

野生動物GPSデータの事前処理： 外れ値以外の怪しいデータの除外方法

Techniques for exposing and removing potentially
problematic data in wildlife GPS data

平川 浩文

(無所属)

瀧井 暁子・泉山 茂之

(信州大学 山岳科学研究所)

村松 大輔

(京大WRC／奈教自然セ)

Marcelo Gordo

(UFAM／Brazil)

2022年8月26-29日 日本哺乳類学会69回大会

はじめに

P2

野生動物GPSデータにはさまざまな精度のデータが含まれる。システムの誤動作や不調に起因するとみられる異常データもある。一方、真の位置データも動物の移動に伴って大きく変化する。つまり、野生動物GPSデータは大きく変動する真の値に大きく変動する測定誤差が重なって構成されている。

的確な行動分析のためには、信頼度が低いデータ（以下、「怪しい」データと呼ぶ）をあらかじめ除外しておくことが望ましい。現在、開発中の野生動物GPSデータ解析システムはそのための仕組みを備えている。この仕組みではまず外れ値を除外する（その方法については注1，2参照）。ただ、信頼性の疑われるデータはまだ残る。本発表ではその除外処理について紹介する。

この処理ではDOP情報を利用せず、連続データ間の位置関係にのみに基づいて怪しいデータを抽出する。抽出したデータを順次地図や属性値の形で分析者に提示して各データ個別に除外可否の判断を求める形をとる。本手法は基本的に種を問わず利用可能である。一方、データが極めて特殊なナマケモノについては、別の抽出方法を用意し、あらかじめ定めた量を一括で除外できるようにした。

注1．日本生態学会69回大会ポスター発表：

<http://wildlife-science.info/publications/220315GAPmethod@ESJ.pdf>

注2．投稿中論文のプレプリント（まもなく改訂版に更新予定）

<https://www.researchsquare.com/article/rs-1567713/v1>

なぜDOPを使わないか

P3

GPSデータの測位精度を示す指標にDOP（Dilution Of Precision：精度低下率）がある。野生動物GPSの位置データにもDOPが必ず付記されている。我々はこれをデータの除外には利用しない。なぜなら、**DOPは真の測位精度ではない**から。

DOPは測位に用いたGPS衛星の数と配置から計算される。しかし、測定精度に影響する要因は他にもいくつかある。たとえば、**マルチパス（衛星からGPS受信機への電波の多重経路）**は分厚い林冠の下などでは大きな精度低下をもたらす。DOP以外の要因による精度低下は動物の動きに伴って状況が随時変化する野外環境下では把握できない。

結局、**DOPは精度の一部（計算可能部分）に過ぎない**。また一般に、精度の低さは必ずしも正確度の低さを意味しない。このため、DOPに頼っても精度の低いデータを十分除外することは期待できず、その一方で、問題の少ないデータを除外してしまう危険もある。

