



OPEN

## A calibrated measure to compare fluctuations of different entities across timescales

Jan Chołoniowski<sup>1</sup>, Julian Sienkiewicz<sup>1</sup>, Naum Dretnik<sup>2</sup>, Gregor Leban<sup>3</sup>, Mike Thelwall<sup>4</sup> & Janusz A. Hołyst<sup>1,5</sup>✉

A common way to learn about a system's properties is to analyze temporal fluctuations in associated variables. However, conclusions based on fluctuations from a single entity can be misleading when used without proper reference to other comparable entities or when examined only on one timescale. Here we introduce a method that uses predictions from a fluctuation scaling law as a benchmark for the observed standard deviations. Differences from the benchmark (residuals) are aggregated across multiple timescales using Principal Component Analysis to reduce data dimensionality. The first component score is a calibrated measure of fluctuations—the *reactivity RA* of a given entity. We apply our method to activity records from the media industry using data from the Event Registry news aggregator—over 32M articles on selected topics published by over 8000 news outlets. Our approach distinguishes between different news outlet reporting styles: high reactivity points to activity fluctuations larger than expected, reflecting a bursty reporting style, whereas low reactivity suggests a relatively stable reporting style. Combining our method with the political bias detector Media Bias/Fact Check we quantify the relative reporting styles for different topics of mainly US media sources grouped by political orientation. The results suggest that news outlets with a liberal bias tended to be the least reactive while conservative news outlets were the most reactive.

Fluctuations occur everywhere and are frequently sources of useful knowledge about intrinsic properties of systems, with diverse applications. When thermal activity<sup>1</sup> or heart rate variability are measured<sup>2</sup>, they can give meaningful information if system-specific features are taken into account, such as the fluctuation–dissipation theorem<sup>3</sup>, or norms for heart rate variability in various groups of people<sup>2,4</sup>. In financial engineering, forecasting the volatility of stock prices<sup>5,6</sup> is central to hedging strategies for stock portfolios<sup>7</sup>. There are multiple ways to quantify fluctuation levels, including the variance (or standard deviation), Fano factor<sup>8</sup>, and coefficient of variation<sup>9</sup>. Nevertheless, studying just absolute (variance) or relative (Fano factor) fluctuations can give misleading conclusions when the system's size is ignored. For example, the variance  $\sigma_E^2$  of energy thermal fluctuations should be proportional to the number of particles in a system<sup>3</sup> while for many non-thermal objects the variance in an ensemble of similar units (e.g. crop volumes in groups of fields) scales with the system size in agreement with Taylor's law<sup>10</sup>. For time-dependent systems the rescaled range of an observable can depend on the length of the observation window  $\Delta$  according to Hurst's law<sup>11</sup>. In many cases time series fluctuation patterns can change with amplitude levels and in these systems generalized Hurst exponents should be calculated with Multifractal Detrended Fluctuations Analysis<sup>6,12</sup>.

In this study we introduce the notion of *reactivity (RA)*, which is related to the residuals of the temporal fluctuation scaling law (or the temporal Taylor power law; TFS)<sup>13,14</sup> that links the mean and the variance of a dynamical process for each observed entity through a power law. In this context, by a residual we mean the measured standard deviation of an observed system variable (henceforth called *activity*) calibrated against values expected from the TFS at a given timescale. The TFS is ubiquitous in complex systems<sup>15</sup>. Analyses of the fitted power law exponent have been applied e.g. to measure the lexical richness of texts in the presence of topical variations<sup>16</sup>, to characterize UK crime<sup>17</sup> and human behavior under high pressure<sup>18</sup>, and to detect strong emotions in physiological timeseries<sup>19</sup>.

<sup>1</sup>Center of Excellence for Complex Systems Research, Faculty of Physics, Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland. <sup>2</sup>Slovenian Press Agency, Tivolska 48, 1000 Ljubljana, Slovenia. <sup>3</sup>Artificial Intelligence Laboratory, Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia. <sup>4</sup>School of Mathematics and Computing, University of Wolverhampton, Wulfruna Street, Wolverhampton WV1 1LY, UK. <sup>5</sup>ITMO University, 49 Kronverkskiy av., Saint Petersburg, Russia 197101. ✉email: janusz.holyst@pw.edu.pl

Differences between entities in the above-mentioned datasets may cause them not to follow the TFS strictly. Although fluctuation scaling residuals cluster in geographical and financial data<sup>15</sup>, properties of residuals seem to have been rarely studied. One somewhat similar application of the TFS is date abnormality quantification by measuring the fluctuation scaling of adjective usage in Japanese blogs<sup>22</sup> but this focused on detecting special dates rather than quantifying activity fluctuations.

The concept of reactivity will be applied to data from the Event Registry news aggregator<sup>23</sup>, containing over 32M articles on 39 topics (we will also call them *keywords* or *concepts*, interchangeably) published by over 8000 news outlets (also *sources*, *publishers*). Understanding media activity is a challenge because of the need to take into account differences in size, scope and political bias, as well as differences in reporting styles by topic and multiple timescales. Due to the complex nature of online news media (many interacting units participating in dynamic information exchanges), it is unsurprising that the activity of news outlets, measured by the number of articles published over time, follows the TFS<sup>24</sup>.

In the paper, we study fluctuations in news outlet activities  $f_{s,\Delta}^{(t)}(k)$ , focusing on multiple topics  $k$  (e.g. *Climate change* or *European Union*) and timescales  $\Delta$  (from a few minutes to a few months). We found that the TFS residual value  $R_{s,\Delta}(k)$  for a given news outlet  $s$  is more characteristic of a timescale  $\Delta$  than of a topic  $k$ . To aggregate properties of fluctuations at different timescales, for each concept  $k$  we considered the 14-dimensional space corresponding to corrections to TFS for  $\Delta \in \{1 \text{ min}, \dots, 60 \text{ days}\}$  and performed Principal Component Analysis to reduce the dimensionality of the data set. The first component score is our calibrated measure of fluctuations: the *reactivity*  $RA_s(k)$  of publisher  $s$  when reporting a topic  $k$ , see Eq. (3). A negative reactivity means smaller-than-typical fluctuations, and a positive reactivity means larger-than-typical fluctuations. We found that the median reactivity of news outlets from a given country was high if significant happenings related to a topic had taken place during the period analyzed. Additionally, we considered news outlet reactivity by political bias provided by [mediabiasfactcheck.com](https://mediabiasfactcheck.com). Analyses showed that news outlets with a liberal bias were, on average, less reactive while conservative news outlets were more reactive.

We use a pipeline of multiple theoretical approaches. The simplified scheme in Fig. 1 summarizes the fundamental concepts of this study for later reference.

## Results

Our reactivity method is described with news publishing data from the Event Registry (ER) dataset<sup>23</sup>. We extracted all articles published in 2018 mentioning one of  $\mathcal{K}$  topics ( $K = |\mathcal{K}| = 39$ , for details see Dataset info subsection of Methods). We divided the news articles into sets corresponding to topics  $\mathcal{K}$  and publishers  $\mathcal{S}$  ( $|\mathcal{S}| = 8155$ ), obtaining a list of article publication dates and times  $f_s(k) = \{\tau_{s,1}(k), \tau_{s,2}(k), \dots\}$  for each keyword  $k$  and publisher  $s$ .

The activity timeseries  $f_{s,\Delta}^{(t)}(k)$  were obtained by aggregating a list of article publication occurrences with timestamps using  $D = 14$  time window sizes  $\Delta \in \{1 \text{ min}, \dots, 60 \text{ days}\}$  for each of the 39 keywords separately (see Extracting timeseries from lists of articles subsection of Methods: Statistical Methods). Figure 1 visualizes the whole process for 1 keyword, 2 publishers and 3 time windows. This is shown in the left column. Next, the timeseries were processed using the TFS procedure (see Temporal fluctuation scaling subsection of Methods: Statistical Methods) to obtain the scaling exponent  $\alpha_\Delta(k)$  and the multiplicative factor  $B_\Delta(k)$ . An example TFS plot is in Fig. 2, corresponding to the middle column of Fig. 1. We then calculated TFS residuals  $R_{s,\Delta}(k)$  for each keyword  $k$  and each relevant combination of publisher  $s$  and time window size  $\Delta$  as a logarithm of a ratio of observed source's activity standard deviation  $\sigma_{s,\Delta}(k)$  and an expected source's activity standard deviation  $\tilde{\sigma}_{s,\Delta}(k)$ . The last value results from the mean activity  $\mu_{s,\Delta}(k) = \langle f_{s,\Delta}^{(t)} \rangle$  and the TFS law, i.e.

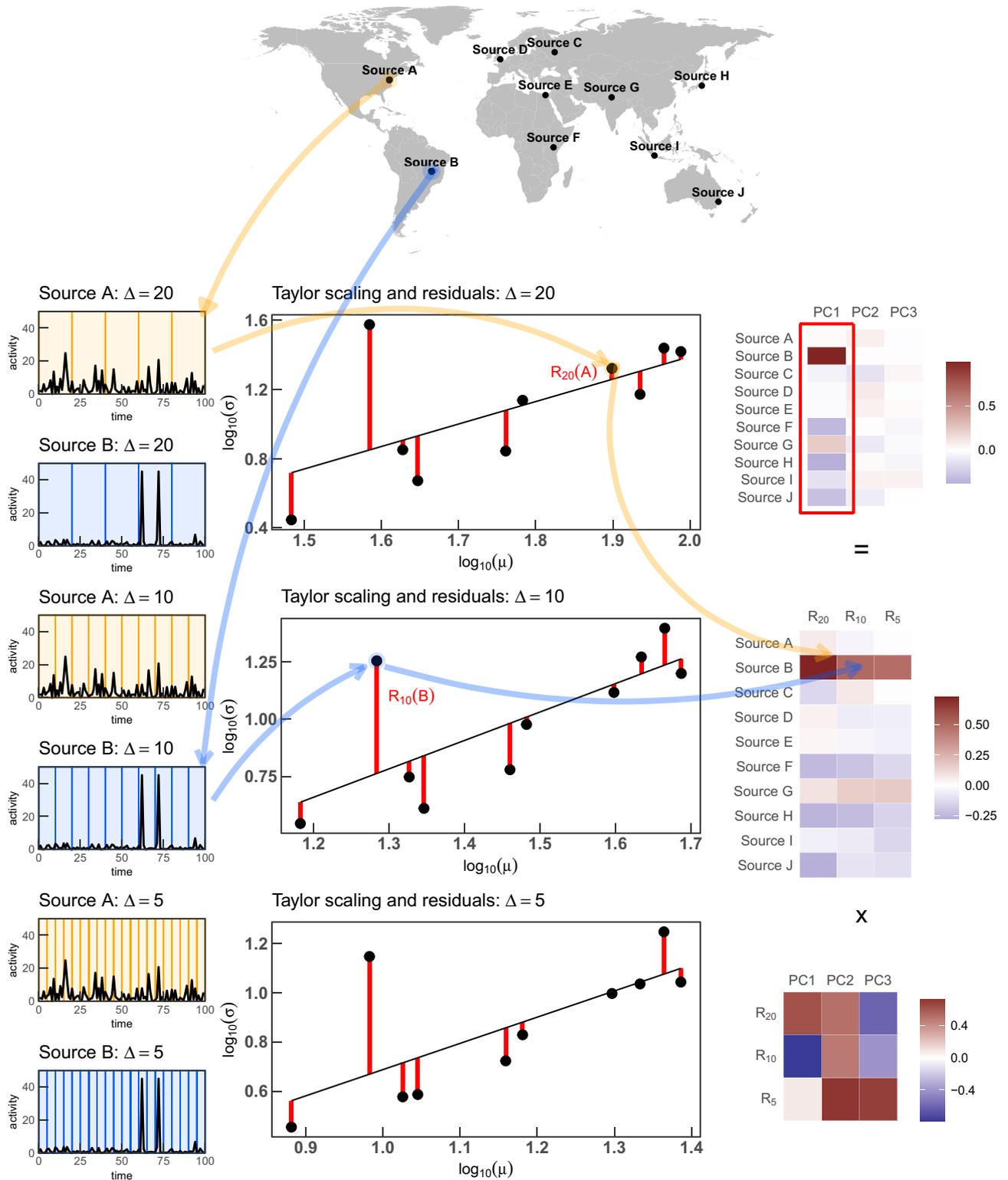
$$R_{s,\Delta} = \log_{10} \left[ \frac{\sigma_{s,\Delta}}{\tilde{\sigma}_{s,\Delta}} \right], \text{ where } \tilde{\sigma}_{s,\Delta} = B_\Delta \langle f_{s,\Delta}^{(t)} \rangle^{\alpha_\Delta} \quad (1)$$

