



Universiteit
Leiden
The Netherlands

When do tweeters tweet about Science? Exploratory analysis of the Twitter dissemination of scientific publications by weekdays and months

Dudek, J.; Bowman, T.D.; Costas, R.

Citation

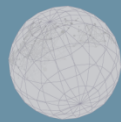
Dudek, J., Bowman, T. D., & Costas, R. (2018). When do tweeters tweet about Science? Exploratory analysis of the Twitter dissemination of scientific publications by weekdays and months. *Sti 2018 Conference Proceedings*, 1566-1578. Retrieved from <https://hdl.handle.net/1887/65294>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/65294>

Note: To cite this publication please use the final published version (if applicable).



STI 2018 Leiden

*23rd International Conference on Science and Technology Indicators
"Science, Technology and Innovation Indicators in Transition"*

STI 2018 Conference Proceedings

Proceedings of the 23rd International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Chair of the Conference

Paul Wouters

Scientific Editors

Rodrigo Costas
Thomas Franssen
Alfredo Yegros-Yegros

Layout

Andrea Reyes Elizondo
Suze van der Luijt-Jansen

The articles of this collection can be accessed at <https://hdl.handle.net/1887/64521>

ISBN: 978-90-9031204-0

© of the text: the authors

© 2018 Centre for Science and Technology Studies (CWTS), Leiden University, The Netherlands



This ARTICLE is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

When do tweeters tweet about Science? Exploratory analysis of the Twitter dissemination of scientific publications by weekdays and months

Jonathan Dudek*, Timothy D. Bowman** and Rodrigo Costas***

*j.dudek@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Wassenaarseweg 62A, Leiden, 2333 AL (The Netherlands)

** timothy.d.bowman@wayne.edu

School of Information Sciences, Wayne State University, 106 Kresge Library, Detroit, MI 48202 (United States)

*** rcostas@cwts.leidenuniv.nl

Centre for Science and Technology Studies (CWTS), Leiden University, Wassenaarseweg 62A, Leiden, 2333 AL (The Netherlands)

DST-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy, Stellenbosch University (South Africa)

Abstract

The potential of Twitter in the dissemination of scientific publications has raised considerable scholarly attention. However, timely patterns of interaction with scientific publications on Twitter have not been covered broadly so far. This study examines the dates of when scientific publications are tweeted, focusing on weekdays and months. For that, data was collected from Altmetric.com, comprising of 25,227,143 tweets spanning the years 2012 to 2016. Next to inquiring general patterns of tweeting dates, subsets of three different Western European countries, several North African states, and South Africa were observed. In general, the highest number of tweets can be observed on Thursdays, with decreasing activity towards the weekend. Concerning months, most tweets can be found in November, with both October and December following close by. January is the month of least activity. Comparing countries, timely variations in tweeting activity were found. The authors conclude that the reception of scholarly publications on Twitter is following certain timely patterns, while geographical as well as cultural backgrounds of tweeters may influence those.

Introduction

Scholarly communication continues to evolve as the online sharing of scientific information flourishes. Scholars, organizations, governments, and the public utilize online platforms to communicate in various ways, including the sharing of, and commenting on, various types of research objects including data, presentations, videos, and journal articles. Scholars can utilize these online contexts in numerous ways including to find out about new publications in their area of interest (filtering), to communicate with their colleagues and the public (networking), and to share their own (or their colleagues') publications (promotion). The speed at which this sharing and communicating occurs introduces many new possibilities for scientists, as traditional measures of tracking the sharing and use of scientific documents (bibliometrics) often took several years.

These online traces of activity are now being recorded and tracked by altmetric data aggregators, such as Altmetric.com, and scholars are studying their potential for measuring various types of scientific and societal impact. Online platforms vary greatly, from short tweets in Twitter, to videos and comments in YouTube, to longer descriptions in blogs and news stories. This heterogeneity presents an interesting problem for social media metrics researchers (Haustein, 2016).

Among all social media platforms currently providing social media metrics, Twitter is one of the most popular ones (Robinson-García, Torres-Salinas, Zahedi, & Costas, 2014). Twitter is a microblogging service that allows users to post short messages (up to 280 characters) to others within the platform. Twitter has approximately 330 million monthly active users, and it is used by scholars across various disciplines (Holmberg & Thelwall, 2014; Bowman, 2015). While there are many interesting aspects of Twitter being investigated within, and outside of, scholarly communication and social media metrics (see Sugimoto, Work, Larivière, & Haustein, 2017), one aspect that hasn't been investigated adequately is temporality. By this we mean the distribution of tweets for scholarly publications over time.

The speed and frequency at which online actions are performed is one of the primary reasons social media metrics are interesting to researchers focusing on scholarly communication and bibliometrics. In addition, the time a user tweets an article is captured very precisely by Twitter, which makes it an interesting and useful factor to investigate. Bruns and Stieglitz (2013) describe the possibilities of investigating tweeting activities not only on the basis of activity indicators but by also including a temporal analysis. The authors suggest that examining account activity on Twitter as well as its respective timely embeddedness may result in a more refined perspective. The current study builds on this idea by contextualizing tweeting activity collected by Altmetric.com according to tweeting dates. Altmetric.com captures tweets containing a DOI or a link to research documents.

Dates of tweeting activity for a scientific document cannot be compared easily to the date this document is made available to the public, as publication dates are not always reported precisely (Haustein, Bowman, & Costas, 2015). Irrespective of the relation to publication patterns, though, the accuracy of tweeting dates provides a reliable way of framing interaction with scientific publications on Twitter. Hence, patterns of timely perception of scientific publications may be revealed by asking:

When do tweeters tweet about scientific publications?

Correspondingly, this study examines the dates of when scientific publications are tweeted, in terms of weekdays and months of the tweeting activity. This was done on the basis of pre-defined criteria, concerning tweets' authors and their geographic location. Proceeding, the results for differing patterns between days of the week and months of the year were examined.

Data and Methods

The data for tweets was obtained from Altmetric.com, updated until early October 2017. This version of Altmetric.com data contains 8.1 million records and 42.9 million tweets. Altmetric.com provides both information about each tweet captured and information about tweeters. A local relational database was created and the Altmetric.com data was stored. The tweet dates were converted in the database and columns were added representing the month, weekday, and year for each tweet. For this work, tweets were considered if they were tweeted from January 2012 through December 2016 (as Altmetric.com began collecting tweets mid-

2011 and this data version ended collection October 2017, which resulted in incomplete data for years 2011 and 2017).

The number of unique tweeter profiles returned from the years 2012 through 2016 was 2,308,976. Altmetric.com also provided the country codes of tweeters as reported by Twitter. These country codes stem from when Twitter users create their profile; note that this is not provided in all cases and may not always be accurate. The percentage of tweeters with country codes as reported by Altmetric.com was 54% (see also Zahedi & Costas, 2017).

For the comparison of tweeting activity according to countries of origin of tweeters, nine countries were chosen from three different time zones. This included the United Kingdom and Morocco (UTC, Coordinated Universal Time), the Netherlands, Spain, Algeria, and Tunisia (CET, Central European Time), and Libya, Egypt and South Africa (EET, Eastern European Time). Here, the focus lay on providing a few exemplary cases with the countries selected located in three neighboring time zones. The latter aspect should eliminate deviating affects due to differing time zones as far as possible. Consequently, North and South American as well as Asian countries were excluded.

