

感染症拡大予測モデルとその考察

廣瀬 英雄^{†1} 松隈 和広^{†1} 作村 建紀^{†1}

疫病感染拡大を予測するパンデミックシミュレーションはシナリオによるシミュレーションとして取り扱われてきたが、実際にパンデミックが起り始めると、観測データを使いながら将来どのようになるかを予測できるかということが重要になってくる。モデルの構造を仮定し、観測データを利用してモデルのパラメータを同定しながら予測を進める方法論は、データ同化とかグレーボックスとも呼ばれているが、パンデミック予測を行う上でもこのことが必要になってくる。ここでは、微分方程式による SIR モデルのパラメータを観測データから正確に推定する BBS 法を提案し、またこれまで実際に観測された、SARS、口蹄疫のデータを用いて予測を行った結果について議論する。また、これを truncated モデルによる予測結果とも比較する。比較の結果、SIR モデルは最悪のケースを早期に予測する可能性があるが、truncated モデルはかなり無力であることが分かった。

Infectious Disease Spread Prediction Models and Consideration

HIDEO HIROSE ^{†1}
and KAZUHIRO MATSUGUMA TATENORI SAKUMURA ^{†1,†1}

For infectious disease spread prediction models, pandemic simulations have been dealt with a kind of simulations by scenario. However, when a pandemic occurs, predicting the future using the observed data becomes crucial. The methodology that assumes the model structure and estimates the model parameters using the observed data, which is called the assimilation or the gray box, is also necessary in pandemic analysis. In this paper, we propose a method to estimate such parameters, called the BBS (best-backward solution) method, and discuss the prediction results for observed real cases such as the SARS and the FMD (foot-and-mouth disease). We compare the results with those using the truncated model. We have found that the SIR model provides the worst case predictions even in earlier stages of pandemics contrary to the truncated model.

1. はじめに

最近の Scientific American¹⁴⁾ に、今後 30 年間に深刻なパンデミックが人類を襲う脅威は 2 回の頻度、と書かれているように、世界のグローバル化が進む中で、未知の伝染病が世界中に蔓延するという不安感は高まってきている¹⁶⁾。例えば、2003 年の春から夏にかけては SARS が世界中を震撼させ、感染者数 8,096 人、死者数 774 人を出している²³⁾。2009 年、メキシコに端を発すると思われる新型インフルエンザは、南北アメリカ、ヨーロッパ、オセアニア、アジアへと感染領域が広範におよび、かなりの感染者が出ている。世界中での感染者数の推定は困難になっているが、日本では 2,000 万人以上が感染したものと考えられている^{17),18)}。重篤な症状にまで至る例が少ないので、通常のインフルエンザとして取り扱われることとなったが、免疫がなく感染率が高い場合の感染拡大の例として、多くの示唆を与えている。2010 年春には、日本（宮崎県）で口蹄疫（FMD, foot-and-mouth disease）の感染が確認され大変な勢いで拡大していった^{20),21)}。感染率が高いため、ネットワークの遮断に加えて、日本ではワクチン投与や殺処分という方法がとられている。FMD は、イギリスでの大きな被害だけでなく、今日世界中で大変な頻度で起こっており、経済的な波及を考えると深刻な問題である。このような時代、感染症拡大の予兆を早期にとらえたり、最終的な被害規模を早期に予測することは極めて重要になってきている。また、初期対策の結果により感染拡大の様相が大きく変化するとと思われることから、感染拡大する中で時々刻々の適切な予測も大きな意味を持つてくる¹⁹⁾。

疫病感染拡大のモデル（パンデミックモデル）としては、1) 局所的な閉鎖状況における感染拡大を数学的にモデル化した微分方程式モデル（SIR モデル）、2) エージェントに個々の感染率や行動パターンを与えて仮想都市を形成し感染伝搬をモデル化するマルチエージェントシステムモデル（MAS モデル）、3) 感染拡大の背後確率分布に不完全データをフィットさせる truncated モデルなどが提案されてきている。パンデミックシミュレーションはシナリオによるシミュレーションの性格を持つため、大規模な計算機資源を利用した MAS によるシミュレーションでいろいろな場面を想定した計算を行うことが多くなっているが、計算と（都市の）モデル化への対応にコストがかかるため、マイクロで精密な計算ができて融通性がない。一方、SIR モデルでは、均質な状態下での計算をマクロに行ってい