For each concept  $k$  we then have a  $S \times D$ -dimensional matrix  $\hat{\mathbf{R}}(k)$  containing TFS residuals  $R_{s,\Delta}(k)$ , i.e., logarithmic differences between observed activity fluctuations of a publisher  $s$  in a given timescale  $\Delta$  and corresponding predictions resulting from the TFS scaling law that contains information about the magnitudes of fluctuations of other publishers.

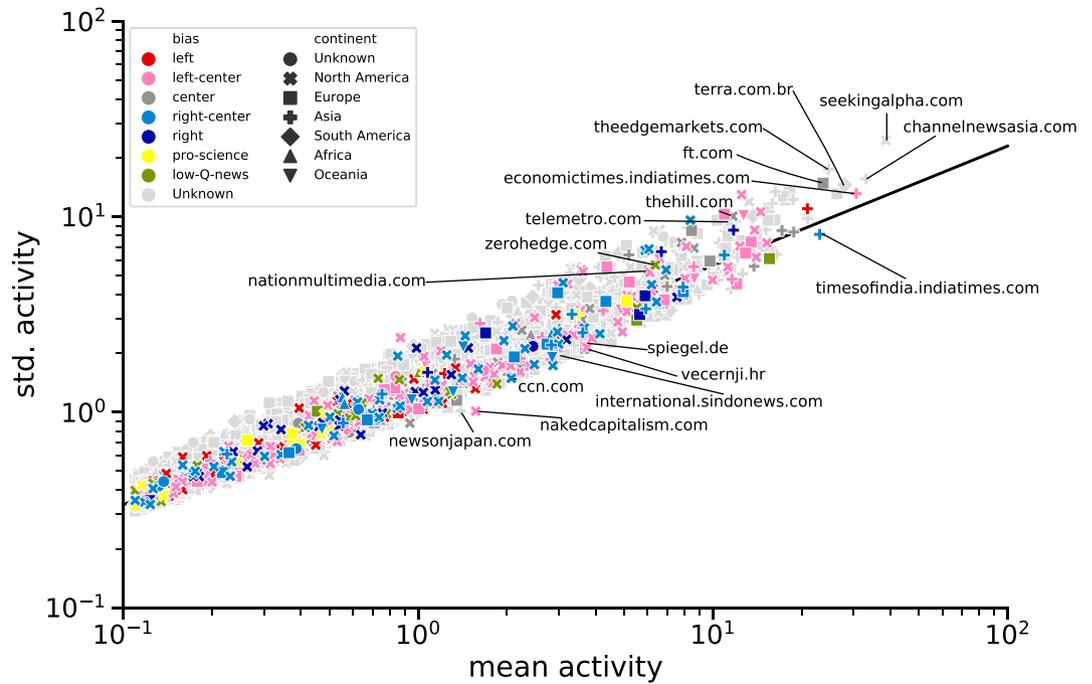
The residuals  $R_{s,\Delta}(k)$  are shown as red segments in the middle column of Fig. 1 and corresponding cells in matrix  $\hat{\mathbf{R}}$  in the middle panel of the right column in the same plot.

**Correlations of residuals across timescales and concepts.** Let us now investigate whether the TFS residuals are similar for a given publisher across all concepts  $k$  and time window sizes  $\Delta$ . For this, we calculate the Pearson correlation between residual values obtained for every pair of  $(k, \Delta)$  combinations, a total of  $N = K \times D = 546$  combinations. The aggregation was performed over all sources publishing on both topics—see the Agglomerative Clustering of correlation matrices subsection of Methods for details.

Figure 3 gives a  $546 \times 546$  matrix of correlations  $\rho(k_1, \Delta_1; k_2, \Delta_2)$  between all possible combinations of  $(k, \Delta)$ . Each row/column stands for one concept-time window size combination. The columns/rows were clustered using Agglomerative Clustering and their order is from the leaves of the resulting dendrogram. The most striking observation is that around 92% of the elements of this matrix are positive valued and the mean value of off-diagonal elements is 0.26. Moreover, the residuals separate roughly into three groups connected to different timescales— $\Delta \leq 1 \text{ h}$  (*short*),  $1 \text{ h} < \Delta \leq 1 \text{ day}$  (*medium*),  $\Delta > 1 \text{ day}$  (*long*). Moreover, inside the groups residuals for a keyword also tend to be adjacent—to see this, compare the sizes of the color clusters and the size of the thinnest color marker. To further quantify this observation, we calculated the means of in-group and between-group correlation matrix elements for the groups (see Supplementary Information: Aggregated correlation matrices).



**Figure 1.** A scheme and a pipeline of methods used in this study. In this example we use 10 hypothetical **sources** (i.e., news outlets): (A)–(J) (top row on the plot) all writing about a given **concept**  $k$  (e.g., “European Union”); each is described with one timeseries reflecting their activity  $f_{s,\Delta}^{(t)}(k)$ , i.e., number of published articles over time (examples shown in the left column for sources A and B). Each series can be divided into **windows of size  $\Delta$**  (here  $\Delta_1 = 20, \Delta_2 = 10, \Delta_3 = 5$ —see left column). In each window the number of articles is summed and then the **mean  $\mu$**  and **standard deviation  $\sigma$**  of these sums are calculated. These pairs for each source and window size are shown on the plot in the middle column and the fit to them in a log-log scale (solid line) is the **temporal fluctuation scaling (TFS, Eq. (9))**. Differences between  $\sigma$  and TFS, given by Eq. (1) are **residuals** and are gathered in a matrix  $\mathbf{R}$  (middle plot in the right column) which is then an input for PCA. As an outcome of PCA we obtain a matrix  $\hat{\mathbf{P}}$  of projections to new dimensions—Principal Components (PCs, top plot in the right column) and a transformation matrix  $\hat{\mathbf{G}}$  (bottom plot). The **first PC**, i.e. the value in  $s$ -th row of the first column of the matrix  $\hat{\mathbf{P}}$  is the key observable in our study and we will use the **reactivity**  $RA_s(k)$  for the keyword  $k$  and source  $s$ . World map has been obtained using `ggplot2` R package<sup>20</sup> using `maps` R package data<sup>21</sup>.



**Figure 2.** An example temporal fluctuation scaling plot for the keyword  $k = \text{China}$  ( $\Delta = 1$  day). Each point represents one publisher  $s$ , its mean  $\mu_{s,\Delta}(k)$  and standard deviation  $\sigma_{s,\Delta}(k)$  of activity timeseries calculated for time window size  $\Delta$ . The black line represents an OLS fit to the points after calculating the logarithm of  $\mu$  and  $\sigma$ . The vertical distance from the line to a point for logarithmic variables is the residual  $R_{s,\Delta}(k)$  as described in the “Results” section. Point colors represent source biases; shapes—sources’ continents. Several sources with known and unknown political bias are annotated.

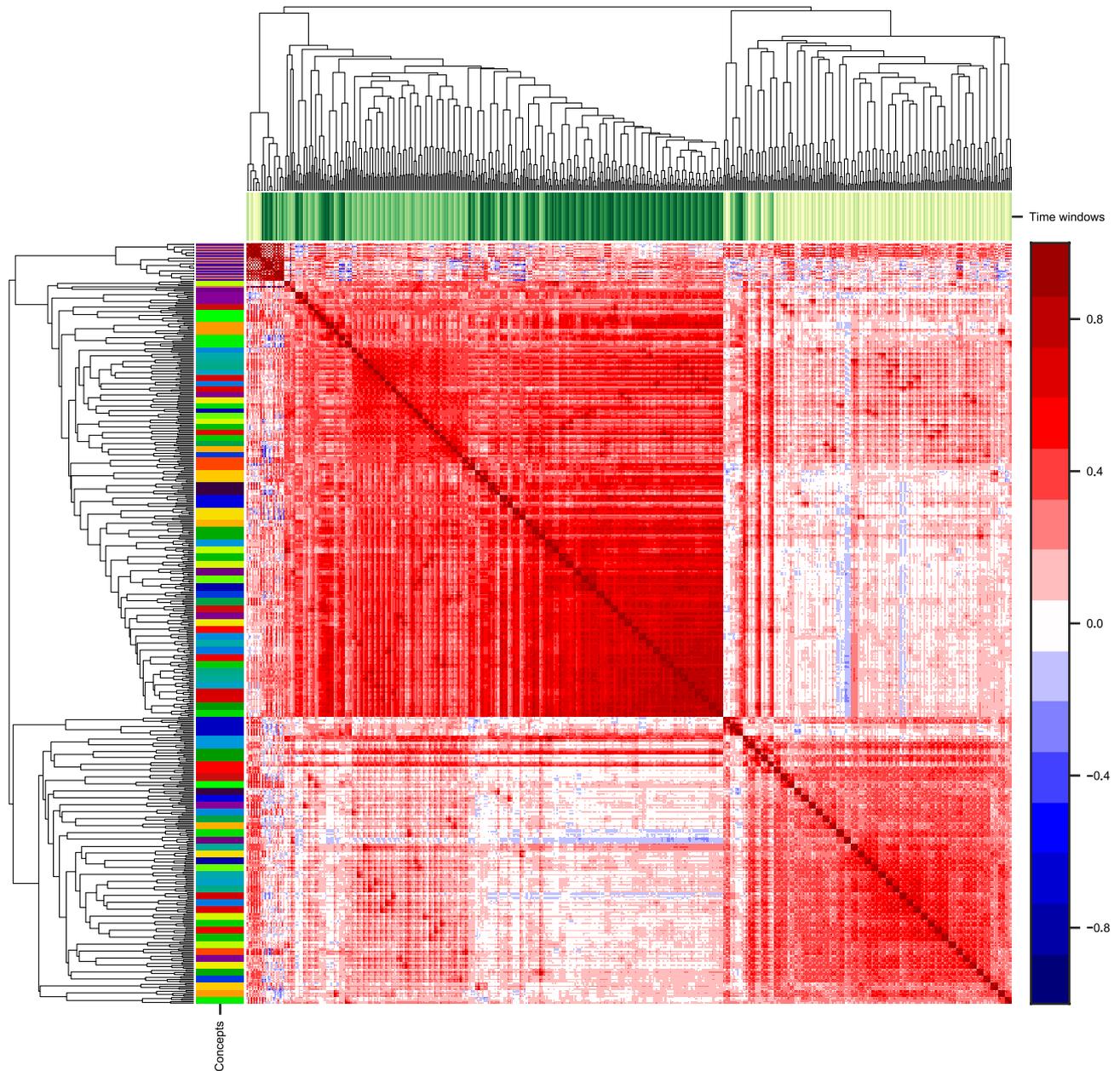
Averaging over all concepts reveals dependencies between  $R_{s,\Delta}(k)$  for various pairs of time window sizes  $\Delta$ . The *long* group had high in-group correlations (0.55) compared to *medium* (0.38) and *short* (0.36); residuals in the *medium* group positively correlated with the *long* group (0.33) and slightly less to the *short* group (0.16). There was no correlation between *short* and *long* timescales (0.06). Similarly, we averaged the correlation matrix elements for pairs of concepts over all time window sizes. The mean of the diagonal elements was 0.52, and the mean of the off-diagonal elements was 0.26.

