



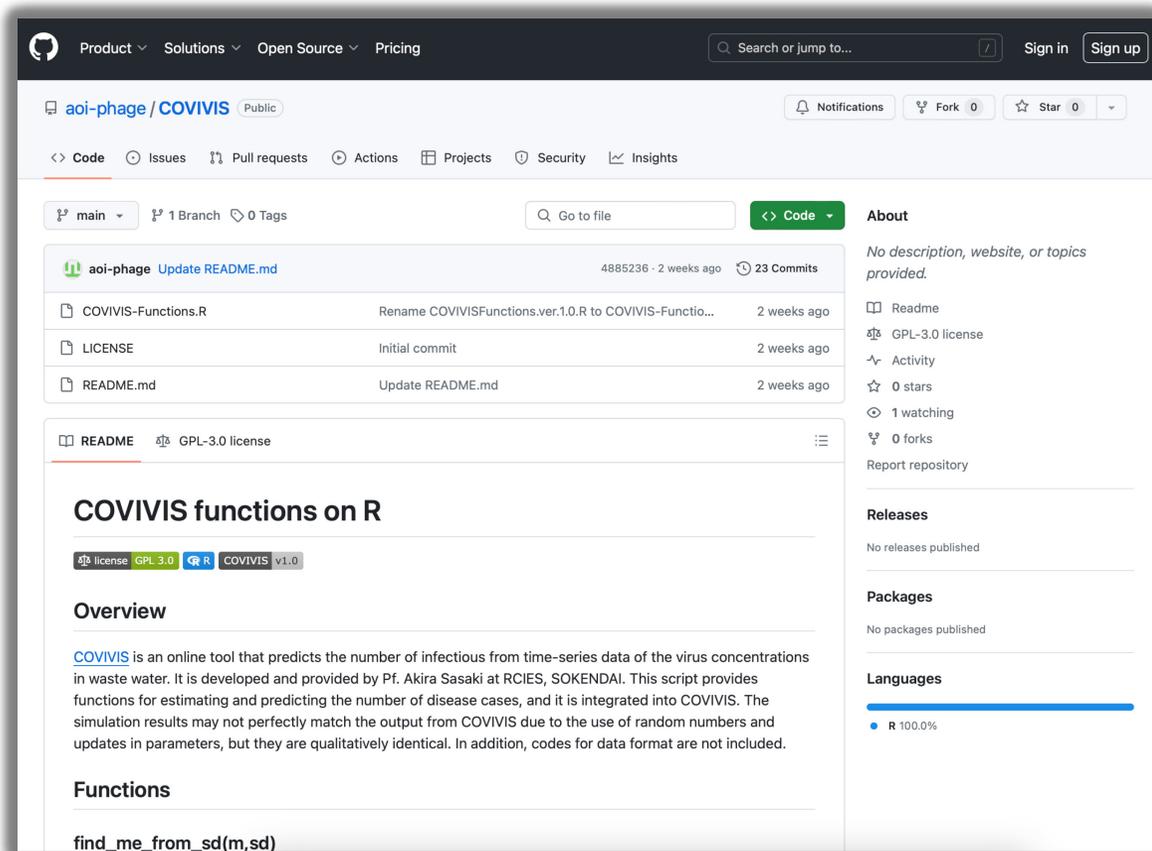
**COVIVIS(コビビス, COVID-19 Virus – infection surveillance tools)は
総合研究大学院大学・統合進化科学研究センター(教授・佐々木顕)が開発・提供する
環境水中のウイルス濃度から流域の感染者数を推定する解析手法
およびウェブツールの総称です。**

このPowerPointファイルまたはその一部について、COVIVISユーザー様が自由に
転載し、報告書やプレゼンにご使用頂くことができます。

ver.1.41 updated! 2024.4.15.

COVIVIS解析ツール Rコード

GitHubにて公開中



<https://github.com/aoi-phage/COVIVIS>

COVIVISの元となったR(統計ソフト)コードでの解析手法をGitHubにて掲載しております。オープンソースとして誰でもご利用頂けます。まずはそちらをご参照ください。

COVIVIS解析ツール ウェブアプリ版



<https://covivis.soken.ac.jp>

2024年4月からβ版を限定公開

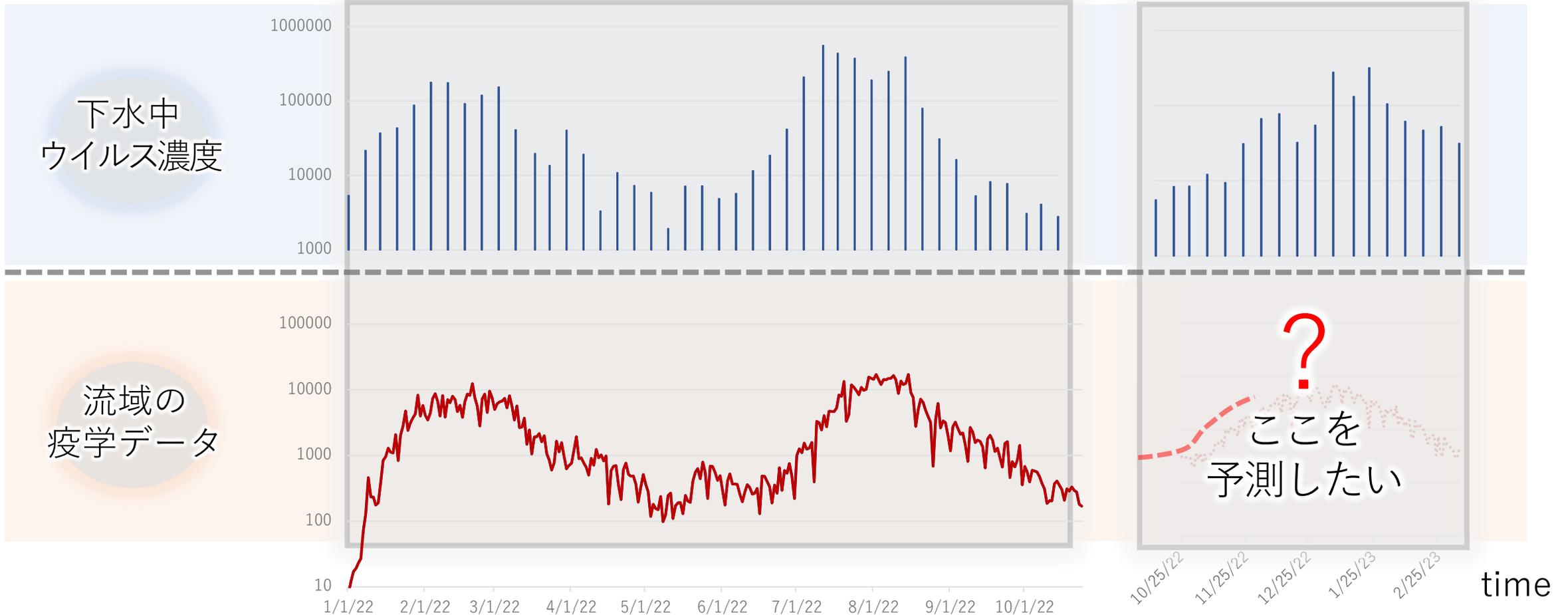
- NIJIs (New Integrated Japanese Sewage Investigation for COVID-19)
- RCIES (総研大統合進化科学研究コース)

COVIVIS解析ツールの目的

※ 架空のデモ用データ使用

双方のデータが十分にある期間

ウイルス濃度データに依存する期間



モデルのパラメータ推定に使える期間

モデル予測する期間

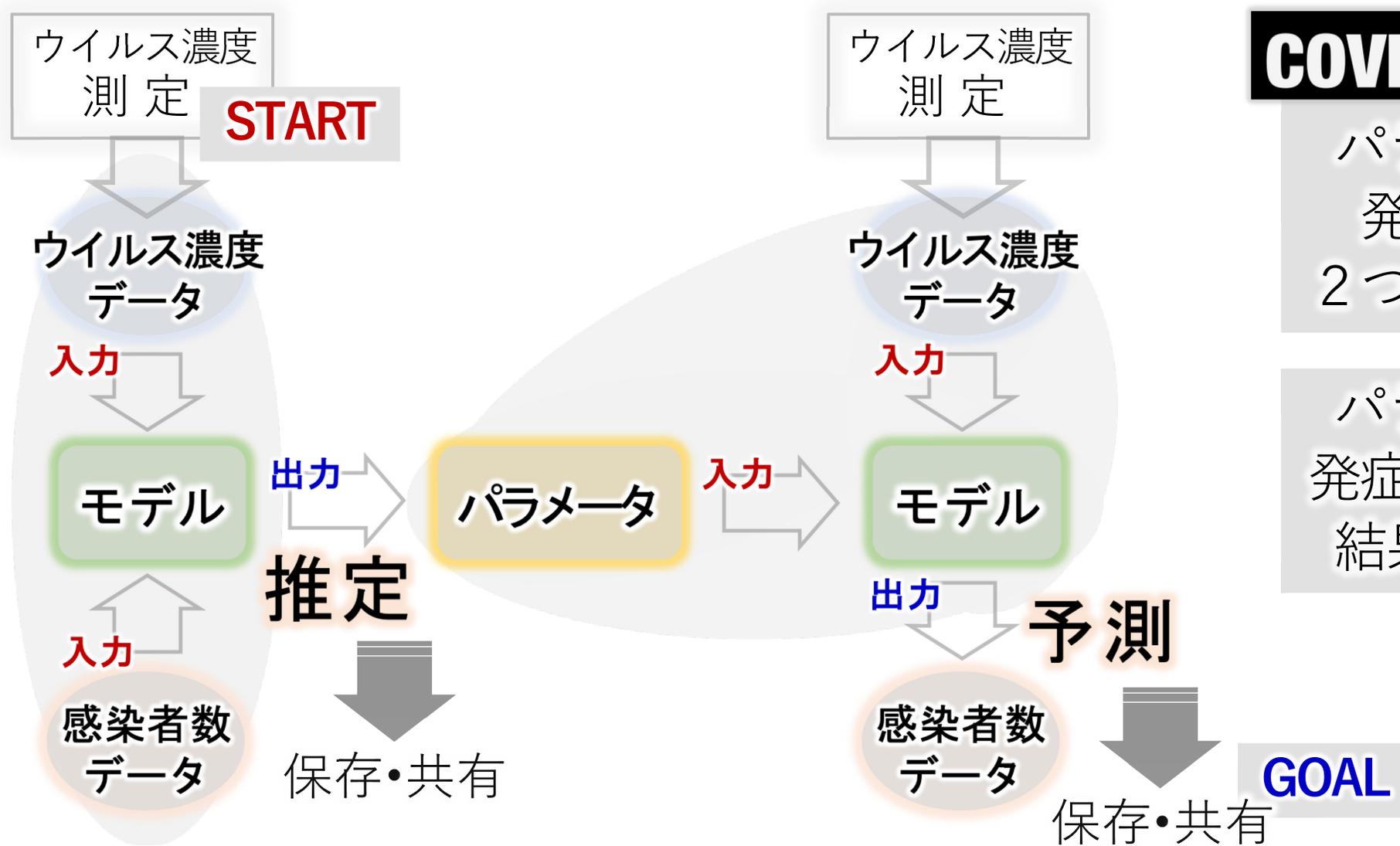
ここを
予測したい

COVIVIS解析ツールとは？

COVIVIS解析ツールでは以下のことを行います。

1. 過去のある期間における、下水中ウイルス濃度とその流域の感染報告数を基に、双方の数値を変換できる数理モデルのパラメータを推定する。
2. さらにこのモデルに、推定したパラメータと予測したい期間の下水中ウイルス濃度を入力し、その期間の発症者数の予測を行う。

