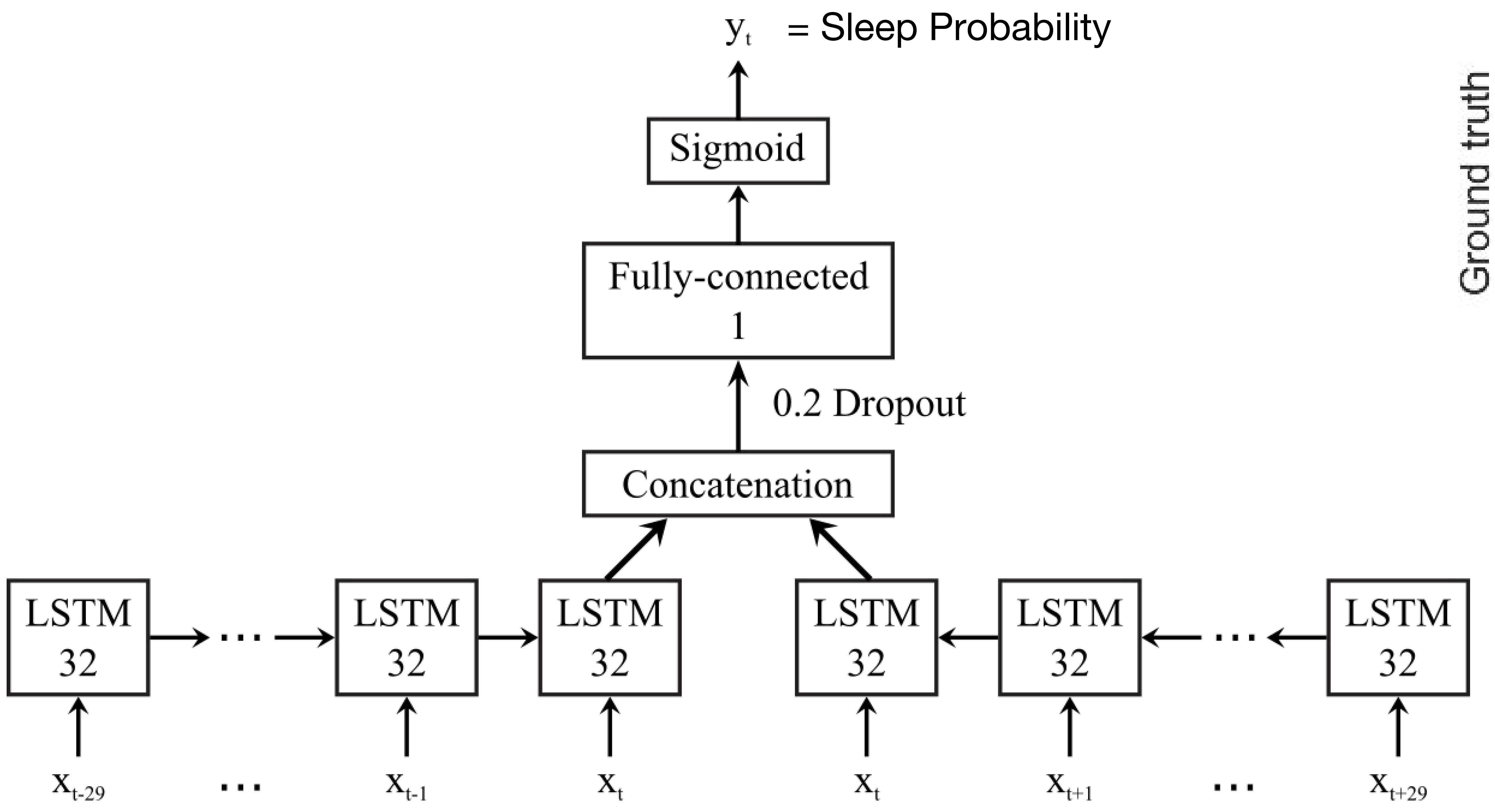
Features

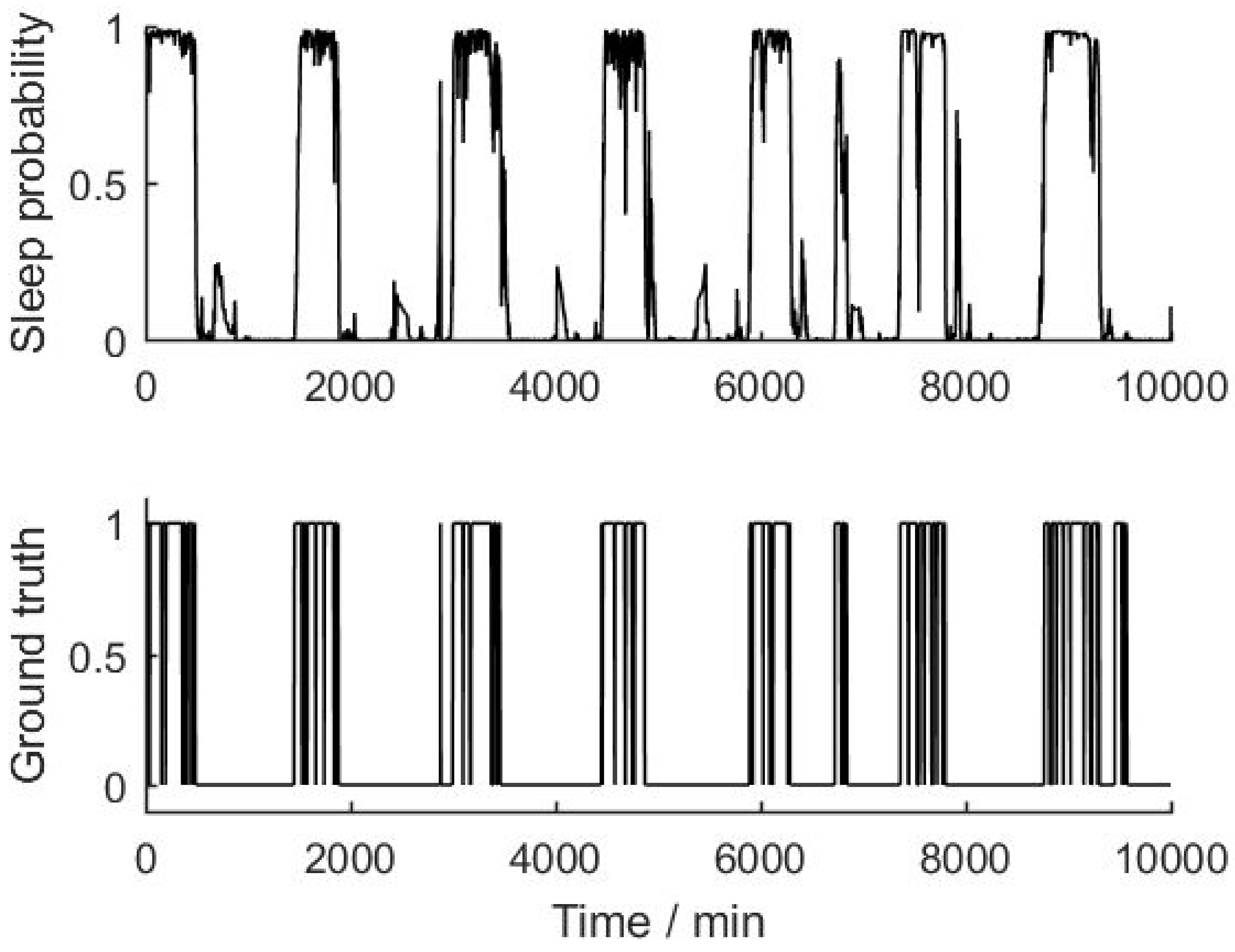
Source	Modality	Feature variables
Wrist sensor	Skin conductance (SC)	Mean, SD, power within 0-0.1, 0.1-0.2, 0.2-0.3, 0.3-0.4, and 0.4-0.5Hz bands, the number of SC responses, storm flag, elapsed time since a storm started
	Acceleration (ACC)	Mean, SD
	Skin temperature (ST)	Mean, SD
Phone	Screen	Screen was on, the time the screen was turned on
	SMS	Sent a message
	Call	On a call, missed a call
	Location	Movement index, connected to WiFi, connected to cellular nets
Time	Time	Elapsed minutes since 12:00 AM