^{†1} 九州工業大学
Kyushu Institute of Technology

るため、地域の特性や感染者の個性を活かすににくいという面があるものの、計算は容易で、計算結果の妥当性も高いため、古典的なモデルでありながら現在でも多用されている。そこで、1)と2)の長所を併せ持ったハイブリッド型のモデル(MADEモデル)¹¹⁾も提案されている。しかし、これも現実のデータを取り扱うのではなくシミュレーションである。

さて、実際にパンデミックが起り始めると、観測データを使いながら将来どのようになるかを予測できるモデルが重要になってくる。シナリオによるシミュレーションは、モデルの構造が明らかになっていて、モデルに使われるパラメータも分かっている状態の計算の類に相当するためこれを直接には利用できない。モデルの構造を仮定し、観測データを利用してパラメータを同定しながら予測を進める方法論は、気象関係ではデータ同化であるとか、制御関係ではグレーボックスなどと呼ばれているが、パンデミック予測を行う上でもこのことが必要になってくる。しかしながら、これまでこのことについてはあまり議論されてきていない。ここでは、SIRのパラメータを観測データから精確に推定する方法について議論する。また、3)のtruncatedモデルによる方法についても、確率分布を仮定して観測データからパラメータを推定するという方法論と共通する面があるので、両者の結果を比較する。

2. SIRモデルによる解法

2.1 SIRの微分方程式モデル

マルサスの人口論に用いられる単純な微分方程式を拡張した、ケルマック・マッケンドリックモデルによるSIRモデル^{1),2),5),10)}は疫病伝搬の代表的なモデルである。このモデルは、 $S(t)$, $I(t)$, $R(t)$ (それぞれ時刻 t における感受性数 (susceptible), 感染数 (infected), 除外数 (回復または死亡, removed)) を、次の微分方程式で表現する。

$$\begin{aligned} S'(t) &= -\lambda S(t)I(t), \\ I'(t) &= \lambda S(t)I(t) - \gamma I(t), \\ R'(t) &= \gamma I(t). \end{aligned} \quad (1)$$

ここで、 λ は感染率、 γ は除外率を表している。感染数の累積総数は最終的には除外数で表される。初期値とパラメータを与えたとき、安定に解が求められる条件下では終息値 (収束する場合) まで決定論的にたどることができる。最近では、これに地域性を取り込んだ、より複雑なモデル^{3),6)}も提案されているが、ここでは最も単純な均一モデルとして取り扱う。

2.2 差分法によるパラメータ同定法

SIRモデルのパラメータは、通常、(1)式から自然に導かれる前進差分方程式から求めている¹²⁾。

$$\lambda(t) = \frac{S(t) - S(t+1)}{S(t)I(t)}, \quad \gamma(t) = \frac{R(t+1) - R(t)}{I(t)}. \quad (2)$$

日毎に値が求められるので、比較的安定した値の平均を固定パラメータとして採用している。前進差分の代わりに後退差分、

$$\lambda(t) = \frac{S(t) - S(t-1)}{S(t)I(t)}, \quad \gamma(t) = \frac{R(t) - R(t-1)}{I(t)}. \quad (3)$$

も用いることができる。

$S(t) + I(t) + R(t) = N$ (N は総数) なので、観測データの $S(t)$, $I(t)$, $R(t)$ のうちいずれか二つが分かればパラメータ λ , γ が計算できる。しかしながら、シミュレーションとは異なり、実際に観測されるのは $R(t)$ だけであることが多い。従って、既感染者が被感染者に影響を与えるまでの期間 (移行日数と呼ぶ) を仮定して $I(t)$ を与え計算を進めることが行われている。移行日数は incubation period とか latent period (いずれも潜伏期間) とか同一視されることもある。過去の例や観測結果から用いられるおおよその平均日数を用いることで対応するが、これを変動させてその影響を見て判断することもある。移行日数が指数分布に従っていると仮定すれば、この移行日数の逆数は除外率 γ に相当する。