In order to observe (assumed) differences in timely patterns, a heterogeneous selection of countries was chosen. This is why – next to three Western European countries (the United Kingdom, the Netherlands, and Spain), a set of Islamic countries was introduced. However, as tweeting activity in these countries is comparatively low, the North African states Morocco, Algeria, Tunisia, Libya and Egypt were combined to one single regional set. As a representative of the southern hemisphere, tweeting activity in South Africa was expected to yield results in line with a differing seasonal cycle.

Concerning tweets observed, the day of the week and month of the year a tweet is sent was considered in order to attain an overall image of the distribution of tweeting activity. Dates of tweets are based on the time-zone selected by Twitter users. Once the overview of tweeting dates was obtained, subsets on the basis of the countries selected were analyzed accordingly.

Results

In order to return a consistent data set, this study focused on tweets referring to publications from 2012 onwards. This stipulation was utilized in order to eliminate tweets that referenced publications older than 2012, mimicking the tweeting activity captured by Altmetric.com. In total, the dataset in this study contained 25,227,143 tweets, which were tweeted from January 2012 through December 2016. Those tweets referred to 3,342,891 distinct scientific publications, which were published from 2012 onwards and were mentioned in the tweets from 2012 through 2016. The overview of tweeting activity in relation to days of the week and months of the year is presented below.

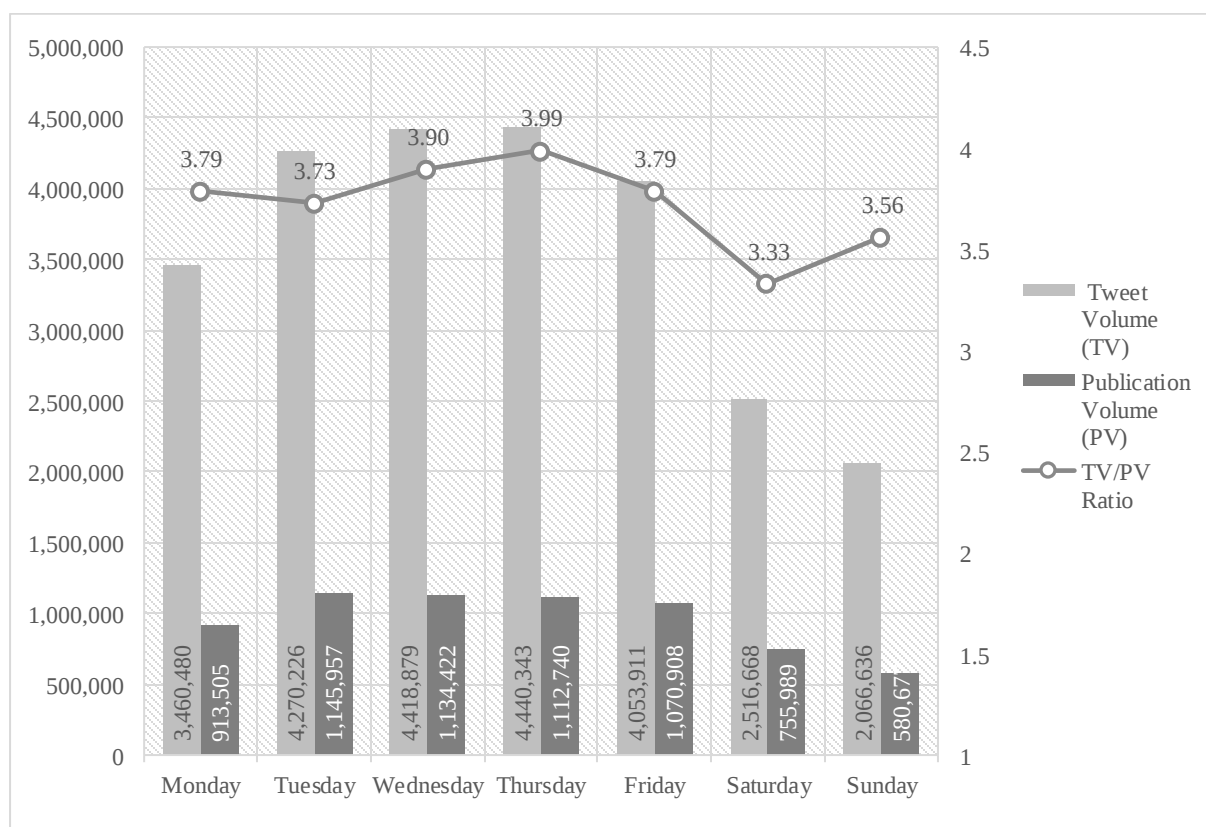
This study regards days of the week or months of the year as separate units in terms of time. Within those units, distinct tweets as well as distinct publications (mentioned by those tweets) are observed. All activity that takes place on a given day is not related to what happens on the following day. As a consequence, tweets occur only once across all days of all months of all years observed. Publications, however, are only distinct within one unit of time, not beyond. Thus, the main focus of this work is on the number of tweets referring to distinct publications on a particular day or in a particular month.

These calculations were completed for the total distribution of tweets, as well as a selection of countries of tweeters. In addition, graphs representing both the single days of the week and months of the year are presented below. Finally, the *tweet volume* (TV)—tweets mentioning scientific publications in a specific time unit—and the *publication volume* (PV)—volume of distinct publications tweeted in a specific time unit—were calculated. Considering these two measures, the TV/PV ratio was calculated in order to measure the *tweeting intensity* in a given time unit. For example, if there are many tweets for only a small number of unique publications, tweeting intensity is high while PV is relatively low. Yet if there are a small number of tweets and a larger number of unique publications, then *tweeting intensity* is low while PV is still high. For example, when examining the TV/PV ratio in Figure 1 for Thursday, one can see that there are on average 3.99 tweets per publication. This is “high” as compared to the value from Wednesday (3.90), although in the latter case there have been more publications tweeted in that day.

Distribution of tweets over days of the week

As shown in Figure 1, the highest number of total tweets (17.60% of the total) could be observed on Thursdays, with a general pattern increasing from the beginning of the week towards this day. Activity then drops, with Friday (16.07%), Saturday (9.98%), and finally Sunday (8.19%) showing a lower tweeting activity. The share of distinct publications tweeted (of the overall sum of publications mentioned) mostly followed this same pattern over the course of the week. Concerning the TV/PV ratio, the graph demonstrates slight changes in the amount of publications covered by tweets on a given day. For example, although Tuesday is the day when most distinct papers have been tweeted (1,145,957; 34% of distinct publications in the data set), the number of tweets per paper is lower on this day when compared to other weekdays.

Figure 1. Total TV, PV, and TV/PV ratio over weekdays from 2012 through 2016 (n=25,227,143 tweets).



Equivalent to the overview of the total distribution of tweets and publications tweeted over the course of the week, subsequent figures present results for the different countries. Those include the United Kingdom (Figure 2), the Netherlands (Figure 3), Spain (Figure 4), a selection of North African countries as defined above (Figure 5), and finally, South Africa (Figure 6). Tweets by tweeters of each of these countries are subsets of the total tweeters and tweets observed.

