

# Airbnb データを用いたローカルな民泊需給の ダイナミック価格形成要因の分析

仙石裕明\*, 秋山祐樹\*, 渡部展也\*\*

\*東京大学空間情報科学研究センター, \*\*中部大学国際 GIS センター

## 1. はじめに

本研究は不動産の複合的かつダイナミックな価格形成要因を解明する分析手法を構築することを目的としている。

不動産の需給構造の大きな構成要素の一つである価格は、経済情勢や社会認知、立地情報などのさまざまな複合要因によって形成される。しかしながら、取得できるデータには限りがあり、同一の財において価格推移の把握は構造的に困難であり、不動産の透明化を妨げる要因となっている（清水，2004）。リピートセールス法により欧米では不動産価格指数化が発展しているが、需給が考慮されていないことや部屋番号まで特定できず時間効果を正確に把握できないことによる課題が報告されている（唐渡ら，2012）。

そこで、本研究では予約情報を含めた過去の民泊情報を公開する Airbnb データ（参考文献・データ

[3]を参照）を用いた。自社サービスにおける民泊登録者や予約状況を定期的に集約し、公開したデータである。宿泊所や予約状況に関する情報が各月単位で公開されている。本データを用いることで、同一物件における価格推移を追跡し続けることが可能となる。

本研究ではオリンピック開催という通常の不動産データでは考慮が困難な経済情勢を対象とし、Airbnb のようにダイナミックなデータの時空間分析の可能性を試みた。延期前後の変化にフォーカスを当てて、その影響評価の分析を行った。

## 2. 方法

### 2.1 データ

本研究では Airbnb データをメインソースとし、表 1 のデータ一覧を用いた。Airbnb は全世界の主要都市を対象とした民泊プラットフォームサービスであり、本データは物件の位置情報（秘匿あり）を含めた物件単位データや予定価格を含めた過去の民泊情報はオープンデータとして公開されて主要諸国の大都市において公開されている。日本の場合では東京都の情報が公開されている。さらに、本データにはホストのスコアなどユーザー評価データも記載されており、口コミ情報も含まれている。時間的粒度が細かく、毎月の動向を追跡することが可能である。ただし、本データでは月次単位でしか実績情報が公開されず、同時に公開される予定価格では予測上、誤差が生じうるため、月次の実績データをモデル学習に用いた。Airbnb データをターゲットとし、国土数値情報などの他の立地情報や動態情報に関する情報を結合した（図 1）。

表 1 利用データ一覧

データソース	データ項目	時点	空間単位	備考
Inside Airbnb Listing, Calendar (メイン)	価格 予約状況 その他物件属性	2019.3~2020.9	物件単位	
国土数値情報 鉄道データ	駅名称・位置情報	2019.12.31	駅単位	物件から最寄駅判定、最寄り駅までの距離計算
新宿区 住民基本台帳による 世帯と人口	人口総数・前月比増減 (外国人住民含む)	2019.3~2020.9	市区町村単位	データを確認し、今回は不使用
Google Trends	「観光関連」 「COVID-19関連」な どのキーワード	2019.3~2020.12	キーワード単位	