2.3 BBS法によるパラメータ同定法

BBS法 (best-backward solution method) では次のようにパラメータを同定する。

- 1) 時刻 T におけるパラメータの初期値 $\lambda^{(0)}$, $\gamma^{(0)}$ を T までのデータを用いて差分法によって求める。
- 2) このパラメータを用いて、通常とは逆の方向に時刻 $T = t_n$ から時刻 t_1 に向かって (1) を解いていく。
- 3) 時刻 T までの観測データ $R(t_1), R(t_2), \dots, R(t_n)$ と、2) で得られた解 $\hat{R}(t_1), \hat{R}(t_2), \dots, \hat{R}(t_n)$ との最小 2 乗誤差

$$S_0 = \sum_j (\hat{R}(t_j) - R(t_j))^2 \quad (4)$$

が最小になるようなパラメータ $\lambda^{(1)}$, $\gamma^{(1)}$ を求める。最適法の計算には、シンプレックス法¹⁵⁾を用いる。

- 4) $\lambda^{(1)} \rightarrow \lambda^{(0)}$, $\gamma^{(1)} \rightarrow \gamma^{(0)}$ として、2), 3) を $\lambda^{(1)}$, $\gamma^{(1)}$ が収束するまで繰り返す。

つまり、BBS法は微分方程式の解に最小 2 乗法を取り入れたパラメータ決定法である。最小 2 乗法では通常解の形が陽に与えられているときに用いられる関数フィッティングであるが、ここでは微分方程式の構造で与えられているときに適用する方法である。微分方程式

の出発点に T を与えている理由は次のとおりである。

1) 初期値問題の微分方程式の性格から、 T に近いところでは観測値と予測値とがより一致しやすいように設定される傾向がある。

2) T では $R(T)$ と $\hat{R}(T)$ とを一致させることで、統計的には truncated モデルと同様の扱いになる。

つまり、微分方程式の出発点を T より前のどの点に設置させるよりも T に置いた方が将来予測の面からは有効であると考えられるからである。図 1 に、香港で起こった SARS の事例²³⁾ について、差分法と BBS 法で決定したパラメータによる累積感染者数の予測終息値の一例を示す。図には観測値も併記している。予測値は観測日までの累積感染者数データを用いた結果である。BBS 法の終息予測値は差分法に比べてパンデミック発生初期でも良

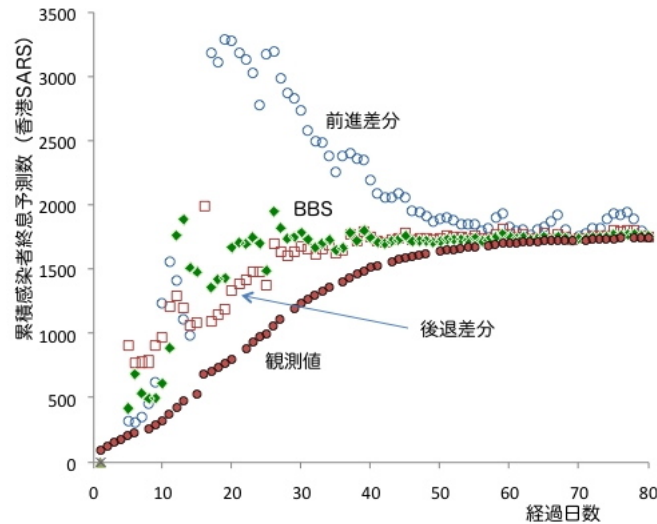


図 1 差分法と BBS 法で決定したパラメータによる累積感染者数終息値例 (SARS)

い値を示していることが分かる。数値計算にはルンゲ・クッタ法を用いている。

さて、 T から遠ざかるにつれて影響力を小さくするために、 S_0 による最小 2 乗法の拡張として、重み付き最小 2 乗法を次のように定義して、この方法でもパラメータを求めることができる。

$$S_p = \sum_j w_j (\hat{R}(t_j) - R(t_j))^2, \quad w_j = \frac{(j/n)^p}{\sum_{j=1}^n (j/n)^p}, \quad p \geq 1. \quad (5)$$