Figure 2. United Kingdom, tweets over weekdays (n=3,021,387 tweets).

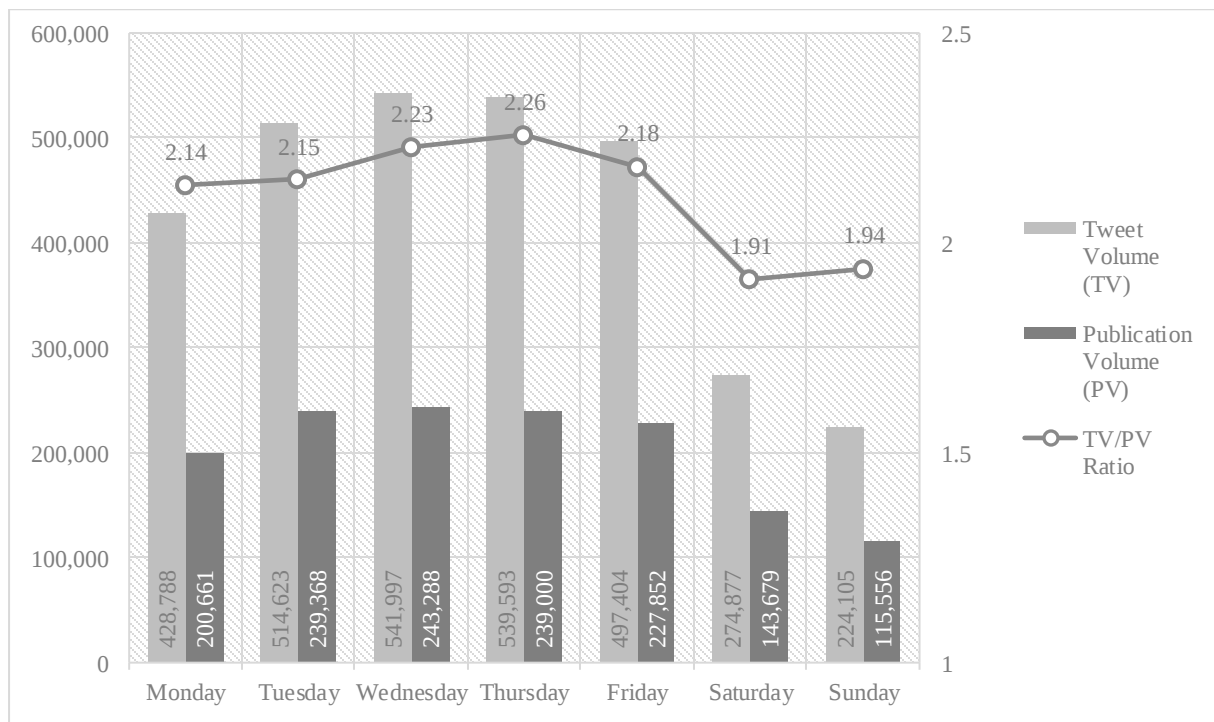


Figure 3. Netherlands, tweets over weekdays (n=245,300 tweets).

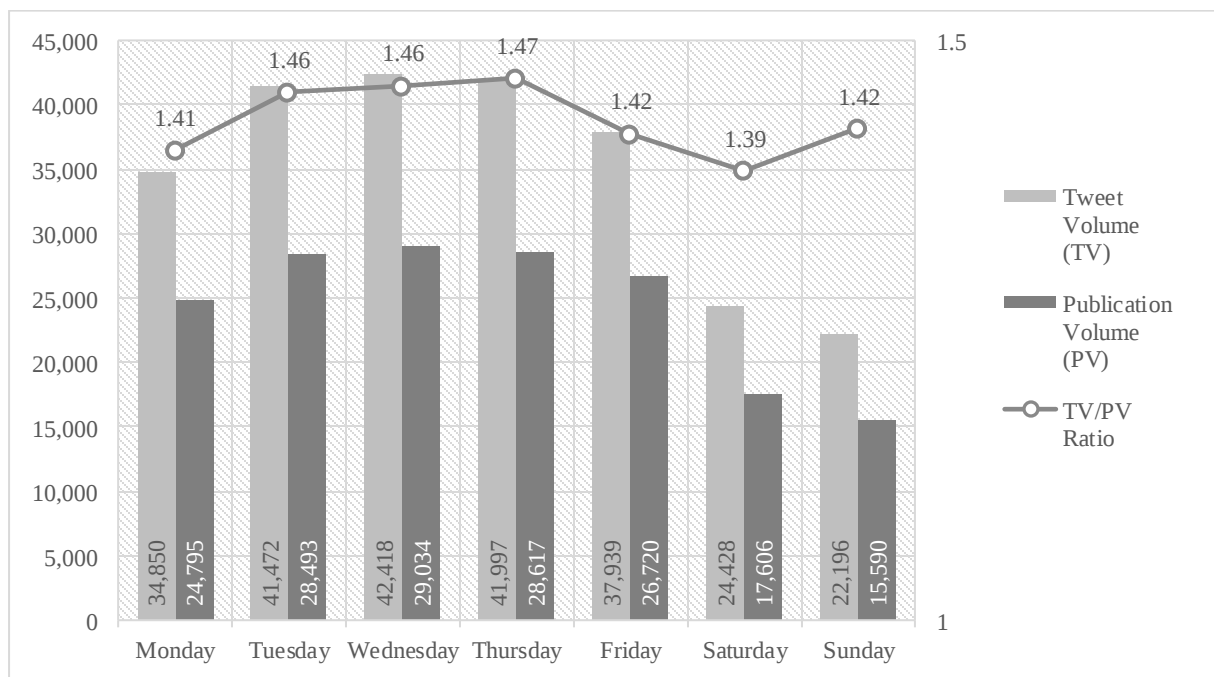


Figure 4. Spain, tweets over weekdays (n=725,233 tweets).