Thus, residuals for the given source and concept positively correlate for time windows of similar size. In the same manner, residuals for a given source and time window positively correlate across different sets of keywords. This is surprising because the residuals for each concept were calculated using different article sets.

**Principal components analysis of publisher residuals for each concept.** So far we have considered standard deviation residuals separately for different lengths of observation windows  $\Delta$ . For every publisher  $s$  and for every topic  $k$  there are 14 standard deviation residuals  $R_{s,\Delta}(k)$  in matrix  $\hat{\mathbf{R}}(k)$  (the middle matrix in the right row of Fig. 1). To reduce the dimensionality of these observations we performed Principal Component Analysis (PCA) on the residuals  $R_{s,\Delta}(k)$  assuming that values for different  $\Delta$  are related to mutually orthogonal directions. PCA produces a matrix of projections of the original variables to Principal Components (PCs)  $\hat{\mathbf{P}}(k)$  (top matrix in the right column) as well as a matrix  $\hat{\mathbf{G}}(k)$  (bottom matrix) of new basis vectors that transforms  $\hat{\mathbf{R}}(k)$  into  $\hat{\mathbf{P}}(k)$  according to

$$\hat{\mathbf{P}}(k) = \hat{\mathbf{R}}(k)\hat{\mathbf{G}}(k) \tag{2}$$

The columns of the transformation matrix  $\hat{\mathbf{G}}(k)$  are eigenvectors of the empirical covariance matrix  $\hat{\mathbf{R}}(k)\hat{\mathbf{R}}^T(k)$  of residuals. The first column  $\mathbf{G}^1(k)$  corresponds to the largest eigenvalue and it defines the direction of the first principal component axis<sup>25</sup>. To make the PCs comparable across keywords, we forced the directions of the axes to have the sign of the contribution from the shortest time window residual (1 min). The first PC (PC1, elements of the first column of the matrix  $\hat{\mathbf{P}}(k)$ ) is almost the mean of the residuals for all time window sizes, i.e., all  $\Delta$  contribute to this PC almost in the same way (see top row in the matrix on the left panel of Fig. 4). This usually explains around 50% of the variance (right panel in Fig. 4). The second PC (PC2, the second column of  $\hat{\mathbf{P}}(k)$ ; 30% of variance) has opposing signs for contributions from residuals for time window sizes below 1 day and over 1 day; contributions for the extreme values were the highest. In the third PC (around 10% of variance), contributions for  $\Delta \leq 30$  min and  $\Delta \geq 3$  days have opposite signs to those for  $\Delta$  over 1 h and below 3 days. PC1 is the residual of a source for a given keyword averaged over the timescales analyzed. Thus, high absolute values of PC1 indicate atypically low/high fluctuations in timeseries of the source for one or more time window sizes. PC2 shows whether the atypical fluctuations were observed for short or long timescales. PC3 seems to describe



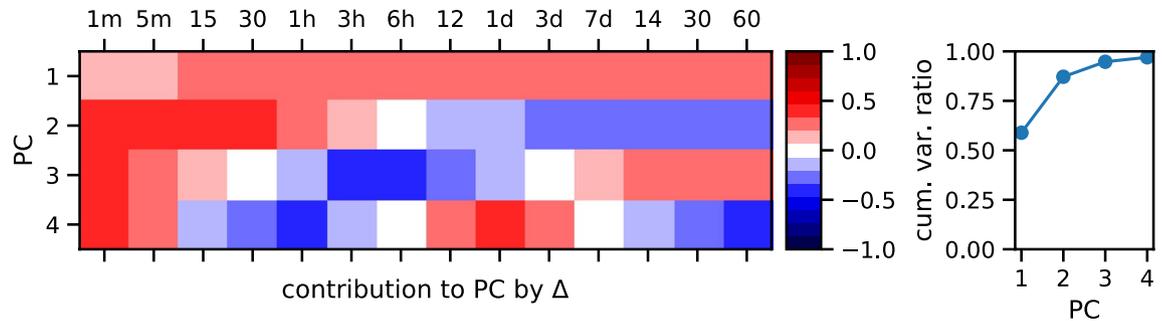
**Figure 3.** Agglomerative Clustering of a correlation matrix of residuals for all concepts and time window sizes. Each row/column stands for one of 546 ( $k, \Delta$ ) combinations. A color bar for the correlation matrix values is on the right hand side. Additional color labels on top of the matrix represent time window size  $\Delta$  (lightest—1 min, darkest—60 days); color labels on the left side of the matrix represent different keywords  $k$ . Residuals for short ( $\Delta < 1$  h), medium ( $1 \text{ h} < \Delta < 1$  day), and long ( $\Delta > 1$  day) time windows are clustered together; inside the three clusters, residuals for most keywords  $k$  tend to be close to each other.

the variance of a source in scales of several hours; it is probably connected to whether a source has a day/night activity cycle. PCs in the cleaned dataset (Fig. 4) have a very similar composition to that in the raw data (see Supplementary Information: Cleaning dataset).

We calculated a correlation matrix of vectors representing each source's selected PC  $p$  for each keyword  $k$ , as for the residuals.

Figure 5 shows that for each PC value, all keywords form a separate cluster. The mean correlation inside a cluster corresponding to results for one PC and all keywords was 0.44.

This finding means that the reporting style is intrinsic for a source rather than a keyword, nevertheless, we believe that slight differences in the values for various keywords might capture subtle reporting style differences between groups of sources when aggregated. We select the first PC as the main indicator of source  $s$  reporting style on a given topic  $k$ —*reactivity*  $RA_s(k)$ —as the first principal component score:



**Figure 4.** Principal Component Analysis of residuals by time window size for the keyword *European Union (clean data)*. (left) Contributions of residuals  $R_{s,\Delta}(k)$  for the  $\Delta$  analyzed to the first four PCs (first four rows of matrix  $\hat{\mathbf{G}}$ ). The first principal component is roughly the arithmetic mean of the residuals over different timescales, the second PC has opposite loadings for long and short timescales. (right) A cumulative explained variance ratio for first four PCs. The first four PCs explain typically around 97% of variance.

$$RA_s(k) \equiv P_{s,1}(k) = \mathbf{R}_s(k)\mathbf{G}^1(k) \quad (3)$$

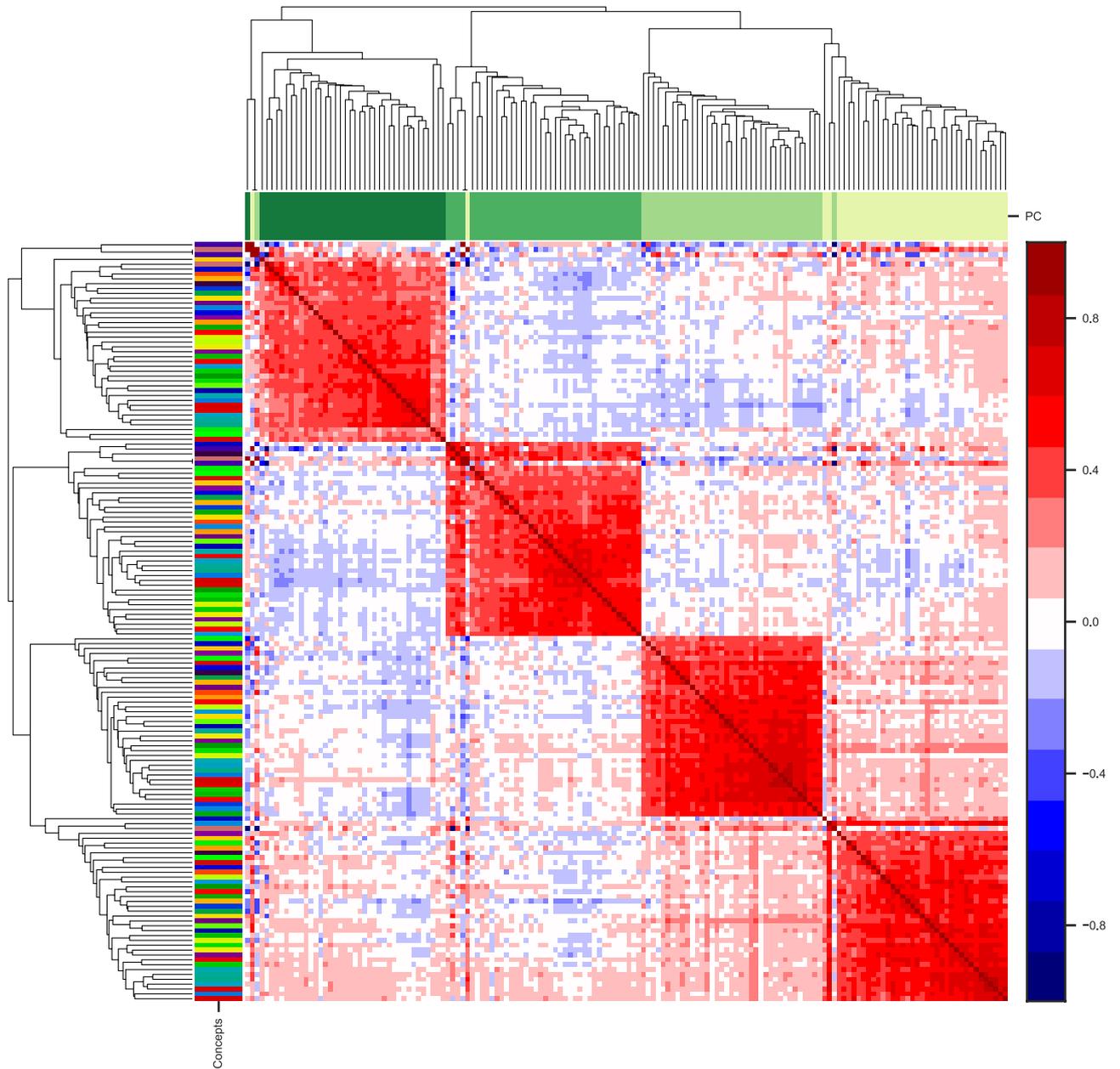
where  $\mathbf{R}_s(k)$  is the  $s$ -th row of matrix  $\hat{\mathbf{R}}(k)$  and  $\mathbf{G}^1(k)$  denotes the first column of transformation matrix  $\hat{\mathbf{G}}(k)$ .

