

グループ演習 最終発表

# 人々の関心度推移モデルの構築

2009/10/2 4班

原田 礼朗

福井 大樹

安田 健一郎

呉 迪

アドバイザー教員: 遠藤 靖典

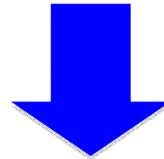
# 発表の流れ

- 1 はじめに
- 2 データ収集
- 3 関心度の分類
- 4 関心度推移モデル
- 5 関心度推移モデルを用いた予測手法の提案
- 6 実際の事例を用いた予測手法の検証
- 7 おわりに

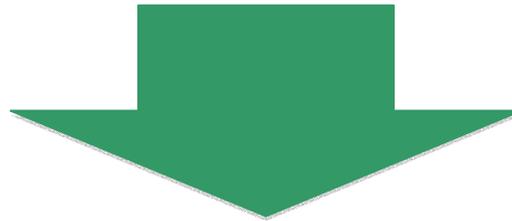
# 1 はじめに

# 1.1 人々の関心について

ニュースなどによる報道



人々の関心

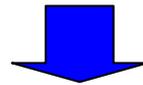


人々の関心について調査することで、  
報道内容や報道量が適切かどうかを判断が可能。

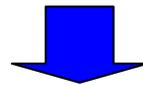
## 1.2.1 研究背景

大規模な社会調査

→社会問題の存在と規模の明確化



インターネット上に存在する  
多くの情報を活用した調査



ブログ

## 1.2.2 研究背景

- 福原ら[1]

ブログ記事を用いて社会的関心動向を分析することで、人々の社会問題についての反応や関心を容易に把握できる。

## 1.3 研究の目的

- 過去の事例の**関心度の推移を分類**。
- 報道量と関心度の推移の関係を微分方程式を用いて**モデル化**。
- 現在進行中の事象に対する、将来の**関心度の推移を予測する手法の提案**。

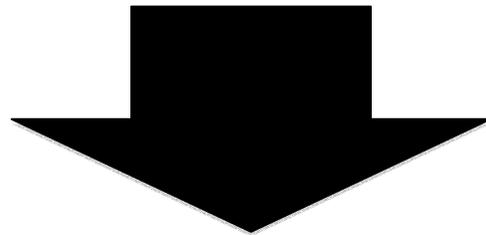
報道量：事例に関するニュースの記事数

関心度：事例に対して関心を持つ人・ブログ記事の数

## 2 実データの収集方法

## 2.1 関心度の推移を示す実データ

- ミクシィ日記キーワードランキング
- kizasi.jp
- YAHOO!JAPANブログ検索

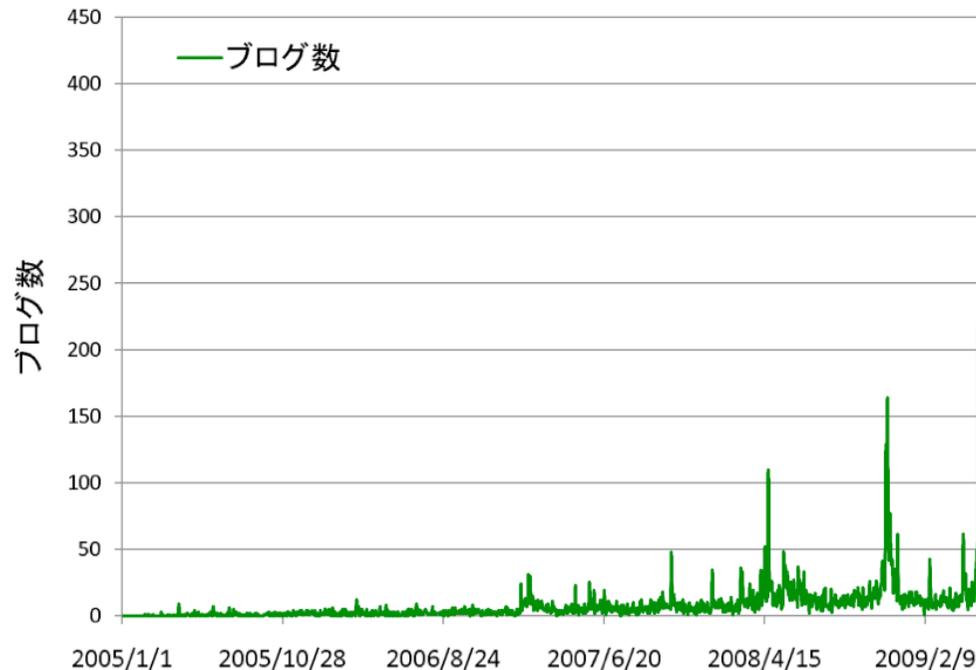


数値データ収集の容易さ、データ収集可能期間の長さからYAHOO!JAPANブログ検索を利用

## 2.2 関心度の推移を示す実データ

### YAHOO!JAPANブログ検索

- キーワードを入力することにより、そのキーワードに関連するネット上のブログを紹介するサービス。
- キーワードに関する注目度(YAHOO!ブログ内のブログ記事の出現頻度)の推移を示すグラフを表示。

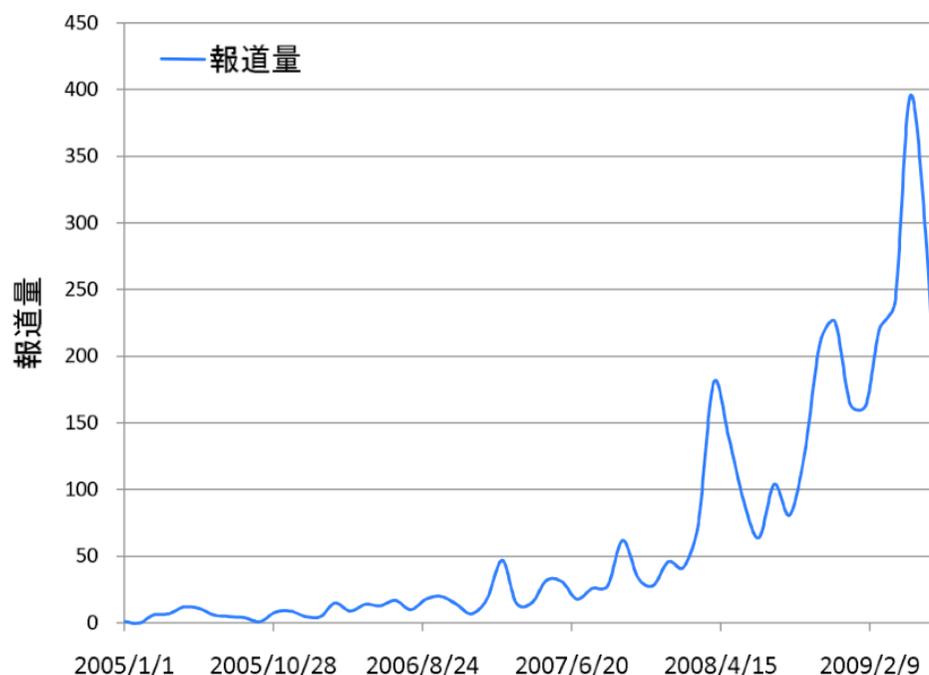


キーワード「裁判員制度」のブログ数の推移 出典文献:[4]

## 2.3 報道量の推移を示す実データ

### Google News Archive Search

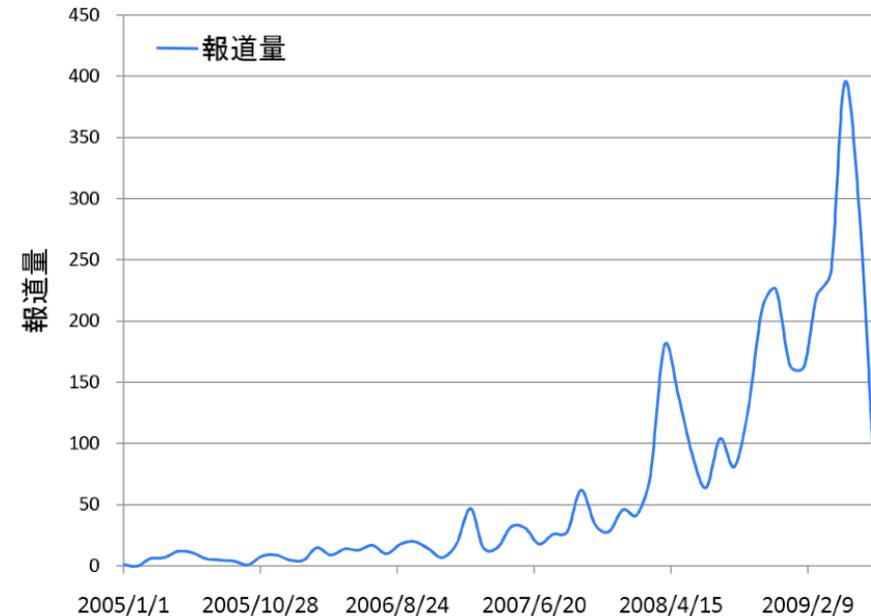
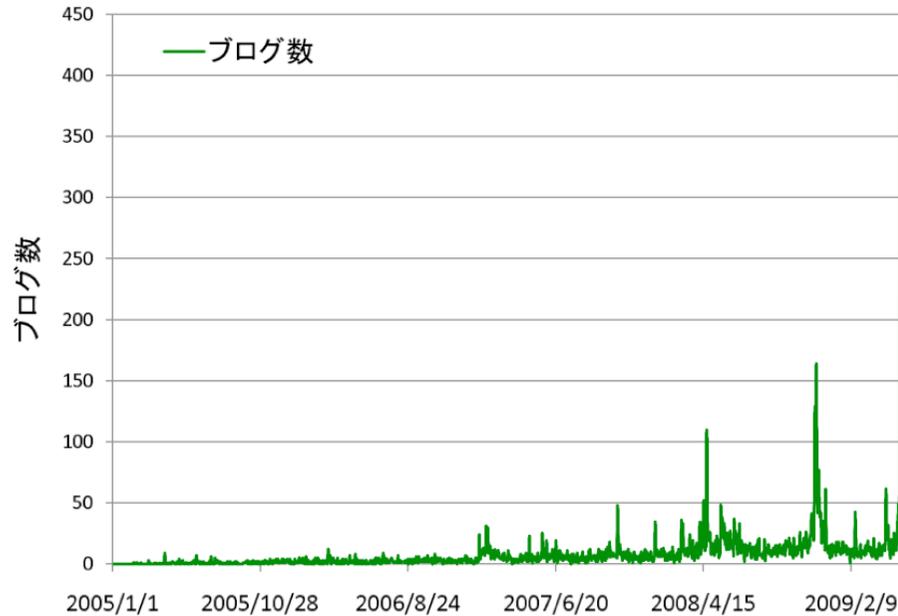
- 米googleによる過去のニュース記事を検索するwebシステム。
- 610以上のニュースソースから記事見出しを収集し・表示。