怪しいデータとは

P4

怪しい（信頼度が低い）データとは

たとえば、

位置関係から信頼度が疑われるもの

1) 前後の測位点から**突出**している

そうでなくとも、

分析者の専門的知見からみて信頼度が疑われるもの

2) 時間に対する距離（動き）が大きい

3) 地形や地況を考慮するとありそうもない

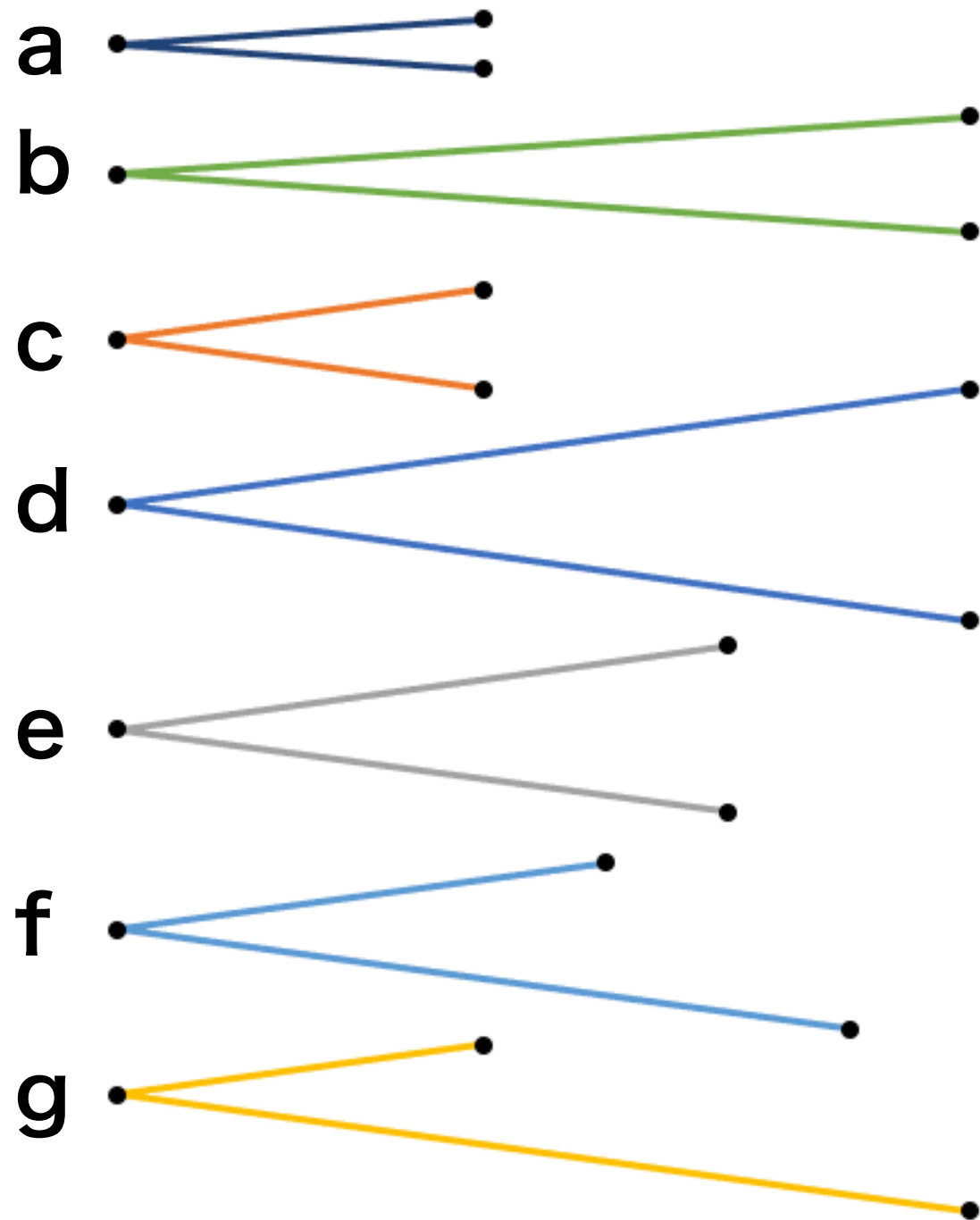
など

問題はこうしたデータをどう評価・抽出して除外するか

位置関係の怪しさ（突出）とは

P5

連続3点の多様な位置関係



位置関係の属性

2辺間の角度

$$(a, b) < (c, d, e, f, g)$$

2辺の長さの和

$$(a, c) < (e, f, g) < (b, d)$$

2辺の長さの差

$$(a, b, c, d, e) < f < g$$

両端2点間の距離

$$a < c < b < e < d < f < g$$

突出をどう定義？

突出をどう定義？

P6

例えば、各属性に条件を設ける

問題

- 1) 動物の種別に条件設定異なる
- 2) 閾値をどう定めるか
- 3) 条件の組み合わせをどうするか
和集合？ 積集合？

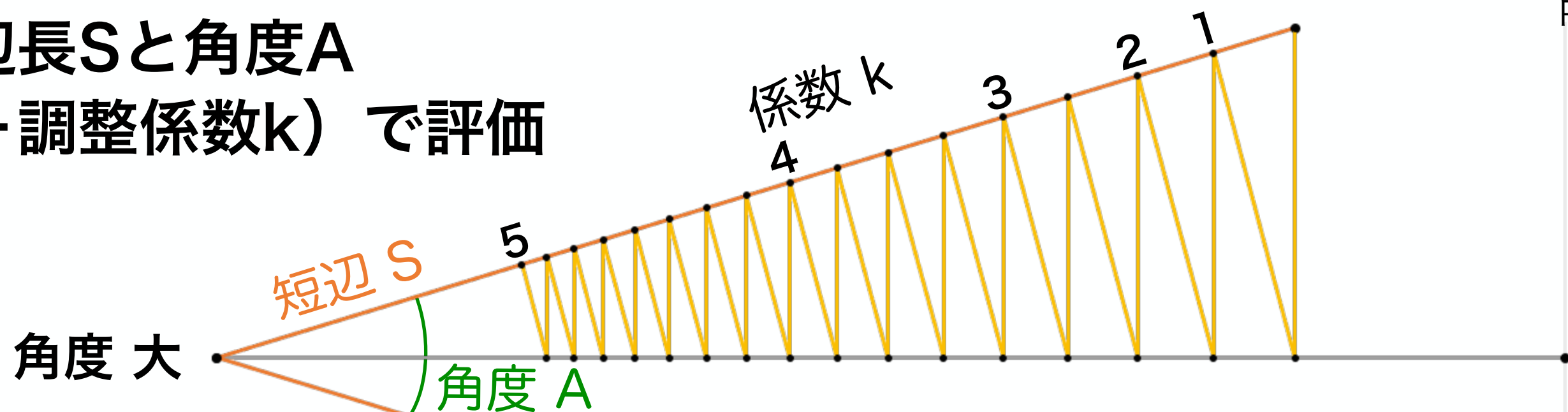
煩雑で非現実的！

突出を長さと角度で相対評価！

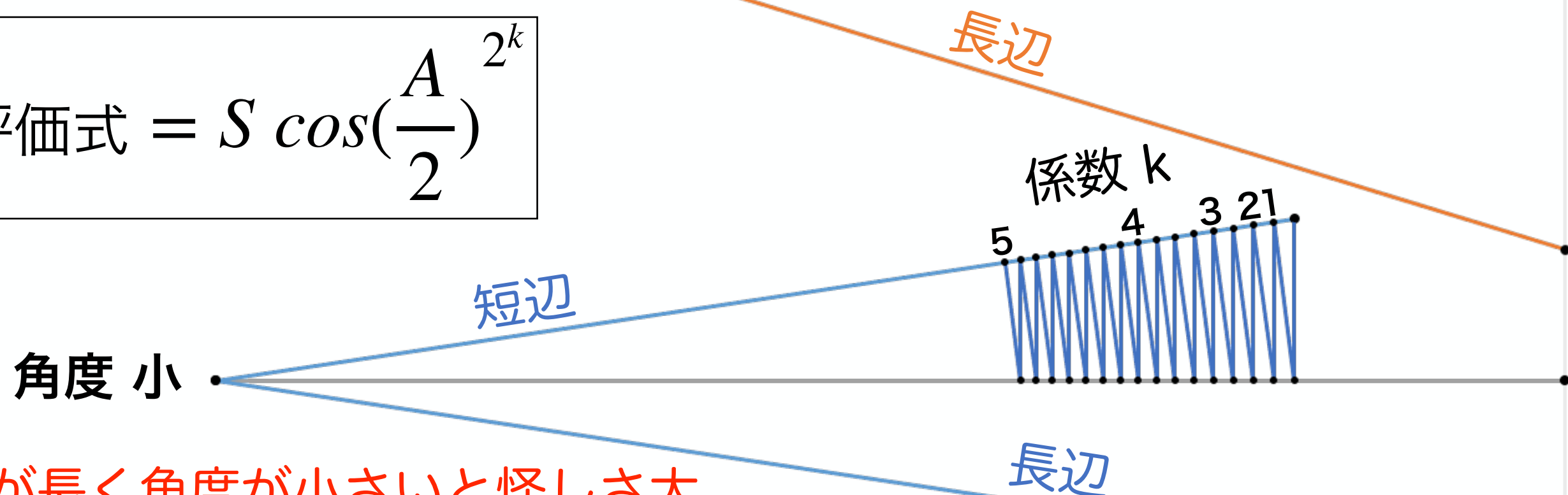
評価の大きい方から順次、測位点を地図表示、属性値も表示
分析者がそれらに基づいて一つ一つ除外の可否を判断
必要時にGoogle Earth表示(kmlファイルの作成)も可能に
そうすると専門的知見による怪しさの判断も併せて可能

突出の評価

短辺長Sと角度A
(+調整係数k) で評価



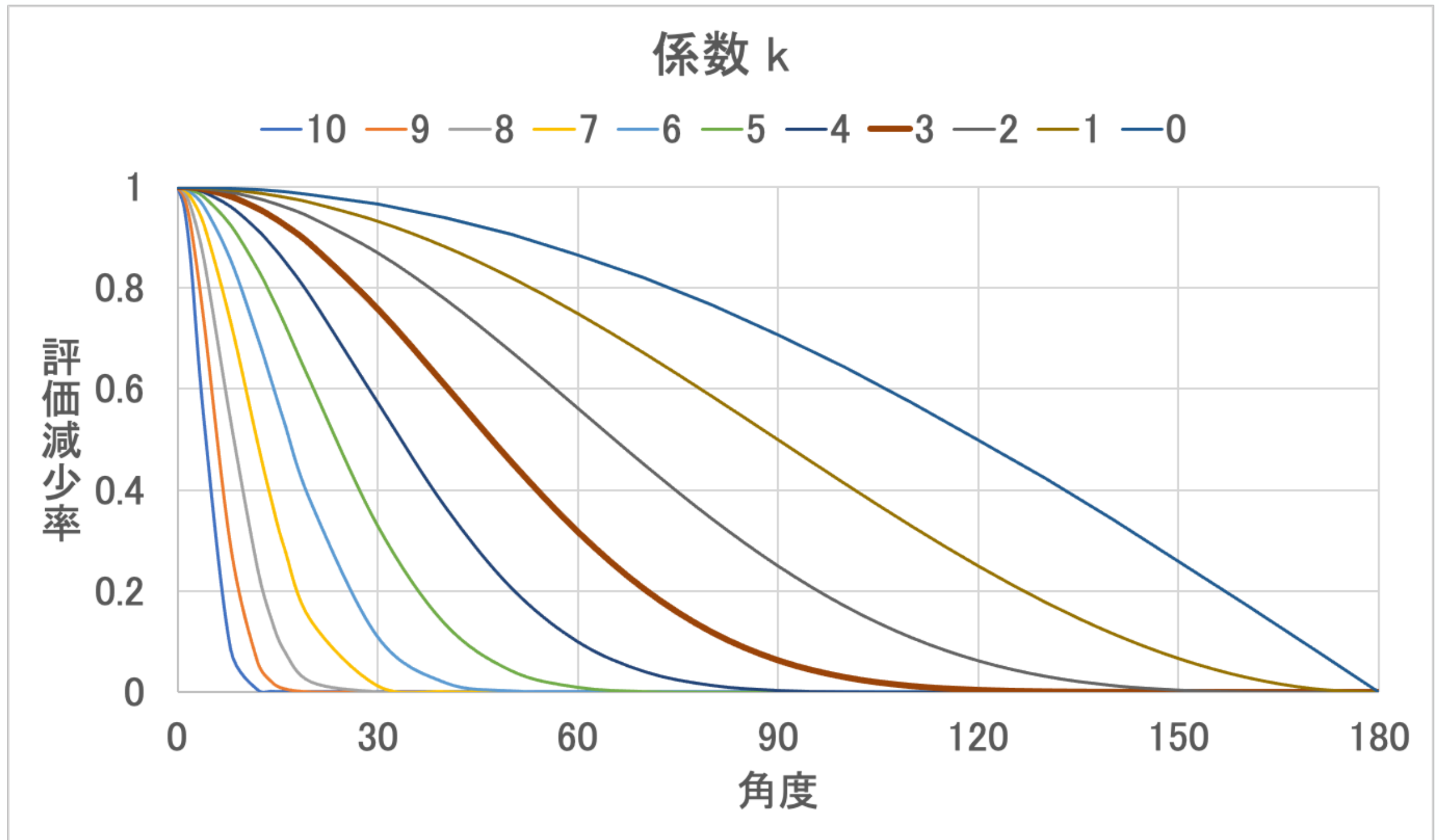
$$\text{評価式} = S \cos\left(\frac{A}{2}\right)^{2^k}$$