For news outlet activities, an intuitive interpretation would be that higher-than-expected fluctuations show a *reactive* reporting style for a given topic by a given source (e.g. activity influenced by external events) and lower fluctuations indicate a *stable* reporting style. These differences are subtle but clearly visible in the Supplementary Information: Additional dataset info. A source's relative burstiness for a topic compared to other news outlets might be useful when interpreting an article as a signal lateral to general interest (number of articles) or sentiment towards the topic.

**Aggregated reactivity.** Having defined a measure of the reactivity  $RA_s(k)$  of publisher  $s$  towards topic  $k$ , we used it to quantify the typical reporting style of groups of publishers. Two features of news outlets that we have used for aggregation are their *political bias* and *country of origin*. Below, we analyze patterns of median reactivity by group.

**Reactivity by country.** The median  $RA_s(k)$  by country for polarizing concepts (see the Concepts subsection of Methods) are in Fig. 6 (for other concepts—see Supplementary Information: Aggregated reactivity). Rows represent countries with the most sources tracked by ER with the highest topping the list; columns stand for concepts. Colors indicate the medians  $\langle RA_s(k) \rangle_{s \in \mathcal{S}_c} = C_c(k)$  of the reactivity for all sources  $s \in \mathcal{S}_c$  from a given country  $c$  when reporting a given topic  $k$ . Recall that reactivity  $RA_s(k)$  has a logarithmic nature (since the original variables were logarithmic), so one unit difference translates to an order of magnitude (factor 10) difference in fluctuations. It is clear that sources from United States, United Kingdom, France, and Australia have fewer-than-typical fluctuations for all the polarizing concepts—colors are very close to white ( $C_c(k) \approx -0.7$ ). Russian, Czech, Polish, and Ukrainian sources also seem to typically cover the topics stably ( $C_c(k) \approx -1.5$ ), Chinese, and Indonesian—reactively ( $C_c(k) \approx 1.0$ ). German and Argentinian publishers most reactively reported on *Abortion* ( $C_c(k) \approx 1.5, 2.1$ , respectively); Canada and Indonesia on *Cannabis* (2.9, 1.8); India and Indonesia on *Capital punishment* (0.9, 0.7); India and China on *Homosexuality* (3.6, 2.7); China and India on *Same-sex marriage* (5.2, 1.6). Interestingly, in Italy the topic *Homosexuality* was reported rather stably ( $-0.5$ ) but *Same-sex marriage*—reactively (1.2). The most notable stable reporting was observed for Russian, Chinese, Spanish and Indian sources for *Abortion* ( $-3.2, -2.1, -2.0, -1.9$ ). Stable reporting was also found for *Cannabis* by Polish and Russian outlets ( $-2.5, -2.11$ ); *Capital punishment* in Czech Republic and Poland ( $-2.2, -2.0$ ); *Homosexuality* in Russia, Ukraine and Canada ( $-2.4, -1.3, -1.2$ ), *Same-sex marriage* in Poland and Mexico ( $-2.2, -2.0$ ). The values seem to reflect the temperature of the national views towards a given concept. For example, in 2018, Canada legalized *Cannabis*<sup>26</sup> and India legalized *Homosexuality*<sup>27</sup>. Publishers from both countries tended to report reactively on the respective topics, presumably driven by the legislative changes. Similarly, in China, there were two major rulings on *Same-sex marriages*<sup>28</sup>. Indonesia reactively reported (1.9, see SI) on *Terrorism* as it suffered six terrorist attacks in 2018<sup>29</sup>; the high reactivity of Italian media for *Shooting* (4.1, see SI) might be an effect of the Macerata shooting<sup>30</sup>. The French media were reactive towards *China* (3.6, see SI) as 2018 was an important year for bilateral relations between the two countries<sup>31</sup>. Recall that high reactivity does not necessarily mean that the given keyword was frequently discussed in the given country. In fact Fig. 6 indicates that *Same-sex marriage* and *Homosexuality* were discussed in China very reactively but less frequently than *Capital punishment*, which was reported very stably (compare square colors and sizes in the above-mentioned figure). Similarly, *Homosexuality* in India was a reactive topic although it attracted less interest than *Abortion* or *Cannabis*, which were reported very stably. More such examples can be found in the Supplementary Information.

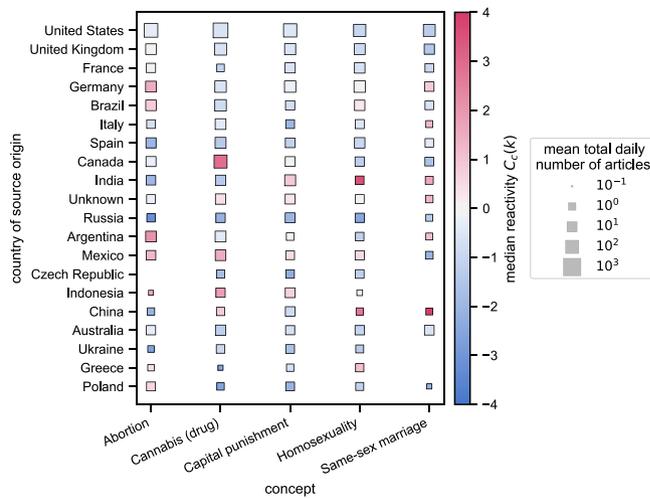
**Reactivity by political bias.** The above analysis has shown that the median value of  $\langle RA_s(k) \rangle_{s \in \mathcal{G}}$  reflects the general attitude of a group of sources  $\mathcal{G}$  (e.g. from the same country) towards a topic  $k$ . Around 10% of the publishers in Event Registry were also annotated by [mediabiasfactcheck.com](https://mediabiasfactcheck.com) for political bias—see the Dataset



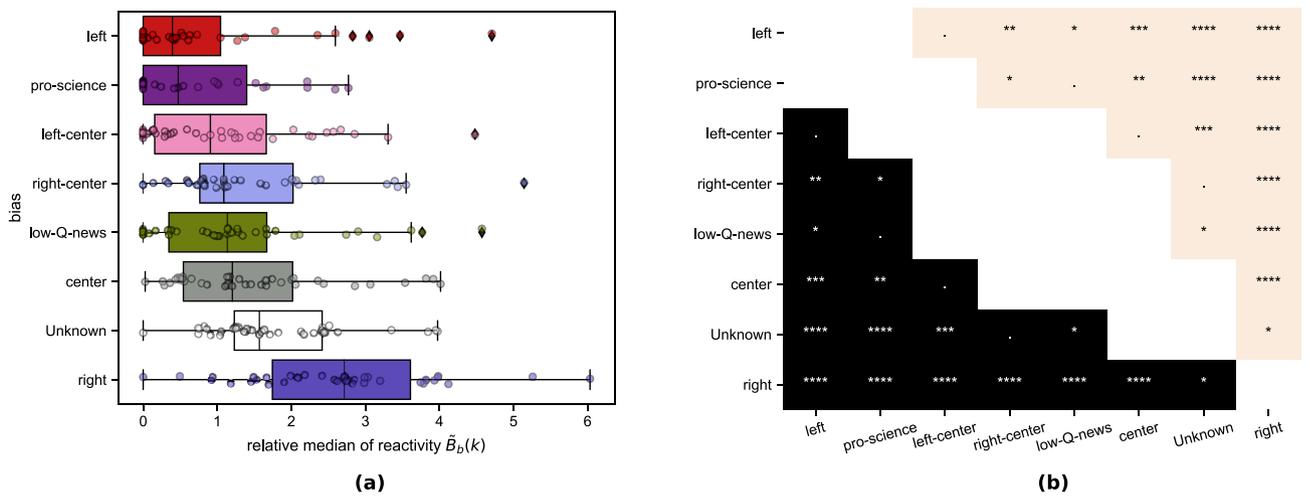
**Figure 5.** Agglomerative Clustering of a correlation matrix of the first four principal components of the residuals  $P_{s,p}(k)$  ( $a = 1, \dots, 4$ ) for all concepts  $k$ . Each row/column stands for one  $(k, p)$  combination. A color bar for the correlation matrix values is on the right hand side. Additional color labels above the matrix represent an order of PC  $a$  (lightest—1st PC, darkest—4th PC); color labels on the left side of the matrix represent different keywords  $k$ . PCs  $P_{s,p}(k)$  of the same order  $p$  are clustered; the mean correlation coefficient  $\langle \rho(k_1, p_1; k_2, p_2) \rangle_{p_1=p_2} = 0.44$ .

info: Political bias of sources subsection of Methods. This is used to investigate the reactivity of sources grouped by political bias.

A general tendency was that sources marked as *left* represent (on average) a stable reporting style towards most concepts, while *right* tended to have a reactive/bursty style. For 30 out of 39 analyzed concepts (see Supplementary Information: Aggregated reactivity), sources in the *right* group had the highest or the second highest reactivity, while the *left* group had the lowest or the second lowest for 26 concepts; often (in 18 cases) all the groups *left-center*, *center*, *right-center* had their median above the *left* and below the *right*. The *right* was relatively nonreactive for the concepts *Kim Kardashian*, *Marvel comics*, and *Same-sex marriage*; the *left* was relatively reactive for *Cannabis (drug)*, *Climate change*, and *Shooting*. The positions of *pro-science* and *low-Q-news* (ie. *low quality news*) did not have a stable pattern, perhaps caused by the low number of sources in these groups. The sources with annotated bias were mostly of US origin.



**Figure 6.** The median reactivity  $C_c(k)$  of the reactivity for all sources from a given country  $c$  for keywords  $k$  related to “polarizing” concepts. Square size is proportional to the mean daily number of articles published by sources from a given country  $c$  on a given topic  $k$ ; color represents the median reactivity  $C_c(k) = \langle RA_s(k) \rangle_{s \in \mathcal{S}_c}$  of sources from the country  $c$ . Missing squares indicate that there were no publishers from the country that published at least 36 articles on the topic in our dataset. Red symbols correspond to topics that were reactively discussed in the country in 2018.



