

# 隠れマルコフモデルに基づく最短経路探索の考え方をを用いたマップマッチング手法の検証

橋口 剛<sup>1</sup>・瀬谷 創<sup>2</sup>・瀬木 俊輔<sup>3</sup>・井料 隆雅<sup>4</sup>・安田 昌平<sup>5</sup>

<sup>1</sup> 学生非会員 神戸大学 大学院工学研究科 (〒657-8501 神戸市灘区六甲台町 1-1)  
E-mail: 84gucci.1218@gmail.com

<sup>2</sup> 正会員 神戸大学准教授 大学院工学研究科 (〒657-8501 神戸市灘区六甲台町 1-1)  
E-mail: hseya@people.kobe-u.ac.jp

<sup>3</sup> 正会員 神戸大学准教授 大学院工学研究科 (〒657-8501 神戸市灘区六甲台町 1-1)  
E-mail: segi@people.kobe-u.ac.jp

<sup>4</sup> 正会員 東北大学教授 大学院情報科学研究科 (〒980-8579 宮城県仙台市青葉区荒巻字青葉 6-6)  
E-mail: iryo@tohoku.ac.jp

<sup>5</sup> 正会員 東京大学助教 大学院工学系研究科 (〒113-8656 東京都文京区本郷 7-3-1)  
E-mail: s.yasuda@civil.t.u-tokyo.ac.jp

マップマッチング (MM) は交通データ処理の基礎技術としてニーズが大きい。近年、隠れマルコフモデル (HMM) を用いた MM 手法の研究が進展しているが、HMM に基づく典型的な方法では、実際には通過していない道路を通過したと判定しやすいという点に課題がある。これに対して瀬木ら (2022) は、HMM における遷移確率 (状態空間モデルにおけるシステムモデル) に最短経路探索の考え方を取り入れた MM 手法を提案し、当該課題の改良を試みた。本研究では、瀬木ら (2022) の手法の適用性を検討するために、世界各国で取得された 100 種類の GPS 軌跡データと真の走行ルートの情報を用いて、道路ネットワークの構造や取得された位置情報の特性が MM の経路推論の正確度にどのように影響するかについて、メタ分析を通して検証を行った。

**Key Words:** map matching, hidden Markov model, shortest path search, logit model

## 1. はじめに

マップマッチング (map matching (MM)) あるいは経路推論 (path inference) とは、GNSS 等によって測位された移動体の移動軌跡より、道路ネットワーク上の移動経路を特定することを意味する。MM はオンライン MM とオフライン (batch) MM に大別される<sup>1)</sup>。前者は主にナビゲーションシステム等に用いられるもので、リアルタイムあるいは準リアルタイムで走行中のネットワークリンクを特定することを目的としている。一方で、後者は、過去の位置情報シーケンスを処理し、単一の車両や人の移動軌跡を連続するネットワークリンクの軌跡として抽出することを目的としている。本研究で着目するのは、オフライン MM である。

GNSS の軌跡情報は、常に測位誤差を含み、マルチパ

ス等の影響で、しばしば測位誤差は大きくなる。したがって、観測された位置を真の位置として扱う MM 手法には限界がある。近年、測位誤差 (観測ノイズ) が多く時系列的に疎な位置情報に対しても頑健な MM 手法として、Newson et al. (2009)<sup>2)</sup>、Low et al. (2009)<sup>3)</sup> が既往研究を改良する形で提案した隠れマルコフモデル (hidden Markov model (HMM)) を用いた MM 手法が広く用いられている。実際に、著名な GIS ソフト (ArcGIS, QGIS) や地図ベンダーが提供する API などでは、現時点で筆者らが知る限りすべて HMM が用いられている。HMM を用いた既往研究として、例えば Jagadeesh and Srikanthan (2017)<sup>4)</sup> は、ノイズが多くスパースな位置情報を正確かつタイムリーに MM するために、HMM とドライバーの経路選択を組み合わせたオンライン MM 手法を提案した。Qi et al. (2019)<sup>5)</sup> は、MM の課題がジャンクション部の

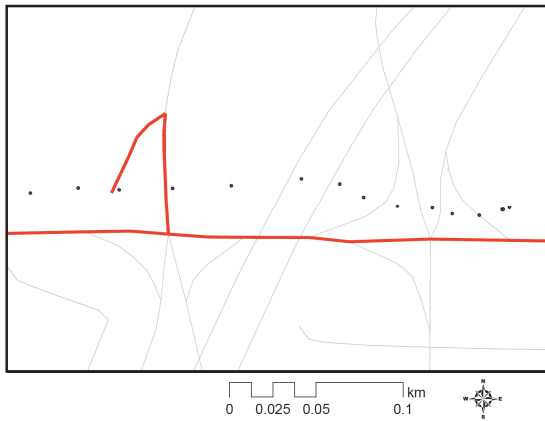


図-1 Reverse movement の例

マッチングにあるとし、ジャンクション領域判定モデルを提案し、当該モデルを用いて HMM による MM の性能を向上させた。Yang and Gidofalvi (2018)<sup>8)</sup>は、Newson et al. (2009)<sup>9)</sup>の HMM にいくつかのモデル上の改良を施しつつ、大規模な実ネットワークにおいて実行可能な方法として実装上の工夫を行った Fast Map Matching (FMM) を提案した。橋口らは<sup>7)</sup>、経路推論の正確度の観点から、FMM の適用性を日本の実データを用いて確認した。しかし、FMM を含む既往の HMM を用いた MM 手法の多くは、reverse movement と呼ばれる、車両が目的地と反対方向に走行し、その後正常な方向に戻るような動き（図-1）を含む非現実的な経路推論結果が得られてしまう、すなわち実際には通っていないリンクが通ったと推定されることが課題であり、実際当該現象は、橋口ら<sup>7)</sup>の検証においても確認されている。Reverse movement については、リンク長を用いたアドホックな補正が行われることが多いが（Yang and Gidofalvi (2018)<sup>8)</sup>）、補正のためのパラメータ設定が難しい点に難がある。

瀬木ら<sup>8)</sup>は、HMM における遷移確率（状態空間モデルにおけるシステムモデル）に最短経路探索の考え方を取り入れた MM 手法を提案し、当該課題の改良を試みた。瀬木ら<sup>8)</sup>の方法は遷移確率のモデル化にロジットモデルを導入しているため、以下“HMM-logit”と呼称することとする。Oyama and Hato<sup>9)</sup>は、GPS の点列から、再帰型ロジット（recursive logit (RL)）モデルを用いて経路推論を行うモデルを提案した。これに対して HMM-logit は、経路選択行動の部分には、RL モデルではなく最短経路探索をベースとしたロジットモデルを用いているため、例えば経路単位の説明変数の導入のためには拡張が必要である。しかし、オフライン MM であることを活かし、ある点の MM に、それ以降の観測データも用いるため、局所的に望ましいリンクへのマッチングを防ぐことができる点に利点がある（link switching の問題が生じない（Oyama and Hato<sup>9)</sup> Appendix A を参照のこと））。

瀬木ら<sup>8)</sup>は、真の経路上にないリンクを通らない、という誤差指標（3.(2) 参照）において、HMM-logit が FMM を上回ることを指摘した。しかし、HMM-logit がどのような状況に強いのか、あるいは弱いのかについてのより詳細な検討を行うことが課題として残っていた。そこで本研究では、Kubička et al.<sup>9)</sup>が提供する世界 100 都市のオープンソースの軌跡情報データセットを用いて MM を行い、メタ分析を通して、道路ネットワークの構造や取得された位置情報の特性が MM の経路推論の正確度にどのように影響するかを明らかにすることを目的とする。

以下、第 2 章では、本検証で用いる FMM および瀬木ら<sup>8)</sup>の手法について述べ、第 3 章ではメタ分析に用いる Kubička et al.<sup>9)</sup>が提供するデータセットについて説明する。続いて第 4 章でメタ分析の結果を示し、最後に第 5 章で本研究のまとめと今後の課題について述べる。

