Received: 10 February 2021 | Revised: 26 October 2021 | Accepted: 6 November 2021

DOI: 10.1111/1468-5973.12385

ORIGINAL ARTICLE

WILEY

Investigating dynamic relations between factual information and misinformation: Empirical studies of tweets related to prevention measures during COVID-19

Yan Wang[1] | Shangde Gao[2] | Wenyu Gao[3]

[1]Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, University of Florida, Gainesville, Florida, USA

[2]Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, College of Design, Construction and Planning, University of Florida, Gainesville, Florida, USA

[3]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, 02115, USA

Correspondence

Yan Wang, Department of Urban and Regional Planning and Florida Institute for Built Environment Resilience, University of Florida, P.O. Box 115706, Gainesville, FL 32611, USA.
Email: yanw@ufl.edu

Funding information

National Science Foundation, Grant/Award Number: 2028012

Abstract

During COVID-19, misinformation on social media has affected people's adoption of appropriate prevention behaviors. Although an array of approaches have been proposed to suppress misinformation, few have investigated the role of disseminating factual information during crises. None has examined its effect on suppressing misinformation quantitatively using longitudinal social media data. Therefore, this study investigates the temporal correlations between factual information and misinformation, and intends to answer whether previously predominant factual information can suppress misinformation. It focuses on two prevention measures, that is, wearing masks and social distancing, using tweets collected from April 3 to June 30, 2020, in the United States. We trained support vector machine classifiers to retrieve relevant tweets and classify tweets containing factual information and misinformation for each topic concerning the prevention measures' effects. Based on cross-correlation analyses of factual and misinformation time series for both topics, we find that the previously predominant factual information leads the decrease of misinformation (i.e., suppression) with a time lag. The research findings provide empirical understandings of dynamic relations between misinformation and factual information in complex online environments and suggest practical strategies for future misinformation management during crises and emergencies.

KEYWORDS

COVID-19, crisis informatics, factual information, misinformation, public health, social media, supervised machine learning

1 | INTRODUCTION

Crisis communication plays a critical role in organizing effective responses and mitigating the impacts of crises (Clark-Ginsberg & Petrun Sayers, 2020). It intends to form people's correct perceptions about preventative measures (Qiu & Chu, 2019) and motivate risk mitigation behaviors in response to crisis events (Gilk, 2007; Utz et al., 2013). Social media platforms facilitate crisis communication by allowing people to seek, interpret, and disseminate information timely during crisis events (Silver & Andrey, 2019). Studies on crisis informatics that leverage social media data have been burgeoning across hazard types including natural hazards (Niles et al., 2019), public health crises (Yu et al., 2020), and extreme events (Stieglitz et al., 2018). In addition to the multiple benefits of using social media in crisis communication, concerns have also been raised for its undesirable role in propagating misinformation (i.e., inaccurate or misleading information) during these crises and emergencies (Vosoughi et al., 2018).

During COVID-19, social media has been ignited with diverse information. The increasing rate of reported incidents along with massive, related dialog has triggered divergent reactions and interactions across stakeholders at various levels (Shimizu, 2020; Wang et al., 2021). Specifically, under the social distancing policy, more people have turned to social media for support (Nabity-Grover et al., 2020). However, the credibility of social media information is worrisome as misinformation spreads widely and quickly (Depoux et al., 2020; Pulido et al., 2020), which posed severe challenges to public health. Based on existing definitions of online misinformation in previous public health studies (Barua et al., 2020; Loomba et al., 2021; Pennycook et al., 2021; Wu et al., 2019), we define misinformation as false information that is against the relevant scientific facts related to the two measures' effects in preventing the infection of SAR-CoV-2. After the worldwide outbreak of SAR-CoV-2 in 2019, 24.8% of tweets about COVID-19 contained misinformation (Kouzy et al., 2020). Unlike factual information, which matched scientific facts towards the correct prevention measures of the crises (Castillo et al., 2011), most misinformation contains false content towards the prevention measures and produces misperceptions about disease prevention (van der Meer & Jin, 2020).

Misinformation during public health crises is harmful because it misdirects people's response behaviors while the effectiveness of intervention policies depends heavily on individuals' response behaviors. Collective individuals' prevention strategies, such as wearing facemasks and social distancing, have the potential to reduce the risk of infection (Chu et al., 2020; Lewnard & Lo, 2020). However, Individuals' crisis response behaviors can be significantly affected by information obtained from the Internet and social media (Swire-Thompson & Lazer, 2020). Additionally, some factors, such as recommendation algorithms and bots, have made misinformation widely propagated in the digital environments (Orabi et al., 2020; Zhang & Ghorbani, 2020). Individuals misled by such misinformation may avoid following correct recommendations and put their health at high risk (Earnshaw & Katz, 2020). For example, widespread online misinformation about coronavirus treatment "injecting disinfectant" has caused 30 poisoning cases in New York City within 18 h (Slotkin, 2020). Specifically, misinformation has a severe impact on vulnerable groups during the COVID-19: mistrust and lack of access to factual information sources have made different vulnerable groups easily to be affected by misinformation (Clark-Ginsberg & Petrun Sayers, 2020). Because of the vast spread and negative societal impacts of misinformation during emergency response, it is urgent to formulate effective strategies to suppress misinformation on social media platforms.

Previous literature in the domain of crisis informatics has proposed several strategies for combating misinformation on social media, including checking information authenticity (Safieddine et al., 2016), controlling bot accounts (Shao et al., 2018), tracking sources of misinformation (Jang et al., 2018), identifying misinformation topics (Vicario et al., 2019), broadening exposure to diverse views (Wang & Song, 2020), and providing news and science literacy education, such as guidelines of social media uses in crisis events (Kaufhold et al., 2019; Trethewey, 2020; Tully, Bode, et al., 2020). The first five strategies can be implemented by social media companies and domain experts, while the last one puts the onus on the public and authoritative agencies. However, the effectiveness of fact-checking and bot control has been limited to suppressing *preknown* misinformation and cannot limit the production or sharing of misinformation that has not been detected (Shao et al., 2018). Controlling bot accounts also cannot mitigate the misinformation generated and shared by human accounts (Silva et al., 2020). In the study field of public health, fact-checking, and literacy education (Walter et al., 2021) have been used as the main strategies to suppress health misinformation on social media. Multiple methods were utilized for the misinformation detection, such as training a Random Forest classifier for examining whether the temporal trend of online information dissemination is similar to the trend of fake news (Previti et al., 2020), and comparing text similarity between reported misinformation and social media posts using a neural network model (Yu et al., 2017).

In comparison with the detection-based "reactive" strategies, literacy education (e.g., news and information literacy) has a greater potential to suppress misinformation "proactively" (Tully, Vraga, et al., 2020). The effectiveness of literacy education has been notable, and experiments have shown that the provision of accurate information made about 20% of the experiment participants change their misperceptions of the research topics (Vraga et al., 2020). Additionally, literacy education is effective in reducing the public's ignorance and misconceptions on other topics, such as climate change (Cook et al., 2014), and helps individuals judge the truthfulness of information (Kahne & Bowyer, 2017). One example of literacy education used on social media platforms is to disseminate factual information about crises and provide correct strategies for crisis prevention (Almaliki, 2019). This approach was also found to correct people's mistrust of misinformation effectively (Vraga et al., 2020). Specifically, in the public health domain, this strategy has been used to suppress misinformation, especially in vaccination promotion and SARS-CoV-2 prevention (Chen et al., 2020; Danielson et al., 2019). The widely generated and disseminated messages containing factual information become "predominant" when the proportion of tweets containing factual information is higher than 50% meaning social media users can be exposed to a higher percentage of factual information.

However, little research has investigated the temporal correlation between misinformation and factual information during crises empirically, and the existing research remains insufficient on whether predominant factual information can effectively suppress misinformation on social media platforms. None has used longitudinal social media data to investigate the temporal relationship between misinformation and factual information quantitatively. It is unclear how effectively the previously predominant factual information (e.g., increased number or proportion) can reduce the overall volume and proportion of misinformation on social media during crises. Considering the existing research gaps and the importance of studying social media misinformation combatting strategies, this manuscript has two primary questions.

#1: What is the temporal relation between the daily volume/ proportion of tweets that contained factual information and misinformation for individual topics of prevention measures during COVID-19?

#2: Can previous predominant factual information suppress misinformation on Twitter during COVID-19?