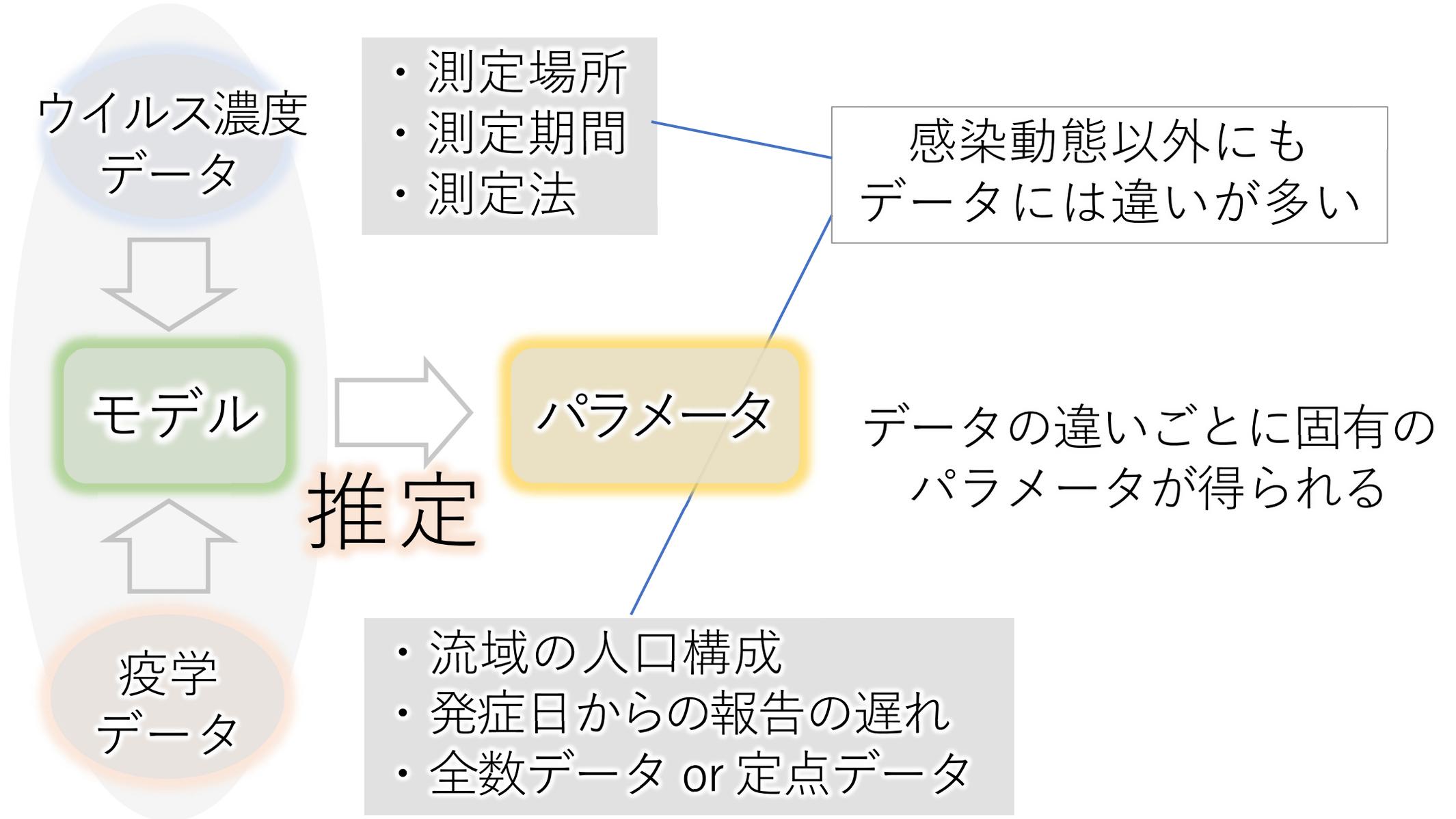
パラメータ推定と発症者数予測とは？



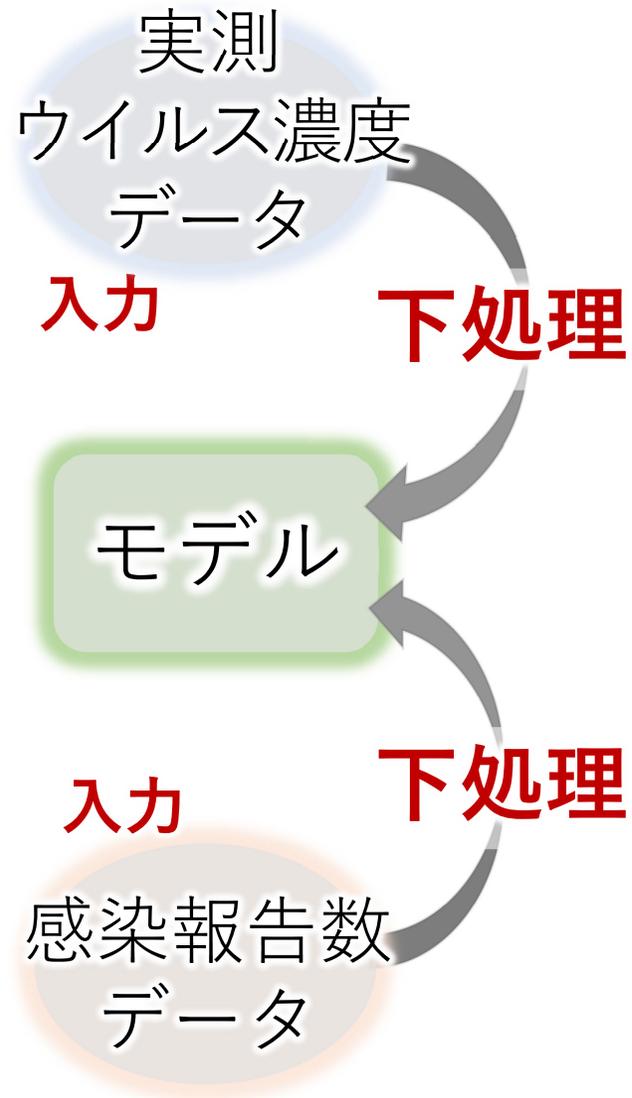
COVIVIS では、
パラメータ推定と
発症者数予測の
2つの計算フォーム

パラメータ推定と
発症者数予測の解析
結果を**保存と共有**

パラメータ推定と発症者数予測とは？



パラメータ推定と発症者数予測とは？



COVIVIS では、
実測データの下処理を全て
計算フォームから設定可能

多くの下処理が必要

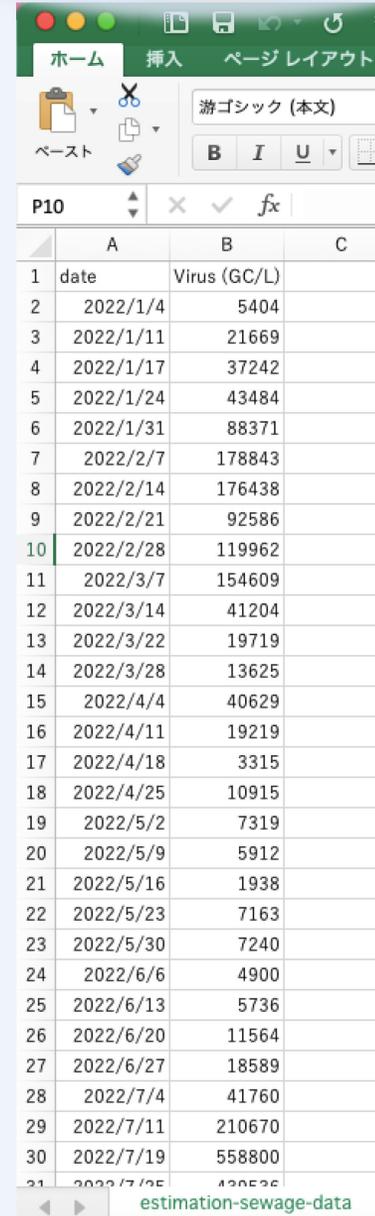
- ・ 移動平均や欠損日の補間
- ・ 1週間の集計データを日次データに変換
- ・ 報告日データから発症日データを推定

入力データ

- .csvもしくは.tsvのファイル入力
とコピー&ペーストのテキスト入力
(ExcelでOK)
- 日付データと数値データの2カラム
- 1行目や1列目に項目名などが
入っていても自動で除去
- ウイルス濃度と疫学データの
日付が一致していなくても自動で処理

※ 詳しくは操作マニュアル参照

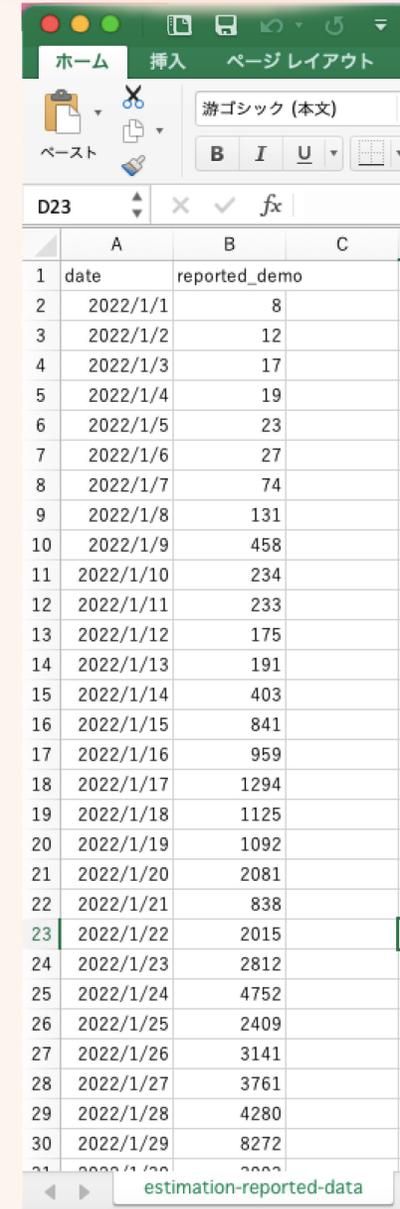
下水中ウイルス濃度



The screenshot shows an Excel spreadsheet with the following data:

	A	B	C
1	date	Virus (GC/L)	
2	2022/1/4	5404	
3	2022/1/11	21669	
4	2022/1/17	37242	
5	2022/1/24	43484	
6	2022/1/31	88371	
7	2022/2/7	178843	
8	2022/2/14	176438	
9	2022/2/21	92586	
10	2022/2/28	119962	
11	2022/3/7	154609	
12	2022/3/14	41204	
13	2022/3/22	19719	
14	2022/3/28	13625	
15	2022/4/4	40629	
16	2022/4/11	19219	
17	2022/4/18	3315	
18	2022/4/25	10915	
19	2022/5/2	7319	
20	2022/5/9	5912	
21	2022/5/16	1938	
22	2022/5/23	7163	
23	2022/5/30	7240	
24	2022/6/6	4900	
25	2022/6/13	5736	
26	2022/6/20	11564	
27	2022/6/27	18589	
28	2022/7/4	41760	
29	2022/7/11	210670	
30	2022/7/19	558800	
31	2022/7/26	420526	

疫学データ



The screenshot shows an Excel spreadsheet with the following data:

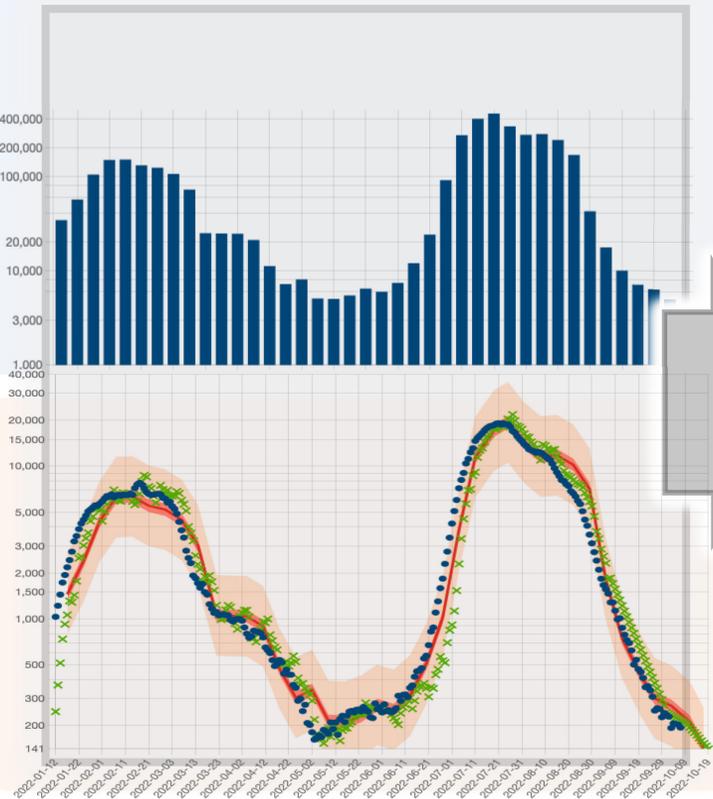
	A	B	C
1	date	reported_demo	
2	2022/1/1	8	
3	2022/1/2	12	
4	2022/1/3	17	
5	2022/1/4	19	
6	2022/1/5	23	
7	2022/1/6	27	
8	2022/1/7	74	
9	2022/1/8	131	
10	2022/1/9	458	
11	2022/1/10	234	
12	2022/1/11	233	
13	2022/1/12	175	
14	2022/1/13	191	
15	2022/1/14	403	
16	2022/1/15	841	
17	2022/1/16	959	
18	2022/1/17	1294	
19	2022/1/18	1125	
20	2022/1/19	1092	
21	2022/1/20	2081	
22	2022/1/21	838	
23	2022/1/22	2015	
24	2022/1/23	2812	
25	2022/1/24	4752	
26	2022/1/25	2409	
27	2022/1/26	3141	
28	2022/1/27	3761	
29	2022/1/28	4280	
30	2022/1/29	8272	
31	2022/1/30	2000	