## 2. 本研究で用いる MM 手法

### (1) HMM

HMM では、出力確率（emission probabilities）と遷移確率（transition probabilities）の積で MM を処理する。以下、GNSS の測位点が得られている各時点  $t = 0, 1, 2, \dots, T$  で表す。時点  $t$  においてマッチングの対象となり得るリンクの集合（候補リンク集合）を  $X_t$  で表す。

#### a) 出力確率

出力確率（状態空間モデルにおける観測モデル）は、トリップ毎の GNSS の点列がリンク集合  $X_t$  内のどのリンク上にあるかを示す確率である。これは、付近の複数あるマッチング候補のリンクとの距離（各リンクへの垂線）に依存した確率である。出力確率は、

$$p(z_t | x_t) \propto \exp \left[ -\frac{1}{2} \left\{ \frac{\text{mdist}(x_t, z_t)}{\sigma_z} \right\}^2 \right] \quad (1)$$

で与えられるとする。ここで、 $z_t$  は時点  $t$  の GPS 測位点を、 $\text{mdist}(x_t, z_t)$  はリンク  $x_t$ （表記をシンプルにするためリンクの ID は省略）と測位点  $z_t$  間の最短距離（射影距離）を表す（図-2）。ここでのポイントは、標準偏差パラメータ  $\sigma_z$  の存在である。GPS の位置正確度が高い場合、これを小さくし、低い場合これを大きくすればよい（すなわち、外生パラメータ）。なお、HMM を用いる場合、時点 0 の状態  $x_0$  の周辺分布（事前分布）を与える必要がある。ここでは、どの候補リンクに存在する確率も等しいという事前分布を用いる。

$$p(x_0) = \frac{1}{|X_0|} \quad (\forall x_0 \in X_0) \quad (2)$$

#### b) 遷移確率

出力確率は、多くの場合正規分布で与えるが、遷移

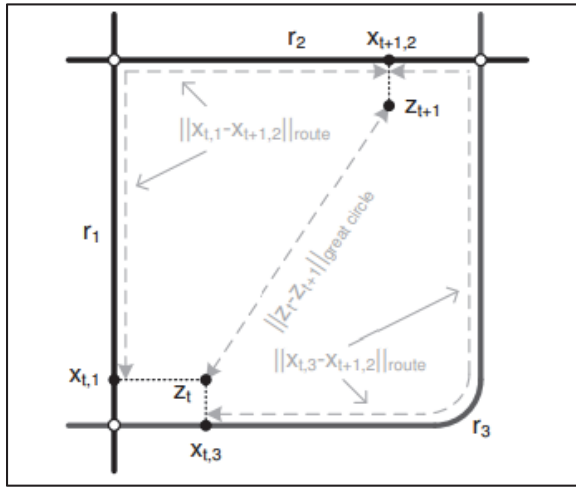


図-2 各変数の意味 (Newson et al. (2009)より)

確率 (状態空間モデルにおけるシステムモデル) は、研究によっていくつかのバリエーションがある。代表的なアプローチは、測定された  $z_t, z_{t+1}$  のような GPS の点同士の直線距離 (厳密には great circle 距離) と、 $x_{t,1}$  と  $x_{t+1,2}$  や  $x_{t,3}$  と  $x_{t+1,2}$  のような、リンク上の対応点間の経路距離との近さによって定義するというものである。Newson et al. (2009)<sup>9)</sup> は、遷移確率  $p(d_t)$  を指数分布として次式により与えている。

$$p(d_t) = \frac{1}{\beta} e^{-d_t/\beta} \quad (3)$$

ただし、

$$d_t = \left| \|z_t - z_{t+1}\|_{\text{great circle}} - \|x_{t,i} - x_{t+1,j}\|_{\text{route}} \right| \quad \text{と}$$

する。ここで、パラメータ  $\beta$  については、推定する必要がある。一方、Yang and Gidofalvi (2018)<sup>9)</sup> の FMM では、 $p(d_t)$  を次式のように定義している。

$$p(d_t) = \frac{\min(\|z_t - z_{t+1}\|_{\text{great circle}} - \|x_{t,i} - x_{t+1,j}\|_{\text{route}})}{\max(\|z_t - z_{t+1}\|_{\text{great circle}} - \|x_{t,i} - x_{t+1,j}\|_{\text{route}})} \quad (4)$$

このように定義することで、パラメータ  $\beta$  を推定する手間がなく、遷移確率が直ちに求められる。ここで、岩館ら (2015)<sup>10)</sup> が述べる通り、これらの手法は、出力確率において観測値が使われており、厳密に言えば HMM とはなっていない。この点は、観測ノイズが大きい場合に実用上問題になり得ると思われる。

一方で、瀬木ら<sup>8)</sup> は、遷移確率を次式によって与えている。時点  $t$  においてリンク  $x_t \in X_t$  にいるとき、それ以降に負担する期待最小費用を  $v_t(x_t)$  で表す。最終時点  $T$  以降に負担する費用は、どの候補リンクにおいても 0 とする： $v_T(x_T) = 0 \ (\forall x_T \in X_T)$ 。時点  $t$  から  $t+1$  にかけての、車両が存在するリンクの遷移確率を

$$p(x_{t+1}|x_t) = \frac{\exp[-\theta\{\text{dist}(x_t, x_{t+1}) + v_{t+1}(x_{t+1})\}]}{\sum_{x'_{t+1} \in A_{t+1}(x_t)} \exp[-\theta\{\text{dist}(x_t, x'_{t+1}) + v_{t+1}(x'_{t+1})\}]} \quad \begin{matrix} (x_{t+1} \in A_{t+1}(x_t)) \\ 0 \quad (x_{t+1} \in X_{t+1} \setminus A_{t+1}(x_t)) \end{matrix} \quad (5)$$

と表す。また、 $v_t(x_t)$  をログサム変数として表す。

$$v_t(x_t) = -\frac{1}{\theta} \ln \sum_{x'_{t+1} \in A_{t+1}(x_t)} \exp[-\theta\{\text{dist}(x_t, x'_{t+1}) + v_{t+1}(x'_{t+1})\}] \quad (6)$$

ここで、 $\text{dist}(x_t, x_{t+1})$  はリンク  $x_t$  から  $x_{t+1}$  への移動に要する距離であり、

$$\text{dist}(x_t, x_{t+1}) = \begin{cases} \text{len}(x_t) + \text{sdist}(\text{to}(x_t), \text{from}(x_{t+1})) & (x_t \neq x_{t+1}) \\ 0 & (x_t = x_{t+1}) \end{cases} \quad (7)$$

と定義する。ここで、 $\text{len}(x_t)$  はリンク  $x_t$  の長さを、 $\text{sdist}(n_1, n_2)$  はノード  $n_1, n_2$  間の最短距離を、 $\text{to}(x_t)$  はリンク  $x_t$  の終点ノードを、 $\text{from}(x_{t+1})$  はリンク  $x_{t+1}$  の始点ノードを表す。同一リンクに留まり続ける場合、移動距離は生じない。 $A_{t+1}(x_t)$  は、時点  $t$  に車がリンク  $x_t$  にいるとき、時点  $t+1$  に車が物理的に存在可能なリンクの集合である。

$$A_{t+1}(x_t) = \{x_{t+1} | x_{t+1} \in X_{t+1} \text{ and } \text{sdist}(\text{to}(x_t), \text{from}(x_{t+1})) \leq \Delta t_s \cdot \bar{v}\}$$

ここで、 $\Delta t_s$  は時点  $t$  と  $t+1$  の間の GPS のタイムスタンプの差分を、 $\bar{v}$  はネットワーク上の最高速度を表す。

以上がシステムモデルとなる。定式化から明らかのように、このシステムモデルは、車両がリンク内のどこに存在するのかは考慮しない。これによって、状態空間の次元を小さくし、計算に要する時間を短縮している。

以上のシステムモデルと観測モデルを前提としてマップマッチングを行う。まず、以下の式により、1期先予測とフィルタリングの計算を  $t=0$  から  $t=T$  まで繰り返す。