キーワード「裁判員制度」の報道量の推移

出典文献:[5]

## 2.4 関心度と報道量の推移の比較 ～裁判員制度～



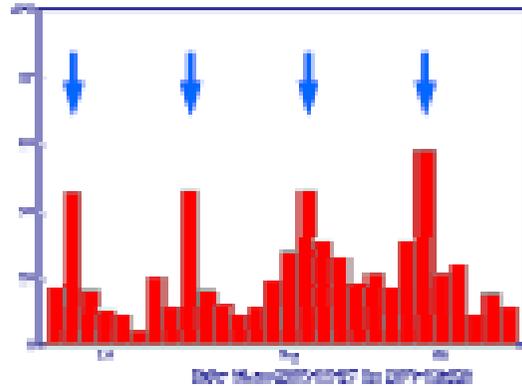
- 報道量と関心度には相関関係。
- 報道量のピークがあるところで関心度もピークとなっているが、報道量と関心度は単なる比例関係ではない。

# 3 関心度推移の分類

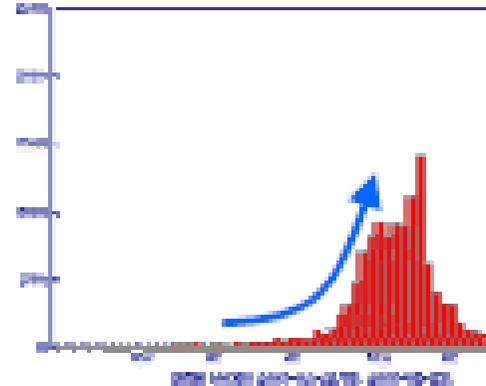
## 3.1 関心度推移の分類

- 福原ら[1]は下記のような社会的関心のパターンを考えている。
- 社会関心のパターン
  - I 周期性
  - II 漸次増加型
  - III 突発型
  - IV 関心持続型
  - V その他

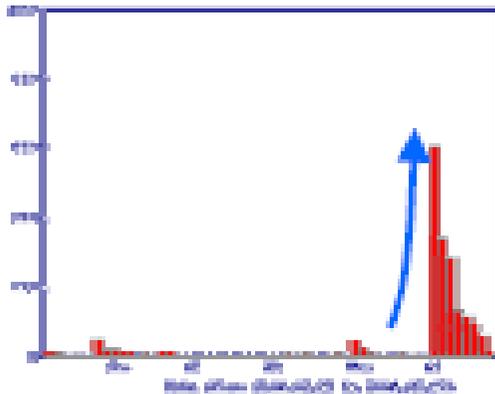
# 3.1 関心度推移の分類



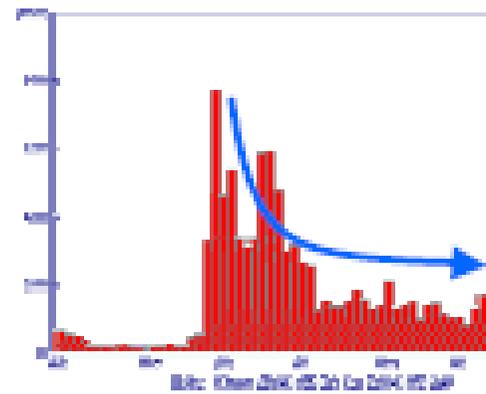
I. 周期型



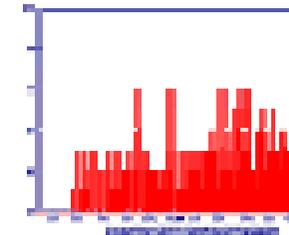
II. 漸次増加型



III. 突発型



IV. 関心持続型



V. その他

出典文献:[1]

## 3.2 関心度推移の分類

- 本研究では、下記のような4つのパターンを提案する。

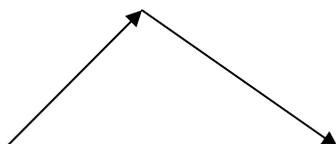
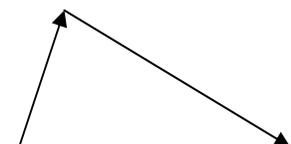
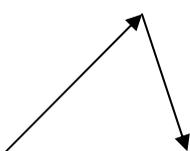
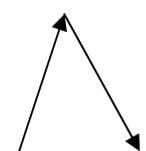
	漸次増加型	突発型
関心持続型		
関心収束型		

図1

## 3.2 関心度推移の分類

- 提案するパターンに属する事例

	漸次増加型	突発型
関心持続型	<p>トリノオリンピック 北京オリンピック ワールドカップドイツ大会 裁判員制度 郵政民営化(保留) GM 崖の上のポニョ 鳥インフルエンザ</p>	<p>四川省地震 新型インフルエンザ チベット問題 世界同時株安 金融不安 おくりびと 大分県 教育委員会</p>
関心収束型	<p>第1回ワールドベースボールクラシック 第2回ワールドベースボールクラシック 洞爺湖サミット開催 リーマン・ブラザーズ タスポ 北朝鮮 ミサイル 冥王星 食品偽装(保留)</p>	<p>スマトラ沖地震 パキスタン地震 能登半島地震 ペルー沖地震 安倍内閣 麻生内閣 小沢辞任 臓器移植改正案 草なぎ 逮捕 マイケル 死 サザンオールスターズ活動休止 福知山線 脱線事故 中央大 教授 刺殺 北朝鮮 核実験 足利事件 ガザ侵攻 インド ムンバイ 同時テロ 秋葉原 無差別殺傷事件 iPhone 日本 若田 宇宙 タスポ ノーベル賞 日本人 4人</p>

図2

## 3.2 関心度推移の分類

- ニュースはカテゴリごとに、提案するパターンに属することが分かる。

	漸次増加型	突発型
関心持続型	スポーツ 映画 病気	映画 病気
関心収束型	スポーツ 生活 政治関連	災害 政治関連 芸能 事件・事故 時事関連

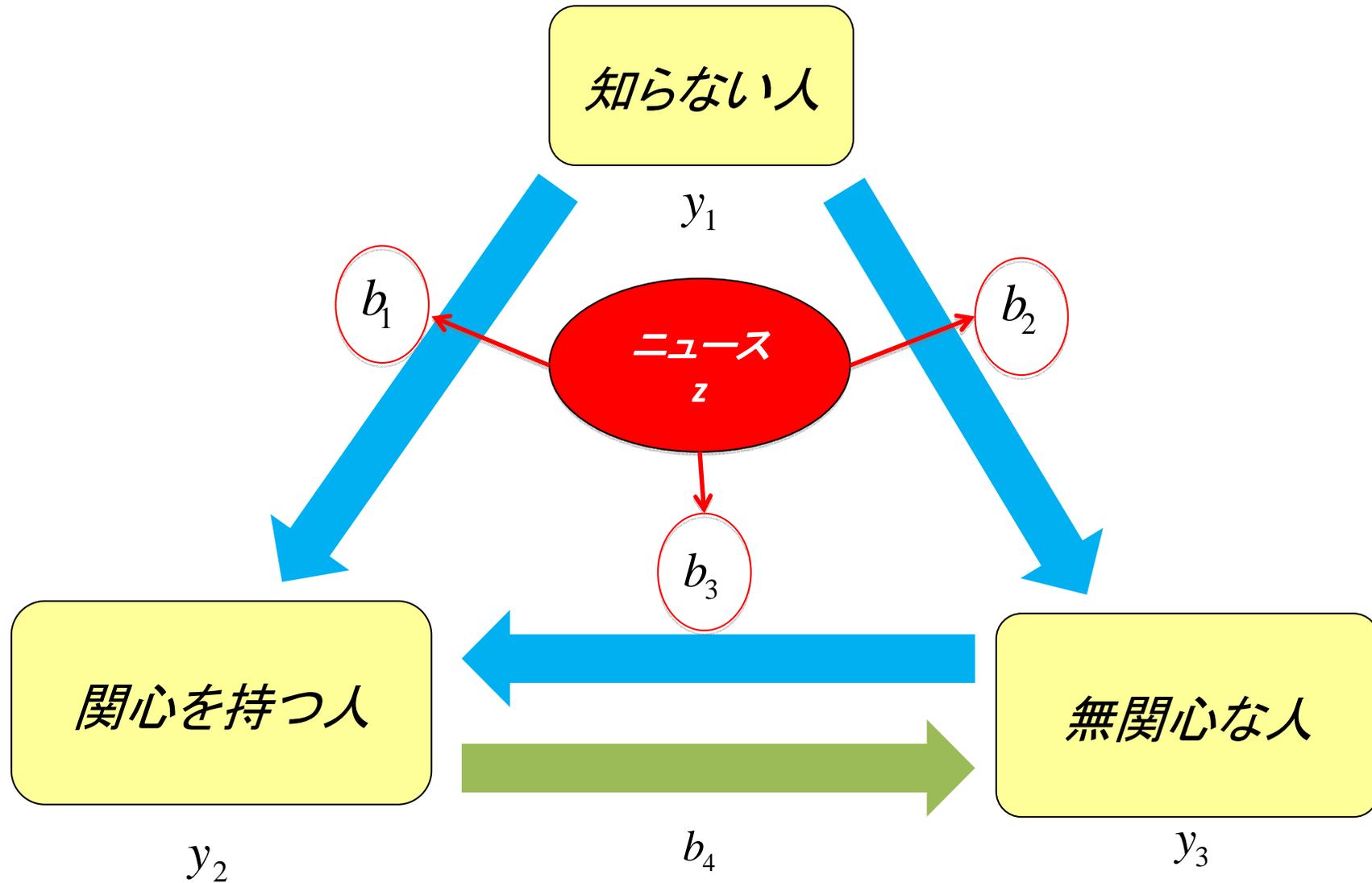
図3

## 3.2 関心度推移の分類

- しかし、提案するパターンに当てはまらない例外がある。
- 四川省地震 ⇒ 突発・関心持続型
- チベット問題 ⇒ 突発・関心持続型
- 裁判員制度 ⇒ 漸次増加・関心持続型
- 大分汚職事件 ⇒ 突発・関心持続型