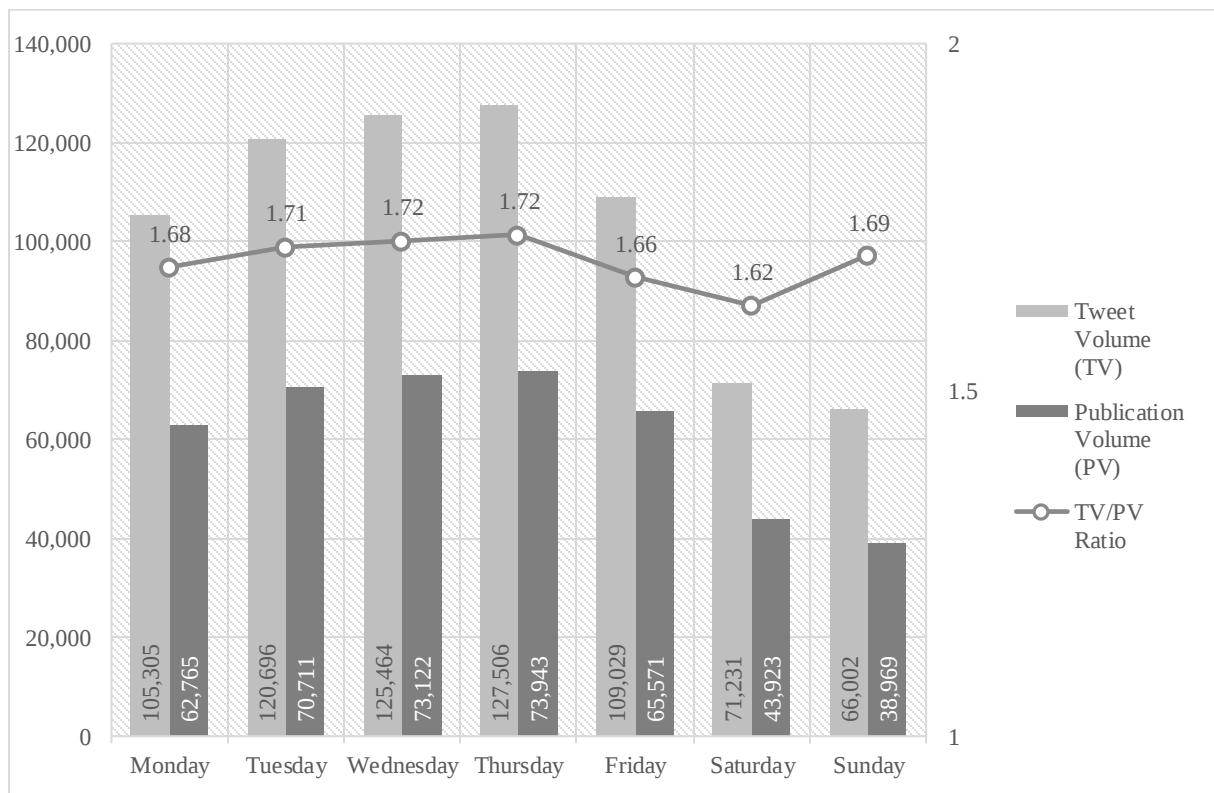


Figure 5. North African states, tweets over weekdays (n=43,671 tweets).

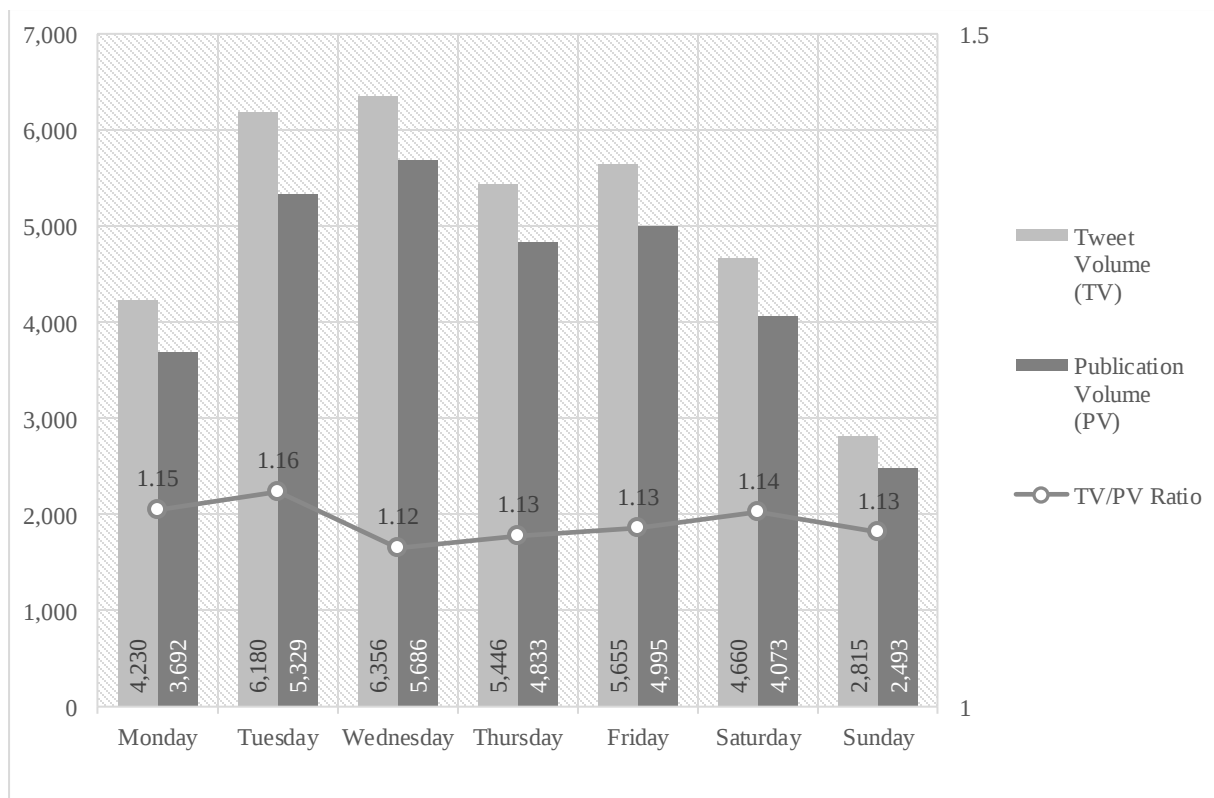
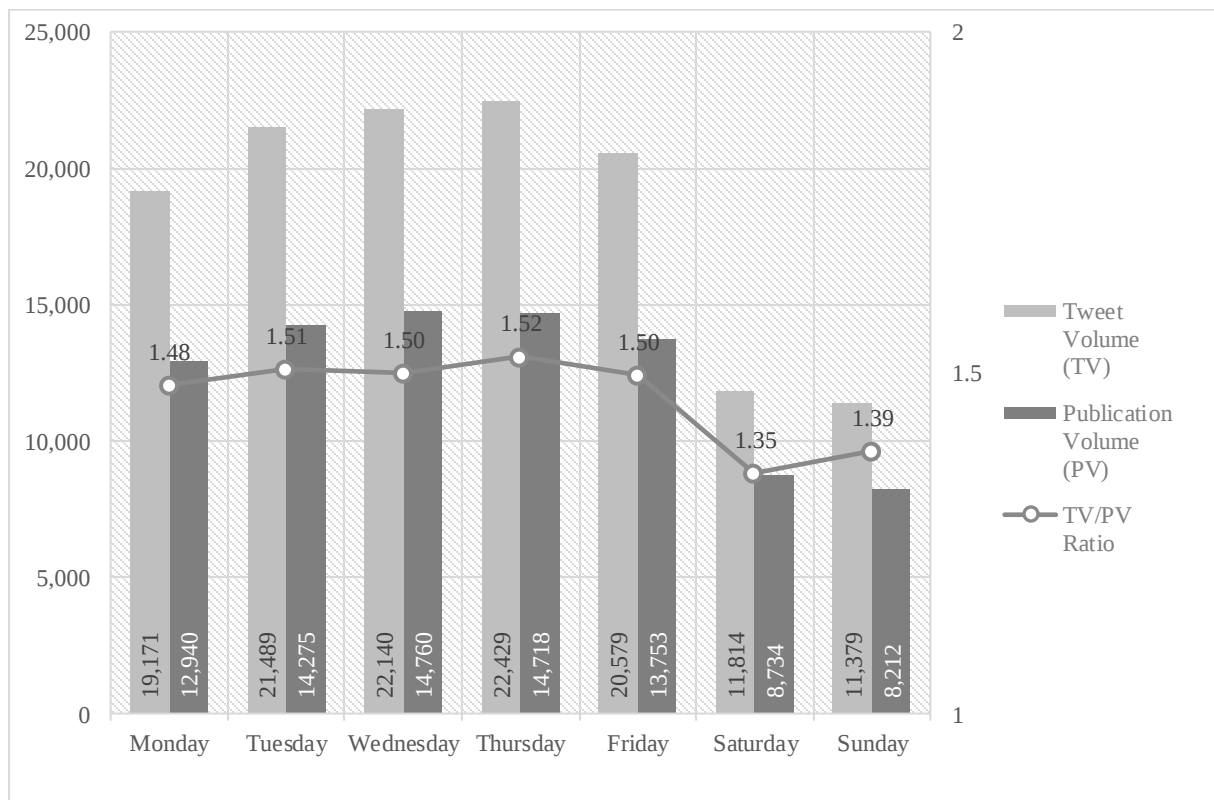