ただし、 $T = t_n$ 。ここでは、 $p = 1, 2$ の場合について求めている。

SIR には感受性数 (susceptible) の $S(0)$ が必要になるが、SARS などでもこの値をどのようにとるかで予測結果が違ってくる。香港の総人口を $S(0)$ としたり、限定的な地域を $S(0)$ としたりしながらいくつかの計算を行うことが考えられる。 $S(0)$ の変化によって R の終息値はそれほど変わらないと思われるが、パラメータ λ, γ は大きく変化する。BBS は拡張性に富んでいるので、パラメータ λ, γ の推定だけでなく、微分方程式の初期値 $S(0)$ や、移行日数も推定できる可能性を持つため、BBS は有用な方法と考える。

3. truncated model

打ち切り時間を T とし、背後にある確率分布の分布関数を $F(t)$ と仮定する。感染数は日毎に報告されるため、 T_{i-1} (日) から T_i (日までの) 区間をグループデータとして取り扱い、その区間での観測数を n_i 、グループの数を k とすれば、尤度関数は次で表すことができる。

$$L_T(\theta) = \prod_{i=1}^k \left\{ \frac{F(T_i) - F(T_{i-1})}{F(T; \theta)} \right\}^{n_i}. \quad (6)$$

最尤推定法により分布のパラメータ $\hat{\theta}$ を推定できれば、感染終息数 \hat{N}_T は、

$$\hat{N}_T = r / F(T; \hat{\theta}), \quad r = \sum_{i=1}^k n_i, \quad (7)$$

から求めることができる。これを truncated モデルと言う。

SARS が発生したときには、感染者数だけでなく回復者と死者との数も報告されており、すべての情報を使う truncated モデル⁸⁾ では更に信頼度の高い確率分布の推定が可能になる。また、truncated モデルでは終息に近くなるにつれて累積感染者数の誤差が 0 になるという不合理が見られるが、truncated モデルでは累積感染者数と非感染者数との 2 項確率から求められる合理的な誤差に収束する誤差も求められている。しかし、推定値そのものはどちらのモデルで解いても同じになるため⁸⁾、ここでは truncated モデルで求めた推定値の結果と比較している。

感染拡大をモデル化する背後分布を、ここでは、3パラメータロジスティック分布に従うと仮定する。この分布は生物成長曲線を表す微分方程式に対応しており、また、ワイブル、ガンマ、対数正規などの、他のよく知られた確率分布よりも観測値に良くフィットすることが知られている⁹⁾。

$$F(t; \mu, \sigma, \beta) = \frac{1}{\{1 + \exp(-(t - \mu)/\sigma)\}^\beta}, \quad (-\infty < t < \infty; -\infty < \mu < \infty; \sigma, \beta > 0).$$

4. 計算例

4.1 口蹄疫 (FMD) の場合

2010年春、宮崎県で口蹄疫 (FMD) の感染が確認され大変な勢いで拡大していった。感染率が高かったため、初期発見、初動体制までの間にも相当の拡大が起こっていたと考えられる。ネットワークの遮断に加えて、ワクチン投与に殺処分という方法により、ようやく沈静化したようにも思われるが、世界の状況を見ると、いつ再発生して感染が拡大するか全く分からない。宮崎県では、累計で約21万頭の家畜が感染およびその恐れありとして殺処分された。この被害の大きさは2001年にイギリスで起こった口蹄疫感染拡大⁶⁾の次に匹敵するものと言われている。

今回用いたデータは農林水産省が提供しているもの²⁴⁾を用いた。感染は牛や豚など偶蹄目の家畜動物に広がるが、ここでは予測方法の比較を主眼にするため対象を牛のみとしている。感染個体数は正確には把握されている訳ではない。家畜を育てる場所毎にサンプル調査を行って、感染が確認されたらその場所の家畜全頭が感染可能性があるとして殺処分対象となっている。そこで、ここでは殺処分対象数を感染個体データとみなした。感染した可能性のある家畜数を推定して予測を行うことも考えられるが、殺処分までには時間がかかりその間にも状況が変化していくと考えたからである。