We chose two topics, that is, "wearing masks" and "social distancing," for detailed empirical investigations (Chu et al., 2020; Lewnard & Lo, 2020) due to their much-identified misinformation on Twitter, which hindered people from following preventative measures (Krause et al., 2020). Recent medical studies have shown the potential positive influence of the two prevention measures on mitigating the risks of the SARS-CoV-2's infection (Chu et al., 2020; Lewnard & Lo, 2020). Thereby, particularly for the two public health topics in the US pandemic response context, our classification criteria are built based on *the potential public-health consequences of the two information categories*. We regard tweets as *factual information* if they regard the two critical prevention measures as effective in mitigating COVID-19 infection, endorse relevant credible information and affirm the negative consequences of not following them; and as *misinformation*, if they oppose or manipulate messages from credible sources (e.g., public health agencies and authorities), manipulating contents relevant to negative impacts of these measures, and contain verified falsehoods (van der Meer & Jin, 2020; Wilson & Starbird, 2020). If people receive and follow the misinformation regarding COVID-19 prevention, they may behave inappropriately in response to the COVID-19 or to share such opinions on social media platforms, and their health status would be highly risky (Earnshaw & Katz, 2020). Acting on misperceptions of these measures, such as not wearing a mask or socially distancing in public places, could potentially accelerate the spread of the virus; one experiment showed that a lack of appropriate prevention measures nearly doubled the number of infections (Lewnard & Lo, 2020). We utilized key-expressions and support vector machine (SVM) to extract relevant tweets from the collected data and categorized them into those containing (a) factual information and (b) misinformation under each topic. We generated the series about the daily volume and proportion of these two information categories, then conducted cross-correlations between the time series of two information categories. The research findings can provide strategies for combating social media misinformation during future public health crises and other extreme events.

2 | DATA COLLECTION AND METHODS

2.1 | Case description and data collection

This study focused on the misinformation about SARS-CoV-2 prevention measures on Twitter and evaluated the influence of factual information on the spread of misinformation in the United States. The study period covers 89 days, from April 3 to June 30, 2020. We chose this period because U.S. CDC published the announcement that encouraged the public to wear masks and keep social distancing (CDC, 2020), and the number of US cases of coronavirus surpassed three million during the second wave of the pandemic (Dong et al., 2020). During this time, a large volume of misinformation spread widely and caused an infodemic (Hernández-García & Giménez-Júlvez, 2020). The misinformation about the prevention measures, such as messages that manipulate agencies' risk communication regarding the effectiveness of the two recommended prevention measures and posts that encourage the public to engage in risky behaviors (Pennycook et al., 2020), may hinder the use of proper prevention measures. Meanwhile, factual information was also disseminated to inform individuals of the proper response measures and to suppress the misinformation.

Over the 3 months, we collected tweets with keywords "coronavirus" and "covid" using an open Twitter streaming API (Twitter, 2021) and retrieved 22,111,831 English tweets from the raw data. In addition, we used Hydrator (Documenting the Now, 2020) to extract the full text of each tweet before further text mining as tweets collected from the streaming API mostly were truncated. With the basic data sets, we conducted analyses in three steps using a set of text mining and machine learning methods: (i) retrieving relevant tweets, (ii) classifying tweets as containing misinformation and as containing factual information, and (iii) investigating the cross-correlation between time series of the two information categories (Figure 1).

2.2 | Retrieving relevant tweets for the two preventative measures

We conducted two steps to extract tweets that are relevant to each topic, including initial keyword-based filtering and supervised classification using SVM. We define "*relevant tweets*" as (i) tweets that directly expressed opinions/beliefs on the studied topics, such as "wearing masks is (not) useful"; (ii) tweets involving suggestions, policies, or opinions in a certain area; or (iii) tweets that endorsed suggestions, policies, and opinions about the preventative measures (including known/demonstrable falsehoods and factual information). First, we used key-expression filtering to retrieve tweets containing the keywords and expressions for each topic (see Table 1). To generate the final keyword list, we first collected key expressions about the topics from the websites of the US CDC (2020) and the WHO (2020). Then we used both the keywords (e.g., "wearing masks") and their expression patterns to collect the potentially relevant tweets. For example, we used the expression pattern "'mask' + 'protect others'" to retrieve tweets containing both "mask" and "protect others," so both "masks can effectively protect others" and "to protect others, masks are necessary" would be retrieved by this expression pattern. Using this list of key patterns and expressions, we retrieved 3000 sample tweets from the collected data and further enriched this list by adding the high-frequency expressions from the retrieved tweet texts.

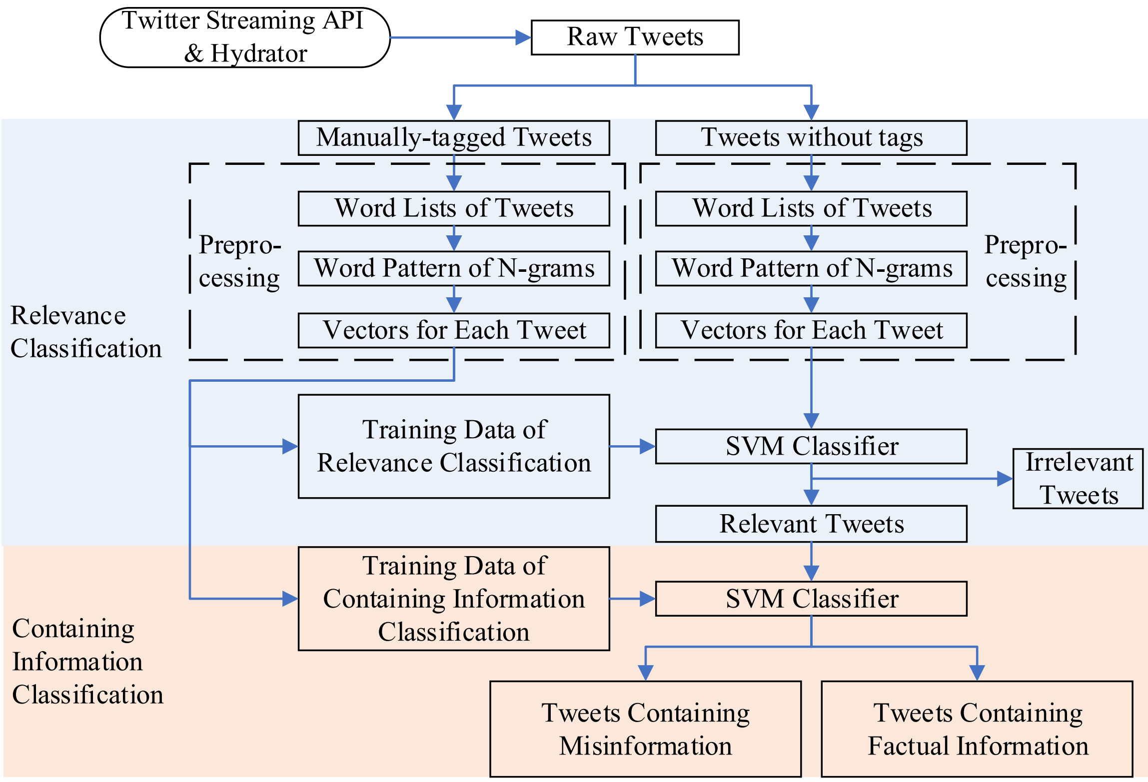

FIGURE 1 Schematic process of tweets analysis

TABLE 1 Final keywords list for each topic and number of filtered tweets

Topics	Key-expressions	Data volume
Mask	'wear a mask', 'wearing a mask', 'wearing face mask', 'wear face masks', 'wear your mask', 'mask-wearing could prevent', 'mask' + 'second waves study', 'mask in public', 'mask protects you', 'mask' + 'please please please', 'mask' + 'prevent the spread', 'mask' + 'prevent you from', 'mask' + 'slow the spread', 'use of facemask', 'mask won't help', 'masks at all times', 'masks are useless', 'mask is useless', 'face coverings', 'facemask use', 'healthy people', 'masks can', 'N95 masks', 'prevent COVID-19', 'please wear', 'mask' + 'protect others', 'mask' + 'protect themselves', 'mask' + 'protect yourself', 'mask' + 'protects you', 'wear mask', 'wearing masks', 'need mask', 'wore mask', 'no mask', 'mask' + 'effectiveness', 'mask' + 'efficiency', 'mask' + 'compulsory', 'WearAMask', 'mask' + 'reduce onward transmission'.	424,566
Social Distancing	'social distancing', '2 arms', '6 feet', '6-foot distance', 'avoid crowded places', 'avoid crowds', 'avoid gathering', 'avoid hugging', 'avoid kissing', 'avoid pooled rides', 'close contact', 'common areas', 'create space between others', 'face-to-face contact', 'increase space between individuals', 'keep a safe space', 'keep distance', 'keep space', 'limit contact', 'limit errands', 'physical distance', 'physical guide', 'safe social activities', 'social distance', 'stay apart', 'stay distanced', 'physical distancing', 'around others'.	100,695

However, the key expressions could still retrieve irrelevant tweets. For example, the tweet "Coronavirus: 3M to Produce 35,000,000 Respirator Masks a Month in the U.S." contains "coronavirus" and "mask", but it is about mask production instead of behaviors of wearing masks, so we regard it as irrelevant. Other examples of irrelevant tweets include: "Coronavirus, social distancing, Floyd protests|Homeland Security Newswire." This tweet contains keywords about social distancing (i.e., "social distancing") but the three rules about relevance classification deem it irrelevant because it does not contain opinions or suggestions about social distancing.