短辺が長く角度が小さいと怪しさ大
係数kで角度の重み調整；kが大きいと角度をより重視

P8

P8



ノイズ除外の方法

目的：測定誤差の大きいデータ（ノイズ）を抽出
分析者が選択して除外

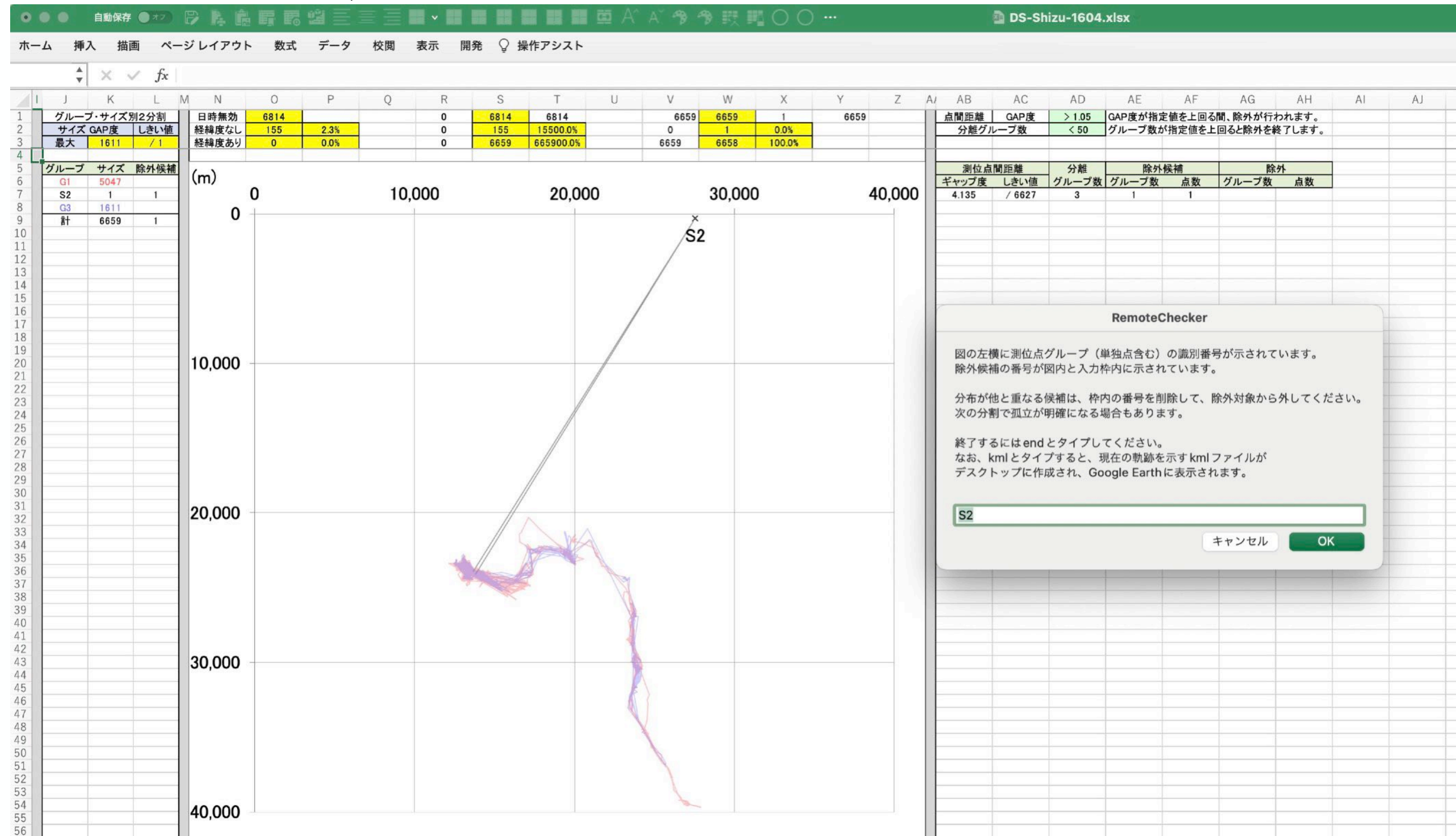
手続き：

- 1) 測定誤差を各測位点の**突出度**として推定
- 2) **突出度**の高い方から指定数 α 個の測位点を
地図に表示
その属性も表示
- 3) 分析者が各データ個別に除外可否を指示
(あるいは処理を終了)
- 4) (1)に戻る

α は適宜指定。実施例は $\alpha = 5$

GAP法による外れ値除外 (前処理)

