

COVID-19新規感染者数の時系列分析：続報

Time Series Analysis on the Number of COVID-19 Newly Infected People : a follow-up study

小松 隆行*

Takayuki Komatsu

概要

本稿では、COVID-19（新型コロナウイルス）に関するオープンデータを、データサイエンスにおける時系列分析的な視点から分析することを試みる。なお、前回の報告以降に発生した感染に関する新データと、それを加えた約1年間分の全データについて分析を行う続報となっている。対象とするデータは、日本の都道府県毎に報告されている新規感染者数の日次データ、及び世界保健機関（WHO）が報告している世界各国の新規感染者数の日次データであり、これらを時系列データと捉えて、これらの移動平均とファイナンス分野でのテクニカル分析を応用した手法を用いた分析結果について報告する。

1. はじめに

2019 年末から COVID-19（新型コロナウイルス）のパンデミックは現在も世界中で継続している新規感染者数の急激な増加と減少により形成される波は、2022 年にも世界各国で発生している。この波は、変異ウィルスの出現や、ワクチン接種、各国のコロナ対策（ゼロコロナ対策、With コロナ対策など）などにも起因すると考えられるが、新規感染者が少ない時期からの増加や、波のピークからの減少期にリバウンドして次の波となることもある。これらの波を含めた新規感染者数の推移は、世界中の様々な機関がインターネット上に公開している感染状況のオープンデータ⁽¹⁾を可視化することで容易に確認できる。日本では厚生労働省や自治体などが、主にCSV形式で日次データを公開してダウンロード可能としており、利用者が独自に可視化と分析ができる⁽²⁾⁽³⁾⁽⁴⁾。様々なサイトや報道などでは、単純なグラフ等での可視化が多く、時系列分析を行ったデータのグラフは未だに殆ど見受けられない。しかしながら、従来の統計学的な手法でも感染状況の様相の把握や予測も可能と考えられる。

このようなことから本報告では、2021 年 6 月までのデータを用いた前回の分析報告と同様にして、日本と世界のオープンデータについて新たに得られた 2023 年 1 月までのデータを加え、時系列分析

における基本的な手法である移動平均と、それを基礎とする様々な分析方法を用いて考察を行う。用いる分析手法は、ファイナンス分野でテクニカル分析⁽⁵⁾として利用されている n 日移動平均、ボリンジャーバンド⁽⁶⁾、そのバンド幅、バンド内の相対的位置をプロットした %b チャート、MACD を COVID-19 の時系列データに応用したもの⁽⁷⁾である。以下では、北海道のデータ、東京都のデータ、諸外国のデータに関してこれらを用いて順に分析を行う。

2. 北海道の新規感染者数に関する分析

まず北海道の新規感染者数のオープンデータ⁽³⁾を分析してみる。今回は、2020/1/28～2023/3/6 までの日次データを使用した。

2.1 移動平均を基礎とした指標による傾向分析

日本の自治体が公開している新規感染者数の日次データは、曜日依存した変化や検査状況などの要因で、連続する時系列データの増減が大きく激しい変化になる傾向がある。曜日毎の前週との比較（倍率など）を使うことが多い。この変動の影響を緩和するため、時系列データ分析の手法である移動平均を使う。一般的には、ある一定期間（ n 日とする、例えば 7 日： n は自然数）の各日のデータの総和を、この期間 n で割った値である単純移動平均（SMA, Simple Moving Average）が用いられる。 n 日移動平

均（ n は自然数）と表現すると、報道などでも多用される 7 日移動平均（ $n=7$ ）をベースにして、その倍数である、 $n=14, 28, 58$ などと、 n の値を大きくして算出した値をプロットしたものが図1の最上段のグラフである。図1の2段目のグラフは、直前のデータほど重みを大きくして、直近のデータに敏感に反応する指数平滑移動平均（EMA）である。ファイナンス分野では、株価や為替の予測でのテクニカル分析で多用されているが、他のいくつかのテクニカル分析を COVID-19 の新規感染者数のデータに適用した結果も3段目以降に示す。上から順に、①日次データとその n 日移動平均（ nMA , $n=7, 14, 28, 56$ ）、②指数平滑移動平均（ $nEMA$, n は①と同様）、③MACDc と MACDc シグナル⁽⁷⁾、④7 日平均基準のボリンジャーバンド⁽⁶⁾、⑤ボリンジャーバンドを基にした %b チャートとその n 日平均（ $n=7, 14$ ）、⑥ボリンジャー