# 4 関心度推移モデル

# 4.1 関心度推移モデル



## 4.2 関心度推移モデル

$$y_1'(t) = -b_1 z(t)y_1(t) - b_2 z(t)y_1(t)$$

$$y_2'(t) = b_1 z(t)y_1(t) + b_3 z(t)y_3(t) - b_4 y_2(t)$$

$$y_3'(t) = b_2 z(t)y_1(t) + b_4 y_2(t) - b_3 z(t)y_3(t)$$

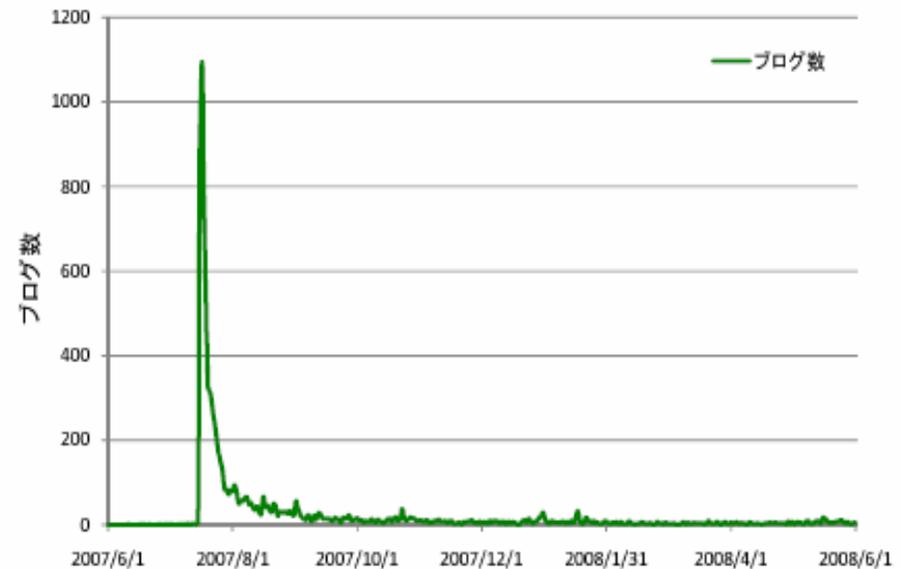
$z$  : ニュース     $y_1$  : 知らない人     $y_2$  : 関心を持つ人     $y_3$  : 無関心な人  
 $b_1, b_2, b_3, b_4$  : パラメータ

yahoo ブログ件数は 2009 年 4 月 29 日現在、200 万人となっているため、 $y_1(t) + y_2(t) + y_3(t) = 2000000$  とする。

## 4.3 関心度推移モデル

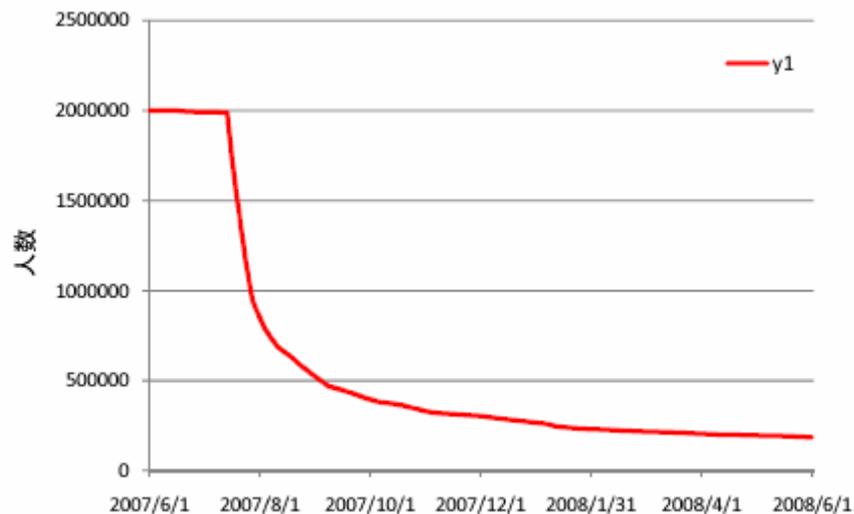


「中越沖 地震」の報道量  $z(t)$



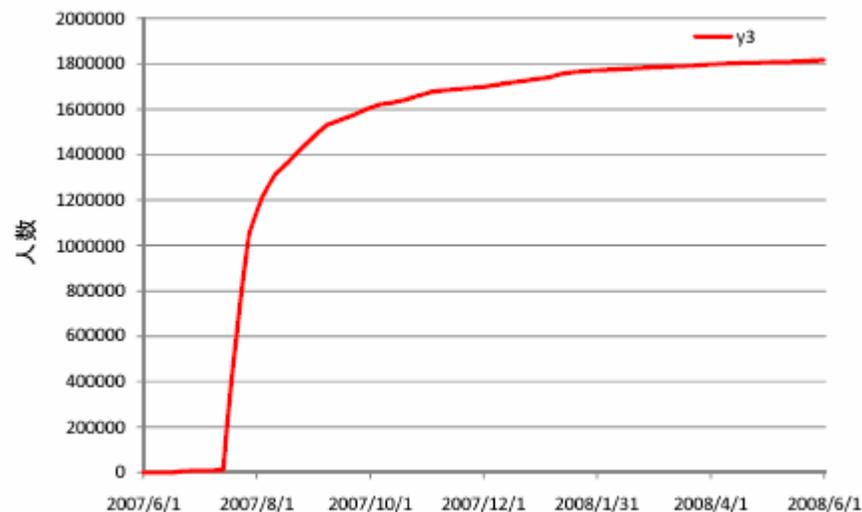
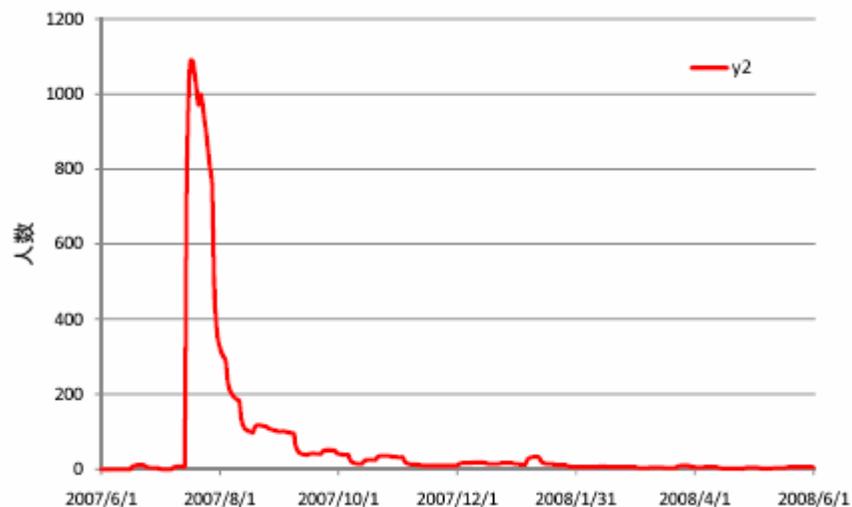
yahooブログによる関心度推移

# 4.4 関心度推移モデル



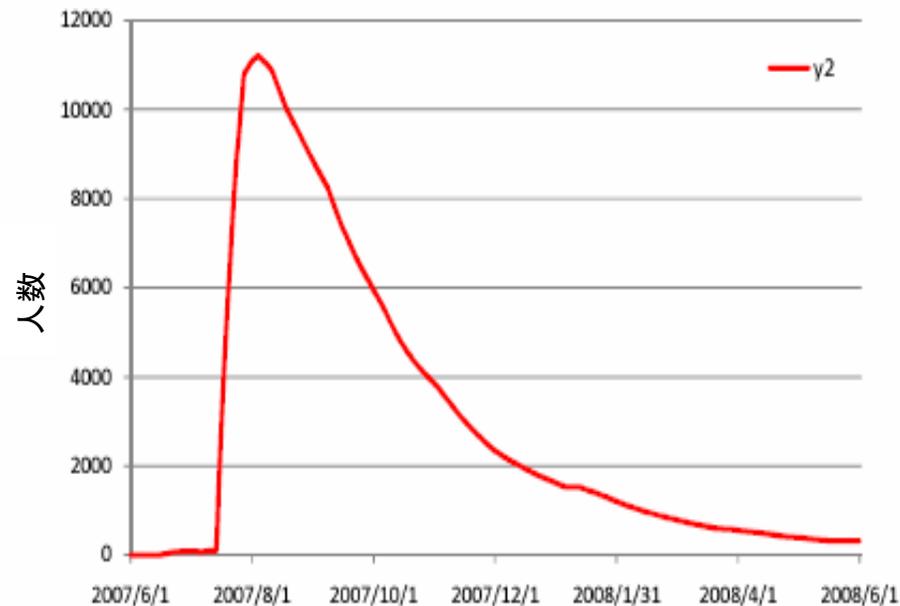
$$b_1 = 0.012, b_2 = 0.988, b_3 = 0.0001, b_4 = 0.9$$