ニホンジカShizu : 2016/4/19 - 2018/4/17 (728日間 : 2時間間隔)
データ数 : 6814; 1回に表示する数 : 5



除外例：初回

ニホンジカShizu：2016/4/19 - 2018/4/17 (728日間：2時間間隔)
データ数：6814; 1回に表示する数：5

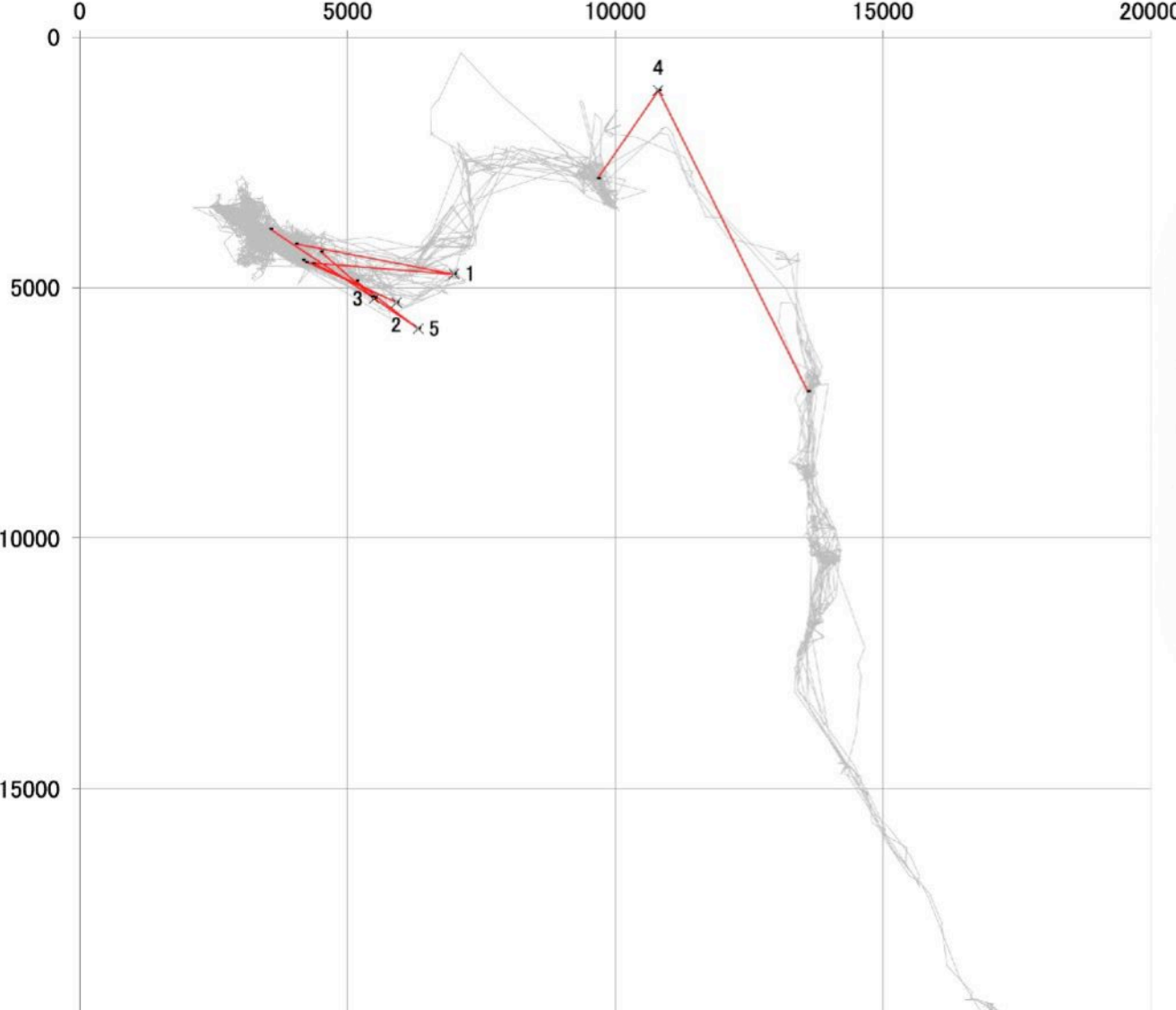
自動保存 DS-Shizu-1604.xlsx

ホーム 挿入 描画 ページレイアウト 数式 データ 校閲 表示 開発 操作アシスト

候補No.	飛躍度	角度 (°)	2辺和 (m)	長短辺比
1	2752	7.1	5780	1.1
2	1750	0.7	3693	1.1
3	1352	7.7	3745	1.7
4	1229	57.5	8707	3.2
5	1080	3.0	2574	1.4

↑

属性表示



LeapChecker (個別選択)

赤線2辺の頂点が除外候補です。

除外したくない点があれば、その番号を枠内から削除してください。
除外せず終了するにはendとタイプしてください。

なお、kmlとタイプすると、現在の軌跡を示すkmlファイルがデスクトップに作成され、Google Earthに表示されます。

5

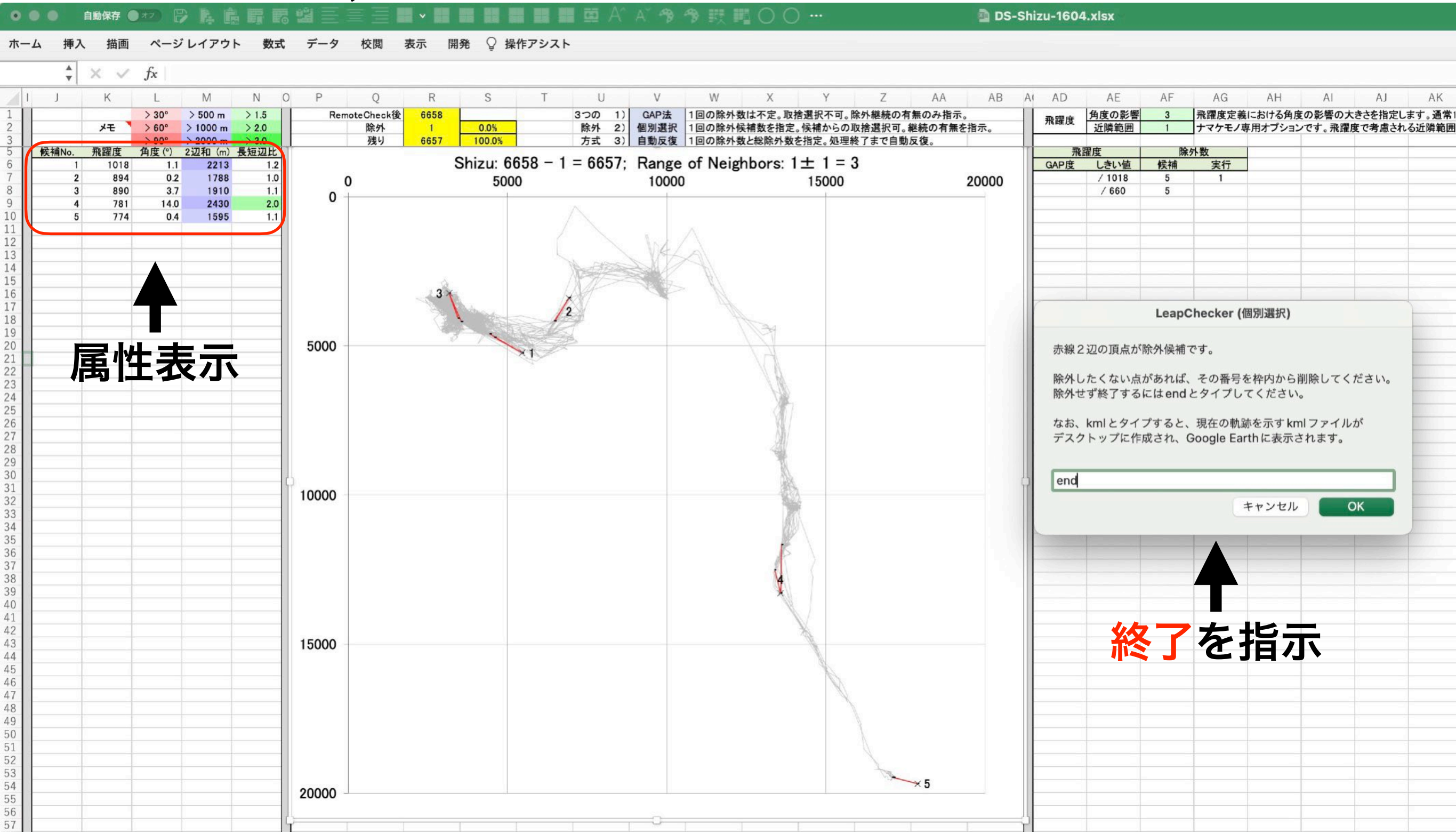
キャンセル OK

↑

番号を選択して除外を指示

除外例：2回目

ニホンジカShizu：2016/4/19 - 2018/4/17 (728日間：2時間間隔)
データ数：6814; 1回に表示する数：5



ナマケモノGPSデータの特殊性

P13

データの特徴

動物の動き

- 1) 1ヶ所の滞在時間が長い（数時間～数日間）
- 2) 動きが遅い（移動速度：一説では最大時速 250 m）
- 3) 行動範囲が狭い（せいぜい数100 m半径）

GPSデータ

- 1) 精度が極めて低い
熱帯の高湿度で密な林冠によるマルチパスが主因
- 2) 機器に起因するデータエラーも

結果

位置データのばらつき >>> 動物の動き

データの特殊性（ノイズまみれ）：別処理が必要

GAP法による外れ値除外 (前処理)

