

# 基于百度指数的上海市气象信息网络关注度分析

郑泽华<sup>1</sup> 穆海振<sup>2</sup> 孔春燕<sup>3</sup> 余延略<sup>1</sup>

(1 上海市青浦区气象局, 上海 200030; 2 上海市气象局, 上海 200030; 3 上海市气象服务中心, 上海 200030)

**摘要:** 利用百度指数和气象资料数据, 将“上海天气”“上海高温”和“上海台风”作为搜索关键词, 对上海气象信息的网络关注度进行统计分析, 研究其网络关注度特征。研究发现: 1) 以“上海天气”为关键词, 2014—2016年关注度呈上升趋势, 2017—2021年变化不明显; 春季和夏季关注度较高, 初秋和冬末较低; 周内变化不明显, 周一相对较高; 法定节假日期间关注时间提前量在2~4天左右, 假期内的变化特征存在差异, 但假期结束3天内仍维持较高的关注度。2) 以“上海高温”为关键词, 公众关注度与日最高气温的变化趋势基本一致。公众对首次40℃的高温及连续高温天气关注度较高, 并在6月底—7月初已对高温天气有所关注。3) 以“上海台风”为关键词, 公众关注度与日极大风速、日降水量的变化趋势基本一致; 公众对“上海台风”的关注起始时间与台风的生成时间相吻合; 关注度显著增多出现在台风本体影响的前2天左右。围绕气象部门如何优化公众气象服务、建立气象舆情数据库、运用新媒体传播等方面提出建议, 以期气象服务提供借鉴。

**关键词:** 百度指数, 上海天气, 上海高温, 上海台风

**中图分类号:** P4

**文献标志码:** A

**DOI:** 10.3969/j.issn.2095-1973.2024.06.008

## Analysis of Network Attention of Shanghai Meteorological Information Based on Baidu Index

Zheng Zehua<sup>1</sup>, Mu Haizhen<sup>2</sup>, Kong Chunyan<sup>3</sup>, Yu Yanlue<sup>1</sup>

(1 Shanghai Qingpu Meteorological Service, Shanghai 200030; 2 Shanghai Meteorological Service, Shanghai 200030;

3 Shanghai Meteorological Service Center, Shanghai 200030)

**Abstract:** Using Baidu index and meteorological data, we analyze the network attention of Shanghai meteorological information with “weather in Shanghai”, “high-temperature in Shanghai” and “typhoon in Shanghai” as search keywords. The results show that the network attention of “weather in Shanghai” shows an overall upward trend from 2014 to 2016, with no significant variation from 2017 to 2021; The attention is mostly high in spring and summer, while low in early autumn and late winter; The weekly variation of the network attention is not obvious, with Monday experiencing relatively higher attention; In case of legal holidays, the attention peaks 2 to 4 days prior to the holiday, the change characteristics differ from each other within the holiday, and the high attention remains within 3 days after the holiday. The network attention of “high-temperature in Shanghai” shows the consistence of the variation trends between that of the public attention and of the daily maximum temperature; The high attention goes to the first 40℃ and the continuous high-temperature weather, with the start of the attention as early as the late June and the early July. The network attention of “typhoon in Shanghai” shows the consistence of the variation trends between that of the public attention and of the daily maximum wind speed and the daily precipitation; The start time of the attention matches with the start time of the typhoon; The attention of the typhoon increases gradually and rises significantly around 2 days prior to the strike of the typhoon itself. Finally, suggestions on how meteorological authorities can optimize public weather services, establish a meteorological public opinion database, and utilize new media for data dissemination are propose in hopes of providing a reference for meteorological services.

**Keywords:** Baidu index, weather in Shanghai, high-temperature in Shanghai, typhoon in Shanghai

### 0 引言

大数据时代, 公众通过网络搜索引擎获取信息, 网络关注度可以作为互联网搜索量的重要衡量指标<sup>[1]</sup>。在各类网络搜索引擎中, 百度是中国最大的搜索引

擎。百度指数<sup>[2]</sup>是基于网民在百度上的搜索量产生的统计数据, 以关键词为统计对象, 科学分析并计算出各个关键词在百度网页搜索中搜索频次的加权。百度指数作为一种研究方法已获广泛认可, 并取得了很多的研究成果, 有基于百度指数研究旅游<sup>[3-4]</sup>、投资<sup>[5-6]</sup>、生态文明<sup>[7-8]</sup>的网络关注情况, 也有基于百度指数开展手足口等疾病<sup>[9-10]</sup>、交通客运量<sup>[11-12]</sup>的预测。同时也