COVIVIS解析ツールでの実際の画像出力

パラメータ推定

発症者数予測

下水中
ウイルス濃度



流域の
疫学データ

モデルパラメータ

V	Omega	Gamma
15.7	2.20	2.09

CSVファイルの
ダウンロードも可



fitting 期間

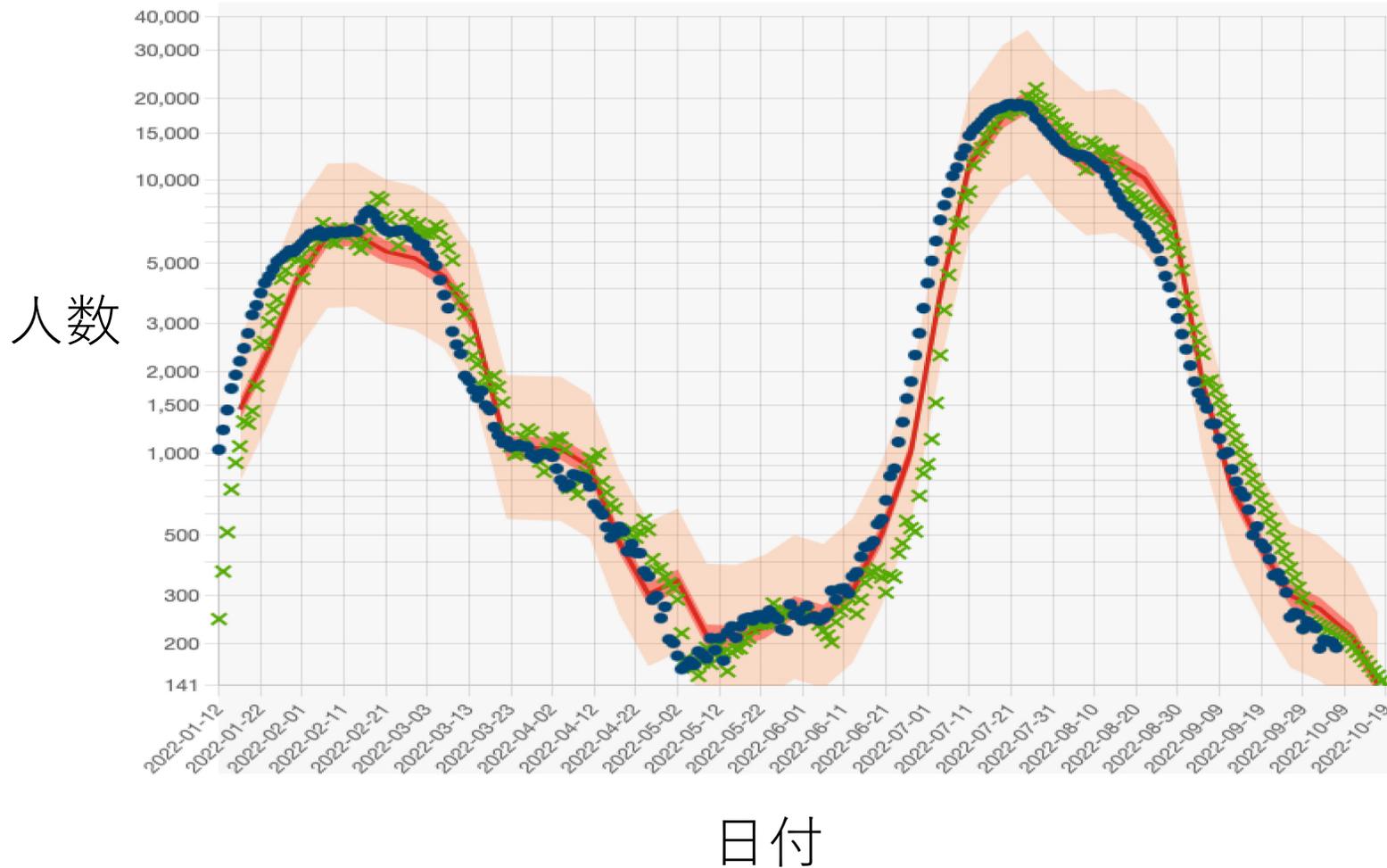
予測 期間

time

time

モデル予測発症者数の出力結果の一例

モデル予測発症者数



× 補間後報告数

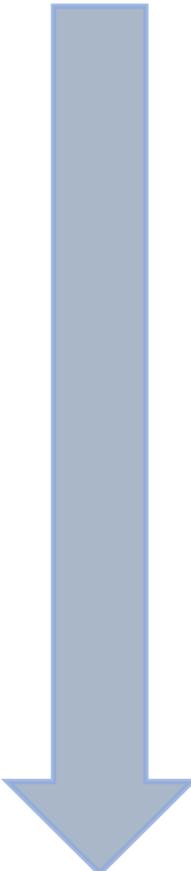
95%信頼区間

モデル予測された曲線が
95% 収まる信頼区間

95%予測区間

基になったデータのプロット
が95% 収まる予測区間

COVIVIS操作マニュアル（※解らないことがある時にみる）



手順 1. ログインする

手順 2. モデルパラメータの推定

手順 3. モデルパラメータ推定結果の確認・保存

手順 4. 発症者数の予測

手順 5. 発症者数予測結果の確認・保存

手順 6. 保存結果の読み出し・編集・削除

手順 1. COVIVIS解析ツールへのログイン

はじめにログインフォームが出ます。



The screenshot shows a login form with the following elements:

- Title: ログイン
- Label: メールアドレス
- Input field: @ メールアドレスを入力してください
- Label: パスワード
- Input field: 🔍 パスワードを入力してください
- Submit button: ログイン

ログイン用メールアドレスとパスワードを入力してください。

アカウント発行はCOVIVIS運営までメールでご連絡ください。

covivis@ml.soken.ac.jp

地衛研、自治体、政府所轄機関、大学等の教育機関を中心に、利用者を募っております。アカウント発行の審査基準はCOVIVIS基本方針をご参照ください。

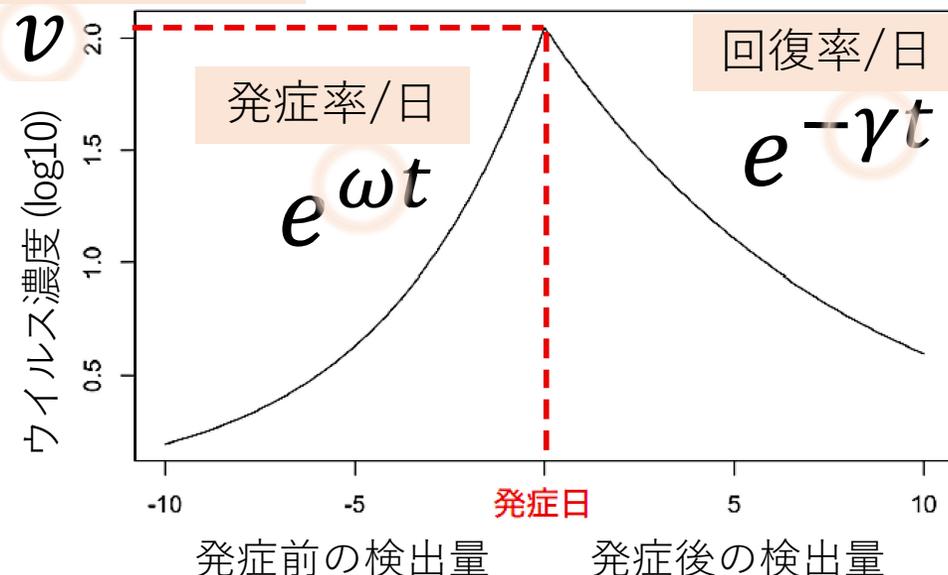
手順 2-1~2-2. 推定モデルの選択

※ これら図は解析ツールでは非表示です

- 「疫学モデル(推奨)」と「回帰モデル」のどちらかを選択します。

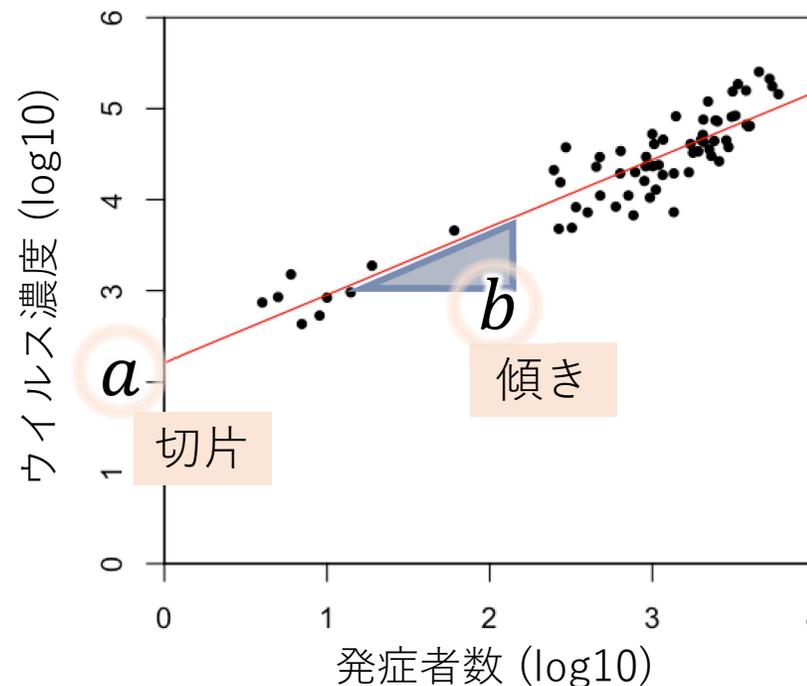
疫学モデル

最大検出量/人



発症者1人あたりのウイルス検出量について、上図のような山形の排出曲線を仮定し、発症者数の時系列データから流域全体のウイルス濃度を計算することで、実測ウイルス濃度の時系列データと最もfittingする最適な**3パラメータ**を推定する。

