

目次

1	概要	3
1.1	はじめに	3
1.2	構成	3
1.3	環境	3
1.4	投稿数収集方法	4
2	日ごとの投稿数データの特徴	5
2.1	概要	5
2.2	一発型モデル	5
2.3	上昇モデル	6
2.4	下降モデル	7
3	ロコミ効果によるヒット現象の数理モデル	8
3.1	概要	8
3.2	数理モデル	8
3.3	まとめ	9
4	ブーム形成・終焉における数理モデル	9
4.1	概要	9
4.2	数理モデル	10
4.3	平均化方法	11
4.4	フィッティング方法	12
4.5	検証	13
4.6	まとめ	17
5	応用事例 1：ブームのピーク値予測	17
5.1	概要	17
5.2	予測	18
5.3	まとめ	19
6	応用事例 2：SNS 上のブームの形成・終焉予測	20
6.1	概要	20
6.2	予測手順	21
6.3	ブームの形成・終焉の予測	22
6.4	まとめ	24

1 概要

1.1 はじめに

近年、情報技術の発展により、誰もが情報端末を身に着け、いつでも情報を受信または、発信できるようになり、情報を簡単に扱うことができるようになった。それに伴い、世界中の情報を、情報端末一つで簡単に入手し拡散することが可能になった。また、情報の入手と拡散には、ソーシャル・ネットワーキング・サービス(以下、本論文では「SNS」と記載する。)が、用いられるようになった。SNSは、インターネット上で他人とコミュニケーションを円滑に行うことができるサービスであり、「Twitter(<https://twitter.com>)」,「Facebook(<https://ja-jp.facebook.com/>)」,「Instagram(<https://www.instagram.com/>)」などが挙げられる。

SNSは、コミュニケーションを円滑に行うだけでなく、世界中の企業が、広告を表示させたり、自社の宣伝を行うなど、ブームの形成にも用いられるようになり、その形成されたブームによって、商品が売れたり、知名度が上がるなど、SNSのブームが現実社会に影響を及ぼすようになった。

そこで、情報のブームはSNS上で、どのように形成され、どのように終焉を迎えていくのか微分方程式としてモデル化する。

本研究に用いるSNSは、「Twitter」のみとする。日本で知名度の高いSNSのうち、社会のどの話題に対しても、ある程度の投稿数を即時に得ることができるからである。本論文で記載する政治分野の話題の投稿数は、「Facebook」や、「Instagram」よりも「Twitter」の投稿数の方が規模が大きかった。また、「Instagram」では、政治分野に関する投稿はほとんど得ることができなかった。そのため、どの話題に対しても、ある程度の投稿数を得ることができる「Twitter」を用いることとする。また、投稿数データは、1日刻みで取得するものとする。

1.2 構成

本論文は、2章まででSNSのブームの特徴について説明した後、3章から4章にわたり、参考文献をもとにSNSのブームの形成・終焉についての数理モデルについて説明する。6章より、応用事例として定義した数理モデルを用いてSNS社会のブームの形成・終焉を予測を行う。

1.3 環境

- Twitter(<https://twitter.com>)
- Yahooリアルタイム検索サービス (<https://search.yahoo.co.jp/realtime>)

1.4 投稿数収集方法

Twitter 上で、話題に対する投稿数データは、「Yahoo リアルタイム検索サービス」を用いて取得する。「Yahoo リアルタイム検索サービス」は、「Twitter」、「Facebook」そして、「Instagram」の投稿数を 1 日ごと、7 日ごと、1 か月ごとで検索することができる。1 日ごとで検索する場合は、15 分刻みで投稿数を検索することができ、7 日ごとでは 2 時間刻み、1 か月ごとでは 1 日刻みで取得することができる。また、Yahoo リアルタイム検索サービスを用いてデータを取得する際に注意することがある。例えば、1 か月ごとで投稿数を検索した時、得られる投稿数は 1 日ごとであるが、その 1 日の投稿数の集計時間が、24 時である。つまり、1 月 1 日と表示されている時の集計時間は、12 月 31 日 0 時 01 分～1 月 01 日 0 時 00 分である。そのため、1 月 01 日中の投稿数データを取得したい時は、1 月 02 日の投稿数データを取得する必要がある。

2 日ごとの投稿数データの特徴

2.1 概要

Twitter の 1 日ごとの投稿数データを，Yahoo リアルタイム検索サービスを用いて取得した．取得した投稿数データを，グラフで表したときに，グラフの形に，ある特徴があることが分かった．1 日毎の投稿数の推移の形が，いくつかに分けることができるということである．本研究では，大きく 3 つに分ける．

1 つ目は，ある 1 日に投稿数が増えていき，すぐに減っていく形である (図 1)．これは，瞬間的にブームが形成され，ピークを迎えた後，すぐに終焉を迎える一発型モデルである．2 つ目は，徐々に投稿数が増加していく上昇モデルである (図 2)．数日前より，ブームが形成され始め，ある日にブームが MAX を迎えるものである．3 つ目は，ピークを過ぎた後に，徐々に投稿数が減少していく下降モデルである (図 3)．ブームが形成され，ピークを迎えた後に，ゆっくりと投稿数が減少する．

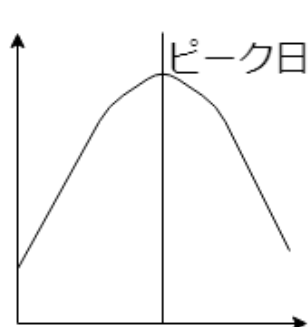


図 1 一発型モデルの模式図.

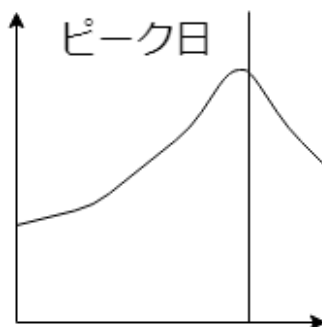


図 2 上昇モデルの模式図.

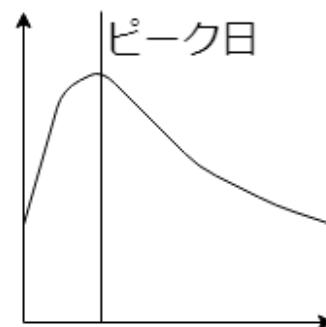


図 3 下降モデルの模式図.

2.2 一発型モデル

一発型モデルとは，瞬間的にブームが形成された後，すぐにブームが去るモデルである，このモデルは，一時的に有名になったニュースや，スポーツ選手に関する投稿数で多く見られる．

図 4 は，2017 年 9 月上旬に発生した，太陽フレアについての，Twitter 上での 1 日ごとの投稿数をグラフにしたものである．約 2 日で，ブームを形成し，25 万規模の投稿数を得ながらも，すぐにブームが去り，投稿数が減少している一発型モデルの代表である．