P14

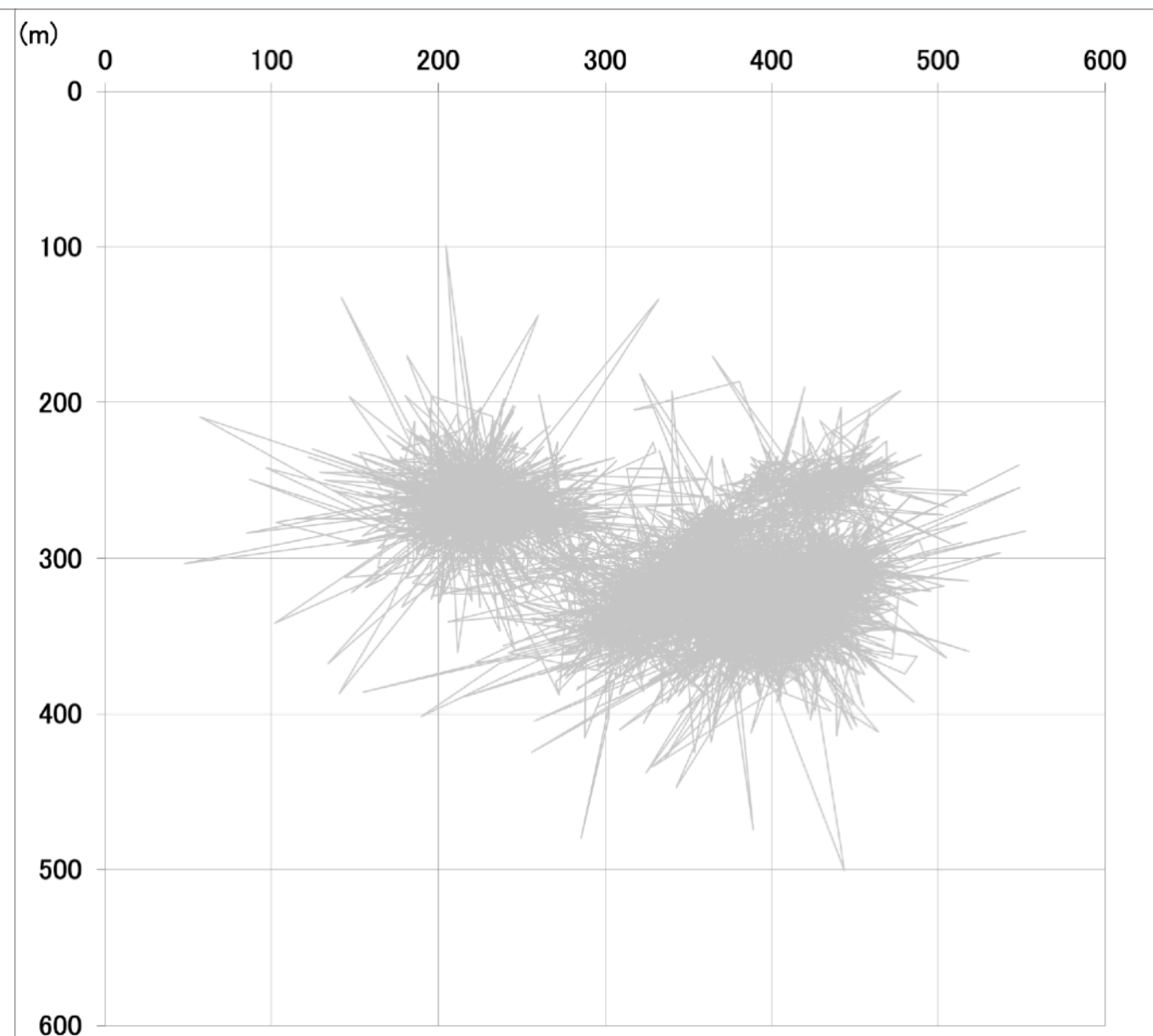
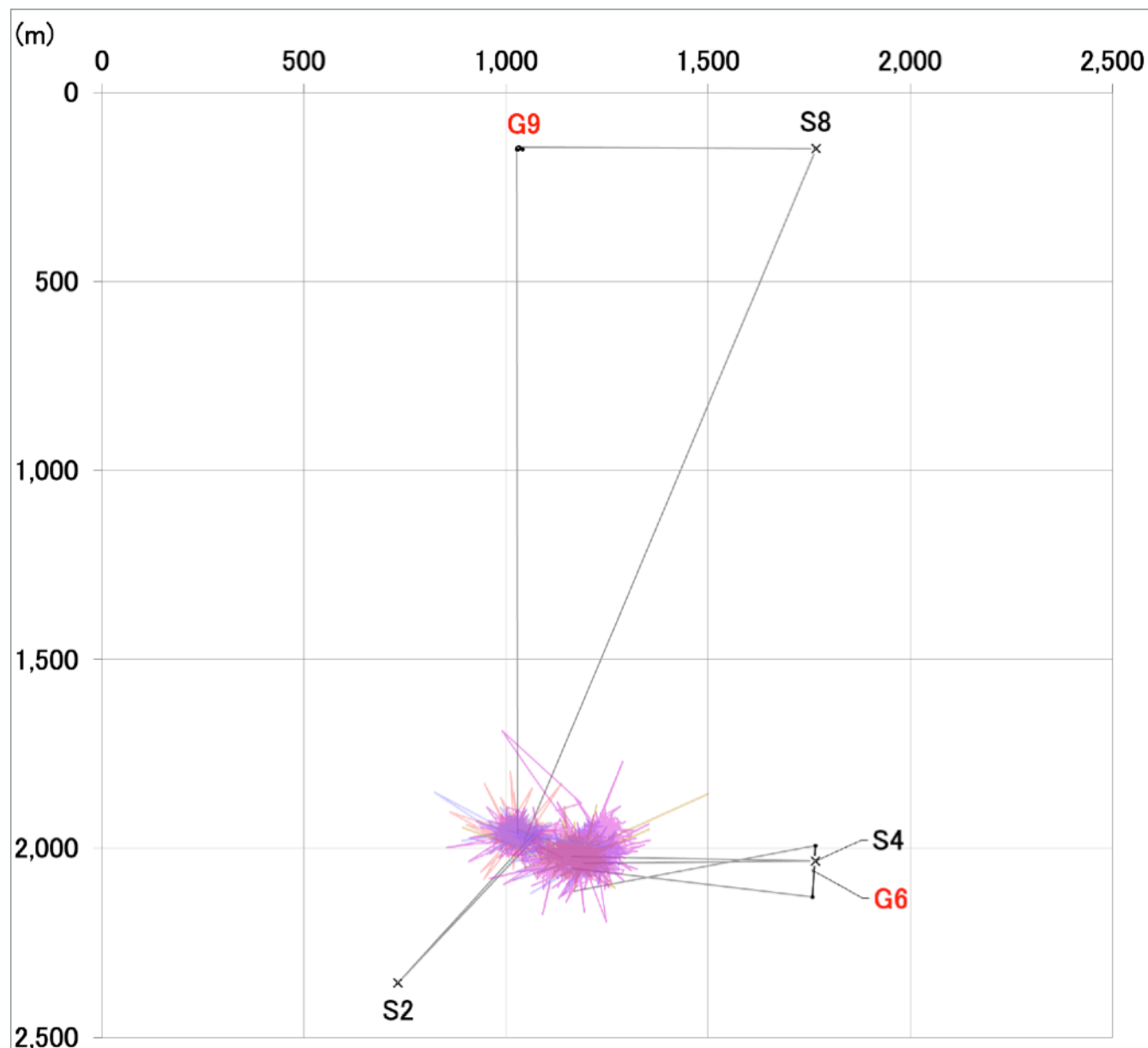
ノドジロミユビナマケモノ BT907