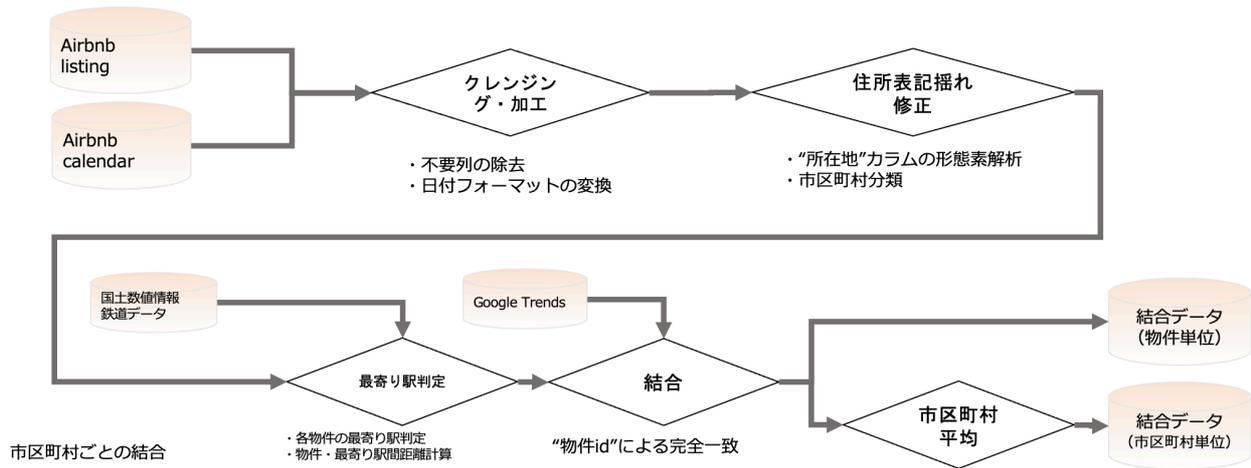


図1 パネルデータ構築手順

## 2.2 分析の流れ

分析の流れは図2のようにになっている。2019年3月-2020年8月を対象とする。オリンピック中止決定前後の予約状況から当初のオリンピック開催期間の前後の価格変化を対象に、立地等の地理情報およびGoogleトレンドデータ等を用いて、状態空間モデルにおける差異を分析する。

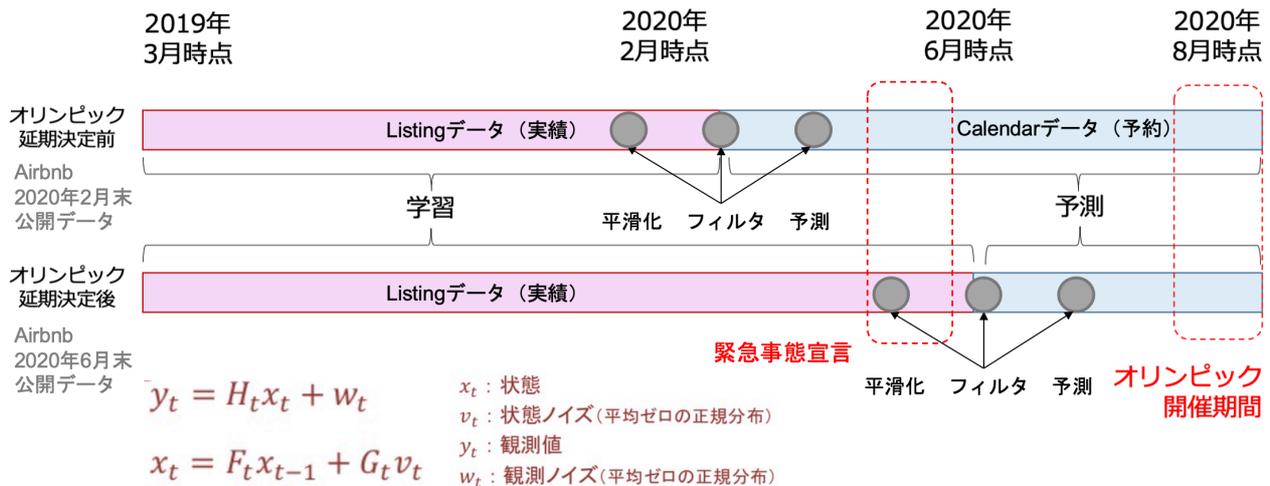


図2 分析の流れ

## 3. 結果

### 3.1 推移

Airbnb データからオリンピック開催期間における予約・価格状況延期決定前の予約は開催期間で満室率・価格ともに他期間と比べて上昇していたが、延期決定後の予約ではその上昇幅が小さくなっている。

人口推移においては2020年2月から12月まで人口減少(夜間人口)が確認された(図4)。2019年は4月に進学や就職・転勤などで移転が多くなる影響で2019/05/01の人口総数前月比が大幅に増加しているが、2020年はその傾向が見られない。コロナ感染拡大による顕著な影響と考えられる。ただし、図3の価格推移とは傾向が一致しないため、モデルの説明変数からは除外することとした。

Googleトレンドからは、こうした年単位での比較においてはCOVID-19流行の影響が顕著に確認できる。価格推移や予約率推移を説明することが延期決定前までは「Olympic」や「Tokyo」というキーワードにおいても緩やかに上昇しつつ、予約率とも弱い相関が確認された。一方で、オリンピック延期決定前後(2020/04/01)では大きな増減の変化は見られなかった。