#### [1期先予測]

$$p(x_t | z_{0:t-1}) = \sum_{x'_{t-1} \in X_{t-1}} p(x'_{t-1} | z_{0:t-1}) p(x_t | x'_{t-1}) \quad (8)$$

#### [フィルタリング]

$$p(x_t | z_{0:t}) = \frac{p(x_t | z_{0:t-1}) p(z_t | x_t)}{\sum_{x'_t \in X_t} p(x'_t | z_{0:t-1}) p(z_t | x'_t)} \quad (9)$$

ただし、 $t=0$  の場合は、 $p(x_0 | z_{0:-1}) = p(x_0)$  となるため、1期先予測の計算は必要ない。次に、平滑化とマッチング確定の計算を  $t=T$  から  $t=0$  まで繰り返す。

#### [平滑化]

$$p(x_t | z_{0:T}) = p(x_t | z_{0:t}) \sum_{x'_{t+1} \in X_{t+1}} p(x'_{t+1} | x_t) \frac{q(x'_{t+1} | z_{0:T})}{p(x'_{t+1} | z_{0:t})} \quad (10)$$

#### [マッチング確定]

$$q(x_t | z_{0:T}) = \begin{cases} 1 & (x_t = \arg \max_{x_t \in X_t} p(x_t | z_{0:T})) \\ 0 & (x_t \neq \arg \max_{x_t \in X_t} p(x_t | z_{0:T})) \end{cases} \quad (11)$$

ただし、 $t=T$  の場合は平滑化の計算は必要ない。通常の状態空間モデルでは、 $x_t$  の値を確定せず、 $(x'_{t+1} | z_{0:T}) = p(x'_{t+1} | z_{0:T})$  として平滑化の計算のみを行

表-1 データセットの例

ID	Area	Spatial-mismatch			Difficulties		
		mad	u-turns	hives	loops	gaps	severe congruence issues
37	Tooyama, Japan	0.571				yes	
97	Hillerød, Denmark	0.582				yes	
82	Saint Petersburg, Russia	0.804					
39	Helsinki, Finland	0.857	yes			yes	
42	Cebu, Philippines	1.029			yes	yes	
8	Huntsville, Alabama, USA	3.795	yes			yes	yes
80	Buenos Aires, Argentina	3.914				yes	yes
43	Amsterdam, Netherlands	5.179				yes	
4	Rathenow, Germany	5.789	yes			yes	
54	Amsterdam, Netherlands	26.01		yes		yes	yes

う。しかし、マップマッチングの場合は1本の経路を確定的に選定する必要があるため、 $x_t$ の値を確定することが望ましい。本手法では、最尤法によって $\theta$ を推定できる。この点は大きな利点である。

### 3. 本研究で用いるデータセット

#### (1) データ

本研究では、Kubička et al.<sup>9)</sup>が提供する世界100都市のオープンソースの軌跡情報データセットを用いてMMの正確度の検証を行う。本データセットは、Open Street Map事業におけるPlanet GPXとGPS点群コレクションを加工したものである。100都市は、都市部から郊外部まで様々である。表-1に、その一例を示す。本データセットには、xml形式の概要説明が付属しており、GPS軌跡と道路ネットワークのずれを示すspatial-mismatchというタグには、ずれの中央値であるmedian absolute deviation (mad)とずれの最大値であるmaxの値が格納されている(表-1にはmadで見た上位と下位5ケースを示している)。Open Street Map事業で収集されたというデータの背景上、GPSの位置正確度は、ID54を除けば、madでみて0.5-5m程度となっており、比較的高い。このようなデータを用いた検証であることを認識することは重要である。

また、データセットの中には、MMが難しくなるケース(difficulties)として、u-turns, hives, loops, gaps, severe congruence issuesそれぞれに該当するか否かの情報が格納されている。それぞれの項目の意味は、u-turns: 車両軌跡にUターンがあるか、hives: 観測点が狭い領域に集中している箇所があるか、loops: ループしている箇所があるか、gaps: 観測点が飛躍している箇所があるか、severe congruence issues: ネットワークと観測点が明らかにずれている箇所があるか、である。

#### (2) 正確度評価指標

本研究で用いるMMの正確度評価指標について、瀬

木ら<sup>9)</sup>同様、以下の3つの指標を採用することとした。

Match ratio =

$$\frac{1 - \text{正解経路のうち特定されなかったリンク長}}{\text{正解経路長}}$$

Undeviating ratio =

$$\frac{1 - \text{誤って特定されたリンク長}}{\text{正解経路長}}$$

Similarity ratio =

$$\frac{1 - (\text{正解経路のうち特定されなかったリンク長} + \text{誤って特定されたリンク長})}{\text{正解経路長}}$$

経路が完全に正しく特定されると、上記の3指標の値はすべて1となる。

### 4. 検証

#### (1) 検証方法

本研究では、全100ケースについて、FMM及び瀬木ら<sup>9)</sup>の方法でMMを行う。両手法では、出力確率における標準偏差パラメータ $\sigma_z$ を外生的に設定する必要がある。本研究では、 $\sigma_z = 5, 10, 20 m$ と与えた3ケースを評価した。MMの正確度は、この設定に加えて、difficultiesの有無や、道路密度等によって異なると考えられる。これらの平均的な傾向をメタ分析によって把握する。メタ分析のためには、回帰分析と決定木を用いる。

#### (2) MMのための設定

FMMおよび瀬木ら<sup>9)</sup>の実装において、計算負荷軽減のために、「最短経路探索のためのGNSS軌跡からの距離上限 $\Delta$ 」を設定する。計算負荷を考慮しつつ十分に大きい値として、本研究では $\Delta=350m$ と設定した。またFMMでは、 $\Delta$ を設定した上でさらなる計算負荷軽減のために、「 $k$ 個の最短経路」のみをMMの対象とする。本研究では、この設定を行っていない瀬木ら<sup>9)</sup>とあわせて、 $k$ を十分に大きい値( $k=10000$ )に設定することとした。

#### (3) メタ分析における説明変数の選定

被説明関数を100種類のMM結果(各正確度評価指標値)とし、メタ分析を行う。説明変数としては、(1)節で述べた出力確率における標準偏差パラメータ( $\sigma = 5, 10, 20 m$  (5mがベースライン))、手法(FMM or HMM-logit (FMMがベースライン))、データセットの概要説明より、mad (3 (1)) および difficulties

ダミー (u\_tums, hives, loops, gaps, severe congruence issues (congruence) (それぞれ No がベースライン)) を用いる。加えて、道路ネットワーク特性として、2 種類の指標を作成した。

#### a) 道路密度

MM において、道路ネットワークの密度による正確度の変化をみるために道路密度を計算した。方法としては、対象の道路ネットワークにおいて、各正解経路から半径 300m のバッファを発生させ、そのバッファ内の総面積 (m<sup>2</sup>) と総リンク長 (m) を GIS を用いて算出する。そして、総リンク長を総面積で割った値を本研究における道路密度 (m/m<sup>2</sup>) と定義し、説明変数の一つとして用いる。

#### b) 道路ネットワークの中心性指標

ネットワークの疎密や形状などの特徴を捉える指標として中心性があげられる。その中で、本研究では道路ネットワーク内でノードに接続しているリンクの数を算出し、単純により多くのリンクとつながっているノードを高く評価する次数中心性を用いる (固有ベクトル中心性についても検討したが、結果は本質的に変わらなかった)。次数中心性の集中度からネットワークの密度のばらつきを捉え、最大値からネットワークの疎密を捉えることができるため、次数中心性の集中度 (degree centralization (centralization)) と最大値 (maximum degree centrality (degree\_max)) を説明変数として格納した。なお、次数中心性の集中度は、ネットワークの次数分布から中心性の集中度を測る指標であり、次数中心性の最大値から各ノードの次数中心性を引いた値の総和を 0-1 になるように基準化したものである。1 に近いほど、少数のノードに集中していることを意味する。