データ：2017/08/15 - 2018/01/02 (140日間：15分間隔)

外れ値除外：12957 - 16 = 12941 (除外率：0.123%)

除外開始 (除外候補10点)

16点除外後



ノイズ除外の方法

P15

目的：測定誤差の大きいデータ（ノイズ）を一定量（ r 個）除外
ただし、時間的に偏らないように

手続き：誤差を推定し、その大きい方から順次除外

- 1) 測定誤差を時間的近接測位点の重心からの距離として推定
 - a) ある点とその前後各 α 個の測位点（ $1+2\alpha$ ）について
 - b) その重心（座標値 X, Y の平均 $\bar{X}, \bar{Y}$ ）を算出
 - c) この重心から当該点への距離（重心距離）を計算
 - d) この距離を期間幅で補正（時間的偏り防止のため：下記参照）
- 2) 推定誤差の大きい方から β 個の測位点を除外
- 3) 1) に戻る（ r 個除外達成まで）

$$\text{推定誤差} = \text{重心距離} / \log \sqrt{\text{前期間幅} \times \text{後期間幅}} \quad \text{補正はまだ工夫の余地}$$

α, β, r は適宜指定（ β は反復数に関与。時間が許す範囲で小さめに）

実施例： $\alpha = 6, \beta = 50, r = 2586$ （元データサイズの20%）

ノイズ除外

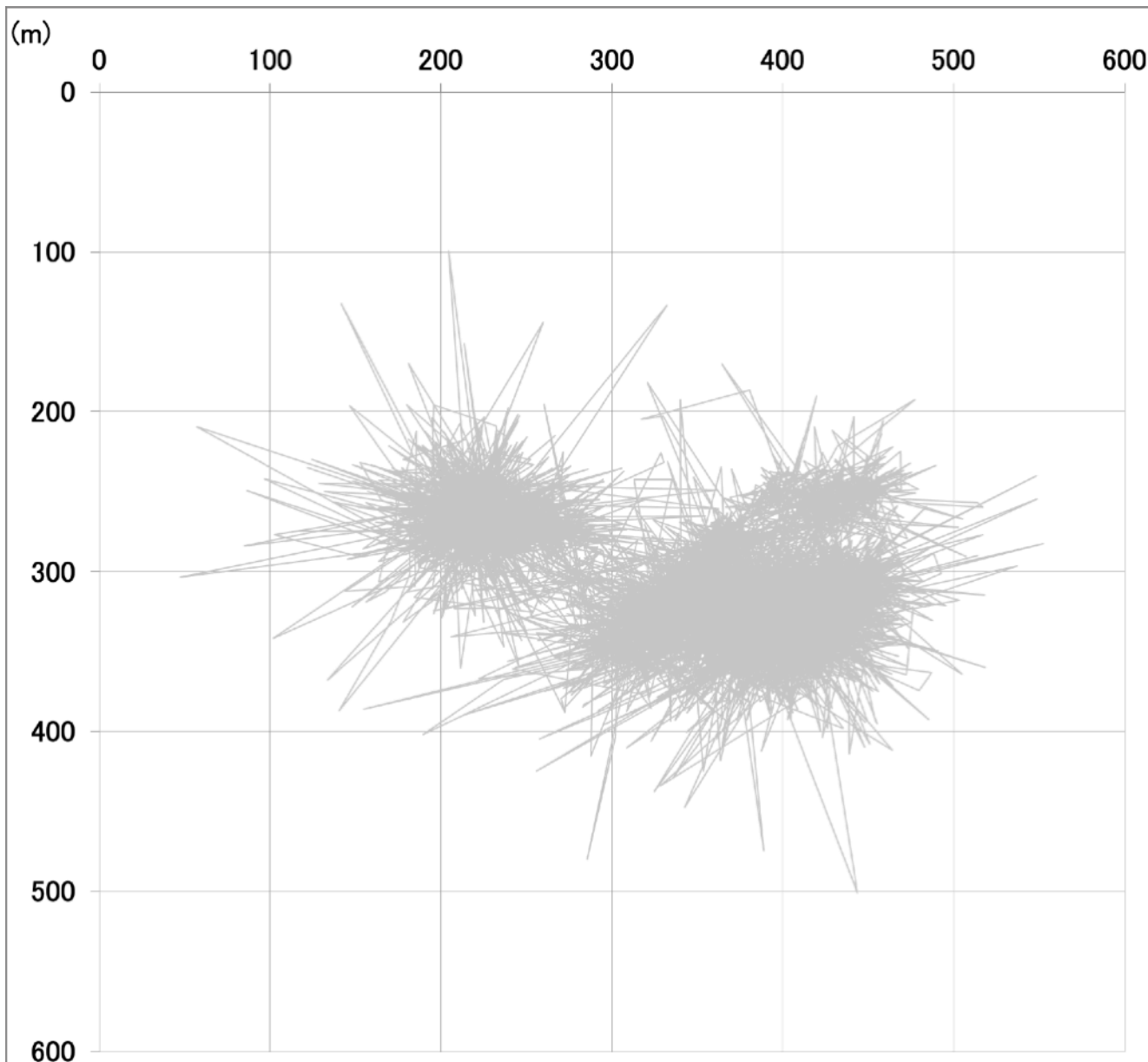
P16

ノドジロミユビナマケモノ BT907

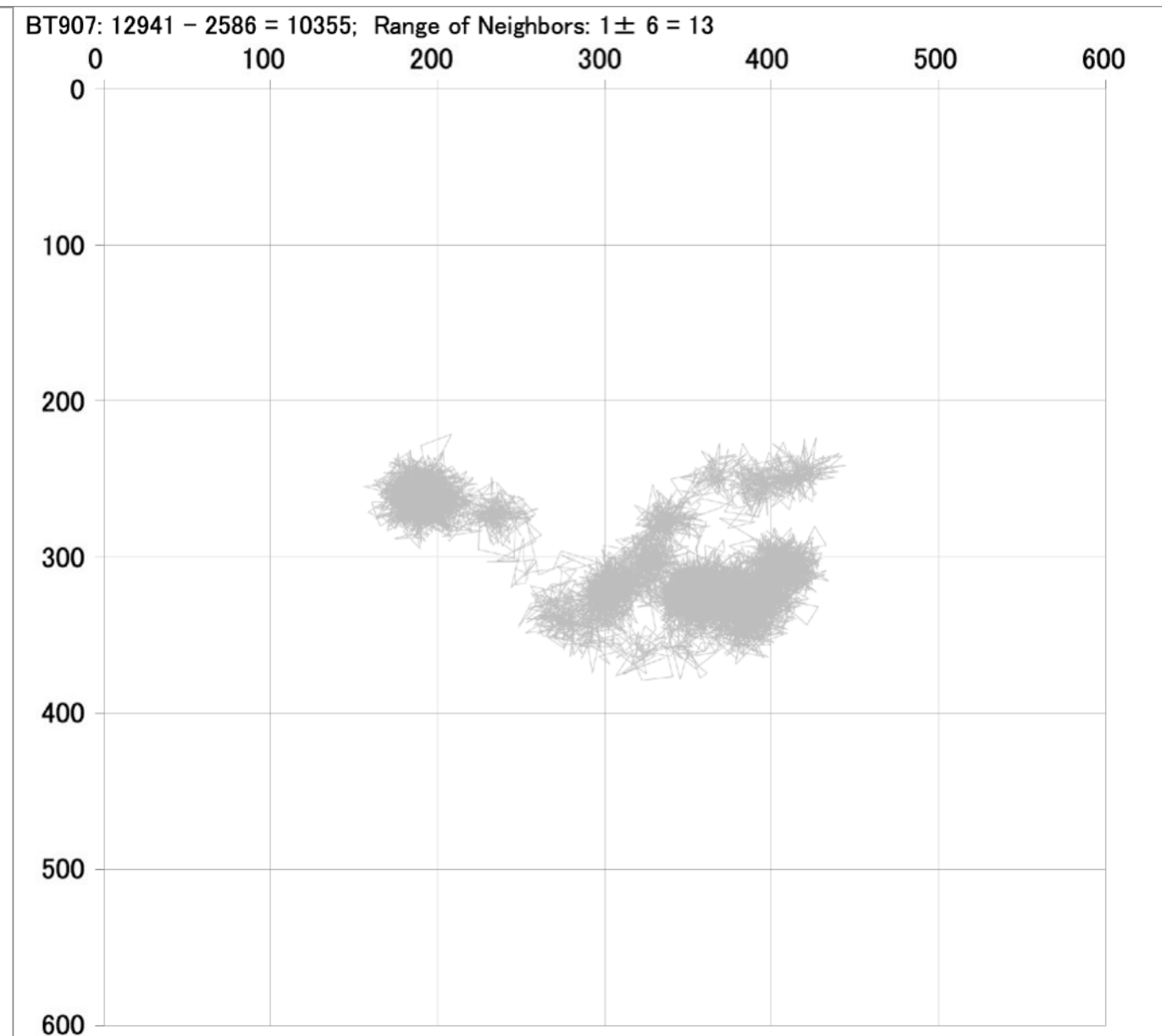
データ：2017/08/15 - 2018/01/02 (140日間：15分間隔)