推定は、宮崎県で最初に FMD 発症が確認された 2010 年 4 月 20 日から最後の埋設処理日の 7 月 5 日までとした。4 月 20 日から発生農場を中心に半径 10km 内の移動制限および 10-20km 以内の搬出制限、5 月 22 日から 10km 内の家畜に対しワクチンの投与（効果が出るまで牛は 1 週間、豚は 2 週間ほどかかる）が行われている。ここでは、SIR で計算を行う際に、4 月 20 日以前は宮崎県全体に広がる可能性があるとして、感染する可能性のある家畜総数を宮崎県家畜牛の全頭数（約 30 万頭）に設定した場合と、4 月 20 日以降は 10km 移動制限内で広がる可能性があるとして、それを宮崎県家畜牛の全頭数（約 6.5 万頭）に設定した場合とで計算した。truncated モデルにはこの設定は不要である。

4.1.1 SIR による予測結果

図 2 に、感染可能性のある牛の頭数上限を 30 万頭、移行日数を 7 日と仮定したときの、打ち切り日までのデータを用いて推定した感染数終息予測値を打ち切り日ごとに示す。図では、重みをつけない BBS 法 (BBS)、 $p = 1$ の重みを付けた場合、 $p = 2$ の重みを付けた場合の予測結果を、打ち切り日までの累積感染数観測結果と併記して示している。重みによる効果はそれほどないことが示されている。このことは香港での SARS の場合でも同様であった。図 3 には、牛の頭数上限を 6.5 万頭と変化させたときのそれらを示している。

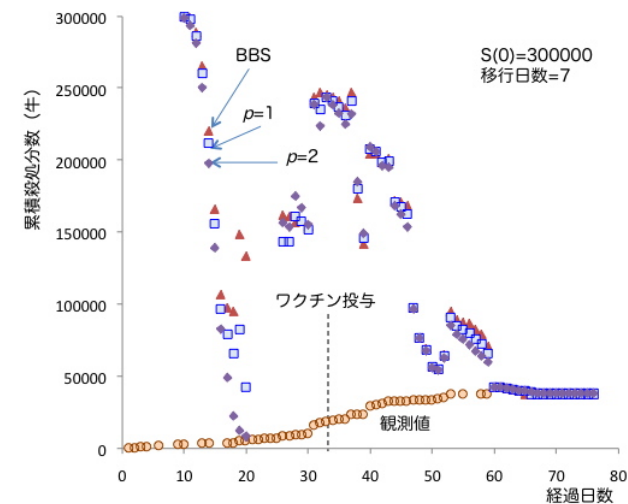


図 2 FMD 感染拡大予測 (SIR, $S(0) = 300,000$)

図 2 は、もしも 10km 制限区域からウイルスが外に出て感染拡大を起こし、宮崎県以外には広がらない場合には、宮崎県の家畜牛の少なくとも 70% 程度までは感染した可能性を示唆している。ワクチン投与して一定期間経った後にはこの可能性は急激に減り、終息に向っていることが分かる。図 3 からは、10km 制限区域からウイルスが外に出て感染拡大を起こさなくても、感染は 10km 制限区域内の家畜牛のほぼ 90% にまで達する危険性があったことを示している。特徴的なのは、どちらの図からも、感染初期から危険側の終息値がある程度予測されていたことが分かる。これらのことから、今回の 10km 制限の対策はかなり

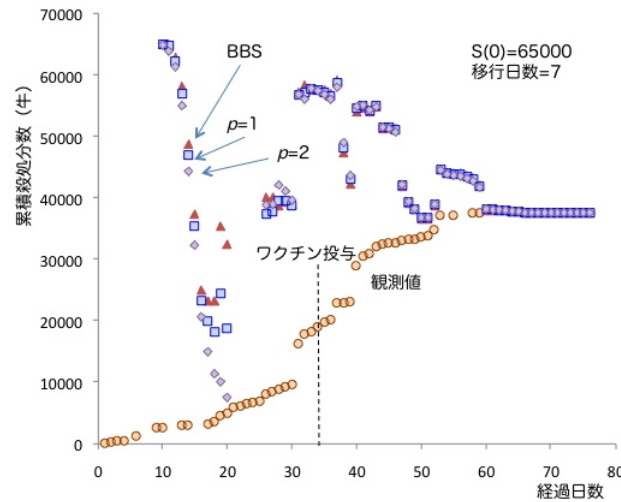


図3 FMD 感染拡大予測 (SIR, $S(0) = 65,000$)

