Feeling the same: Investigating emotion dynamics on Twitter during COVID-19

Data Science exam

Stine Nyhus Larsen (SNL): 201808306

Sara Møller Østergaard (SMØ): 201808643

M.Sc. Cognitive Science, Aarhus University

May 31st, 2022

Abstract

In this paper, we investigate emotion dynamics on Danish Twitter before and during the COVID-19 pandemic. The analysis extends upon previous work where windowed relative entropy has been used to calculate information signals on emotion classification of tweets. Change point detection is applied to identify relevant time periods where the dynamics change. Relying on these windows, the varying dynamics of the system are examined. The results of the study indicate a stronger positive coupling between the entropy measures *resonance* and *novelty* during the COVID-19 pandemic. Moreover, visually different correlations of individual emotions are detected in the change point periods. The findings of the study suggest a change in the dynamics of emotions on Twitter before and after COVID-19 appeared in Denmark. We discuss the implications of the emotion dynamics during the periods and suggest further analysis using the emotion correlations.

Link to GitHub: https://github.com/saraoe/data_science_exam

1 Introduction (SNL)

The emergence of the Coronavirus Disease (COVID-19) caused abrupt changes to all of modern society. Governments all over the world introduced strict lockdowns and closed their borders. The shock experienced by the population has given rise to interesting but also important studies. Bor, Jørgensen, & Petersen (2021) studied how this uncommon governmental involvement in the lives of citizens has affected structural relationships i.e., relationships between citizens and between the state and citizens. Nielbo et al. (2021) looked at how information dynamics of topics in Danish newspapers changed during the first part of the pandemic.

In this paper, we investigate whether the societal changes and events that occurred during the COVID-19 pandemic correlate with information dynamics of emotional content on Danish Twitter. This is done using time series signals extracted from a large corpus of Danish tweets from January 2019 to April 2021.

1.1 Research on Twitter data (SMØ)

Twitter has in recent years increasingly been used to assess public mental health and study public opinion (Chancellor & De Choudhury, 2020; Thelwall, 2014). Twitter data is well suited for these kinds of analyses due to a range of reasons; the user group is large and quite diverse with 436 million active users around the world as of October 2021 (Kemp, 2021) and it includes both the general public and high-profile opinion leaders. Hence, Twitter is a valuable source of information that is continuously updated with real-time reactions to events, and it is freely available for scientific investigation. Finally, it is time-stamped, which means that Twitter can be analyzed as a complex time series system with changing dynamics.

Data from Twitter has thus been used to study a range of topics. A popular example is politics and elections; Buccoliero, Bellio, Crestini, & Arkoudas (2020) studied how the two candidates used Twitter in their campaigns in the 2016 US election, Sharma & Moh (2016) investigated the extent to which sentiment on Twitter could be used to predict the outcome of an Indian election and Vliet, Törnberg, & Uitermark (2020) developed a database of parliamentarian Twitter-use across 26 countries. Other examples of topics studied using Twitter include Bitcoin tradings, where Twitter has been found to be predictive of trading volumes the following day (Shen, Urquhart, & Wang, 2019), and conspiracies, where large differences in emotions and topics were found between groups being mainly exposed to conspiratorial and scientific context, respectively (Fong, Roozenbeek, Goldwert, Rathje, & van der Linden, 2021). Several studies have also analyzed changes in sentiment and emotion expressed in tweets in relation to the COVID-19 pandemic (Dubey, 2020; Li et al., 2020; Yin, Yang, & Li, 2020; Zhou, Yang, Xiao, & Chen, 2020).

1.2 Windowed relative entropy (SNL)

In a study by Barron, Huang, Spang, & DeDeo (2018), the windowed relative entropy measures *novelty*, *transience*, and *resonance* were defined and were found useful for examining dynamics of time series systems. Novelty and transience are very similar measures, with novelty being defined as in equation 1 and transience as in equation 2.

$$N_w(j) = \frac{1}{w} \sum_{d=1}^{w} D\left(s^{(j)} | s^{(j-d)}\right) \tag{1}$$

$$T_w(j) = \frac{1}{w} \sum_{d=1}^{w} D\left(s^{(j)} | s^{(j+d)}\right) \tag{2}$$

In both equations, w is the window size, s^j is the probability distribution at the j'th time point, and $D(P|Q)$ is a function for relative entropy (e.g. Kullback-Leibler divergence) where the inputs P and Q are probabilistic representations of documents. Resonance is defined as the difference between novelty and transience (equation 3).

$$R_w(j) = N_w(j) - T_w(j) \tag{3}$$

In other words, novelty is a measure of the average amount of relative surprise between the probability distribution at a given time point and the probability distributions in a window with the w previous time point. Likewise, transience is comparing the probability distribution at the j'th time point with the w following probability distributions. Resonance is high if novelty is high while transience is low, meaning the document is very surprising compared to previous documents but not the following. In Barron et al. (2018), the measures and the relation between them were used to examine the dynamics in parliament speeches during the French revolution, but the same methods have been used to study other time series systems (Murdock, Allen, & DeDeo, 2017).

In a study be Nielbo et al. (2021), windowed relative entropy was used to study the response in Danish newspapers to the first COVID-19 lockdown. The study relied on Latent Dirichlet allocation for creating a topic representation for each front paper article of six major Danish newspapers over a time period from December 2019 to July 2020. The study examined the coupling between resonance and novelty and found a general positive resonance-novelty coupling, meaning more novel topics tended to resonate more (Nielbo et al., 2021). However, a decoupling between resonance and novelty in Danish broadsheet newspapers was identified during the first lockdown (Nielbo et al., 2021). The news information decoupling suggests that in this period in Denmark, less novel topics resonated more, thus, the same topics were presented in the newspapers. These findings indicate that the changes imposed on society with the arrival of COVID-19 correlated with changes in the information dynamics of newspapers (Nielbo et al., 2021).

1.3 Emotion dynamics on Twitter (SMØ)

In our natural language processing exam (see appendix A.2), we applied windowed relative entropy to Danish tweets inspired by previous studies (Barron et al., 2018; Murdock et al., 2017; Nielbo et al., 2021). We used an emotion classifier to construct document representations of the tweets rather than topic modeling as previously seen (Nielbo et al., 2021; Barron et al., 2018). The results of the study showed that certain dynamics of Twitter can be captured using this approach. For example, there were observable changes in both the resonance and novelty of emotional content around Christmas and New Year (see appendix A.1). To validate the signal, the same methods were applied to polarity probability distributions for the tweets, given the conceptual overlap between polarity and emotions. The signals calculated using the two different document representations had strong correlations (appendix A.2, page 12), thus indicating that the dynamics found using emotion predictions as document representations could generalize across emotion classifiers.

1.4 Aim of paper (SNL & SMØ)

In this paper, we extend on previous work (appendix A.2) with the aim of investigating changing dynamics on social media during the COVID-19 pandemic using windowed relative entropy signals based on emotion distributions. This will be done by first using change point detection to identify relevant shifts in resonance. This is followed by analyses on the signal in the different time windows specified by the change points, using both linear regression models and a system characterisation based on variable correlations. The latter approach is inspired by the methods applied in Münnix et al. (2012), where correlation matrices were used to identify different states of complex financial time series systems.

2 Methods

2.1 Information signals (SNL)

For the analysis of the present study, we investigated the time series that were the result of our natural language processing exam. The following section will include short descriptions of the methods used to obtain these signals. For a more elaborate description, we refer to the methods section of the exam paper (appendix A.2).

2.1.1 Data and preprocessing (SMØ)

The corpus consisted of 26,024,959 tweets scraped from Twitter from January 2019 to April 2021 using a manually defined set of common Nordic words. The tweets were subsequently filtered to only include Danish tweets. Before computing the signal, the tweets were preprocessed to remove hashtags, mentions, emojis, and URLs, and all retweets were excluded from the analysis.

2.1.2 Emotion classification (SNL)

To obtain the emotion classifications of the tweets in the corpus, we used the Danish BERT Emotion model (Pauli, Barrett, Lacroix, & Hvingelby, 2021). This model includes eight emotion categories (see table 2.1) and outputs a predictive distribution over the emotion categories. We use this predictive distribution as document representation.

Danish Labels	**Translations**
Glæde/Sindsro	Happiness/Calmness
Tillid/Accept	Trust/Acceptance
Forventning/Interrese	Expectation/Interest
Overasket/Målløs	Surprised/Speechless
Vrede/Irritation	Anger/Irritation
Foragt/Modvilje	Contempt/Reluctance
Sorg/Trist	Grief/Sadness
Frygt/Bekymret	Fear/Worry

Table 2.1: The eight emotion categories in the Danish BERT Emotion model together with our English translations.

Subsequently, the emotion distributions were summarized by averaging over the value for each of the eight emotion categories across all tweets in a day. This resulted in a single emotion distribution for each day. The subsequent analysis relied on these summarized emotion distributions, and they will henceforth simply be referred to as *emotion distributions*.

2.1.3 Windowed relative entropy (SMØ)

Windowed relative entropy was used to compute the information signals novelty, transience, and resonance. The Jensen-Shannon divergence was used as the relative entropy measure, and the information signals were calculated using the emotion distributions as latent variables and a window size of $w = 3$.

2.2 Change point detection (SNL)

From visually inspecting the resonance signal, it was apparent that the time series was not stationary. The signal seemed to have a stabilized mean but not a stabilized variance (see appendix A.1). Taking advantage of this variance shift, we wanted to identify time points where the underlying state of the time series shifted. We used offline change point detection to identify these relevant changes in the information signal.

The search method Pruned Exact Linear Time (PELT) was used to identify the change points. This method not only finds the relevant change points but also determines the number of change points. By using a linear penalization on the number of change points, the PELT algorithm identifies the number of change points while aiming to minimize overfitting (Truong, Oudre, & Vayatis, 2020). PELT is an optimal search method, meaning that it is guaranteed to find the optimal segmentation of the signal given the cost function and penalization (Truong et al., 2020).

We used the radial basis function (rbf) as the cost function, which is a cost function based on a Gaussian kernel. The kernel k for rbf is defined as

$$k(x, y) = \exp(-\gamma\|x - y\|^2) \tag{4}$$

Here $\|\cdot\|$ is the Euclidian norm and $\gamma > 0$ is the bandwidth parameter which is defined as the inverse of the median of all pairwise distances (Truong et al., 2020). When fitting the model, we used a smoothing parameter $\beta = 4$. A low β would result in an increased segmentation of the signal while a higher value for β would make the algorithm disregard more change points. Thus, setting this parameter can be thought of as a trade-off between complexity and goodness-of-fit for the model. The model was fitted using the `ruptures` python package (Truong et al., 2020).

2.3 Analysis (SMØ & SNL)

Using the detected change points, we divided the resonance signal into the corresponding number of periods where the beginning and end of each period was determined by a change point. We investigated the changing dynamics within each of these windows.

2.3.1 Resonance-novelty coupling (SMØ)

To describe the changes in the signal between the different change point periods, we investigated the coupling between resonance and novelty following Nielbo et al. (2021). This was implemented as a linear regression model predicting resonance from novelty within each of the change point periods,

$$R_i = \beta_0 + \beta_1 N_i + \epsilon_i \tag{5}$$

where R_i and N_i refer to resonance and novelty at the i'th day, respectively. β_0 is the intercept, β_1 the $N \times R$ slope, and ϵ_i the error term. To make the estimate of the $N \times R$ slope more interpretable, both resonance and novelty were z-scored before fitting the model.

2.3.2 Individual emotions (SNL)

The emotion probability for each of the eight emotion labels can by themselves be conceptualized as time series. Thus, we had eight distinct time series consisting of the daily average of the emotions. In order to investigate whether changes in any emotion were particularly predictive of changes in the resonance signal, we used the time series for the individual emotions as predictors of resonance in a linear regression. Thus, the following regression was fitted within each of the change point windows,

$$\begin{aligned} R_i = \beta_0 + \beta_1 happiness_i + \beta_2 trust_i + \beta_3 expectation_i + \beta_4 surprised_i \\ + \beta_5 anger_i + \beta_6 contempt_i + \beta_7 grief_i + \beta_8 fear_i + \epsilon_i \end{aligned} \tag{6}$$

Here, the subscript i refers to time point i in the signal. As for the resonance-novelty coupling model, both resonance and the emotion signals were z-scored before fitting the model.