バンド幅とその nMA となっている。これらの指標は、前回の報告である文献(8)における 2021/12/29 までのデータ分析において使用されている。今回使用したデータは、第5波の収束後から第8波までのデータである。各波のピークでの最大感染者は数倍に激増している。短期移動平均線が長期の移動平均線をダウクロスするより早く見える MACDc (③) のダウクロスは、感染拡大期のピークアウトを捉えているように見えるが、変動幅が大きく明確ではない。ボリンジャーバンド幅(⑥)は、感染の急拡大縮小に伴い同様に増加減少している。一方で、感染拡大の兆候（起点）や下げ止まりのサインとしては、文献(8)と同様に⑤の %b チャートの移動平均線の動きや状況が、より早期のサインである考えられる。次節では、このようなサインの可能性について改めて詳しく考察を試みる。

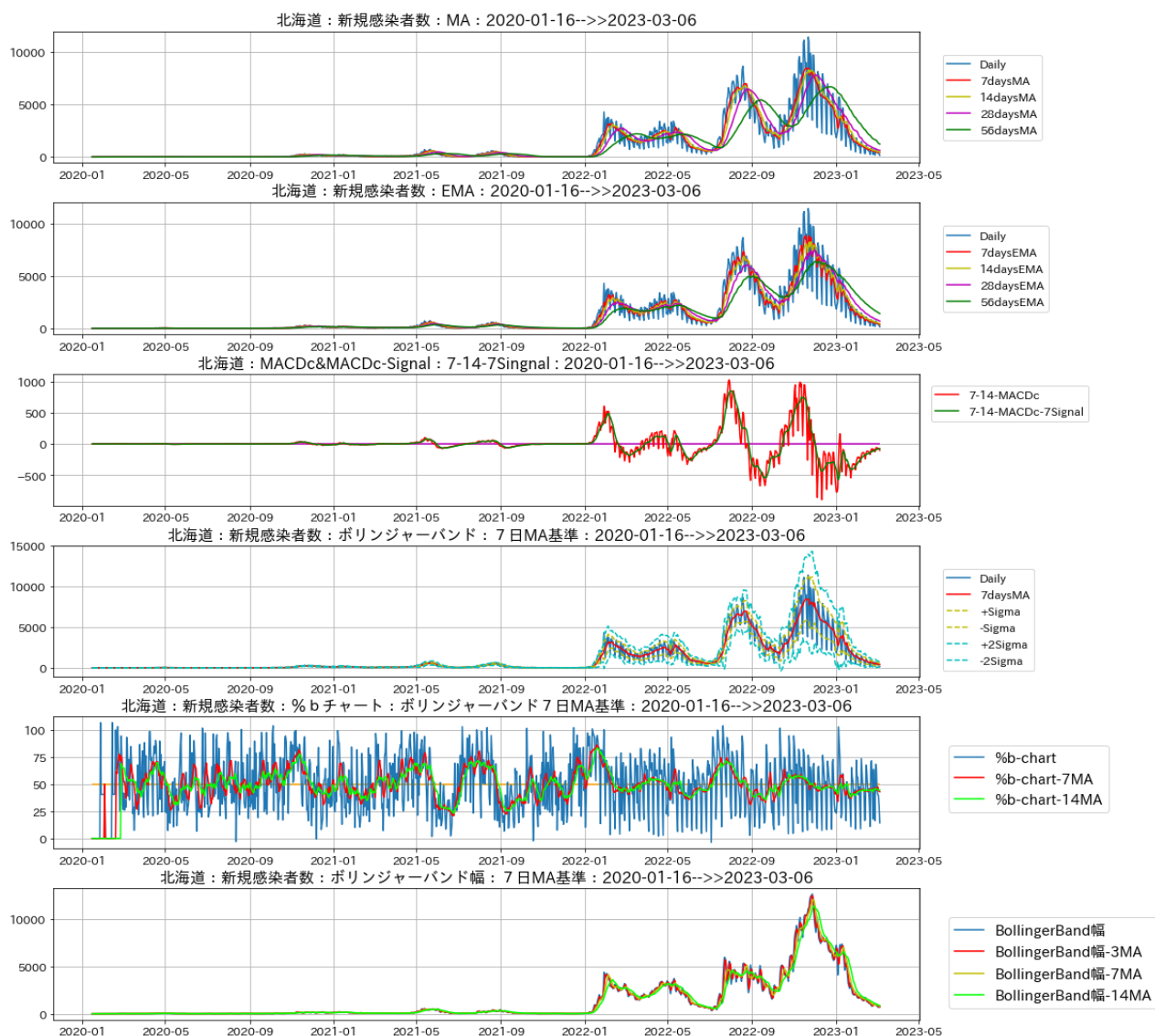


図1 北海道における新規感染者数の移動平均線为基础とした指標のグラフ（2020/1/16～2023/3/6）

2.2 感染拡大期毎の主要グラフを複数用いた考察

ここでは、第6波から第8波までを含む2021年12月1日から2023年3月6日までの北海道のデータの可視化で(図2)を用いた詳しい考察について述べる。図2の3番目のグラフは14日EMAを基準としたMACDcであるが、このダウングロス(MACDcがそのシグナルを上から下へ突き抜ける)は感染拡大期のピークアウト(2022年2月、8月、11月)を捉えていると言える。7日EMA基準を使用すると「揺れ」が大きく明確ではなかった(図1)。一方、6月や9月のMACDcのアップクロス(ダウングロスの逆)は、下げ止まりの兆候を示している。6番目のグラフでは、ボリンジャーバンド幅の増減が、急拡大縮小に追従したまたそれ以外の期間である安定期では小さい。ボリンジャーバンド幅のMAのダウングロスがピークアウト付近を捉えているが、サインとして不安定で明確ではない。このボリンジャーバンドに

おいて、ある時点の日次データの標準偏差を基準にしたバンド内での位置をプロットしたものが5番目のグラフの%bチャートである。その7日移動平均と14日移動平均もプロットした。この2つが各波の各拡大開始点付近で7MAと14MAが50%線をアップクロスし、各波のピーク付近まで50%~80%間に停留していることが多い。その後ピークアウトの少し前から減少し始めダウングロスし、縮小期後半で50%線をダウングロスし、安定期では20%~50%の間で停留している。%bチャートのアップクロス直後にMACDcがアップクロスし、%bチャートMAが50%線をアップクロスするサインが急拡大の開始点の強いサインの可能性がある。このようにして、複数のグラフの変化や動きを捉えることが、新規感染者の増減の様相を把握し、将来の感染者数の予測の一助になると考えられる。