**Figure 7.** Comparison of relative median reactivity  $\tilde{B}_b(k)$  between political bias groups for all keywords. *Left*-oriented sources are generally less reactive than the *right*-oriented sources. **(a)** values of relative median reactivity  $\tilde{B}_b(k)$  (see Eq. (4)) for all keywords by political bias. **(b)** Results of pairwise Dunn median tests with a two-step Benjamini–Krieger–Yekutieli FDR adjustment<sup>32</sup>; the adjusted  $p$ -values describe the likelihood of the observed difference in medians of two samples assuming there is no difference between the medians of their populations. Thus, the lower the  $p$ -value, the more statistically significant the difference between the two group medians.  $**** - p \in [0, 0.001)$ ,  $*** - p \in [0.001, 0.005)$ ,  $** - p \in [0.005, 0.01)$ ,  $* - p \in [0.01, .05)$ ,  $. - p \in [0.05, 0.1)$ , otherwise  $p \in [0.1, 1]$ . The color indicates which group had the higher median—black if the median of the group from the corresponding row was higher than the median of the group from the corresponding column; white—the opposite case.

To compare reactivity values across all of the analyzed keywords, we calculated the median reactivity of each bias group  $b$  for each keyword  $k$  as  $(RA_s(k))_{bias(s)=b} = P_b(k)$ , where  $s$ —publisher,  $bias(s)$ —a political bias of a source  $s$ . We defined the relative median reactivity of a bias group for a keyword as follows:

$$\tilde{B}_b(k) = B_b(k) - \min_{b'} B_{b'}(k) \tag{4}$$

Thus, a bias group with the lowest  $B_b(k)$  for a given  $k$  will indicate  $\tilde{B}_b(k) = 0$ . In Fig. 7 (left panel), reports  $\tilde{B}_b(k)$  by a bias group for all the keywords. To quantify the general tendencies of the relative median reactivity  $\tilde{B}_b(k)$ , we performed a Kruskal–Wallis test with the FDR-adjusted Dunn’s test for pairwise comparisons<sup>32</sup> of the bias

groups; the results are in Fig. 7 (right panel). Both *left* and *pro-science* typically had lower  $\tilde{B}_b(k)$  values than the remaining groups but are statistically indistinguishable from each other. On the other hand, the relative reactivity  $\tilde{B}_b(k)$  for *right* was significantly higher than the remaining groups; *Unknown* scored similarly but even this group had lower average reactivity than *right*. *Low-Q-news* and *left-center* were distinguishable from *right* and *Unknown* only; the relative median reactivity  $\tilde{B}_b(k)$  of *center* and *right-center* are over *left* and below *right*. The statistical tests confirmed that the *left* is significantly less reactive than the *right*.

## Discussion

This paper reports a novel two-step method to quantify the temporal fluctuations of a variable  $f_{s,\Delta}^{(t)}(k)$  describing the dynamics of entities  $s$  belonging to a complex system, such as news outlets in the media space. In this case  $f_{s,\Delta}^{(t)}(k)$  is the number of articles published by the outlet  $s$  about a keyword  $k$  during a time window  $\Delta$  around a time moment  $t$ . The first step consists of checking whether the Temporal Fluctuation Scaling law (TFS) is valid for the system of entities  $s$  and, if so, determining TFS residuals for various time window sizes  $\Delta$ . The residuals are ratios of empirical standard deviations  $f_{s,\Delta}^{(t)}(k)$  (measured from a time series) to the value predicted by the TFS law<sup>15,18</sup>. The second step is dimensionality reduction using Principal Component Analysis (PCA) to obtain a value independent of the time window size. The First Principal Component  $PC1$  contains aggregated information about fluctuations in all timescales and we call this *reactivity*. To give a practical application, we used a case study of the online news aggregator Event Registry<sup>23</sup>.

We started with calculating TFS for activity timeseries in several timescales  $\Delta$  for each keyword  $k$  separately (Fig. 2) and then obtained TFS residuals  $R_{s,\Delta}(k)$  (see Eq. (1)) in each timescale  $\Delta$  for every source  $s$ . Using a correlation matrix of the residual values (Fig. 3), we showed that the residuals of a given source positively correlate across different keywords for similar time window sizes. Moreover, using this method there are three timescales for fluctuations of publishers' activities (*short*— $\Delta \leq 1$  h, *medium*— $1 \text{ h} < \Delta \leq 1$  day, *long*— $1 \text{ day} < \Delta$ ) reported in<sup>24</sup>. We conclude that reporting style is intrinsic to a publisher but small differences in the values for keywords capture subtle differences in reporting style. A similar observation was reported in<sup>15</sup> for a spatial clustering of world precipitation data and a clustering of financial stocks from the same sectors.

To extract the most substantial factors that could aggregate the residuals over different timescales we utilized PCA. PCA reveals linear combinations of residuals at different timescales to form an uncorrelated orthogonal basis set where consecutive PCA components correspond to the highest variances between news sources (Eq. (2)). From the results,  $PC1$ , later called reactivity  $RA_s(k)$  (Eq. (3)), is a good summary of source  $s$  deviation from the typical variation of activity on a topic  $k$ . In fact, the reactivity value for the dataset has nearly the same loadings from residuals  $R_{s,\Delta}(k)$  at each timescale  $\Delta$ . On the other hand, the  $PC2$  has opposing loadings from residuals in short ( $\Delta < 1$  day) and long ( $\Delta > 1$  day) timescales. The 1st and 2nd PCs together explain about 80% of the cumulative variance for the dataset (Fig. 4) and therefore can be used to reduce the system dimensionality. A correlation matrix of the PC values  $\hat{P}(k)$  (Fig. 5) upholds our observation—PC values for a given source obtained from various keywords positively correlate and therefore seem to be intrinsic to the source.

We then aggregated reactivity values grouped by political bias or country of origin to assess the relative reporting style of the groups towards different topics. First, we showed that the reactivity measure can detect which topics were reported reactively by comparing the results grouped by source country (Fig. 6), e.g., legalization of Cannabis in Canada and Homosexuality in India. Such topics would not be highlighted and could be disregarded if one relied only on activity, i.e., taking into account only the number of articles about the given keyword. Then, we analyzed differences in reporting style between (mostly American) sources grouped by political bias (Fig. 7). Groups *left* and *pro-science* had been typically less reactive than others for the concepts. On the other hand, news outlets annotated as *right* were usually the most reactive.

The reactivity value is dimensionless and normalized which enables comparisons of source reactivity across different context. For example, it could be applied to compare the relative volatility of stocks in various stock exchanges. Although our reactivity  $RA_s(k)$  coefficient (Eq. 3) is similar in concept to Fano factor<sup>8</sup> and the coefficient of variation<sup>9</sup> it has several advantages compared to them. (i) It is based on a comparison of activity fluctuations of a given object to activity fluctuations in an ensemble of other objects that obeying temporal fluctuation scaling, so it has a well-defined *zero level*. (ii) Its single value takes into account fluctuations at multiple timescales  $\Delta$  due to Principal Component Analysis. (iii) It is not correlated ( $\rho = -0.05$ ) with activity and so can be used to compare fluctuations in entities of different sizes. In fact, the correlation between the Fano factor and activity for the ER data set is about  $\rho = 0.3$ . Thus, this factor overestimates fluctuations for large publishers and the correlation between the coefficient of variation and activity is about  $\rho = -0.57$ , overestimating fluctuations for small and medium outlets (for details—see Supplementary Information: Measure comparisons).

Another way of comparing reactivity to the Fano factor and the coefficient of variation is to use them instead of TFS residuals in the first step of our method and perform the second step and analyses without changes (see Supplementary Information: Measures comparisons). Correlation matrices obtained with these alternative variability measures could not be clustered into three timescales. For the Fano factor, two regimes were observed (for  $\Delta$  below or above 15 min) and a typical correlation inside each cluster was lower than for the reactivity  $RA$ . For the coefficient of variation, two regimes also exist (for  $\Delta$  below or above 1 day) and almost all values positively correlated. Our reactivity approach finds clustering in *three timescales* (Fig. 3) that cover both divisions found by the Fano (Supplementary Fig. S18) and coefficient of variation (Supplementary Fig. S17) approaches separately. It follows that while the Fano is sensitive to the short timescale division and the coefficient of variation to the long timescale, reactivity takes into account all the timescales in the case analyzed. Moreover, aggregating variability by the *political bias* of a source (Fig. 7) was most successful when reactivity  $RA$  was used, because ordering political biases by median relative variability was similar to ordering on a political spectrum—from *left* through *center* to *right*. For the coefficient of variation (Supplementary Fig. S23), the ordering seemed uninformative—the

Continent	Sources	(With known bias)
Europe	3065	76
North America	2650	526
Asia	1015	71
South America	655	4
Africa	426	7
Unknown	190	7
Oceania	154	18

**Table 1.** Sources by continents. Only sources which published at least one article per month about at least one of the analyzed keywords.

*left-center* group had the lowest relative fluctuations and *left* had the highest. For the Fano factor, only *left* biased sources had a median significantly different from other bias groups.

Our reactivity parameter,  $RA_s(k)$  is related to fluctuations of a single unit  $s$  at different timescales but it should not be mistaken for the generalized Hurst  $H(q)$  exponent that is calculated from a scaling law describing  $q$ -th order fluctuation as the function of time window size  $\Delta$ . In fact, *reactivity*  $RA$  is related to the scaling behavior (TFS) that comes from comparing different units. This means that *reactivity* takes into account features following from a set of entities while the generalized Hurst exponent  $H(q)$  reflects only scaling for the timeseries of a single entity. The relation between the TFS scaling exponent and the Hurst exponent  $H(2)$  was discussed in<sup>15</sup>. Since the TFS scaling law takes into account fluctuations of order  $q = 2$ , we are not considering how the  $RA$  measure is dependent on the multifractality of a timeseries when the Hurst exponent is  $q$ -dependent.

We have summarized above the properties of the reactivity measure  $RA$ , which is related to the first principal component PC1 of the residuals matrix  $\hat{\mathbf{R}}(k)$ . The first two PCs (PC1 and PC2) have also been used to detect corrupted time series (see SI: Cleaning Dataset) in our dataset. The first type of data corruption is related to an abnormally large number of published articles by a given outlet in a short time period (a few minutes to a few hours) and is likely to be result of changes in the website shape that misleads the web crawler. This can be detected by a high value of the PC1 for this publisher. The second type of data corruption is related to missing records in a period that can be caused by the lack of a functional crawler. This can be detected by a value of PC2 indicating large corrections to TFS in long timescales for this publisher.

We believe the new approach is universal and can be applied to other fields where a proper assessment of the relative amount of signal variation is critical (e.g. financial markets<sup>33–35</sup>, or neuroscience<sup>36,37</sup>). Additionally, the number of sources annotated by political bias was relatively low and skewed towards the US which significantly decreased the reliability and generality of the results, although giving a more homogeneous dataset to analyze. Unfortunately, mass-scale automatic bias classification with high accuracy seems out-of-reach for now<sup>38,39</sup>. Moreover, more information could be extracted from similar datasets by inspection of consecutive PCs. For example PC2 indicates whether the fluctuations described by PC1 are biased towards short or long timescales. A method to adjust the mean and variance of corrupted timeseries might be developed using the new method to compensate for the two described disturbances (artificial activity spike, periods of missing data). One of the biggest disadvantages of the new measure is its high computational complexity, but this could be resolved using approximate methods<sup>40</sup>.