ノイズ除外：12957 - 2586 = 10355 (除外率：20.0%)

除外前



除外後

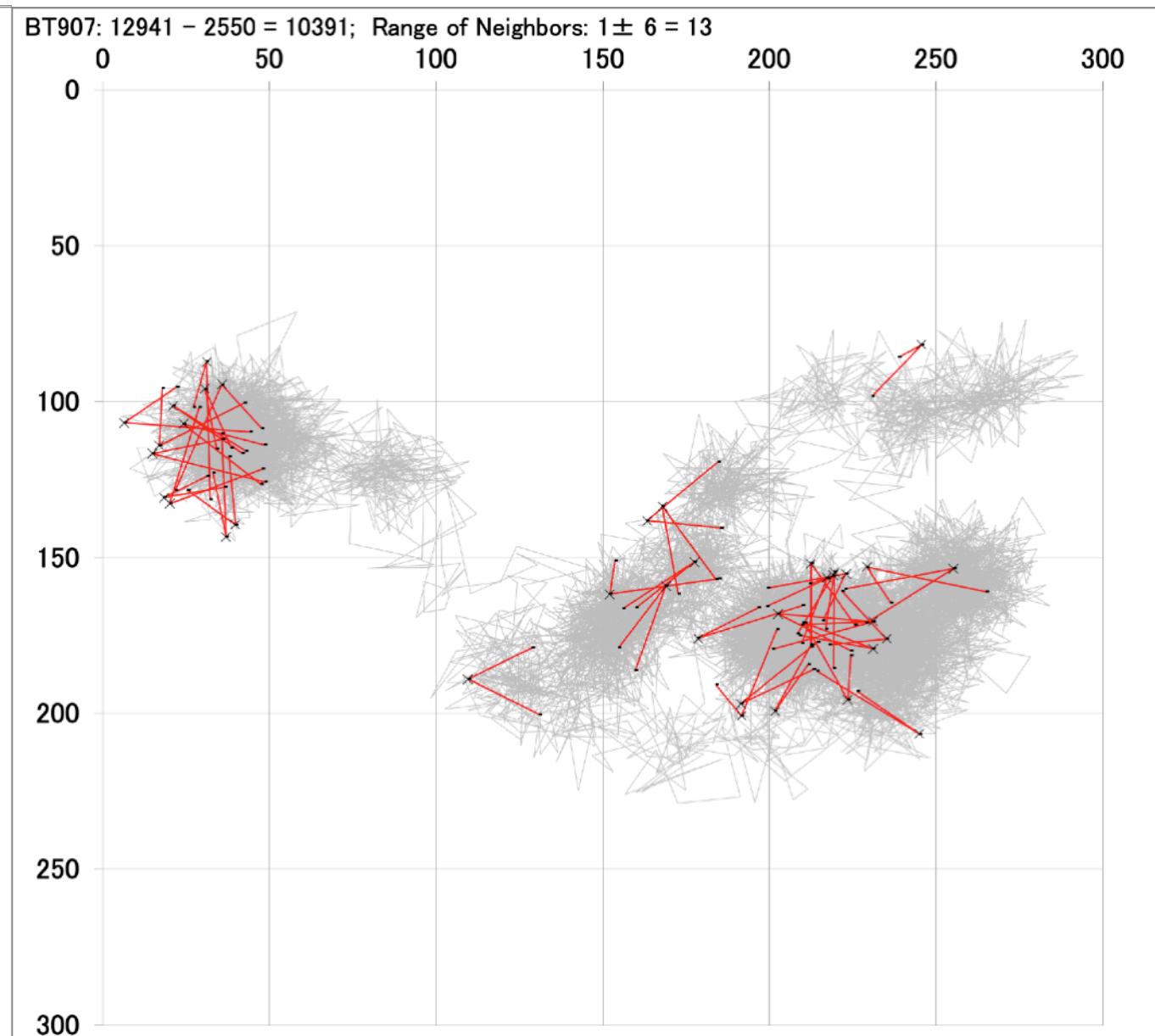
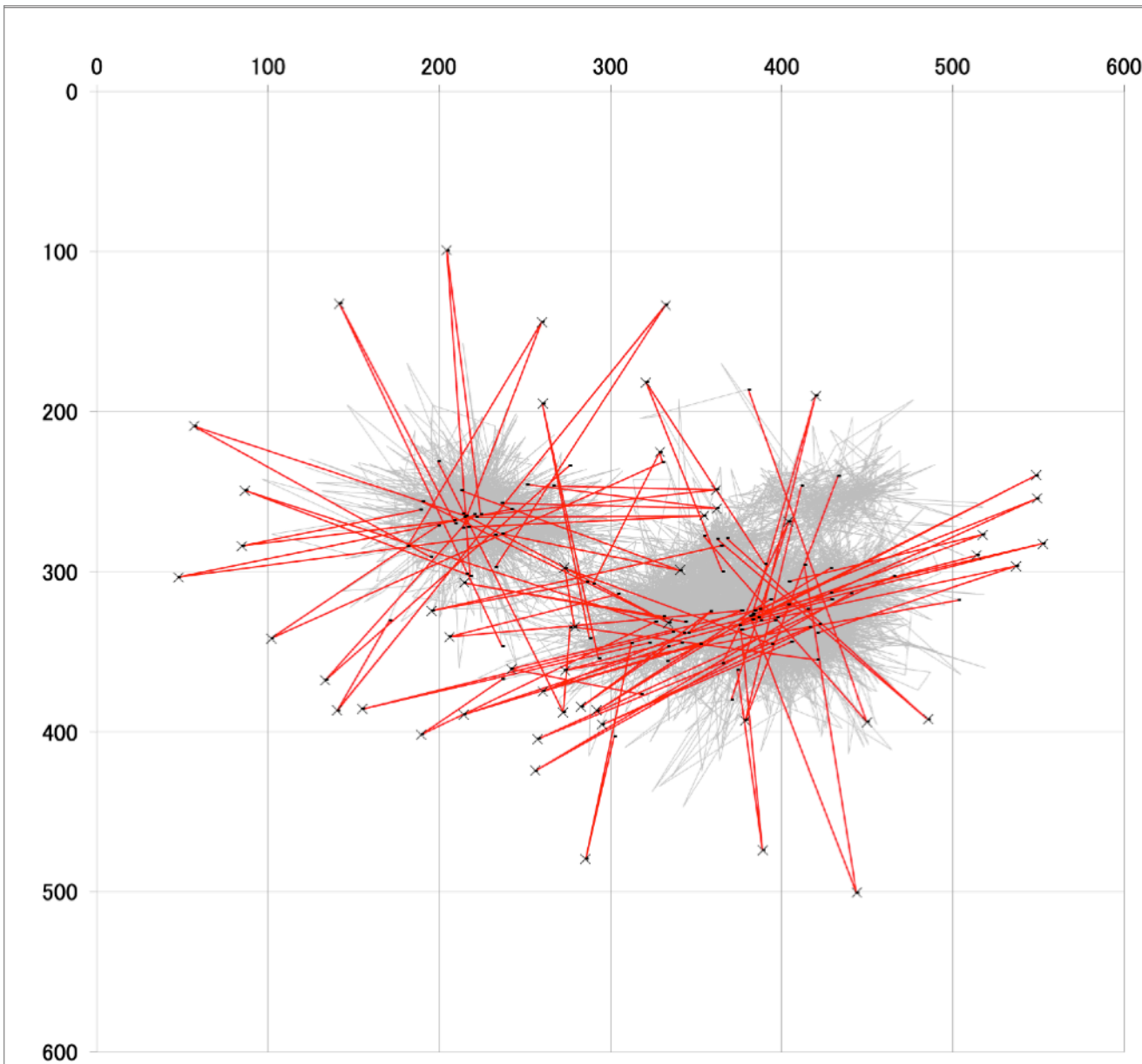