Figure 6. South Africa, tweets over weekdays (n=129,001 tweets).



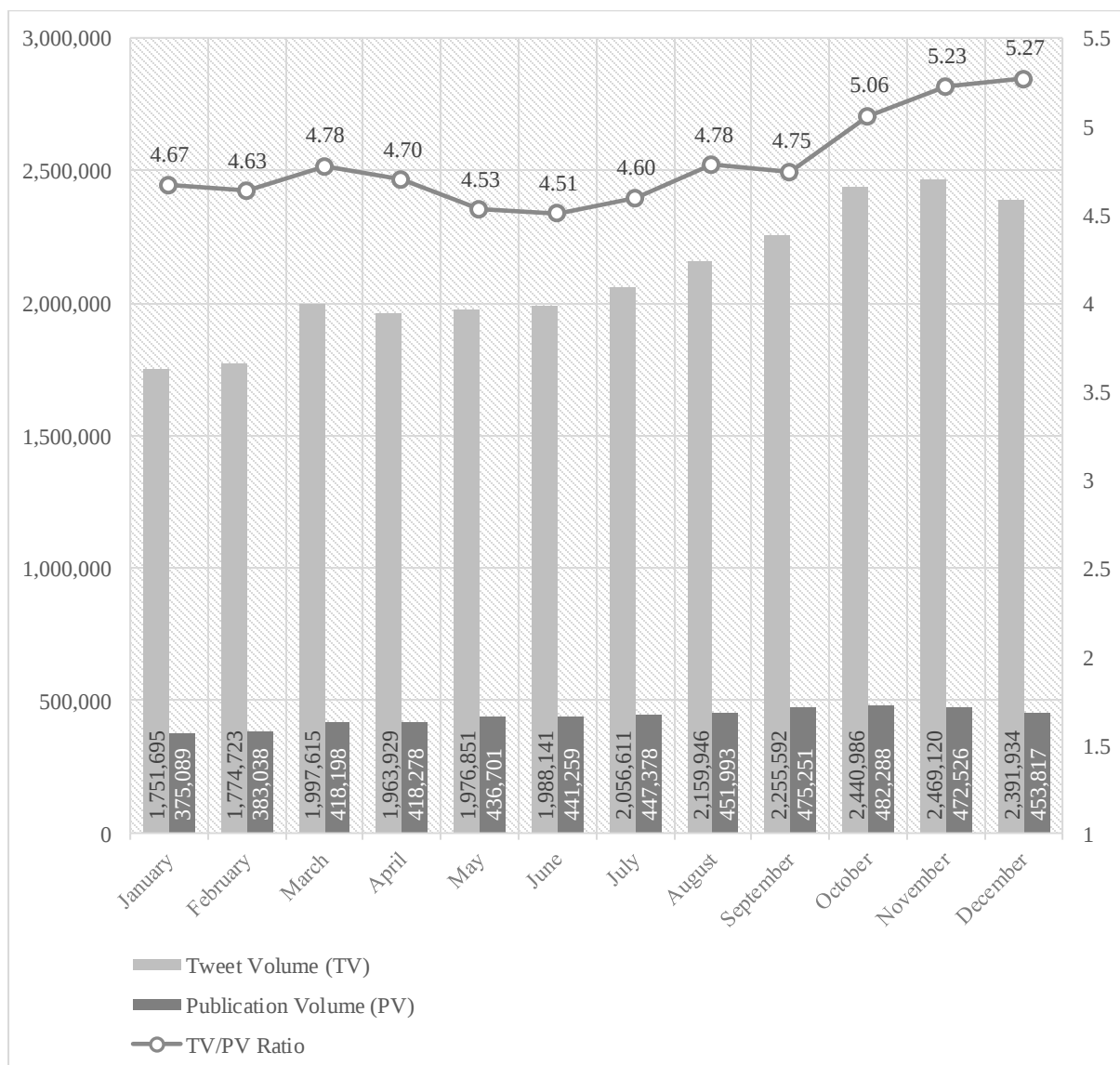
Overall, tweeting activity from Western European countries (the United Kingdom, the Netherlands, and Spain) as well as South Africa demonstrates a similar trend to the overall tweeting activity of the dataset, with tweeting activity increasing in the beginning of the week (Monday, Tuesday, and Wednesday), and a decrease following, typically from Thursday onwards towards the weekend. Weekends are also days with the lowest intensity of tweets per publication, thus meaning fewer papers and fewer tweets, but also a smaller share of tweets per publication.

North African countries, however, exhibit a different pattern, with Sundays and Mondays being the lowest days in terms of TV and PV, and Mondays and Tuesdays being the days of higher intensity of tweeting publications. This different pattern can be related to the different structuration of the working week in these countries, with weekends on Fridays and Saturdays (Workweek and weekend, n.d.).

Distribution for months of the year

Equivalent to days of the week, this study took into account the distribution of tweets over months of the year. Concerning the total of tweets over months, most tweets were sent in November (9.79%), closely followed by October and December (9.68% and 9.48% respectively). The least activity can be observed for January (6.94%), then constantly increasing towards the end of the year. (See Figure 7.)

Figure 7. Total number of tweets over months of the year (n=25,227,143 tweets).



Similar to the days of the week, the distribution of tweets over the course of the year is presented for the various countries observed. The underlying set of tweets remained the same, therefore, numbers of tweets were the same as in the cases above. Included are the United Kingdom (Figure 8.), the Netherlands (Figure 9.), Spain (Figure 10.), the predefined selection of North African countries (Figure 11.), and South Africa (Figure 12.).

Figure 8. United Kingdom, tweets over months of the year (n=3,021,387 tweets).