のときの $y_1(t)$ ,  $y_2(t)$ ,  $y_3(t)$ の時系列変化



## 4.5 関心度推移モデル

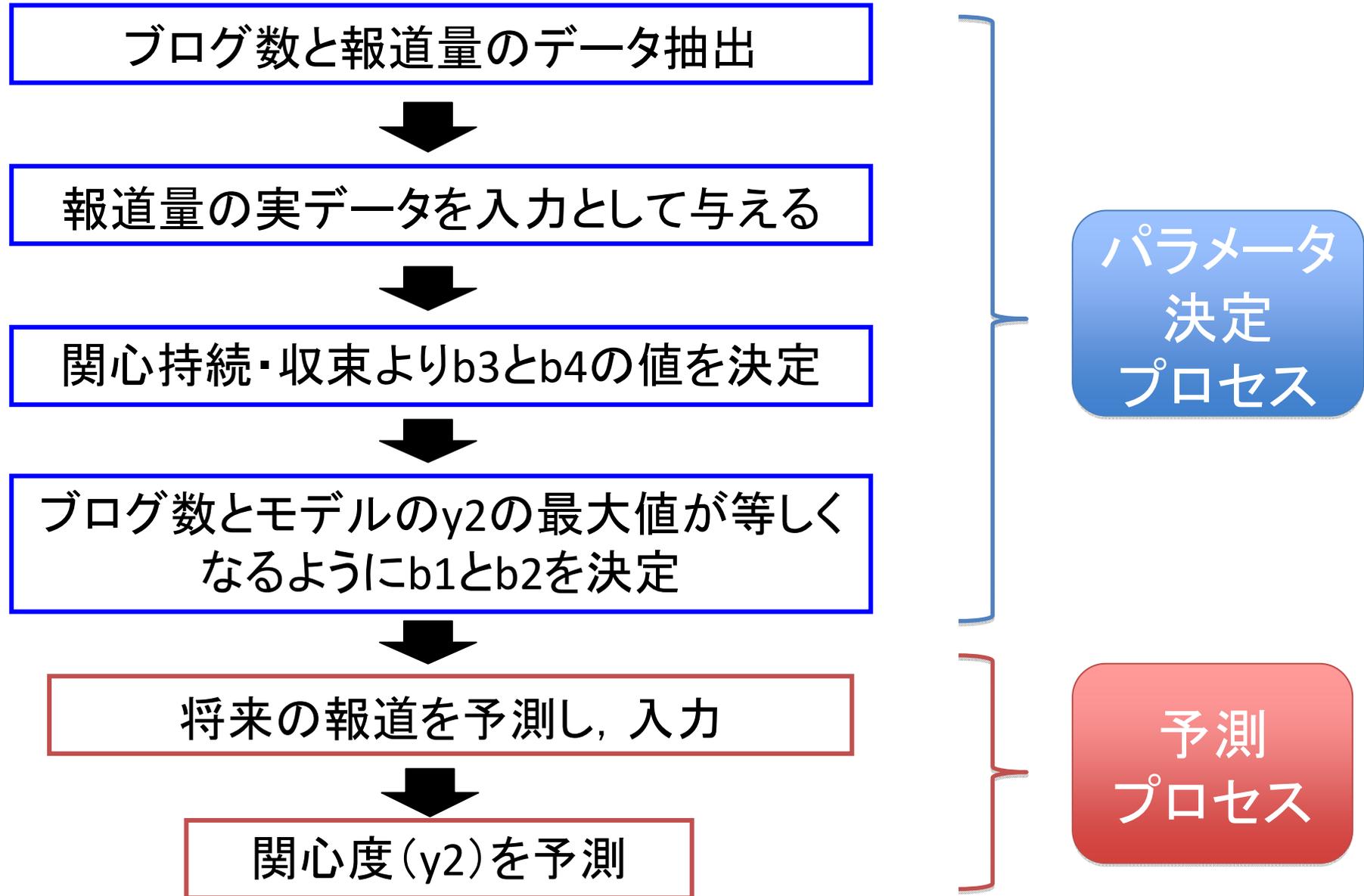
- さらに、パラメータを変化させたとき、 $y_2$ は関心持続型に変化する。 $z$



$$b_1 = 0.012, b_2 = 0.988, b_3 = 0.0001, b_4 = 0.02$$

# 5 関心度推移モデルを用いた 予測手法の提案

# 関心度予測手法の流れ

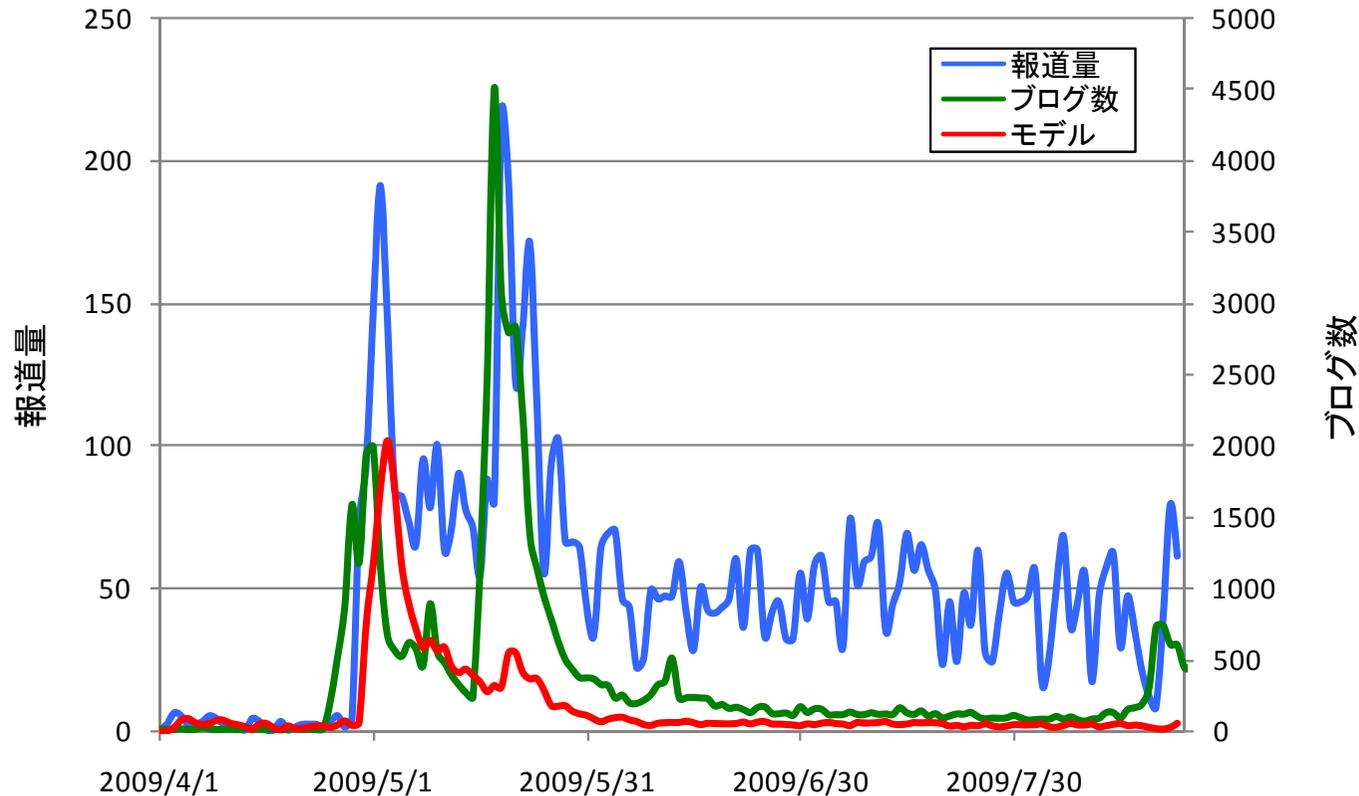


# 6 実際の事例を用いた 予測手法の検証

## 6.1 キーワード「新型インフルエンザ」

- 新たに人から人に伝染する能力を有することとなったウイルスを病原体とするインフルエンザ。
- 2009年4月下旬頃から世界中で流行。

# 関心度推移モデルのパラメータの決定

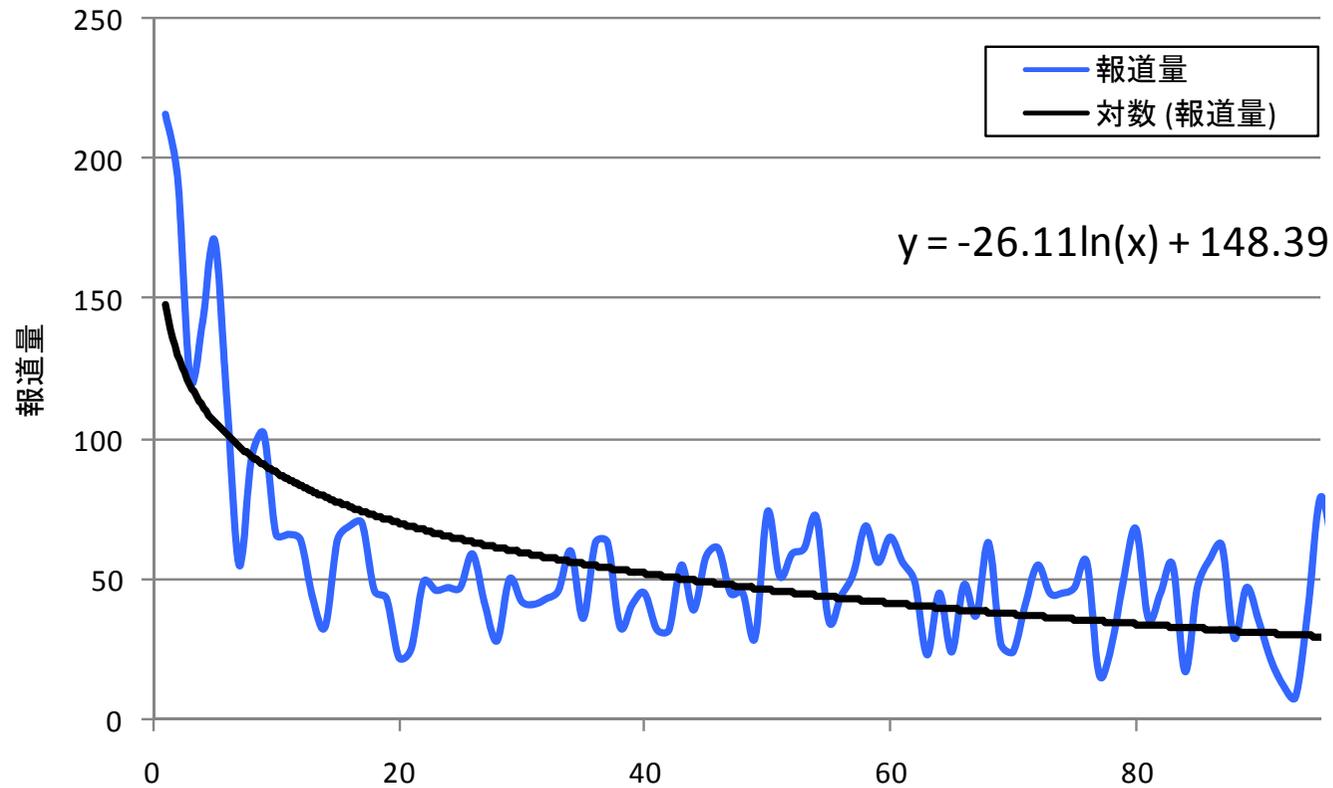


1つめのピークに合わせた場合の同定結果

( $b_1=0.0085$ ,  $b_2=0.9915$ ,  $b_3=0.0004$ ,  $b_4=0.9$ )