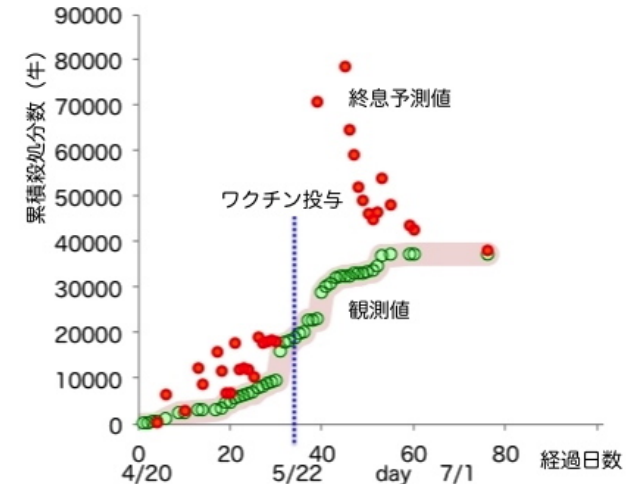


図4 FMD 感染拡大終息予測値 (truncated)

効果的に働いていたと考えることができる。また、ワクチン投与の時期は効果的であった³⁾ことも示される。

4.1.2 truncated モデルによる予測結果

図4に、打ち切り時刻 T ごとに推定された終息数 \hat{N}_T と、観測データ $N_j = \sum_{i=1}^j n_i$ との比較を示す。初期段階 ($T = 30$ あたりまで) は終息数は小さめに推定され、中盤 $T = 30 - 50$ では変動が大きくなり、中盤以降 ($T = 50$ 以降) は落ち着いている。つまり、観測初期段階では終息数は実際よりも低めに推定される傾向があり、終息段階に近くなって正確な終息数に近づいてきていることが分かる。この初期段階の振る舞いは SIR モデルによるそれとは大きく異なっている。こういった傾向は FMD の場合に限らず、香港の SARS の例や、2009 年に世界中を襲った新型インフルエンザ (A(H1N1)) の場合でも全く同様であった。

5. ま と め

本論文では、微分方程式による SIR モデルのパラメータを観測データから正確に推定する BBS 法を提案し、またこれまで実際に観測された、SARS、口蹄疫のデータを用いて終息状況での感染数の予測を行った結果を示した。疫病感染拡大予測モデルとしては SIR モ

デルの他に統計的な truncated モデルを用いることができるので、この結果とも比較した。比較の結果、SIR モデルは古典的で単純であるにもかかわらず、感染拡大の初期段階での将来予測に有効であることが示唆された。一方で、統計的な truncated モデルでは初期データだけでは危険側の予測ができないことが分かった。

宮崎県で発生した FMD の場合について、牛の感染可能性のある頭数上限を宮崎県内の全頭数 30 万頭と仮定したときと、それを 10km 制限区域内での全頭数 6.5 万頭と仮定したときについて、それぞれ SIR モデルを用いて FMD 感染拡大予測を行った。その結果、1) もしも 10km 制限区域からウイルスが外に出て感染拡大を起こし、宮崎県以外には広がらない場合には、宮崎県の家畜牛の少なくとも 70% 程度までは感染した可能性があった、2) 10km 制限区域からウイルスが外に出て感染拡大を起こさなくても、感染は 10km 制限区域内の家畜牛のほぼ 90% にまで達する危険性があった、3) ワクチン投与して一定期間経った後には感染は急激に縮小し、終息に向っていることが分かった。従って、今回の 10km 制限の対策はかなり効果的に働いていたと考えることができ、また、ワクチン投与の時期は効果的であったことが分かる。

最近では、いち早く予兆をとらえるため、Google やソーシャルネットを使って疫病の出

現を見いだそうという報告も出ているので^{4),13)}、今後はこの方面の検討も加えたい。推定値の信頼度についてはこれまである程度の知見は得ているが、今回紙面の制約上それには言及しなかった。今後、確率微分方程式を用いたり²²⁾、さまざまな方法を用いながら、そのことについても報告する予定である。