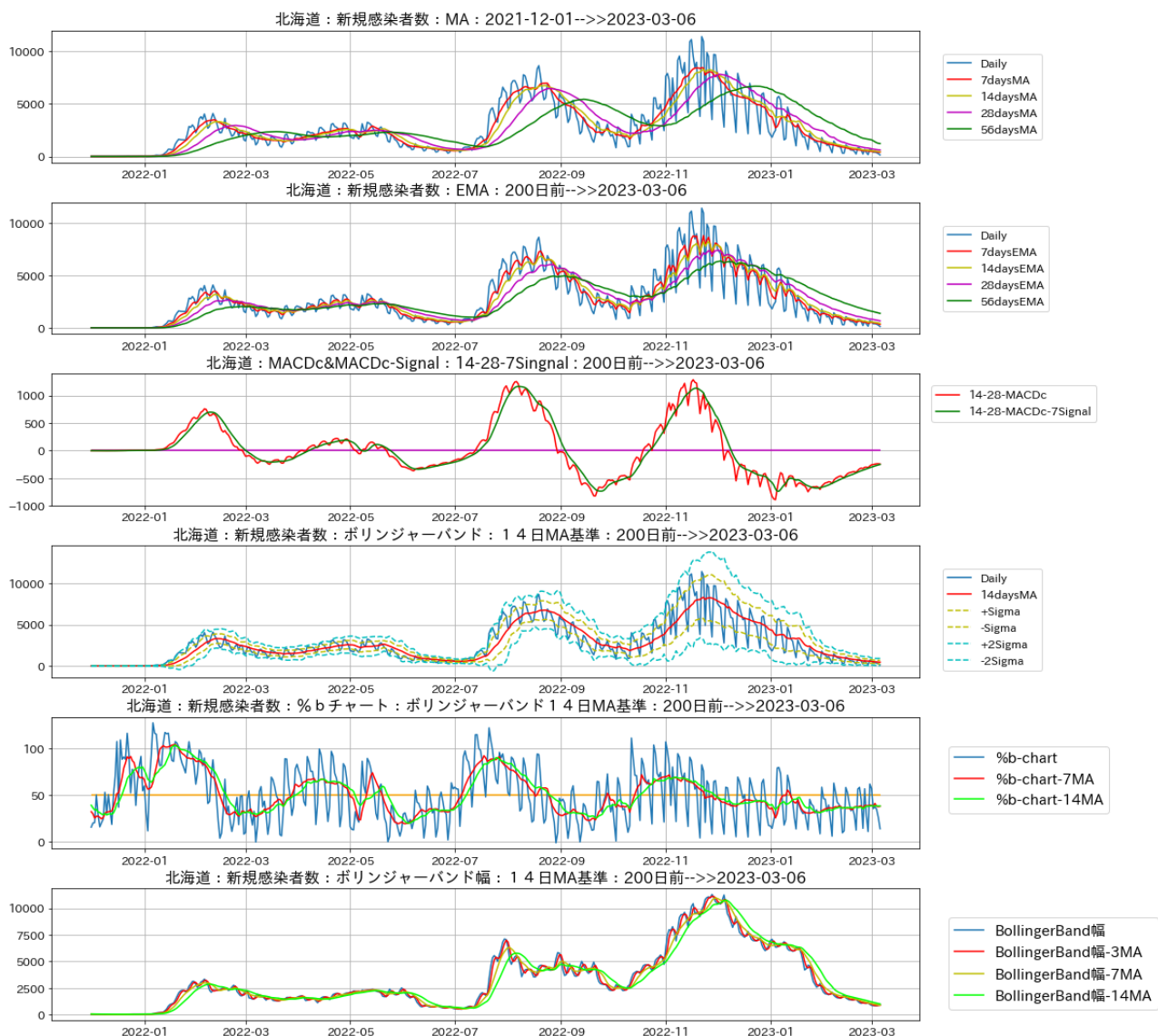


図2 北海道における新規感染者数の移動平均線を基礎とした指標のグラフ (2021/12/1~2023/3/6)

次に、前回の報告同様に、7日平均の値を加工した指標を考察してみる。具体的には $(7MA_{i-6}+7MA_{i-5}+7MA_{i-4}+7MA_{i-3}+7MA_{i-2}+7MA_{i-1}+7MA_i)/7$ で定義された「7日平均の7日平均」(ここで、 $7MA_i$ は日付*i*の7日平均)、及び7日平均の変化率を考え、これまでと同様の考察を試みる。図3において上から順に、①日次データとその移動平均およびMACDcとそのシグナル(これのみ14日平均基準)、②%bチャートとその移動平均、③7日平均の7日平均とそのMACDc、④7日平均の変化率(前日との差)とその7日平均、⑤7日平均の変化率のMACDc、⑥7日平均の変化率の7日平均とそのMACDcである。④と⑤の7日平均の変化率は「日次データにおける当日と7日前の差(の7分の1)」であり「曜日毎の1週間での増減(の7分の1)」という意味でもある。また、⑥の7日平

均の変化率の7日平均は「7日平均における当日と7日前の差(の7分の1)」であり「曜日の起点に無関係な1週間での7日平均の増減(の7分の1)」という意味である。③では、7日平均の7MAのMACDc(赤線)が日次データピーク付近でシグナル(黄線)を揺れずに明確にダウncロcsしている。第6波と第7波では、⑤ではそのMACDc(赤線)のシグナル(黄色)との強いダウncロcsが、感染拡大ピーク付近で見られる。⑥では、7日平均の変化率の7日平均(緑線)のMACDcのシグナルとのダウncロcsが、ピークアウトを捉えている。しかし、日次データの変動幅が非常に大きい第8波ではサインとは言い難い。また緑線の0との明確なアッpcロcsが7日平均の意味の増減の転換点を捉えていると言える。