### (4) 結果と考察

#### a) メタ分析の方針

以下では、まず b), c) に、回帰分析によるメタ分析結果を示す。b) で全体的な傾向をつかみ (表-2~4), その後 c) で交差項を用いたより詳細な分析を行う (表-5~22)。続いて d), e) では、それぞれ hives (図-3~5) と mad (図-6~8) に着目した分析を行う。最後に、f) で、説明変数の非線形効果を捉えるために、決定木を用いた分析を行う (図-9~11)。

#### b) 回帰分析を用いたメタ分析

表-2 に Match ratio, 表-3 に Undeviating ratio, 表-4 に Similarity ratio をそれぞれ被説明変数とした回帰分析結果を示す (N=100)。

まず、表-2 より、道路密度 (road\_density) や次数中心性の集中度 (centralization), 最大値 (degree\_max) が高いことが Match ratio を押し下げることが分かる。HMM-logit を使用した場合は FMM を使用した場合と比較して

表-2 Match ratio

	Estimate	t value	Pr(> t )
(Intercept)	1.00935917	383.59313	0
u_tums	-0.0008048	-0.45657	0.64815
hives	0.00965356	1.96541	0.04984
loops	0.0015479	0.91273	0.36176
gaps	-0.000593	-0.37651	0.70668
congruence	-0.0030136	-1.65214	0.09905
mad	-0.0001722	-0.53671	0.59167
road_density	-0.3122094	-3.98748	0.00008
centralization	-1.5976488	-2.98026	0.00300
degree_max	-1.31E-10	-9.45760	0
sigma=10	-0.0009303	-0.56498	0.57231
sigma=20	-0.0042292	-2.56844	0.01046
logit	-0.0036154	-2.69657	0.00721
自由度修正済み決定係数: 0.18170			

表-3 Undeviating ratio

	Estimate	t value	Pr(> t )
(Intercept)	1.01101814	105.27510	0
u_tums	0.00027651	0.04298	0.96573
hives	-0.07605842	-4.24283	0.00003
loops	0.00694887	1.12268	0.26204
gaps	-0.00618439	-1.07583	0.28245
congruence	-0.06050538	-9.08871	0.00000
mad	-0.00209785	-1.79142	0.07375
road_density	-0.92529667	-3.23798	0.00127
centralization	-4.42114344	-2.25969	0.02421
degree_max	-7.21E-11	-1.42876	0.15361
sigma=10	0.00103295	0.17188	0.86359
sigma=20	0.00538983	0.89687	0.37016
logit	0.01869076	3.81961	0.00015
自由度修正済み決定係数: 0.26372			

Match ratio が 0.36% 低くなる。このことから、正解経路を網羅的に特定したい場合、FMM を使用した方が良いということが分かる。

続いて、Undeviating ratio に関する分析結果を表-3 に示す。road\_density や centralization, mad が負の影響を与えていることが分かる。Match ratio とは逆で、HMM-logit は、

表-4 Similarity ratio

	Estimate	t value	Pr(> t )
(Intercept)	1.02038	94.74153	0
u_turns	-0.00053	-0.07322	0.94165
hives	-0.06640	-3.30309	0.00102
loops	0.00850	1.22408	0.22142
gaps	-0.00678	-1.05130	0.29356
congruence	-0.06352	-8.50794	0
mad	-0.00227	-1.72851	0.08443
road_density	-1.23751	-3.86148	0.00013
centralization	-6.01879	-2.74306	0.00627
degree_max	-2.03E-10	-3.58466	0.00037
sigma=10	0.00010	0.01523	0.98785
sigma=20	0.00116	0.17221	0.86333
logit	0.01508	2.74708	0.00620
自由度修正済み決定係数:0.24124			

FMM と比較して相対的に高い Undeviating ratio を与えている。すなわち、HMM-logit は誤った経路を特定することが FMM と比べて少ないと言える。また、Similarity ratio の分析結果を表-4 に示す。これは Matchratio と Undeviating ratio の総合的な指標である。road\_density や centralization, 次数中心性の最大値 (degree\_max), mad が大きい場合に Similarity ratio は低くなる傾向を意味する。この指標においても、HMM-logit を使用した場合は FMM を使用した場合と比較して評価値は高くなる傾向がある。すなわち、総合的な評価指標においても、HMM-logit は FMM よりも正確度が平均的に見て高いと解釈できる。

### c) 交差項を用いた分析

各正確度評価指標の結果において、説明変数の影響が HMM-logit を用いた場合と FMM を用いた場合や、sigma の設定値でどう変化するかをみるために、交差項を用いて回帰分析を行った。説明変数 hives と logit の交差項を含む回帰の結果を表-5 と表-6 に、congruence と logit の交差項を含む結果を表-7 と表-8 に、mad と logit の交差項を含む結果を表-9 と表-10 に、road\_density と logit の交差項を含む結果を表-11 と表-12 に、mad と sigma の交差項を含む結果を表-13 と表-14 に示す。

加えて、HMM-logit と FMM では説明変数の影響が異なる可能性があるため、それぞれにおける mad と sigma の交差項を含む結果を表-15 から表-18 に、mad の大きさを分位点をもとに 4 分割した説明変数 mad\_4cut と sigma の交差項を含む結果を表-19 から表-22 に示す。なお、表において“変数 A:変数 B”は、変数 A と変数 B の交差

項を示している。

まず、hives と HMM-logit における交互作用の分析結果を表-5 と表-6 に示す。観測点が密になっている箇所である hives があることは、Undeviating ratio, Similarity ratio に負の影響を及ぼしているが、HMM-logit を使用した場合、この負の影響が緩和されていることが確認できた。次に表-7 と表-8 に、congruence と HMM-logit における交互作用の分析結果を示す。hives の場合と同様の傾向が確認できる。表-9 と表-10 に結果を示した mad の場合も同様である。

続いて表-11, 表-12 に、road\_density と HMM-logit における交互作用の分析結果を示す。road\_density が大きいサンプルは小さいサンプルと比較して、Undeviating ratio, Similarity ratio が悪化する傾向にある。一方で、road\_density が大きい場合に HMM-logit を使用した場合、FMM を使用した場合と比較して、Undeviating ratio, Similarity ratio への負の影響が緩和されていることが確認できる。

最後に、mad と sigma における交互作用の分析結果を表-13, 表-14 に示す。mad の大きさと sigma の設定値は、Undeviating ratio, Similarity ratio の両方に関係があることが分かった。また、mad の大きさが大きくなると sigma の設定値が大きい方が、観測点とネットワークの誤差による負の影響をやや小さくできる可能性があるといえる。

ここで、表-15~表-18 のように HMM-logit と FMM ごとに Undeviating ratio, Similarity ratio の結果をメタ分析してみる。すると、FMM では sigma の設定値が正確度に影響する一方で、HMM-logit ではほとんど影響がないことが分かる。また、観測点とネットワークの誤差の大きさと sigma の設定値の関係は、FMM には影響するが、HMM-logit にはほとんど影響しないといえる。ただし、本データセットは、3 (1) で述べた通り位置正確度が比較的高い GPS からなるものであるため、この分析結果の一般性については、より詳細な検討が必要である。

さらに、表-19~表-22 のように mad の大きさを分位点をもとに 4 分割して分析したところ、FMM は観測点とネットワークの誤差の大きさに対して sigma の値を適切に指定すると Undeviating ratio, Similarity ratio に正の影響を与えられる可能性を示している。一方で HMM-logit は sigma の設定値に対してロバストであり、実用上使いやすいと考えられる。

表-5 Undeviating ratio (hives × logit)

	Estimate	t value	Pr(> t )
(Intercept)	1.01436	111.22834	0
u_tums	0.00027	0.04425	0.96472
hives	-0.18567	-8.54360	1.15E-16
loops	0.006931	1.18052	0.23828
gaps	-0.00619	-1.13471	0.25697
congruence	-0.06053	-9.58536	2.67E-20
mad	-0.00210	-1.89058	0.05918
road_density	-0.92578	-3.41509	0.00068
centralization	-4.43211	-2.38795	0.01726
degree_max	-7.25E-11	-1.51415	0.13053
sigma=10	0.001068	0.18731	0.85148
sigma=20	0.005425	0.95155	0.34172
logit	0.012043	2.55474	0.01088
hives : logit	0.219355	8.10067	3.24E-15
自由度修正済み決定係数 : 0.33742			