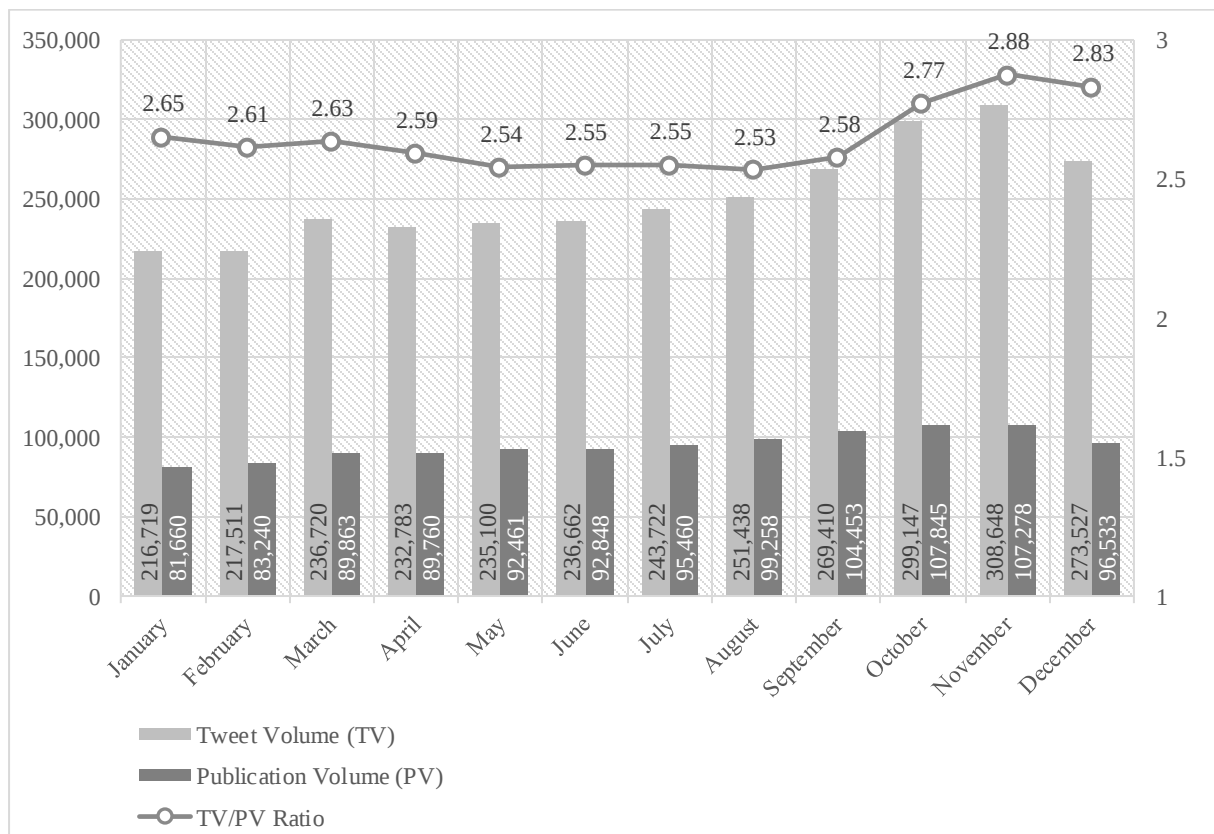


Figure 9. Netherlands, tweets over months of the year (n=245,300 tweets).

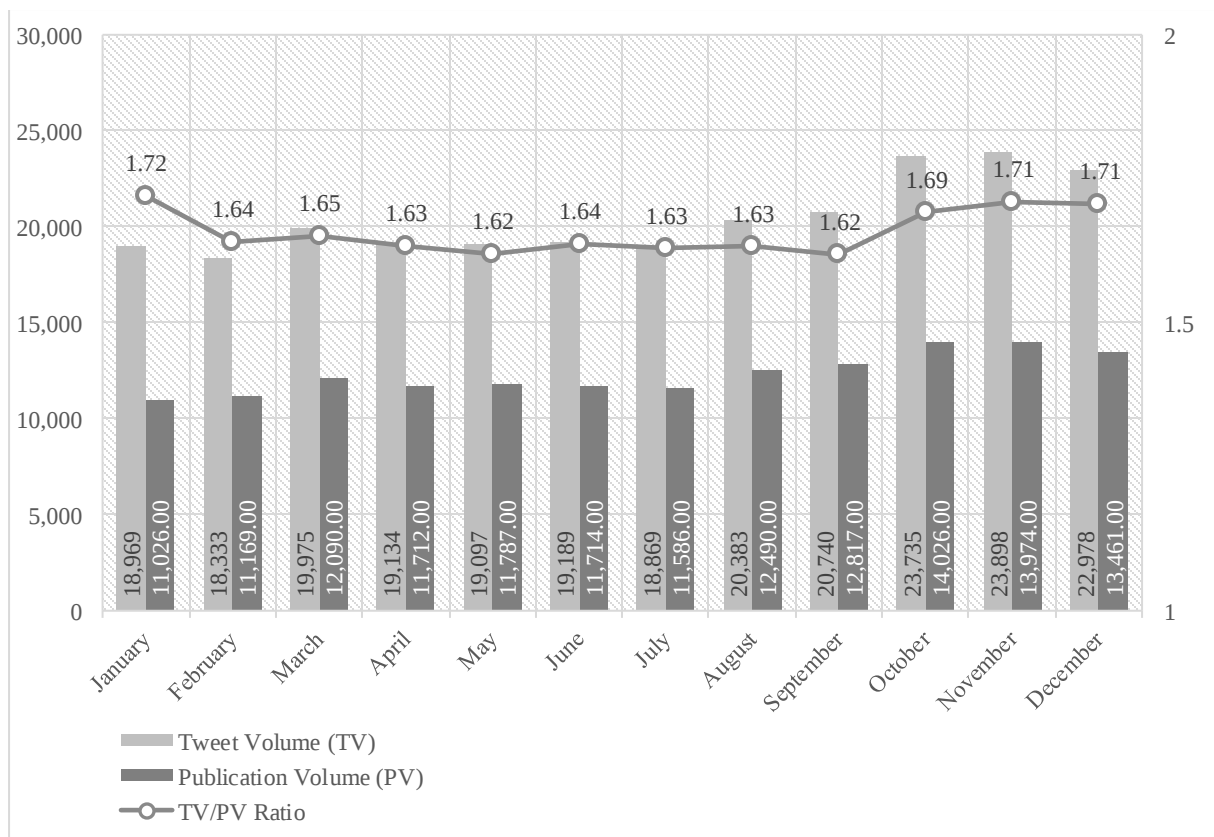


Figure 10. Spain, tweets over months of the year (n=725,233 tweets).