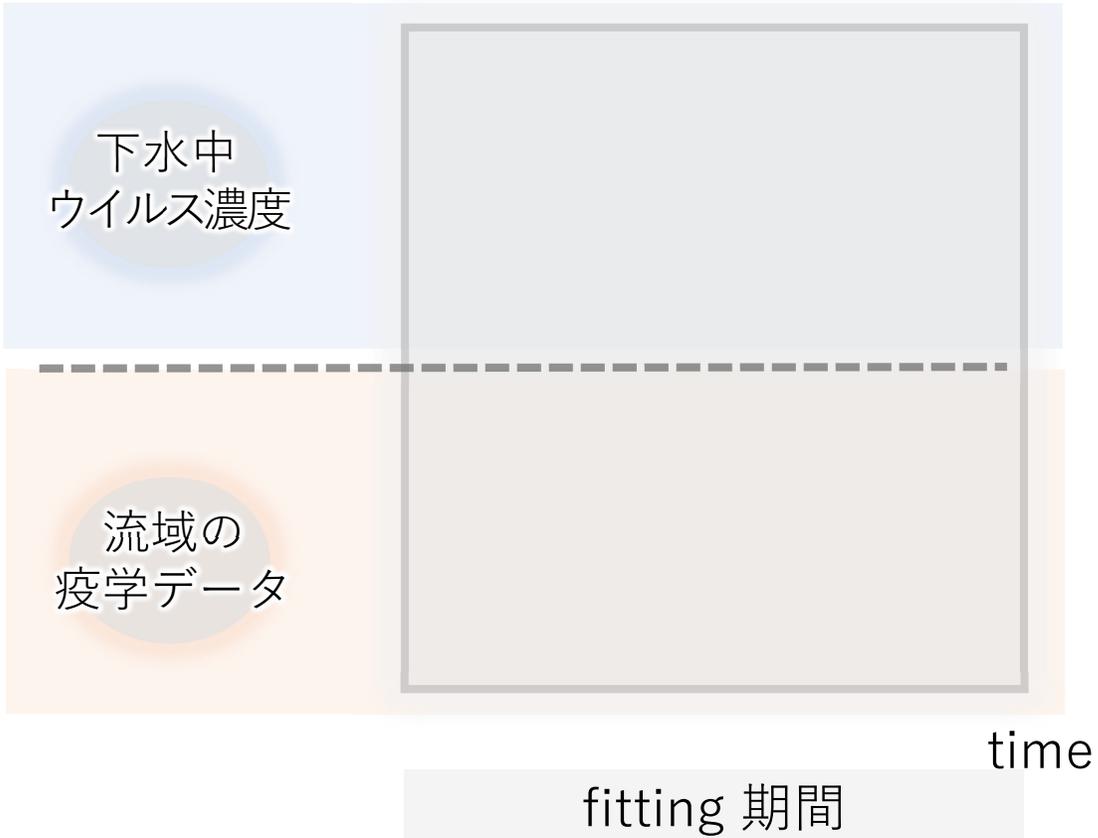
線形回帰モデル



実データを基にした発症者数とウイルス濃度の関係を両対数プロットした散布図。これを直線で回帰し、その傾き(b)と切片(a)をパラメータとして推定する。

手順 2-3. 実測ウイルス濃度データ (fitting期間) の入力

パラメータ推定

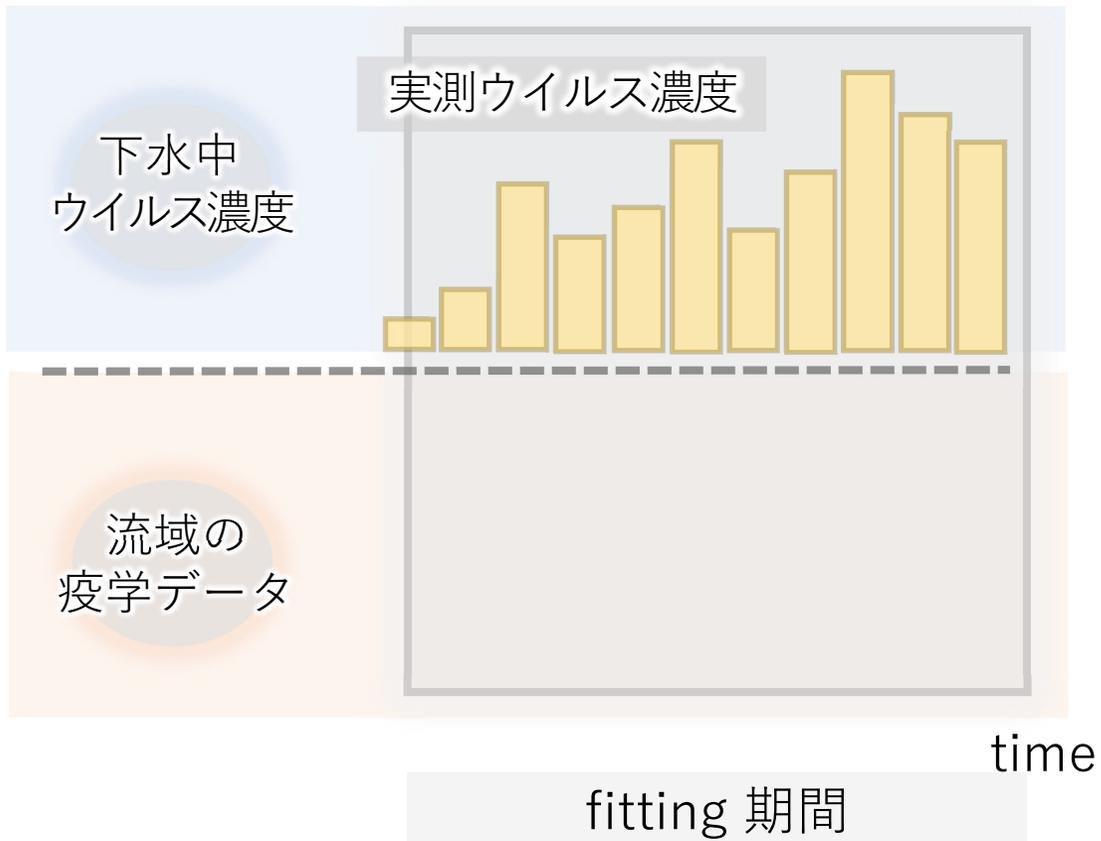


モデルパラメータの推定 → をクリック



手順 2-3. 実測ウイルス濃度データ (fitting期間) の入力

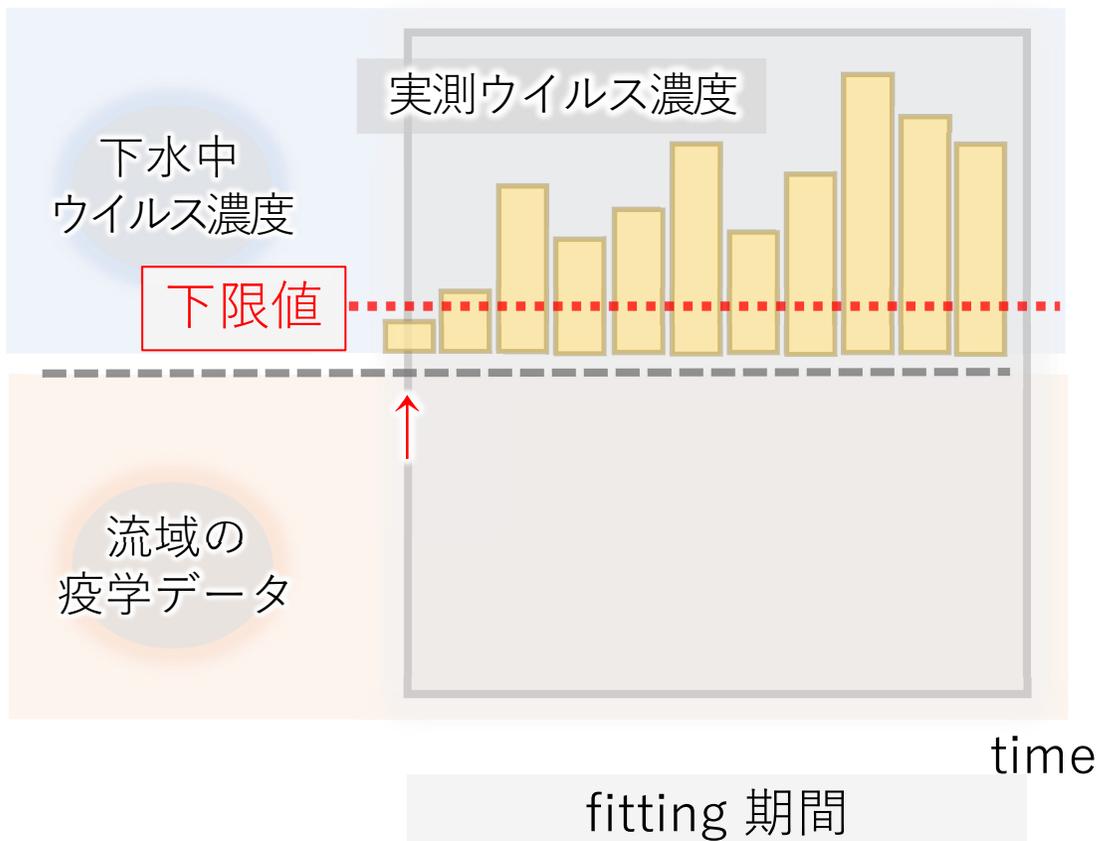
パラメータ推定



- fitting期間実測ウイルス濃度データを入力します.

手順 2-3. 実測ウイルス濃度データ (fitting期間) の入力

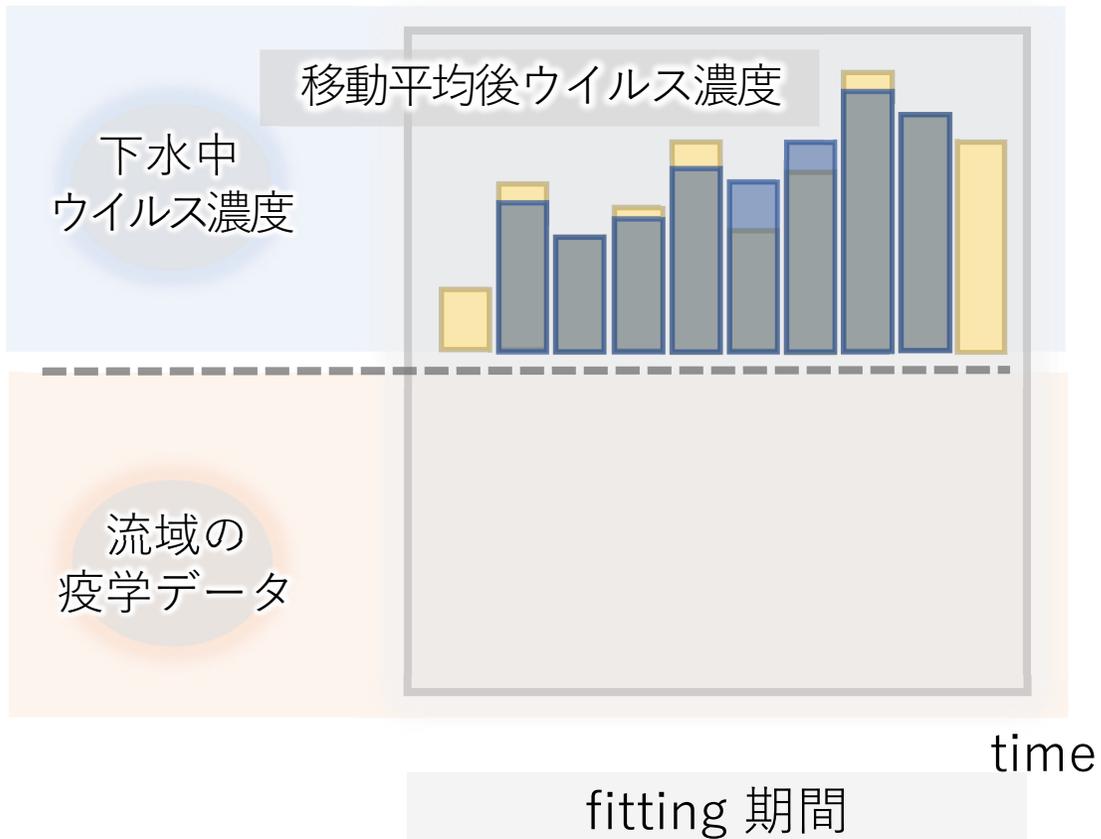
パラメータ推定



- fitting期間実測ウイルス濃度データを入力します.
- 推定に使うウイルス濃度の下限値を設定します.
→ 下限値以下のデータは除去されます.

手順 2-4. 実測ウイルス濃度データ (fitting期間) に移動平均を適用

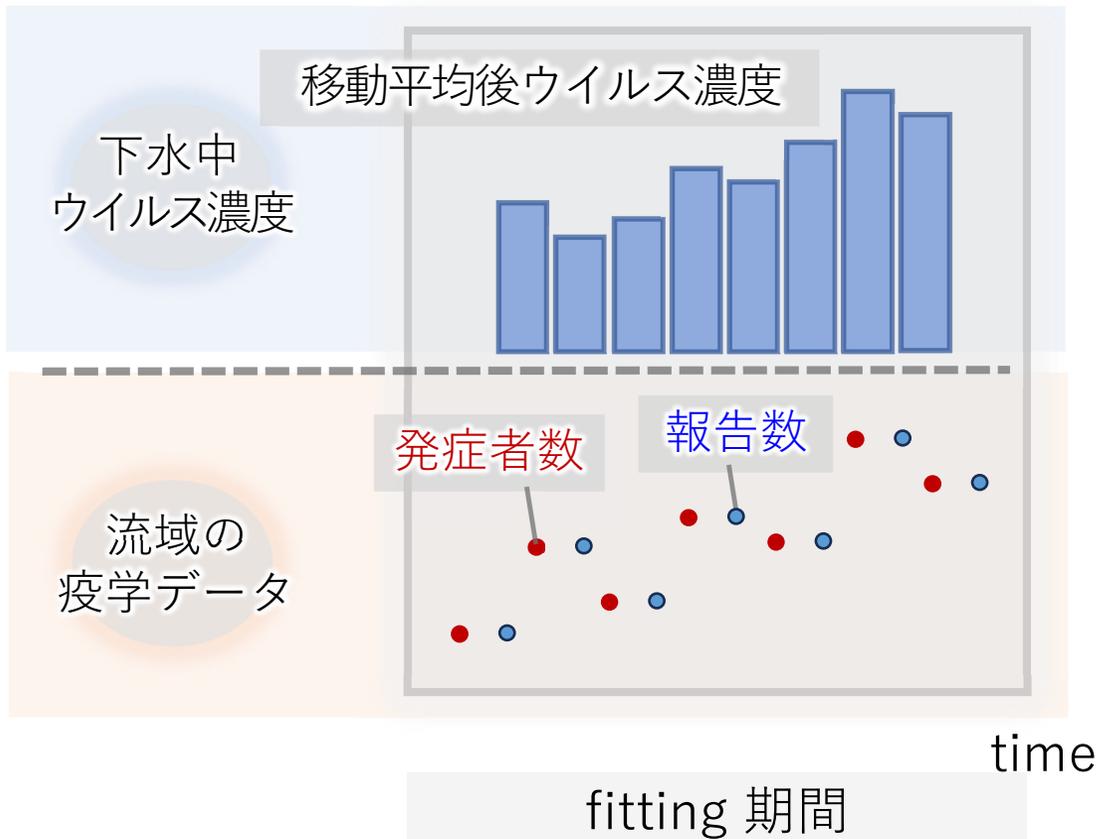
パラメータ推定



- fitting期間実測ウイルス濃度データを入力します.
- 推定に使うウイルス濃度の下限値を設定します.
→ 下限値以下のデータは除去されます.
- 入力された実測ウイルス濃度データに移動平均を適用するか選択できます.
→ 適用しない、3項移動平均、5項移動平均
→ 移動平均を掛けると両端のデータは消えます