To overcome the limitations of key expressions in retrieving relevant tweets, considering the high-level performance of SVM in text classification, of which the accuracy was higher than 90% (Gopi et al., 2020; Liu et al., 2013), we conducted the second step of relevance classification using an SVM-based classifier. We set the parameters of SVM models by default. This process was illustrated in Figure 2, showing how the example tweet was transformed to the vector for relevance classification. The training data sets were randomly extracted from the raw data set over the whole study period. As shown in Figure 2, tweets were firstly cleaned by removing all stopwords such as "the", "is", "which" and meaningless characters like emojis, "@", and "#". The cleaned text was tokenized to unigrams, bigrams, and trigrams using the NLTK Tokenizer. Then, we normalized the terms by transforming them to lower case and stemming

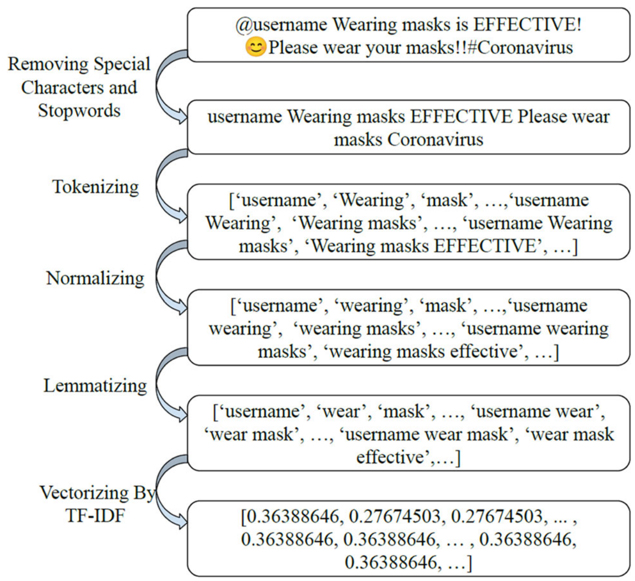

FIGURE 2 Preprocessing steps for tweets

TABLE 2 Training and testing data set's sizes and classification performance for relevant tweets

	Wearing mask	Social distancing
Training data set	1192	2099
Testing data set	300	300
Accuracy	0.8833	0.9133
Precision	0.9099	0.9462
Recall	0.9380	0.9535
Mean accuracy	0.8879	0.9243
Standard deviation	0.0194	0.0119

these terms. For example, "EFFECTIVE" was normalized and stemmed to "effect". Finally, we used Term Frequency-Inverse Document Frequency (TF-IDF) (Ramos, 2003) to vectorize the list of tokens. The output vectors of the training data were used to train the SVM-based classifier.

Two annotators (i.e., two authors) labeled the tweets in the training data independently regarding (i) whether the tweets were relevant to wearing masks or social distancing, and (ii) whether the tweet contained factual information or misinformation (for Section 2.3) based on the aforementioned criteria and definitions. The annotation outcomes were compared, if some labels were not consistent, the annotators would discuss and determine the final labels together. The labeled tweets were then utilized to train the SVM classifier. We also utilized *k-folder* cross-validation to evaluate the performance of the SVM model by dividing each training data set into 10 subsets of the same size. The outcome of the training model and the cross-validation are shown in Table 2. The row of "mean accuracy" represents the average level of accuracies of training processes with the 10 subsets of training data. The outcome indicates that the trained SVM model is accurate and reliable.

2.3 | Classifying tweets containing factual information and misinformation

After manually annotating tweets containing the two categories of information based on the criteria described in Section 1, we trained another SVM model to classify tweets under each topic over the 4.5 months for further temporal correlation analyses. Table 3 shows the criteria and examples of factual information and misinformation when annotating the tweets. The training outcome of the SVM-based classifier is shown in Table 4. We utilized the same *k-folder* cross-validation to evaluate the performance of the SVM model. The results are shown in Table 4. The row of "mean accuracy" represents the average level of accuracy of the training process with the ten subsets of training data. The outcome indicates that the SVM model trained for classifying information categories is reliable.

2.4 | Cross-Correlation analysis of factual and misinformation time series

Time-series analyses have been widely used in analyzing data and information mined from social media platforms (e.g., Wang & Taylor, 2018). We employed a cross-correlation analysis of two time series to identify lags (h) of the predominant daily volume/proportion of factual information (f_{t+h}/ fp_{t+h}) that might be useful predictors of daily volume/proportion of misinformation (m_t/mp_t) for tweets relevant to "wearing masks" and "social distancing" topics separately. For example, when one or more f_{t+h}, with h negative, are predictors of m_t, it is sometimes said that f leads m; when one or more, f_{t+h} with h positive, are predictors of m_t, it is sometimes said that f lags m.

The cross-correlation analysis is performed based on the plot of cross-correlation function (CCF) (e.g., Figures 5 and 6) between the time series of factual information and misinformation (i.e., daily tweet count and daily proportion) for each topic. Values of the x-axis (i.e., time lags) of the peaks in CCF plots indicate potential significant time lags on the predictor (i.e., factual information). Before running CCF, a prewhitening procedure using an autoregressive integrated moving average (ARIMA; Box et al., 2015) model is used to remove the common trends of time series of two information categories and to help better interpret the CCF. The final model is constructed with the final chosen lags based on the CCF plot and the ARIMA model.

To perform prewhitening, we fit the ARIMA model to the predictor (f_{t+h}/ fp_{t+h}) and use the fitted model structure to filter out the response (m_t/mp_t). The ARIMA also requires stationarity (i.e., the mean and variance do not change over time). We conducted the Augmented Dickey–Fuller (ADF) test (Said & Dickey, 1984) and the analysis is performed using the *adf.test* function from the "tseries" R package (Trapletti & Hornik, 2020). If p value of the ADF test is less than 0.05, the time series is stationary; if p value of ADF test is equal

TABLE 3 Annotation criteria and example tweets of factual information and misinformation

Information category	Criteria	Example tweets for "wearing masks"	Example tweets for "social distancing"
Factual information	Containing information about the (potential) positive influence of wearing masks (or social distancing) in reducing the infection risk of SARS-CoV-2	"The study found that if people wear masks whenever they are in public it is twice as effective at reducing the R value than if masks are only worn after symptoms appear"	"Coronavirus plea from Johns Hopkins: please take social distancing seriously to save lives."
	Endorsing information about the prevention measures from credible sources (e.g., public health agencies)	"@CDCgov The vulnerable are not 100% protected from COVID-19 even if they just stay home because those they live with can bring it home to them when they buy groceries. My article emphasizes the importance of wearing masks in public."	"Important information on COVID-19 from @cdcgov Wash your hands, stay at home if you're feeling ill. Practice social distancing and avoid large crowds to stop the spread!!"
	Describing the potential public health consequence(s) of not conducting wearing masks (or social distancing)	"If masks have been found to save lives, not wearing a mask does the opposite, You could literally be killing people by refusing to take this simple step to protect others."	"The longer we do not comply with social distancing, then the longer we will have to do it."
Misinformation	Manipulating information about the potential negative influence of wearing masks (or social distancing) on preventing SARS-CoV-2.	"A Surgical Mask Won't Protect You From #Coronavirus."	"Social distancing won't stop 'accelerating' coronavirus pandemic, WHO warns."
	Manipulating information about health authorities' messages to opposing the effectiveness of these measures.	"An email proves Fauci knew masks were ineffective for COVID-19."	"Bill Gates and White House health advisor Dr Anthony Fauci violating social distancing norms and not wearing masks."
	Containing other falsehood information verified by the fact-checking websites (e.g., Poynter)	"Criminals give contaminated masks from door-to-door to make people fall asleep and rob them."	"There was no real scientific basis for believing that" social distancing would be necessary, "since it had never been studied."