途中経過と除外後の有効測位数

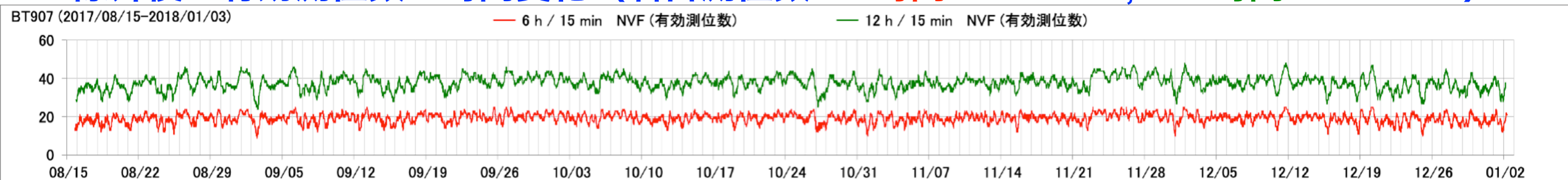
P17

除外開始時（除外予定50点）

19.7%除外後（残り除外36点）



除外後の有効測位数の時間変化（名目測位数：6時間で 24 + 1; 12時間で 48 + 1）



軌跡分割と重心軌跡化 (後処理)

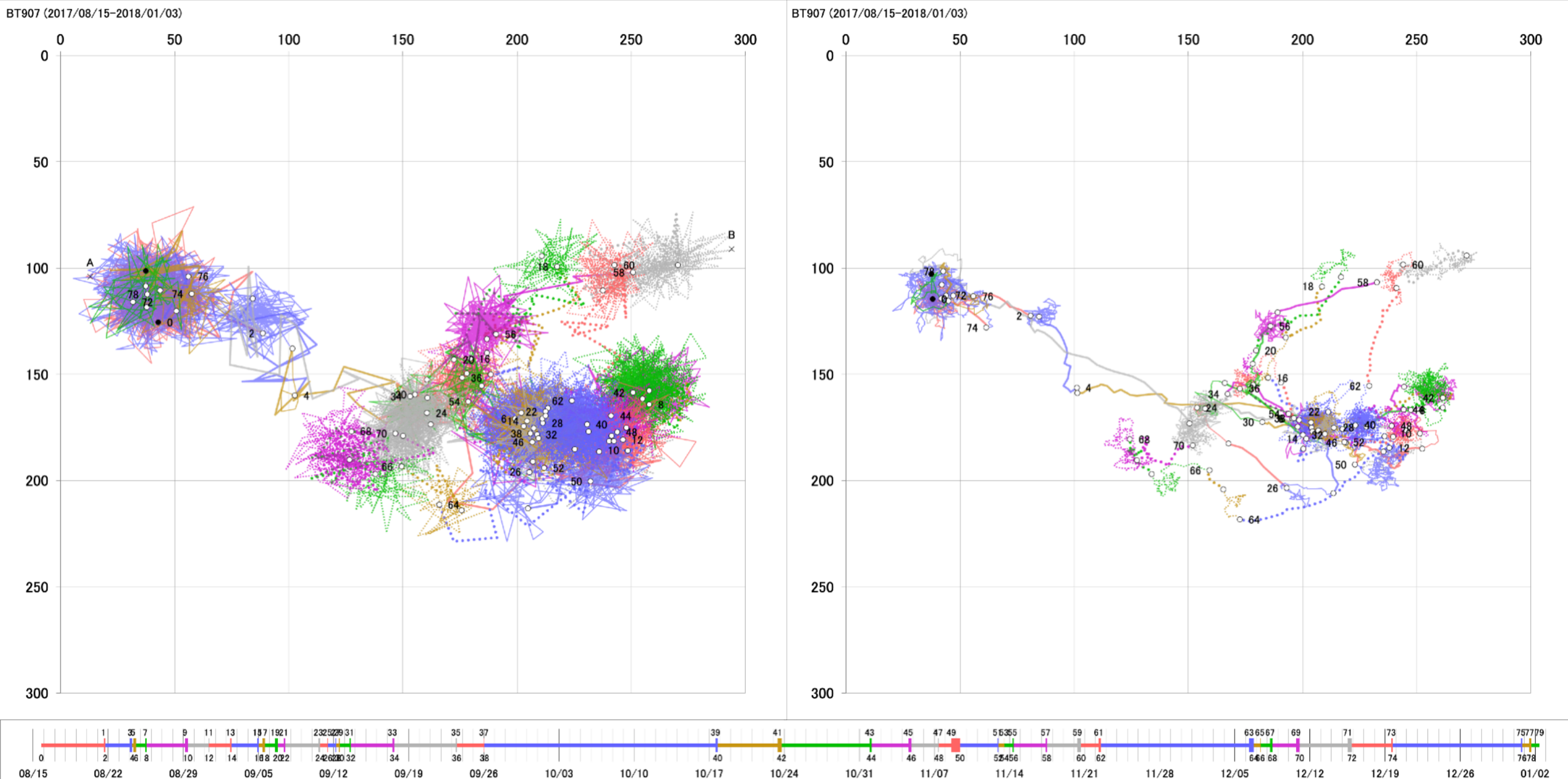
P18

時空間密度法による軌跡分割：**滞在と移動40反復**

重心軌跡：3時間（15分間隔で13点：ただし、除外等による欠損あり）

軌跡分割（色分け）

重心軌跡化



今後の課題

P19

技術課題

除外候補の属性値表示に次を追加。

●前後の測位点への時間間隔

その他

- 1) 開発技術の論文化（年内にもう 1 本を目標）
- 2) 今年一杯はもう少しシステム整備
- 3) 来年以降

国内でシステム公開、共同研究の枠組み拡大
将来的には英語版で海外展開

謝辞

P20

ニホンジカのGPS調査における捕獲は、
有限会社あかつき動物研究所の望月隆史氏にご協力いただきました。

ナマケモノのGPS調査は下記による支援を受けました。

JST/JICA, SATREPS

JSPS Core-to-Core Program

Grant Number JPJSCCA20170005

JSPS KAKENHI

Grant Number JP19KK0189