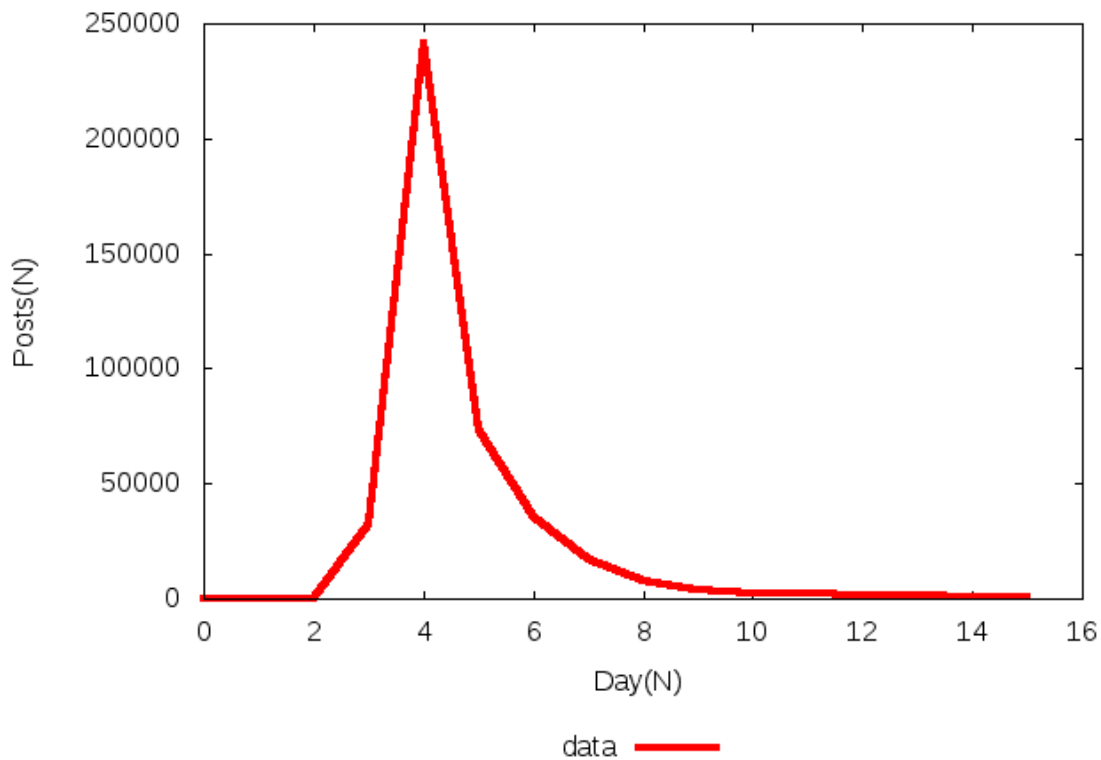


図4 2017年9月上旬に発生した、太陽フレアについての日ごとの投稿数データである。初日は9月4日。

2.3 上昇モデル

上昇モデルとは、ある日に向かって、ブームをなだらかに形成し、特定の日付近で最も盛り上がるモデルである。このモデルは、2017年10月の衆議院選挙時の、自由民主党に関する投稿数で見られた。

図5は、2017年10月の衆議院選挙時の、自由民主党についてのTwitter上での一日ごとの投稿数をグラフにしたものである。徐々にブームを形成していることがわかる。

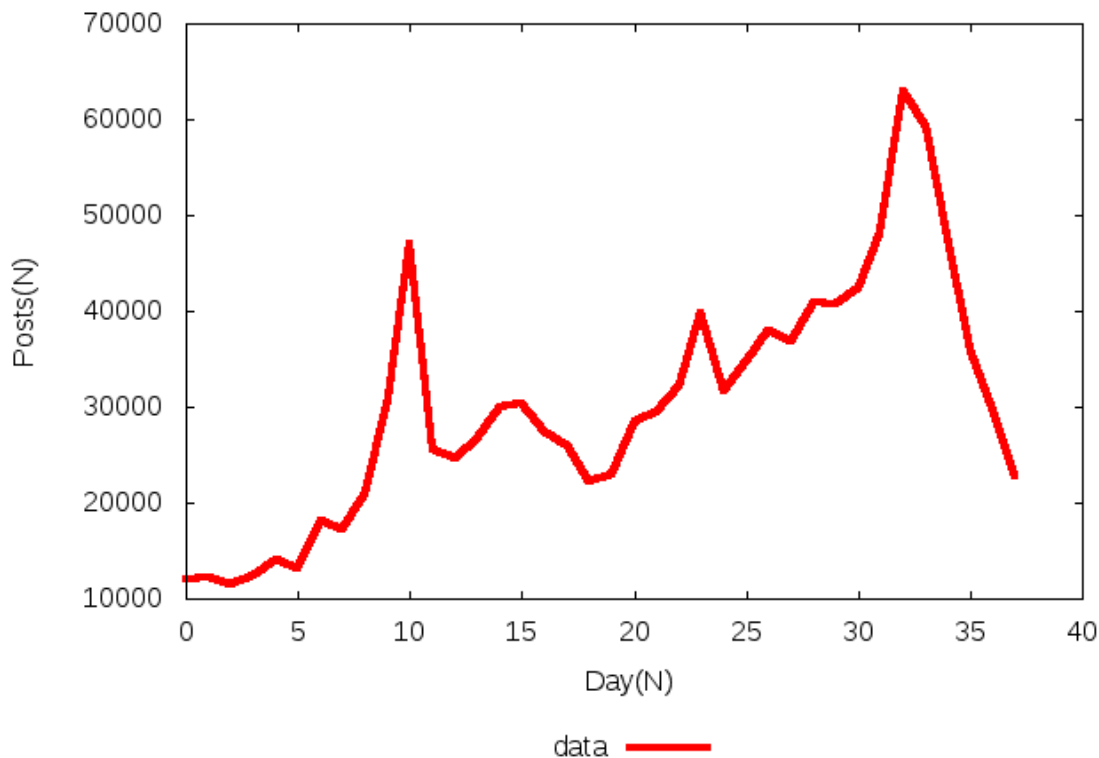


図5 2017年10月の衆議院選挙時の、自由民主党についての日ごとの投稿数データである。検索ワードは、自由民主党のみである。初日は9月19日。

2.4 下降モデル

下降モデルとは、ピークを過ぎた後に、徐々に投稿数が減っていくモデルである。これはブームが形成され始め、ピークを過ぎた後、ゆっくりと興味が薄れていくものである。このモデルは、2017年10月の衆議院選挙時の希望の党に関する投稿数で見られた。希望の党は、2017年に新たに設立された党であり、設立当初でかなりの投稿数を得た後、投票日に向かってゆっくりと投稿数が減少していった。

図6は、2017年の10月の衆議院選挙時の、希望の党についてのTwitter上での1日ごとの投稿数をグラフにしたものである。設立当初(図6のグラフで盛り上がっている付近)は、7万規模もの投稿数を得ていたが、投票日には、2万~3万規模まで減少していた。

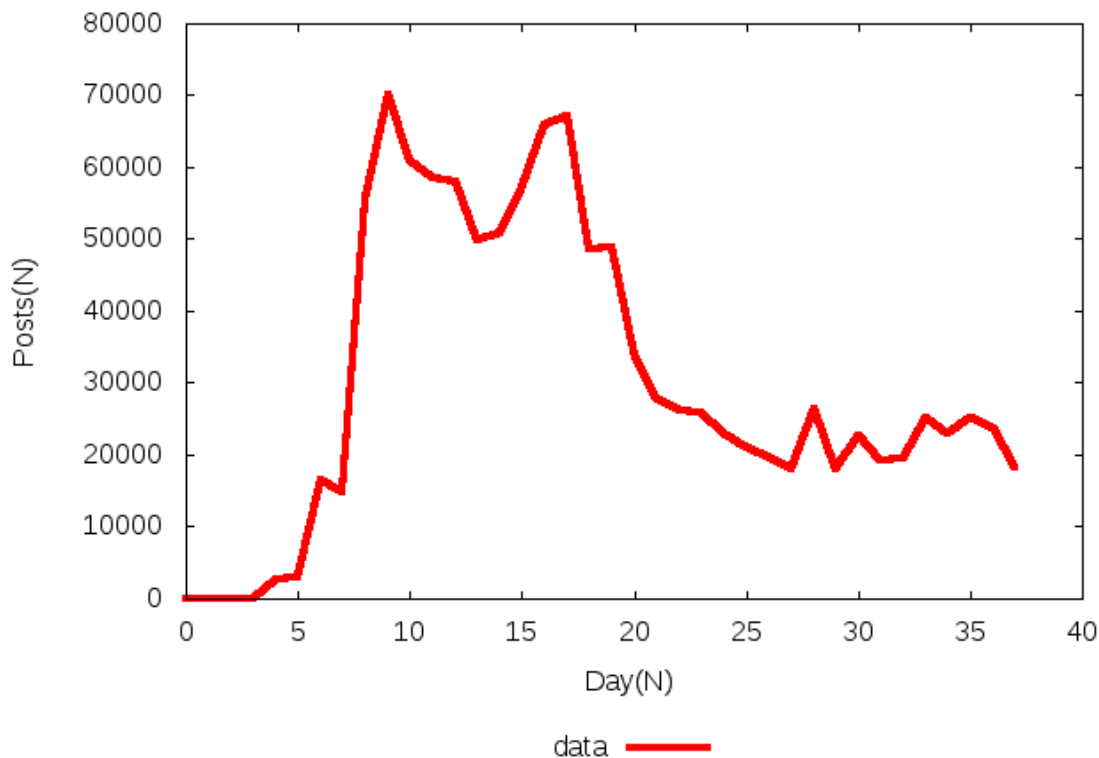


図6 2017年10月の衆議院選挙時の、希望の党についての日ごとの投稿数データである。検索ワードは、希望の党のみである。初日は9月19日。

3 口コミ効果によるヒット現象の数理モデル

3.1 概要

Ishii らは、口コミ効果による、映画のようなエンタテインメント事業の経済効果を求める数理モデルを定義した [1]。私は、この論文をもとに、次章から SNS の投稿数のブームの形成と終焉をモデル化する。そのために、この章で、Ishii らが定義した数理モデルを説明する。