TABLE 4 Training and testing datasets and classification performance for information categories

	Wearing mask	Social distancing
Training data set	1684	941
Testing data set	300	300
Accuracy	0.9641	0.8333
Precision	0.9447	0.8507
Recall	0.9305	0.9495
Mean accuracy	0.9404	0.8210
Standard deviation	0.0148	0.0108

to or larger than 0.05, the time series is not stationary. As the function cannot pass missing values, we imputed missing data using *Kalman* smoothing (Bishop & Welch, 2001; Harvey, 1990; Grewal et al., 2020), a nonparametric method without model assumptions and is conducted purely from the data so the imputed values are stick to the observed data. This process employed a *na_kalman* function from the package "imputeTS" (Moritz & Bartz-Beielstein, 2017). If the time series is stationary, the ARIMA model is fitted using the *sarima* function from the "astsa" package (Stoffer, 2020). To be noted, the functions from the "astsa" package can handle the missing values themselves. They employ the basic ACF and ARIMA functions, where missing values are dealt with by Intervention Detection and the model structure (Kalman filtering), respectively.

To select the best ARIMA and final cross-correlation models, we start the fitting with all the candidate time lags, then use backward selection. The ARIMA model selection criteria are based on Akaike information criterion (AIC; Akaike, 1998), and Bayesian information criterion (BIC; Schwarz, 1978) (For AIC and BIC, the smaller the better), and ensure the residuals to be independent (ACF around zero) and random (Ljung-Box test with $p > 0.05$) (Ljung & Box, 1978). The final cross-correlation model is linear, so our selection is based on adjusted R^2 (Draper & Smith, 1998), ensuring the residuals to be independent (ACF around zero) in model validation. Essential time series plots including CCF, autoregressive function (ACF), and partial autoregressive function (pACF), were made using the R built-in package "stats." All the statistical analyses were performed in R language (R Core Team, 2020).

3 | RESULTS

3.1 | Tweets containing factual information and misinformation

We utilized the key-expressions (Table 1) and the trained SVM-based classifiers to retrieve "relevant" tweets for case topics (i.e., wearing masks and social distancing) from the tweets collected from April 3 to June 30 (using methods described in Sections 2.2 and 2.3). We have 12 days with missing data from April 21–28 and June 6–9 due to the tropical-storm-incurred power outages in Florida and computer resetting, which has a very minor impact on the following analyses based on 3-month data. The changes in the daily volume of tweets that contain misinformation and factual information are plotted in Figure 3.

Based on the daily data volume of the classified tweets (Figure 3), we found that tweets relevant to "wearing masks" kept growing over the period April 3–May 30, potentially caused by the increasing public attention on the reasonability and implementation of wearing masks. The second period of growth might be intensified by the event of George Floyd on May 25 (Dave et al., 2020), when people protested for the policemen's violence in Minneapolis. Additionally, as the number of US cases surpassed 100,000 on May 28 (Dong et al., 2020), CDC highly recommended individuals wearing masks in public space, which might have contributed to the increased discussions as well. In comparison, the number of "social distancing" tweets did not change drastically and grew from April 3 to June 6 steadily, then decreased gradually. Based on the health literature (e.g., Lewnard & Lo, 2020), social distancing was proved as an effective strategy and continuously promoted by public health agencies, and the discussion of social distancing on Twitter was growing from early April to early June. The prevention measure was promoted by the persistent recommendation of the related public health policies (Chui et al., 2020), but the popularity level of the discussion was not as high as "wearing masks." To understand the proportion of factual information and misinformation in the two data sets containing topic-relevant tweets, we also calculated the daily proportion for each topic (Figure 4).

3.2 | Cross-correlation between "wearing masks" factual information and misinformation time series

To explore the relation between misinformation and factual information for the "wearing masks" topic over time, we employed

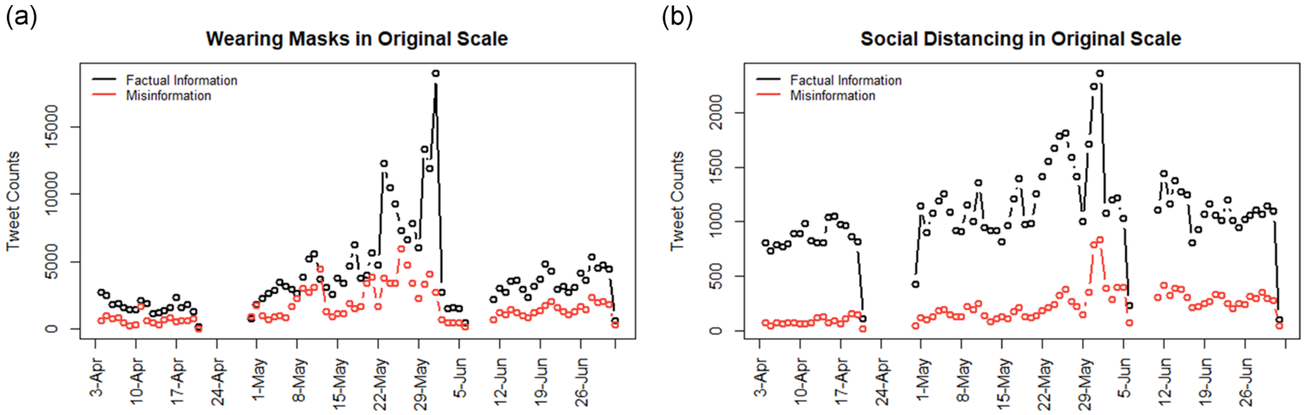

FIGURE 3 Daily number of tweets containing misinformation and factual information (a: wearing masks; b: social distancing)

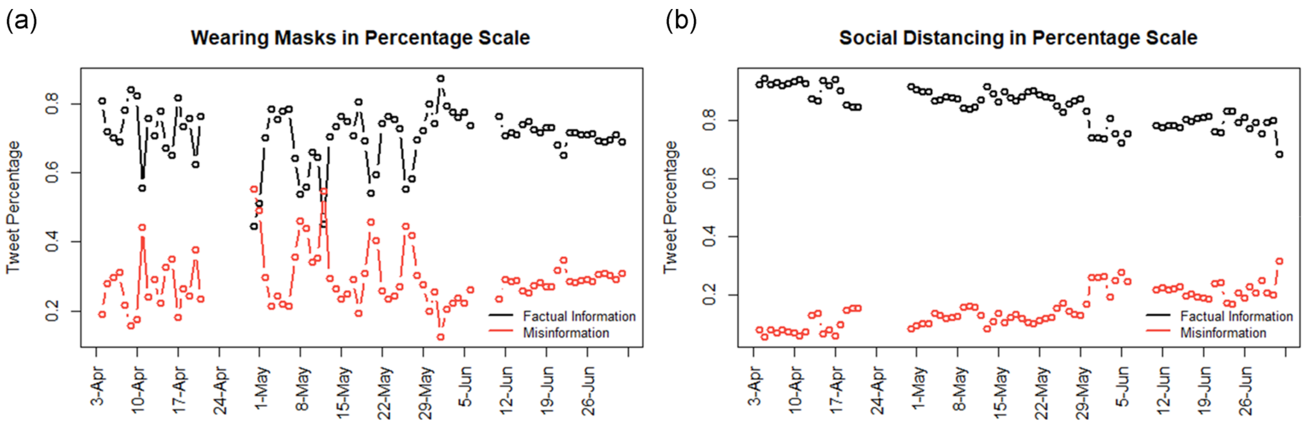

FIGURE 4 Daily percentage of tweets containing misinformation and factual information (a: wearing masks; b: social distancing)