- 報道内容や報道量が過剰すぎるため、このモデルではうまく表現できない。

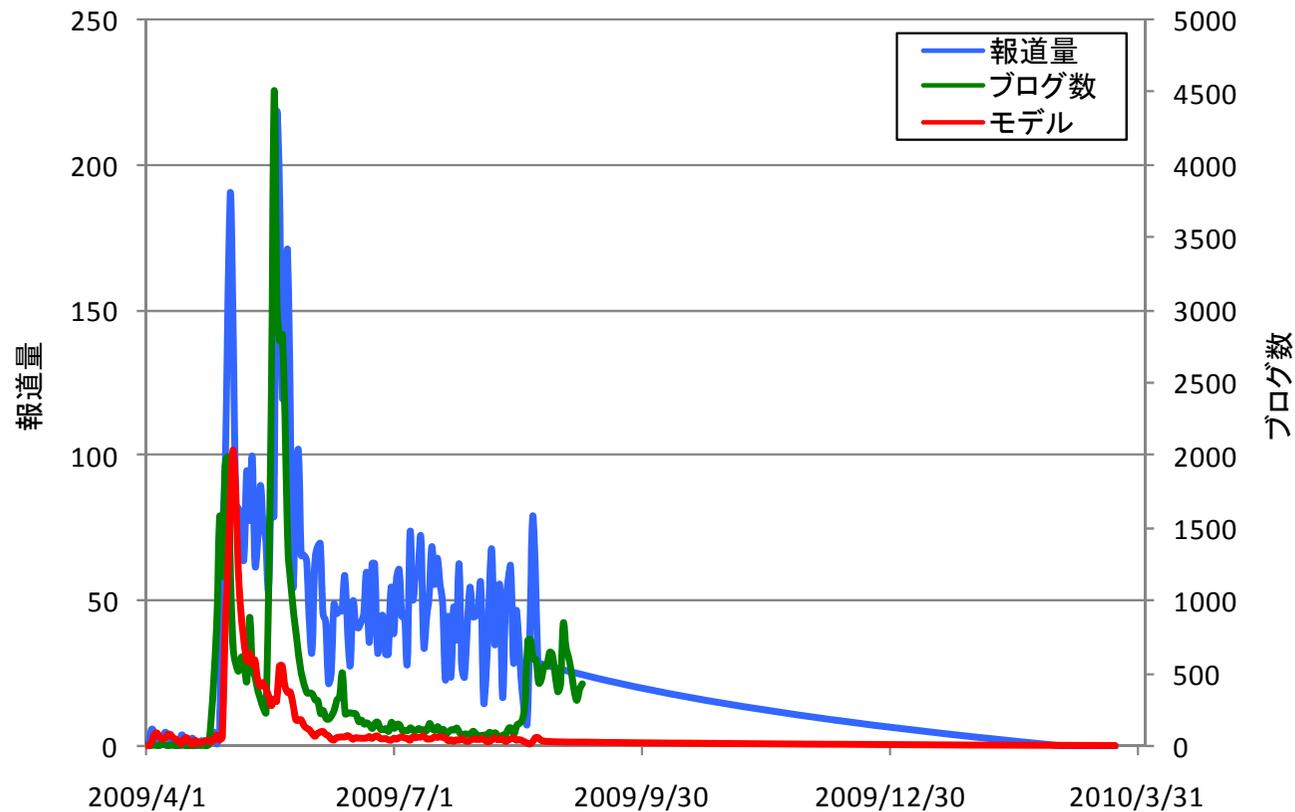
# 関心度推移の予測1



報道量の対数近似

# 関心度推移の予測1

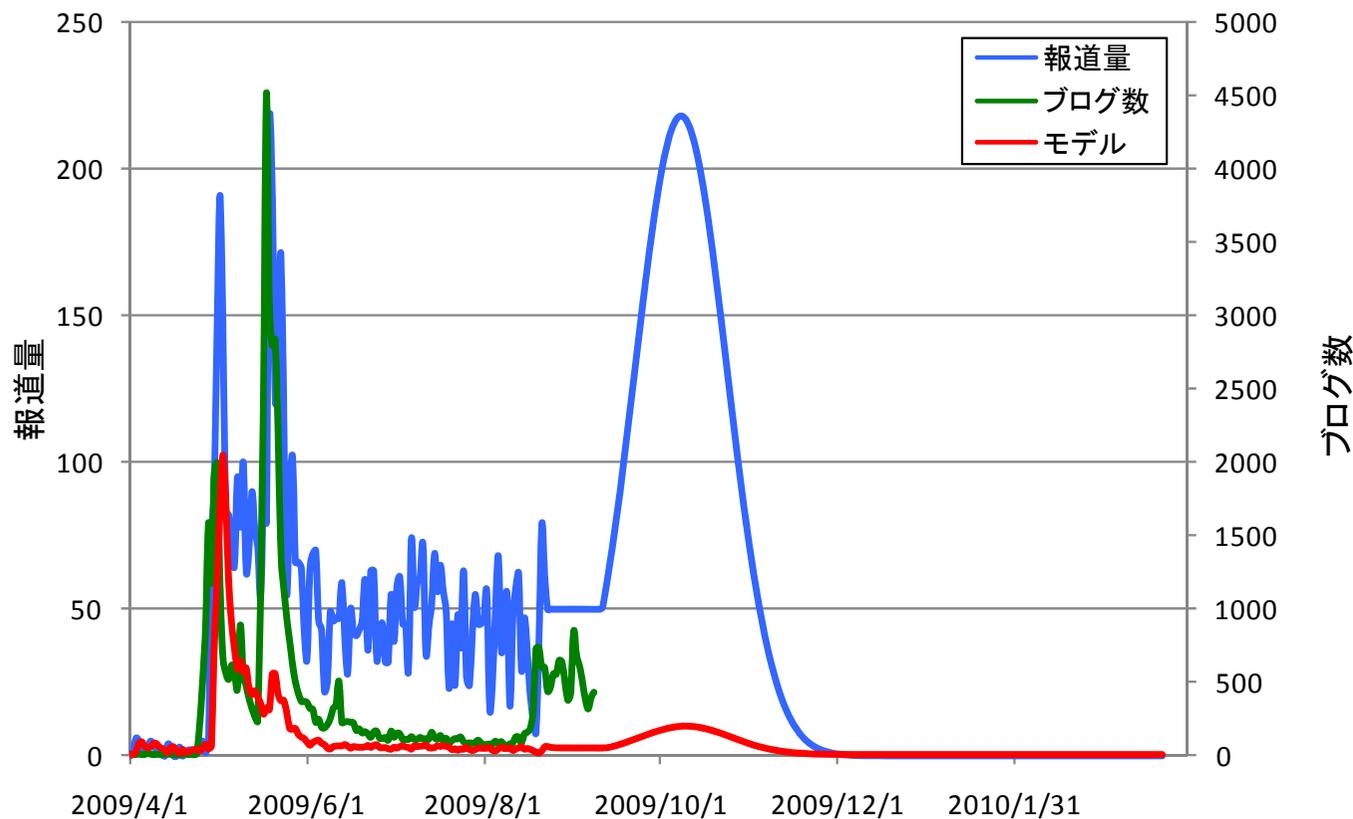
報道が収束した場合



- モデルはほとんどゼロのまま。

# 関心度推移の予測2

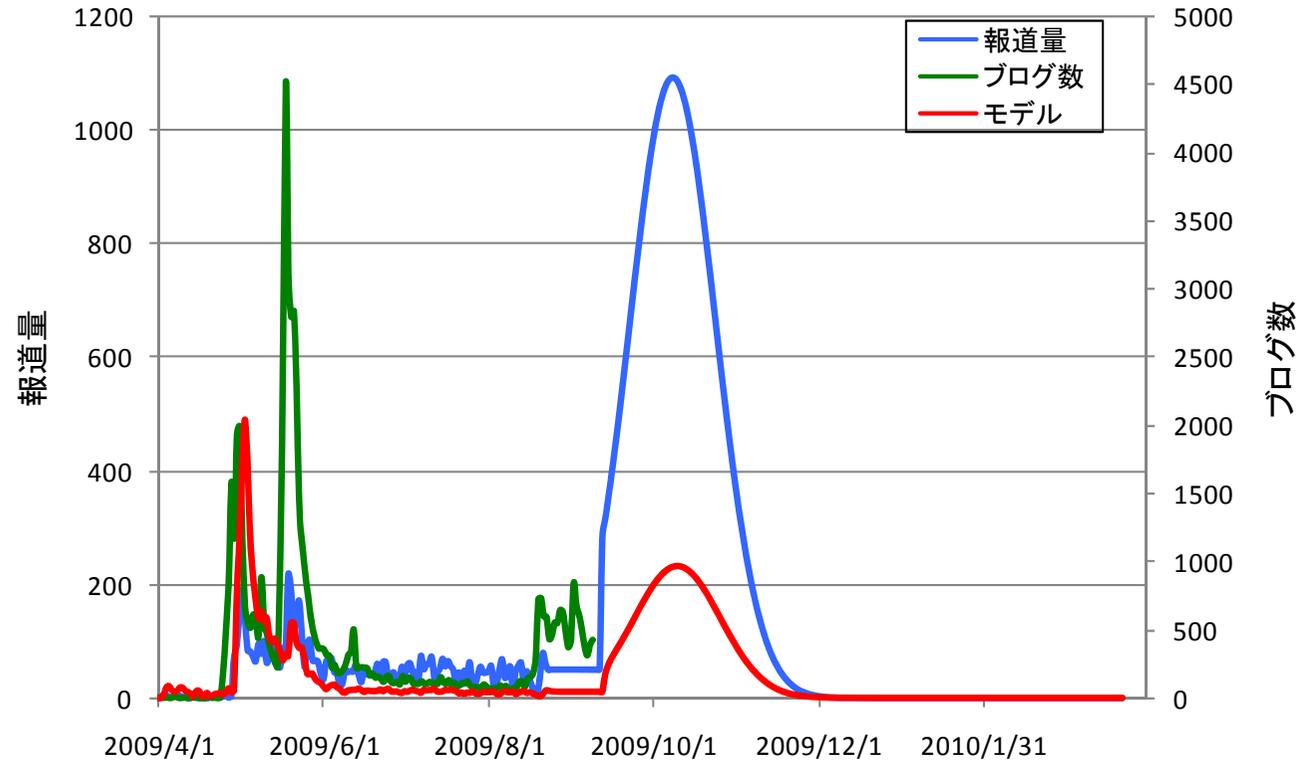
流行予測に合わせて報道量を与えた場合



- 報道量が再び生ずれば関心度も再び生ずる。

# 関心度推移の予測3

流行予測に合わせて大きな報道量を与えた場合

