Sleepless Cities under Social Isolation: Geographically and Temporally Revealing Circadian Rhythm Disorders through Social Media

Mingzheng Yang, Lei Zou, Wanhe Li, Heng Cai, Binbin Lin, Bing Zhou

ABSTRACT: The circadian rhythm typically follows a 24-hour cycle and is primarily influenced by the daily light-dark cycle. Animal experiments have shown that social isolation could elevate the stress on the intrinsic clock that modulates the circadian hormone release, leading to circadian rhythm disorders. However, previous research on human circadian rhythms has largely depended on survey-based methods, limited by sample sizes and geographical and temporal coverage. The widespread social isolation caused by COVID-19 in 2020 provided a unique opportunity to explore the spatial-temporal dynamics of the relationship between social isolation and human circadian rhythms. This research developed circadian rhythm disorder metrics based on data collected from X. Employing these metrics, the study delineates the spatial-temporal patterns and variations of circadian rhythms across 50 major cities in the United States during the pre-pandemic norm (2019), in-pandemic isolation (2020), and post-pandemic recovery (2021) phases. The proposed framework provides valuable insights into revealing the pattern of circadian rhythm disorders based on social sensing and inform experimental biology.

KEYWORDS: *Circadian Rhythm, Social Sensing, Social Isolation, COVID-19, Time-series Analysis*

1. Introduction

Circadian rhythm is a biological term that describes the intrinsic period of the human sleep/wake clock with small inter- and intra-individual ranges (Czeisler et al., 1999). It is a pivotal component of the human system, as it regulates sleep patterns, which substantially influence daily life, such as work efficiency, sleep quality, and overall health (Vitaterna et al., 2001). Despite its significance, millions of people worldwide are experiencing circadian rhythm disorders. The high prevalence of circadian rhythm disorders has raised concerns about their contribution to the development of chronic diseases, e.g., obesity (Rácz et al., 2018), schizophrenia (Bromundt et al., 2011), and heart diseases (Thosar et al., 2018).

These circadian rhythm disorders can be triggered by different social-environmental stressors, such as social isolation, which refers to the state of being separated from society, community, or social interaction (Mistlberger and Skene, 2004). Animal experiments have shown that social isolation could decrease plasma prolactin and growth hormone, increase plasma leptin and corticosterone, and decrease adrenal corticosterone. These changes can elevate the stress on the intrinsic clock that modulates the circadian hormone release, leading to circadian rhythm disorders (Perelló et al., 2006). However, conducting experiments on the etiology of sleep and circadian disorders for humans is challenging due to the limited sample sizes, spatial-temporal coverage, and suitable scenarios.

The widespread social isolation caused by COVID-19 in 2020 provided a unique opportunity to explore the spatial-temporal dynamics of the relationship between social isolation and human circadian rhythms. Meanwhile, diverse social media platforms have enabled billions of users to share feelings and discuss "what's happening" anytime at any place. These user-generated data are inherently time-stamped and sometimes geo-tagged, making them suitable for time series and geographical analyses.

The goal of this research is to utilize location-based social media data from X (formerly Twitter), referred to as tweets, to quantitatively assess the effect of social isolation on circadian rhythms from both spatial and temporal perspectives. Our analysis concentrates on the circadian rhythm dynamics across 50 major cities in the United States from 2019 to 2021, encapsulating the phases of pre-pandemic norm (2019), in-pandemic isolation (2020), and post-pandemic isolation recovery (2021). The objectives are threefold: (1) to develop a framework and a set of indices for measuring city-level, temporally evolving circadian rhythm characteristics and variations using social media data; (2) to evaluate the impact of COVID-19-triggered social isolation on circadian rhythm disturbances; and (3) to reveal the disparities of circadian rhythm disorders and recoveries among the 50 cities. We hypothesize that (a) the pandemic-triggered social isolation disrupted human circadian rhythms, which is reflected by an increase in nighttime social media activity during the in-pandemic isolation phase; (b) the circadian rhythm disorders in most cities recovered in the post-pandemic phase when the social isolation was alleviated; (c) human circadian rhythm disorders and recoveries observed from social media data are geographically and temporally disparate.