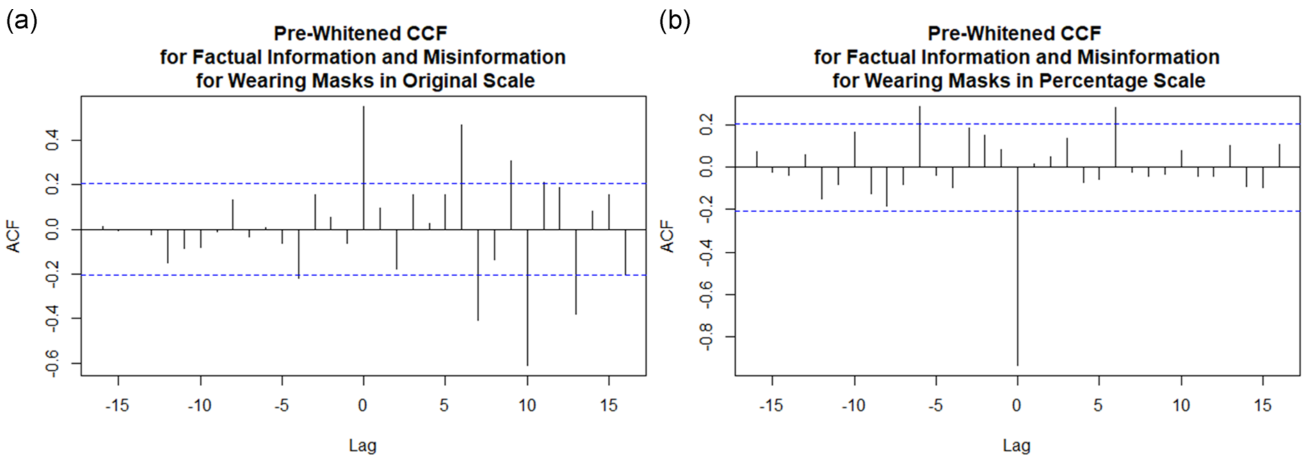

FIGURE 5 Cross-correlation of m_t and f_t of wearing masks after prewhitening based on different lags (a: original scale; b: percentage scale)

cross-correlation analysis of time series in both original-number and percentage scales, including (a) daily tweet number containing misinformation (m_t) and factual information (f_t); and (b) daily proportion of misinformation (mp_t) and factual information (fp_t). The initial CCF plots are included in Figure S1 in the Supporting Information Material for time series in both scales showed unclear peaks of time lags, so prewhitening is conducted. The CCF plots after the prewhitening process are in Figure 5.

Specifically, for "wearing mask," the ADF test ($p = .495 > .05$) indicates that the time series (f_t) is not stationary (Fuller, 2009), so we took the first-order difference of predictor between the daily values of adjacent dates ($f_t - f_{t-1}$). The purpose is to remove the trend and seasonality in the time series which stabilizes the mean of the time series (Rasheed, 2020). Then the ADF test ($p < .01$) shows the time series of the predictor's first-order difference is stationary. Thus, we considered integration with Order 1. Our final ARIMA model chose AR with Order 6, because of the integration of Order 1, we considered time lag ($t = 1, 2, 3, 4, 6, 7$) eliminating $t = 5$ after model selection (see Section 2). The final cross-correlation model is listed in Equation (1) with detailed coefficients and significance levels in Table S1 and satisfactory ACF and pACF tests for model validation in Figure S3. The adjusted R^2 for the final model is 0.4879 with $p = .0007726$.

For fp_t and mp_t, after imputing missing values, the ADF test has a $p < .05$, indicating stationary. The prewhitened CCF plot (Figure S1b) indicated potential important time lags (h) at 0 and −12. Based on the fitting outcomes of the ARIMA model, time lags at −1 and −12 were also considered. Notably, time lag at 0 is omitted due to the collinearity with the response variable. Thus, we chose the final model based on adjusted R^2 and ACF tests (Figure S4) and the model is listed in Equation (2) with detailed coefficients and significance in Table S2 and satisfactory ACF and pACF tests for model validation (Figure S4). For the final model, the adjusted R^2 is .1824, and the $p = .0007726$.

Based on the final fitted cross-correlation models (Equations 1 and 2) and the significance of coefficients in Tables S1 and S2, we find evidence that predominant factual information (i.e., tweet number and percentage) leads to the decrease of misinformation significantly when lag (h) is −1 (1 day) or 7 (1 week). However, we also find that factual tweets from the previous 3 days and the same day have a positive significant correlation with the number of tweets containing misinformation; the number of misinformation tweets can also negatively impact the number of factual tweets in the future with a time lag at 7. Additionally, the number of tweets containing misinformation is also positively related to the time (t) significantly. For the daily percentage of misinformation tweets, previous dominant factual information in percentages with a time lag at −1 can all significantly decrease the percentage of misinformation for wearing masks tweets.

$$\begin{cases} m_t = 982.691 - 0.09132 m_{t-3} + 0.36578 f_t + 0.12195 f_{t-3} - 0.046 f_{t-4} \\ \quad -0.16853 f_{t-7} - 0.04395 f_{t+7} + 5.67228 t + \epsilon_t \\ \quad \epsilon_t \sim N(0, \sigma^2) \end{cases} \tag{1}$$

$$\begin{cases} mp_t = 0.6484 - 0.377 fp_{t-1} - 0.1379 fp_{t-12} + \epsilon_t \\ \quad \epsilon_t \sim N(0, \sigma^2) \end{cases} \tag{2}$$

3.3 | Cross-correlation between "social distancing" factual and misinformation time series

Similarly, we conducted cross-correlation analyses and an ADF test for the time series of "social distancing" tweets containing misinformation and factual information. The CCF plots for the two scales (daily number and proportion) shown in Figure S2 indicate that prewhitening is necessary. The CCF plots after the prewhitening process are in Figure 6.

For original tweet number under each information category (m_t and f_t), the ADF test on c_t has a p *value* of .3532 (nonstationary).

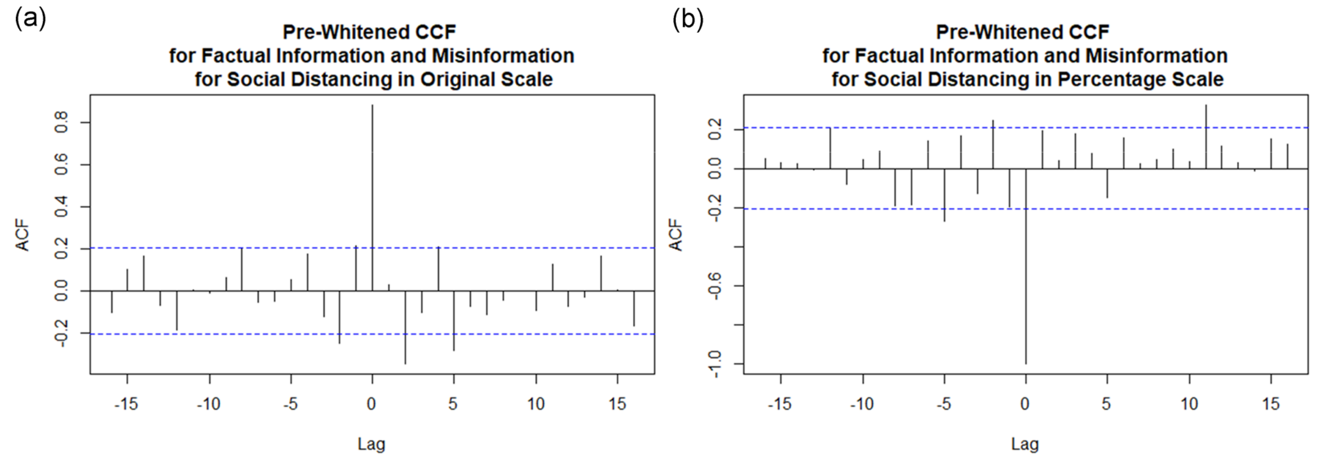

FIGURE 6 Cross-correlation of m_t and f_t of social distancing after prewhitening based on different lags (a: original scale; b: percentage scale)

After taking the first-order difference of the predictor, the p value of the ADF test is smaller than .01. Thus, an order (1) integration is considered. The $ARIMA(12,1,0)$ model omitting orders 1, 8, 9, and 10 are used. The preliminary model based on adjusted R^2 is listed in Equation (3). Coefficients and significance in Table S3, the adjusted R^2 is 0.7671, and the model $p = 1.922e-06$.