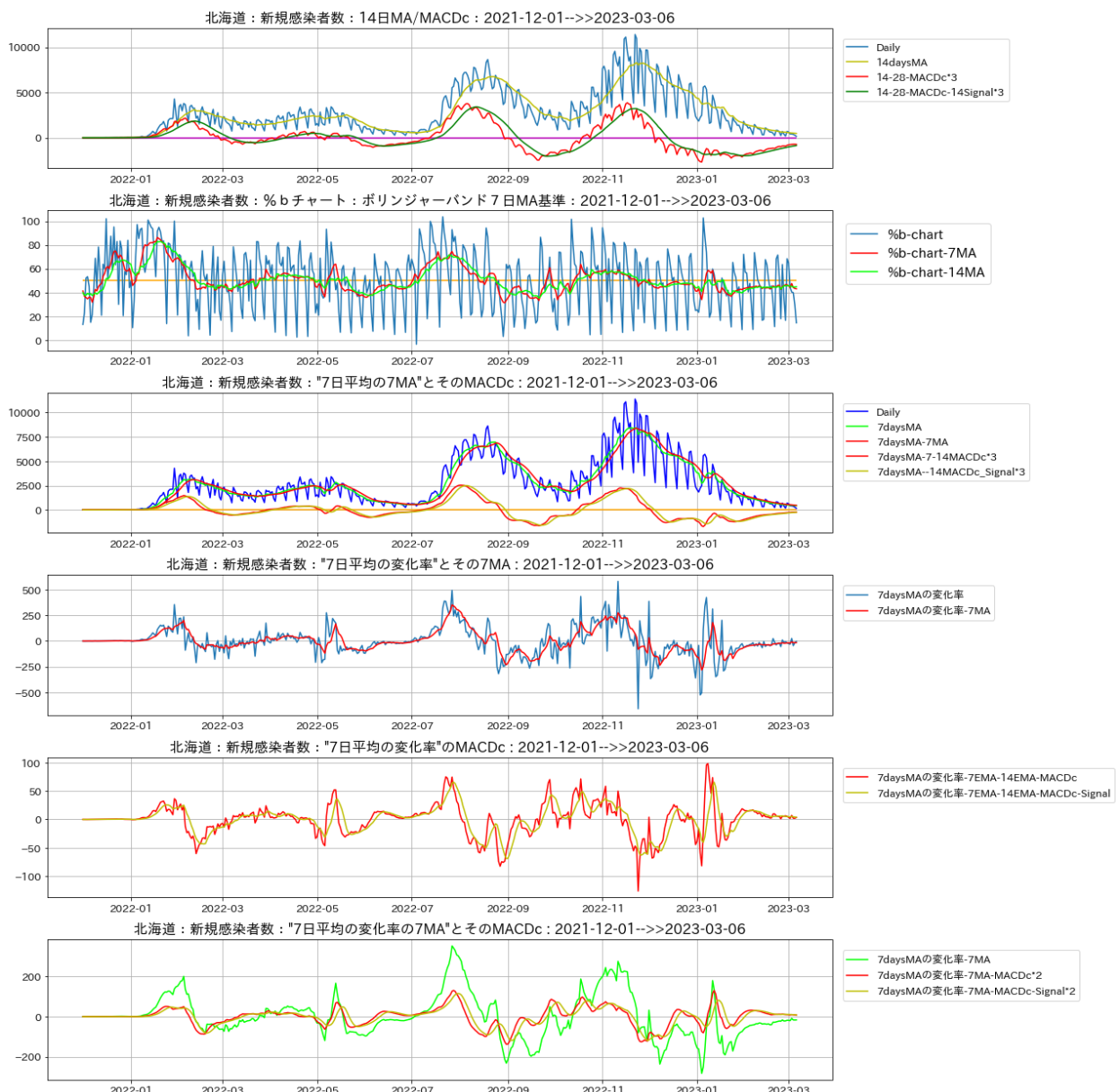


図3 北海道における新規感染者数の7日移動平均に関わる指標のグラフ(2021/12/1~2023/3/6)

3. 大都市圏の新規感染者数の分析と考察

次に、他の都府県として東京都のデータについて分析する。図4は、図3と同様のグラフにボリンジャーバンドのグラフを2番目に加えたものである。波の開始時期では、%bチャートの2つの移動平均が50%をアップクロスし、7日平均の変化率とその7日平均が正値で増加傾向になり、また7日平均の変化率の7日平均が波の開始時期に同時に増加しはじめ、また増減の転換点では0とアップクロスして

いる。さらに、7日平均の7日平均のMACDcでのダウンクロスがピークアウトを捉えている。一方で、7日平均の変化率に関するグラフは変動が激しく大きく、明確なサインを得られない。ピークアウトを捉えているとは言い難い。このように、他のデータにも前章での知見が同様に適用できると考えられるが、データの変動幅が大きい場合は、明確なサインとはならないことも同様に言えることが分かる。