SC is more likely to have periods of high frequency activity called “storms” during NREM2 and SWS sleep
Movement index = (var(latitude) + var(longitude)) / 2

Methods

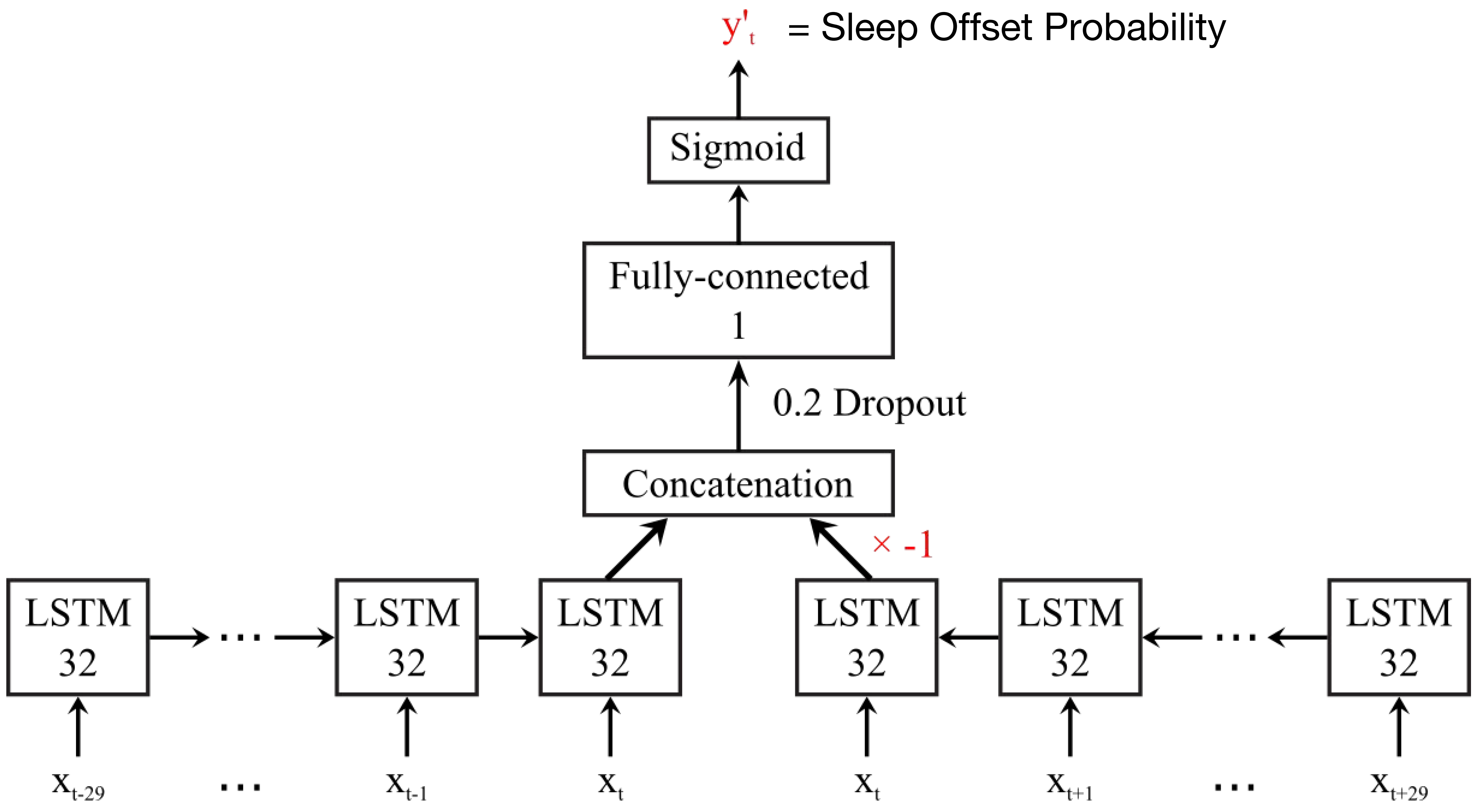
Sleep detection:
Bidirectional long short-term memory neural network model