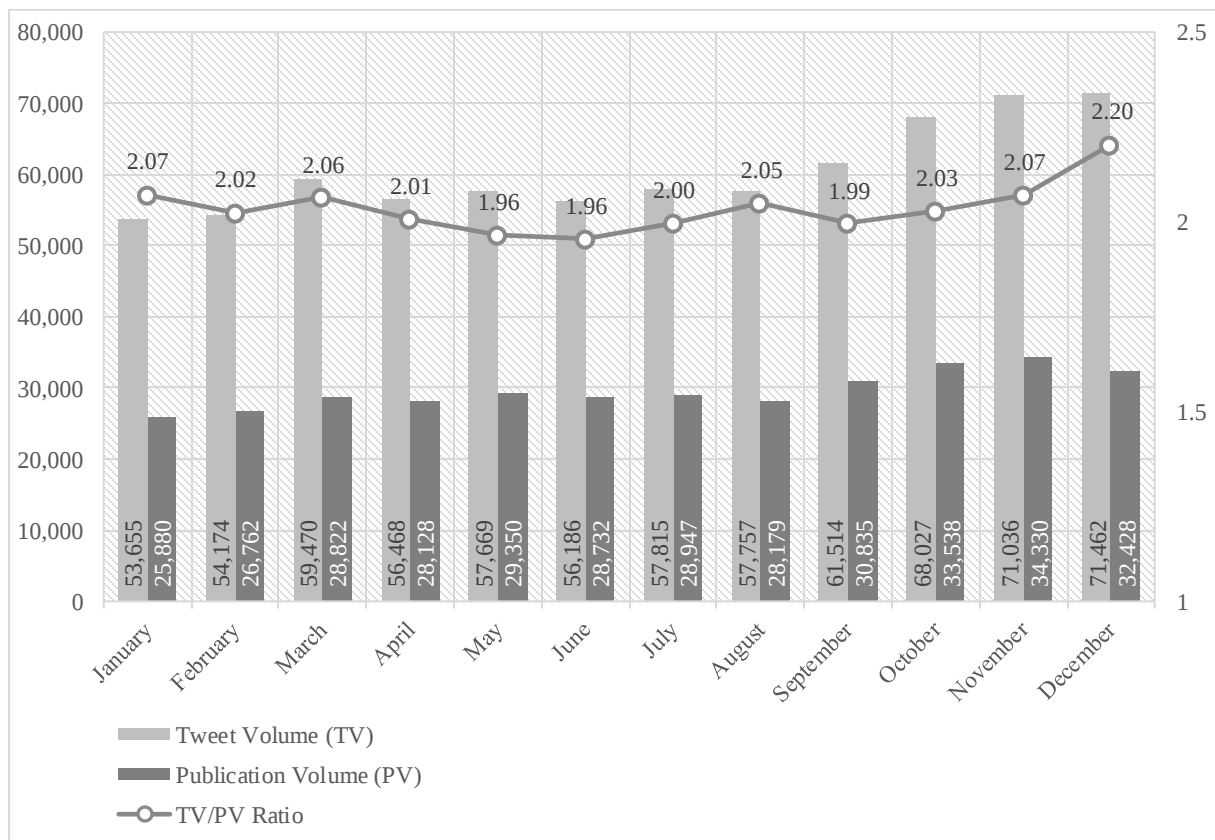


Figure 11. North African countries, tweets over months of the year (n=43,671 tweets).

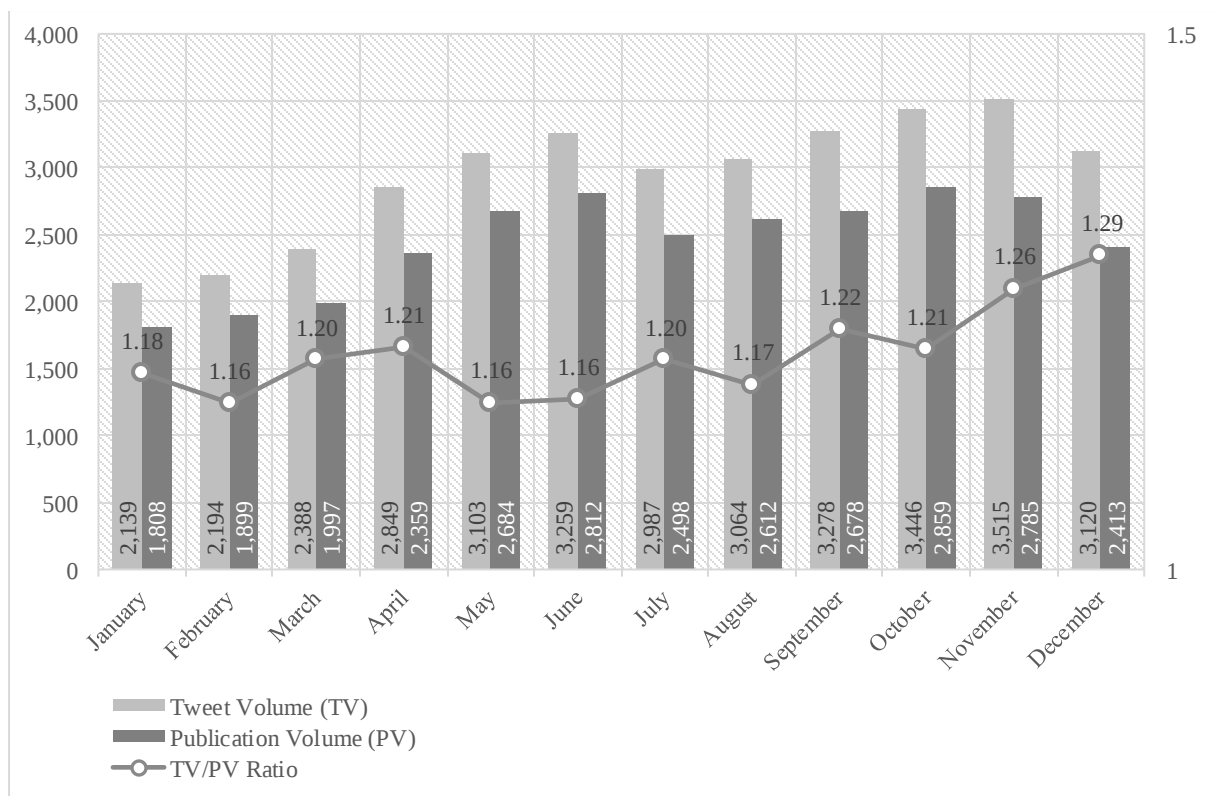
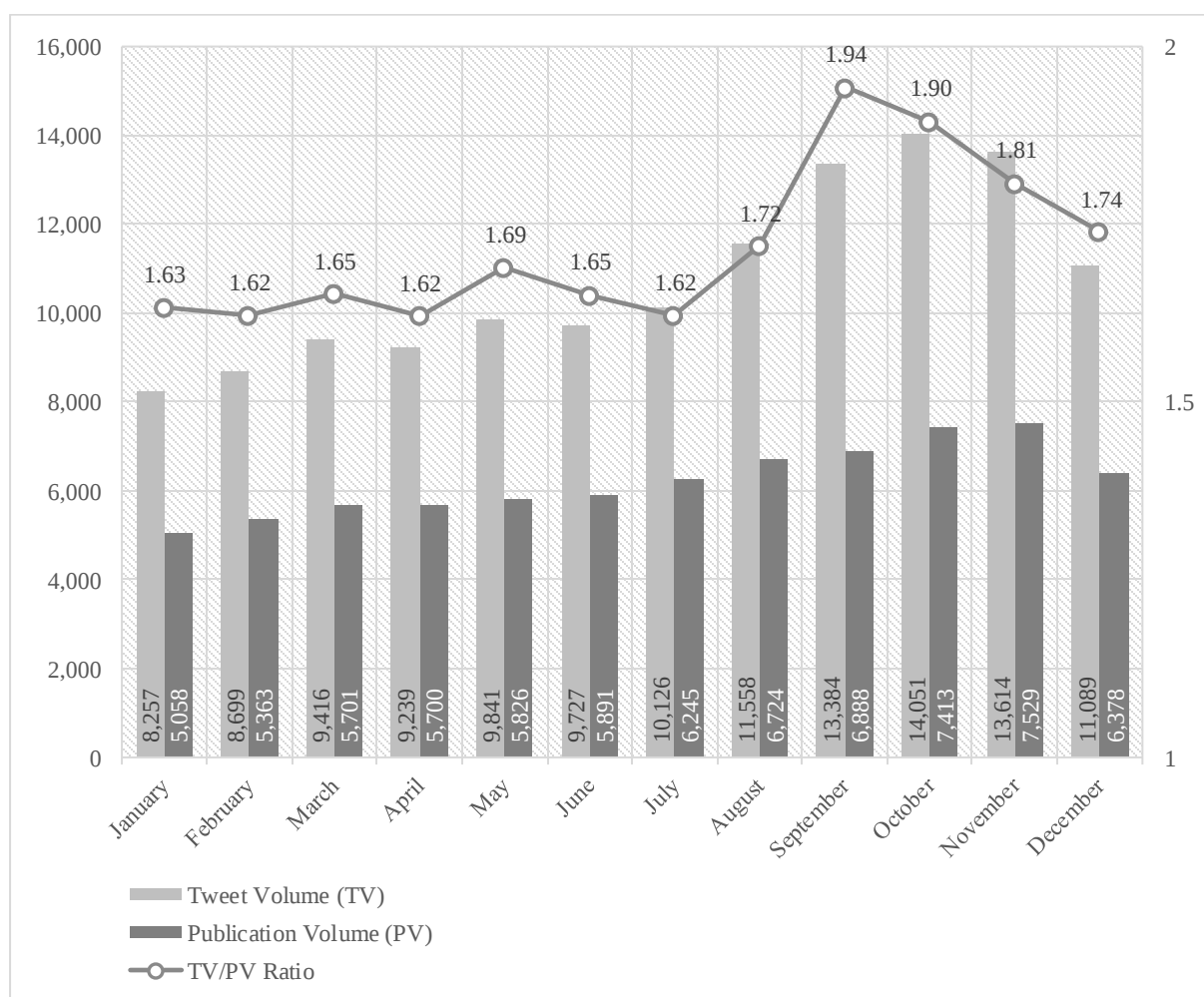