We were also interested in investigating whether the interplay between the emotion signals could characterize the resonance signal within each change point period. Inspired by methods conveyed in Münnix et al. (2012), we used correlation matrices with Pearson correlations between all emotion signals to describe how the interplay between emotions differed between windows. Pearson's correlation coefficient is defined such that $r \in [-1, 1]$, thus spanning from a perfect negative correlation to a perfect positive correlation.

3 Results (SMØ)

Three change points in resonance of the tweets were detected by the PELT algorithm, hence dividing the signal into four time periods. An overview of the date ranges and resonance in each of the associated periods can be seen in table 3.1 below. Figure 3.1 shows the densities of resonance within each of the change point periods along with 94% high density intervals.

Change point period	**Dates**	**Resonance**
1	2019-01-01 to 2020-02-04	0.000 $[-0.001, 0.002]$
2	2020-02-05 to 2020-12-20	0.000 $[-0.001, 0.001]$
3	2020-12-21 to 2021-01-14	0.000 $[-0.007, 0.006]$
4	2021-01-15 to 2021-04-30	0.000 $[-0.001, 0.001]$

Table 3.1: Change points identified, the dates included in each change points, and the mean of resonance together with the 94% high density intervals.

Mean resonance in the four change point periods stayed the same, however, variance in the signal changed between the different periods. The variance in

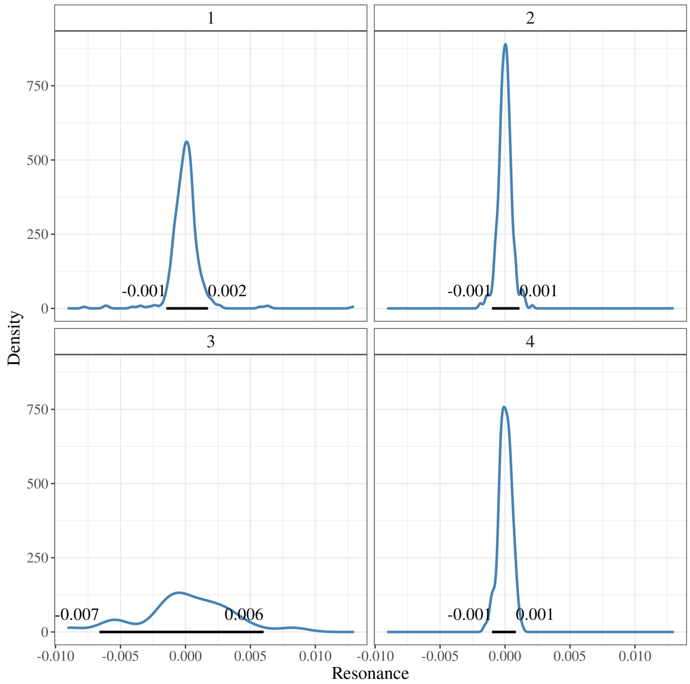


Figure 3.1: Density of resonance and 94% high density intervals in the four change point periods.

resonance decreased between the first and the second change point period. Between the second and the third change point period the variance had a large increase and the variance was considerably bigger in the third period compared to the first. Finally, the variance in the fourth change point period decreased to a similar level as expressed in the second change point period.

Figure 3.2 shows the resonance time series together with the change points periods and selected events related to COVID-19 from a timeline published by Statens Serum Institut (SSI) (Statens Serum Institut, 2022). The changes in variance between the change point periods are visually apparent when inspecting the signal. A description of the events can be found in table 3.2.

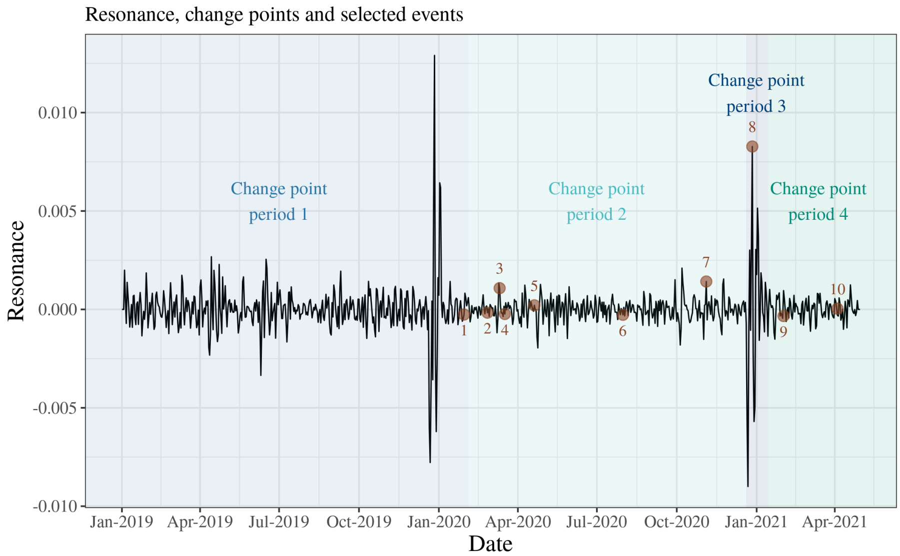


Figure 3.2: Resonance signal with background colors separating the change point periods. Event descriptions can be found in table 3.2, and change point start- and end dates in table 3.1

3.1 Resonance-novelty coupling (SNL)

Change point period	$N \times R$ **slope**
1	$0.511\ [0.425, 0.598]$
2	$0.875\ [0.747, 1.003]$
3	$0.648\ [0.299, 0.997]$
4	$0.741\ [0.540, 0.942]$

Table 3.3: Coefficients of the $N \times R$ slopes together with the 95% confidence intervals.

Table 3.3 below shows $N \times R$ slopes from the linear regression models predicting resonance from novelty. In change point period 1, the estimated coefficient was $\beta_1 = 0.511$ and was the smallest out of all of the four time periods, while the estimate of the $N \times R$ slope in the second change point period was the largest with a coefficient of $\beta_1 = 0.875$. The coefficient for the slope in change point period 4 was $\beta_1 = 0.741$, thus, it lay between the coefficients estimated in change

No.	**Period**	**Date**	**Description**
1	1	2020-01-30	The outbreak of the virus is declared a threat to global health by WHO.
2	2	2020-02-26	The first Danish citizen is tested positive for COVID-19.
3	2	2020-03-11	The Danish prime minister has her second press conference and she announces a two-week lockdown in Denmark. All schools, daycares and institutions are closing. Assembly ban for more than 100 people is introduced. Public employees with a non-critical functionality are sent home.
4	2	2020-03-17	The Queen of Denmark speaks to the public about the COVID-19 crisis.
5	2	2020-04-20	Partial reopening. Driving schools, hair dressers, research laboratories, and certain other liberal professions together with youngest grade levels and outdoor sport activities without body contact is allowed to reopen.
6	2	2020-07-31	Danish health authority recommend wearing face masks in public transportations if there are many people.
7	2	2020-11-04	The government decides to put down all mink on Danish mink farms due to an outbreak of a COVID-19 mutation.
8	3	2020-12-27	The first Danish citizens are vaccinated using the Pfizer/BioNTech vaccine.
9	4	2021-01-28	The lockdown in Denmark, which was introduced in December 2020, is prolonged until February 28, 2021.
10	4	2021-04-14	The AstraZeneca vaccine is withdrawn completely from the Danish vaccination program.

Table 3.2: Date and translated description for the relevant events from the timeline published by SSI (Statens Serum Institut, 2022)

point period 1 and 2. The confidence intervals of the estimates in change point periods 2 and 4 had a larger overlap, compared to the confidence intervals of the estimates in periods 1 and 4. Finally, the $N \times R$ slope fitted in change point period 3 had an estimate of $\beta_1 = 0.648$, however, the confidence intervals of the estimate overlapped with the estimated coefficient in the three other change point periods.

The $N \times R$ slopes of the linear models are visualized in figure 3.3. Notice that the scale of both axes differ between change point periods. This is due to variations in the distribution of the data points, which makes visualizations using the same scales difficult to interpret. The figure shows a general positive coupling between resonance and novelty. Moreover, it can be seen that $N \times R$ slope in change point periods 2 was steeper compared to the same slope in the first change point period.

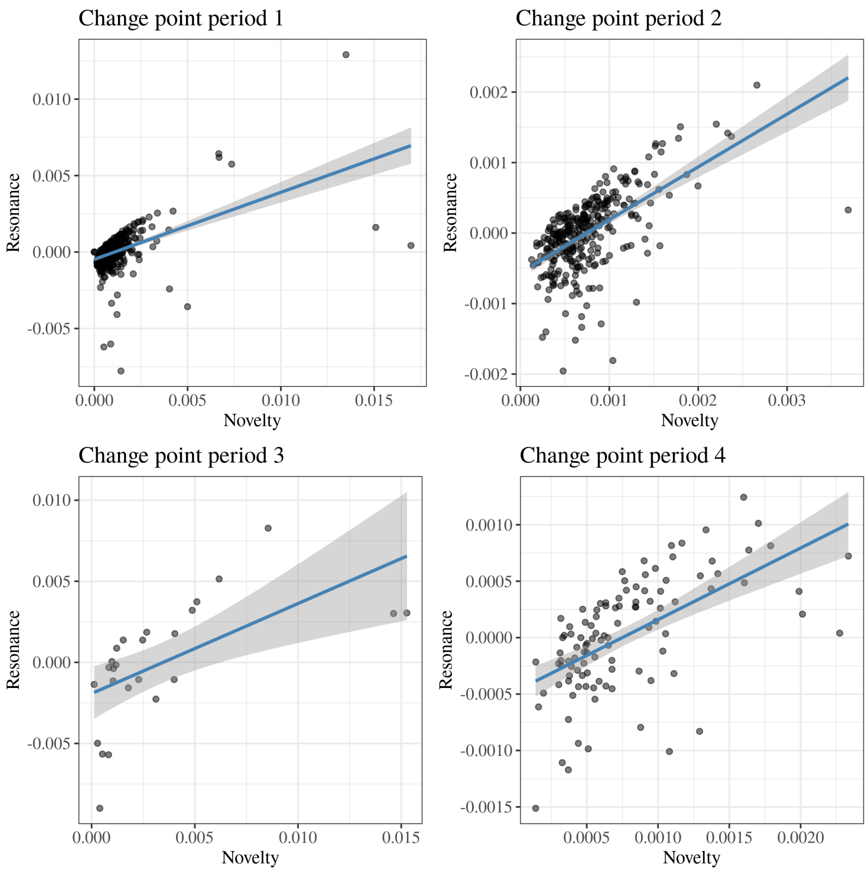


Figure 3.3: $N \times R$ slopes in the four change point periods. Notice that both axes vary between plots.

3.2 Individual emotions (SMØ)

Figure 3.4 shows emotion distributions as time series signals. Variations in the dynamics of the different signals can be visually detected, e.g. *grief* seems stationary, while *expectation* and *contempt* show negative and positive trends, respectively. The figure also depicts sudden changes in values for some of the emotions, for example, the spike in *fear* in the first half of 2020.

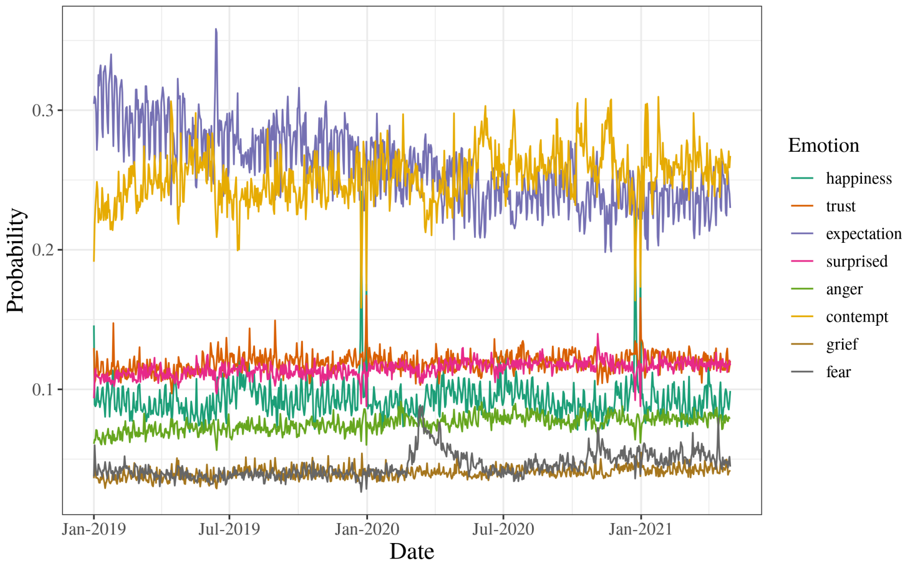


Figure 3.4: Mean value for each of the summarized emotion categories.