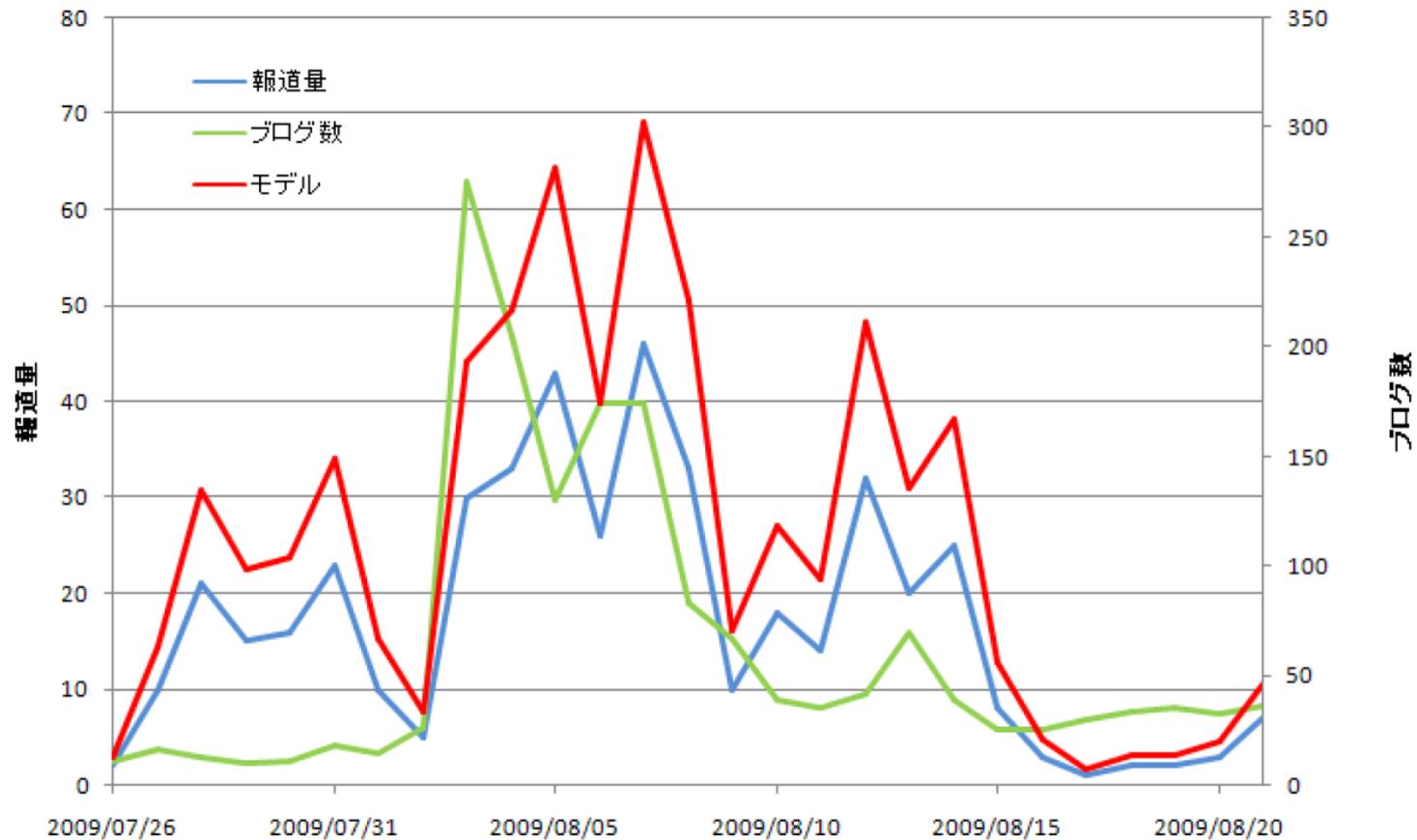


- 報道量の増加に伴い、関心度も増加。
- 感染率の高まる10月に関心が上がらず、感染リスクが高まる恐れ。

## 6.2 キーワード「裁判員制度」

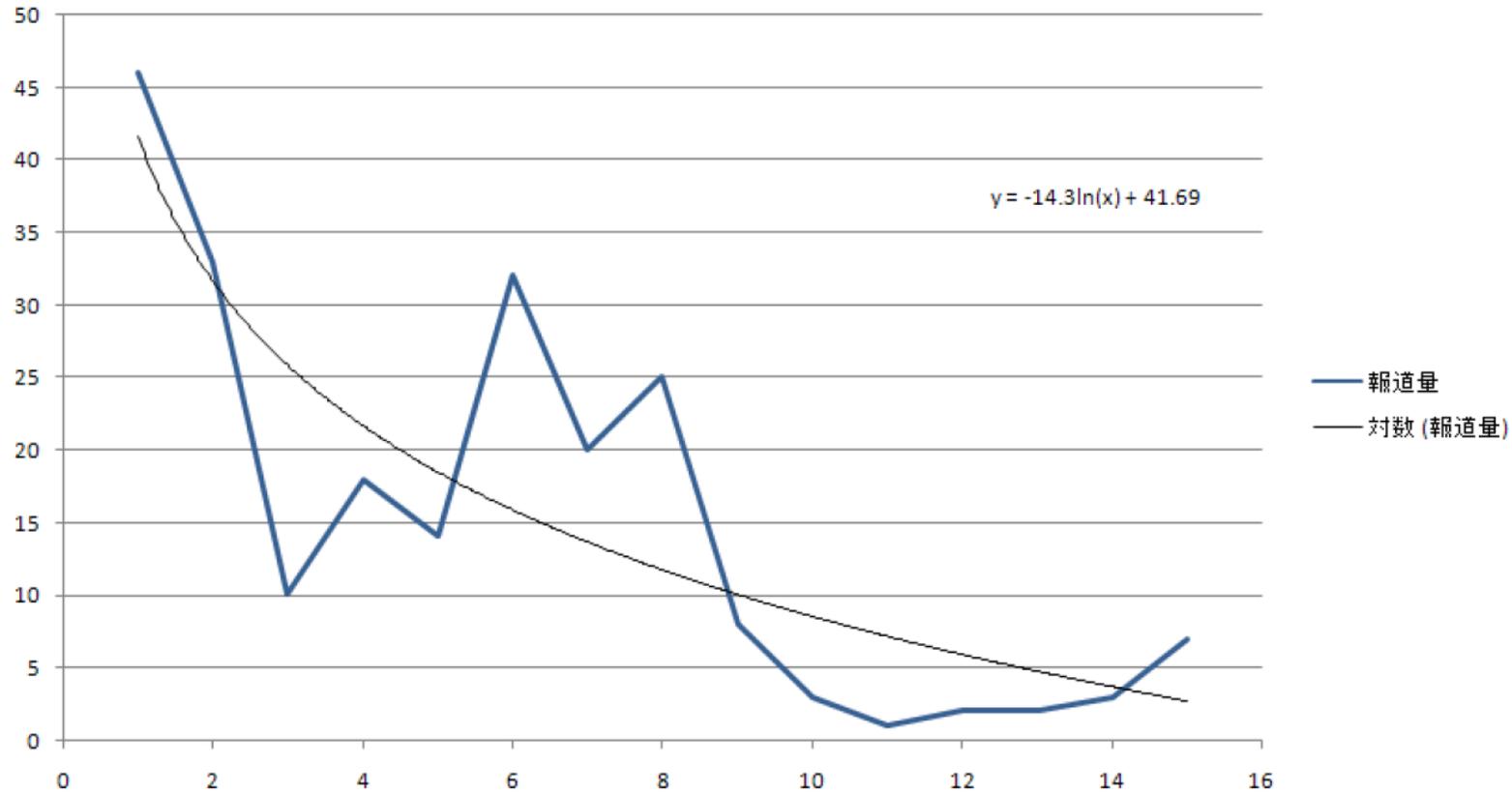
- 国民が刑事裁判に参加する制度。
- 2004年5月21日「裁判員の参加する刑事裁判に関する法律」が成立し、2009年5月21日から始まった。

# 関心度推移モデルのパラメータの決定



$y_1=20$ 万,  $y_2=11$ ,  $y_3=180$ 万-11とした時の同定結果  
( $b_1=0.001$ ,  $b_2=0.999$ ,  $b_3=0.015$ ,  $b_4=4.2$ )