收稿日期: 2023年6月28日; 修回日期: 2023年12月26日  
第一作者: 郑泽华(1992—), Email: 1006520538@qq.com

有学者将气候资料与网络关注度相结合，王硕等<sup>[13]</sup>分析城市气候舒适度与游客网络关注度的变化特征，并构建相关模型。刘璐等<sup>[14]</sup>通过分析蓬莱市旅游气候舒适度与旅客网络关注度关联性，发现旅客网络关注度的改变一定程度上受气候影响。马丽君等<sup>[15]</sup>分析中国30个城市气候舒适度和游客网络关注度，并建立了两者的时空模型。周威等<sup>[16]</sup>基于气象观测数据及百度指数，研究天气与旅游的相关性。

本文基于百度指数，将“上海天气”“上海高温”和“上海台风”作为搜索关键词，研究气象相关关键词的变化趋势，结合气象要素数据掌握公众对气象的关注特点，为更好地开展气象服务提供参考依据。

## 1 数据与方法

### 1.1 研究设计

关键词搜索是网络搜索引擎信息检索的主要方法之一。公众通过在搜索引擎上输入特定的关键词来寻找合适的目标信息。通过研究公众获取气象信息的行为规律，有助于气象部门更加具有针对性、更加高效地向公众提供气象服务信息。本研究选取“上海天气”“上海高温”和“上海台风”作为搜索关键词。

为剔除年份差异的影响，计算日相对搜索指数：

$$E = \frac{m}{M}, \quad (1)$$

式中， $E$ 表示相对搜索指数； $m$ 表示日搜索指数； $M$ 表示当年日均搜索指数。

### 1.2 数据收集

通过在百度指数输入上文所述关键词，获取2014—2021年逐日上海本地的用户关注度指数，利用统计工具对搜索数据进行统计分析。

## 2 结果与分析

### 2.1 基于“上海天气”关键词的公众关注度分析

#### 2.1.1 年度变化趋势

以2014—2021年为统计时间段，分析以“上海天气”为关键词的百度搜索指数值(图1)，将逐日数据绘制成折线图。总体上呈先升后降的趋势，2015年和2016年搜索指数较高，这两年的年总搜索指数达62301540和62307786。2017年起降低且趋于平稳，各年份最高值差异逐渐缩小，关注度的降低可能与基于移动端的天气预报软件兴起有关，公众获取信息的渠道越来越多。

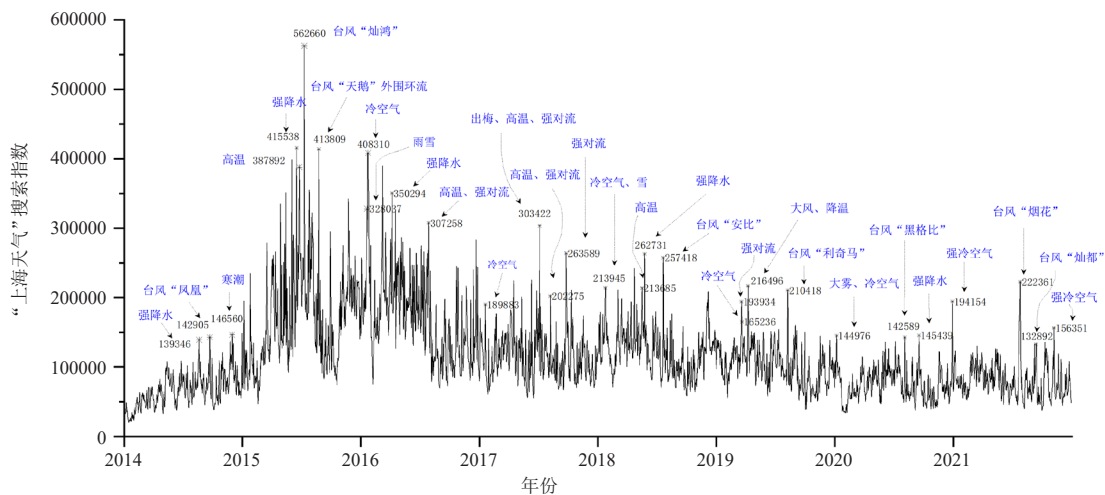


图1 2014—2021年“上海天气”网络关注度

Fig. 1 The network attention of “weather in Shanghai” from 2014 to 2021

#### 2.1.2 月变化趋势

根据日搜索指数累加得到月搜索指数，分析月变化趋势表明：1) 公众的关注程度与当月的天气状况有较大的联系，3—7月，11—12月的相对搜索指数均位于多年均值之上，其中以7月为最高，其余月份数值接近。其原因为3—5月，处于季节转换期，天气形势多变，气温、降水变化幅度比较大有关；上海一般自6月起进入汛期，特别是6月、7月为梅汛期，强降水、