Table 3.4 shows the β-estimates of the linear models predicting resonance from the emotion probabilities in each of the change point periods. Neither of the emotions could significantly predict resonance in any of the change point periods (with a significance threshold of $\alpha = 0.05$). Uncertainty of all the estimates was large with confidence intervals crossing zero for all estimates.

Correlation matrices showing the correlation between the individual emotion time series signals in each of the change point periods can be found in figure

	Change point period	
	1	**2**
Happiness	$30.97\ [-128.17, 190.12]$	$-24.73\ [-109.17, 59.72]$
Trust	$15.71\ [-64.74, 96.16]$	$-12.62\ [-55.32, 30.08]$
Expectation	$62.74\ [-259.63, 385.11]$	$-50.07\ [-221.11, 120.98]$
Surprised	$12.55\ [-51.41, 76.51]$	$-9.91\ [-43.87, 24.04]$
Anger	$12.88\ [-53.6, 79.36]$	$-10.34\ [-45.62, 24.94]$
Contempt	$43.76\ [-180.93, 268.46]$	$-34.91\ [-154.14, 84.31]$
Grief	$9.07\ [-37.74, 55.87]$	$-7.24\ [-32.08, 17.6]$
Fear	$20.52\ [-82.26, 123.31]$	$-15.99\ [-70.57, 38.58]$
	3	**4**
Happiness	$-1864.6\ [-5184.42, 1455.22]$	$61.04\ [-120.52, 242.61]$
Trust	$-1864.6\ [-5184.42, 1455.22]$	$61.04\ [-120.52, 242.61]$
Expectation	$-3778.61\ [-10500.3, 2943.09]$	$123.2\ [-244.46, 490.85]$
Surprised	$-748.9\ [-2084.76, 586.97]$	$24.52\ [-48.46, 97.5]$
Anger	$-776.96\ [-2162.06, 608.14]$	$25.33\ [-50.47, 101.13]$
Contempt	$-2633.35\ [-7318.82, 2052.13]$	$85.92\ [-170.35, 342.2]$
Grief	$-549.29\ [-1524.85, 426.28]$	$17.85\ [-35.53, 71.22]$
Fear	$-1204.24\ [-3348.63, 940.14]$	$39.26\ [-78.06, 156.58]$

Table 3.4: Coefficients of the β estimates for the eight different emotions together with the 95% confidence intervals of the estimates. None of the estimates are significant ($p > 0.05$ for all the estimates).

3.5. The correlations between the emotions are generally stronger in change point period 3 compared to the three other periods. This period contains very clear clusters; *happiness* and *trust* are positively correlated with each other while negatively correlated with *surprised*, *anger*, *contempt*, and *fear*. The four latter emotions are all positively correlated with each other. All these correlations have a correlation coefficient $r > .5$.

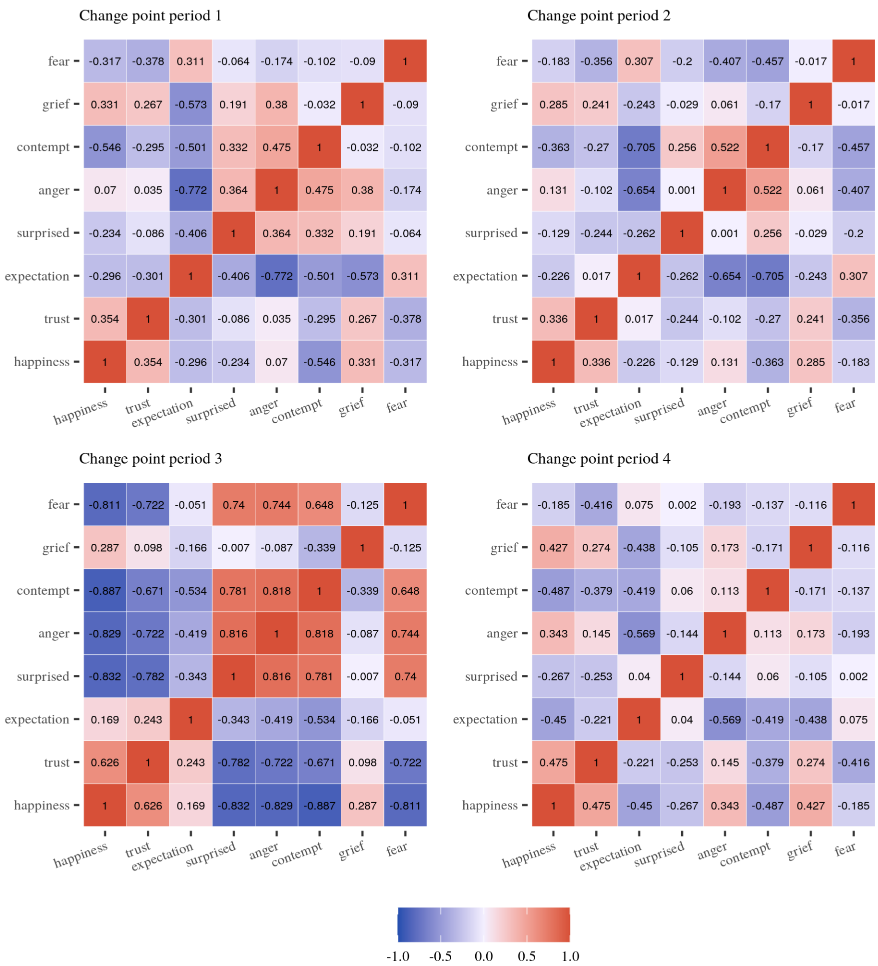


Figure 3.5: Correlation matrices for the eight emotion time series signals within the periods defined by change points.

Patterns in correlations are also visually detectable in change point periods 1, 2, and 4. A similar positive correlation between *happiness* and *trust* was found in these three time windows. Negative correlations between *expectation* and the three emotions *anger*, *contempt*, and *grief* were also found in all three periods, however, with varying sizes of the correlation coefficients (e.g. the correlation coefficient between *expectation* and *grief* was $r = -0.573$ in period

1, while the correlation between the same emotions in period 2 had a coefficient of $r = -0.243$). Finally, the correlations between *surprised* and the same three emotions (i.e. *anger*, *contempt*, and *grief*) seems to change noticeably between change point period 1 and 2 (where the first period generally shows stronger correlations) but less so between change point period 2 and 4.

4 Discussion (SMØ & SNL)

In the following, we will discuss the results of the analysis along with the limitations of the applied methods. We will first look into the detected change points; how they match with intuitions from visual inspection of the signal and how they relate to the COVID-19 pandemic. Then, we will discuss what the change in resonance-novelty coupling might indicate about the dynamics of emotional content on Twitter. Lastly, we compare the linear regression models of resonance predicted by emotions and the correlation matrices of emotions to investigate how the individual emotions and the interplay between them characterize the information signal in the change point periods.

4.1 Change points (SNL)

Three change points were detected in the resonance time series segmenting the signal into four time periods (Table 3.1). The first split happened on February 4th, 2020. The second change point was on December 20th, 2020, just before Christmas, while the third was on January 14th, 2021, just two weeks into the new year. The time windows were of different lengths and most noticeably the period between the second and the third change point was only 24 days, leading to higher uncertainty in estimates for this period.

The first change point occurred shortly before the pandemic reached Denmark, as the first Danish citizen tested positive on February 26th and the lockdown is announced on March 11th (Table 3.2). This change point thus separated the resonance signal in periods before and after COVID-19 in Denmark. There seems to be a general change in how emotions resonate after the pandemic had reached Denmark, as periods 2 and 4 exhibited similar dynamics with slightly lower variance compared to period 1 (Table 3.1 and Figure 3.1). They were shortly interrupted by the third change point period in which a much larger variance in resonance was expressed, but otherwise, variance in resonance was lower in the first year of the pandemic compared to the year before. Given resonance is defined as how future information conforms with the novelty of the current day, for resonance to have a higher value, novelty must be correspondingly high. Thus, the variance shift indicates the emotion distributions after the first change point were more similar compared to those in the previous time period.

Intuitively, there seems to be a strong seasonal effect around Christmas and New Year in both years even though the algorithm only detected a change in the signal in 2020 and not in 2019. However, the effect of the holidays was different between the two years. In 2020, a longer period with increased variance in resonance was captured while the increase in variance in 2019 was very abrupt and quickly returned to previous levels. This makes the spiky behavior in 2019 seem more like an event than a new period, which might explain why it was not detected by the algorithm. If the time series had included multiple Christmas holidays one could have adjusted the signal for the seasonal component, however, given only two seasons were included this was deemed unfeasible.

While it can be discussed whether the first holiday spike should be considered its own period, and thus whether penalization on the number of change points was too high, the other change points fit well with our intuition on visually detectable changes in the signal. Choosing a penalty parameter will always have some degree of arbitrariness (Haynes, Fearnhead, & Eckley, 2017), but we decided on the rather high penalization on the number of change points. We did this as we were interested in investigating changes in overall dynamics throughout longer periods, in which case fewer change points with more data in each period are better suited for the subsequent analysis. Furthermore, a lower penalty could lead to a higher risk of finding change points that were only an expression of random fluctuations in the signal rather than actual changes in the overall dynamics (Truong et al., 2020).

4.2 Changes in emotion dynamics

4.2.1 Resonance-novelty coupling (SMØ)

To elucidate how the dynamics on Twitter changed in the four change point periods, we looked at the coupling between resonance and novelty, formalized by the $N \times R$ slope in the linear regression model (see table 3.3 and figure 3.3). This approach was inspired by a study investigating the changing dynamics of information in Danish newspapers during the COVID-19 pandemic (Nielbo et al., 2021). The largest difference in $N \times R$ slope was found between the first and second periods. Moreover, the 95% confidence intervals of these two estimates were not overlapping, supporting a change in the coupling from the first to the second change point period.

The estimated $N \times R$ slope in change point period 4 lay between the estimated slopes in change points periods 1 and 2. However, the confidence intervals overlapped more with the estimation of the slope in the second change point period than in the first. This could indicate that the dynamics from the second change point period persisted after the Christmas holidays. Since the slope estimate had decreased, it could also indicate that the dynamics were gradually returning to levels similar to what they were before COVID-19.

The third change point period only included a few days, which would explain the great uncertainty of the estimated $N \times R$ slope in this time window. With the large overlap of confidence intervals, it is difficult to draw reasonable conclusions from this estimate.

Taken together, these findings indicate a stronger coupling between resonance and novelty of emotional content on Danish twitter after the COVID-19 pandemic struck in Denmark. Thus, when more novel emotional content was expressed in the tweets, the more the use of similar emotions persisted in tweets within a window of 3 days. This finding stands in contrast to the findings of topics on news media, where a decoupling between resonance and novelty was found during the first lockdown hit in Denmark (Nielbo et al., 2021).

This effect could be due to a number of things. Firstly, reactions on social media usually happen rapidly after an event (Saeed et al., 2019) and the signal in the current study was calculated using a window of 3 days. One could imagine that under normal circumstances, the reaction to an event would have disappeared after 3 days and the emotional content at this time would already have changed. The stronger coupling between resonance and novelty of the emotion distributions found in the second change point period could indicate that novel emotional content (e.g. reactions to a specific event) lasted for longer during the first months after the COVID-19 pandemic hit in Denmark.

Another factor that could have influenced the results is the timing of the change points. As mentioned, the first change point period includes the extreme values expressed around Christmas and New Year, while these extremes are not included in the second change point period. The more extreme values in the first change point period are also visually apparent in figure 3.3, and these values could have affected the estimation of the coupling between resonance and novelty.

The windowed relative entropy signals were calculated on tweets up until April 2021, however, important events related to COVID-19 occurred after this time point. On September 10th, 2021 the Danish health ministry declared that COVID-19 no longer was a critical threat to society after a long period with few cases (Statens Serum Institut, 2022). Later the same year, the number of infected with COVID-19 in Denmark began to rise due to the new Omicron variant, resulting in the virus being classified as a critical threat to society again in November 2021. In 2022, this classification was once again removed and most restrictions lifted, thus, society went back to a state resembling the one before the pandemic (Statens Serum Institut, 2022).