表-7 Undeviating ratio (congruence × logit)

	Estimate	t value	Pr(> t )
(Intercept)	1.01654	106.99637	0
u_tums	0.00032	0.05106	0.95930
hives	-0.07623	-4.33100	1.75E-05
loops	0.00710	1.16834	0.24315
gaps	-0.00620	-1.09792	0.27269
congruence	-0.08901	-10.03938	5.64E-22
mad	-0.00211	-1.83179	0.06749
road_density	-0.92058	-3.28092	0.00110
centralization	-4.36866	-2.27404	0.02333
degree_max	-7.22E-11	-1.45612	0.14590
sigma=10	0.00094	0.15968	0.87319
sigma=20	0.00530	0.89804	0.36954
logit	0.00737	1.37430	0.16988
congruence: logit	0.05749	4.75865	2.46E-06
自由度修正済み決定係数 : 0.29017			

表-6 Similarity ratio (hives × logit)

	Estimate	t value	Pr(> t )
(Intercept)	1.02368	98.85046	0
u_tums	-0.00053	-0.07715	0.93853
hives	-0.17473	-7.08037	4.17E-12
loops	0.00848	1.27178	0.20396
gaps	-0.00678	-1.09501	0.27397
congruence	-0.06355	-8.86129	9.66E-18
mad	-0.00227	-1.80139	0.07216
road_density	-1.23798	-4.02161	6.54E-05
centralization	-6.02963	-2.86086	0.00438
degree_max	-2.03E-10	-3.73887	0.00020
sigma=10	0.00014	0.02118	0.98311
sigma=20	0.00120	0.18461	0.85360
logit	0.00851	1.58895	0.11262
hives : logit	0.21678	7.04989	5.11E-12
自由度修正済み決定係数 : 0.29992			

表-8 Similarity ratio (congruence × logit)

	Estimate	t value	Pr(> t )
(Intercept)	1.02549	95.65397	0
u_tums	-0.00049	-0.06814	0.94570
hives	-0.06657	-3.35141	0.00086
loops	0.00864	1.25944	0.20838
gaps	-0.00679	-1.06593	0.28690
congruence	-0.08990	-8.98626	3.58E-18
mad	-0.00228	-1.75556	0.07969
road_density	-1.23314	-3.89470	0.00011
centralization	-5.97021	-2.75402	0.00607
degree_max	-2.03E-10	-3.62913	0.00031
sigma=10	1.87E-05	0.00280	0.99777
sigma=20	0.00108	0.16170	0.87160
logit	0.00459	0.75941	0.44792
congruence: logit	0.05322	3.90362	0.00011
自由度修正済み決定係数 : 0.25939			

表-9 Undeviating ratio (mad - logit)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.02285	107.14335	0
u_tums	0.00026	0.04227	0.96630
hives	-0.07617	-4.37147	1.46E-05
loops	0.00689	1.14594	0.25229
gaps	-0.00616	-1.10282	0.27056
congruence	-0.06059	-9.36350	1.68E-19
mad	-0.00761	-5.17321	3.17E-07
road_density	-0.92721	-3.33819	0.00090
centralization	-4.42156	-2.32504	0.02041
degree_max	-7.07E-11	-1.43993	0.15043
sigma=10	0.00097	0.16682	0.86757
sigma=20	0.00533	0.91270	0.36178
logit	-0.00492	-0.79285	0.42819
mad:logit	0.01106	5.91395	5.71E-09
自由度修正済み決定係数:0.30440			

表-11 Undeviating ratio (road\_density × logit)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.02189	96.76439	0
u_tums	0.00029	0.04456	0.96448
hives	-0.07610	-4.26304	2.35E-05
loops	0.00698	1.13205	0.25808
gaps	-0.00619	-1.08060	0.28033
congruence	-0.06046	-9.11996	1.22E-18
mad	-0.00210	-1.79971	0.07243
road_density	-1.54576	-4.04135	6.03E-05
centralization	-4.41019	-2.26356	0.02397
degree_max	-7.21E-11	-1.43334	0.15230
sigma=10	0.00101	0.16889	0.86594
sigma=20	0.00537	0.89693	0.37013
logit	-0.00312	-0.30490	0.76055
road_density:logit	1.24268	2.42774	0.01550
自由度修正済み決定係数:0.26987			

表-10 Similarity ratio (mad × logit)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.03264	96.02972	0
u_tums	-0.00054	-0.07680	0.93881
hives	-0.06652	-3.38924	0.00075
loops	0.00844	1.24547	0.21346
gaps	-0.00675	-1.07315	0.28365
congruence	-0.06361	-8.72659	2.78E-17
mad	-0.00798	-4.81685	1.86E-06
road_density	-1.23949	-3.96169	8.37E-05
centralization	-6.01922	-2.80996	0.00512
degree_max	-2.01E-10	-3.64419	0.00029
sigma=10	4.20E-05	0.00639	0.99491
sigma=20	0.00110	0.16719	0.86728
logit	-0.00939	-1.34276	0.17987
mad:logit	0.01146	5.44009	7.86E-08
自由度修正済み決定係数:0.27683			

表-12 Similarity ratio (road\_density × logit)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.02984	86.77925	0
u_tums	-0.00052	-0.07230	0.94239
hives	-0.06644	-3.31213	0.00098
loops	0.00852	1.23034	0.21907
gaps	-0.00678	-1.05377	0.29243
congruence	-0.06348	-8.52109	1.37E-16
mad	-0.00227	-1.73287	0.08365
road_density	-1.77736	-4.13522	4.07E-05
centralization	-6.00926	-2.74468	0.00624
degree_max	-2.03E-10	-3.59137	0.00036
sigma=10	8.33E-05	0.01239	0.99012
sigma=20	0.00114	0.16971	0.86529
logit	-0.00390	-0.33948	0.73437
road_density:logit	1.08125	1.87977	0.06064
自由度修正済み決定係数:0.24453			

表-13 Undeviating ratio (mad × sigma)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.00423	100.38419	0
u_tums	0.00026	0.04087	0.96742
hives	-0.07614	-4.26005	2.39E-05
loops	0.00691	1.12031	0.26305
gaps	-0.00616	-1.07570	0.28251
congruence	-0.06054	-9.12076	1.22E-18
mad	0.00109	0.60863	0.54301
road_density	-0.92668	-3.25265	0.00121
centralization	-4.42169	-2.26682	0.02377
degree_max	-7.13E-11	-1.41676	0.15709
logit	0.01871	3.83567	0.00014
sigma=10	0.01171	1.49915	0.13438
sigma=20	0.01504	1.92519	0.05470
mad: sigma=10	-0.00501	-2.13189	0.03344
mad: sigma=20	-0.00453	-1.92744	0.05441
自由度修正済み決定係数: 0.26817			

表-15 HMM-logit Undeviating ratio (mad × sigma)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.01696	107.34882	1.00E-230
u_tums	0.00202	0.32295	0.74697
hives	0.01214	0.69507	0.48758
loops	0.00046	0.07587	0.93958
gaps	-0.00544	-0.97232	0.33173
congruence	-0.04516	-6.94038	2.69E-11
mad	-6.78E-05	-0.03854	0.96928
road_density	-0.44847	-1.61103	0.10829
centralization	-4.37468	-2.28989	0.02277
degree_max	-1.09E-10	-2.20682	0.02813
sigma=10	-0.00222	-0.29094	0.77131
sigma=20	-0.00208	-0.27285	0.78517
mad: sigma=10	-0.00023	-0.10207	0.91877
mad: sigma=20	-8.01E-05	-0.03483	0.97224
自由度修正済み決定係数: 0.15059			