2. Data and Method

2.1 Social Media Data Collection and Processing

Figure 1 illustrates the workflow of social media data collection, processing, and analysis. We harvested all geotagged tweets from 01/01/2019 to 12/31/2021 in the United States using the Twitter academic Application Programming Interface (API) v2. The Twitter API for academic researchers grants authorized users' free access to full-archive search with a monthly cap of 10 million tweets per account. The initial collection contains a total of 730 million tweets. Each tweet contains information such as tweet ID, user ID, timestamp, and geolocation.

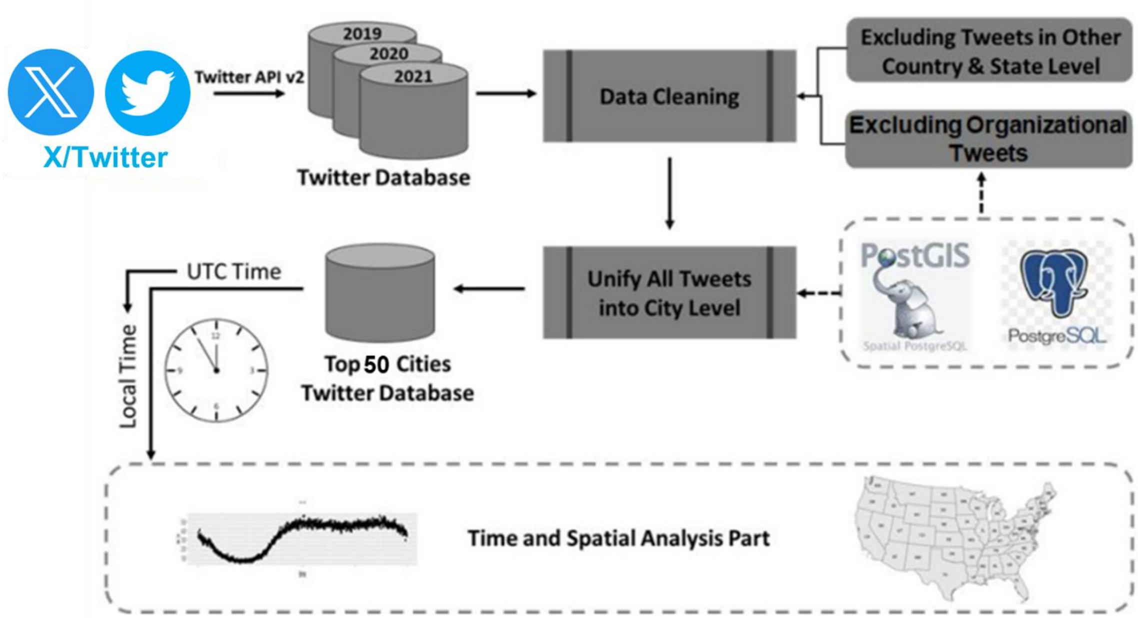

Figure 1: The workflow of Twitter data processing.

2.2 Circadian Rhythm Disorders Metrics

This investigation compared the monthly circadian rhythms from 2019 to 2021. The first step is to generate monthly-aggregated circadian patterns of users' tweeting activity over 24 hours at a 1-minute temporal resolution in each city. Equation 1 is employed to calculate the normalized Social-media Activity Intensity (SAI), which is the proportion of social media posts during time bin *t* in all social media posts in city *i* ($SAI_i(t)$).

$$SAI_i(t) = \frac{\sum_\theta f_i(t)}{\sum_\theta \sum_t f_i(t)} \quad (1)$$

Where θ represents the day in a specific month. For example, if there were 1,000 tweets posted in one city during the 00:00 a.m. to 00:01 a.m. time bin in January 2020, and the total number of tweets posted in the same city for January 2020 was 50,000, then the percentage of tweets during the 1-min bin in January 2020 ($SAI_{Los\ Angeles}(00{:}00 - 00{:}01\ am, January\ 2020)$ is 0.02.

The second step is to generate Social-media Activity Rhythm (SAR) by smoothing the SAI with a 150-min Gaussian kernel window and 60-min standard deviation (Figure 2). The purpose of the smoothing process is to mitigate the influences of abnormal values in SAI when evaluating the circadian rhythm variations.