手順 2-5. 疫学データ(fitting期間)を入力

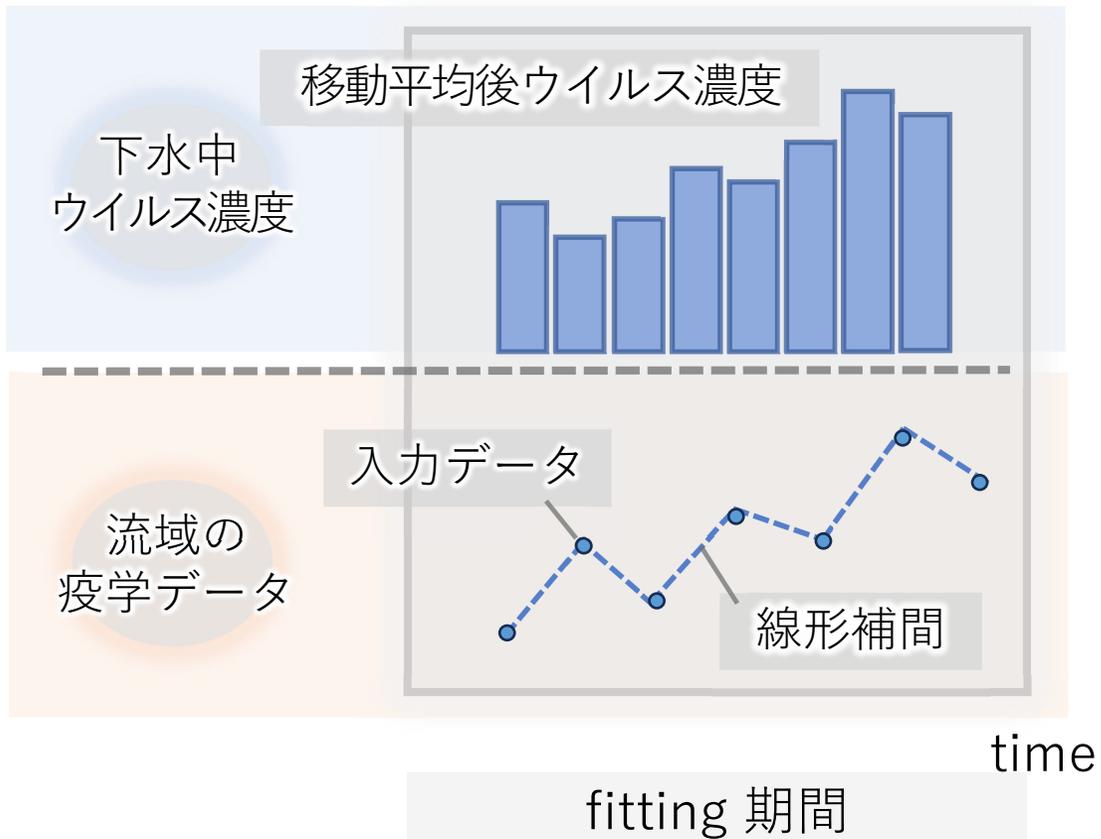
パラメータ推定



- 報告数データ(公表日ベース)の場合は、陽性報告数タグから入力します。
- 旧Her-sysなどの発症日データがある場合は、発症者数タグから入力してください。

手順 2-5. 疫学データ(fitting期間)を入力

パラメータ推定



- ・ 集計入力モードの選択

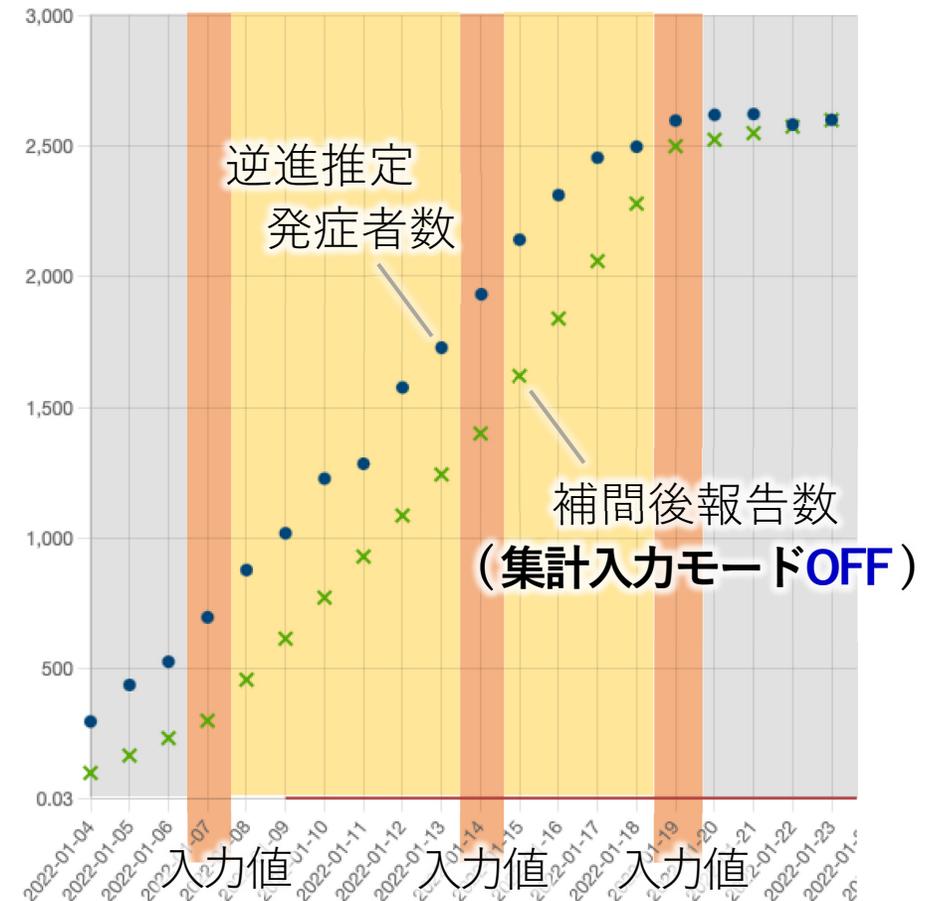
OFFの場合

→ 入力データはそのまま入力され、データ間のブランクは**線形補間**されます。

補足. 集計入力モードOFFの場合

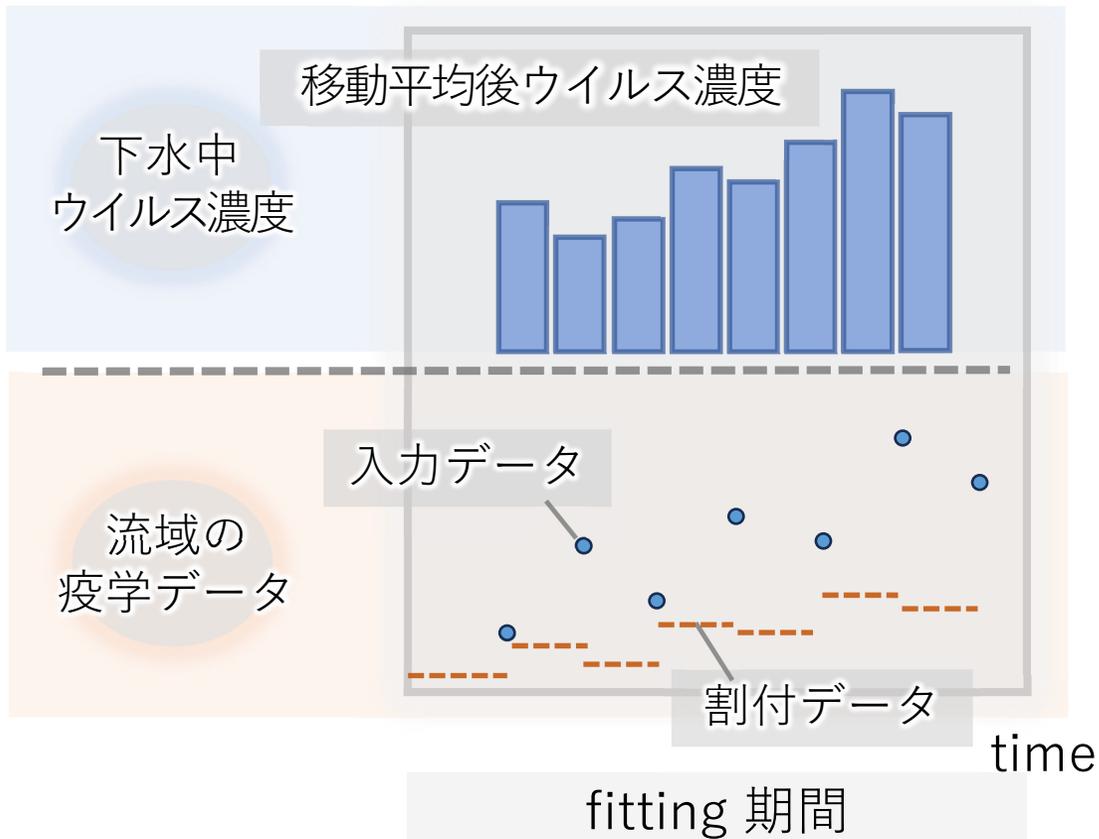
	報告数 日次データ			報告数 補間後データ	
当日分のみの数値	1月7日	300	データの欠損期間	1月7日	300
				1月8日	457
				1月9日	614
			1月10日	771	線形で補間
			1月11日	929	
			1月12日	1086	
当日分のみの数値	1月14日	1400	データの欠損期間	1月13日	1243
				1月14日	1400
				1月15日	1620
			1月16日	1840	線形で補間
			1月17日	2060	
			1月18日	2280	
当日分のみの数値	1月19日	2500		1月19日	2500
				1月20日	
				1月21日	
			1月22日		
			1月23日		

日付に**空白**がある場合、**集計入力モードを切る**と、入力データの間の空白を前後の数値で**線形補間**して空白を埋める入力となります。



手順 2-5. 疫学データ(fitting期間)を入力

パラメータ推定



- ・ 集計入力モードの選択

OFFの場合

→ 入力データはそのまま入力され、データ間の空白は線形補間されます。

ONの場合

→ 入力値を集計データとみなしてデータ後方(過去)の空白を含む日数で**均等割**されます。

→ 例えば1週間分の報告数をまとめた週次データは**日次データ(1/7の値)に変換**。

→ 空白が等間隔のインターバルである必要はありません。

補足. 集計入力モード ON の場合

報告数 集計データ		報告数 補間後データ	
		1月6日	100
1月7日	300	1月7日	100
		1月8日	200
		1月9日	200
		1月10日	200
		1月11日	200
		1月12日	200
		1月13日	200
1月14日	1400	1月14日	200
		1月15日	500
		1月16日	500
		1月17日	500
		1月18日	500
1月19日	2500	1月19日	500
		1月20日	
		1月21日	
		1月22日	
		1月23日	

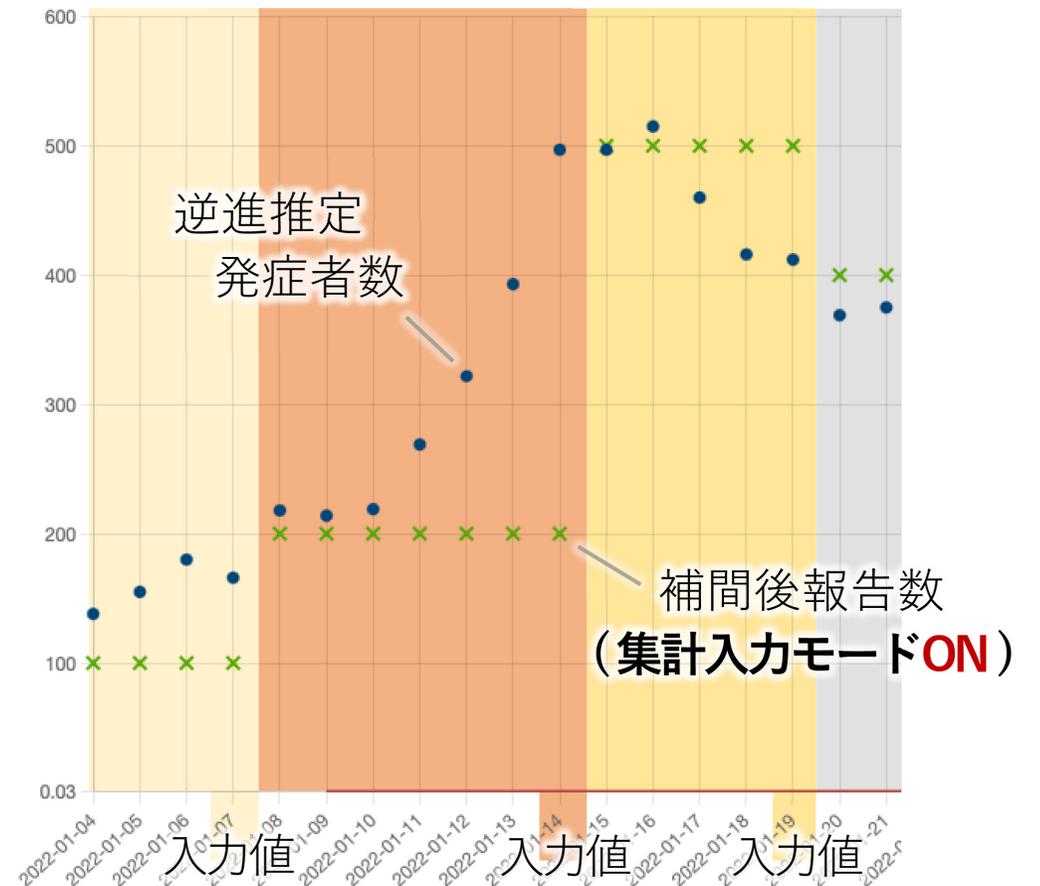
1/7で
均等割

1/5で
均等割

7日分の
合計値

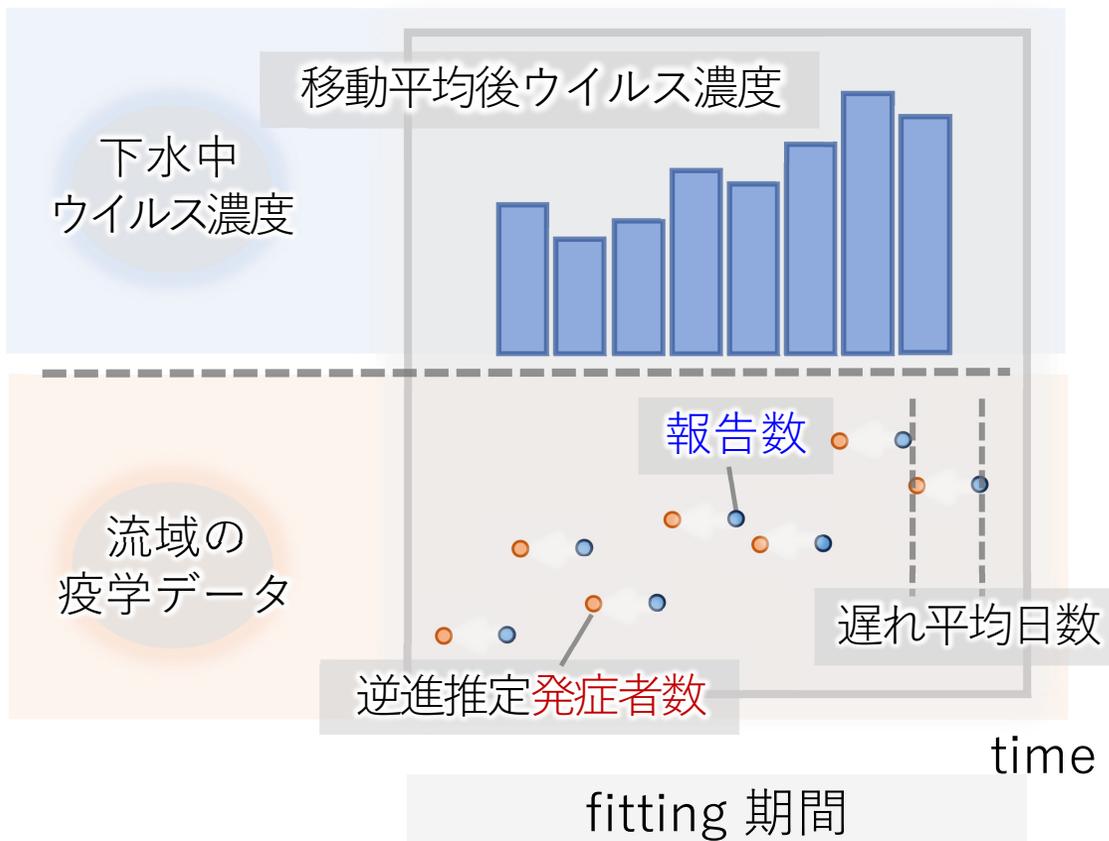
5日分の
合計値

日付に**空白**がある場合、**集計入力モード**を**使う**と、データの空白を含む日数で**数値を均等割**にし、直前の空白を埋める入力を行うことができます。