3.2 数理モデル

映画の購入意欲の減少は、指数関数的な減少で表すことが可能である。また、第2章の分類でいうと、購入意欲は上昇モデルに近い。これをもとに口コミ効果によるヒット現象の数理モデルを定義するにあたり、映画を見ようとする、購入意欲に影響する要因がある。それは、「直接コミュニケーション」と呼ばれる、直接に友人から勧められることと、「間接コミュニケーション」と呼ばれる、ネット検索や、ブログのやりとり、街中でのうわさが存在する。また、映画のような事業では、視聴した人と、視聴していない人が存在する。

上記のことを考慮すると、直接コミュニケーションでは、まだ視聴していない人同士の影響と、視聴した人から、視聴したことのない人への影響の二つが考えられる。また、間接コミュニケーションでは、視聴していない人同士からの影響、視聴した人同士からの影響、そして、視聴した人と視聴していない人の会話からの影響の三つが考えられる。この事と、購入意欲は、指数関数的に減少していき、そして、映画には広告出稿量による影響が考えられる。これらを考慮し、 a は正の定数、時間 t あたりの購入意欲を $I(t)$ 、 N_p を、映画タイトルの潜在的な観客の総数、 $N(t)$ を、時刻 t の時点で、既にその映画を視聴した観客の数 D^{nn} を、視聴していない人同士の直接コミュニケーションの係数、 D^{ny} を、視聴した人から、視聴していない人への直接コミュニケーションの係数、 P^{nn} を、視聴していない人同士からの間接コミュニケーションの係数、 P^{yy} を、視聴した人同士からの間接コミュニケーションの係数、 P^{ny} を、視聴した人と、視聴していない人からの間接コミュニケーションの係数、 $f(t)$ を時間 t の広告出稿量とおくと、

$$\begin{aligned} \frac{dI(t)}{dt} = & -aI(t) + f(t) + \frac{(N_p - N(t))^2}{N_p} D^{nn} I + \frac{N(t)}{N_p} (N_p - N(t)) D^{ny} I \\ & + \frac{(N_p - N(t))^3}{N_p} P^{nn} I^2 + \frac{(N(t))^2 (N_p - N(t))}{N_p} P^{yy} I^2 \\ & + \frac{N(t) (N_p - N(t))^2}{N_p} P^{ny} I^2 \end{aligned} \quad (1)$$

と定義することができる。また、 I は、

$$I = \frac{1}{N_p} \sum_i I_i(t) \quad (2)$$

と表すことができ、 N_p 人の消費者が皆同じ行動をするということを意味している。

3.3 まとめ

式 (1) は、広告出稿量を入力することで、興業収入を再現することができる。また、広告出稿量の時間配分を変更することで、興行収入の予測を変更することができる。

また、式 (1) を見ると、購入意欲は、指数関数で影響する項、2 乗で影響する項、そして、広告からの影響つまり、外力で影響する項がある。この 3 項をもとに、本研究である。Twitter 上のブームの形成・終焉にも適応していく。

4 ブーム形成・終焉における数理モデル

4.1 概要

この章では、3 章で記載した式 (1) のアイデアをもとに、SNS に適応することができるブーム形成・終焉における数理モデルを説明する。そして、数理モデルが実際の SNS の

ブームの形成・終焉を表すことができるか検証する。

4.2 数理モデル

3章の Ishii らの口コミ効果についての論文 [1] より、2乗で影響を与える項と、ニュース報道など、一時的な外因を表す $f(t)$ (本論文において「外力」と記載する。) があることがわかる。外力 $f(t)$ は、時刻 t で影響を与える。そこで、 p を、コミュニケーションの係数と置き、投稿数 y を時間 t の関数とすると、

$$\frac{dy}{dt} = a(t)y + py^2 + f(t) \quad (3)$$

と、表すことができる。この時 $f(t) = 0$, a を定数とすると、

$$\frac{dy}{dt} = ay + py^2 \quad (4)$$

と表すことができ、これはロジスティック曲線と呼ばれる曲線を解にもつ、生物の種が一定になる様子などがフィッティングできる式と知られている。式 (4) を変数分離法を用いて解くと、

$$y(t) = \frac{a_0 C e^{at}}{1 - p C e^{a_0 t}} \quad (5)$$

となる。 C は積分定数であり、本研究では、あるキーワードに対する計測初日の投稿数を代入する。

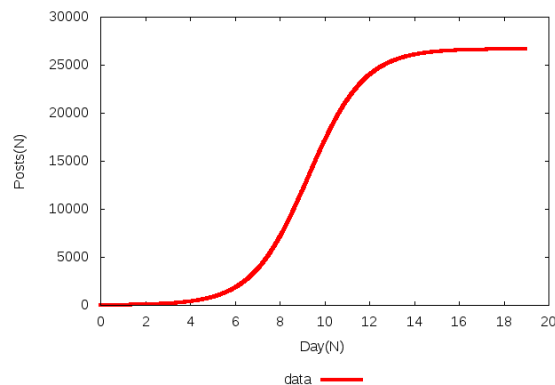


図7 式 (5) の解の一例。 $a_0 = 0.8$, $p = -0.00003$, $c = 20$ としたときの $y(t)$ の時間変化を表したグラフ。

また、式 (3) の a の符号は、ブーム形成時・終焉時で正から負へ変化するため、

$$a(t) = -a_0 \tanh K(t - t_0) \quad (6)$$

と、仮定する． a_0 と K は正の定数であり， t_0 はブームが落ち着く時間である．

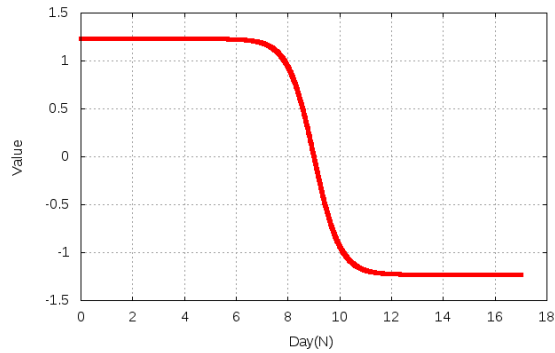


図8 式(6)の解の一例． $a_0 = 1.23$, $K = 0.01$, $y_0 = 9$ としたときの $a(t)$ の時間変化を表したグラフ．

4.3 平均化方法

実際の SNS の投稿数には，なるべく同じようなふるまいをするように 3 点平均を用いる．本研究で用いる 3 点平均は，元のデータを $y_1, y_2, y_3 \cdots y_f$ として， n 日目のデータ \tilde{y}_n を 3 点平均して，

$$\tilde{y}_n = \frac{y_{n-1} + y_n + y_{n+1}}{3} \quad \text{ただし, } \tilde{y}_1 = y_1, \tilde{y}_f = y_f \quad (7)$$

また，平均化を行う手法として 3 点平均のほかに，5 点平均，7 点平均といった平均化する手法があるが，本研究では 3 点平均のみを用いる．以下に実際の投稿数データに，3 点平均，5 点平均，7 点平均を行った図を示す．