For the daily proportion of each information category (mp_t and fp_t), the ADF test has p value of .5337 after imputation, and $p < .01$ after taking first-order differences (indicating stationarity). Thus, Order 1 integration is considered. An $ARIMA(19,1,0)$ keeping orders at 1 to 4, and 19 is chosen. The prewhitened CCF plot is shown in Figure 6. Lags to be considered include −2, −5, −12, and 11. The final model considering both adjusted R^2 and p values is listed in Equation (4) with coefficients and significance in Table S4 and satisfactory ACF and pACF tests for model validation (Figure S6). The adjusted R^2 is 0.8654, and the model $p < 2.2e-16$.

Based on the final fitted cross-correlation models (Equation 4 and 5) and the significance of coefficients in Tables S3 and S4 we find evidence that for the topic of "social distancing," predominant factual information (i.e., tweet number and proportion) leads the decrease of misinformation significantly when lag (h) is −2, −12 for the number of tweets and −1 for the proportion of tweets. However, we find that the proportion of misinformation tweets can also negatively impact the proportion of factual tweets in the future with a time lag of 1 day. The number of tweets containing misinformation is also positively related to the time (t) significantly.

$$\begin{cases} m_t = -13.00137 + 0.41524m_{t-1} - 0.2104f_{t-2} + 0.14772f_{t-4} \\ \quad + 0.01694f_{t-5} + 0.15307f_{t-7} \\ +0.21147f_{t-8} - 0.44504f_{t-12} + 0.12427f_{t-14} + 3.26323t + \epsilon_t \\ \epsilon_t \sim N(0,\sigma^2) \end{cases} \tag{3}$$

$$\begin{cases} mp_t = 1.0749 - 0.45844fp_{t-1} - 0.5834fp_{t+1} \\ \quad - 0.03397fp_{t+11} - 0.0001962t + \epsilon_t \\ \epsilon_t \sim N(0,\sigma^2) \end{cases} \tag{4}$$

4 | DISCUSSION AND CONCLUSION

4.1 | Findings and contributions

COVID-19, the worldwide drastic pandemic, has ignited online platforms and caused an "infodemic" on various channels; misinformation about prevention measures of coronavirus also spread widely and has affected the adoption of proper prevention measures. Although studies (e.g., Wang, et al., 2020) have found that effective risk and crisis communication with factual information can positively impact the performance of public health campaigns and government agencies have also disseminated factual messages on social media platforms actively, the temporal relationship and the potential suppression effects of factual information on misinformation have not been investigated empirically.

This study analyzed large-quantity longitudinal social media data using supervised machine learning methods and cross-correlation time-series analyses. It quantitatively investigated the temporal correlation between factual information and misinformation over days and whether predominant factual information can suppress misinformation on Twitter. Our analyses found evidence about the suppression effects of previously predominant factual information on misinformation for the two preventive-measure topics on Twitter. Specifically, in tweets relevant to topics of "wearing masks" and "social distancing," we found that the increasing percentage of factual information from the previous day led a decrease in the percentage of misinformation significantly. The increasing number of tweets containing factual information from a previous day led a decrease in the number of tweets containing misinformation significantly, while the significant time lags (h) for the two topics are different. In addition to the "suppression" effect of factual information (in scales of number and percentage) on misinformation, we also found that; (a) the number of misinformation-relevant tweets increased significantly over time for both topics; (b) the number of factual tweets from the same day had a positive significant correlation with the number of misinformation tweets; and (c) the number of misinformation tweets also had significant correlations with the number of factual tweets in future days but the effects varied when the time lags were different.

This study advances the existing knowledge body of crisis communication and misinformation, especially for studies focused on public health crises. Although high-volume/proportion factual information has the potential to reduce the misinformation on social media platforms (Iosifidis & Nicoli, 2020; Jin et al., 2020), little research has found empirical evidence to support the strategies of leveraging predominant factual information in suppressing misinformation. To the best of our knowledge, we are among the first to quantify the suppression effect of factual information over time by analyzing real-world social media posts (tweets) during the COVID-19 pandemic. Compared to the survey outcomes of previous research, this longitudinal data set reflects the information change of general Twitter users towards COVID-19 prevention measures in real-world situations rather than in experimental scenarios. The research findings can guide public health authorities, emergency responders, and other crisis managers to actively disseminate and endorse factual information in online platforms to suppress misinformation increases aggregately over time and to achieve crisis and risk communication goals more effectively. Specifically, benign social bots may be used to communicate factual information to social media users (Hofeditz et al., 2019) and tailor messages based on users' misinformation exposure status and risk perception (Reyes et al., 2021; Tully, Bode, et al., 2020). Crisis managers could also encourage the public to endorse (e.g., repost) useful factual information and actionable knowledge to increase their diffusion speed and effects (Gao et al., 2021). Strategies could also be implemented at critical timings to ensure effectiveness (Rich & Zaragoza, 2020). Message frames, wording, and other visual techniques can also be studied to achieve literacy education success on social media (Tully, Vraga, et al., 2020).

Additionally, the research framework provides insights into data-driven methods for studying different information campaigns during crises and emergencies on social media, as we mined and revealed the temporal patterns of both factual information and misinformation on Twitter during COVID-19. Crisis and risk communication studies and practices focused on disaster preparedness would also benefit from the findings to adopt proactive strategies to suppress misinformation as misinformation (e.g., rumors and myths) are also a parcel of disaster responses, such as the disaster mythology during Hurricane Katrina (Jacob et al., 2008). The research also contributes to a better understanding of social media's complex role in emergency and crisis management as well as the dynamic and divergent public responses in digital platforms.

4.2 | Limitations and future work

There are a few potential limitations of this study that have opened opportunities for future research. First, it focused on English tweets collected by a keyword-based Twitter Streaming API. Future work might use accurate translation algorithms to process tweets in other languages before conducting English-based natural language processing. Data from other social networking platforms could also be considered if they become available. Second, the existing supervised machine learning methods, including SVM, cannot achieve 100% accuracy when classifying data. We have put considerable effort into raising the classifier's accuracy to the level between 85% and 95%, such as increasing the volume of training data, comparing classification algorithms, and manually annotating the training data, and overall, our final classifiers outperformed the existing classifier used in similar tasks (e.g., Yao & Wang, 2020). With further development of text mining techniques, researchers could use more advanced machine learning techniques to classify tweets containing different information categories and reveal the real-world situation of information dissemination more accurately. Third, we classified tweets containing misinformation or factual information mainly based on whether the tweets containing falsehoods of the effectiveness of the preventions measures and opposing/manipulating public health agencies' guidance or endorsing/matching health agencies guidance towards prevention measures (i.e., wearing masks and social distancing), whose potential positive impacts on reducing the infection risk of the COVID-19 have been discussed in the up-to-date scientific findings (Chu et al., 2020; Lewnard & Lo, 2020). Our definitions of factual information and misinformation are based on both the latest scientific facts and authorities' risk messages. Fourth, we regarded English tweets containing “covid” or “coronavirus” as generated by US Twitter users, while a small portion of tweets from users of other English-speaking countries may also be collected by the Twitter API; however, such information cannot be verified if users choose not to post their profile location or geotag/include location information in contents. These tweets may affect the accuracy of the research outcome in the spatial perspective. Future research can improve the data collection process by restraining the location features of tweets. Lastly, the research focuses on suppressing misinformation in a social media platform (Twitter), but digital divides still exist. Not everyone has equal access to online resources nor the same ability to respond to misinformation during hazards; and Twitter users are younger, more educated than the general public (Pew Research Center, 2019). Future research can extend the work by investigating the effectiveness of disseminating factual information in suppressing misinformation across platforms and collaborating strategies (e.g., fact-checking and literacy education) to reduce and combat misinformation (Previti et al., 2020; Tully, Vraga, et al., 2020). It is also critical to understand the different effects of misinformation on people with distinct digital skills as they may perceive misinformation differently. Digital divides at multiple levels should also be considered in studying the overall strategies and impacts of misinformation suppression across communities.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 2028012. Any opinions, findings, and

conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in figshare at https://doi.org/10.6084/m9.figshare.13562519

REFERENCES

Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle, *Selected Papers of Hirotugu Akaike* (pp. 199–213). Springer. https://doi.org/10.1007/978-1-4612-1694-0_15

Almaliki, M. (2019). Online misinformation spread: A systematic literature map. In *Proceedings of the 2019 3rd International Conference on Information System and Data Mining* (pp. 171–178). https://doi.org/10.1145/3325917.3325938.