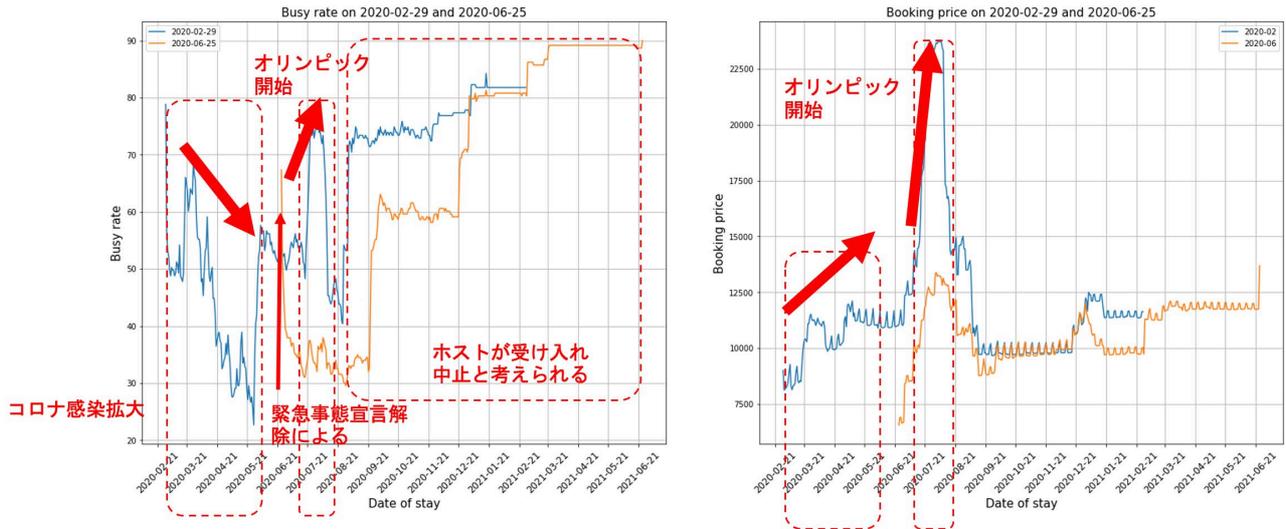


図3 オリンピック開催期間における予約・価格状況

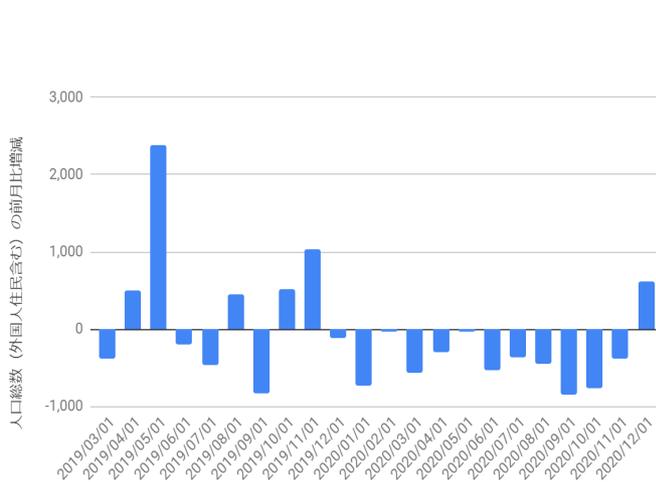


図4 人口総数前月比増減の推移

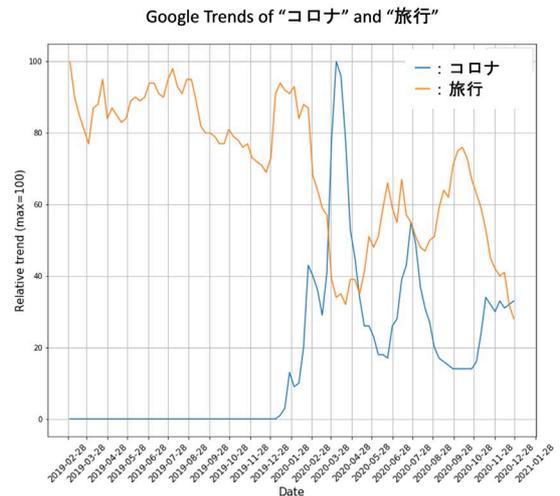


図5 Google Trends “コロナ”と”旅行”