In conclusion, we have introduced a method that uses predictions from a fluctuation scaling law to benchmark observed standard deviations and our reactivity measure  $RA$  aggregates information about fluctuations at different timescales. We show that the approach makes it possible to compare news outlet reporting styles by geography and political orientation.

## Methods

**Dataset info.** Online news media are hard to analyze on a large scale due to their distributed character compared to social media. Social media data is concentrated around a few important platforms (like Twitter or Facebook) but the news industry has a huge number of news outlets. Nevertheless, mass-scale surveillance of the online news industry is possible with news aggregators such as Event Registry (ER, [www.eventregistry.org](http://www.eventregistry.org)). ER is an online system which collects, annotates, clusters, and stores news items from online sources around the globe<sup>23</sup>. The system has been maintained by the Artificial Intelligence Laboratory (Jožef Stefan Institute, Ljubljana, Slovenia) since 2014 and tracks tens of thousands of publishers in near real-time. While its purpose is to track world events by clustering news items (across different languages), access to this database was provided by one of the coauthors. We queried the system to download all articles containing one of selected concepts (see the “Concepts” subsection of “Materials and methods”). Each article was represented as a tuple: *publisher*, *keyword* (*topic*), *timestamp*.

**Sources.** When our data set was created there were 8155 sources in ER that published at least three articles per month about at least one of the analyzed concepts. To examine geographical and political differences between the publishers, we considered the sources’ countries of origin and political biases, where applicable. Source counts by continent can be found in Table 1.

Source countries of origin are reported by ER. Sources from all continents were present in the dataset; over two thirds were in Europe (37%) or North America (33%). Asia (13%) and Africa (5%) are underrepresented in the set compared to their populations. The relatively small number of publishers from South America (8%) and Oceania (2%) seems to be justified by their relatively small populations. Geographic data was missing for 2% of sources.

**Political bias of sources.** Political bias data was gathered from Media Bias/Fact Check (MBFC, [www.media-biasfactcheck.com](http://www.media-biasfactcheck.com)). At the moment of writing, the website listed 2896 online media bias annotations, but only 709 overlapped with our list. While its credibility is sometimes questioned, it has been regarded as accurate enough to be used as ground-truth for e.g. media bias classifiers<sup>39,41,42</sup>, fake news studies<sup>43–45</sup>, and automatic fact-checking systems<sup>46–48</sup>.

MBFC groups sources into 8 categories: *left* (92 sources in our list), *left-center* (255), *center* (147), *right-center* (125), *right* (53), *pro-science* (27), *fake-news* (21), *conspiracy* (17), *satire* (2). We merged the last three categories to one—*low-Q-news* (41). Following the website, sources marked as *left* are “moderately to strongly biased toward liberal causes through story selection and/or political affiliation”, *left-center* have “slight to moderate liberal bias”, sources marked as *right* are “moderately to strongly biased toward conservative causes” and *right-center* “media sources are slightly to moderately conservative in bias”. *Center* media are considered to have a minimal bias, provide factual and sourced reporting, and to be “the most credible media sources”. Sources with an extreme liberal/conservative bias are listed mainly as *fake-news* or *conspiracy*. Sources from all continents can be found in the MBFC list but the distribution is very US-centered—the vast majority of sources tracked by ER and annotated by MBFC are in North America (75%); the rest consist of a few European (10%), Asian (10%), Oceanian (4%), African (1.5%), and South American (1%) news outlets.

**Concepts.** One of natural language processing techniques implemented in ER is the extraction of concepts and entities (called here *topics* or *keywords*, despite often being longer than one word) involved in a given piece of news. ER identifies keywords using Wikipedia and so every keyword has an associated Wikipedia article. The set of concepts selected for the study consisted of 19 countries and 20 current topics. We chose a few countries from all the continents: North America (*United States, Mexico*), Europe (*United Kingdom, France, Germany, Austria, Poland, Slovenia, and European Union*), Asia (*China, India, Japan, North Korea, Iran*), South America (*Brazil, Argentina*), Africa (*Egypt, South Africa, Morocco*), and Oceania (*Indonesia*). The current topics set was divided into “polarizing” and “other” topics. The first group consists of concepts which we expect to be perceived in different way by various political groups (*Cannabis (drug), Capital punishment, Abortion, Homosexuality, Same-sex marriage*). These concepts seem likely to differentiate between liberal and conservative views. For example, they were a vital part of questionnaire in the NetSense experiment<sup>49–51</sup>; unfortunately, Event Registry had too few articles on other topics used in the study (*Premarital sex, Euthanasia*) for statistical analyses. The “other” group contains keywords related to violence (*Terrorism, Shooting*), sports (*Real Madrid C.F., Roger Federer, Usain Bolt*), celebrities (*Kim Kardashian, British royal family*), and other concepts covered nowadays (*Facebook, Islam, Climate change, Universe, Bitcoin, Blockchain, Marvel Comics*).

Numbers of sources and articles for each concept are in the Supplementary Information: Additional dataset info.

**Statistical methods.** *Extracting timeseries from lists of articles.* The raw data for this study consists of lists of article publication dates and times  $f_s(k) = \{\tau_{s,1}(k), \tau_{s,2}(k), \dots\}$  for each keyword  $k$  and publisher  $s$  that had at least three articles with the keyword. To obtain timeseries, we divided the observation period (01.01–31.12.2018,  $T = 1$  year) into non-overlapping windows of length  $\Delta$  and counted occurrences of publications in window  $t$  as follows:

$$f_{s,\Delta}^{(t)}(k) = \left| \{ \tau_{s,i}(k) : t\Delta \leq \tau_{s,i}(k) < (t+1)\Delta \} \right| \quad (5)$$

where  $|\mathcal{X}|$  is the number of elements of a set  $\mathcal{X}$ . For a given  $\Delta$ , we have  $W = \lceil T/\Delta \rceil$  time windows.

To allow statistical analyses, the longest period analyzed was 60 days. There was not much information in timescales below a few minutes as the exact position of the timestamp was often determined by the crawl time. We chose 14 time window sizes  $\Delta \in \{1 \text{ min}, 5 \text{ min}, 15 \text{ min}, 30 \text{ min}, 1 \text{ h}, 3 \text{ h}, 6 \text{ h}, 12 \text{ h}, 1 \text{ day}, 3 \text{ days}, 7 \text{ days}, 14 \text{ days}, 30 \text{ days}, 60 \text{ days}\}$  for the study. When the window length  $\Delta$  did not divide  $T$  without a remainder, we treated the last time window equally with the rest.

**Temporal fluctuation scaling.** The following introduces temporal fluctuation scaling (TFS), which was been applied to our data and served as a starting point to define TFS residuals.

First, we calculate means of the timeseries:

$$\langle f_{s,\Delta}^{(t)}(k) \rangle = W^{-1} \sum_{t=1}^W f_{s,\Delta}^{(t)}(k) \quad (6)$$

then variances:

$$\sigma_{s,\Delta}^2(k) = \left\langle \left[ f_{s,\Delta}^{(t)}(k) \right]^2 \right\rangle - \left\langle f_{s,\Delta}^{(t)}(k) \right\rangle^2 \quad (7)$$

where

$$\left\langle \left[ f_{s,\Delta}^{(t)}(k) \right]^2 \right\rangle = W^{-1} \sum_{t=1}^W \left( f_{s,\Delta}^{(t)}(k) \right)^2 \quad (8)$$

It has been observed for router activity and email traffic (but also stock markets, river flows, or printing activity)<sup>15</sup> that the standard deviation of the  $i$ -th entity's activity (calculated in windows of size  $\Delta$ ) scales with its mean with an exponent  $\alpha$  (and a multiplicative constant  $B$ ):

$$\tilde{\sigma}_{s,\Delta}(k) = B_{\Delta} \left\langle f_{s,\Delta}^{(t)}(k) \right\rangle^{\alpha_{\Delta}}. \quad (9)$$

Artificial examples of this procedure are shown in Fig. 1, explaining how a set of sources publishing about a single concept gives rise to TFS (middle column in Fig. 1) for different time windows.

Recent theoretical advances have greatly improved our understanding of mechanisms that underlie TFS in a given system<sup>40,52–54</sup>. One of the generative mechanisms is that entity activities in this class of systems can be described with Tweedie distributions; exponents between 1 and 2 suggests the compound Poisson-gamma distribution<sup>55</sup>. This model arises when in each time step activity of an entity is a sum of i.i.d. gamma-distributed random variables and the number of the random variables is drawn from a Poisson distribution.

In the case of news outlet activities, each random variable might be an occurrence of a real world event which causes a news outlet to write articles describing it; the number of articles per event follows a Gamma distribution. Each outlet can have a different threshold for an event importance or relevance (influencing the Poisson process rate) and/or attitude towards a topic (influencing parameters of the distribution of articles per event) which can account for heterogeneity in the number of articles and temporal activity fluctuations amount of the entities creating the system.

*Agglomerative clustering of correlation matrices.* To analyze relationships between residuals  $R_{s,\Delta}(k)$  (see Eq. (1)) and Principal Components  $P_{s,p}(k)$  (see the subsection below) for various combination of keywords  $k$  and time window sizes  $\Delta$  or Principal Components  $p$ , we calculated correlation matrices. Each row/column of the matrix represents correlations between one pair  $(k, \Delta)$  or one pair  $(k, p)$  with the remaining pairs. As keywords were reported by different numbers of publishers, it was necessary to calculate correlation coefficients using only publishers that covered both keywords. We used Pearson's correlation coefficient; here for residuals, similarly for PCs:

$$\rho(k_1, \Delta_1; k_2, \Delta_2) = \frac{\sum_{s \in \mathcal{S}_{k_1, k_2}} (R_{s,\Delta_1}(k_1) - \bar{R}_1)(R_{s,\Delta_2}(k_2) - \bar{R}_2)}{\sqrt{\sum_{s \in \mathcal{S}_{k_1, k_2}} (R_{s,\Delta_1}(k_1) - \bar{R}_1)^2} \sqrt{\sum_{s \in \mathcal{S}_{k_1, k_2}} (R_{s,\Delta_2}(k_2) - \bar{R}_2)^2}} \quad (10)$$

where  $R_{s,\Delta}(k)$ —the residual of publisher  $s$  calculated for keyword  $k$  and time window size  $\Delta$ ,  $\mathcal{S}_k$ —a set of publishers that published articles about a keyword  $k$ ,  $\mathcal{S}_{k_1, k_2} = \mathcal{S}_{k_1} \cap \mathcal{S}_{k_2}$ , and  $\bar{R}_x = \frac{1}{|\mathcal{S}_{k_1, k_2}|} \sum_{s \in \mathcal{S}_{k_1, k_2}} R_{s,\Delta_x}(k_x)$ .

