ETH zürich (but in Basel)

Spatial modelling of viral loads in wastewater: an evaluation of sampling strategies for comprehensive surveillance programs

James D Munday, Julien Riou,
Moritz Wagner, Tim Julien, Christoph Ort, Tanja Stadler

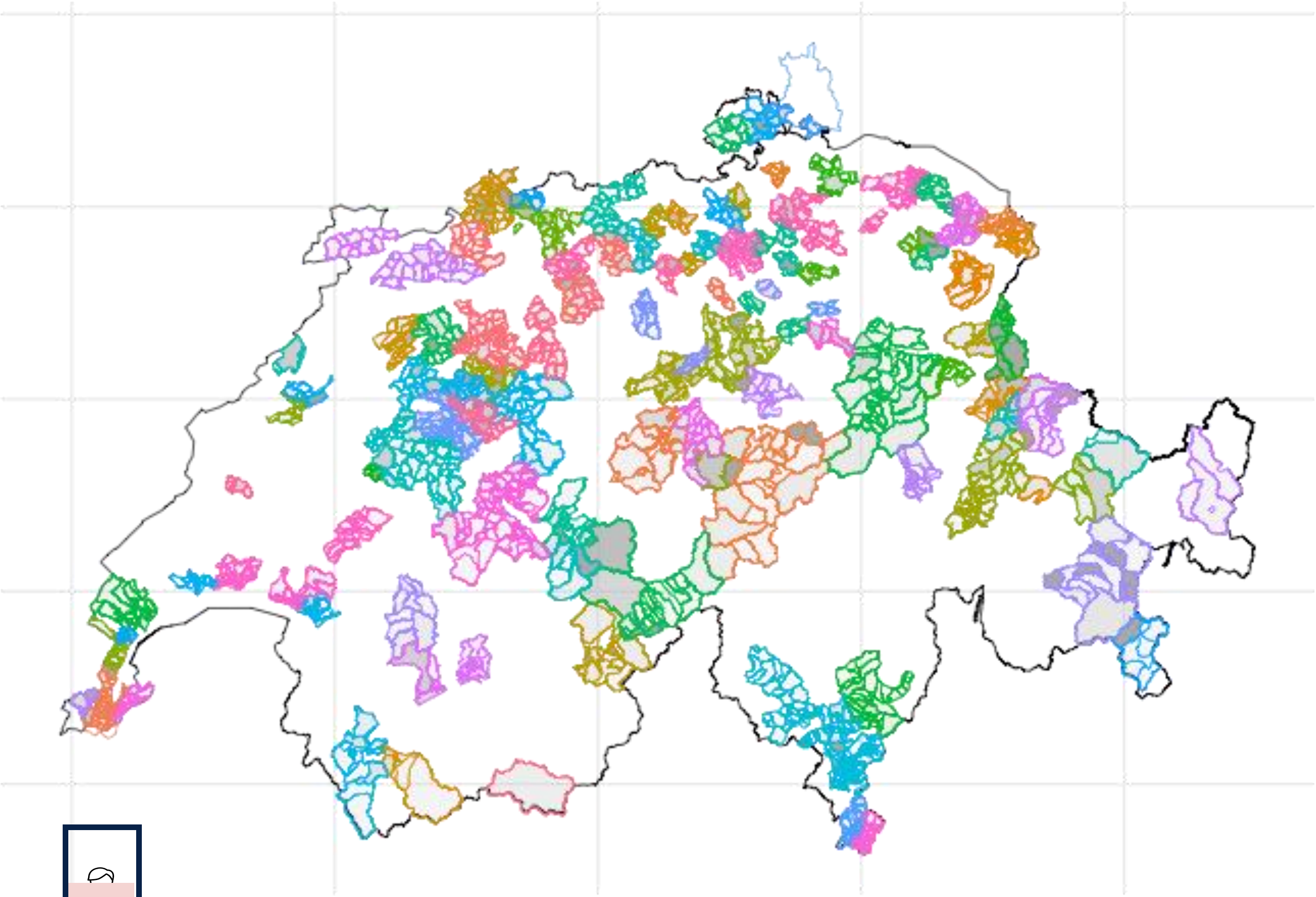
GEOMED | September 2024

Overview

1. Wastewater (WW) based pathogen monitoring in switzerland
2. How might we choose a more sustainable sampling strategy?
   a) Adminastrative regions
   b) Geographical proximity
   c) Mobility data
3. Evaluation of selection strategies with spatio-temporal model
   a) Back validation
   b) Spatial projections to administrative regions (Cantons)
4. Summary