强对流等灾害性天气频发；12月处于冬季，多寒潮和降温天气。2) 2月、9月和10月的相对搜索指数较集中，且低于多年均值，可能与该时段内天气相对平稳，灾害性天气相对较少有很大的关系。3) 从四季的划分来看，春季和夏季关注度较高，初秋和冬末较低。

#### 2.1.3 周变化趋势

以周一至周日为统计，剔除节假日影响，根据日相对搜索指数，累加求平均值得到多年周一至周日的

百度搜索指数均值。一周的变化趋势基本呈现先降后升的趋势,可能与公众的生活工作安排有较大联系。周一作为周内的首个工作日,相应的搜索指数为最高,可能和市民多在周一根据天气趋势做本周工作生活安排相关;之后关注度逐渐降低,周六最少,周日关注度又明显增多。

#### 2.1.4 法定节假日变化趋势

选取“元旦”“春节”“清明”“五一”“端午”“中秋”和“国庆”节假日相对搜索指数数据分析网络关注度趋势特点,因假期长短差异,选取最长时段作为时间序列,故图中有些年份出现线段中断的情况。研究发现节假日期间公众对天气信息的网络关注具有一定提前量,故统计时段选取假期前5天和假期后3天(图2)。

“元旦”假期,2014、2016、2018、2019和2020年假期前5天网络关注度已达较高值,之后开始下降。所有年份中一般假期首日的关注度最低,之后呈增加趋势。假期后网络关注度仍位于较高值,一般在假期结束的1~2天。“春节”假期前4天左右网络关注度较高,可能与公众安排出行计划有关,且在假期内第1~2天降至最低。假期内基本呈现升—降—升的趋势,其中4年在假期内第3天左右达到峰值,可能与市民开始走亲访友,故关注天气信息较多有关。之后有小幅下降,基本在假期内最后1天关注度回升至较高值。假期后结束的3天内网络关注度仍维持较高值。

“清明”假期前2~4天网络关注度较高,之后逐步降低。其中5年的假期首日关注度相对较高,可能与清明节安排祭扫活动有关。“五一”假期前4天左右网络关注度较高,假期首日关注度较低。其中假期为3天的年份,假期内呈增加趋势。假期为5天的年份,假期内呈先升后降的趋势。假期结束后首日关注度仍维持高值。“端午”前2~3天网络关注度较高,假期内第1~2天关注度相对较低。假期后关注度回升,在结束后的1~2天增加至较高值。

“中秋”假期因2017年和2020年中秋假期处在国庆假期内,故仅选取假期当日的的数据。假期前3~4天网络关注度较高,假期首日的关注度基本较低,假期结束后第1天关注度回升至较高值。“国庆”假期前2~3天网络关注度较高,之后到假期首日呈下降趋势。假期内基本呈现先降后升的趋势,所有年份自假期内第6天起关注度增加,在假期结束后的2天关注度仍较高。

#### 2.2 基于“上海高温”关键词的公众关注度分析

选取2013—2022年每年5—9月的徐家汇站点气象数据,对比分析“上海高温”搜索指数和日最高气温的变化趋势(图3)。因高温天气多集中在7—8月,其

间搜索指数波动较集中,日搜索指数最高值基本出现在7月,仅有2018年出现在5月,2020年出现在8月。

#### 2.3 基于“上海台风”关键词的公众关注度分析

以2012—2021年为统计时间段,分析以“上海台风”为关键词,重点分析每年6—10月的搜索指数变化趋势,如图4。搜索趋势与台风的影响时段高度匹配,且影响的程度大,搜索量相应较高。