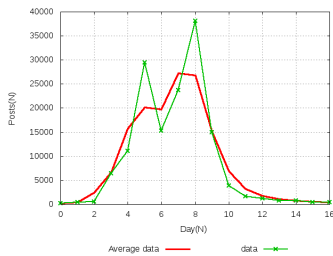


図9 1日毎の投稿数データと 3 点平均を行ったデータの比較．

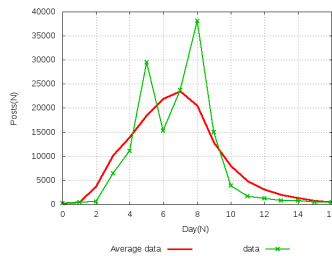


図10 1日毎の投稿数データと 5 点平均を行ったデータの比較．

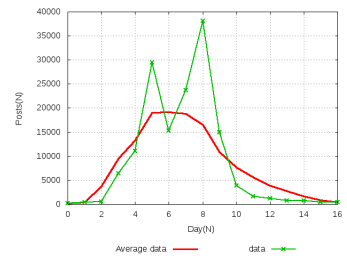


図11 1日毎の投稿数データと 7 点平均を行ったデータの比較．

図9から、図11を見ると、3点平均を行った図では、急激な実際の投稿数の増減をかすかに残している。しかし、5点平均、7点平均では、残らずなだらかな図となっていることがわかる。このため、ある程度の増減を残しつつ、投稿数の変化をなだらかにすることができるため、3点平均を用いる。

4.4 フィットting方法

本研究で用いたフィッティング方法を記述する。フィッティングとは、ある SNS 上のキーワードについての日毎の投稿数をもとに、定義した数理モデル式 (3), 式 (6) の a_0 と、 p の値を求めることである。

フィッティングを行うにあたっての注意点は、次節から、式 (3), 式 (6) の数理モデルを検証していくが、その際用いられるデータは、ブームの形成から終焉までのトータルデータである (図 12)。しかし、フィッティング作業では、ブームの形成時のデータのみを用いる (図 13)。ここで言うブームの形成時とは、ブームが形成し始めた時から、ブームが落ち着いた時までである。

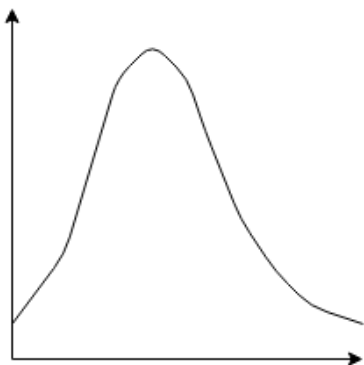


図 12 ブーム形成から終焉までのトータルデータの模式図。

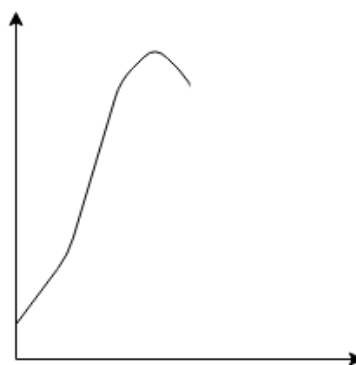


図 13 ブームの形成時の模式図。

次に、フィッティングの方法を記載する。式 (5) を用いて、数理モデル式 (3), 式 (6) に代入する最適な a_0 と p の組み合わせを探索する。まず、SNS 上のあるキーワードについて 1 日毎の投稿数を取得する。そして、取得した投稿数に 3 点平均を行う。そして式 (5) を複数回計算を行い、投稿数を平均化したデータと比較し誤差が最も小さい時の a_0 と p をフィッティング結果とする。その際に、 a_0 は正の定数をとるため、0.0 から任意の値までを繰り返す。 p は、最終日のデータを y_f とおくと、 $\frac{-a_0}{\frac{1}{2}y_f}$ から $\frac{-a_0}{2y_f}$ まで繰り返して計算を行う。本研究では小数点第 5 桁まで探索を行った。また、誤差は 1 日毎の平均データとフィッティングデータの差を 2 乗したものの総和をそれぞれ比較し、誤差が最小のものをフィッティング結果とする。

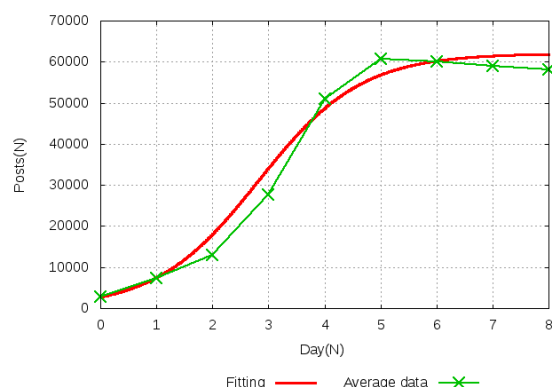


図 14 フィッティングの一例. フィッティング結果と, 実際の投稿数を平均化したものの比較. $a_0 = 1.090740$, $p = -0.00001$, 誤差 = 10502.060369. 投稿数データの検索ワードは希望の党であり, 初日は 9 月 23 日である.

4.5 検証

4.2 節で定義した, 式 (3) を用いて, 実際の SNS の投稿数と一致するかを調べる. 本研究は, Twitter 上の様々な分野の話題に対応すると仮定する. そのため, 扱うデータの分野はばらばらとなるものとする. また, この章では 1 週間から 2 週間程度のデータを扱う.

まず最初に, $f(t) = 0$ つまり, 外力を考えないものとして検証を行う. 計算に用いるパラメータは, フィッティングで求めた値であり, 表 1 に示す. また, 外力を考慮することでどの程度精度が増すかも検証するため, 誤差も表記する.

誤差の求め方は, ある検索ワードの 1 日毎の投稿数を取得し, 取得したデータに対して 3 点平均を行う. そして, 総投稿数と, 1 日毎の平均データと計算データの差を比で表す.

平均化を行ったデータを $y_1, y_2, y_3 \cdots y_f$ として, n 日目のデータ \tilde{y}_n とし, 計算データを $t_1, t_2, t_3 \cdots t_f$ として, n 日目のデータ \tilde{t}_n とし, S_1, S_2 を用いて表すと,

$$S_1 = y_1 + y_2 + y_3 + \cdots + y_f \quad (8)$$

$$S_2 = \sqrt{(y_1 - t_1)^2 + (y_2 - t_2)^2 + (y_3 - t_3)^2 + \cdots + (y_f - t_f)^2} \quad (9)$$

$$\text{誤差} = \frac{S_2}{S_1} \times 100 \quad (10)$$

となり, 誤差を % として表す.

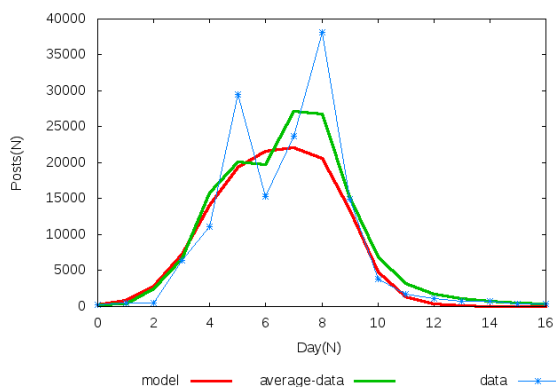


図 15 2017 年 9 月に発生した台風 18 号の時の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。検索ワードは、台風 18 号のみである。初日は 2017 年 09 月 09 日。

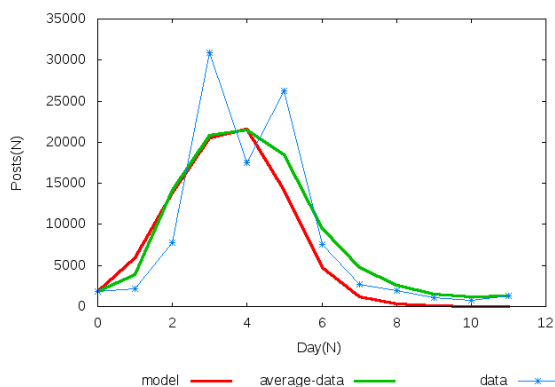


図 16 2017 年 6 月に開催された AKB48 総選挙時の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。検索ワードは、AKB48 総選挙のみである。初日は 2017 年 06 月 14 日。