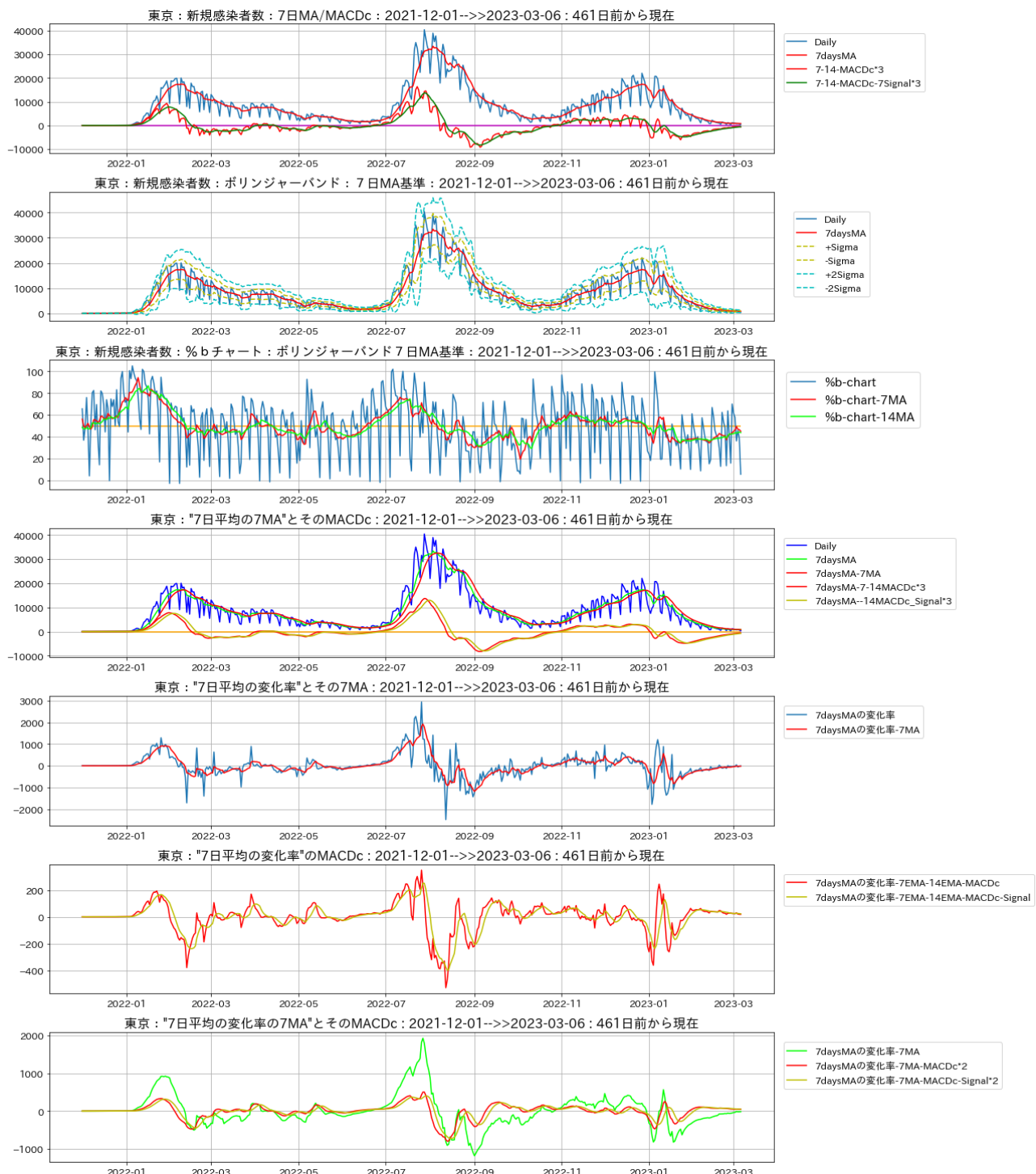


図4 東京都における新規感染者数に関する主要グラフ (2021/12/1~2023/3/6)