Given the findings of this study, it would be interesting to see how these events affected the emotion dynamics on Twitter. Extending the analysis to the rest of 2021 and the beginning of 2022 could reveal if the dynamics in emotions around September 2021 would resemble those expressed in the first change point period. Moreover, it would be interesting to investigate if the increased number of cases and stringency in restrictions at the end of 2021 would result in a similar change in the dynamics as was found when comparing the first two change point periods in the current study.

4.2.2 Changes in individual emotions (SNL)

Linear regression models and correlations were used to investigate the extent to which the resonance signal can be characterized by the individual emotions and the interplay between the emotions. The confidence intervals of the β-estimates for the linear regression models were all very wide and crossing zero (Table 3.4), indicating that none of the emotions were predictive of resonance in any of the four change point periods. This finding suggests that dynamics in one emotion are never single-handedly driving the dynamics of the resonance signal but that the signal in all four change point periods is a result of the interplay between the emotions.

In a paper by Münnix et al. (2012), correlations of time series signals were used to identify points of drastic changes in complex systems. Inspired by this approach, we calculated correlations between the eight emotion time series signals to investigate changes in the states of the system in the four change point periods. The correlations in the third change point period were generally stronger than the correlations found in the three other time periods (Figure 3.5). This time period showed strong positive correlations between the two emotions *happiness* and *trust*, and a positive correlation between the four emotions *surprised*, *anger*, *contempt*, and *fear*, while the emotions were negatively correlated across these two clusters.

More interestingly, the correlation matrices seem to uncover changing dynamics in the interplay between the eight different emotions in the three other change point periods too. While most of the correlations were weak to moderate (Schober, Boer, & Schwarte, 2018), there were interesting variations in the patterns across periods. The patterns were similar for some emotions; for example, a negative correlation between *expectation* and each of the three emotions *anger*, *contempt*, and *grief*, and a positive correlation between *trust* and *happiness* were found in all change point periods. For other emotions, the patterns varied between periods. An example of this is the correlation between *anger* and *surprised*; they were positively correlated in period 1, and even more so in period 3. However, they didn't show any correlation in change point period 2, and in the fourth change point period, the direction had changed to a negative correlation.

Even though these findings are more exploratory than definitive, they indicate that changes in the dynamics of emotions in tweets can be captured using correlations. In the analysis of the present study, the correlation matrices were used to describe the signal within already defined time windows. In the study by Münnix et al. (2012), the correlations between time series were clustered and used to identify points of change in the underlying state of the overall system. Using a similar approach on the time series signals presented in the current study might reveal interesting dynamics that are hidden within the periods defined by the change points, or they might serve as a way to confirm the change points found with the PELT algorithm.

4.3 Future studies (SMØ)

We have already outlined a few interesting directions for the analysis to go from here, namely using the correlation matrices as input for a clustering algorithm, and extending the analysis to more recent data. Another interesting approach would be to take advantage of the fact that a lot of data from COVID-19 including temporal progressions is available (Statens Serum Institut, 2022; Ritchie et al., 2020; Google, 2021).

Zhou et al. (2020) used information on the daily number of tests, the daily number of positive cases, and the most likely source of infection in the analysis of Twitter sentiment, and Li et al. (2020) investigated the difference in social media sentiment between Chinese and American citizens, including the number of COVID-19 related tweets and the number of positive cases in the analysis. Inspired by these studies, including temporal information on the daily number of positive cases or COVID-19 related hospitalizations and deaths could illuminate some of the mechanisms in the dynamics of the emotional content on Twitter.

It could also be interesting to look at more fine-grained analyses of the signal, possibly with a top-down approach to identifying changes in resonance. This could be done by using important events from the timeline of COVID-19 events (Statens Serum Institut, 2022) for separating the signal into different windows, and thus investigate emotional discourse on Twitter within externally defined periods. Another interesting approach could be dividing the tweets into subsets, e.g. based on whether they talk about COVID-19, whether they were posted by a member of the Danish government, or looking at politicians from different positions on the political spectrum. Looking at only a subset of Twitter has previously been done (Li et al., 2020; Buccoliero et al., 2020) and one might expect the emotional dynamics of social media to vary between these different subgroups.

5 Conclusion (SMØ & SNL)

The aim of this paper was to characterize the windowed relative entropy measure resonance calculated using emotion distributions of Danish tweets and identify dynamics on social media during the COVID-19 pandemic in Denmark. Using offline change point detection, three change points were found, thus dividing the signal into four distinct periods. The detected change points indicated a change in the dynamics of emotions when the COVID-19 pandemic appeared in Denmark, only interrupted by a short change in dynamics around the Christmas holidays in 2020. Stronger coupling between the entropy measures resonance and novelty was found in the period during the COVID-19 pandemic compared to that before, indicating that in this period the reactions to novel events lasted longer. Moreover, using correlations between the time series signals of the probability of the individual emotions, visually different states of the systems were identified. This effect was most pronounced in the time window containing the

Christmas season of 2021, however, the findings also indicated shifting patterns before and after the COVID-19 pandemic. Further studies of the correlation matrices would be necessary to draw a more definitive conclusion and could possibly support or reveal new insights to the findings of the current paper.

References

Barron, A. T. J., Huang, J., Spang, R. L., & DeDeo, S. (2018, May). Individuals, institutions, and innovation in the debates of the French Revolution. *Proceedings of the National Academy of Sciences*, *115*(18), 4607–4612. Retrieved 2021-12-17, from http://www.pnas.org/lookup/doi/10.1073/pnas.1717729115 doi: 10.1073/pnas.1717729115

Bor, A., Jørgensen, F. J., & Petersen, M. B. (2021, December). *The COVID-19 Pandemic Eroded System Support But Not Social Solidarity* (Tech. Rep.). PsyArXiv. Retrieved 2022-05-29, from https://psyarxiv.com/qjmct/ doi: 10.31234/osf.io/qjmct

Buccoliero, L., Bellio, E., Crestini, G., & Arkoudas, A. (2020, January). Twitter and politics: Evidence from the US presidential elections 2016. *Journal of Marketing Communications*, *26*(1), 88–114. Retrieved 2021-12-13, from https://doi.org/10.1080/13527266.2018.1504228 doi: 10.1080/13527266.2018.1504228

Chancellor, S., & De Choudhury, M. (2020, March). Methods in predictive techniques for mental health status on social media: a critical review. *npj Digital Medicine*, *3*(1), 1–11. Retrieved 2022-05-15, from https://www.nature.com/articles/s41746-020-0233-7 doi: 10.1038/s41746-020-0233-7

Dubey, A. D. (2020). Twitter Sentiment Analysis during COVID19 Outbreak. *SSRN Electronic Journal*. Retrieved 2022-05-15, from https://www.ssrn.com/abstract=3572023 doi: 10.2139/ssrn.3572023

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & van der Linden, S. (2021, June). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606–623. Retrieved

2021-12-13, from https://doi.org/10.1177/1368430220987596 doi: 10.1177/1368430220987596

Google. (2021). *Google COVID-19 community mobility reports.* Retrieved 2021-12-21, from https://www.google.com/covid19/mobility/

Haynes, K., Fearnhead, P., & Eckley, I. A. (2017, September). A computationally efficient nonparametric approach for changepoint detection. *Statistics and Computing*, *27*(5), 1293–1305. Retrieved 2022-05-28, from https://doi.org/10.1007/s11222-016-9687-5 doi: 10.1007/s11222-016-9687-5

Kemp, S. (2021, October). *Digital 2021 October Global Statshot Report.* Retrieved 2022-05-18, from https://datareportal.com/reports/digital-2021-october-global-statshot

Li, X., Zhou, M., Wu, J., Yuan, A., Wu, F., & Li, J. (2020, June). *Analyzing COVID-19 on Online Social Media: Trends, Sentiments and Emotions* (Tech. Rep. No. arXiv:2005.14464). arXiv. Retrieved 2022-05-15, from http://arxiv.org/abs/2005.14464

Murdock, J., Allen, C., & DeDeo, S. (2017, February). Exploration and exploitation of Victorian science in Darwin's reading notebooks. *Cognition*, *159*, 117–126. Retrieved 2021-12-17, from https://linkinghub.elsevier.com/retrieve/pii/S0010027716302840 doi: 10.1016/j.cognition.2016.11.012

Münnix, M. C., Shimada, T., Schäfer, R., Leyvraz, F., Seligman, T. H., Guhr, T., & Stanley, H. E. (2012, December). Identifying States of a Financial Market. *Scientific Reports*, *2*(1), 644. Retrieved 2022-05-17, from http://www.nature.com/articles/srep00644 doi: 10.1038/srep00644

Nielbo, K. L., Haestrup, F., Enevoldsen, K. C., Vahlstrup, P. B., Baglini, R. B., & Roepstorff, A. (2021, February). *When no news is bad news – Detection of negative events from news media content* (Tech. Rep.

No. arXiv:2102.06505). arXiv. Retrieved 2022-05-14, from http://arxiv.org/abs/2102.06505

Pauli, A. B., Barrett, M., Lacroix, O., & Hvingelby, R. (2021, May). DaNLP: An open-source toolkit for Danish Natural Language Processing. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)* (pp. 460–466). Reykjavik, Iceland (Online): Linköping University Electronic Press, Sweden. Retrieved 2021-12-15, from https://aclanthology.org/2021.nodalida-main.53

Ritchie, H., Mathieu, E., Rodés-Guirao, L., Appel, C., Giattino, C., Ortiz-Ospina, E., ... Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*.

Saeed, Z., Abbasi, R. A., Maqbool, O., Sadaf, A., Razzak, I., Daud, A., ... Xu, G. (2019, June). What's Happening Around the World? A Survey and Framework on Event Detection Techniques on Twitter. *Journal of Grid Computing*, *17*(2), 279–312. Retrieved 2021-12-18, from http://link.springer.com/10.1007/s10723-019-09482-2 doi: 10.1007/s10723-019-09482-2

Schober, P., Boer, C., & Schwarte, L. A. (2018, May). Correlation Coefficients: Appropriate Use and Interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768. Retrieved 2022-05-24, from https://journals.lww.com/anesthesia-analgesia/fulltext/2018/05000/correlation_coefficients__appropriate_use_and.50.aspx doi: 10.1213/ANE.0000000000002864

Sharma, P., & Moh, T.-S. (2016, December). Prediction of Indian election using sentiment analysis on Hindi Twitter. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 1966–1971). doi: 10.1109/BigData.2016.7840818

Shen, D., Urquhart, A., & Wang, P. (2019, January). Does twitter predict Bitcoin? *Economics Letters*, *174*, 118–122. Retrieved 2021-12-13, from https://www.sciencedirect.com/science/article/pii/S0165176518304634 doi: 10.1016/j.econlet.2018.11.007

Statens Serum Institut. (2022, April). *Tidslinje for covid-19.* SSI. Retrieved 2022-05-14, from https://www.ssi.dk/-/media/arkiv/subsites/covid19/presse/tidslinje-over-covid-19/covid-19-tidslinje-for-2020-2022-lang-version---version-1---april-2022.pdf?la=da

Thelwall, M. (2014). Sentiment analysis and time series with Twitter. In *Twitter and Society* (Katrin Weller, Axel Bruns, Jean Burgess, Merja Mahrt and Cornelius Puschmann ed., Vol. 89). New York: Peter Lang. Retrieved 2021-12-13, from https://core.ac.uk/reader/19541510#page=83

Truong, C., Oudre, L., & Vayatis, N. (2020, February). Selective review of offline change point detection methods. *Signal Processing*, *167*, 107299. Retrieved 2022-05-13, from https://linkinghub.elsevier.com/retrieve/pii/S0165168419303494 doi: 10.1016/j.sigpro.2019.107299

Vliet, L. v., Törnberg, P., & Uitermark, J. (2020, September). The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries. *PLOS ONE*, *15*(9), e0237073. Retrieved 2021-12-13, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237073 doi: 10.1371/journal.pone.0237073

Yin, H., Yang, S., & Li, J. (2020). Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media. In X. Yang, C.-D. Wang, M. S. Islam, & Z. Zhang (Eds.), *Advanced Data Mining and Applications* (pp. 610–623). Cham: Springer International Publishing. doi: 10.1007/

978-3-030-65390-3_46

Zhou, J., Yang, S., Xiao, C., & Chen, F. (2020, June). *Examination of Community Sentiment Dynamics due to COVID-19 Pandemic: A Case Study from A State in Australia* (Tech. Rep.). Retrieved 2022-05-30, from http://arxiv.org/abs/2006.12185 doi: 10.1007/s42979-021-00596-7