表-14 Similarity ratio (mad × sigma)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.01326	90.26322	0
u_tums	-0.00054	-0.07550	0.93985
hives	-0.06649	-3.31521	0.00097
loops	0.00846	1.22166	0.22233
gaps	-0.00676	-1.05069	0.29384
congruence	-0.06355	-8.53287	1.25E-16
mad	0.00107	0.53227	0.59474
road_density	-1.23896	-3.87542	0.00012
centralization	-6.01937	-2.75001	0.00615
degree_max	-2.02E-10	-3.57810	0.00038
logit	0.01510	2.75796	0.00600
sigma=10	0.01140	1.30055	0.19393
sigma=20	0.01116	1.27295	0.20355
mad: sigma=10	-0.00530	-2.01001	0.04489
mad: sigma=20	-0.00469	-1.77973	0.07565
自由度修正済み決定係数: 0.24492			

表-16 FMM Undeviating ratio (mad × sigma)

	Estimate	tvalue	Pr(> t )
(Intercept)	1.00705	66.41532	3.17E-174
u_tums	-0.00068	-0.06770	0.94608
hives	-0.17132	-6.12770	0.00000
loops	0.01572	1.62719	0.10481
gaps	-0.00665	-0.74235	0.45850
congruence	-0.07207	-6.91990	3.04E-11
mad	0.00263	0.93500	0.35059
road_density	-1.33683	-3.00030	0.00294
centralization	-3.20533	-1.04825	0.29542
degree_max	1.05E-12	0.01331	0.98939
sigma=10	0.02268	1.85546	0.06457
sigma=20	0.02920	2.38868	0.01757
mad: sigma=10	-0.01010	-2.74268	0.00648
mad: sigma=20	-0.00929	-2.52359	0.01217
自由度修正済み決定係数: 0.40497			

表-17 HMM-logit Similarity ratio (mad × sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.02547	85.38350	1.58E-203
u_tums	-0.00020	-0.02486	0.98019
hives	0.01814	0.81935	0.41328
loops	0.00198	0.25905	0.79579
gaps	-0.00691	-0.97425	0.33077
congruence	-0.04976	-6.03179	5.08E-09
mad	0.00044	0.19732	0.84372
road_density	-0.82487	-2.33725	0.02013
centralization	-5.84505	-2.41330	0.01645
degree_max	-2.27E-10	-3.62179	0.00035
sigma=10	-0.00366	-0.37776	0.70589
sigma=20	-0.01076	-1.11112	0.26747
mad: sigma=10	-0.00082	-0.28017	0.77956
mad: sigma=20	-0.00045	-0.15587	0.87625
自由度修正済み決定係数: 0.15349			

表-18 FMM Similarity ratio (mad × sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.01148	62.45735	3.53E-167
u_tums	8.42E-05	0.00787	0.99373
hives	-0.16159	-5.41155	1.34E-07
loops	0.01752	1.69823	0.09057
gaps	-0.00604	-0.63135	0.52832
congruence	-0.07305	-6.56723	2.46E-10
mad	0.00270	0.89910	0.36937
road_density	-1.58167	-3.32365	0.00101
centralization	-4.49663	-1.37685	0.16965
degree_max	-1.15E-10	-1.35950	0.17507
sigma=10	0.02288	1.75259	0.08076
sigma=20	0.02950	2.25940	0.02462
mad: sigma=10	-0.01069	-2.71789	0.00698
mad: sigma=20	-0.00983	-2.50107	0.01295
自由度修正済み決定係数: 0.38058			

表-19 HMM-logit Undeviating ratio  
(mad\_4cut × sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.01658	89.21259	1.49E-205
u_tums	0.00370	0.57552	0.56541
hives	0.01987	1.30132	0.19423
loops	-0.00134	-0.22040	0.82572
gaps	-0.00279	-0.48755	0.62626
congruence	-0.04418	-6.77954	7.30E-11
road_density	-0.46022	-1.65839	0.09838
centralization	-4.91566	-2.53368	0.01184
degree_max	-1.35E-10	-2.58607	0.01022
sigma=10	-0.00751	-0.64548	0.51915
sigma=20	-0.00510	-0.43811	0.66165
mad_2	0.00446	0.36702	0.71389
mad_3	0.00362	0.30659	0.75939
mad_4	-0.00898	-0.74693	0.45574
mad_2: sigma=10	0.01185	0.72365	0.46989
mad_2: sigma=20	0.00831	0.50773	0.61205
mad_3: sigma=10	0.00944	0.57361	0.56670
mad_3: sigma=20	0.00158	0.09628	0.92337
mad_4: sigma=10	-0.00154	-0.09293	0.92603
mad_4: sigma=20	0.00208	0.12540	0.90030
自由度修正済み決定係数: 0.16389			

## 各正確度評価指標と hives との関係

ダミー変数 hives が 1 であった 3 つのケース (この 3 ケースしかない) について、各正確度評価指標に外生標準偏差がどのように影響するのかを調べた。図-3 に Match ratio, 図-4 に Undeviating ratio, 図-5 に Similarity ratio の結果を示している。

観測点が密な箇所(hives)が該当する 3 つのケースで sigma が同じ値の場合に注目する。図-3 より Match ratio は HMM-logit を用いた場合よりも FMM を用いた方が正確度が高く、一方で図-4, 図-5 より Undeviating, Similarity ratio は FMM を用いた場合よりも HMM-logit を用いた方が正確度が高いことがわかる。これらのことから、hives がある場合でも FMM は正解経路を特定することに長けており、HMM-logit は余分な経路を特定する割合が少ないということが言える。

表-20 FMM Undeviating ratio  
(mad=4cut × sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.03027	55.57720	6.67E-152
u_tums	-0.00248	-0.23760	0.81237
hives	-0.18886	-7.60449	4.51E-13
loops	0.01650	1.67181	0.09570
gaps	-0.00431	-0.46286	0.64383
congruence	-0.07063	-6.66187	1.46E-10
road_density	-1.31940	-2.92253	0.00376
centralization	-2.43756	-0.77231	0.44059
degree_max	-3.25E-11	-0.38216	0.70264
sigma=10	-0.02003	-1.05783	0.29106
sigma=20	-0.01934	-1.02132	0.30799
mad_2	-0.01313	-0.66440	0.50699
mad_3	-0.01589	-0.82639	0.40930
mad_4	-0.04925	-2.51884	0.01234
mad_2: sigma=10	0.02566	0.96335	0.33622
mad_2: sigma=20	0.02951	1.10772	0.26895
mad_3: sigma=10	0.03320	1.23987	0.21608
mad_3: sigma=20	0.04152	1.55050	0.12217
mad_4: sigma=10	0.02731	1.01415	0.31140
mad_4: sigma=20	0.04536	1.68441	0.09323
自由度修正済み決定係数: 0.39493			

表-21 HMM-logit Similarity ratio  
(mad\_4cut - sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.02526	71.27043	9.20E-180
u_tums	0.00292	0.35941	0.71956
hives	0.03135	1.62691	0.10490
loops	-0.00120	-0.15653	0.87573
gaps	-0.00250	-0.34579	0.72976
congruence	-0.04841	-5.88395	1.16E-08
road_density	-0.84914	-2.42375	0.01600
centralization	-6.96832	-2.84504	0.00477
degree_max	-2.72E-10	-4.13003	4.81E-05
sigma=10	-0.01014	-0.69002	0.49076
sigma=20	-0.01345	-0.91543	0.36076
mad_2	0.00994	0.64795	0.51755
mad_3	0.00373	0.24977	0.80295
mad_4	-0.00993	-0.65412	0.51358
mad_2: sigma=10	0.01351	0.65378	0.51380
mad_2: sigma=20	0.00995	0.48132	0.63067
mad_3: sigma=10	0.01153	0.55502	0.57933
mad_3: sigma=20	-0.00081	-0.03900	0.96892
mad_4: sigma=10	-0.00514	-0.24608	0.80580
mad_4: sigma=20	-0.00128	-0.06149	0.95101
自由度修正済み決定係数: 0.173756			