4. 世界各国の分析と考察

ここでは、同様の分析を他国(ドイツ、フランス、韓国、ロシア)のデータで試みた。それらの結果を以

下の図 5 と図 6 に示す。データは同時期の 2021 年 12 月 1 日から 2023 年 3 月 6 日までを使用した⁽¹⁾。

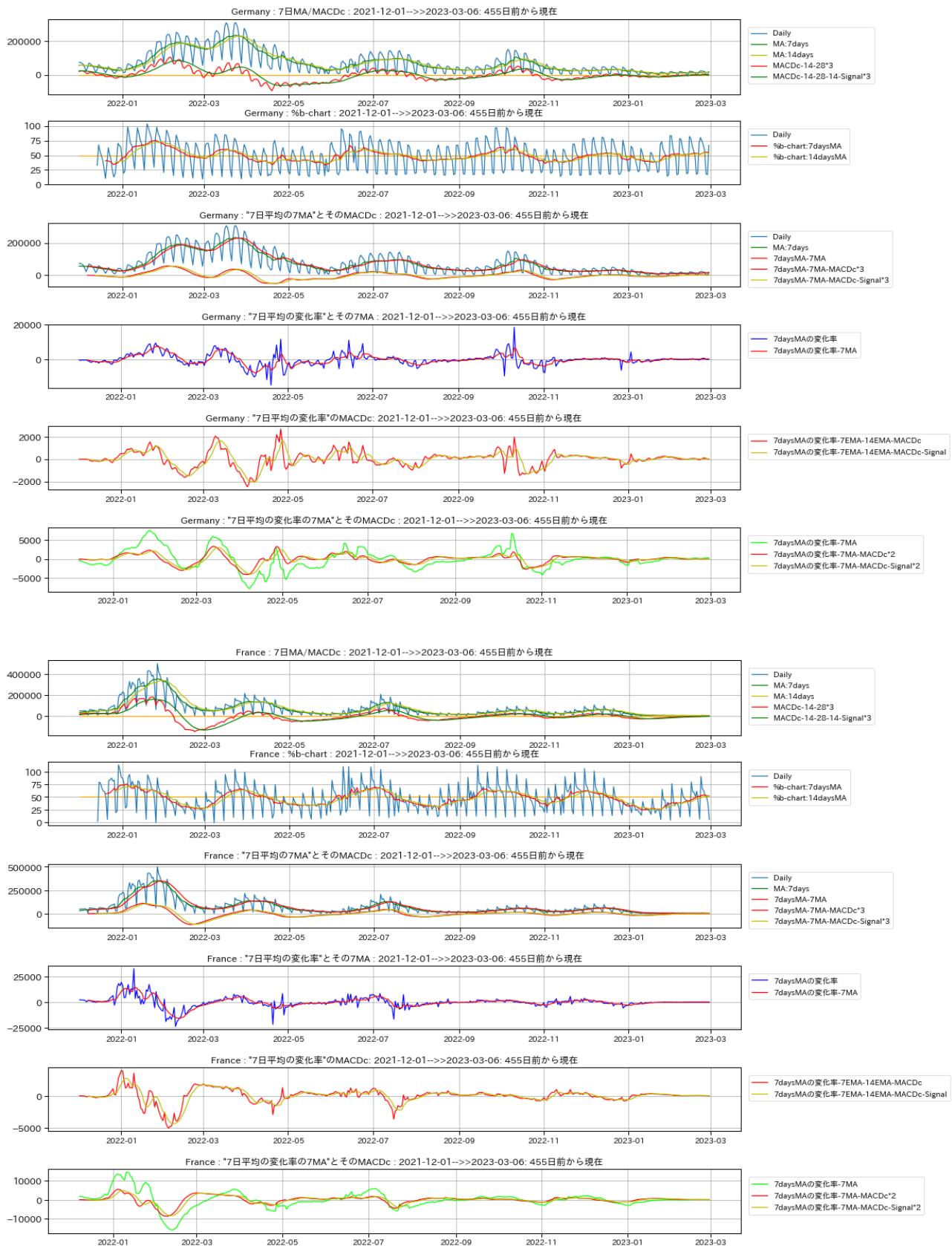


図 5 世界の新規感染者数に関する主要グラフ (上: ドイツ, 下: フランス, 2021/12/1~2023/3/6)