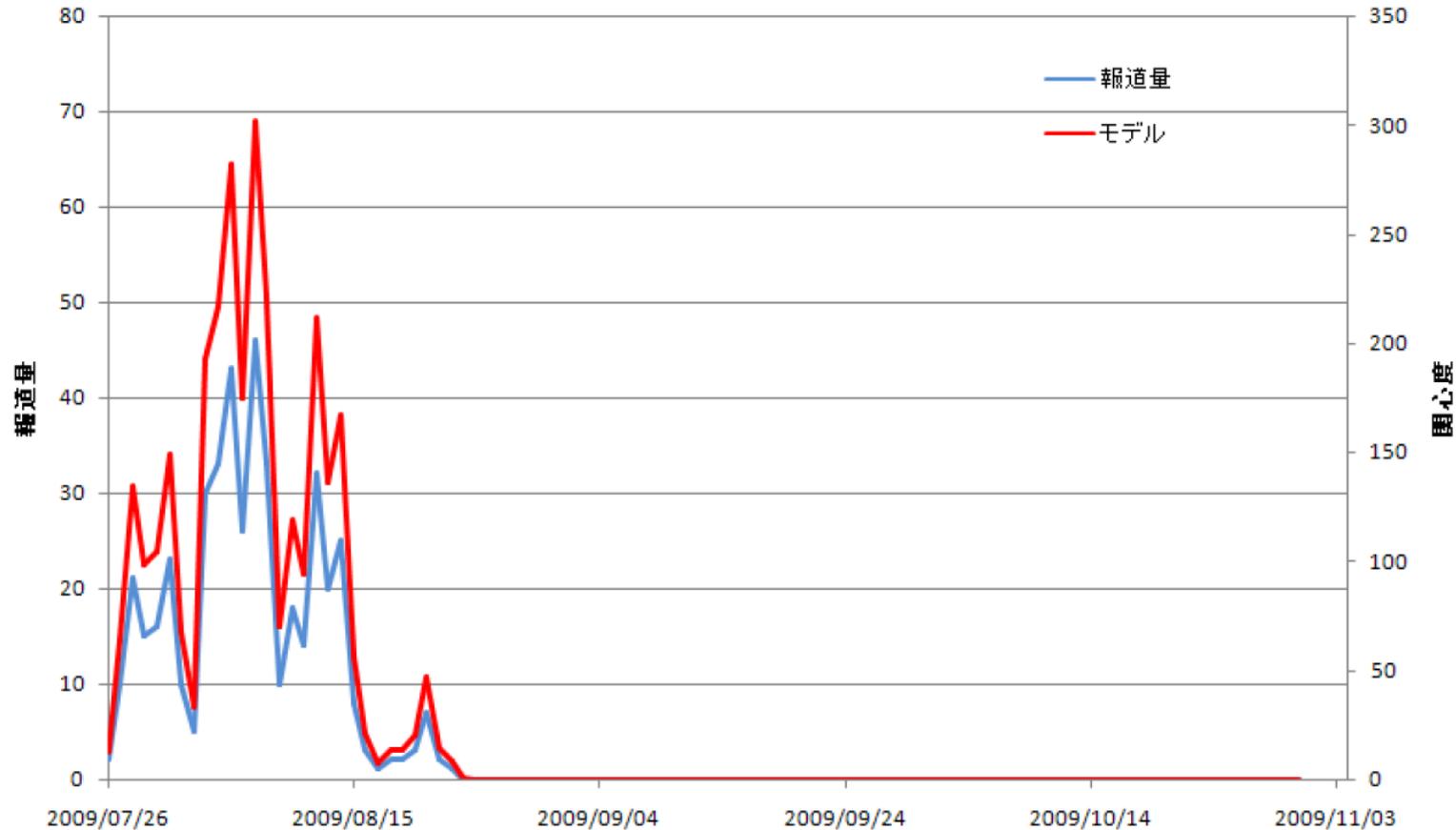
# 関心度推移の予測1



報道量の対数近似

# 関心度推移の予測1

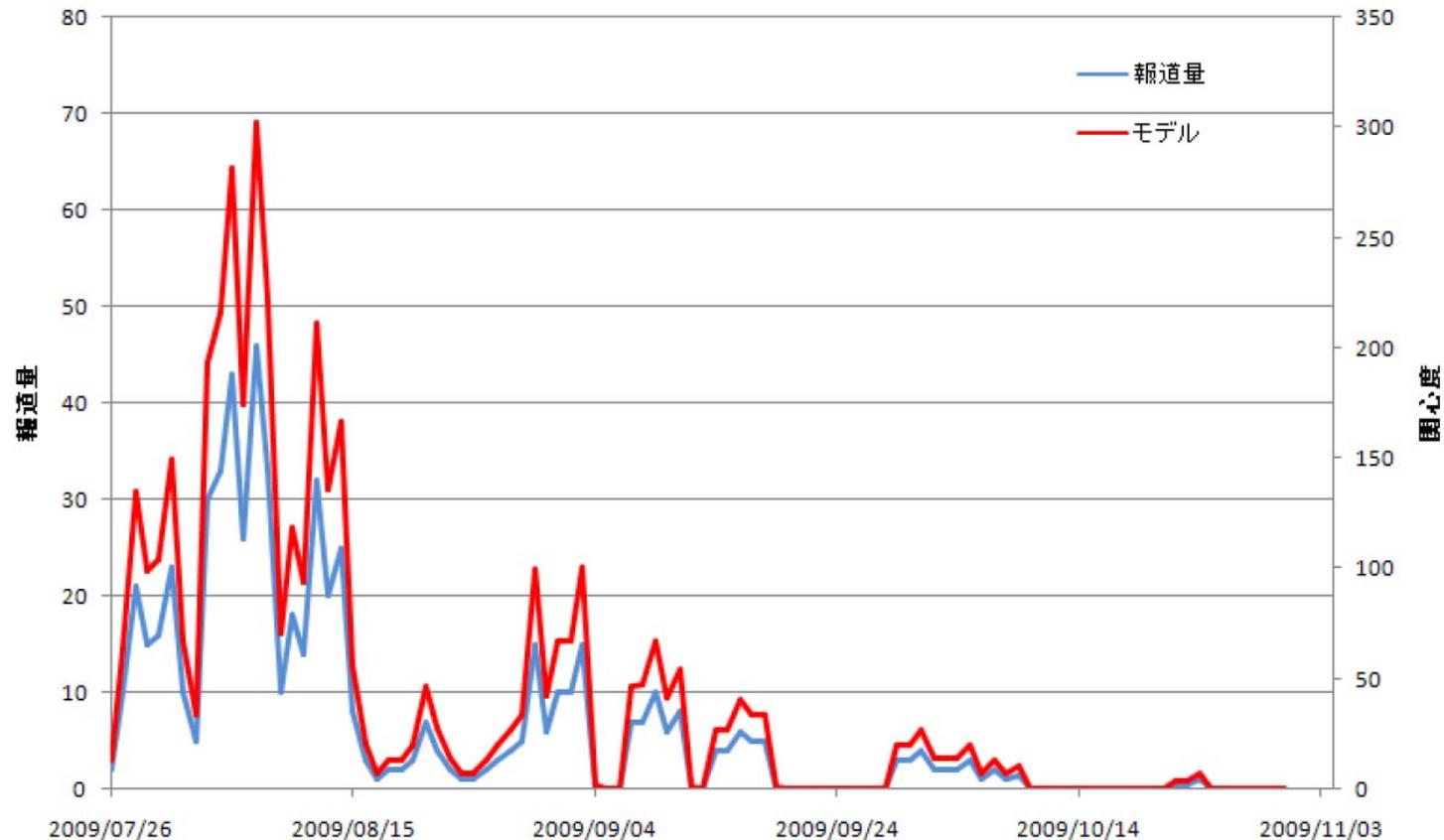
報道が収束した場合



- 関心度はすぐに収束。

# 関心度推移の予測2

今後の公判日程から報道予測を与えた場合

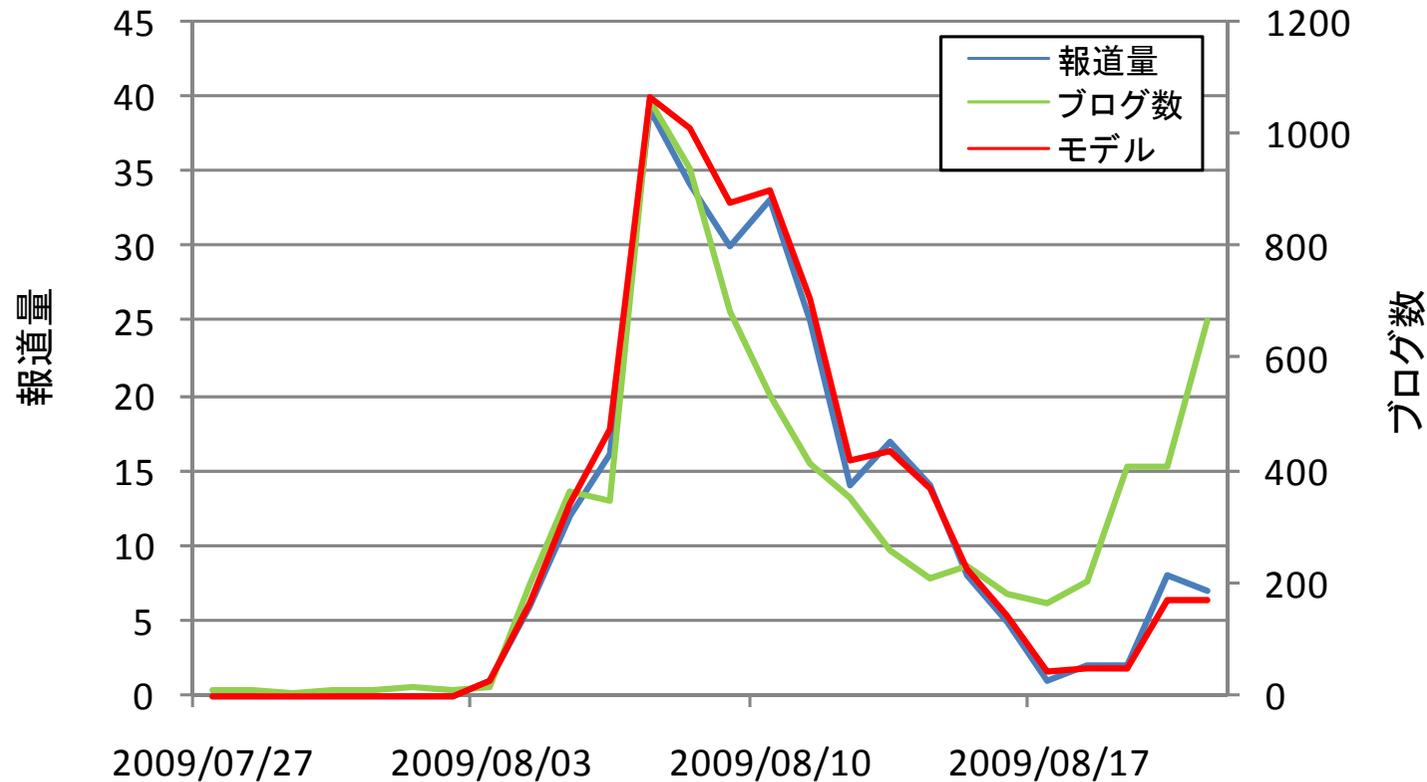


- 公判を重ねるごとに関心度の上昇は少なくなる。

## 6.3 キーワード「酒井法子」

- 2009年8月3日に女優の酒井法子が覚せい剤取締法違反容疑で逮捕された事件。
- 逮捕状が取られた時点で多くの報道があり、通常の芸能人の犯罪では異例の注目を集めた。