A Appendix

A.1 Resonance and novelty signals

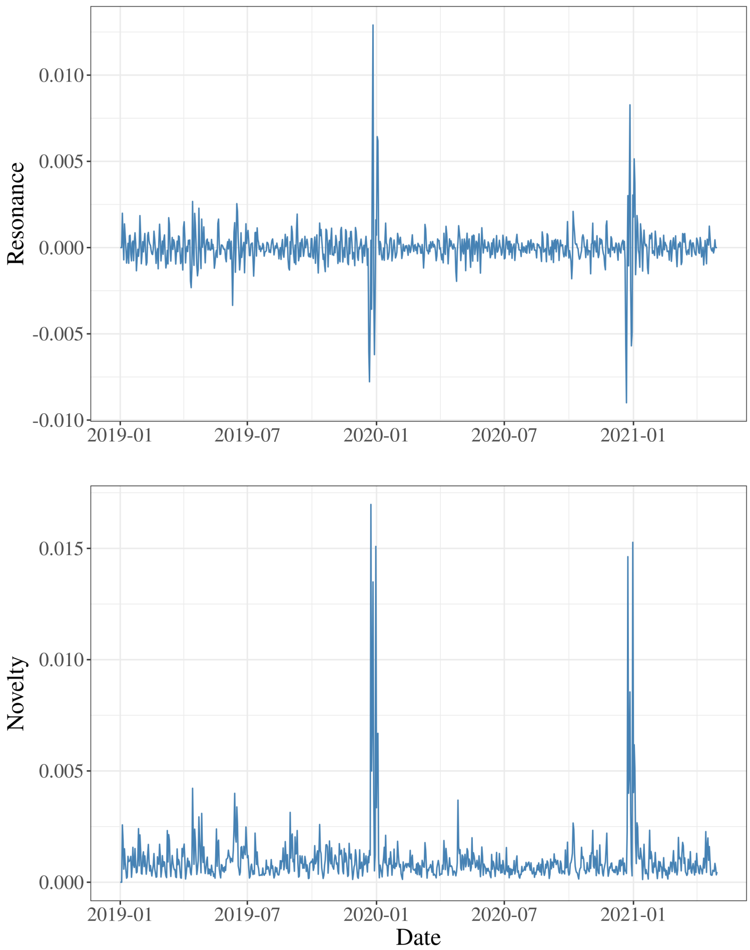


Figure A.1: Raw resonance and novelty signals.

A.2 Natural language processing exam

Characterizing Dynamics on Twitter using BERT Emotion Classifications

Synopsis Natural Language Processing

Stine Nyhus Larsen - SNL (201808306)
Sara Møller Østergaard - SMØ (201808643)

M.Sc. Cognitive Science, Aarhus University

January 3rd, 2022

Abstract

Research using Twitter data has become increasingly popular in the last few years. This study investigates whether the temporal dynamics of Twitter can be characterized with windowed relative entropy measures using probability distributions based on the Danish BERT Emotion model as latent variables. We find that this method does seem to capture changes in discourse on Twitter over time. The validity of the method is supported by repeating the analysis using BERT Tone polarity distributions as latent variables. Advantages of the approach are discussed along with possible limitations. Hereafter, the signal is explored using linear models prediction resonance from novelty, and the influence of individual emotions is assessed using Pearson's r. The findings of this study lays the ground for future studies of dynamics of emotions on Twitter, and possible directions of this is considered.

Link to GitHub: https://github.com/saraoe/emoDynamics

1

1 Introduction (SNL & SMØ)

Twitter has in the past years become increasingly popular for research (Karami, Lundy, Webb, & Dwivedi, 2020), and has been used as a database for studying e.g. politics and elections (Buccoliero, Bellio, Crestini, & Arkoudas, 2020; Vliet, Törnberg, & Uitermark, 2020), conspiracy theories (Fong, Roozenbeek, Goldwert, Rathje, & van der Linden, 2021; Visentin, Tuan, & Di Domenico, 2021), Bitcoin tradings (Shen, Urquhart, & Wang, 2019), reactions to COVID-19 outbreaks and restrictions (Lwin et al., 2020; Rufai & Bunce, 2020), to just mention a few. The advantage of using a Twitter-based dataset is that Twitter is publicly available and commonly used by both high-profile politicians, opinion leaders, and the general public (Thelwall, 2014). It is therefore relatively easy to acquire a large dataset generated by diverse users. Furthermore, tweets are time-stamped, meaning that tweets gathered over a period of time can be analyzed as a time series system. Twitter data can thus be used to assess public opinion towards different topics, and related changes in sentiment, over time (Thelwall, 2014).

In this study, we take our starting point in the dynamic nature of Twitter data and in combination with methods from the field of natural language processing (NLP), we investigate the following research question:

> *Can a windowed relative entropy approach based on emotions be used to characterize the dynamics of Twitter?*

We apply BERT models trained for Danish corpora to Danish tweets scraped between 01/01 2019 and 04/30 2021 to get probability distributions from two different emotion classifiers; *emotion* (using eight emotion categories) and *polarity* (using categories negative, neutral, and positive) to hold the findings for emotion against an alternative. We then investigate the characteristics and dynamics of these distributions using the information dynamic measures *novelty* and *resonance*. The characteristics and dynamics are visualized and further quantified using linear regression models and Pearson correlation coefficients.

1.1 Information Dynamics (SMØ)

In recent studies, windowed relative entropy has been applied to text corpora to explore the information dynamics in unstructured data (Barron, Huang, Spang, & DeDeo, 2018; Nielbo, Enevoldsen, Baglini, Fano, & Gao, 2021). This method captures information *novelty*, how unexpected the content is compared to previous content, and *resonance*, which represents the degree to which the novel content persists into future information.

Windowed relative entropy has been applied to quite diverse systems, with examples such as sensory information (Itti & Baldi, 2009), the development of Darwin's thinking (Murdock, Allen, & DeDeo, 2017), parliament speeches (Barron et al., 2018), and newspaper articles (Nielbo, Haestrup, et al., 2021). Though the systems in these examples are quite different, having a quantitative measure of information makes it possible to apply the same mathematical definition of entropy to all of them. The digitization of text allows for the use of NLP methods to measure the information in texts at different times. Commonly, for the studies of the dynamics of Darwin's thinking, parliament speeches, and newspaper articles, topic models were fitted to the texts and used as the latent variable for windowed relative entropy. Thus, relative surprise was modelled using a distribution of topics. The quantitative results of this approach have been found to align well with previously qualitative findings, indicating that windowed relative entropy does capture the intended dynamics of the system (Barron et al., 2018). In addition, this method could illuminate dynamics that might previously have been overlooked.

In this study, we wish to examine whether windowed relative entropy can elucidate the dynamics that characterize social media. Sentiment analysis is a commonly used method to characterize different states on social media (Thelwall, Buckley, & Paltoglou, 2011), due to the more pronounced emotional valence found in social media corpora compared to e.g. newspaper articles (Welbers & Opgenhaffen, 2019). With this in mind, the emotions of tweets have been chosen as the measure of information used in the present study. Previously, information dynamics in newspapers during the first phase of COVID-19 have been used to examine national response strategies to the pandemic in Denmark and Sweden (Nielbo, Enevoldsen, et al., 2021), giving rise to the assumption that equally interesting dynamics could be found in social media data during the COVID-19 outbreak. The focus of this study is not concerned with interpreting and investigating the actual dynamics, but rather combining relative entropy and NLP methods to investigate whether it is possible to quantify the dynamics on Twitter using emotion classification in the first place.

1.2 Introducing BERT (SNL)

BERT (Bidirectional Encoder Representations from Transformers) is a pioneering language model developed by Devlin, Chang, Lee, and Toutanova (2019). It builds on the *Transformer* model developed by Vaswani et al. (2017), which uses an encoder-decoder architecture to process the input data and produce the desired output. A key feature of the Transformer model is the use of multi-headed self-attention mechanisms in the encoder. This attention mechanism makes recurrent elements of previous language models redundant, thus, allowing for more computationally efficient models (Vaswani et al., 2017). The encoder architecture of Transformers consists of 6 identical modules stacked on top of each other of, each containing both multi-head self-attention and feed-forward layers. Because all 6 layers of the encoder include an attention mechanism, each position can attend to inputs at all positions of the previous layer (Vaswani et al., 2017). The advantage of this approach is that the model can take all previous tokens into account simultaneously, thus the context of each token is not limited to words positioned near by.

Furthermore, as BERT uses WordPiece tokenization, a token does not necessarily refer to an actual word, but can refer to parts of words, e.g. common endings such as ”ing” (Devlin et al., 2019). This means that multiple instances of ”ing” can be treated as the same token, which makes good conceptual sense and greatly reduces the number of token representations to learn. It creates a model with a sophisticated and general understanding of language that can be decoded to produce an output.

Broadly speaking, BERT models consist of a number of layers (12-24), each containing Transformer encoder blocks as described above. The attention blocks of the BERT model are bidirectional, meaning that attention to context following a token is also included when training the token representations (Devlin et al., 2019).

For the training procedure, a `[clf]` token is added at the beginning of each sentence, and a `[sep]` token at the end. The `[clf]` token will attend to the entire sequence, thus, this token can later be used for classification of an entire sentence. During training, the model receives a sequence as input and is optimized to output the same sequence. The issue is that the mapping between the tokens of the input and the output sequence is direct. In order to avoid this problem, Masked Language Modelling (MLM) is used. MLM is implemented by masking 15% of the final hidden vectors and feeding them to an output softmax for prediction. The masking consists 80% of the time of replacing the original token with a `[MASK]` token, 10% of the time of replacing the original token with a random token, and 10% of the time of leaving the token unchanged. This prevents the model from learning identical weights for all `[MASK]` tokens, but instead a representation for each unique token.

Together with MLM, the training of BERT consists of Next Sentence Prediction (NSP). This is done in order for the model to learn relationships between sentences and paragraphs and not just within. NSP is a binary classification task in which the model is presented with a number of sentence pairs, sentences A and B, generated from the data corpus. 50% of the time, sentence B does indeed follow sentence A, and 50% of the time, it does not.

After training, BERT has a very complex and advanced representation of language. It can then finally, and relatively easy, be fine-tuned to specific tasks, such as question answering, sentence completion and a range of classification tasks by exploiting the model's general understanding of language from MLM and NSP. BERT have shown to reach state-of-the-art performance on a wide range of NLP tasks (Devlin et al., 2019).

2 Methods

2.1 Data and Preprocessing (SNL)

The corpus used for the analysis consisted of tweets from the period between 01/01 2019 and 04/30 2021. The tweets were scraped from Twitter using a set of common Nordic words (see appendix A.1). Subsequently, Twitter's language classifier was used to separate the Tweets into languages. For this analysis only the Danish tweets were used.

A few preprocessing steps were applied before running the BERT models. Firstly, all retweets were excluded from the analysis. Secondly, emojis, hashtags, mentions, and urls were removed from the remaining tweets, so the tweets only included the actual text. The total number of tweets included in the final analysis was 32,499,007.

2.2 BERT models: Emotion and Tone (SMØ)

We used two different BERT models trained by DaNLP (Pauli, Barrett, Lacroix, & Hvingelby, 2021) to classify the emotion probabilities of the tweets: the Danish BERT Emotion and the Danish BERT Tone. Both models rely on the pre-trained Danish BERT model (*Nordic BERT*, 2021) on which they are fine-tuned to their specific classification tasks. The Danish BERT model is trained on all Danish text from Common Crawl, Danish Wikipeida, Danish OpenSubtitles, together with custom scraped data from *dindebat.dk* and *hestenettet.dk* (*Nordic BERT*, 2021).

The Danish BERT Emotion model is fine-tuned to classify the emotion of Danish texts. The model is trained for two classification tasks: binary classification of whether the text is emotional or not, and subsequently classifying the emotion label (*Sentiment Analysis — DaNLP documentation*, 2020). The model is fine-tuned for these two tasks using a manually annotated dataset consisting of social media data (*Sentiment Analysis — DaNLP documentation*, 2020). For the classification of the emotion labels, the model classifies between eight different emotions (see table 2.1). Thus, it is possible to extract a probability distribution of eight probabilities from the model, one probability for each emotion.

