Noting that there have been attempts to make predictions about external real-world events based on internet behaviour/discussion etc., Hannak et al. (2021) ask (in the other direction) whether meteoroligical events and patterns have significant impacts on online behaviour.

The method used to estimate sentiment involves counting co-occurences of tokens with emoticons to construct a sentiment-scored word list.

The dataset study consists of over 1.3 billion tweets. Many of the twitter users have supplied a location string, which the authors convert to latitude and longitude using Google Maps.

Using meteorlogical data (cloud cover, humidity, precipitation, wind speed) from WeatherData, the authors can estimate the weather experienced by the twitter user by assuming the twitter user is tweeting from their self-disclosed location.

The paper sort of trails off from its original objective of exploring patterns of sentiment as influenced by weather. Instead the authors report the accuracy of the Bagged Decision Model in predicting sentiment, without compelling discussion of the meaning of the model's predictions or accuracy thereof.

References

Hannak, A., Anderson, E., Barrett, L. F., Lehmann, S., Mislove, A., & Riedwald, M. (2012). Tweetin' in the Rain: Exploring Societal-scale Effects of Weather. In J. Breslin, N. B. Ellison, J. G. Shanahan, & Z. Tufekci (Eds.), Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media (pp. 479–482). AAAI