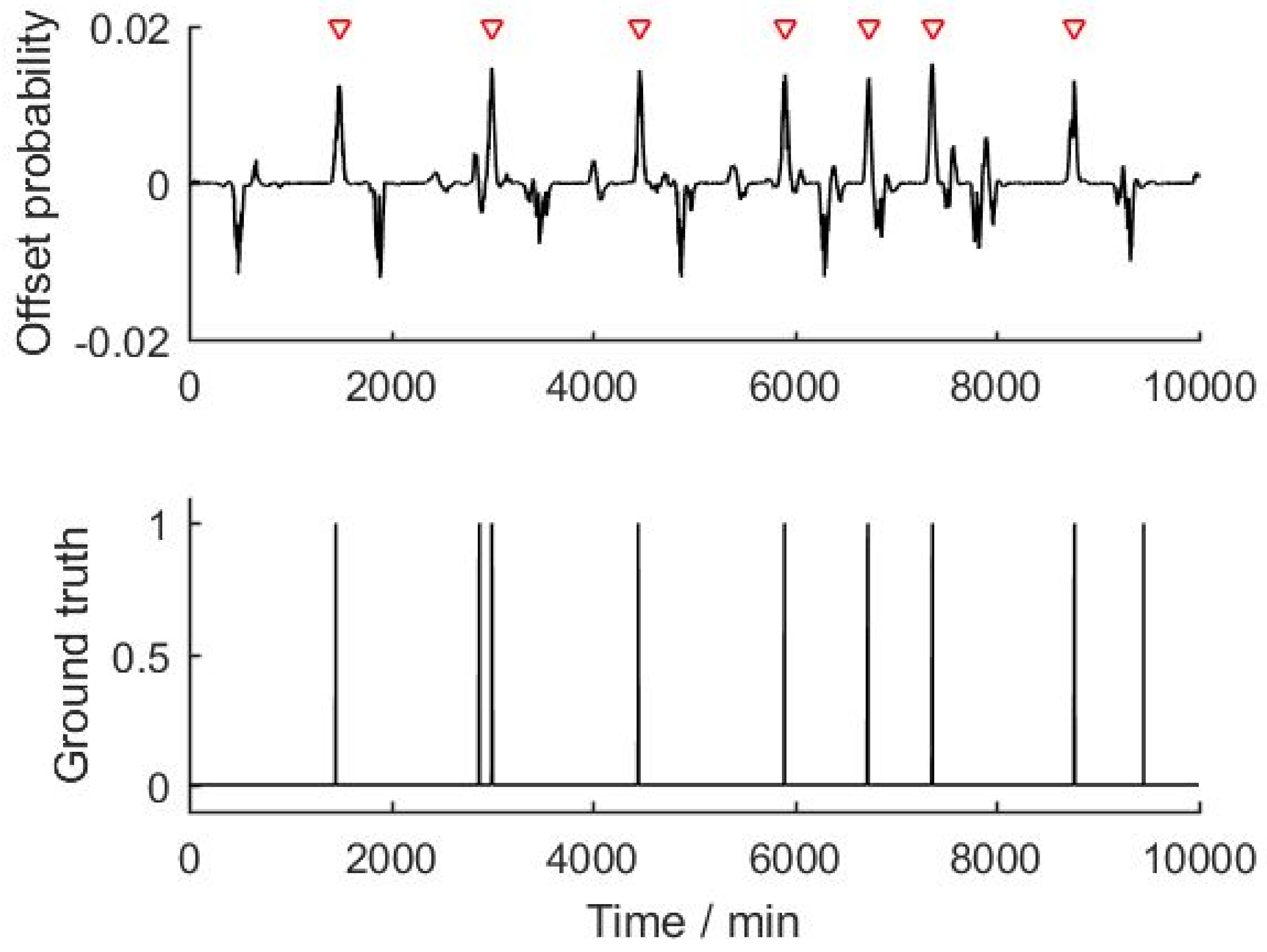




Methods

Sleep episode onset/offset detection:
Bidirectional long short-term memory neural network model
+ Peak detection





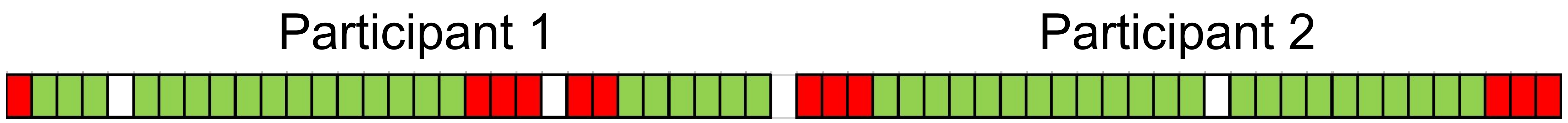
Results

Sleep/wake classification accuracy: **96.5%**
(**Acceleration** + **Skin temperature** + **Time**)

Sleep episode onset detection
F_1 scores: **0.86**, mean errors: **5.0 min**

Sleep episode offset detection
F_1 scores: **0.84**, mean errors: **5.5 min**

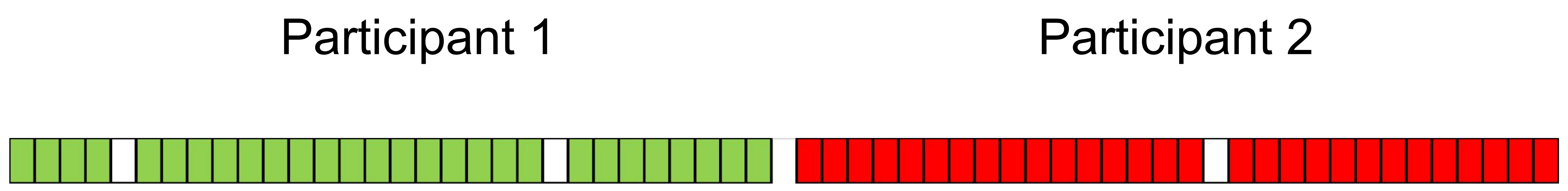
For each participant,
80% of days - training set,
20% of days - test set