Danish Labels	**Translations**
Glæde/Sindsro	Happiness/Calmness
Tillid/Accept	Trust/Acceptance
Forventning/Interrese	Expectation/Interest
Overasket/Målløs	Surprised/Speechless
Vrede/Irritation	Anger/Irritation
Foragt/Modvilje	Comtempt/Reluctance
Sorg/Trist	Grief/Sadness
Frygt/Bekymret	Fear/Worry

Table 2.1: The eight emotion categories in the Danish BERT Emotion model together with our English translations.

The Danish BERT Tone model is fine-tuned on manually annotated datasets consisting of tweets and sentences from the Europe Parlament (*Sentiment Analysis — DaNLP documentation*, 2020), and it too consists of two classifiers: one classifying the polarity of a Danish text (i.e. positive, negative, neutral), and one giving a binary classification of either subjective or objective (*Sentiment Analysis — DaNLP documentation*, 2020). In this study, only the polarity scores were used.

We summarized the probability distributions by averaging the probability for each emotion over one day, thus getting a single distribution of emotions for each day. This summarization allows for daily changes in emotional content to be captured, while also ensuring a roughly equal number of tweets in each summarized distribution. A shorter window could lead to higher differences in number of tweets, as it would be assumed that activity on Twitter is not constant during a day. Identical summarization was applied to the polarity distributions, making it possible to compare the signals from the two classifiers.

2.3 Windowed relative entropy (SNL)

We used Jensen-Shannon divergence (JSD) to quantify the amount of surprise between two probability distributions. The advantage of JSD over the closely related Kullback-Leibler divergence (KLD) is that it is symmetrical and smoothed, making it a distance metric (Reis, 2020). JSD is calculated as

$$JSD(s^{(j)}|s^{(k)}) = \frac{1}{2}D(s^{(j)}|M) + \frac{1}{2}D(s^{(k)}|M)$$

Here, $s^{(j)}$ is the probability distribution at the j'th day (and similarly for $s^{(k)}$), $M = \frac{1}{2}(s^{(j)} + s^{(k)})$, and D is KLD, which is defined as

$$D(s^{(j)}|s^{k}) = \sum_{i=1}^{K} s_i^{(j)} log \frac{s_i^{(j)}}{s_i^{(j)}}$$

K corresponds to the number of labels in the probability distribution $s^{(j)}$.

The emotion probability distributions of the BERT models were used as latent variables for the information dynamics measures *novelty*, *transience*, and *resonance*. These measures were calculated following previous definitions (Barron et al., 2018; Nielbo, Haestrup, et al., 2021). Novelty of the j'th distribution was calculated as

$$N_w(j) = \frac{1}{w}\sum_{d=1}^{w} JSD\left(s^{(j)}|s^{(j-d)}\right)$$

Here, w is the window size. Novelty of the probability distribution of a given day is thus the mean of the entropy between that distribution and the w previous distributions.

Similarly, transience for the jth distribution was calculated as

$$T_w(j) = \frac{1}{w} \sum_{d=1}^{w} JSD\left(s^{(j)} | s^{(j+d)}\right)$$

Transience of the probability distribution of a given day is thus the mean of entropy between that distribution and the w subsequent distributions. Finally, resonance was calculated as

$$R_w(j) = N_w(j) - T_w(j)$$

2.4 Assessing the Signal (SMØ)

The information signals *novelty* and *resonance* were calculated for the probability distribution from the emotion classification averaged over one day. A window of three days ($w = 3$) was chosen for the analysis, meaning that resonance for each day was calculated relative to three previous and three following days. The chosen window can be thought of as deciding the granularity of the analysis. The longer the window, the less fine-grained the analysis. In newspapers, a typical cycle is seven days (Nielbo, Haestrup, et al., 2021), but dynamics change much more quickly on social media (Saeed et al., 2019). By setting a window of three days, we aim to capture the main fluctuations in emotional discourse related to external events.

2.4.1 Linear models (SMØ)

Using the information signals calculated from the emotion distributions, we fitted a linear regression using ordinary least squares that predict resonance from novelty:

$$R = \beta_0 + \beta_1 N + \epsilon$$

Here β_1 is the $N \times R$ slope, β_0 is the intercept, and ϵ the error of the fit.

Regression lines were fitted for 1) the full signal, 2) the signals before and after March 11th, 2020 where the first lockdown in Denmark was announced, and 3) the signals where data in the period between December 1st to January 2nd was removed for both years. The latter regression was fitted in order to assess whether the signal was affected by a seasonal component around Christmas and New Year.

2.4.2 Polarity (SNL)

To assess the validity of using emotions to characterize the dynamics on Twitter, we calculated novelty and resonance from the polarity probabilities using the same window size ($w = 3$) and compared the information signals. To quantify this comparison, we calculated Pearson correlation coefficient (Pearson's r) between the raw resonance signals calculated from the emotion distributions and the polarity distributions, respectively. The Pearson correlation coefficient is a measure of co-variance over time for continuous signals (Cheong, 2019), and $r = -1$ indicates perfect negative correlation, $r = 1$ indicates perfect positive correlation while $r = 0$ indicates no correlation.

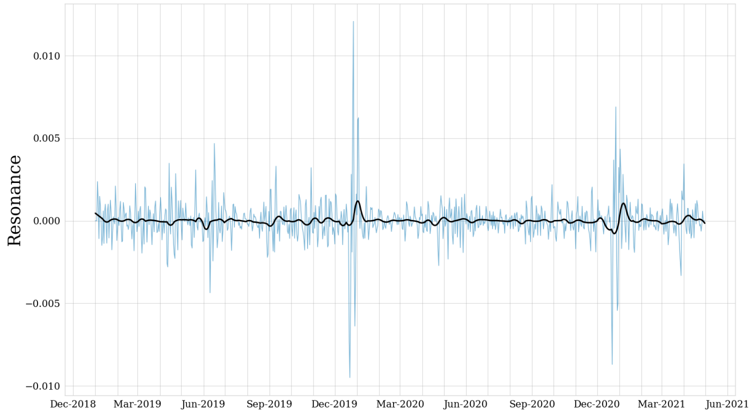
2.4.3 Individual emotions (SNL)

In order to investigate the extend to which individual emotions drive the effects, we re-calculated novelty and resonance signals leaving out one emotion at a time. We ensured that the distributions still summed to one by calculating the percentage of the remaining probabilities every emotion represented. By assessing the information signal when leaving out one emotion, we could get a sense of the degree to which each emotion was important for the overall signal. To quantify this, we calculated Pearson's r between the resonance signals of the full emotion distribution and the distribution where one emotion was left out.

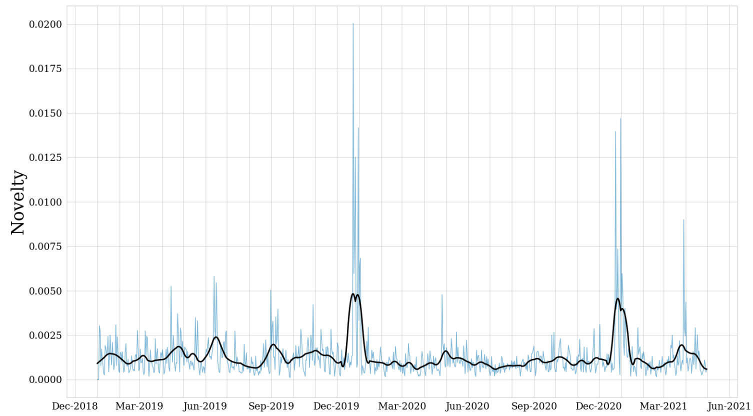
3 Results

3.1 Novelty and resonance signals (SMØ)

Resonance and novelty signals using window $w = 3$ calculated from probability distributions of emotions from the Danish BERT Emotion model are visualized in figure 3.1. We observe clear and easily detectable tendencies in the signals.



(a) Resonance



(b) Novelty

Figure 3.1. Resonance and novelty calculated with a window $w = 3$. The blue line represents the raw signal, and the black line is the signal smoothed using a non-linear adaptive filter (Gao, Hu, & Tung, 2011).

3.2 Linear models (SNL)

Figure 3.2 shows the $N \times R$ slopes β_1 in the respective time windows. The slopes was found to be positive for all fitted regressions, indicating that novel emotions resonate more than non-novel emotions. There seems to be a stronger association between resonance and novelty when excluding Christmas and New Year ($\beta_1 = 0.62$, SE $= 0.028$) compared to the association of the two information signals when looking at the entire time period ($\beta_1 = 0.53$, SE $= 0.029$). When comparing the $N \times R$ slopes before and after the first COVID-19 lockdown in Denmark (March 11th, 2020), we found a similar association between resonance and novelty before the first lockdown ($\beta_1 = 0.54$, SE $= 0.040$) and after the first lockdown ($\beta_1 = 0.53$, SE $= 0.042$).

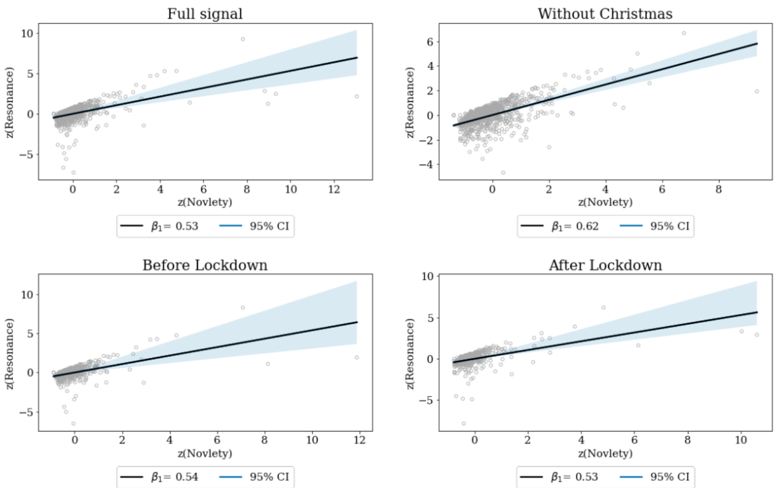


Figure 3.2. $N \times R$ slopes β_1 from 4 different time windows: for the full signal (upper left), without Christmas (upper right), before the first COVID-19 lockdown (lower left), and after the first COVID-19 lockdown (lower right).

3.3 Polarity (SMØ)

Resonance and novelty signals using window $w = 3$, calculated from probability distributions of both emotion and polarity, are visualized in figure 3.3. The two signals seem to capture similar dynamics when eyeballing the figure.

Figure 3.3. Resonance (top) and novelty (bottom) calculated with a window $w = 3$. The blue line represents the emotion signal, and the red line is the polarity signal

To quantify the similarity between the two signals, we calculated Pearson's r. The correlation between the emotion resonance signal and the polarity resonance signal was $r = 0.79$ ($p < .0056$). This threshold is the Bonferroni correction of a 0.05 threshold with 9 comparisons.

3.4 Individual emotions (SNL)

In order to delineate the effects of individual emotions, we calculated correlation coefficients between the full resonance signal (i.e. the signal calculated from the probabilities of all eight emotions) and the resonance signal with one emotion left out. This was done for all eight emotions. The correlation coefficients can be found in table 3.1. Resonance signals for the distributions with each emotion excluded are visualized in figure A.1 in the appendix.

Emotion excluded	Foragt	Forventning	Frygt	Glæde	Overrasket	Sorg	Tillid	Vrede
Correlation coefficient	0.937	0.850	0.990	0.850	0.998	0.997	0.994	0.997

Table 3.1: Pearson's r between resonance signal for the full emotion distribution and the signal for the re-scaled distribution where the emotion in question has been excluded. All correlations are significant when corrected for multiple comparisons using Bonferroni correction (significance level of 0.0056).

4 Discussion (SMØ)

In this study, we investigated whether a windowed relative entropy approach applied to emotion distributions can be used to characterize the dynamics of Twitter. We do indeed find that the information signals novelty and resonance calculated from BERT Emotion probability distributions capture visually observable dynamics in Twitter discourse. For example, large spikes in novelty are evident around Christmas and New Year, indicating that novel distributions of emotions are introduced around these events.

The analysis was repeated using BERT Tone polarity distributions as latent variables for the information dynamic measures. This was done to address whether the signal found for BERT Emotion probability distributions would generalize to a conceptually similar classification. Visually, the two signals are very similar (see figure 3.3), and the Pearson correlation coefficient between the two resonance signals was 0.79 with significant p-value. This indicates that the dynamics found are generalizable across models and that the methods applied in this study are appropriate for detecting the dynamics of Twitter.