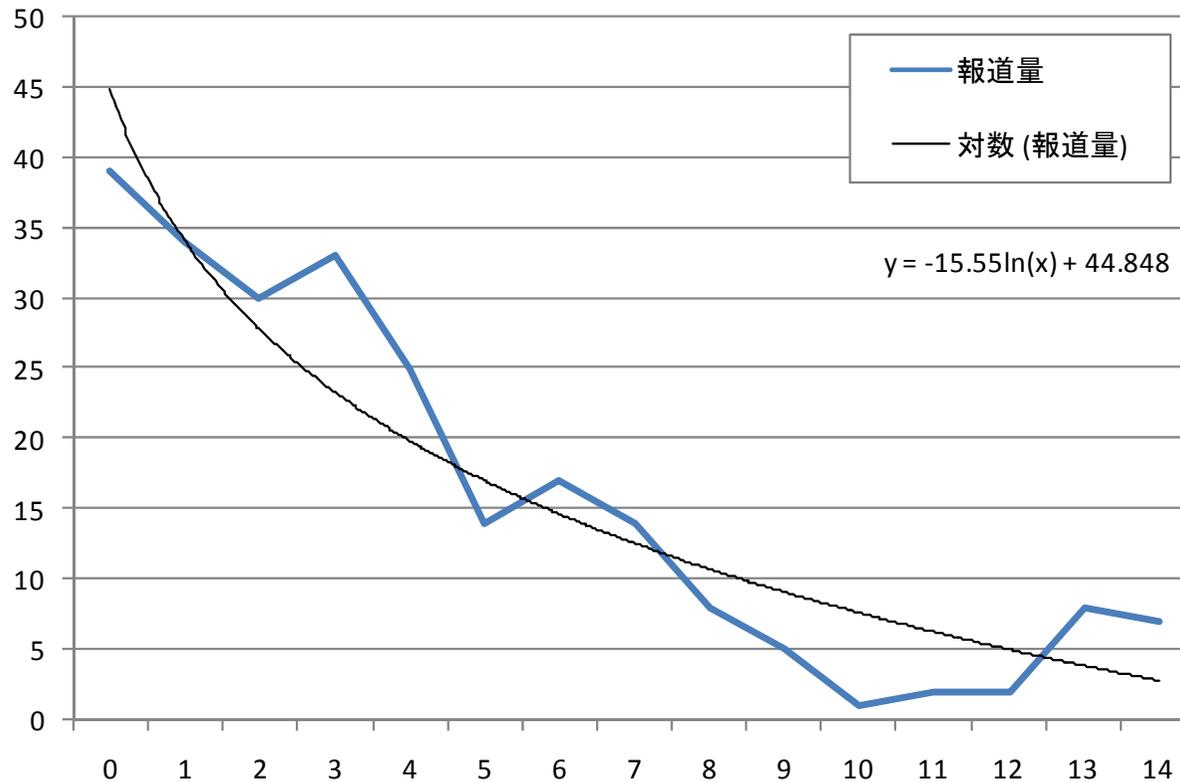
# 関心度推移モデルのパラメータの決定



同定結果 ( $b_1=0.032$ ,  $b_2=0.968$ ,  $b_3=0.001$ ,  $b_4=2$ )

- ピーク値とグラフの形は近い。

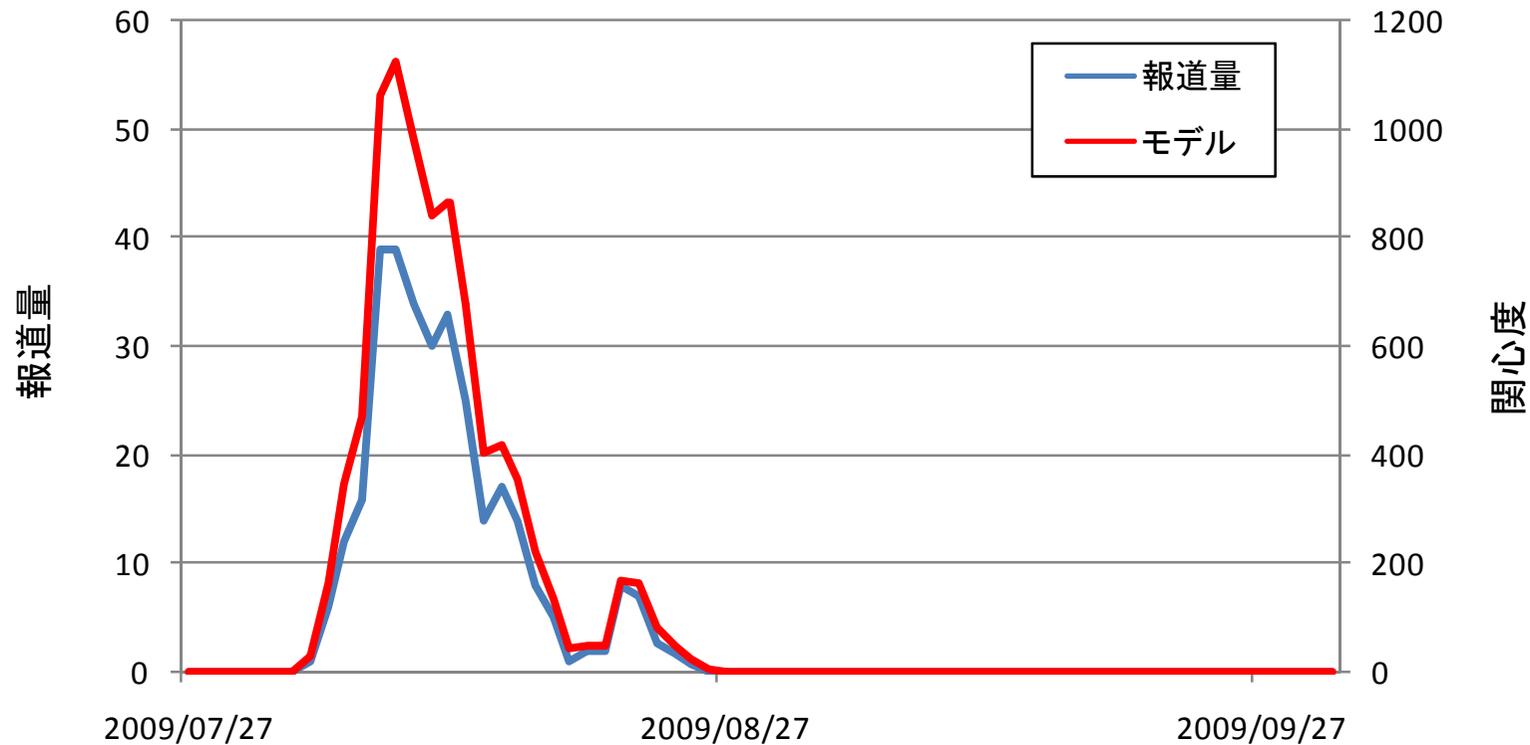
# 関心度推移の予測1



報道量の対数近似

# 関心度推移の予測1

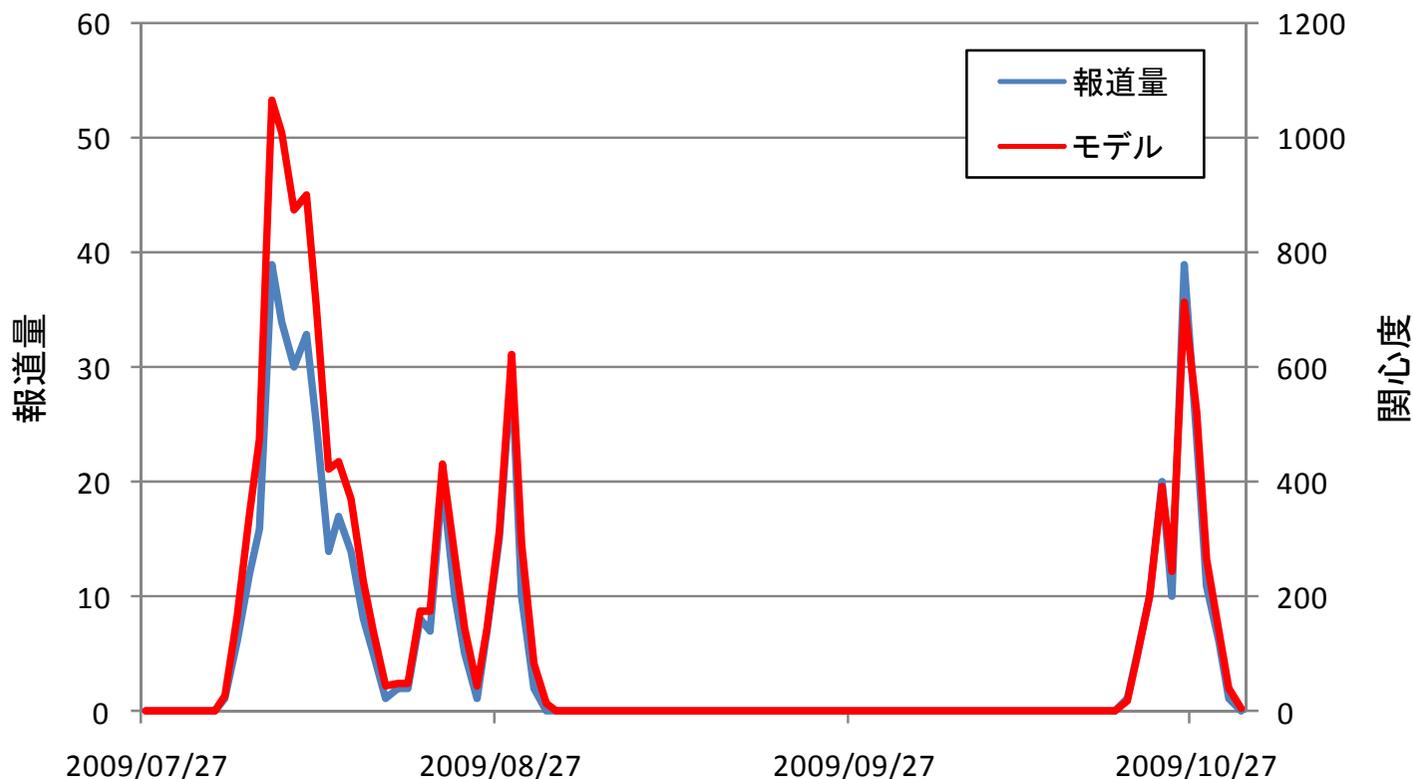
報道が収束した場合



- 関心度はすぐに収束。

# 関心度推移の予測2

今後の公判日程から報道予測を与えた場合



- 1度目のピークと同量の報道量を入力しても、関心度は1回目ほど増加しない。

# 7 おわりに

# 本研究のまとめ

- ブログの記事数を関心度推移の指標として用いた。
- 関心度の推移のパターンを分類した。
- 関心度推移モデルを構築した。
- 関心度推移モデルを用いて、現在進行中の事例に対する関心度推移の予測手法を提案した。
- 実際の事例を用いて、提案手法の検証を行った。

# 課題

- モデルのパラメータと関心度の分類の関係をより明確にする。
- モデルに報道の内容の影響などを考慮して、より適切な予測ができるように改良する。

# 参考文献

- [1] 福原知宏, 村山敏泰, 西田豊明, “ウェブログ記事を用いた社会事象に関する関心動向の分析”, 人工知能学会第6回AI若手の集い, 2005.
- [2] 中桐裕子・栗田治, “社会的なブームの微分方程式モデル”, 日本オペレーションズ・リサーチ学会和文論文誌, Vol.47, 83-105, 2004.
- [4] YAHOO!ブログ検索, <http://blogsearch.yahoo.co.jp/>
- [5] Google News Archive Search, <http://news.google.com/archivesearch/>