Barua, Z., Barua, S., Aktar, S., Kabir, N., & Li, M. (2020). Effects of misinformation on COVID-19 individual responses and recommendations for resilience of disastrous consequences of misinformation. *Progress in Disaster Science*, *8*, 100119. https://doi.org/10.1016/j.pdisas.2020.100119

Bishop, G., & Welch, G. (2001). An introduction to the kalman filter. *Proc of SIGGRAPH*, *Course*, *8*(27599-23175), 41.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: Forecasting and control*. John Wiley & Sons.

Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on twitter. In *Proceedings of the 20th International Conference on World Wide Web*. https://doi.org/10.1145/1963405.1963500

CDC. (2020, August 17). What's New. https://www.cdc.gov/coronavirus/2019-ncov/whats-new-all.html

Chen, E., Lerman, K., & Ferrara, E. (2020). Covid-19: The first public coronavirus twitter dataset. *JMIR Public Health and Surveillance*, *6*(2), e19273. https://doi.org/10.2196/19273

Chu, D. K., Akl, E. A., Duda, S., Solo, K., Yaacoub, S., & Schünemann, H. J. (2020). Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. *The Lancet*, *395*(10242), 1973–1987. https://doi.org/10.1016/S0140-6736(20)31142-9

Clark-Ginsberg, A., & Petrun Sayers, E. L. (2020). Communication missteps during COVID-19 hurt those already most at risk. *Journal of Contingencies and Crisis Management*, *28*(4), 482–484. https://doi.org/10.1111/1468-5973.12304

Cook, J., Bedford, D., & Mandia, S. (2014). Raising climate literacy through addressing misinformation: Case studies in agnotology-based learning. *Journal of Geoscience Education*, *62*(3), 296–306. https://doi.org/10.5408/13-071.1

Danielson, L., Marcus, B., & Boyle, L. (2019). Special feature: Countering vaccine misinformation. *The American Journal of Nursing*, *119*(10), 50–55. https://doi.org/10.1097/01.NAJ.0000586176.77841.86

Dave, D. M., Friedson, A. I., Matsuzawa, K., Sabia, J. J., & Safford, S. (2020). *Black Lives Matter protests, social distancing, and COVID-19*. National Bureau of Economic Research. https://doi.org/10.3386/w27408

Depoux, A., Martin, S., Karafillakis, E., Preet, R., Wilder-Smith, A., & Larson, H. (2020). The pandemic of social media panic travels faster than the COVID-19 outbreak. *Journal of Travel Medicine*, *27*(3). https://doi.org/10.1093/jtm/taaa031

Documenting the Now. (2020, October 25). *Hydrator [Computer Software]*. https://github.com/docnow/hydrator

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, *20*(5), 533–534. https://doi.org/10.1016/S1473-3099(20)30120-1

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (Vol. 326). John Wiley & Sons.

Earnshaw, V. A., & Katz, I. T. (2020). Educate, Amplify, and Focus to Address COVID-19 Misinformation. *JAMA Health Forum*, *1*(4), e200460. https://doi.org/10.1001/jamahealthforum.2020.0460.

Fuller, W. A. (2009). *Introduction to statistical time series* (Vol. 428). John Wiley & Sons.

Gao, S., Wang, Y., & Platt, L. S. (2021). Modeling US health agencies' message dissemination on twitter and users' exposure to vaccine-related misinformation using system dynamics. In *Proceedings of the 18th International Conference on Information Systems for Crisis Response and Management (ISCRAM)*. Blacksburg, Virginia, U.S.

Gopi, A. P., Jyothi, R. N. S., Narayana, V. L., & Sandeep, K. S. (2020). Classification of tweets data based on polarity using improved RBF kernel of SVM. *International Journal of Information Technology*, 1–16. https://doi.org/10.1007/s41870-019-00409-4

Grewal, M. S., Andrews, A. P., & Bartone, C. G. (2020). Kalman filtering, *Global Navigation Satellite Systems, Inertial Navigation, and Integration* (pp. 355–417). Springer.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Hernández-García, I., & Giménez-Júlvez, T. (2020). Assessment of health information about COVID-19 prevention on the internet: Infodemiological study. *JMIR Public Health and Surveillance*, *6*(2), e18717. https://doi.org/10.2196/18717

Hofeditz, L., Ehnis, C., Bunker, D., Brachten, F., & Stieglitz, S. (2019). Meaningful use of social bots? Possible applications in crisis communication during disasters. In *Proceedings of the 27th European Conference on Information Systems*.

Iosifidis, P., & Nicoli, N. (2020). The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation. *International Communication Gazette*, *82*(1), 60–81. https://doi.org/10.1177/1748048519880729

Jacob, B., Mawson, A. R., Marinelle, P., & Guignard, J. C. (2008). Disaster mythology and fact: Hurricane Katrina and social attachment. *Public Health Reports*, *123*(5), 555–566. https://doi.org/10.1177/003335490812300505

Jang, S. M., Geng, T., Li, J. Y. Q., Xia, R., Huang, C. T., Kim, H., & Tang, J. (2018). A computational approach for examining the roots and spreading patterns of fake news: Evolution tree analysis. *Computers in Human Behavior*, *84*, 103–113. https://doi.org/10.1016/j.chb.2018.02.032

Jin, Y., van der Meer, T. G., Lee, Y. I., & Lu, X. (2020). The effects of corrective communication and employee backup on the effectiveness of fighting crisis misinformation. *Public Relations Review*, *46*(3), 101910. https://doi.org/10.1016/j.pubrev.2020.101910

Kahne, J., & Bowyer, B. (2017). Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation. *American educational research journal*, *54*(1), 3–34. https://doi.org/10.3102/0002831216679817

Kaufhold, M. A., Gizikis, A., Reuter, C., Habdank, M., & Grinko, M. (2019). Avoiding chaotic use of social media before, during, and after emergencies: Design and evaluation of citizens' guidelines. *Journal of Contingencies and Crisis Management*, *27*(3), 198–213. https://doi.org/10.1111/1468-5973.12249

Kouzy, R., Abi Jaoude, J., Kraitem, A., El Alam, M. B., Karam, B., Adib, E., Zarka, J., Traboulsi, C., Akl, E. W., & Baddour, K. (2020). Coronavirus goes viral: Quantifying the COVID-19 misinformation epidemic on Twitter. *Cureus*, *12*(3), e7255. https://doi.org/10.7759/cureus.7255

Krause, N. M., Freiling, I., Beets, B., & Brossard, D. (2020). Fact-checking as risk communication: the multi-layered risk of misinformation in times of COVID-19. *Journal of Risk Research*, *23*(7-8), 1052–1059. https://doi.org/10.1080/13669877.2020.1756385

Lewnard, J. A., & Lo, N. C. (2020). Scientific and ethical basis for social-distancing interventions against COVID-19. *The Lancet Infectious Diseases*, *20*(6), 631–633. https://doi.org/10.1016/S1473-3099(20)30190-0

Liu, S., Li, F., Li, F., Cheng, X., & Shen, H. (2013). Adaptive co-training SVM for sentiment classification on tweets. In *Proceedings of the 22nd ACM International Conference on Information & Knowledge Management* (pp. 2079–2088). https://doi.org/10.1145/2505515.2505569

Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297–303. https://doi.org/10.1093/biomet/65.2.297

Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021). Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. *Nature Human Behavior*, *5*(3), 337–348. https://doi.org/10.1038/s41562-021-01056-1

van der Meer, T. G., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication*, *35*(5), 560–575. https://doi.org/10.1080/10410236.2019.1573295

Moritz, S., & Bartz-Beielstein, T. (2017). imputeTS: Time series missing value imputation in R. *The R Journal*, *9*(1), 207–218. https://doi.org/10.32614/RJ-2017-009

Nabity-Grover, T., Cheung, C., & Thatcher, J. B. (2020). Inside out and outside in: How the COVID-19 pandemic affects self-disclosure on social media. *International Journal of Information Management*, *55*, 102188. https://doi.org/10.1016/j.ijinfomgt.2020.102188

Niles, M. T., Emery, B. F., Reagan, A. J., Dodds, P. S., & Danforth, C. M. (2019). Social media usage patterns during natural hazards. *PLOS One*, *14*(2), e0210484. https://doi.org/10.1371/journal.pone.0210484

Orabi, M., Mouheb, D., Al Aghbari, Z., & Kamel, I. (2020). Detection of bots in social media: A systematic review. *Information Processing & Management*, *57*(4), 102250. https://doi.org/10.1016/j.ipm.2020.102250