While the validity of the method applied in this study is generally supported by our findings, there are also some limitations to the approach. Generally, windowed relative entropy, as used in this study and in previous studies (Barron et al., 2018; Murdock et al., 2017; Nielbo, Enevoldsen, et al., 2021; Nielbo, Haestrup, et al., 2021), involves a number of more or less arbitrary choices. One of these is the size of the window. In the study of information dynamics in newspaper articles, a window size of 7 days was used, corresponding to a newscycle (Nielbo, Haestrup, et al., 2021). This study used a window of 3 days based on the fact that reactions on Twitter happen in real-time (Saeed et al., 2019). However, using a window of a different size could yield slightly different results, and experimenting with this could be interesting. In line with this, the probability distributions were summarized for one day, but if one wanted to look at more fine-grained dynamics, summarizing the tweets per hour and looking at hourly changes in dynamics might be interesting for social media.

Summarizing the emotion distributions on a daily basis also introduced another potential loss of information. When using the mean probability of each emotion for the analysis, variance is not taken into account. Potentially, some emotions could have similar means between two days, but very dissimilar variance due to e.g. more polarized use of those emotions. Taking the standard deviation of the daily averages into account, for example through weighting, might elucidate other interesting dynamics.

The Danish BERT Emotion model was used to extract the emotion probability distribution for tweets. This model is trained on two classification tasks: binary emotion classification (i.e., emotional or not) and then classifying the emotional texts into one of eight emotion categories. However, the binary classification of emotional or non-emotional was not used in this study, meaning that all tweets received a probability for all eight emotions regardless of the binary label. This is potentially a problem, both due to how the emotion category classifier is trained, and as non-emotional tweets would be characterized by probabilities of different emotions (see labels in table 2.1). However, the BERT Tone polarity distributions do not suffer from this problem because the labels include a neutral category. As we found the two signals to capture similar dynamics, it could be argued that discarding the binary classification in this study is not a problem.

4.1 Exploring the signal (SNL)

Our analysis indicated that Twitter can actually be described using emotion distributions. We now move on to discussing the exploration of the signal.

4.1.1 Linear models (SNL)

By modeling a linear relationship between novelty and resonance, we generally found positive trends, indicating that a change in novelty of emotions results in a corresponding change in resonance of the novel emotions. The positive relationship between the two variables has also been found in studies using topic modelling to calculate the information signals (Barron et al., 2018; Nielbo, Haestrup, et al., 2021).

Comparing the $N \times R$ slope for linear models fitted to the data in different time windows allows for assessing changes in the dynamics of emotions in the data. As Christmas and New Year give rise to new emotional responses with high spikes in both signals around the 24th and 31st of December (see figure 3.1), a linear slope was fitted on the data after removing all data points from December 1st to January 2nd in both 2019 and 2020. The value of the new slope was higher than the value of the slope for the entire dataset, which suggests that novel emotions are introduced over Christmas, however, they don't resonate to a corresponding extent. Thus, the change in the $N \times R$ slope might be caused by the extremely high novelty in emotions during December, and while this does lead to an increase in resonance (see figure 3.1a), it is a proportionally lower resonance score. This finding suggests the existence of a seasonal component in the dynamics of emotions on Twitter during Christmas and New Years. Taking steps to correct for this seasonal component would be optimal if one wants to conclude something about the general emotion dynamics on Twitter during other events.

To further investigate the dynamics of emotions during the COVID-19 pandemic, a $N \times R$ slope was fitted on the data before and after the first lockdown in Denmark (i.e. March 11th, 2020). The association between resonance and novelty was found to be similar for the two time periods. It would be interesting to calculate the $N \times R$ slopes on smaller windows based on specific events to minimize confounding factors from events unrelated to COVID-19 (such as Christmas). Previous studies have used change point detection (Nielbo, Haestrup, et al., 2021) or event detection (Saeed et al., 2019) to determine relevant windows.

4.1.2 Individual emotions (SMØ)

The use of emotion probabilities, rather than topic modelling, as the latent variable allows for a more direct interpretation of the labels driving the dynamic. In topic modelling, the latent variables are not directly interpretable (Bashar & Li, 2017). However, the same isn't the case for the emotion labels, thus, allowing for further discussion of the effect of the different labels on the dynamics of the signal.

In order to assess the degree to which individual emotions are driving the effect in this study, we excluded one emotion at a time and recalculated the information signals for the new emotion distributions. The correlation coefficients between the recalculated resonance signals and the resonance signal for the full distribution were all generally high. The lowest values were found for emotions '*Forventning/Interesse*' and '*Glæde/Sindsro*', which both have a correlation coefficient of $r = 0.85$. The remaining six emotions all had correlation coefficients $r > 0.9$. This could indicate that these two emotions created the largest change in the signal compared to the other emotions (compare with figures A.1 and A.2). This also suggests that the signal is generally generated by the interplay of the emotions, thus, one emotion alone can't explain much of the full signal.

The methods in this study display a simple way of examining the role of individual emotions on resonance. A more mathematical approach to assessing the importance of specific emotions could be the use of backpropagation and gradients to calculate what change in the emotion distribution would lead to the largest change in novelty. This could reveal which emotions are particular important for the dynamics.

4.2 Future Research (SNL)

The methods applied in this study have proven to capture interesting characteristics of Twitter. As this study was mainly concerned with actually applying the method and capturing the signal in the first place, there is number of natural next steps to take.

Some of the preliminary findings in our exploration of the overall signal suggest that zooming in on shorter windows could be interesting. This approach could also be combined with some way of quantifying significant changes within the system, for example event detection (Saeed et al., 2019) or change point detection (Nielbo, Haestrup, et al., 2021). Other studies have investigated sentiment on Twitter in relation to COVID-19 pandemic (Yin, Yang, & Li, 2020; J. Zhou, Yang, Xiao, & Chen, 2020) and expression of emotions on Twitter in relation to terroist attacts (Becker, Harb, & Ebeling, 2019), indicating a relevance to look at dynamics of emotions during extraordinary events.

In this study, the windowed relative entropy measures were calculated based on emotion probabilities from the Danish BERT Emotion model. Under the assumption that this approach captures general emotion dynamics, it would be interesting to compare our findings with other Danish emotion classifiers (e.g. Munch (n.d.)), expecting that findings will be similar. Furthermore, models for emotion classification exists for other languages (e.g Huang et al. (2021); Illendula and Sheth (2019); Mao, Chang, Shi, Li, and Shi (2019); Q. Zhou, Wu, and Zhang (2018); Savani (n.d.)), thus, it would be possible to apply the methods described in this study for other languages and compare dynamics between countries.

5 Conclusion

The aim of this study was to investigate whether using emotion probabilities as latent variables for windowed relative entropy could capture dynamics on Twitter. The results of the analysis reveals detectable signals of both novelty and resonance using emotion classification of the Danish BERT Emotion model on Danish tweets from 01/01 2019 to 04/30 2021. These findings lay the ground for using windowed relative entropy together with emotion classification to identify important dynamics on Twitter. For future studies, looking at the signal in relation to specific events (e.g. in the COVID-19 pandemic) could reveal important insights to the public response to those.

References

Barron, A. T. J., Huang, J., Spang, R. L., & DeDeo, S. (2018, May). Individuals, institutions, and innovation in the debates of the French Revolution. *Proceedings of the National Academy of Sciences*, *115*(18), 4607–4612. Retrieved 2021-12-17, from http://www.pnas.org/lookup/doi/10.1073/pnas.1717729115 doi: 10.1073/pnas.1717729115

Bashar, M. A., & Li, Y. (2017). Random Set to Interpret Topic Models in Terms of Ontology Concepts. In W. Peng, D. Alahakoon, & X. Li (Eds.), *AI 2017: Advances in Artificial Intelligence* (pp. 237–249). Cham: Springer International Publishing. doi: 10.1007/978-3-319-63004-5_19

Becker, K., Harb, J. G., & Ebeling, R. (2019, October). Exploring Deep Learning for the Analysis of Emotional Reactions to Terrorist Events on Twitter. *Journal of Information and Data Management*, *10*(2), 97–115. Retrieved 2021-12-20, from https://periodicos.ufmg.br/index.php/jidm/article/view/14711 (Number: 2)

Buccoliero, L., Bellio, E., Crestini, G., & Arkoudas, A. (2020, January). Twitter and politics: Evidence from the US presidential elections 2016. *Journal of Marketing Communications*, *26*(1), 88–114. Retrieved 2021-12-13, from https://doi.org/10.1080/13527266.2018.1504228 (Publisher: Routledge _eprint: https://doi.org/10.1080/13527266.2018.1504228) doi: 10.1080/13527266.2018.1504228

Cheong, J. H. (2019, May). *Four ways to quantify synchrony between time series data.* Retrieved 2021-12-10, from https://towardsdatascience.com/four-ways-to-quantify-synchrony-between-time-series-data-b99136c4a9c9

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019, May). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv:1810.04805 [cs]*. Retrieved 2021-12-08, from http://arxiv.org/abs/1810.04805 (arXiv: 1810.04805)

Fong, A., Roozenbeek, J., Goldwert, D., Rathje, S., & van der Linden, S. (2021, June). The language of conspiracy: A psychological analysis of speech used by conspiracy theorists and their followers on Twitter. *Group Processes & Intergroup Relations*, *24*(4), 606–623. Retrieved 2021-12-13, from https://doi.org/10.1177/1368430220987596 (Publisher: SAGE Publications Ltd) doi: 10.1177/1368430220987596

Gao, J., Hu, J., & Tung, W.-w. (2011, September). Facilitating Joint Chaos and Fractal Analysis of Biosignals through Nonlinear Adaptive Filtering. *PLOS ONE*, *6*(9), e24331. Retrieved 2022-01-02, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0024331 (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0024331

Huang, C., Trabelsi, A., Qin, X., Farruque, N., Mou, L., & Zaïane, O. (2021, June). Seq2Emo: A Sequence to Multi-Label Emotion Classification Model. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 4717–4724). Online: Association for Computational Linguistics. Retrieved 2021-12-18, from https://aclanthology.org/2021.naacl-main.375 doi: 10.18653/v1/2021.naacl-main.375

Illendula, A., & Sheth, A. (2019, May). Multimodal Emotion Classification. In *Companion Proceedings of The 2019 World Wide Web Conference* (pp. 439–449). San Francisco USA: ACM. Retrieved 2021-12-18, from https://dl.acm.org/doi/10.1145/3308560.3316549 doi: 10.1145/3308560.3316549

Itti, L., & Baldi, P. (2009, June). Bayesian surprise attracts human attention. *Vision Research*, *49*(10), 1295–1306. Retrieved 2021-12-20, from https://www.sciencedirect.com/science/article/pii/S0042698908004380 doi: 10.1016/j.visres.2008.09.007

Karami, A., Lundy, M., Webb, F., & Dwivedi, Y. K. (2020). Twitter and Research: A Systematic Literature Review Through Text Mining. *IEEE Access*, *8*, 67698–67717. (Conference Name: IEEE Access) doi: 10.1109/ACCESS.2020.2983656

Lwin, M. O., Lu, J., Sheldenkar, A., Schulz, P. J., Shin, W., Gupta, R., & Yang, Y. (2020, May). Global Sentiments Surrounding the COVID-19 Pandemic on Twitter: Analysis of Twitter Trends. *JMIR Public Health and Surveillance*, *6*(2), e19447. Retrieved 2021-12-13, from https://publichealth.jmir.org/2020/2/e19447 (Company: JMIR Public Health and Surveillance Distributor: JMIR Public Health and Surveillance Institution: JMIR Public Health and Surveillance Label: JMIR Public Health and Surveillance Publisher: JMIR Publications Inc., Toronto, Canada) doi: 10.2196/19447

Mao, X., Chang, S., Shi, J., Li, F., & Shi, R. (2019, January). Sentiment-Aware Word Embedding for Emotion Classification. *Applied Sciences*, *9*(7), 1334. Retrieved 2021-12-18, from https://www.mdpi.com/2076-3417/9/7/1334 (Number: 7 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/app9071334

Munch, N. (n.d.). *NikolajMunch/danish-emotion-classification · Hugging Face.* Retrieved 2021-12-18, from https://huggingface.co/NikolajMunch/danish-emotion-classification

Murdock, J., Allen, C., & DeDeo, S. (2017, February). Exploration and exploitation of Victorian science in Darwin's reading notebooks. *Cognition*, *159*, 117–126. Retrieved 2021-12-17, from https://linkinghub.elsevier.com/retrieve/pii/S0010027716302840 doi: 10.1016/j.cognition.2016.11.012

Nielbo, K. L., Enevoldsen, K., Baglini, R., Fano, E., & Gao, J. (2021). Pandemic news information uncertainty - News dynamics mirror differential response strategies to COVID-19. , 16.