手順 2-5. 疫学データ(fitting期間)を入力

パラメータ推定



- 報告数→発症者数の**逆進推定パラメータ**の入力
→ 感染者の発症日から報告日への**遅れ平均日数**とその**標準偏差**を入力します。

ワイブル分布に従って逆進推定を行い、
発症者数の時系列データが生成されます。

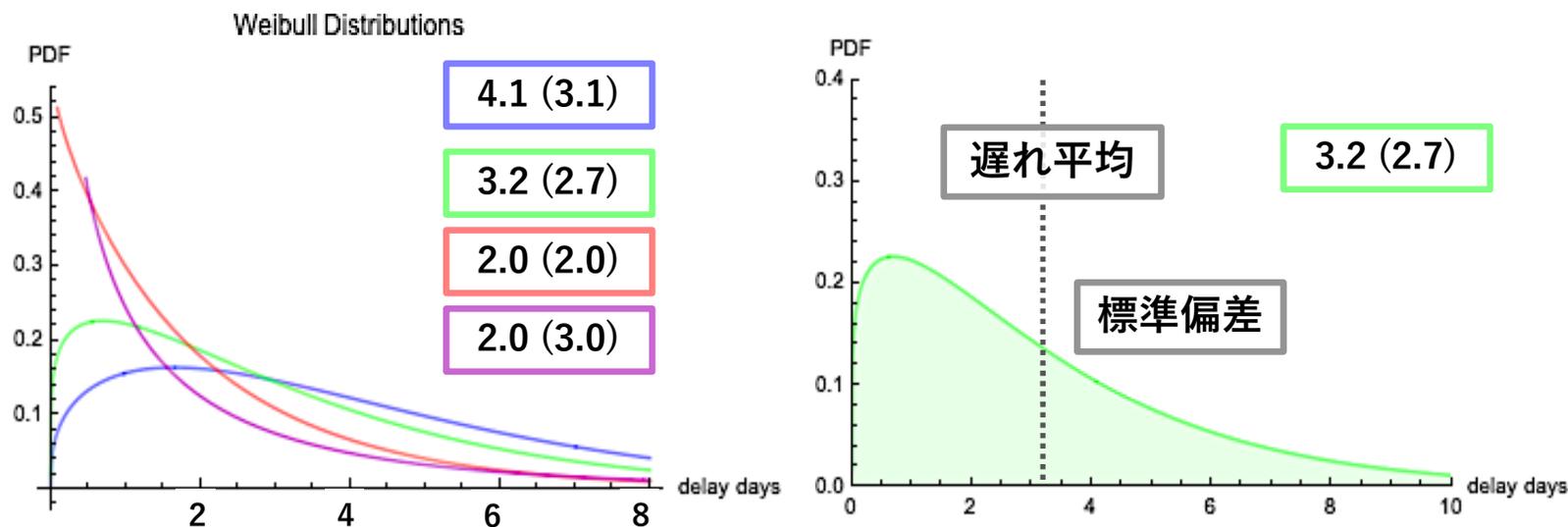
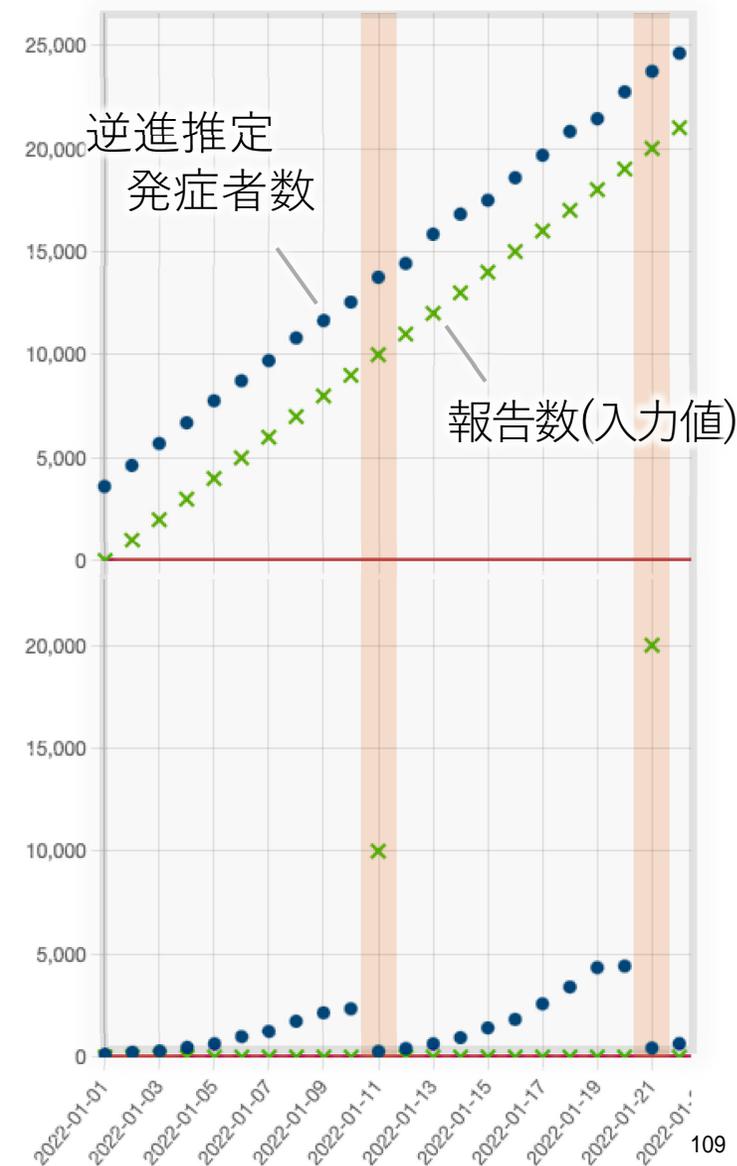
補足. ワイブル分布を用いた逆進推定

公表日ベースの報告数データから発症日ベースの疫学データを得るために、**ワイブル分布**を用いた**逆進推定**を行います。

発症から報告までの**遅れ平均日数**と**標準偏差**を逆進推定パラメータとして、そのワイブル分布の確率密度に従うモンテカルロシミュレーションを行い、報告数を逆進方向へばらします。

連続したデータの場合、遅れ平均日数分だけプロットが逆進方向に並行移動して見えます。孤立した報告数データでは、逆進推定発症者数が(左右反転した)ワイブル分布の形状となります。

遅れ平均日数 (標準偏差)
= 3.2 (2.7)



補足. 遅れ平均日数の算出について

右はエクセルで下処理する場合の一例です。

発症日(B列)から公表日(C列)を引き算し、遅れ日数(D列)を出します。

D列をコピー → 形式を選択してペースト → 数値を選択して昇降順に並べ替えたものがE列です。

発症日が「調査中」や「空欄」の場合は下部に“#VALUE!”として弾かれ、上部に数値だけが昇降順に並びます。さらに、マイナスの数値や明らかに大きな外れ値は使用範囲から外します。

平均値と標準偏差の関数を使い、その引数として遅れ日数(E列)の適切な数値のみ範囲指定します。

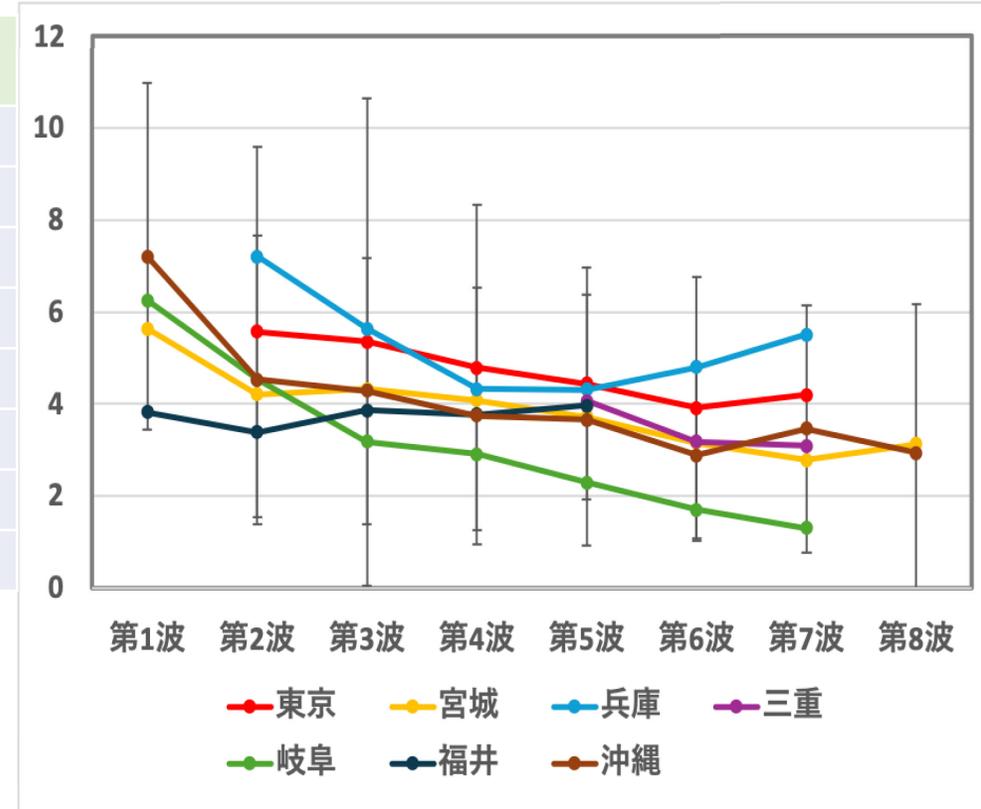
オープンデータを用いて遅れ日数(E列)を0～100の範囲で集計した例が次のスライドになります。

	A	B	C	D	E	F
1	No	公表_年月日	発症_年月日	遅れ日数	sorted	
2	888779	9月4日	調査中	#VALUE!	1	
3	888780	9月4日	調査中	#VALUE!	1	
4	888781	9月4日	調査中	#VALUE!	1	
5	888782	9月4日	9月3日	1	1	
6	888783	9月4日	9月2日	2	1	
7	888784	9月4日	9月3日	1	1	
8	888785	9月4日	9月3日	1	1	
9	888786	9月4日	9月3日	1	2	
10	888787	9月5日	調査中	#VALUE!	2	
11	888788	9月5日	調査中	#VALUE!	2	
12	888789	9月5日	調査中	#VALUE!	2	
13	888790	9月5日	9月2日	3	2	
14	888791	9月5日	9月4日	1	2	
15	888792	9月5日	9月2日	3	3	
16	888793	9月5日	9月3日	2	3	
17	888794	9月5日	9月1日	4	3	
18	888795	9月6日	8月27日	10	4	
19	888796	9月6日	9月5日	1	4	
20	888797	9月6日	9月4日	2	5	
21	888798	9月6日	調査中	#VALUE!	10	
22	888799	9月6日	8月22日	15	15	
23	888800	9月6日	9月2日	4	#VALUE!	
24	888801	9月6日	9月4日	2	#VALUE!	
25	888802	9月6日	9月1日	5	#VALUE!	
26	888803	9月6日	調査中	#VALUE!	#VALUE!	
27	888804	9月6日	9月5日	1	#VALUE!	
28	888805	9月6日	9月3日	3	#VALUE!	
29	888806	9月6日	9月4日	2	#VALUE!	
30	888807	9月6日	9月4日	2	#VALUE!	

補足. 県ごとの遅れ平均日数と標準偏差

都県のオープンデータを基にした流行波ごとの遅れ平均日数(標準偏差日数)

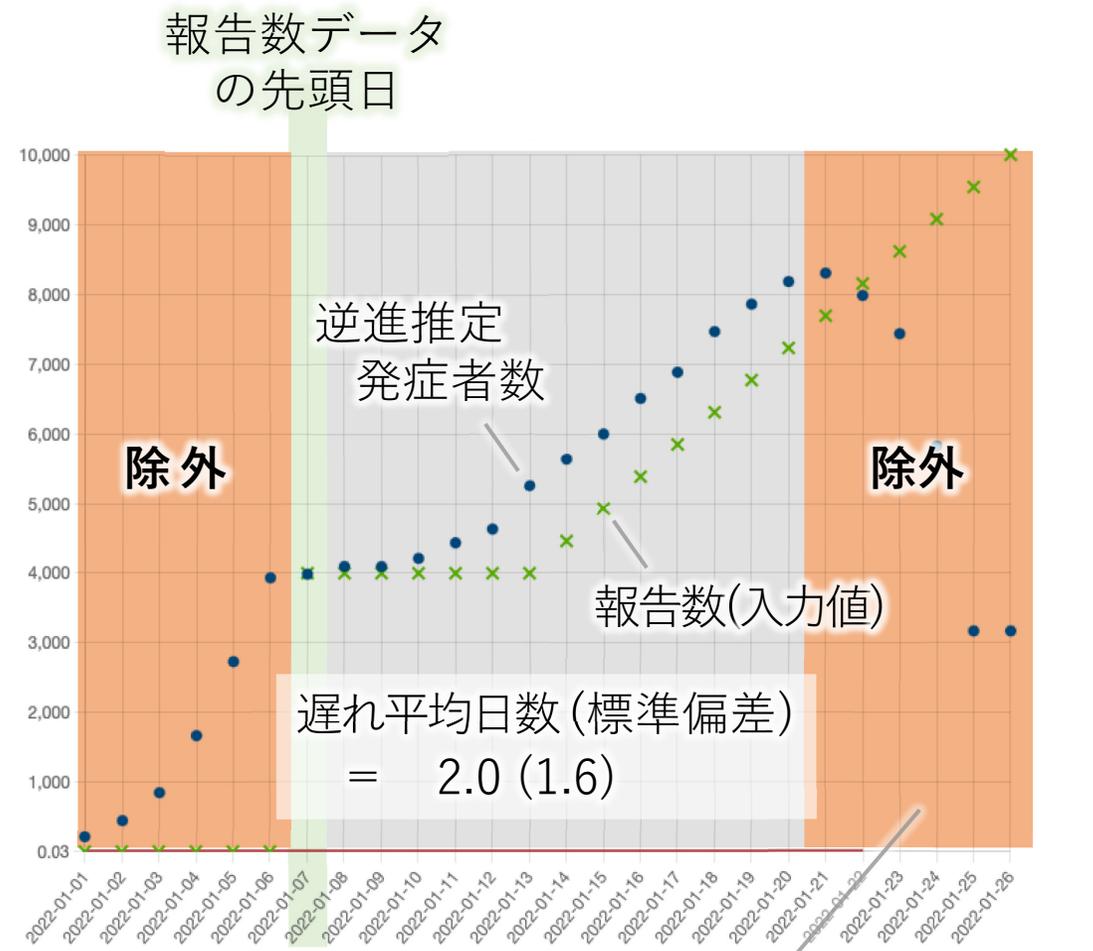
	東京	宮城	兵庫	三重	岐阜	福井	沖縄
第1波		5.63 (3.14)			6.25 (3.74)	3.82 (2.86)	7.21 (3.77)
第2波	5.57 (4.38)	4.21 (3.52)	7.20 (4.28)		4.52 (3.19)	3.38 (2.35)	4.52 (3.15)
第3波	5.35 (4.03)	4.32 (3.39)	5.63 (3.26)		3.17 (3.02)	3.86 (1.73)	4.28 (2.89)
第4波	4.79 (5.31)	4.07 (2.91)	4.31 (3.06)		2.91 (2.82)	3.77 (1.77)	3.74 (2.79)
第5波	4.44 (3.53)	3.71 (2.41)	4.31 (2.58)	4.06 (2.31)	2.28 (2.24)	3.96 (2.18)	3.65 (2.73)
第6波	3.91 (2.52)	3.15 (1.84)	4.79 (3.82)	3.18 (1.66)	1.70 (1.66)		2.88 (1.87)
第7波	4.19 (2.85)	2.78 (1.67)	5.51 (3.59)	3.08 (2.00)	1.29 (1.24)		3.46 (2.69)
第8波		3.12 (2.79)					2.93 (3.25)



期間の定義	
第1波	2020-02-01 ~ 2020-06-30
第2波	2020-07-01 ~ 2020-10-31
第3波	2020-11-01 ~ 2021-02-28
第4波	2021-03-01 ~ 2021-06-30
第5波	2021-07-01 ~ 2021-12-31
第6波	2022-01-01 ~ 2022-06-30
第7波	2022-07-01 ~ 2022-09-30
第8波	2022-10-01 ~