To group combinations that were positively correlated, we performed Agglomerative Clustering (AC) of the matrices<sup>56</sup>. The usual inputs to clustering algorithms are dissimilarity (distance) matrices so we calculated distances between the  $(k, \Delta)$  pairs as follows:

$$d(k_1, \Delta_1; k_2, \Delta_2) = 1 - \rho(k_1, \Delta_1; k_2, \Delta_2) \quad (11)$$

AC is a bottom-up clustering method—the algorithm is initialized with each  $(k, \Delta)$  combination in a separate cluster. Then, at each step, the nearest pair of clusters is merged. We selected the unweighted pair group method with arithmetic mean (UPGMA) which defines the distance between clusters  $A$  and  $B$  as the arithmetic mean of the distances between objects in the clusters:

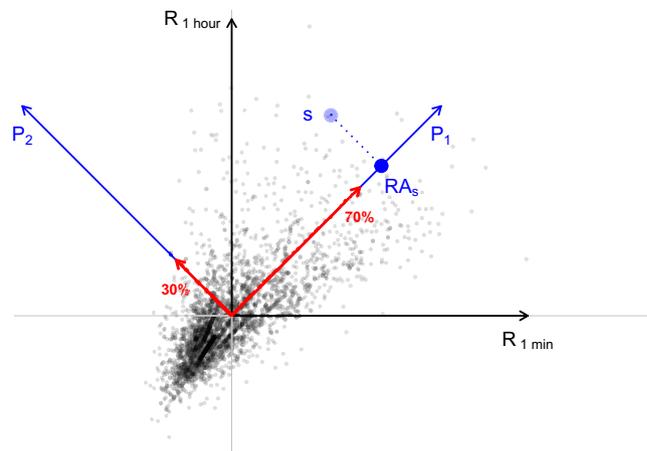
$$d(\mathcal{A}, \mathcal{B}) = \frac{1}{|\mathcal{A}| |\mathcal{B}|} \sum_{a \in \mathcal{A}} \sum_{b \in \mathcal{B}} d(a, b). \quad (12)$$

The result of the clustering can be drawn as a rooted tree (a dendrogram) and used to sort the rows/columns of a correlation matrix.

*Principal component analysis (PCA).* Despite being over a hundred years old<sup>57</sup>, PCA is commonly used in contemporary exploratory analyses of highly dimensional datasets<sup>25,58–60</sup>. This method transforms an original set of variables to so-called Principal Components (PC). PCs are mutually orthogonal linear combinations of the original variables. The first PC is a linear combination of the original variables so that the variance of the data along the axis is maximized; the remaining PCs are constructed in a similar way with additional requirements of orthogonality to all the previous PCs. The main PCA applications are visualization and dimensionality reduction. Nowadays, the procedure is usually carried out using Singular Value Decomposition (SVD) of the data matrix (or, equivalently, Eigenvalue Decomposition of the data covariance matrix). The *scikit-learn*, a Python machine learning library<sup>61</sup> which we performed all the PCA procedures in the study with, uses the LAPACK implementation of the SVD<sup>62</sup>.

In the paper, we denote the decomposition of a matrix  $\hat{\mathbf{R}}$  that includes the original data vectors into PCs as follows:

$$\hat{\mathbf{R}} = \hat{\mathbf{P}} \hat{\mathbf{G}}^T \iff \hat{\mathbf{P}} = \hat{\mathbf{R}} \hat{\mathbf{G}} \quad (13)$$



**Figure 8.** Illustration of PCA dimension reduction. Points are original data—residuals coming from the *Terrorism* keyword for two selected time windows: 1 min ( $R_{1 \text{ min}}$ ) and 1 h ( $R_{1 \text{ h}}$ ). Blue vectors show directions of the new set of variables obtained by the PCA method: first ( $P_1$ ) and second ( $P_2$ ) Principal Components. While red arrows mark the amount of variance explained by these variables (the length of each arrow is proportional to the variance explained by the respective PC). The blue point is a selected source  $s$  and the blue dotted line shows the projection of this data point onto the first PC which is this source's reactivity  $RA_s$ .

where a matrix  $\hat{P}$  contains projections of the original variables to principle components, and  $\hat{G}$  is a matrix of new basis vectors: a transformation matrix for old variables to principle components.

One of the most important features of PCA is its ability to detect linear correlations among the components of data vectors and to indicate the amount of variance explained by each of the newly-introduced variables, i.e., PCs. If the consecutive PCs explain a considerable amount of variance, e.g., 90%, the remaining variables can be ignored. This is known as dimensionality reduction. In the case of our data we make use of the first PC which usually explains over 60% of variance (cf Fig. 4).

As it is impossible to visualize 14 dimensions, Fig. 8 shows residuals for two selected dimensions—1 min ( $R_{1 \text{ min}}$ ) and 1 h ( $R_{1 \text{ h}}$ ) for the keyword *Terrorism*. The plot shows that  $R_{1 \text{ min}}$  and  $R_{1 \text{ h}}$  are not independent: the larger  $R_{1 \text{ min}}$  the larger  $R_{1 \text{ h}}$ , so they should not be treated as a valid set of coordinates. As an outcome of PCA we obtain two new directions— $P_1$  and  $P_2$  that explain, respectively, 70% and 30% of the total variance and create a better setting to describe this data. The value of  $P_1$ , i.e., the projection of the original variables onto the first PC, is the *reactivity*, examined in this study.

**Kruskal–Wallis test and Dunn's post-hoc test.** To check for differences in relative median reactivity  $\tilde{B}_b(k)$  between bias groups, we performed Kruskal–Wallis tests<sup>63</sup>. After a significant Kruskal–Wallis test, pairwise comparisons were performed to order bias groups by increasing median  $\tilde{B}_b$ . The  $p$ -values were FDR adjusted using the two-step Benjamini–Krieger–Yekutieli method<sup>32</sup>. In our study, we used the *scikit-posthocs*<sup>64</sup> implementation of these tests.

### Data availability

The dataset analyzed in the study (lists of news publication timestamps by keyword and publisher) is available from the corresponding author on reasonable request.

### Code availability

Python scripts for calculating TFS and reactivity from a list of timestamps as well as scripts that aggregate reactivity are available from the corresponding author on reasonable request.

Received: 29 June 2020; Accepted: 13 November 2020

Published online: 26 November 2020

### References

- Kondepudi, D. & Prigogine, I. *Modern Thermodynamics: From Heat Engines to Dissipative Structures* 2nd edn. (Wiley, Hoboken, 2015).
- Shaffer, F. & Ginsberg, J. P. An overview of heart rate variability metrics and norms. *Front. Public Health* **5**, 258. <https://doi.org/10.3389/fpubh.2017.00258> (2017).
- Marconi, U. M. B., Puglisi, A., Rondoni, L. & Vulpiani, A. Fluctuation-dissipation: response theory in statistical physics. *Phys. Rep.* **461**, 111–195. <https://doi.org/10.1016/j.physrep.2008.02.002> (2008).
- An, C. E. *et al.* Heart rate variability as an index of resilience. *Mil. Med.* <https://doi.org/10.1093/milmed/usz325> (2019).
- Liu, Y. *et al.* Statistical properties of the volatility of price fluctuations. *Phys. Rev. E* **60**, 1390–1400. <https://doi.org/10.1103/PhysRevE.60.1390> (1999).
- Kwapien, J. & Drożdż, S. Physical approach to complex systems. *Phys. Rep.* **515**, 115–226. <https://doi.org/10.1016/j.physrep.2012.01.007> (2012).