### 3.2 予測結果

以上のデータを用いて、ローカル線形トレンドモデルを作成した結果を示す(表2)。物件単位においてオリンピック延期決定前(2月時点)と決定後(6月時点)でのそれぞれの任意の物件での価格推移結果を示す(図6)。コロナ感染拡大の影響が顕在的に確認された。本研究で用いているGoogleトレンドのような説明変数では、このような急激な変化が同時期にないために、本モデルには適切なトレンド情報を考慮できていないと考えられる。たとえば緊急事態宣言やGoToトラベルキャンペーンなどの政策等が影響していると考えられ、同時期における計画情報・予想情報・中規模以上のイベント情報を加味できることが望ましい。また、物件単位では物件に応じて、価格を変更していない場合や延期前から価格を下げている場合が確認され、ホストの意向によって個別具体的となるため、傾向を集約できることが望ましい。

最寄駅からの距離などの地理的な特徴については、モデルへの寄与は確認されなかった。この点は居住物件とは傾向が異なる点といえる。また、物件そのもののグレードやホストの人気は必ずしも価格と連動していないことも確認された。

表 2 状態空間モデルのモデル精度

Dep. Variable:	price	No. Observations:	16
Model:	deterministic trend	Log Likelihood	-0.264
Date:	Tue, 23 Feb 2021	AIC	76.528
Time:	16:04:30	BIC	100.813
Sample:	03-01-2019	HQIC	74.281
	- 06-01-2020		
Covariance Type:	opg		

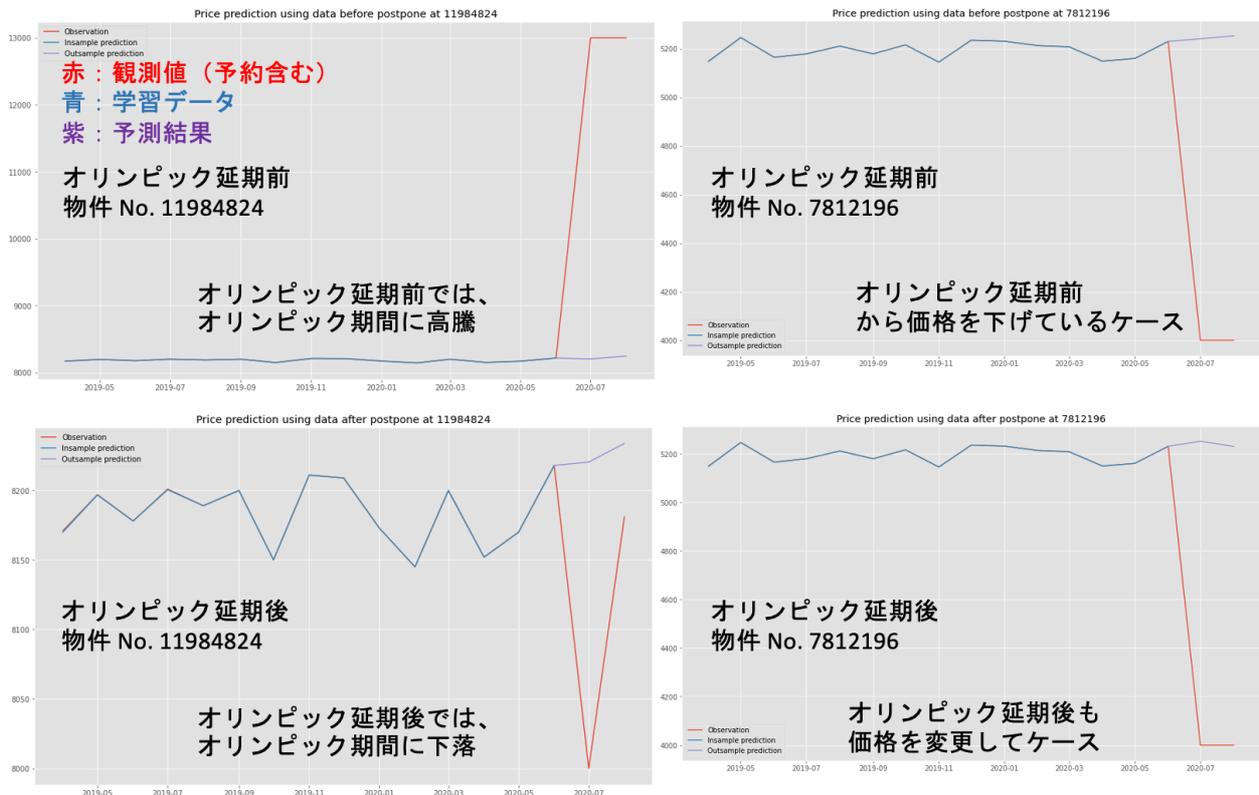


図 6 予測結果 (物件単位)