Figure 12. South Africa, tweets over months of the year (n=129,001 tweets).



Overall, tweeting activity in all countries, except of Spain and South Africa, follows a trend similar to the overall tweeting activity of the dataset, with higher numbers demonstrated in September, October, and November, and the lowest activity in January and February.

South Africa demonstrates higher tweet volume in August and September, while Spain demonstrated the most tweeting activity in December, which varied from all other countries observed. With regards to publication volume, North African countries demonstrated a high variety of publications in June, which is unlike any other country observed.

Discussion and outlook

This study presents the first large-scale analysis of the distribution of tweets to scientific publications over weekdays and months. Results point to a general “western” pattern of reception, with higher levels of activity (both in terms of volume and intensity) during the middle workdays (i.e. Wednesday and Thursday) and lower activity during the weekends, particularly on Sundays. In terms of tweeting activity over months, tweets to scholarly publications seem to be larger during the last months of the year (e.g. particularly October, November and December), although there are differences for South Africa (with larger activity during the months of August, September, October and November) and for the North African countries, thus pointing to geographical (e.g. southern/northern hemisphere differences) as well

as cultural differences (e.g., different workdays in Islamic countries and distribution of months (Workweek and weekend, n.d.)).

These results highlight the relevance of considering time dynamics of activity when studying social media metrics stemming from sources with a strong cultural or geographical component, as in the case of Twitter. Thus, in contrast to citations that are more “universal” and less dependent on very precise time measures (typically citation indicators are calculated at the year-level), Twitter metrics of publication mentions may be more strongly influenced by the day of the week or the month in which they occur. This suggests that the exact moment a publication is made available (i.e., in terms of day of the week/month combination) (cf. Haustein et al, 2015) or when it is tweeted for the first time, could be seen as critical elements with a much stronger impact on how the publication would eventually be received on Twitter (or other social media platforms) than initially expected. This supports the idea that studies on the temporality of social media metrics may need to focus on unveiling how cultural and geographical factors (e.g. southern hemisphere or northern hemisphere activity, as well as different week dynamics across countries) may play a role in the dissemination and reception of scholarly outputs on social media.

Limitations

Finally, there were some limitations with the data used in this study. As mentioned above, the Twitter user profile geographic code is not mandatory or always accurate, thus all users do not have geographic codes and there is no guarantee that all codes are accurate. In addition, there were unpredictable results from the North African states observed, which was potentially due to the small number of tweets collected, as well as variation in work days found in the countries. Finally, the publication dates range from 2012 to approximately mid-2017, whereas tweets span from 2012 through 2016. However, it is difficult to imply a direct connection between tweeting dates and publication dates.

Outlook

This explorative study supports the idea that there are still important elements to understand and explore further. One must consider the timeliness and geographic effects when examining social media metrics. Moreover, there is a strong need for understanding tweeting habits across different countries, regions, and time-zones. Such studies would have to validate results found here by contrasting them to patterns of general social media usage in different countries or regions. Further research should focus on developing a more comprehensive understanding of what tweeting activity—in terms of temporality—actually comprehends and how different time-related elements (days of the week, months, etc.) may influence the social media metrics accrued by scientific publications and their overall visibility on social media.

References

- Bowman, T. D. (2015). Differences in personal and professional tweets of scholars. *Aslib Journal of Information Management*, 67(3), 356- 371. doi: 10.1108/AJIM-12-2014-0180
- Bruns, A. & Stieglitz, S. (2013). Towards more systematic Twitter analysis: Metrics for tweeting activities. *International Journal of Social Research Methodology*, 16(2), pp. 91-108. doi: 10.1080/13645579.2012.756095
- Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics*, 108, 412-423. doi: 10.1007/s11192-016-1910-9

- Haustein, S., Bowman, T. D., & Costas, R. (2015). When is an article actually published? An analysis of online availability, publication, and indexation dates. Paper presented at *Proceedings of the 15th International Society of Scientometrics and Informetrics Conference*, Istanbul, Turkey. Available at: <https://arxiv.org/abs/1505.00796>
- Holmberg, K., & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, 101(2), 1027- 1042. doi: 10.1007/s11192-014-1229-3
- Robinson-García, N., Torres-Salinas, D., Zahedi, Z., & Costas, R. (2014). New data, new possibilities: Exploring the insides of Altmeter.com. *Profession de la Información*, 23(4), 359-366. doi: 10.3145/epi.2014.jul.03
- Sugimoto, C. R., Work, S., Larivière, V. and Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, 68: 2037-2062. doi:10.1002/asi.23833
- Workweek and weekend (n.d.). In *Wikipedia*. Retrieved April 15, 2018, from https://en.wikipedia.org/wiki/Workweek_and_weekend
- Zahedi, Z., Costas, R. (2017). How visible are the research of different countries on WoS and Twitter? an analysis of global vs. local reach of WoS publications on Twitter. Paper presented at *16th International Conference on Scientometrics & Informetrics (ISSI)*, Wuhan, China. doi: 10.6084/m9.figshare.5481283.v1