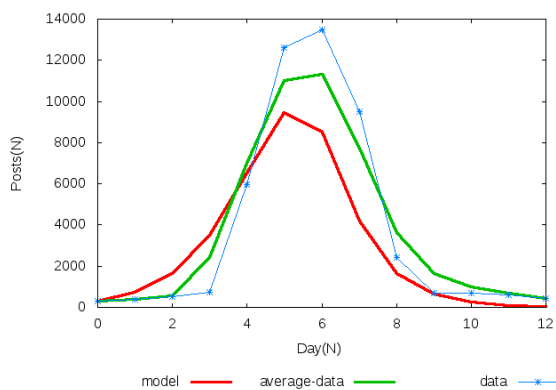


図 17 2017 年 11 月に開催されたポケモン GO イベント時の鳥取砂丘の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。検索ワードは、鳥取砂丘のみである。初日は 2017 年 11 月 20 日であり、イベントは 24 日から 26 日の 3 日間。

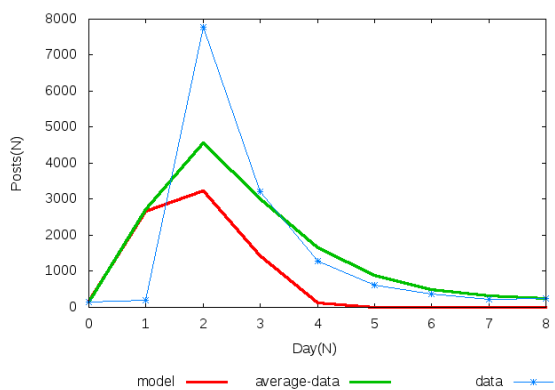


図 18 2017 年 10 月のノーベル賞で話題になった重力波の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。検索ワードは、重力波のみである。初日は 2017 年 10 月 01 日。

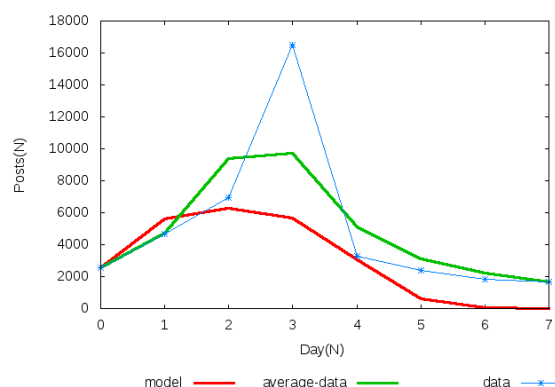


図 19 2017 年 9 月に発生したハリケーンの Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。検索ワードは、ハリケーンのみである。初日は 2017 年 9 月 14 日。

表 1 $f(t) = 0$ の時のパラメータ

検索ワード	a_0	p	t_0	K	誤差 (%)
台風 18 号	1.231790	-0.000054	9	0.01	17.84
AKB48 総選挙	1.365790	-0.000054	5	0.01	21.28
鳥取砂丘	0.883180	-0.000051	6	0.01	32.55
重力波	4.072930	-0.001048	3	0.01	45.90
ハリケーン	2.282700	-0.000350	4	0.01	42.70

図 15 から図 19 より、式 (3) で数理モデル化した結果と 20% から 40% ほどで一致させることができた。しかし、どの図も頂点付近で大きく外れていることがわかる。そのため、次に外力 $f(t)$ を考慮して検証する。本研究では、外力は正の整数値を任意に決め代入し、誤差がどれくらい小さくなるか検証した。計算に用いるパラメータは表 1 と同じものとする。また、ハリケーンや重力波については、徐々にブームが形成されたのではなく、急激にブームが形成されたため、誤差が大きくなっている。

次に、計算値と、平均データを比較したとき、大きく外れてしまった部分に $f(t)$ を考慮して、再計算を行う。

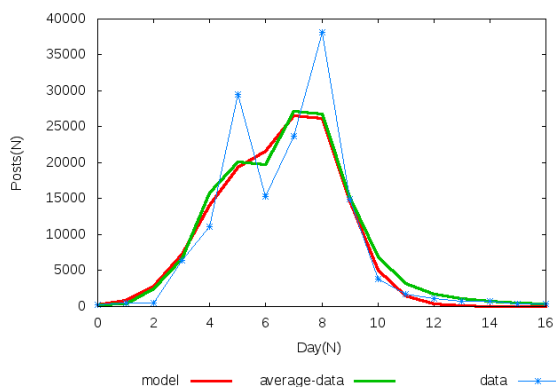


図 20 2017 年 9 月に発生した台風 18 号の時の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。7 日目に 8000、8 日目に 9000 の外力を挿入している。

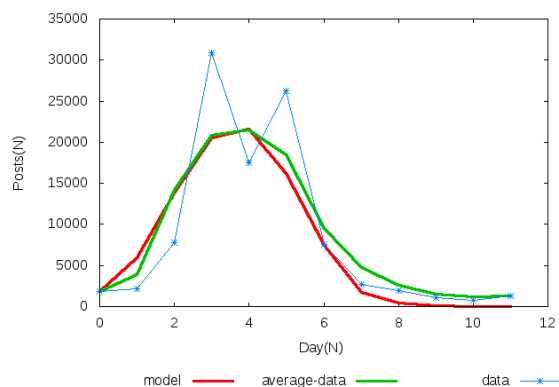


図 21 2017 年 6 月に開催された AKB48 総選挙時の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。5 日目に 4000、6 日目に 5000 の外力を挿入している。

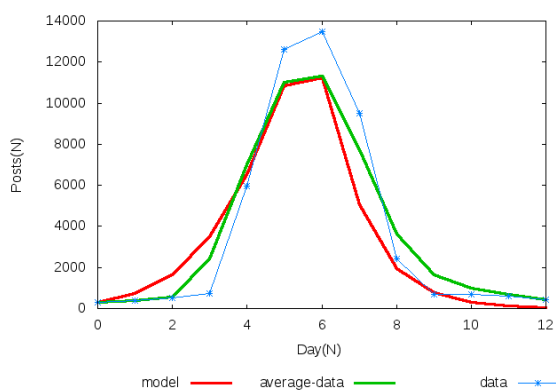


図 22 2017 年 11 月に開催されたポケモン GO イベント時の鳥取砂丘の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。5 日目に 1500、6 日目に 3000 の外力を挿入している。

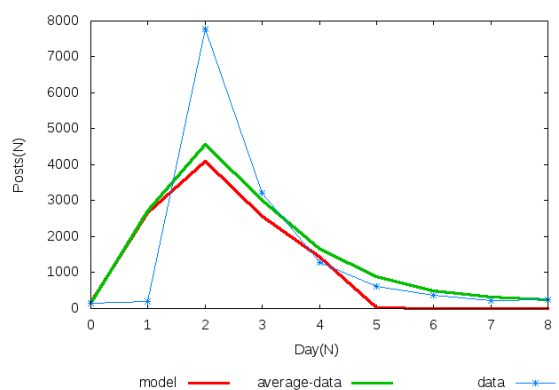


図 23 2017 年 10 月のノーベル賞で話題になった重力波の Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。2 日目に 4000、3 日目に 5000、4 日目に 6000 の外力を挿入している。

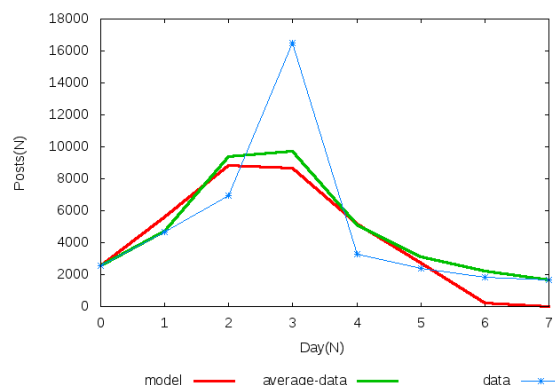


図 24 2017 年 9 月に発生したハリケーンの Twitter 上の投稿数を 3 点平均を行ったものと、計算値の比較である。外力は 2 日目に 8000, 3 日目に 10000, 4 日目に 6000, 5 日目に 6000 の外力を挿入している。