参 考 文 献

- 1) R. Anderson and R. May, *Infectious diseases of humans: Dynamics and control*, Oxford University Press, 1991.
- 2) F. Brauer, P. van den Driessche and J. Wu (ed.), *Mathematical Epidemiology*, Lecture Notes in Mathematics, Springer, 2008.
- 3) G. Chowell, A.L. Rivas, N.W. Hengartner, J.M. Hyman, and C. Castillo-Chavez, "Critical response to post-outbreak vaccination against foot-and-mouth disease, Mathematical studies on human disease dynamics: emerging paradigms and challenges," *AMS Contemporary Mathematics* 47, 73-87, 2006.
- 4) N.A. Christakis and J.H. Fowler, "Social Network Sensors for Early Detection of Contagious Outbreaks," *PLoS ONE* Vol.5, No.9 (2010)
- 5) O. Diekmann and J.A.P. Heesterbeek *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, New York: Wiley, 2000.
- 6) N.M. Ferguson, C.A. Donnelly, R.M. Anderson, "The Foot-and-Mouth Epidemic in Great Britain: Pattern of Spread and Impact of Interventions," *Science*, 292 1155-1160, 11 May 2001.
- 7)
- 8) H. Hirose, "The truncated model and its Applications to lifetime analysis: unified censored and truncated models," *IEEE Transactions on Reliability*, Vol.54, pp.11-21 (2005).
- 9) H. Hirose, "The mixed truncated model with applications to SARS," *Mathematics and Computers in Simulation*, vol. 74, pp. 443-453, 2007.
- 10) W.O. Kermack and A.G. McKendrick, "Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity," *Proceedings of the Royal Society*, vol. 141A, pp. 94-122. 1933.
- 11) Y. Toyosaka and H. Hirose, "Pandemic simulations by MADE: a combination of multi-agent and differential equations," *The 2009 International Conference on Parallel and Distributed Processing Techniques and Applications (PDPTA2009)*, pp. 331-335, 2009.
- 12) Y. Toyosaka and H. Hirose, "Pandemic Simulations by MADE: the Hybrid Method of Multi-Agent and Differential Equations," *Asia Simulation Conference 2009 (JSST2009)*, October, 2009.
- 13) J. Ginsberg1, M.H. Mohebbi1, R.S. Patel1, L. Brammer, M.S. Smolinski1 & L. Brilliant1, "Detecting influenza epidemics using search engine query data," *Nature* 457, 1012-1014, 2009.
- 14) , J. Matson and J Pavlus, "Laying Odds on the Apocalypse: Experts Assess Doomsday," *Scientific American* September 1, 2010.
- 15) J.A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, Vol.7, pp.308-313, 1965.
- 16) 廣瀬, 豊坂, " パンデミックシミュレーション, " 日本信頼性学会誌, Vol.33, No.6, pp.270-277, 2009.
- 17) 廣瀬: A(H1N1) インフルエンザパンデミックシミュレーション, 日本計算機統計学会第 23 回シンポジウム論文集, pp.127-130, 2010 .
- 18) 廣瀬, " パンデミックシミュレーションの信頼性 " , 2009 年度第 1 回日本信頼性学会フォーラム「医療の信頼性・安全性」 - 安心につながる医療に向けて - 2010.2.19.
- 19) 松隈, 廣瀬: インフルエンザ感染拡大モデルとその周辺, 日本信頼性学会第 18 回春季信頼性シンポジウム報文集, pp.37-40, 2010 .
- 20) 作村, 廣瀬, " truncated モデルを用いた口蹄疫拡大予測とパラメータ推定 " , 電気関連学会九州, 02-2P-09 2010.
- 21) 河野, 廣瀬, " SIR モデルを用いた口蹄疫拡大予測とパラメータ決定法 " , 電気関連学会九州, 02-2P-10 2010.
- 22) 牧, 廣瀬, " 確率微分方程式を用いたパンデミック解析の基礎モデル " , 電気関連学会九州, 02-2P-11 2010.
- 23) WHO, <http://www.who.int/csr/sars/country/en/>
- 24) MAFF, http://www.maff.go.jp/j/syouan/douei/katiku_yobo/k.fmd/