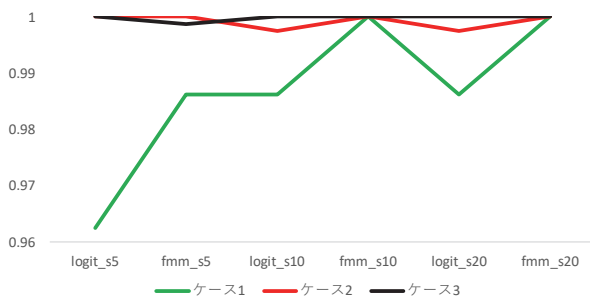


図-3 hives 有における Match ratio 結果

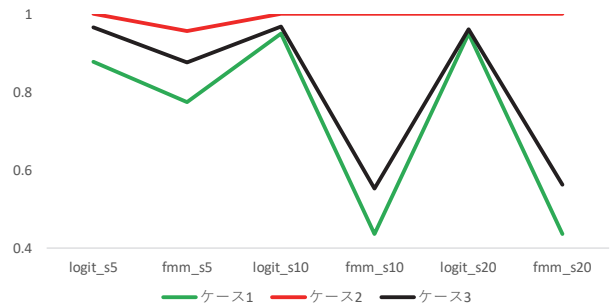


図-4 hives 有における Undeviating ratio 結果

表-22 FMM Similarity ratio  
(mad\_4cut - sigma)

	Estimate	t value	Pr(> t )
(Intercept)	1.03490	52.58395	7.75E-146
u_turns	-0.00073	-0.06595	0.94747
hives	-0.17730	-6.72418	1.01E-10
loops	0.01721	1.64276	0.10157
gaps	-0.00197	-0.19887	0.84251
congruence	-0.07127	-6.33197	9.78E-10
road_density	-1.57329	-3.28247	0.00116
centralization	-4.15392	-1.23965	0.21616
degree_max	-1.69E-10	-1.86841	0.06276
sigma=10	-0.02065	-1.02692	0.30536
sigma=20	-0.02073	-1.03105	0.30342
mad_2	-0.00813	-0.38767	0.69856
mad_3	-0.01790	-0.87702	0.38124
mad_4	-0.05195	-2.50263	0.01291
mad_2 : sigma=10	0.02654	0.93851	0.34880
mad_2 : sigma=20	0.03100	1.09600	0.27403
mad_3 : sigma=10	0.03511	1.23477	0.21796
mad_3 : sigma=20	0.04520	1.58967	0.11305
mad_4 : sigma=10	0.02311	0.80844	0.41953
mad_4 : sigma=20	0.04267	1.49245	0.13672
自由度修正済み決定係数:0.37762			

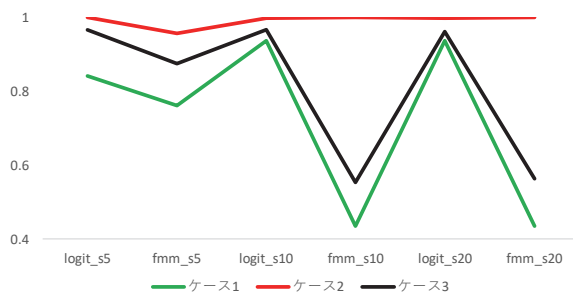


図-5 hives 有における Similarity ratio 結果

d) 各正確度評価指標に対する mad の影響

続いて、各正確度評価指標に対する mad の影響を調べた。図-6に Match ratio, 図-7に Undeviating ratio, 図-8に Similarity ratio の結果を示している。具体的には、mad が

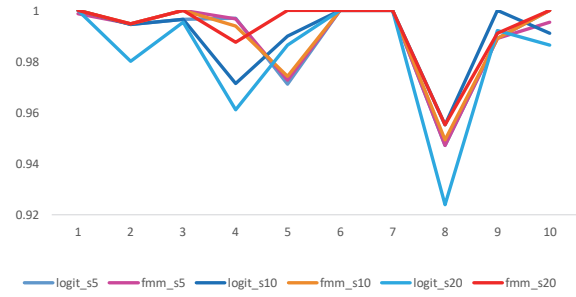


図-6 Match ratio (mad の影響)

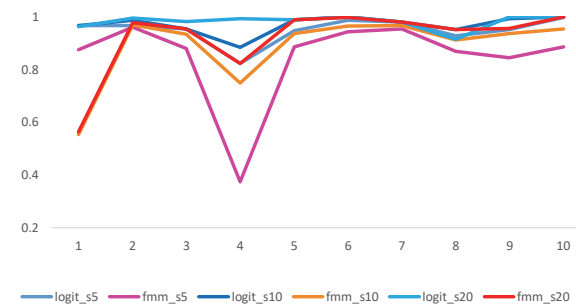


図-7 Undeviating ratio (mad の影響)

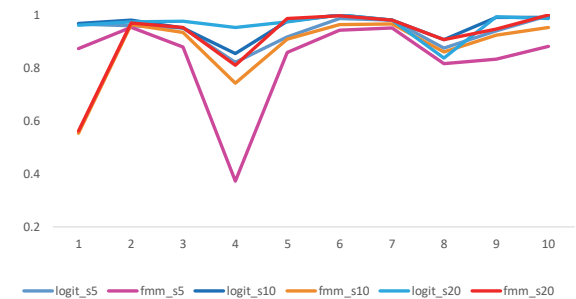


図-8 Similarity ratio (mad の影響)

ワーストの 10 ケースを取り出し、外生標準偏差パラメータとの関係を調べた。図-6 から図-8 の横軸は、当該 10 ケースを示しており、左側のほうが mad の値が大きい (GPS の位置正確度が低い)。図-6 より Match ratio において、HMM-logit を用いるよりも FMM を用いた方が正確度が高いことがわかる。一方で、図-7, 図-8 より Undeviating ratio, Similarity ratio においては、FMM を用いるよりも HMM-logit を用いた方が、正確度が高いことがわかる。観測点の位置正確度が低いデータであっても、FMM は正解経路を特定することには強みがあるが、余

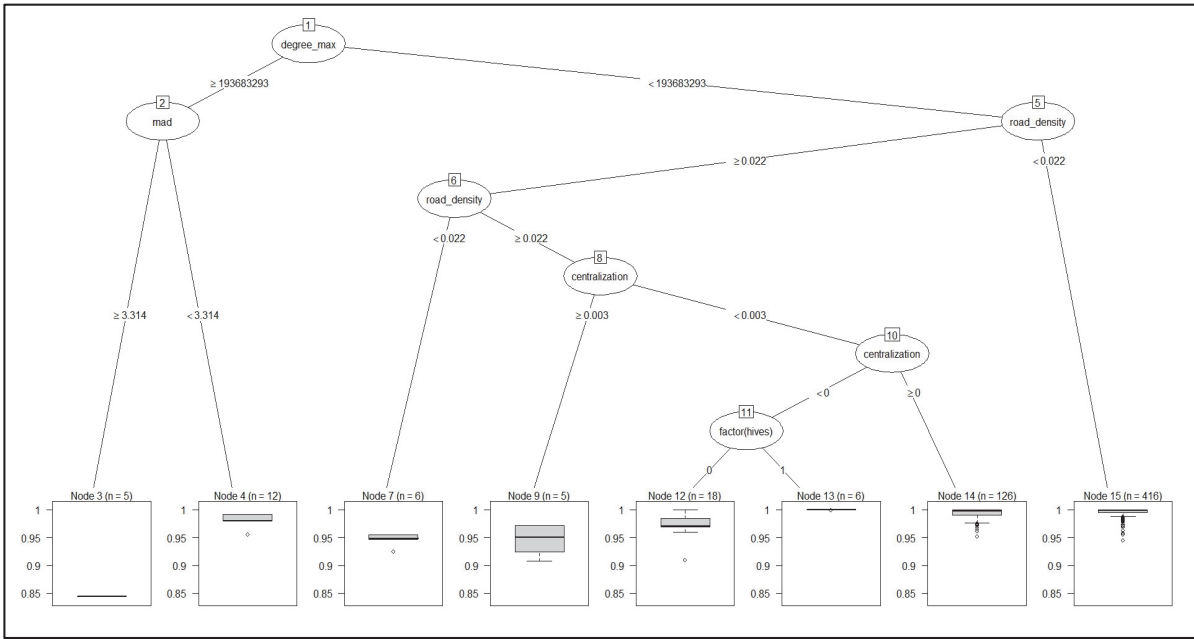


図-9 決定木分析結果 (Match Percentage)