外力 $f(t)$ を考慮することで、図 20 から図 24 のように、外力を考慮しなかったときよりも、平均化したデータと比較した時の誤差を減らすことができた。また、外力を考慮することで計算値が変化するため、外力を考慮したときの誤差を 2 に示す。

表 2 外力を考慮したときの平均データと計算値の誤差 (%)

検索ワード	誤差 (%)
台風 18 号	10.20
AKB48 総選挙	15.75
鳥取砂丘	21.20
重力波	22.11
ハリケーン	17.90

4.6 まとめ

SNS の 1 週間から 2 週間程度の、1 日毎の投稿数データの変化を微分方程式で数理モデルを定義することができた。そして、外力 $f(t)$ を用いることで誤差を減らすことができる。しかし、ブームの形成時に、急激にピーク値まで投稿数が増加するようなブームのふるまいは、定義した数理モデルでは誤差が大きくなることがわかった。

5 応用事例 1：ブームのピーク値予測

5.1 概要

SNS 上のブームは大きく 2 種類ある。1 つはピークに到達する日が予測できるブーム。2 つ目はピークに到達する日が予測できないブームである。ピークが予測できるブームとは、ある日に向かって投稿数が増え、ある日が終わると投稿数が減っていくものである。例え

ば、台風のように日本列島に上陸した時にブームのピークを迎えるなどである。予測できないブームとは、投稿数が増えてもピークを迎えるであろう日を予測できないものである。例えば、ある製品が発売してブームが形成され、投稿数が増え続けている時などである。定義した数理モデル式 (3), 式 (6) を用いるには, t_0 として, ピーク日を代入する必要がある。そのため, ピーク日を予測できないブームには用いることができない。そのため, 4.4 節のフィッティングを用いて, ピークに到達する日と, 値を予測する。

5.2 予測

フィッティングを用いて, ピーク値と, ピークになる日を予測する手順を示す。

1. 予測したい SNS 上のキーワードの 1 日毎の投稿数を取得する。
2. 3 点平均を行う。
3. 取得した日までの投稿数と, 式 (4) を用いてフィッティングを行い計算する。
4. 取得した日数毎のデータで計算したものと平均データと比較を行う。

ピーク値の予測とピーク日を予測するデータは, ゲームソフト「MONSTER HUNTER WORLD」(CAPCOM) についての SNS 上の投稿数と, 2018 年成人式の日話題になった「はれのひ」についての SNS 上の投稿数, そして, 立憲民主党設立時の SNS 上の投稿数についてである。いずれも 1 日毎に取得する。ブーム形成から数日間毎のデータを, 式 (4) を用いてフィッティングを行い, 平均化したデータを比較し, 精度を検証する。

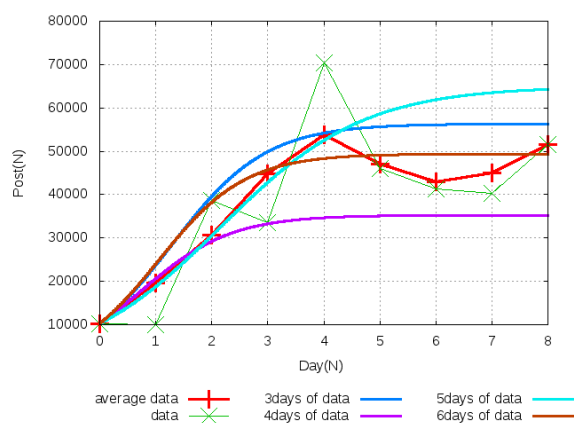


図 25 2018 年 1 月に発売した「MONSTER HUNTER WORD」の SNS 上の実際の投稿数を平均化したデータと 3 日, 4 日, 5 日, 6 日間の投稿数からのピーク値の予測の比較。検索ワードは「MHW」のみである。初日は 2018 年 1 月 22 日である。

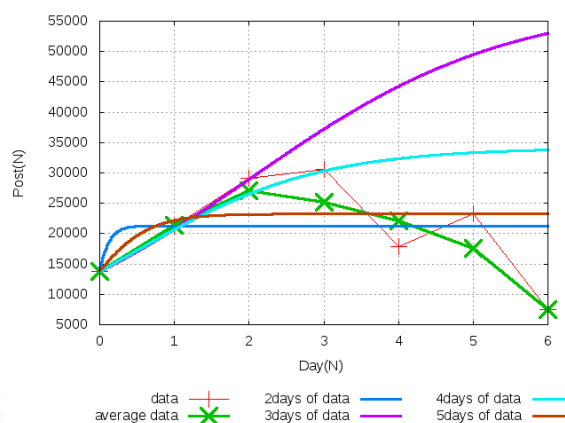


図 26 2018 年 1 月に話題になった「はれのひ」事件の SNS 上の実際の投稿数を平均化したデータと 2 日, 3 日, 4 日, 5 日間の投稿数からのピーク値の予測の比較。検索ワードは「はれのひ」のみである。初日は 2018 年 1 月 08 日である。

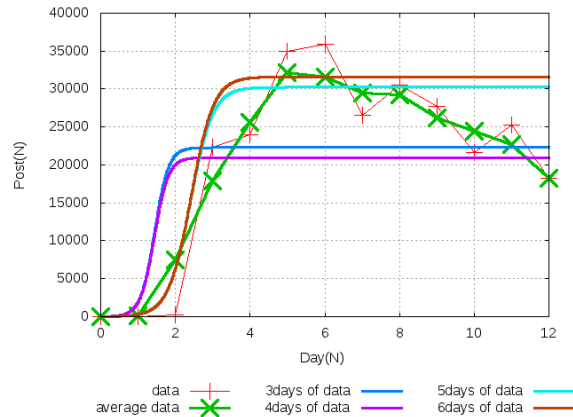


図 27 2017 年立憲民主党が設立した際の SNS 上の実際の投稿数を平均化したデータと、3 日、4 日、5 日、6 日間の投稿数からのピーク値を予測の比較。検索ワードは「立憲民主党」のみである。初日は、2017 年 10 月 29 日である、

図 25 から図 27 は、式 (4) を用いた予測と平均化データの比較結果である。精度を評価するために、平均化データのピーク値をもとに、それぞれの予測データのピーク値を比較し誤差を差として表す。

$$\text{誤差} = \sqrt{(\text{平均化データのピーク値} - \text{予測データのピーク値})^2} \quad (11)$$

表 3 それぞれの日数で予測を行った予測データのピーク値と平均データのピーク値の比較

検索ワード	2 日間	3 日間	4 日間	5 日間	6 日間
MHW	null	1909.36	18668.21	10532.64	4654.31
はれのひ	5764.86	25960.20	6761.66	3863.2	null
立憲民主党	null	9879.68	11200	2059.77	724.02

表 4 それぞれの日数で予測を行った予測データのピーク日と平均データのピーク日の誤差

検索ワード	2 日間	3 日間	4 日間	5 日間	6 日間
MHW	null	1 日	1 日	4 日	1 日
はれのひ	1 日	4 日	4 日	1 日	null
立憲民主党	null	2 日	2 日	1 日	1 日

表 3,4 より、よりたくさんの日数毎のデータを予測に用いることで、ピーク値も、ピーク日の誤差も少なくなることがわかる。

5.3 まとめ

表 3, 4 より、よりたくさんの日数毎のデータを予測に用いることで、ピーク値も、ピーク日の誤差も数なくなることがわかる。このことより、ピーク日がわからないブームに対し

ては、このような手法を用いることでピークに達すると予測できる値と、日を求めることができる。

6 応用事例 2：SNS 上のブームの形成・終焉予測

6.1 概要

4章で説明した数理モデルを用いて、SNS 社会のブームの形成・終焉について予測を応用事例として行う。本章で取り扱うデータは、2017 年中に発生した台風、台風 18 号、台風 21 号、そして、台風 22 号の発生日からのブームの形成・終焉時の投稿数を取り扱う。

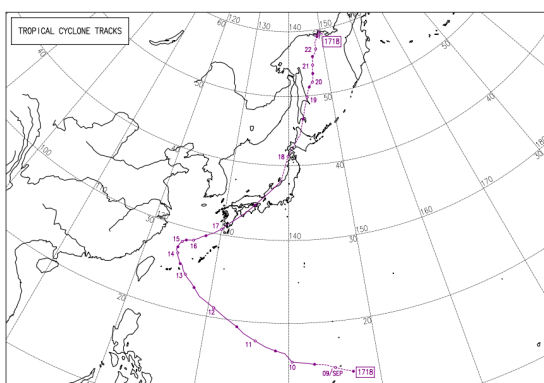


図 28 台風 18 号経路図 平成 29 年 (2017 年). 図中の○は、その日の午前 9 時, ●は、その日の午後 9 時を示す [2].