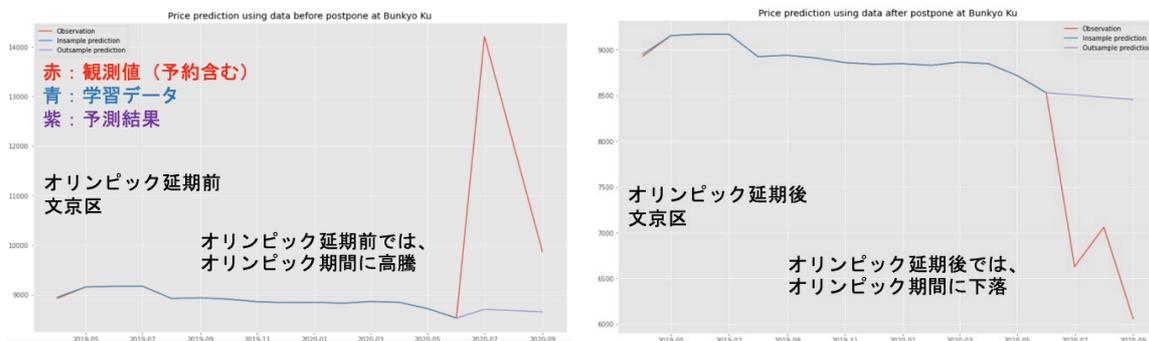


図 7 市区町村単位平均予測結果 (文京区)

市区町村単位で価格を平均化した場合の結果を示す(図7)。通常時では大きな変動がなくほぼ横ばいとなっているが、オリンピック期間中では他期間にはない突出があることを確認できる。物件単位のとくと同様に大きな残差となっている。寄与している変数としては、Google Trend - Tokyo, Google Trend - 五輪, Google Trend - 旅行などがあげられた。外生変数よりも自己回帰性があるデータであるが、月次に集約してしまったことにより、モデルで表現しづらくしてしまったと考えられる。

#### 4. 考察

モデル学習において、2020年7月以前の実績値での変動はほとんど見られなかったため、オリンピック開催期間における高騰・下落を表現できなかつたと考えられる。今回はデータの制約上、月次単位でモデルを構築したが、日次変化を考慮する以外にはモデルでの表現が難しい。Inside Airbnb データだけでは、オリンピック延期やバブル崩壊のような大規模な経済事象の表現には、モデル学習に十分とはいえず、別途日次で実績データを取得しつづけること、日別で利用できるデータリソースが必要である。同データにある予定価格となるカレンダーデータの活用もありうるが、経済状況を正確に表すことができないことに留意する必要がある。長期のデータ保有というよりは細かい単位でのデータ保持が重要と考えられる。

モデルに寄与している変数のなかには、内生変数ではなく、外生変数が大半であった。満室率と価格には一定の関係性があることがデータから確認できるが、物件単位では満室率を計算することが構造上できないため、モデルでは考慮できていなかった。たとえば、一週間前から現在に至るまでというように、一定期間内における平均満室率または、地域一帯の平均満室率などというかたちで説明変数に加えることが考えられる。また、本研究では部屋タイプの区分を分けていなかったが、部屋タイプが異なることで価格への影響が大きいいため、本来的には区分するべきであった。

物件によってはグレードやホストの意向が異なるなど、データ化されていない Airbnb 特有の価格不均質性が存在することを確認できた。

#### 5. まとめ

本研究では、Airbnb データをメインソースとしたパネルデータを生成し、状態空間モデルを用いて複合的な分析を行った。オリンピック期間における影響は時間単位を日次に変更することや、満室率をモデルに反映させる必要が考えられる。オリンピック期間外においても、物件グレードやホストの意向などデータ化されていない項目の存在が課題となっている。

#### 6. 謝辞

本研究は中部大学問題複合体を対象とするデジタルアース共同利用・共同研究 IDEAS202003 の助成を受けたものです。

#### 参考文献・データ

1. 清水千弘, 不動産市場分析, 住宅新報社, 2004.
2. 唐渡広志・清水千弘・中川雅之・原野啓, 「リピー トセールス不動産価格指数における集計バイアス」『日本経済研究』No66 , pp22 -50, 2012.
3. Inside Airbnb <http://insideairbnb.com/get-the-data.html>