### 3 结论与讨论

以“上海天气”“上海高温”和“上海台风”关键词为例,基于百度指数数据,分析了公众对气象的网络关注度趋势,及其与相应气象资料数据的关系。

预报预警、气象服务在公众关注的视野里占据的位置越来越重要,气象部门应重点根据公众的关注特点和频次。一是在春节、元旦、清明、五一等法定节假日期间前2~4天左右和假期结束后,分别制作天气提醒。二是在本文研究的40℃的高温,及其他极端气候事件出现前及时开展科普解读、提供防御指引,提前制定天气展望材料。三是预计台风将影响本地,及时发送台风生成信息及跟踪监测服务。随着台风的移动趋势,提供预测信息,并在预测台风本体影响的前2天左右,适当增加服务频次,及时提供实况及未来预测信息。

新媒体时代下,舆情信息传播的深度与广度随之扩大,一旦传播的负面舆情信息影响力超出可控范围,即会引起突发性舆情危机事件<sup>[17]</sup>。2016年,一条“据气象局预报,预计23—26日上海市和浙江省将出现1992年以来最寒冷的天气”“23—26日为极端最低气温,将降至-13~-9℃”……的消息被频频转发。而该消息内容又经篡改,在2017年再次出现。幸而气象部门和相关单位通过专家解读、发布辟谣文章等方式迅速平息谣言。可见极端天气事件新媒体传播过程中,因公众难以科学掌握事件情况,出现碎片化、大范围的传播。因此对于气象部门,一是要借助大数据等技术,建立气象舆情信息数据库,定期抓取公众对气象相关信息的搜索数据,提高舆情研判的准确性和预见性,当监测到相关舆情动态及时发声。二是优化公众气象服务内容,凝练关键气象预报预警信息,以更易理解的表述及展现形式,满足公众对气象信息的需求。三是拓宽传播渠道,与主流媒体建立机制,对于公众的讨论热点事件,进行专业解读,借助其权威发布平台发布进行回应。