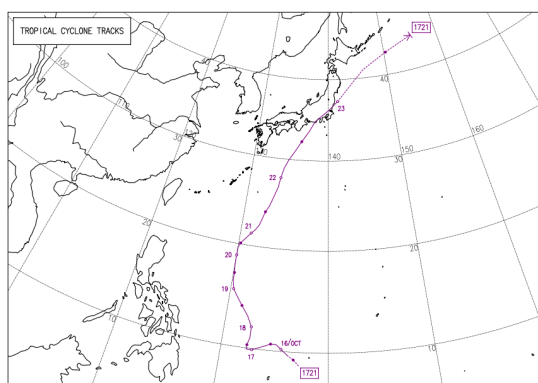


図 29 台風 21 号経路図 平成 29 年 (2017 年). 図中の○は、その日の午前 9 時, ●は、その日の午後 9 時を示す [2].

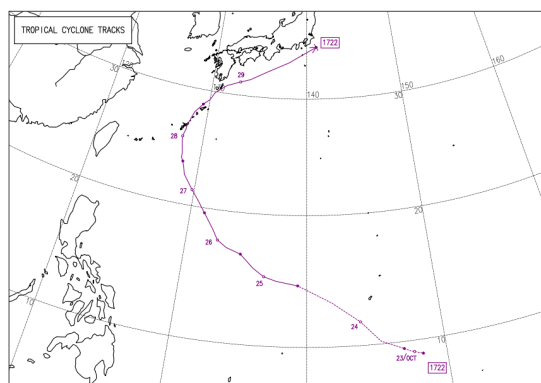


図 30 台風 22 号経路図 平成 29 年 (2017 年). 図中の○は、その日の午前 9 時, ●は、その日の午後 9 時を示す [2].

図 28 から図 30 は、それぞれの台風の経路図である。台風 18 号 (図 28) は、日本大陸に上陸しないと思われていた。しかし、急な進路変更で日本大陸に上陸した。また、台風 22 号 (図 30) は、台風 21 号 (図 29) が日本大陸に上陸した後、すぐ発生したという特徴がある。

6.2 予測手順

台風 18 号, 台風 21 号, そして, 台風 22 号の投稿数を取得し, 式 (3) を用いて, ブームの形成・終焉を予測する. 各台風の発生日からのデータを 1 日毎に取得し, 取得したデータより予測を行う. 予測手順を以下に示す.

1. 予測したい台風の Twitter 上の 1 日毎の投稿数を取得する.
2. 3 点平均を行う.
3. 取得した日までの投稿数と, 式 (4) を用いてフィッティングを行い a と p のパラメータを決定する.
4. ブームの落ち着く時間 t_0 を任意に決定する.
5. 式 (3) に決定したパラメータを代入し計算する.
6. 式 (3) の $f(t)$ を任意に代入することで, 様々な予測を行う.

ただし, 上記の手順で予測を行うにあたり, 初日もしくは日数が少ないと式 (3) の数理モデルでは予測が困難である. 最低でも 5 日から 7 日間の投稿数データが必要である.

予測手順 1 より, 台風 18 号, 台風 21 号, そして, 台風 22 号の 1 日毎の Twitter 上の投稿数を取得する. 台風発生日を $Day = 0$ とし, 図 31 から図 33 に発生日からの投稿数のグラフを記載する.

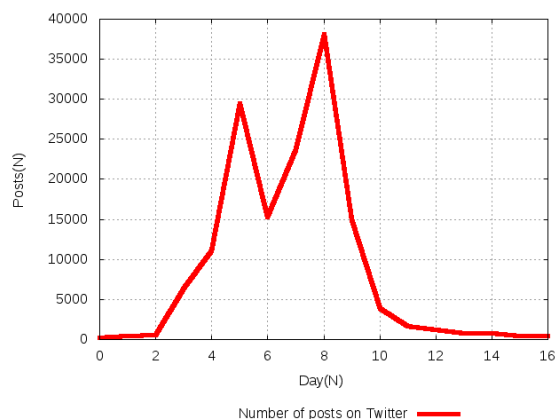


図 31 台風 18 号についての Twitter 上の投稿数. $Day = 0$ は, 発生日を示す. 台風発生日は 9 月 09 日.

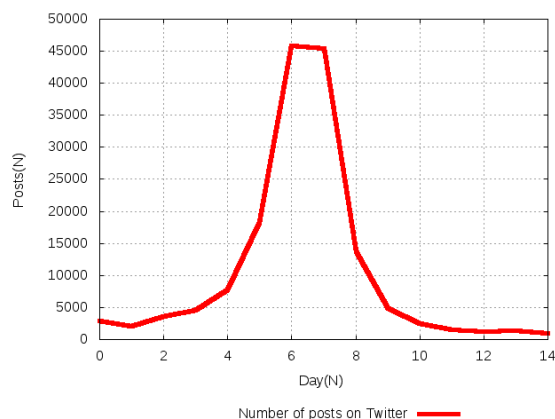


図 32 台風 21 号についての Twitter 上の投稿数. $Day = 0$ は, 発生日を示す. 台風発生日は 10 月 16 日.

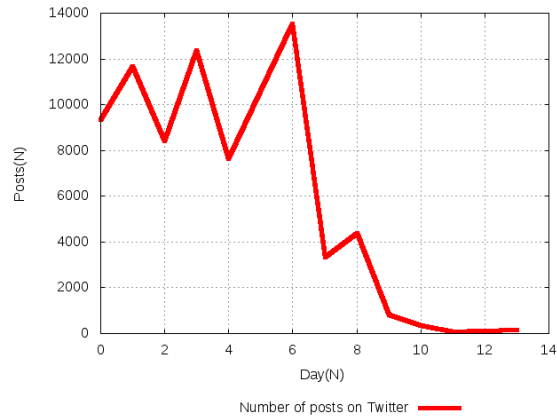


図 33 台風 22 号についての Twitter 上の投稿数. $Day = 0$ は, 発生日を示す. 台風発生日 11 月 23 日.

6.3 ブームの形成・終焉の予測

台風 18 号, 台風 21 号, そして, 台風 22 号と 5 日間もしくは, 6 日間のデータを用いて予測する. それぞれのデータを表 5 に示す.

表 5 各台風の発生日からの 5 日間の投稿数.

Day	台風 18 号	台風 21 号	台風 22 号
0	267	2870	9331
1	479	2104	11686
2	546	3543	8424
3	6440	4630	12367
4	11102	7759	7641
5	null	18235	10575

まず, 表 5 のデータを用いてフィッティング作業を行う. フィッティング作業で得られる, パラメータ a_0 と p を用いて式 (3) に代入し, ブームの形成・終焉の予測を行う. フィッティング作業で得られたパラメータを表 6 に示す. また, t_0 はブームが落ち着いた時間を示すため, 予測者が t_0 を任意に決定することができる. そのため, 多様な予測をすることが可能である.

表 6 表 5 のデータを用いたフィッティング作業から得られるパラメータ.