県や流行波ごとに固有の遅れ日数となっております。波を追うごとに遅れ日数が縮小する傾向はありますが、全国的な標準値があるわけではないようです。

補足. 逆進推定発症者数データの末端の扱い



遅れ平均日数 + 標準偏差日数の2倍
= $2.0 + 2 \times 1.6 = 5.2$
≒ 6日分の右側末尾データを除外

逆進推定の原理上、**データの新しい側の末端**では、発症者数を**正確に推定できません**。
(まだ報告されていない発症者が多くなるため)

左図のように、右端のデータ領域(橙色部分)は明らかに**少ない値**となっており、正しく逆進推定されている発症者数のデータ領域は全体的に報告数のおおよそ真左にずれています。

COVIVISでは、この**右側末尾は自動的に除外**されますのでご注意ください。除外範囲は「**遅れ平均 + 標準偏差の2倍**」を整数で繰り上げた日数となります。

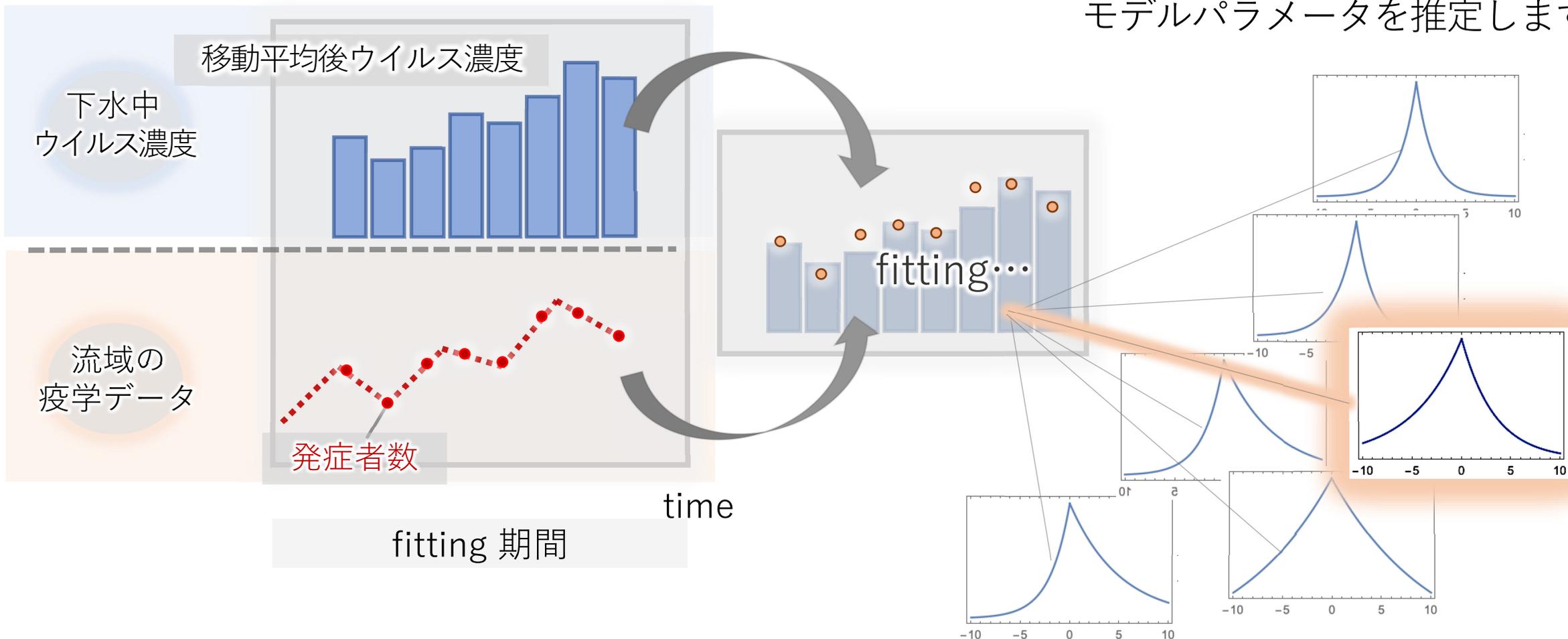
ウイルス濃度データが除外範囲にならないように、**報告数データは長い目に入力して**してください。