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*(7855), 590–595. https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., McPhetres, J., Zhang, Y., Lu, J. G., & Rand, D. G. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science*, *31*(7), 770–780. https://doi.org/10.1177/0956797620939054

Previti, M., Rodriguez-Fernandez, V., Camacho, D., Carchiolo, V., & Malgeri, M. (2020). Fake news detection using time series and user features classification. In *International Conference on the Applications of Evolutionary Computation* (pp. 339–353). https://doi.org/10.1007/978-3-030-43722-0_22

Pulido, C. M., Villarejo-Carballido, B., Redondo-Sama, G., & Gómez, A. (2020). COVID-19 infodemic: More retweets for science-based information on coronavirus than for false information. *International Sociology*, *35*(4), 377–392. https://doi.org/10.1177/0268580920914755

Qiu, W., & Chu, C. (2019). Clarification of the concept of risk communication and its role in public health crisis management in China. *Disaster Medicine and Public Health Preparedness*, *13*(5-6), 834–836. https://doi.org/10.1017/dmp.2019.10

R Core Team. (2020). R: A language and environment for statistical computing. https://www.R-project.org/

Ramos, J. (2003). Using TF-IDF to determine word relevance in document queries. *In Proceedings of the First Instructional Conference on Machine Learning*, *242*, 133–142.

Rasheed, R. (2020, July 9). Why does stationarity matter in time series analysis? Learn the fundamental rule of time series analysis. https://towardsdatascience.com/why-does-stationarity-matter-in-time-series-analysis-e2fb7be74454

Reyes, L. M., Ortiz, L., Abedi, M., Luciano, Y., Ramos, W., & Reyes, P. J. D. J. (2021). Misinformation on COVID-19 origin and its relationship with perception and knowledge about social distancing: A cross-sectional study. *PLOS One*, *16*(3), e0248160. https://doi.org/10.1371/journal.pone.0248160

Rich, P. R., & Zaragoza, M. S. (2020). Correcting misinformation in news stories: An investigation of correction timing and correction durability. *Journal of Applied Research in Memory and Cognition*, *9*(3), 310–322. https://doi.org/10.1016/j.jarmac.2020.04.001

Said, S. E., & Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, *71*, 599–607. https://doi.org/10.2307/2336570

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K. C., Flammini, A., & Menczer, F. (2018). The spread of low-credibility content by social bots. *Nature Communications*, *9*(1), 1–9. https://doi.org/10.1038/s41467-018-06930-7

Shimizu, K. (2020). 2019-nCoV, fake news, and racism. *The Lancet*, *395*(10225), 685–686. https://doi.org/10.1016/S0140-6736(20)30357-3

Silva, M., Ceschin, F., Shrestha, P., Brant, C., Fernandes, J., Silva, C. S., Grégio, A., Oliveira, D., & Giovanini, L. (2020). Predicting misinformation and engagement in covid-19 twitter discourse in the first months of the outbreak. arXiv preprint arXiv:2012.02164.

Silver, A., & Andrey, J. (2019). Public attention to extreme weather as reflected by social media activity. *Journal of Contingencies and Crisis Management*, *27*(4), 346–358. https://doi.org/10.1111/1468-5973.12265

Slotkin, J. (2020, October 16). NYC poison control sees uptick in calls after Trump's disinfectant comments. https://www.npr.org/sections/coronavirus-live-updates/2020/04/25/845015236/nyc-poison-control-sees-uptick-in-calls-after-trumps-disinfectant-comments

Stieglitz, S., Bunker, D., Mirbabaie, M., & Ehnis, C. (2018). Sense-making in social media during extreme events. *Journal of Contingencies and Crisis Management*, *26*(1), 4–15. https://doi.org/10.1111/1468-5973.12193

Stoffer, D. (2020). astsa: Applied statistical time series analysis. https://CRAN.R-project.org/package=astsa

Swire-Thompson, B., & Lazer, D. (2020). Public health and online misinformation: Challenges and recommendations. *Annual Review of Public Health*, *41*, 433–451. https://doi.org/10.1146/annurev-publhealth-040119-094127

Trapletti, A., & Hornik, K. (2020). tseries: Time series analysis and computational finance. https://CRAN.R-project.org/package=tseries.

Trethewey, S. P. (2020). Strategies to combat medical misinformation on social media. *Postgraduate Medical Journal*, *96*, 4–6. https://doi.org/10.1136/postgradmedj-2019-137201

Tully, M., Bode, L., & Vraga, E. K. (2020). Mobilizing users: Does exposure to misinformation and its correction affect users' responses to a health misinformation post? *Social Media+ Society*, *6*(4), 1–12. https://doi.org/10.1177/2056305120978377

Tully, M., Vraga, E. K., & Bode, L. (2020). Designing and testing news literacy messages for social media. *Mass Communication and Society*, *23*(1), 22–46. https://doi.org/10.1080/15205436.2019.1604970

Twitter. (2021, October 25). *Filter realtime tweets*. https://developer.twitter.com/en/docs/twitter-api/v1/tweets/filter-realtime/overview

Utz, S., Schultz, F., & Glocka, S. (2013). Crisis communication online: How medium, crisis type and emotions affected public reactions in the Fukushima Daiichi nuclear disaster. *Public Relations Review*, *39*(1), 40–46. https://doi.org/10.1016/j.pubrev.2012.09.010

Vicario, M. D., Quattrociocchi, W., Scala, A., & Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. *ACM Transactions on the Web (TWEB)*, *13*, 1–22. https://doi.org/10.1145/3316809

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Vraga, E. K., Bode, L., & Tully, M. (2020). Creating news literacy messages to enhance expert corrections of misinformation on Twitter. *Communication Research*, 1–23. https://doi.org/10.1177/0093650219898094

Walter, N., Brooks, J. J., Saucier, C. J., & Suresh, S. (2021). Evaluating the impact of attempts to correct health misinformation on social media: A meta-analysis. *Health Communication*, *36*(13), 1776–1784. https://doi.org/10.1080/10410236.2020.1794553

Wang, X., & Song, Y. (2020). Viral misinformation and echo chambers: The diffusion of rumors about genetically modified organisms on social media. *Internet Research*, *30*(5), 1547–1564. https://doi.org/10.1108/INTR-11-2019-0491

Wang, Y., Hao, H., & Platt, L. S. (2021). Examining risk and crisis communications of government agencies and stakeholders during early-stages of COVID-19 on Twitter. *Computers in Human Behavior*, *114*, 106568. https://doi.org/10.1016/j.chb.2020.106568

Wang, Y., & Taylor, J. E. (2018). Coupling sentiment and human mobility in natural disasters: A Twitter-based study of the 2014 South Napa Earthquake. *Natural Hazards*, *92*(2), 907–925. https://doi.org/10.1007/s11069-018-3231-1

WHO. (2020, August 17). *Coronavirus disease (COVID-19)*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019

Wilson, T., & Starbird, K. (2020). Cross-platform disinformation campaigns: Lessons learned and next steps. *Harvard Kennedy School Misinformation Review*, *1*(1). https://doi.org/10.37016/mr-2020-002

Wu, L., Morstatter, F., Carley, K. M., & Liu, H. (2019). Misinformation in social media: Definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter*, *21*(2), 80–90. https://doi.org/10.1145/3373464.3373475

Yao, F., & Wang, Y. (2020). Domain-specific sentiment analysis for tweets during hurricanes (DSSA-H): A domain-adversarial neural-network-based approach. *Computers, Environment and Urban Systems*, *83*, 101522. https://doi.org/10.1016/j.compenvurbsys.2020.101522

Yu, F., Liu, Q., Wu, S., Wang, L., & Tan, T. (2017). A Convolutional Approach for Misinformation Identification. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence* (pp. 3901–3907). https://doi.org/10.24963/ijcai.2017/545

Yu, M., Li, Z., Yu, Z., He, J., & Zhou, J. (2020). Communication related health crisis on social media: A case of COVID-19 outbreak. *Current Issues in Tourism*, 1–7. https://doi.org/10.1080/13683500.2020.1752632

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management*, *57*(2), 102025. https://doi.org/10.1016/j.ipm.2019.03.004

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Wang, Y., Gao, S., & Gao, W. (2021). Investigating dynamic relations between factual information and misinformation: Empirical studies of tweets related to prevention measures during COVID-19. *Journal of Contingencies and Crisis Management*, 1–13. https://doi.org/10.1111/1468-5973.12385