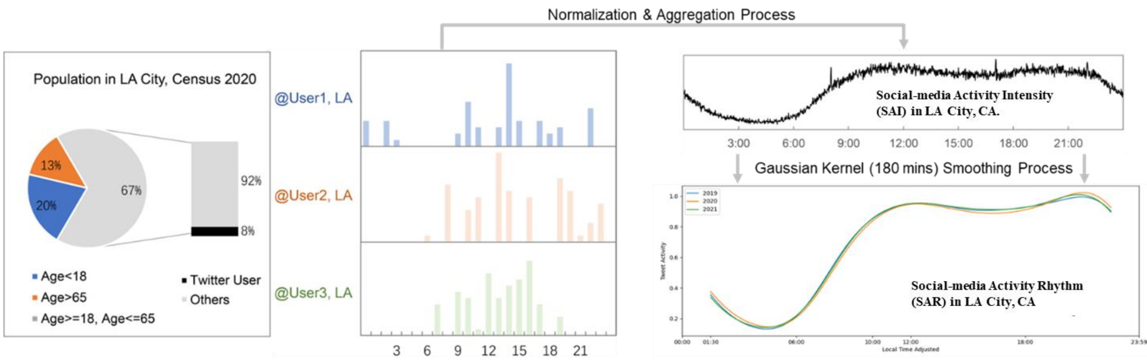

Figure 2. The process of calculating Social-media Activity Intensity (SAI) and Social-media Activity Rhythm (SAR) with Los Angeles City, CA, as a case study.

The third step involves constructing indices to measure circadian rhythms and disorders in each city. As suggested by Zhou et al. (2023), this study defines the time frame from 1:30 a.m. to 10:00 a.m. as a homogeneous period for delineating circadian rhythms and their variations in the pre-, during, and post-isolation phases (Figure 3a). We calculated three circadian rhythm disorder metrics based on social media activities in the homogeneous period, including the Lowest Activity Time (LAT), Rhythm Disorder Area (RDA), and Rhythm Shifting Time (RST), as indicated in Figure 3b. LAT represents the time when a city exhibits the least active Twitter activity (smallest SAI value; green and red points in Figure 3b). RDA describes the circadian rhythm differences during nocturnal and matutinal periods (the sum of red and green areas in Figure 3b). Finally, we employed the Time

Domain Cross-Correlation Function (Chao and Chung, 2019) to quantify RST (the differential value between the red and green points in Figure 3b).

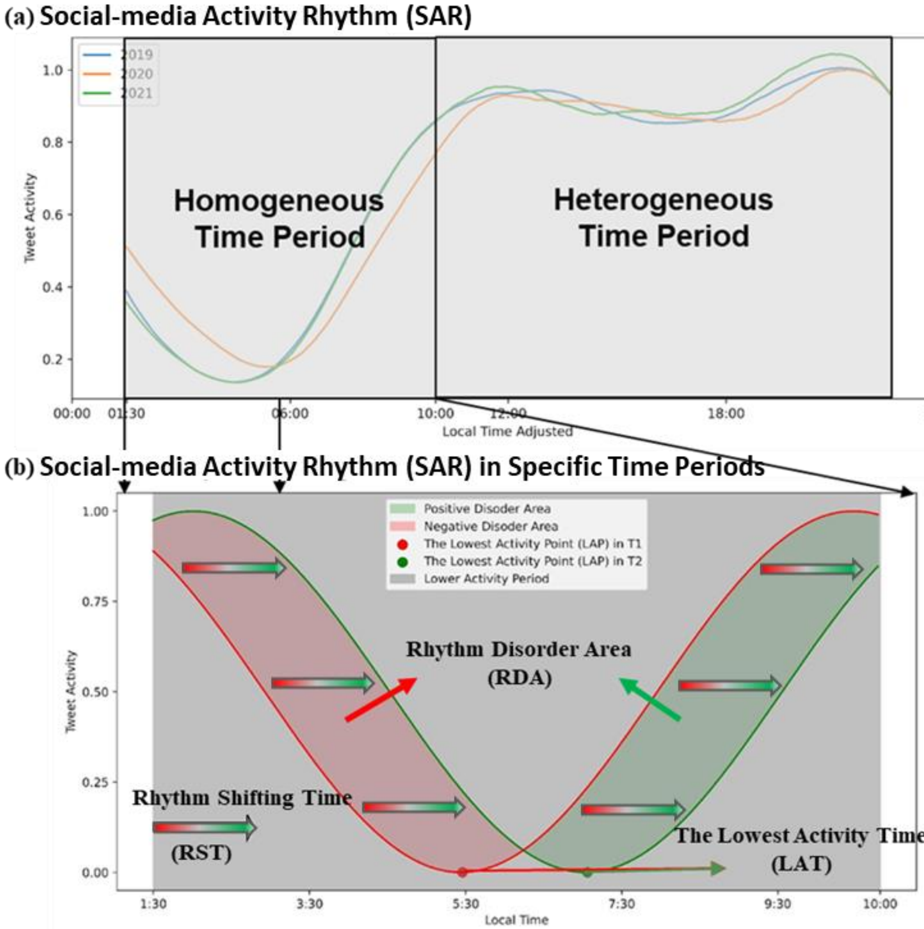

Figure 3. Calculation of circadian rhythm disorders using Social-media Activity Rhythm (SAR)