Nielbo, K. L., Haestrup, F., Enevoldsen, K. C., Vahlstrup, P. B., Baglini, R. B., & Roepstorff, A. (2021, February). When no news is bad news – Detection of negative events from news media content. *arXiv:2102.06505 [cs]*. Retrieved 2021-12-09, from http://arxiv.org/abs/2102.06505 (arXiv: 2102.06505)

Nordic BERT. (2021, December). Certainly. Retrieved 2021-12-15, from https://github.com/certainlyio/nordic_bert (original-date: 2019-12-03T09:48:29Z)

Pauli, A. B., Barrett, M., Lacroix, O., & Hvingelby, R. (2021, May). DaNLP: An open-source toolkit for Danish Natural Language Processing. In *Proceedings of the 23rd Nordic Conference on Computational Linguistics (NoDaLiDa)* (pp. 460–466). Reykjavik, Iceland (Online): Linköping University Electronic Press, Sweden. Retrieved 2021-12-15, from https://aclanthology.org/2021.nodalida-main.53

Reis, T. R. d. (2020, February). *Measuring the statistical similarity between two samples using Jensen-Shannon and Kullback-Leibler....* Retrieved 2021-12-09, from https://medium.com/datalab-log/measuring-the-statistical-similarity-between-two-samples-using-jensen-shannon-and-kullback-leibler-8d05af514b15

Rufai, S. R., & Bunce, C. (2020, August). World leaders' usage of Twitter in response to the COVID-19 pandemic: a content analysis. *Journal of Public Health*, *42*(3), 510–516. Retrieved 2021-12-13, from https://doi.org/10.1093/pubmed/fdaa049 doi: 10.1093/pubmed/fdaa049

Saeed, Z., Abbasi, R. A., Maqbool, O., Sadaf, A., Razzak, I., Daud, A., ... Xu, G. (2019, June). What's Happening Around the World? A Survey and Framework on Event Detection Techniques on Twitter. *Journal of Grid Computing*, *17*(2), 279–312. Retrieved 2021-12-18, from http://link.springer.com/10.1007/s10723-019-09482-2 doi: 10.1007/s10723-019-09482-2

Savani, B. (n.d.). *bhadresh-savani/bert-base-uncased-emotion · Hugging Face.* Retrieved 2021-12-18, from https://huggingface.co/bhadresh-savani/bert-base-uncased-emotion

Sentiment Analysis — DaNLP documentation. (2020). Retrieved 2021-12-20, from https://danlp-alexandra.readthedocs.io/en/latest/docs/tasks/sentiment_analysis.html#bert-emotion

Shen, D., Urquhart, A., & Wang, P. (2019, January). Does twitter predict Bitcoin? *Economics Letters*, *174*, 118–122. Retrieved 2021-

12-13, from https://www.sciencedirect.com/science/article/pii/S0165176518304634 doi: 10.1016/j.econlet.2018.11.007

Thelwall, M. (2014). Sentiment analysis and time series with Twitter. In *Twitter and Society* (Katrin Weller, Axel Bruns, Jean Burgess, Merja Mahrt and Cornelius Puschmann ed., Vol. 89). New York: Peter Lang. Retrieved 2021-12-13, from https://core.ac.uk/reader/19541510#page=83

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406–418. Retrieved 2021-12-17, from https://onlinelibrary.wiley.com/doi/abs/10.1002/asi.21462 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/asi.21462) doi: 10.1002/asi.21462

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., . . . Polosukhin, I. (2017). Attention is All you Need. In I. Guyon et al. (Eds.), *Advances in Neural Information Processing Systems* (Vol. 30). Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf

Visentin, M., Tuan, A., & Di Domenico, G. (2021). Words matter: How privacy concerns and conspiracy theories spread on twitter. *Psychology & Marketing*, *38*(10), 1828–1846. Retrieved 2021-12-13, from https://onlinelibrary.wiley.com/doi/abs/10.1002/mar.21542 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mar.21542) doi: 10.1002/mar.21542

Vliet, L. v., Törnberg, P., & Uitermark, J. (2020, September). The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries. *PLOS ONE*, *15*(9), e0237073. Retrieved 2021-12-13, from https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237073 (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0237073

Welbers, K., & Opgenhaffen, M. (2019, January). Presenting News on Social Media. *Digital Journalism*, *7*(1), 45–62. Retrieved 2021-12-20, from https://doi.org/10.1080/21670811.2018.1493939 (Publisher: Routledge _eprint: https://doi.org/10.1080/21670811.2018.1493939) doi: 10.1080/21670811.2018.1493939

Yin, H., Yang, S., & Li, J. (2020). Detecting Topic and Sentiment Dynamics Due to COVID-19 Pandemic Using Social Media. In X. Yang, C.-D. Wang, M. S. Islam, & Z. Zhang (Eds.), *Advanced Data Mining and Applications* (pp. 610–623). Cham: Springer International Publishing. doi: 10.1007/978-3-030-65390-3_46

Zhou, J., Yang, S., Xiao, C., & Chen, F. (2020, June). Examination of Community Sentiment Dynamics due to COVID-19 Pandemic: A Case Study from A State in Australia.

Zhou, Q., Wu, H., & Zhang, Z. (2018, October). NLP at IEST 2018: BiLSTM-Attention and LSTM-Attention via Soft Voting in Emotion Classification. In *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 189–194). Brussels, Belgium: Association for Computational Linguistics. Retrieved 2021-12-18, from https://aclanthology.org/W18-6226 doi: 10.18653/v1/W18-6226

A Appendix

A.1 Scraping keywords

Below follows the list of top words from Danish, Swedish and Norwegian used to scrape Twitter:

> aften, aldrig, alltid, altid, andet, arbejde, bedste, behöver, behøver, beklager, berätta, betyr, blev, blevet, blir, blitt, blive, bliver, bruge, burde, bättre, båe, bør, deim, deires, ditt, drar, drepe, dykk, dykkar, där, död, döda, død, døde, efter, elsker, endnu, faen, fandt, feil, fikk, finner, flere, forstår, fortelle, fortfarande, fortsatt, fortælle, från, få, fået, får, fått, förlåt, första, försöker, før, først, første, gick, gikk, gillar, gjennom, gjerne, gjorde, gjort, gjør, gjøre, godt, gå, gång, går, göra, gør, gøre, hadde, hallå, havde, hedder, helt, helvete, hende, hendes, hennes, herregud, hjelp, hjelpe, hjem, hjälp, hjå, hjælp, hjælpe, honom, hossen, hvem, hvis, hvordan, hvorfor, händer, här, håll, håller, hør, høre, hører, igjen, ikkje, ingenting, inkje, inte, intet, jeres, jävla, kanske, kanskje, kender, kjenner, korleis, kvarhelst, kveld, kven, kvifor, känner, ledsen, lenger, lidt, livet, längre, låt, låter, længe, meget, menar, mycket, mykje, må, måde, många, mår, måske, måste, måtte, navn, nogen, noget, nogle, noko, nokon, nokor, nokre, någon, något, några, nån, når, nåt, nødt, också, også, pengar, penger, pratar, prøver, på, redan, rundt, rätt, sagde, saker, samma, sammen, selv, selvfølgelig, sidan, sidste, siger, sikker, sikkert, själv, skete, skjedde, skjer, skulle, sluta, slutt, snakke, snakker, snill, snälla, somt, stadig, stanna, sted, står, synes, säger, sätt, så, sådan, såg, sånn, tager, tiden, tilbage, tilbake, tillbaka, titta, trenger, trodde, troede, tror, två, tycker, tänker, uden, undskyld, unnskyld, ursäkta, uten, varför, varit, varte, veldig, venner, verkligen, vidste, vilken, virkelig, visste, väg, väl, väldigt, vän, vår, våra, våre, væk, vær, være, været, älskar, åh, år, åt, över.

A.2 Resonance for distributions with excluded emotions

Figure A.1 shows resonance for emotion distributions with one emotion excluded.

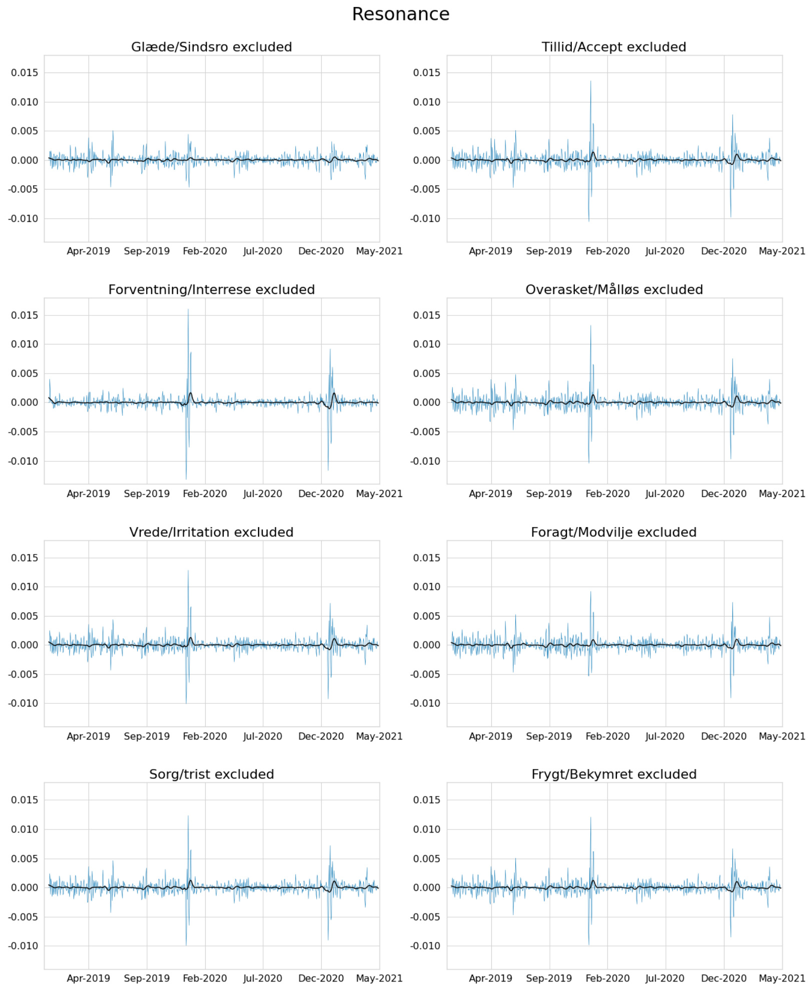


Figure A.1. Resonance calculated with a window $w = 3$ for the re-scaled distributions with one emotion excluded. The blue line represents the raw signal, and the black line is the signal smoothed using a non-linear adaptive filter (Gao et al., 2011)

A.3 Raw distributions of emotions

Figure A.2 shows the mean probability of each emotion in emotion probability distributions summarized over one day

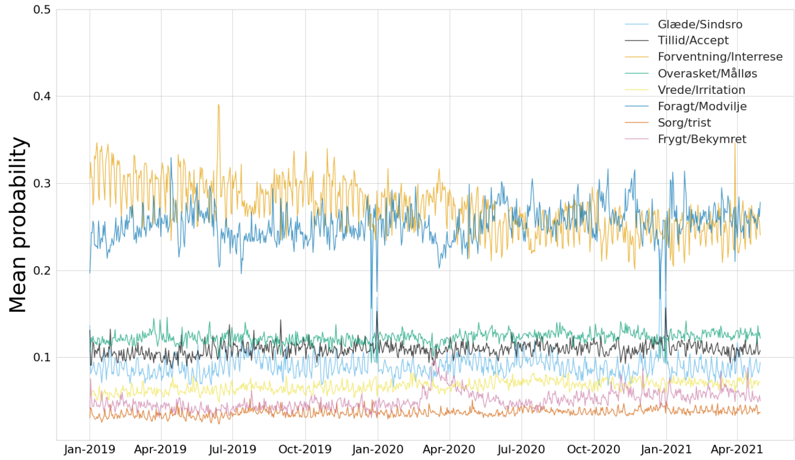


Figure A.2. Mean probability of each emotion in the daily summarized distributions