#### 参考文献

[1] 陈选. 基于GTWR的中国生态文明网络关注度时空特征及影响因

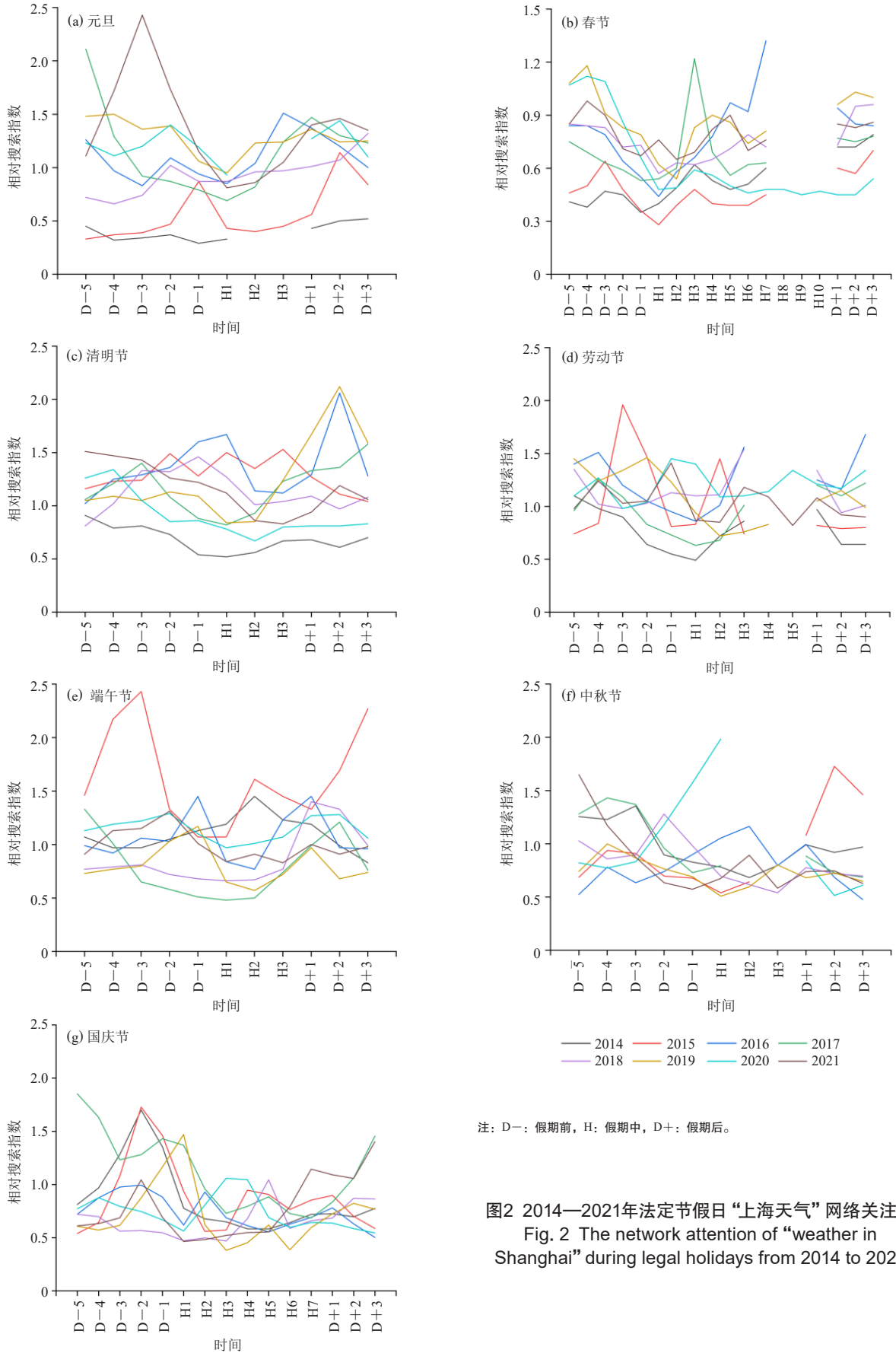


图2 2014—2021年法定节假日“上海天气”网络关注度  
Fig. 2 The network attention of “weather in Shanghai” during legal holidays from 2014 to 2021

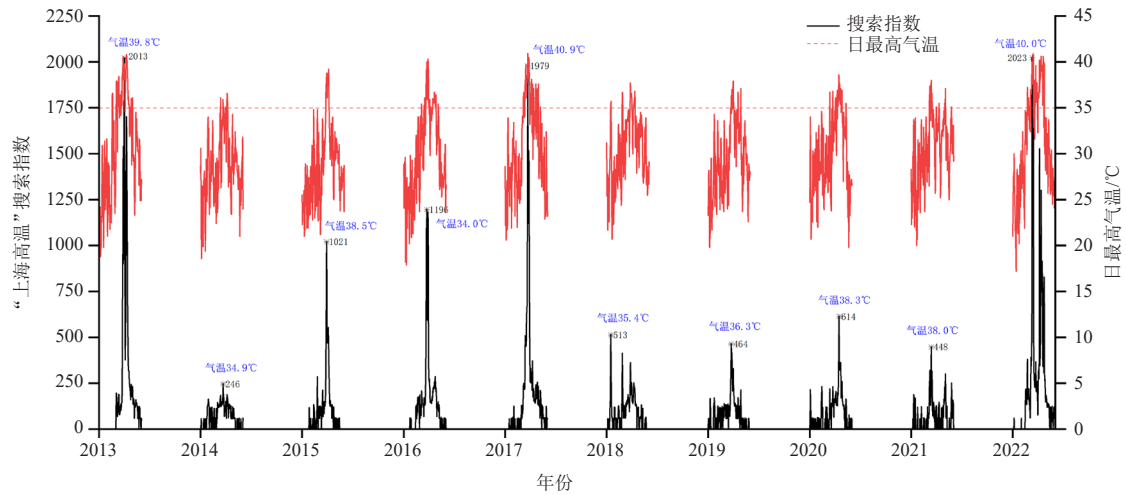


图3 2013—2022年5—9月“上海高温”网络关注度与日最高气温对比

Fig. 3 Comparison of the network attention of “high-temperature in Shanghai” and daily maximum temperature from May to September of 2013—2022