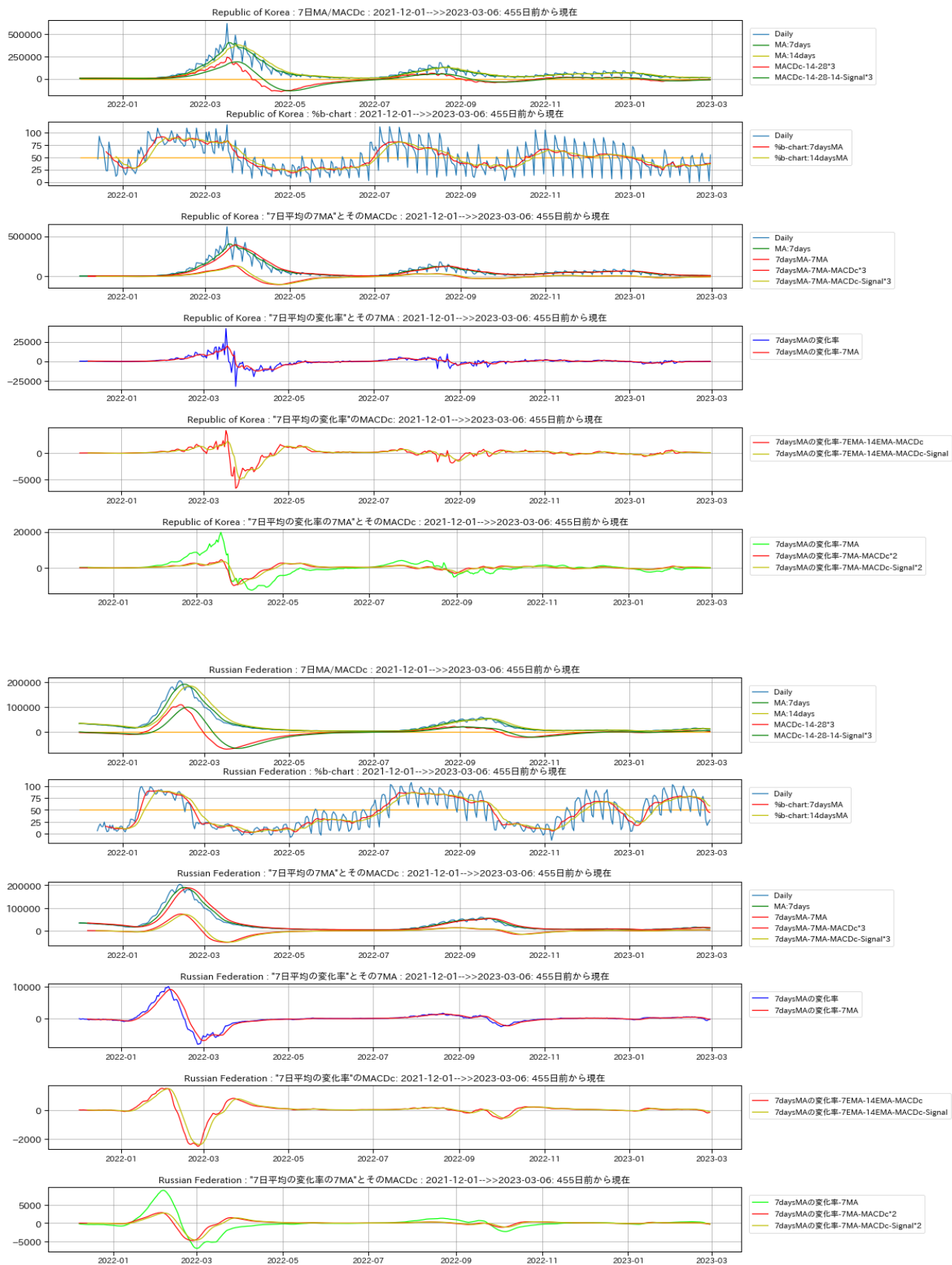


図6 世界の新規感染者数に関する主要グラフ（上：韓国，下：ロシア，2021/12/1～2023/3/6）

図5と図6から、国毎に感染の急拡大縮小期の波の大きさや期間は異なっているが、前章までで得られた知見と同様の知見が適用できると考えられる。図5と図6から、国毎に感染の急拡大縮小期の波の大きさや期間は異なっているが、前章までで得られた知見と同様の知見が適用できると考えられる。他国のデータで確認してみると、感染者数の急拡大期としては、図5のドイツとフランスの2022年1月、3月、6月、9月頃、図6の韓国の2022年2月、7月、10月頃、図6のロシアの2022年1月、8月頃、である。また、ピークアウト付近では、図5のドイツの2022年2月、5月頃図6の韓国の2022年3月、8月頃、図6のロシアの2月、9月頃に同様の知見が確認できる。

5. まとめ

前回の報告と同様に、COVID-19の新規感染者数のオープンデータを時系列データと捉えて、それを時系列分析の手法やその応用手法で分析し、いくつかの知見が再確認された。日本では感染ピークの波の規模が数倍になった2022年の約1年分のデータを中心に分析したが、メディアで多用される7日平均を基にした指標とそれらの可視化により、時系列分析の視点からの理解がしやすくなり、感染者数の増減の動向を理解したり予測したりする手がかりやサインが再確認できたと考えられる。ただし、日本での第8波のような日次変動の幅が極端に大きくなった場合では、7日移動平均では明確なサインにならないこともあり、14日移動平均よりも長いスパンの移動平均を用いることが必要であろうと考えられる。