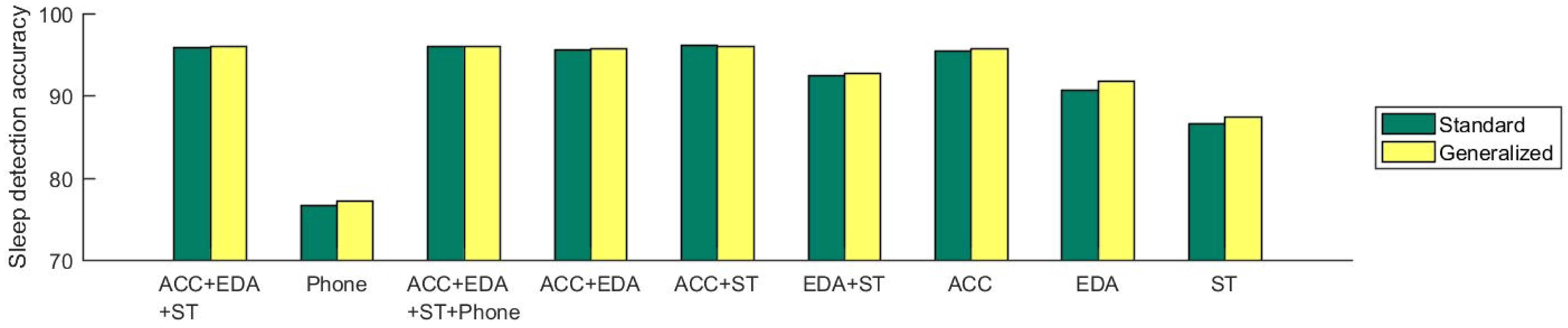


Feature combinations	Sleep detection accuracy		Sleep episode on/offset detection							
			Without time				With time			
	Without time	With time	Sleep onset		Sleep offset		Sleep onset		Sleep offset	
			F_1	ME	F_1	ME	F_1	ME	F_1	ME
Wrist sensor	95.9	96.5	0.85	5.3	0.82	6.3	0.84	5.3	0.84	5.5
Phone	76.7	89.1	0.27	12.4	0.21	15.6	0.43	10.3	0.36	13.2
Wrist + Phone	96.0	96.3	0.84	5.3	0.82	6.0	0.84	5.1	0.82	6.3
ACC + EDA	95.6	96.3	0.84	5.3	0.83	6.4	0.84	5.1	0.83	6.0
ACC + ST	**96.2**	**96.5**	**0.86**	**5.0**	**0.84**	**5.6**	**0.86**	**5.0**	**0.84**	**5.5**
EDA + ST	92.5	94.5	0.74	6.9	0.73	7.2	0.74	6.7	0.74	7.1
ACC	95.5	96.3	0.85	5.4	0.81	6.9	0.84	5.1	0.81	6.5
EDA	90.8	93.7	0.71	8.0	0.68	7.7	0.70	6.4	0.70	7.1
ST	86.7	90.7	0.59	9.8	0.56	11.5	0.52	10.3	0.60	11.0