3. Results

3.1 The temporal variations of circadian rhythms

The averaged temporal variation of the LAT in the 50 cities from 2019 to 2021 is displayed in Figure 4. In 2019 and 2021, LAT was highest in March (T1) and lowest in November (T3). This is because daylight saving time (DST) in the United States starts in March and ends in November (Kantermann et al., 2007). However, there is a discrepant temporal pattern of LAT during the COVID-19 pandemic outbreak in 2020. The LAT in April 2020, when the first outbreak of COVID-19 occurred in the United States, is later than any other month of this year. This observation signifies the impact of COVID-19-induced social isolation on changing the circadian rhythm derived from social media users' activities.

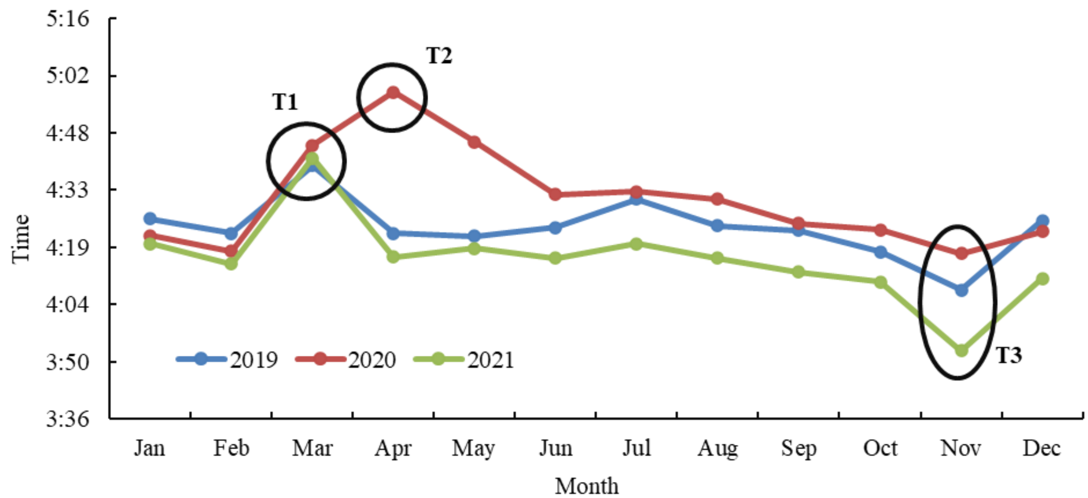

Figure 4. The average lowest Twitter activity time in 50 cities

Since there was no COVID-19 pandemic in 2019, the circadian rhythms in 2019 were selected as the control group. Figure 5 shows the Twitter-derived averaged monthly RST in 2019-2020 and 2019-2021 in the 50 cities. In January and February 2020, when COVID-19 was a regional infectious disease, a slight difference in circadian rhythms between 2019 and 2020 was detected in major cities in the United States. However, from March to December 2020, the circadian rhythms have been significantly delayed, especially in April (delayed by 31 minutes) and in May (delayed by 23 minutes). In contrast to the circadian rhythm disorders in 2019-2020, the circadian rhythms in 2021 were similar to that in 2019, with a shifting time of fewer than 10 minutes in every month. It demonstrates the first hypothesis that the social isolation during the early outbreak of the COVID-19 pandemic had a significant effect on the circadian rhythms of social media users.