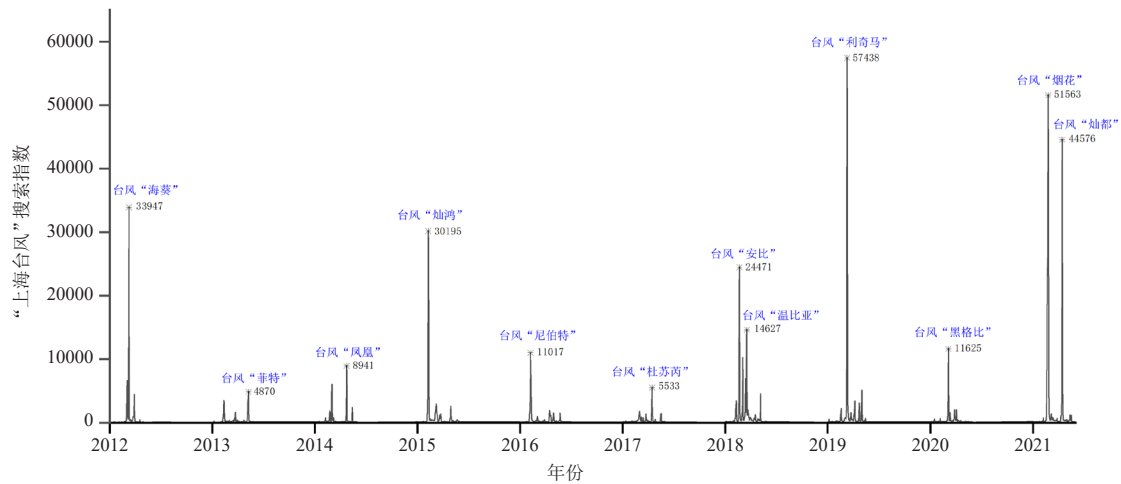


图4 2012—2021年“上海台风”网络关注度

Fig. 4 The network attention of “typhoon in Shanghai” from 2012 to 2021

素研究[D]. 武汉: 武汉大学, 2020.

- [2] 百度指数[EB/OL]. [2023-06-20]. <https://index.baidu.com/v2/index.html#/>.
- [3] 王璐, 皮常玲, 郑向敏. 中国海洋旅游安全事故时空特征分析[J]. 热带地理, 2022, 42(7): 1148-1157.
- [4] 张宇飞, 任艳路, 梁丽敏. 中国冰雪旅游网络关注度的时空差异与影响因素——基于百度指数的实证数据[J]. 世界地理研究, 2024, 33(10): 155-167.
- [5] 张学勇, 唐国梅. 行业关注度与股票横截面收益率——基于百度行业搜索指数的研究[J]. 经济学(季刊), 2022, 22(3): 773-794.
- [6] 梁超, 魏宇, 马锋, 等. 投资者关注对中国黄金价格波动率的影响研究[J]. 系统工程理论与实践, 2022, 42(2): 320-332.
- [7] 冯霞, 张目. 基于百度指数的贵州地级市生态文明建设水平评价研究[J]. 环境保护科学, 2020, 46(4): 8-15.
- [8] 李智山, 邓一荣. 中国生态文明网络关注度区域差异及其影响因素[J]. 生态经济, 2020, 36(2): 208-213.
- [9] 周骥, 彭丽, 付世华, 等. 气象因素和百度指数在手足口病预报模型中的应用[J]. 兰州大学学报(医学版), 2022, 48(5): 24-29.
- [10] 吴昊澄, 吴晨, 鲁琴宝, 等. 基于百度指数的诺如病毒感染暴发疫情预警研究[J]. 中国预防医学杂志, 2021, 22(2): 120-124.
- [11] 孙阳, 张落成, 姚士谋. 长三角城市群“空间流”网络结构特征——基于公路运输、火车客运及百度指数的综合分析[J]. 长江流域资源与环境, 2017, 26(9): 1304-1310.
- [12] 程利莎, 王士君, 杨冉. 基于交通与信息流的哈长城市群空间网络结构[J]. 经济地理, 2017, 37(5): 74-80.
- [13] 王硕, 童洁, 刘超, 等. 武汉气候舒适度与游客网络关注度年内变化相关性分析[J]. 亚热带资源与环境学报, 2013, 8(3): 24-28.
- [14] 刘璐, 王一然, 于言, 等. 蓬莱气候舒适度与旅客网络关注度相关研究[J]. 西南师范大学学报(自然科学版), 2018, 43(5): 57-63.
- [15] 马丽君, 孙根年, 杨睿, 等. 城市气候舒适度与游客网络关注度时空相关分析[J]. 地理科学进展, 2011, 30(6): 753-759.
- [16] 周威, 陈朝晖, 李仁鹏, 等. 基于网络关注度的天气与旅游相关性分析——以长沙为例[J]. 气象科技, 2020, 48(4): 607-614.
- [17] 谢媛, 李本乾. 新媒体环境下突发环境事件网络舆情风险信息感知模型[J]. 现代情报, 2023, 43(6): 158-165.

(编辑: 卢冰)