一方、逆進推定発症者数データの**左側末端(最も過去)**は、**報告数データの先頭日**を境に除外されます。

手順 2-6～3-1. モデルパラメータ推定の実行

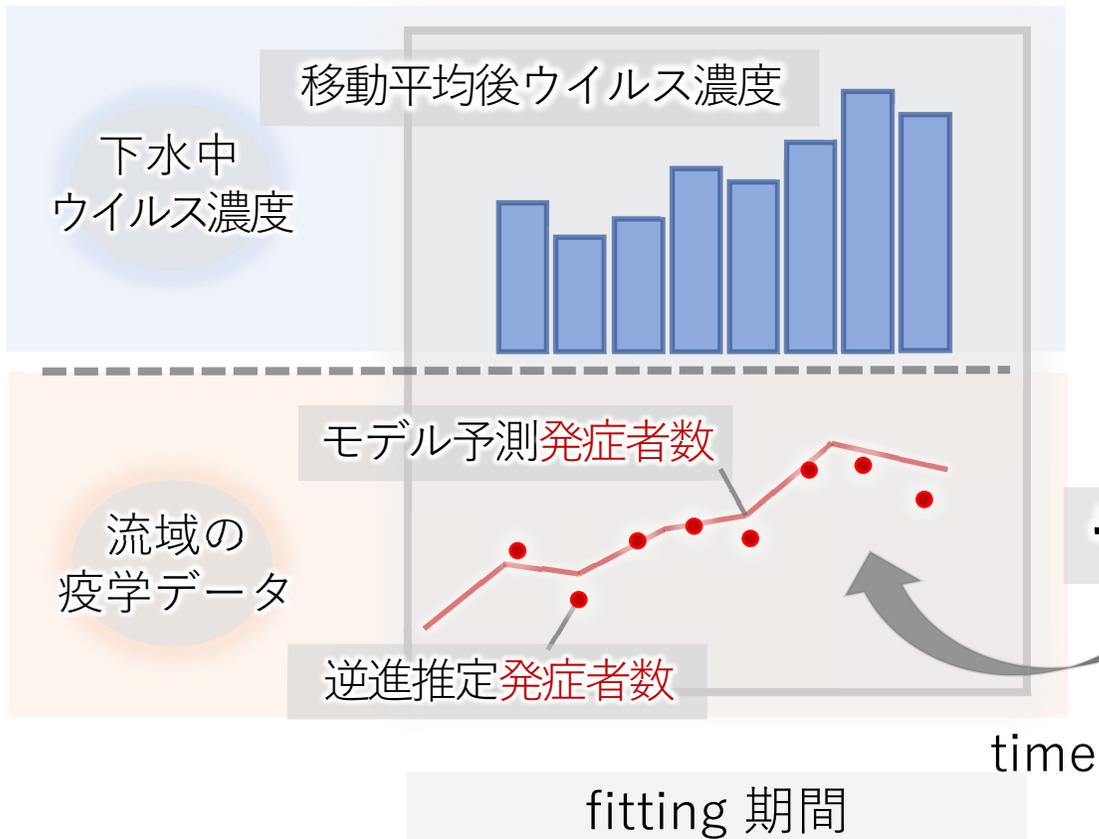
パラメータ推定

- ・ウイルス濃度と発症者数が最も一致するモデルパラメータを推定します。

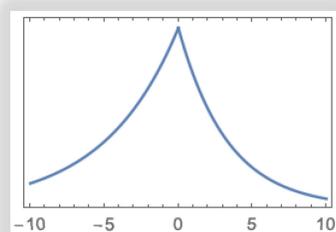


手順 2-6～3-1. モデルパラメータ推定の実行

パラメータ推定



- ウイルス濃度と発症者数を最もフィッティングさせるモデルパラメータを推定します。



ν	ω	γ
15.7	2.20	2.09

モデルパラメータ

逆進推定発症者数

報告数をワイブル分布に従って逆進させたもの

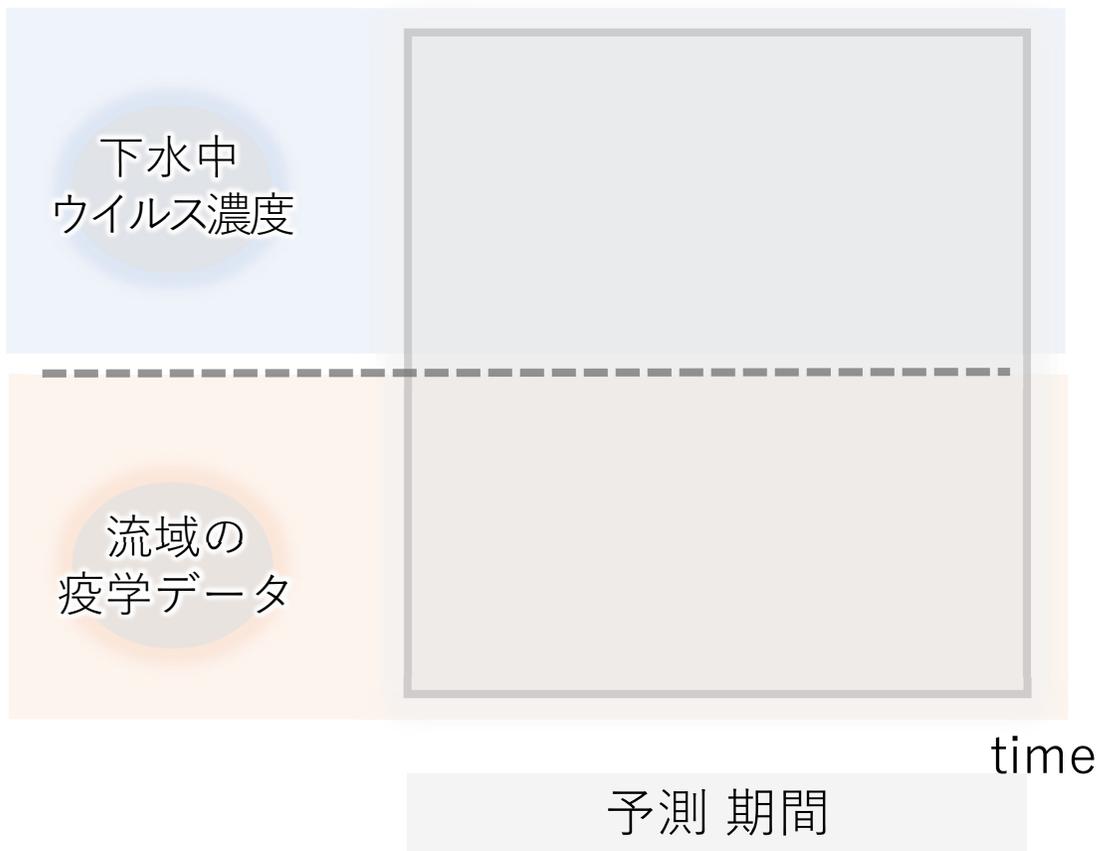
モデル予測発症者数

モデルによって下水中ウイルス濃度から予測されたもの

- ウイルス濃度と推定したモデルパラメータからモデル予測された発症者数の時系列データも合わせて出力されます。

手順 4-1~4-4. 発症者数の予測をする

発症者数予測



発症者数の予測→をクリック



手順 4-1~4-4. 発症者数の予測をする

発症者数予測

モデルパラメータ

ν	ω	γ
15.7	2.20	2.09

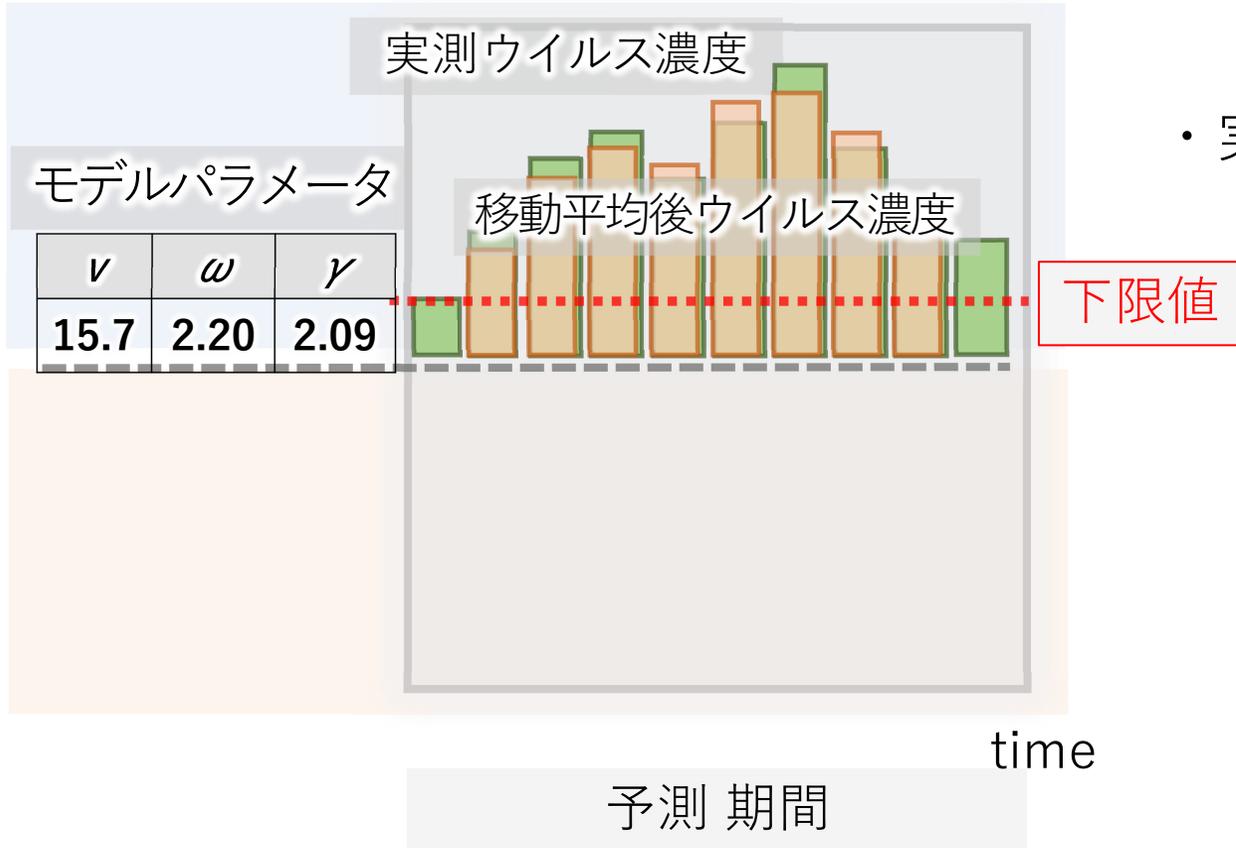
time

予測 期間

- ・ 推定と同じモデルを選択し、
モデルパラメータを入力します。

手順 4-1~4-4. 発症者数の予測をする

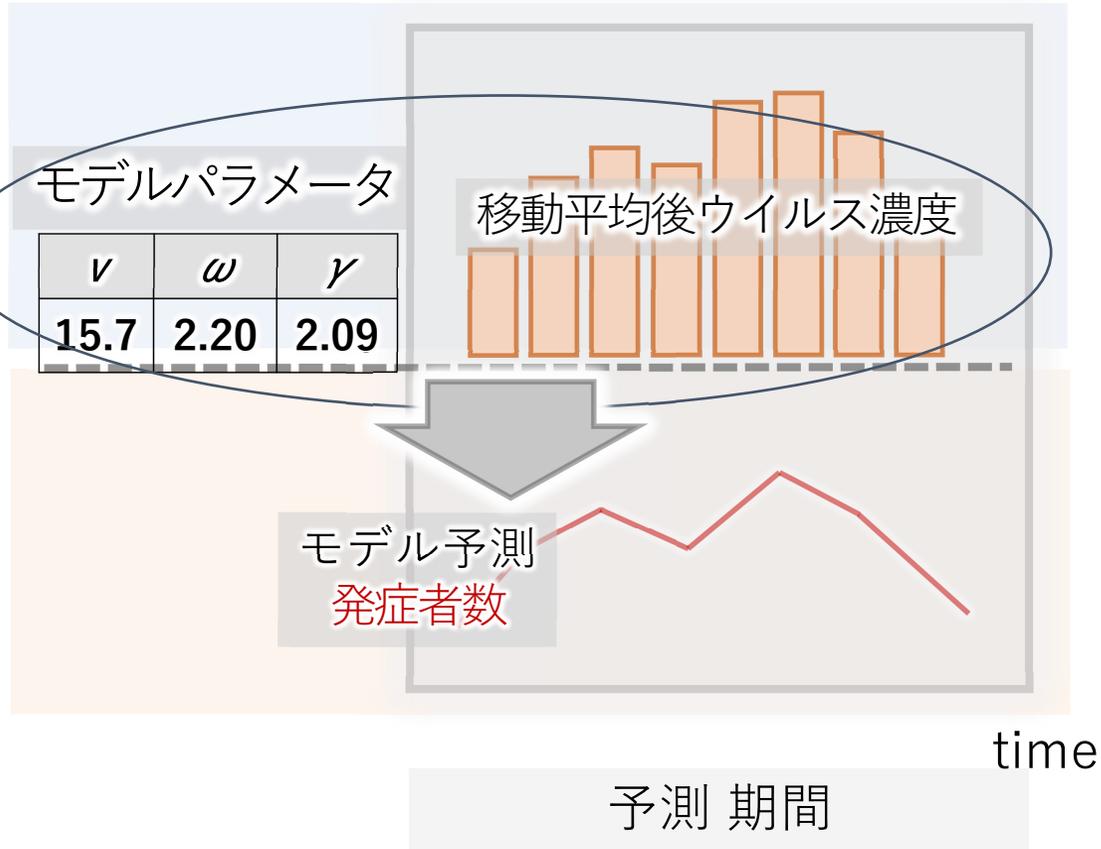
発症者数予測



- 推定と同じモデルを選択し、
モデルパラメータを入力します。
- 実測ウイルス濃度データを入力し、
ウイルス濃度下限値、移動平均の設定をします。

手順 4-1~4-4. 発症者数の予測をする

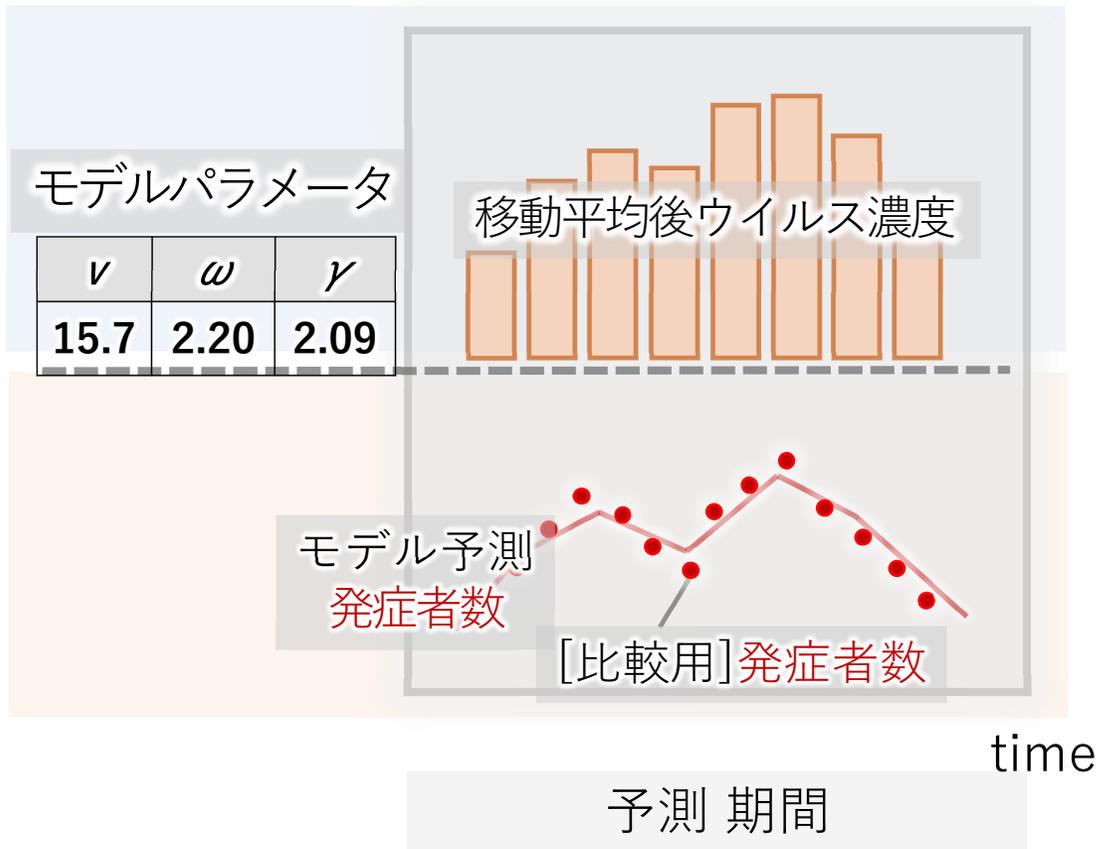
発症者数予測



- 推定と同じモデルを選択し、
モデルパラメータを入力します。
- 実測ウイルス濃度データを入力し、
ウイルス濃度下限値、移動平均の設定をします。
- モデル予測発症者数(予測期間)を出力します。

手順 4-1~4-4. 発症者数の予測をする

発症者数予測



- 推定と同じモデルを選択し、
モデルパラメータを入力します。
- 実測ウイルス濃度データを入力し、
ウイルス濃度下限値、移動平均の設定をします。
- モデル予測発症者数(予測期間)を出力します。
- モデル予測と実際の発症者数を比較したい時は、
追加で予測期間報告数/発症者数データを入力可能。

(※ 詳細は手順 2 を参照)

補足. 報告数データの扱いについて

上記のように、COVIVISでは採水日ベースの下水中ウイルス濃度データと発症日ベースの疫学データをフィッティングします。



COVIVISでは、報告数データに発症日から公表日までの遅れ平均日数とその標準偏差を設定することで、その地域、その流行時期ごとの発症日ベースの疫学データとして解析に使用することができます。

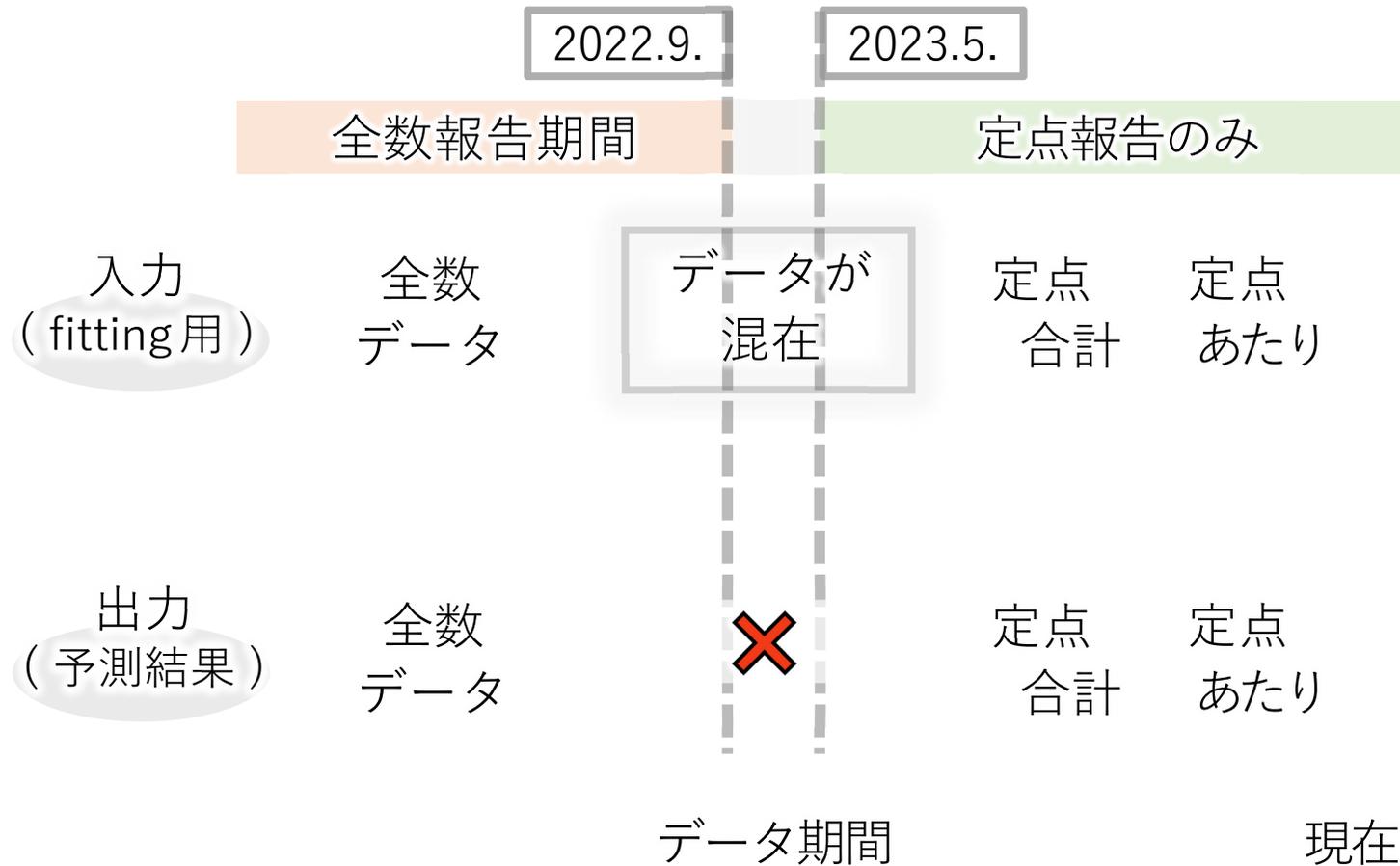
感染者数に関連する日付は多く、定義が曖昧なデータも散見されます。使用する疫学データの日付が何を指し示しているのかご注意ください。ほとんどの疫学データは、公表日をベースに陽性報告数を集計しています。

旧Her-sys等のデータを利用する際には、発症日で集計したデータが十分なサンプルサイズであれば、発症日データを**直接入力**すれば良いです。操作マニュアル2-5)参照

ただし、旧Her-sysにおいて発症日が判明しているデータは全体の半分程度です。期間によっては発症日がほぼ空欄の場合もあります。また、定点報告データには発症日の属性データはありません。

これらの場合、報告数データから発症日を**逆進推定する必要**があります。

補足. 全数データと定点データについて



データの時期によって**全数報告**か**定点報告**か種類が異なります。

COVIVISでは、**どのデータの種類のでも使用可能**です。同じように入力してください。

ただし、全数データで推定したモデルパラメータで予測すると、全数換算での発症者数が出ます。定点あたりデータで推定したパラメータでは定点あたり換算の発症者数の予測となります。

また、**全数期間と定点期間を跨ぐデータ**を使用する場合はデータの統一か変換が必要
Q.13, Q.14を参照

補足. 週次集計の定点データの扱い

定点報告数 実際の日次データ			定点報告数 公表データ			定点報告数 集計入力モード	
月曜日	70		月曜日		月曜日	100	
火曜日	90		火曜日		火曜日	100	
水曜日	100		水曜日		水曜日	100	
木曜日	110		木曜日		木曜日	100	
金曜日	130		金曜日		金曜日	100	
土曜日	110		土曜日		土曜日	100	
日曜日	90	報告日	日曜日	700	日曜日	100	
月曜日		↑	月曜日		月曜日		
火曜日			火曜日		火曜日		
水曜日		公表日	水曜日	700	水曜日		
木曜日			木曜日		木曜日		
金曜日			金曜日		金曜日		
土曜日			土曜日		土曜日		
日曜日			日曜日		日曜日		

日付を
付け替え

1週間分
を均等割

1週間の合計数が
次の週に公表される

定点報告データは、**公表日の前の週の1週間分**の報告数の合計となっています。(これをここでは**週次集計データ**と呼びます)

つまり**公表日ベースでは数日以上遅れた日付**として紐づいています。そのため、週次集計データの報告日の代表日として**公表日の前の週の最終日(日曜日)**の日付に付け直してデータを整理してください。

このような日付設定をしておくと、**集計入力モード**(操作マニュアル手順2-5参照)が正しく処理されます。