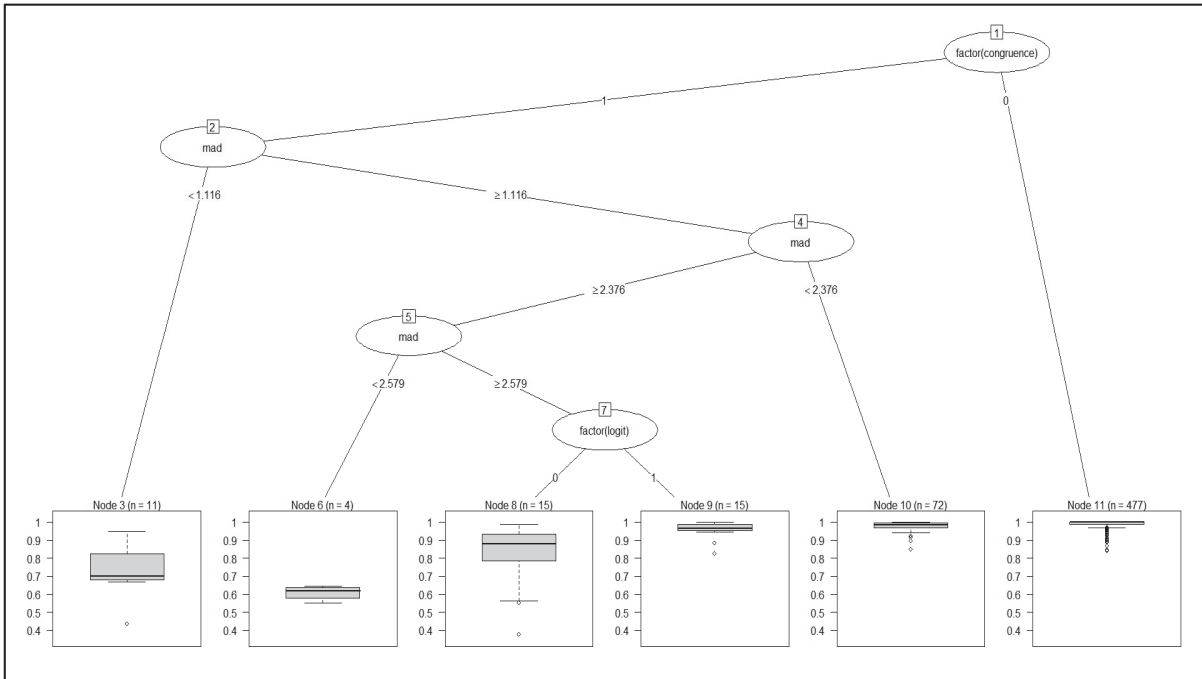


図-10 決定木分析結果 (Undeviating ratio)

分な経路を特定してしまうことが弱みであることが言える。

e) 決定木分析

各正確度評価指標に説明変数が与える非線形効果を調べるために決定木を用いて分析した。なお、分割数はクロスバリデーションによって決定している。図-9に Match ratio, 図-10に Undeviating ratio, 図-11に Similarity ratio の結果を示している。図-9 より Match ratio は, degree\_max の大きさや, road\_density の大きさによって大きく評価値が分かれており, その次に mad や centralization の大きさで分かれていることが分かる。次に図-10, 図-

11より Undeviating ratio と Similarity Percentage については, congruence の有無と mad の大きさにより評価指標値が大きく分かれていることが分かる。

4. おわりに

本研究では, 近年研究が蓄積されている隠れマルコフモデル (HMM) に基づくマップマッチング (MM) と, 瀬木ら<sup>9)</sup>の HMM に最短経路探索の考え方を取り入れた MM の経路推論のそれぞれの強みや特徴を分析し

た. 具体的には, 世界各国で取得された 100 種類の GPS データと真の走行ルートの情報を用いて, HMM に基づく MM 手法である FMM と, HMM における遷移確率 (状態空間モデルにおけるシステムモデル) に最短経路探索の考え方を取り入れた MM 手法 (HMM-logit) に対し, 道路特性や取得された位置情報データの特性がそれぞれの MM の経路推論の正確度にどのように影響するかについて, 回帰分析と決定木を用いて検証した.

結果として, 正解経路を漏れなく特定することに強いのは FMM, 特定経路に連続性があり無駄な経路の特定が少ないのは HMM-logit という傾向がみられた. また, 道路の密度や観測点の密度, さらに観測点と道路ネットワークの距離誤差が大きい場合では HMM-logit は FMM よりも正確度が良いことが確認された. さらに, 観測モデルの標準偏差パラメータ  $\sigma$  の設定に対し, HMM-logit は頑健であった.

今後の課題として, HMM-logit において正解経路の特定率をより向上させる必要がある. また, 2 次元的な MM だけでなく, 高度情報を盛り込んだ 3 次元的な MM 手法にも対応できるモデルに拡張していく必要があると考えられる.

## 謝辞

本研究は, 国土交通省による道路政策の質の向上に資する技術研究開発『ETC2.0 データの活用と評価を通じた次世代 ETC の基本設計提案』の助成を受けて実施されたものである.

## 参考文献

- 1) Quddus, M.A., Ochieng, W.Y. and Noland, R.B.: Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*, Vol.15, No.5, pp.312–318, 2007.
- 2) Newson, P. and Krumm, J.: Hidden Markov map matching through noise and sparseness, In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp.336–343, 2009.
- 3) Lou, Y., Zhang, C., Zheng, Y., Xie, X., Wang, W. and Huang, Y.: Map-matching for low-sampling-rate GPS trajectories, In Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp.352–361, 2009.
- 4) Jagadeesh, G.R. and Srikanthan, T.: Online map-matching of noisy and sparse location data with hidden Markov and route choice models, *IEEE Transactions on Intelligent Transportation Systems*, Vol.18, No.9, pp.2423–2434, 2017.
- 5) Qi, H., Di, X. and Li, J.: Map-matching algorithm based on the junction decision domain and the hidden Markov model, *PLoS*

*one*, Vol.14, No.5, e0216476, 2019.

- 6) Yang, C. and Gidofalvi, G.: Fast map matching, an algorithm integrating hidden Markov model with precomputation. *International Journal of Geographical Information Science*, Vol.32, No.3, pp.547–570, 2018.
- 7) 橋口剛, 瀬谷創, 安田昌平, 井料隆雅: 隠れマルコフモデルに基づくマップマッチングの適用性の検証, 土木学会論文集 D3 (土木計画学), Vol.77, No.5, pp.I 595–I 602, 2022.
- 8) 瀬木俊輔, 瀬谷創, 井料隆雅, 安田昌平: 状態空間モデルと recursive logit モデルを用いたマップマッチング, 広島大学セミナー発表資料, 2022年4月15日.
- 9) Oyama, Y. and Hato, E.: Link-based measurement model to estimate route choice parameters in urban pedestrian networks, *Transportation Research Part C: Emerging Technologies*, Vol.93, pp.62–78, 2018.
- 10) Kubička, M., Cela, A., Moulin, P., Mounier, H. and Niculescu, S.I.: Dataset for testing and training of map-matching algorithms, Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV), pp.1088–1093, 2015.
- 11) 岩館洗太, 和佐州洋, 有村博紀: 隠れマルコフモデルを用いた軌跡データに対するカーネルの設計, 第7回データ工学と情報マネジメントに関するフォーラム (<https://db-event.jpn.org/deim2015/paper/128.pdf>), 2015.