検索ワード	a_0	p
台風 18 号	1.257820	-0.000087
台風 21 号	0.416270	-0.000016
台風 22 号	1.447700	-0.000144

台風 18 号では, 5 日間のデータ, 台風 21 号・台風 22 号は 6 日間のデータを用いて予測

を行った。 $f(t) = 0$ として予測を行うと図 34 から図 36 のようになる。

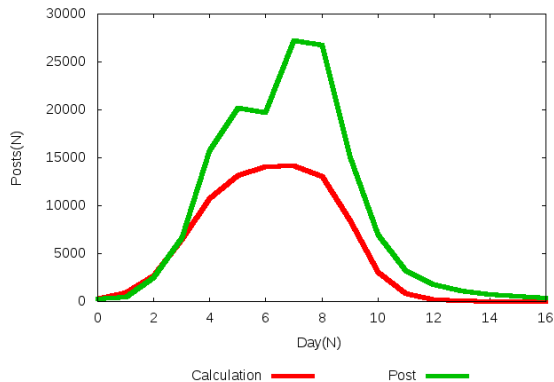


図 34 台風 18 号の 5 日間の投稿数を用いた予測と、実際の投稿数に 3 点平均を行った値との比較図。

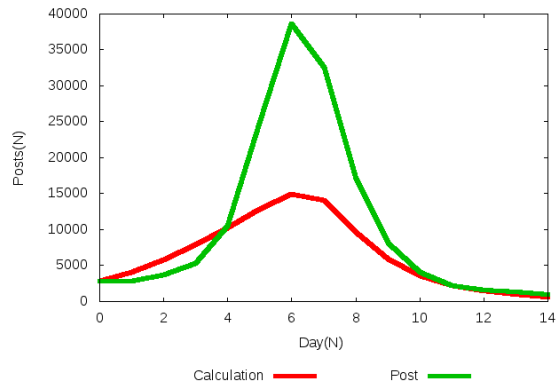


図 35 台風 21 号の 6 日間の投稿数を用いた予測と、実際の投稿数に 3 点平均を行った値との比較図。

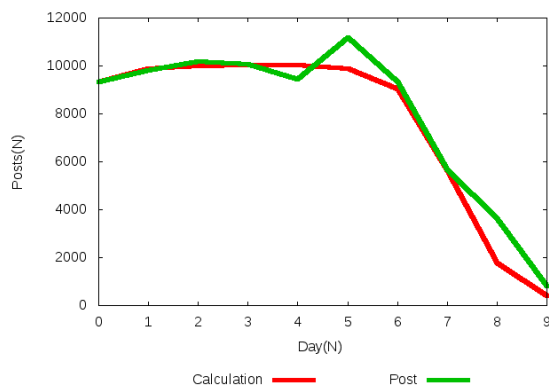


図 36 台風 22 号の 6 日間の投稿数を用いた予測と、実際の投稿数に 3 点平均を行った値との比較図。

図 34 から図 36 と、6.1 節の図 28 から図 30 とを比較すると、各台風の予測の一致していない部分と、各台風の進路図の日本大陸に接近・上陸している日に対応していることがわかる。このことは、各台風が日本に接近・上陸したことが外力として、Twitter 上の投稿数に影響しているといえる。

また、台風 18 号 (図 28) は、最初日本大陸に上陸しないと予測されていた。もし、台風 18 号が日本大陸に来なかったと考え、ブームの形成・終焉の予測は、図 37 のようになったであろうと予測することが可能である。

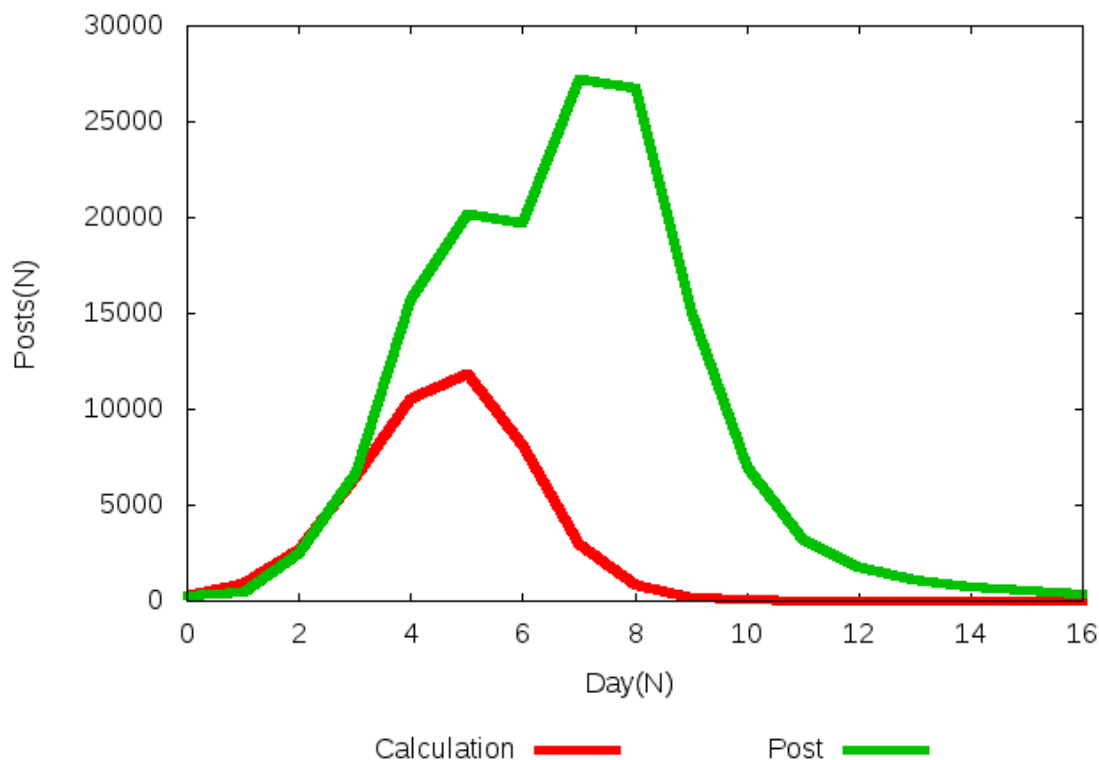


図 37 台風 18 号の 5 日間の投稿数を用いた、もし日本大陸に接近・上陸しなかった場合の予測と、実際の投稿数に 3 点平均を行った値との比較図。

6.4 まとめ

本章では、式 (3) の SNS における数理モデルを用いて台風 18 号、台風 21 号、そして、台風 22 号についての、Twitter 上のブームの形成・終焉の予測を行った。5 日間から、7 日間のデータを用いることで予測を立てることが可能であるが、外力を考慮しないと上手く予測を立てることができない。しかし、外力や、ブームが落ち着く時間を任意に設定することで、多様な予測が可能である。そして、台風 18 号と、台風 21 号は、日本列島に台風が上陸しているが、台風 22 号は日本列島に接近しただけである。このことと、図 34 から図 36 より、接近と上陸で外力の値が大きく変わることがわかる。

7 結論

本研究は、SNS 上の日毎の投稿数の増減を微分方程式で数理モデル化することが目標であった。そこで、論文 [1] を参考に、3 項の微分方程式で、SNS 上の投稿数の増減を数理モデルで定義することができた。本研究では 1 週間から 2 週間の SNS 上のデータのみを用い

た。それは、数日で投稿数が急増し、ピークを迎え、投稿数が急減を行うようなブームは、定義した数理モデルで予測できず、また1か月程度のブームでは、投稿数の盛衰が激しく、予測を行うことができなかつたためである。また、1日ごとではなく1時間ごとで収集すると、21時から1時にかけて投稿数が増加し、2時から6時までは、投稿数が減少するなど、本研究の目的とは異なるため、投稿数データの取得は1日毎に取得した。

参考文献

- [1] Ishii A, Arakaki H, Matsuda N, Umemura S, Urushidani T, Yamagata N, and Yoshida N, "The 'hit' phenomenon: a mathematical model of human dynamics interactions as a stochastic process", *New Journal of Physics* 14(2012) 063018
- [2] 気象庁ホームページ (http://www.data.jma.go.jp/fcd/yoho/typhoon/route_map/bstv2017.html) 参照日 (2018-02-12)