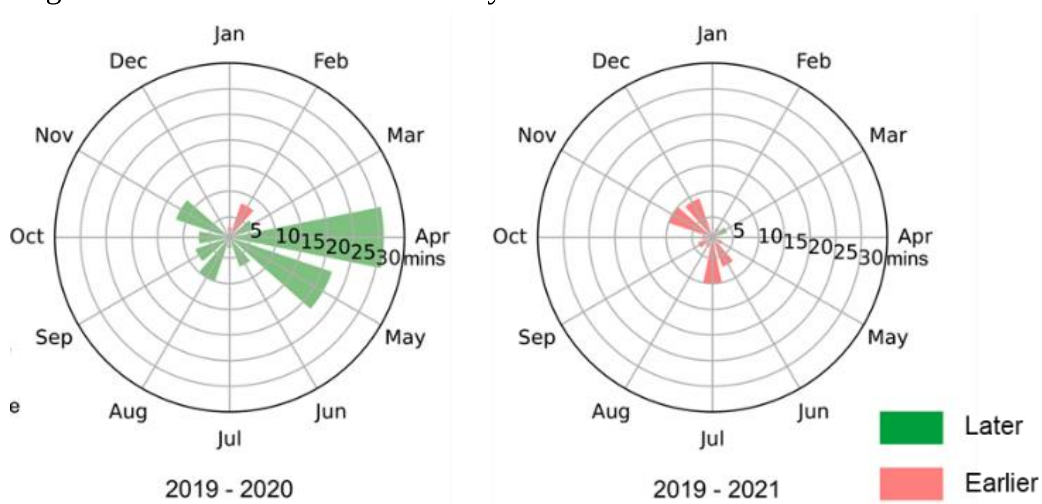

Figure 5. The temporal pattern of circadian rhythm disorders

3.2 Geographic disparities of circadian rhythm disorders

The geographical disparities of city-level RST (the delay of the time when X users' activity is the lowest) in April between 2020 and 2019 and between 2021 and 2019 are shown in Figure 6. When comparing the social media-derived circadian rhythms in April 2020 with

April 2019, 47 out of the 50 major cities displayed an evident delay. On the other hand, no obvious disorders were detected in most cities, except for Tallahassee, FL, and Kansas City, MO, when comparing the Twitter-computed circadian rhythm differences between April 2021 and April 2019.