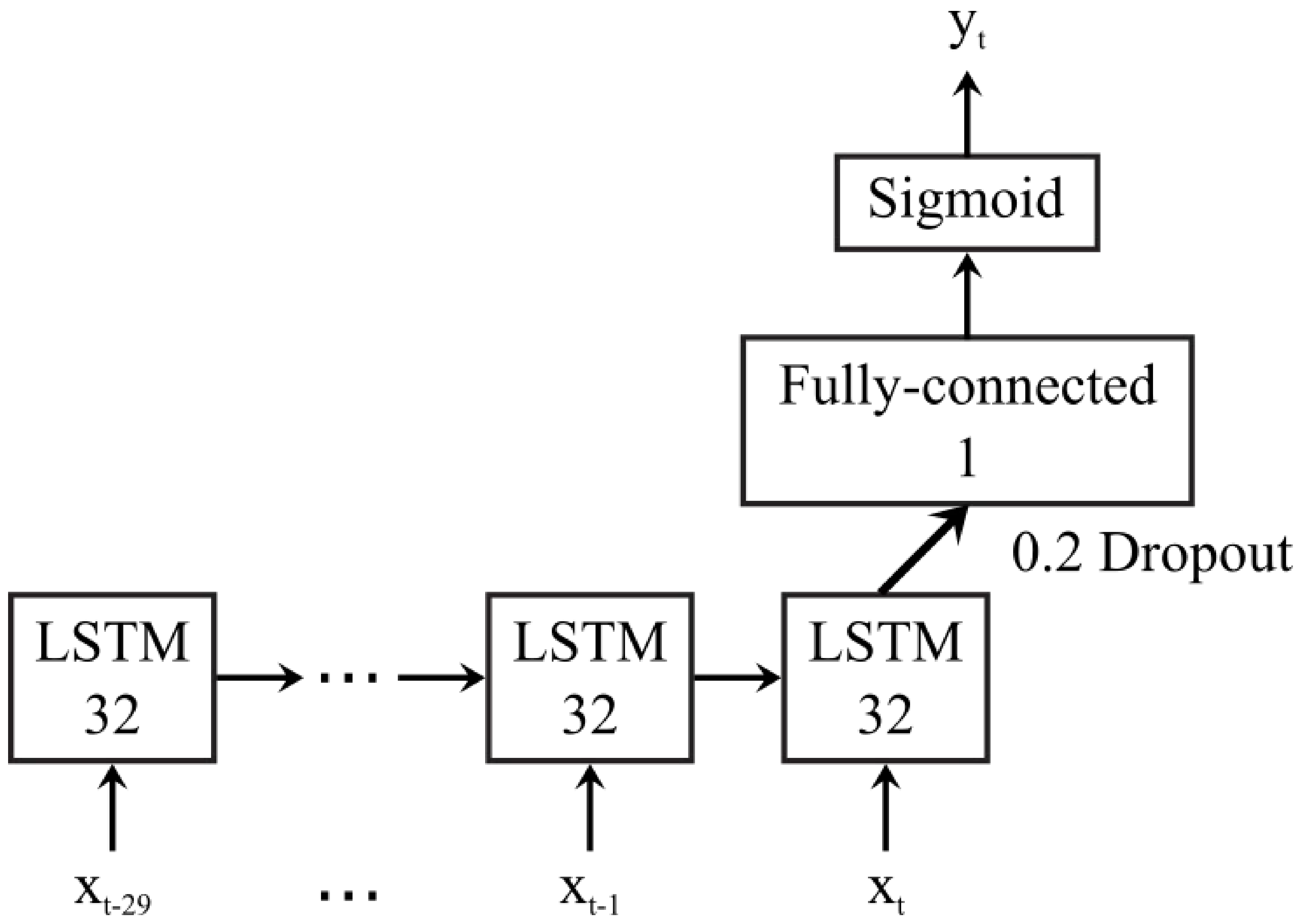
Generalized to different participants

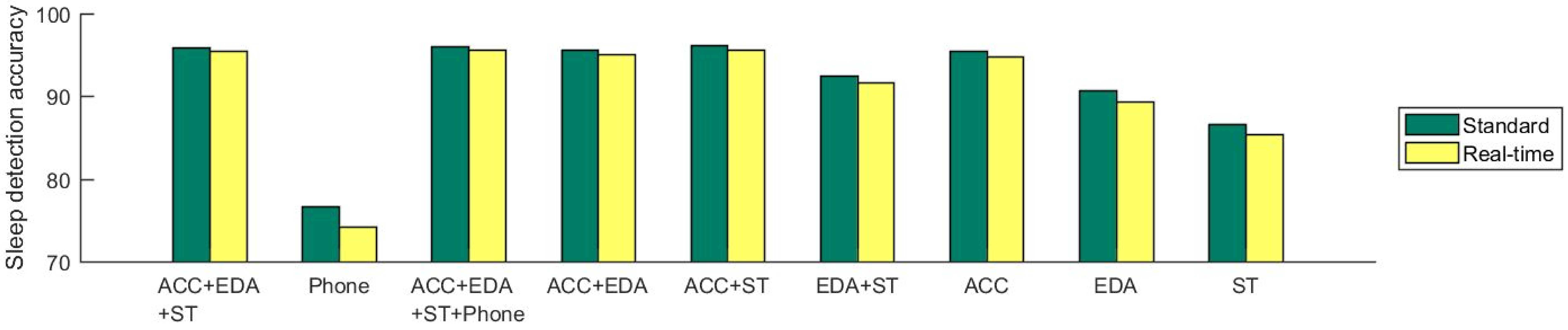
80% of participants - training set
20% of participants - test set





Real-time implementation





Conclusion

We showed

Sleep/wake classification accuracy: **96.5**% with features from **Acceleration** + **Skin temperature** **+** **Time**

Sleep episode onset detection (F_1 scores: **0.86**, mean errors: **5.0 min)**

Sleep episode offset detection (F_1 scores: **0.84**, mean errors: **5.5 min)**

Our results indicate that long-term ambulatory sleep/wake records from large populations can be measured unobtrusively and accurately by exploiting the ubiquity of smartphones and wearable sensors and the power of deep learning.