7. Mantegna, R. N. & Stanley, H. E. *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, 2007).
8. Fano, U. Ionization Yield of Radiations. II. The Fluctuations of the Number of Ions. *Phys. Rev.* **72**, 26–29. <https://doi.org/10.1103/PhysRev.72.26> (1947).
9. Allison, P. D. Measures of inequality. *Am. Sociol. Rev.* **43**, 865. <https://doi.org/10.2307/2094626> (1978).
10. Smith, H. F. An empirical law describing heterogeneity in the yields of agricultural crops. *J. Agric. Sci.* **28**, 1–23. <https://doi.org/10.1017/S0021859600050516> (1938).
11. Hurst, H. E. The problem of long-term storage in reservoirs. *Int. Assoc. Sci. Hydrol. Bull.* **1**, 13–27. <https://doi.org/10.1080/02626665609493644> (1956).
12. Matia, K., Ashkenazy, Y. & Stanley, H. E. Multifractal properties of price fluctuations of stocks and commodities. *Europhys. Lett. (EPL)* **61**, 422–428. <https://doi.org/10.1209/epl/i2003-00194-y> (2003).
13. Taylor, L. R. Aggregation, variance and the mean. *Nature* **189**, 732–735. <https://doi.org/10.1038/189732a0> (1961).
14. Taylor, R. A. J. *Taylor's Power Law: Order and Pattern in Nature* (Academic Press, London, 2019).
15. Eisler, Z., Bartos, I. & Kertesz, J. Fluctuation scaling in complex systems: Taylor's law and beyond. *Adv. Phys.* **57**, 89–142. <https://doi.org/10.1080/00018730801893043> (2008).
16. Gerlach, M. & Altmann, E. G. Scaling laws and fluctuations in the statistics of word frequencies. *New J. Phys.* **16**, 113010. <https://doi.org/10.1088/1367-2630/16/11/113010> (2014).
17. Hanley, Q. S., Khatun, S., Yosef, A. & Dyer, R.-M. Fluctuation scaling, Taylor's law, and crime. *PLoS ONE* <https://doi.org/10.1371/journal.pone.0109004> (2014).
18. Wang, Y., Zhang, Q., Zhu, C., Hu, M. & Duong, V. Human activity under high pressure: a case study on fluctuation scaling of air traffic controllers communication behaviors. *Phys. A Stat. Mech. Appl.* **441**, 151–157. <https://doi.org/10.1016/j.physa.2015.08.040> (2016).
19. Choloniewski, J. *et al.* Temporal Taylor's scaling of facial electromyography and electrodermal activity in the course of emotional stimulation. *Chaos Solit. Fract.* **90**, 91–100. <https://doi.org/10.1016/j.chaos.2016.04.023> (2016).
20. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis (Version 3.3.4)* (Springer, New York, 2016).
21. Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P. & Deckmyn, A. maps: Draw Geographical Maps (Version 3.3.0). <https://CRAN.R-project.org/package=maps> (2018).
22. Watanabe, H., Sano, Y., Takayasu, H. & Takayasu, M. Statistical properties of fluctuations of time series representing appearances of words in nationwide blog data and their applications: an example of modeling fluctuation scalings of nonstationary time series. *Phys. Rev. E* **94**, 052317. <https://doi.org/10.1103/PhysRevE.94.052317> (2016).
23. Leban, G., Fortuna, B., Brank, J. & Grobelnik, M. Event registry: learning about world events from news. In *Proceedings of the 23rd International Conference on World Wide Web, WWW '14 Companion*, 107–110, <https://doi.org/10.1145/2567948.2577024> (Association for Computing Machinery, New York, NY, USA, 2014).
24. Choloniewski, J., Sienkiewicz, J., Leban, G. & Holyt, J. A. Modelling of temporal fluctuation scaling in online news network with independent cascade model. *Phys. A Stat. Mech. Appl.* **523**, 129–144. <https://doi.org/10.1016/j.physa.2019.02.035> (2019).
25. Ringnér, M. What is principal component analysis?. *Nat. Biotechnol.* **26**, 303–304. <https://doi.org/10.1038/nbt0308-303> (2008).
26. Wikipedia. Cannabis in Canada. [https://en.wikipedia.org/w/index.php?title=Cannabis\\_in\\_Canada&oldid=926464860](https://en.wikipedia.org/w/index.php?title=Cannabis_in_Canada&oldid=926464860) (2019).
27. Wikipedia. Homosexuality in India. [https://en.wikipedia.org/w/index.php?title=Homosexuality\\_in\\_India&oldid=925310796](https://en.wikipedia.org/w/index.php?title=Homosexuality_in_India&oldid=925310796) (2019).
28. Wikipedia. Recognition of same-sex unions in China. [https://en.wikipedia.org/w/index.php?title=Recognition\\_of\\_same-sex\\_unions\\_in\\_China&oldid=925518643](https://en.wikipedia.org/w/index.php?title=Recognition_of_same-sex_unions_in_China&oldid=925518643) (2019).
29. Wikipedia. Terrorism in Indonesia. [https://en.wikipedia.org/w/index.php?title=Terrorism\\_in\\_Indonesia&oldid=926611352](https://en.wikipedia.org/w/index.php?title=Terrorism_in_Indonesia&oldid=926611352) (2019).
30. Wikipedia. Macerata shooting. [https://en.wikipedia.org/w/index.php?title=Macerata\\_shooting&oldid=920525256](https://en.wikipedia.org/w/index.php?title=Macerata_shooting&oldid=920525256) (2019).
31. Ministère de l'Europe et des Affaires. France and China. <https://www.diplomatie.gouv.fr/en/country-files/china/france-and-china/> (2018).
32. Pike, N. Using false discovery rates for multiple comparisons in ecology and evolution. *Methods Ecol. Evol.* **2**, 278–282. <https://doi.org/10.1111/j.2041-210X.2010.00061.x> (2011).
33. Næs, R. & Skjeltorp, J. A. Order book characteristics and the volume-volatility relation: empirical evidence from a limit order market. *J. Financ. Mark.* **9**, 408–432. <https://doi.org/10.1016/j.finmar.2006.04.001> (2006).
34. Linsley, P. M. & Lawrence, M. J. Risk reporting by the largest UK companies: readability and lack of obfuscation. *Account. Audit. Account. J.* **20**, 620–627. <https://doi.org/10.1108/09513570710762601> (2007).
35. Dentcheva, D. & Stock, G. J. On the price of risk in a mean-risk optimization model. *Quant. Finance* **18**, 1699–1713. <https://doi.org/10.1080/14697688.2018.1436765> (2018).
36. Nawrot, M. P. *et al.* Measurement of variability dynamics in cortical spike trains. *J. Neurosci. Methods* **169**, 374–390. <https://doi.org/10.1016/j.jneumeth.2007.10.013> (2008).
37. Deco, G. & Hugues, E. Neural network mechanisms underlying stimulus driven variability reduction. *PLoS Comput. Biol.* <https://doi.org/10.1371/journal.pcbi.1002395> (2012).
38. Hamborg, F., Donnay, K. & Gipp, B. Automated identification of media bias in news articles: an interdisciplinary literature review. *Int. J. Digit. Libr.* **20**, 391–415. <https://doi.org/10.1007/s00799-018-0261-y> (2019).
39. Baly, R., Karadzhov, G., Alexandrov, D., Glass, J. & Nakov, P. Predicting Factuality of Reporting and Bias of News Media Sources. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, 3528–3539, <https://doi.org/10.18653/v1/D18-1389> (Association for Computational Linguistics, Brussels, Belgium, 2018).
40. Zhao, L., Sheppard, L. W., Reid, P. C., Walter, J. A. & Reuman, D. C. Proximate determinants of Taylor's law slopes. *J. Anim. Ecol.* **88**, 484–494. <https://doi.org/10.1111/1365-2656.12931> (2019).
41. Dinkov, Y., Ali, A., Koychev, I. & Nakov, P. Predicting the Leading Political Ideology of YouTube Channels Using Acoustic, Textual, and Metadata Information. In *Interspeech 2019*, 501–505, <https://doi.org/10.21437/Interspeech.2019-2965> (ISCA, 2019).
42. Stefanov, P., Darwish, K. & Nakov, P. Predicting the Topical Stance of Media and Popular Twitter Users. [arXiv:1907.01260](https://arxiv.org/abs/1907.01260) (2019).
43. Shu, K., Wang, S. & Liu, H. Beyond news contents: the role of social context for fake news detection. In *Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM'19*, 312–320, <https://doi.org/10.1145/3289600.3290994> (Association for Computing Machinery, New York, NY, USA, 2019).
44. Bovet, A. & Makse, H. A. Influence of fake news in Twitter during the 2016 US presidential election. *Nat. Commun.* **10**, 1–14. <https://doi.org/10.1038/s41467-018-07761-2> (2019).
45. Badawy, A., Lerman, K. & Ferrara, E. Who falls for online political manipulation? In *Companion Proceedings of The 2019 World Wide Web Conference on—WWW'19*, 162–168, <https://doi.org/10.1145/3308560.3316494> (ACM Press, San Francisco, USA, 2019).
46. Barrón-Cedeño, A., Jaradat, I., Da San Martino, G. & Nakov, P. Propopy: organizing the news based on their propagandistic content. *Inform. Process. Manag.* **56**, 1849–1864. <https://doi.org/10.1016/j.ipm.2019.03.005> (2019).
47. Ye, J. & Skiena, S. Mediarank: computational ranking of online news sources. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD'19*, 2469–2477, <https://doi.org/10.1145/3292500.3330709> (Association for Computing Machinery, New York, NY, USA, 2019).

48. Nadeem, M., Fang, W., Xu, B., Mohtarami, M. & Glass, J. FAKTA: an automatic end-to-end fact checking system. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations)*, 78–83, <https://doi.org/10.18653/v1/N19-4014> (Association for Computational Linguistics, Minneapolis, Minnesota, 2019).
49. Bahulkar, A., Szymanski, B. K., Chawla, N., Lizardo, O. & Chan, K. Influence of personal preferences on link dynamics in social networks. *Complexity* **1–12**, 2017. <https://doi.org/10.1155/2017/4543563> (2017).
50. Bahulkar, A., Szymanski, B. K., Chan, K. & Lizardo, O. Impact of Attributes on Group Formation. In *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, 1250–1257, <https://doi.org/10.1109/ASONAM.2018.8508658> (IEEE, Barcelona, 2018).
51. Nigam, A. *et al.* ONE-M: modeling the co-evolution of opinions and network connections. In: Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N. & Ifrim, G. (eds.) *Machine Learning and Knowledge Discovery in Databases*, vol. 11052, 122–140, [https://doi.org/10.1007/978-3-030-10928-8\\_8](https://doi.org/10.1007/978-3-030-10928-8_8) (Springer International Publishing, Cham, 2019).
52. Reuman, D. C., Zhao, L., Sheppard, L. W., Reid, P. C. & Cohen, J. E. Synchrony affects Taylors law in theory and data. *Proc. Natl. Acad. Sci.* <https://doi.org/10.1073/pnas.1703593114> (2017).
53. James, C., Azaele, S., Maritan, A. & Simini, F. Zipfs and Taylors laws. *Phys. Rev. E* **98**, 032408. <https://doi.org/10.1103/PhysRevE.98.032408> (2018).
54. Sakoda, G., Takayasu, H. & Takayasu, M. Tracking Poisson parameter for non-stationary discontinuous time series with Taylors abnormal fluctuation scaling. *Stats* **2**, 55–69. <https://doi.org/10.3390/stats2010005> (2019).
55. Kendal, W. S. & Jørgensen, B. Taylors power law and fluctuation scaling explained by a central-limit-like convergence. *Phys. Rev. E* **83**, 066115. <https://doi.org/10.1103/PhysRevE.83.066115> (2011).
56. Liu, X., Zhu, X.-H., Qiu, P. & Chen, W. A correlation-matrix-based hierarchical clustering method for functional connectivity analysis. *J. Neurosci. Methods* **211**, 94–102. <https://doi.org/10.1016/j.jneumeth.2012.08.016> (2012).
57. Pearson, K. Liii. On lines and planes of closest fit to systems of points in space. *J. Sci.* **2**, 559–572. <https://doi.org/10.1080/14786440109462720> (1901).
58. Bollen, J., Van de Sompel, H., Hagberg, A. & Chute, R. A principal component analysis of 39 scientific impact measures. *PLoS ONE* **4**, 1–11. <https://doi.org/10.1371/journal.pone.0006022> (2009).
59. Gajewski, L. G., Choloniewski, J. & Holyst, J. A. Key courses of academic curriculum uncovered by data mining of students grades. *Acta Physica Polonica A* **129**, 1071–1076. <https://doi.org/10.12693/APhysPolA.129.1071> (2016).
60. Sienkiewicz, J., Soja, K., Holyst, J. A. & Słot, P. M. A. Categorical and geographical separation in science. *Sci. Rep.* **8**, 8253. <https://doi.org/10.1038/s41598-018-26511-4> (2018).
61. Pedregosa, F. *et al.* Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
62. Anderson, E. *et al.* Lapack: a portable linear algebra library for high-performance computers. In *Proceedings of the 1990 ACM/IEEE Conference on Supercomputing*, 2–11 (IEEE Computer Society Press, 1990).
63. Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **47**, 583–621. <https://doi.org/10.1080/01621459.1952.10483441> (1952).
64. Terpilowski, M. scikit-posthocs: pairwise multiple comparison tests in python. *J. Open Source Softw.* **4**, 1169. <https://doi.org/10.21105/joss.01169> (2019).

## Acknowledgements

J.Ch., J.S. and J.A.H. are thankful to Slovenian Press Agency STA and especially to Aljoša Rehar for the hospitality during secondments in the framework of RENOIR Project. The political bias of the news outlets was courtesy of Dave van Zart, the owner of Media Bias/Fact Check ([www.mediabiasfactcheck.com](http://www.mediabiasfactcheck.com)). Helpful insights from Eduardo Altmann, Stanisław Drożdż, and Przemysław Biecek are gratefully acknowledged. The work was partially supported as RENOIR Project by the European Union Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie Grant Agreement No. 691152 and by Ministry of Science and Higher Education (Poland), Grant Nos. 34/H2020/2016, 329025/PnH/2016 and by National Science Centre, Poland Grant No. 2015/19/B/ST6/02612. J.A.H. was partially supported by the Russian Scientific Foundation, Agreement #17-71-30029 with co-financing of Bank Saint Petersburg and by POB Research Centre Cybersecurity and Data Science of Warsaw University of Technology within the Excellence Initiative Program—Research University (IDUB).

## Author contributions

J.C.—conception, design, data acquisition, data analysis, interpretation, software, draft, revision; J.S.—design, interpretation, draft, revision; N.D.—interpretation; G.L.—data acquisition; M.T.—revision; J.A.H.—conception, design, interpretation, draft, revision, supervision.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-77660-4>.

**Correspondence** and requests for materials should be addressed to J.A.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020