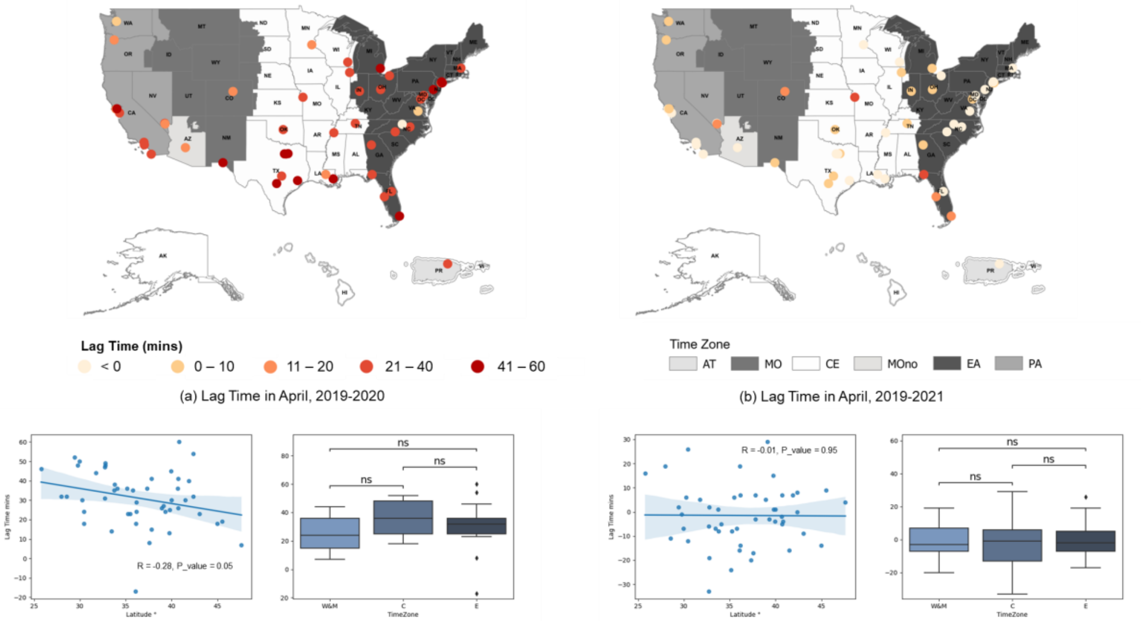

Figure 6. The geographical pattern of circadian rhythm disorders

Table 1 shows the 10 cities with the largest circadian rhythm disorders during the pandemic based on the RDA and RST indices. Bronx, NY, ranked first, with the highest RDA and RST of 58.04 and 60 minutes in April 2020, respectively. Meanwhile, all of the most severe city-level circadian rhythm disorders occurred in April 2020. Five cities in Texas, including Arlington, Dallas, San Antonio, Fort Worth, and Houston, are ranked in the top ten. The result implies that the biological clocks of people living in Texas are likely to be more vulnerable to social isolation compared with residents in other cities. Further research to identify the potential social and environmental determinants of this observation, e.g., climate and built-environment environments as well as social isolation degrees and durations, are necessary.

Table 1. Top 10 cities' circadian rhythm disorders.

	Rhythm Disorder Area			**Rhythm Shifting Time**		
Rank	City	RDA (Norm Activity *Min)	Month	City	RST (Mins)	Month
1	Bronx, NY	58.04	April	Bronx, NY	60	April
2	Arlington, TX	56.11	April	Detroit, IL	54	April
3	New Orleans, LA	53.21	April	San Antonio, TX	52	April
4	Miami, FL	50.22	April	New Orleans, LA	50	April

5	Detroit, IL	49.07	April	Fort Worth, TX	49	April
6	Milwaukee, WI	48.95	April	Houston, TX	48	April
7	Dallas, TX	47.41	April	Dallas, TX	48	April
8	San Antonio, TX	47.17	April	Arlington, TX	47	April
9	Houston, TX	46.01	April	Miami, FL	46	April
10	Philadelphia, PA	45.55	April	Philadelphia, PA	45	April

4. Conclusions

This study innovatively utilizes nocturnal and matutinal social media activities to approximate city-level circadian rhythms and investigate circadian disorders during pandemic-induced social isolations. Using the designed indices, this research has examined circadian rhythm disorders and recoveries in the 50 major cities in the United States from 2019 to 2021. Three hypotheses have been tested and confirmed. The results demonstrate that social media activities can accurately depict the circadian rhythms of residents in different cities. Furthermore, social isolation during the pandemic lockdown significantly impacts the circadian rhythms of human populations, with the extent of these effects varying by location. This study confirms the phenomenon observed in animal experiments with real-world human-generated data and sheds light on experimental biology for an improved understanding of the etiology of circadian rhythm disorders.

References

Cacioppo, John T., Louise C. Hawkley, L. Elizabeth Crawford, John M. Ernst, Mary H. Burleson, Ray B. Kowalewski, William B. Malarkey, Eve Van Cauter, and Gary G. Berntson, 2002. Loneliness and health: Potential mechanisms. *Psychosomatic medicine*, 64(3): 407-417.

Cellini, N., Canale, N., Mioni, G. and Costa, S., 2020, Changes in sleep pattern, sense of time and digital media use during COVID-19 lockdown in Italy. *Journal of sleep research*, 29(4), p.e13074.

Chao, B.F. and Chung, C.H., 2019, On estimating the cross correlation and least squares fit of one data set to another with time shift. *Earth and Space Science*, 6(8), pp.1409-1415.

Hwang, T. J., Rabheru, K., Peisah, C., Reichman, W., & Ikeda, M., 2020. Loneliness and social isolation during the COVID-19 pandemic. *International psychogeriatrics*, 32(10), 1217-1220.

Kantermann, T., Juda, M., Merrow, M. and Roenneberg, T., 2007. The human circadian clock's seasonal adjustment is disrupted by daylight saving time. *Current Biology*, 17(22), pp.1996-2000.

Zhou, K., Constantinides, M., Quercia, D. and Šćepanović, S., 2023, June. How Circadian Rhythms Extracted from Social Media Relate to Physical Activity and Sleep. *In Proceedings of the International AAAI Conference on Web and social media*, Vol. 17, pp. 948-959.

Mingzheng Yang, Ph.D. Student, Department of Geography, Texas A&M University, TX 77843

Lei Zou, Professor, Department of Geography, Texas A&M University, TX 77843

Wanhe Li, Professor, Department of Biology, Texas A&M University, TX 77843

Heng Cai, Professor, Department of Geography, Texas A&M University, TX 77843

Binbin Lin, Ph.D. Student, Department of Geography, Texas A&M University, TX 77843

Bing Zhou, Ph.D. Candidate, Department of Geography, Texas A&M University, TX 77843