データサイエンスの視点からの分析であることから、分析における可視化はデータを定量的に処理している。しかしながら、その可視化のグラフから感染の様相を定性的に判断しているにすぎない。データサイエンス的な感染の様相の定義付け（例えば、拡大期、縮小期、停滞期など）を定量的に行った上で、定量的な分析結果を判断することが必要と考えられる。また、分析方法の最適化（例えば、移動平均の日数などのパラメータ最適化）ということも課題となると考えられる。他の感染症などの分析結果との比較も必要と考えられる。世界各国や経済圏、日本国内の都道府県や地方など）についてのデータ分析結果との比較検討も重要と考えられる。

参考文献

- (1) World Health Organization: WHO Coronavirus (COVID-19) Dashboard, 2023年3月7日, <https://covid19.who.int/info/>.
- (2) 厚生労働省: データからわかる－新型コロナウイルス感染症情報－, 2023年3月7日, <https://covid19.mhlw.go.jp/>.
- (3) 北海道データポータル: データカタログ: 新型コロナウイルス感染症に関するデータ【北海道】: 2023年3月7日, <https://www.harp.lg.jp/opendata/dataset/1369.html>.
- (5) Gerald Appel: Technical Analysis: Power Tools for Active Investors, Financial Times Prentice Hall, p.166, 2005.
- (6) John Bollinger's Official Bollinger Band Website: 2023年3月7日, <https://www.bollingerbands.com/>.
- (7) 機械学習を用いたCOVID-19新規感染者数の分析, 小松隆行, 北海道科学大学研究紀要, (49), 49-61 (2021-09-30).
- (8) COVID-19新規感染者数の時系列分析, 小松隆行, 北海道科学大学研究紀要, (50), 53-60 (2022-09-30).