A comprehensive WW monitoring programme

A history of pathogen monitoring in municipal wastewater in Switzerland



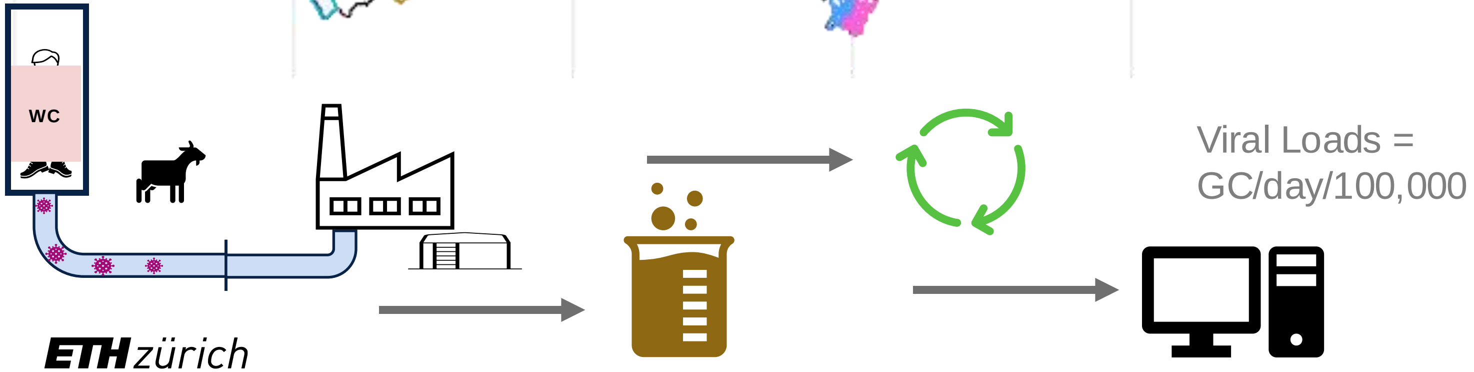
120 Wastewater Treatment Plants (WWTP)

monitored for **SARS-CoV-2** at various times

since March 2020 (most since February 2022 – January 2023) .

Quantification of viral concentrations via

qPCR or ddPCR across **8 Laboratories**.



ETH zürich

February 2020

Shall we track nCoV19 in Wastewater?

Yes

A history of pathogen monitoring in municipal wastewater in Switzerland

Sampling consistency varied

between Feb 2022 – Jan 2023

between WWTPs – here between

104 and 118 WWTP were sampled

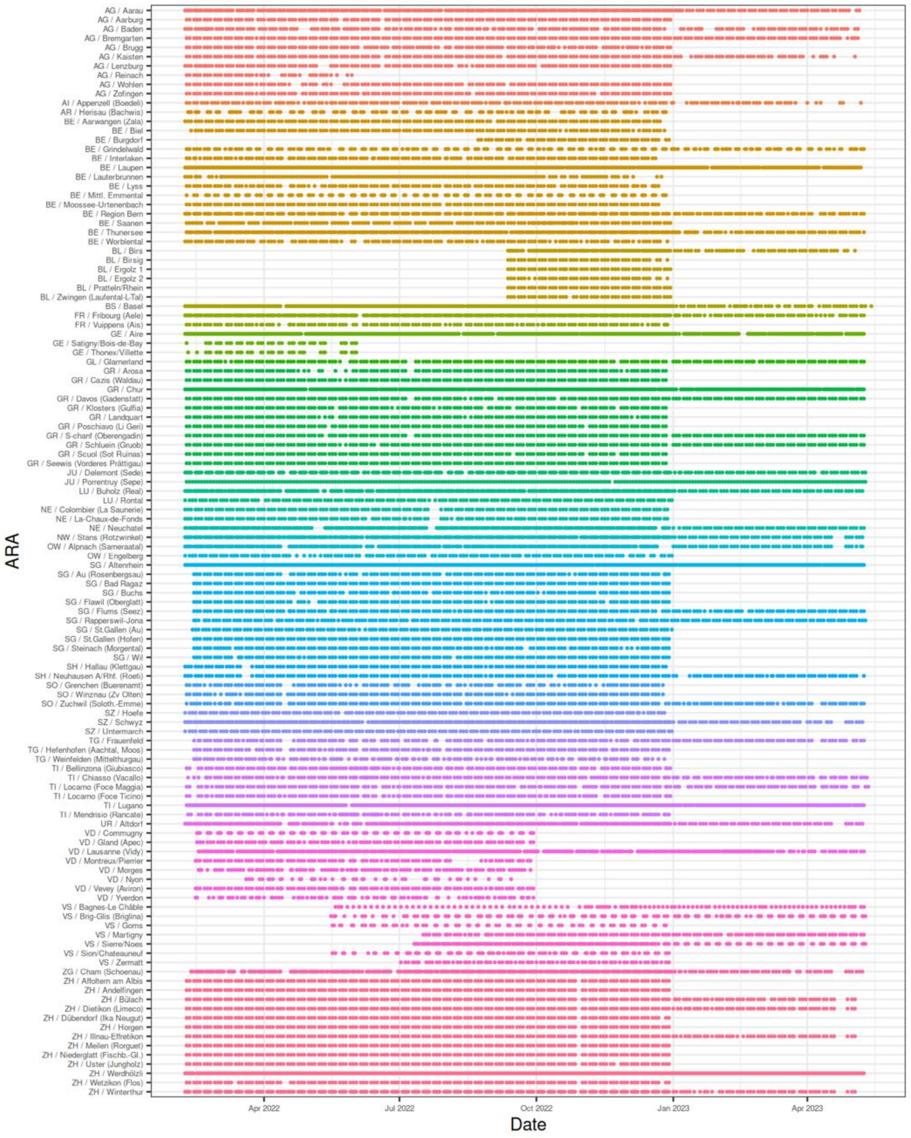
at any given time. After **Jan 2023** the

number was **roughly halved to 50**.

After **July 2023** this was **reduced**

again to 15 and all processing

moved to **one lab (Eawag).**



Site selection criteria

Geographical based selection

Administrative regions + population size

Rational: easily accessible + understood

Likely to hold some geographical information which may relate to transmission

Approach: Highest population catchment in each “NUTS2” region (n=7) plus 3 highest population remaining catchments (10 total)

Spatial cluster based + population size

Rational: captures groups of WWTPs that are highly proximal so likely to share information

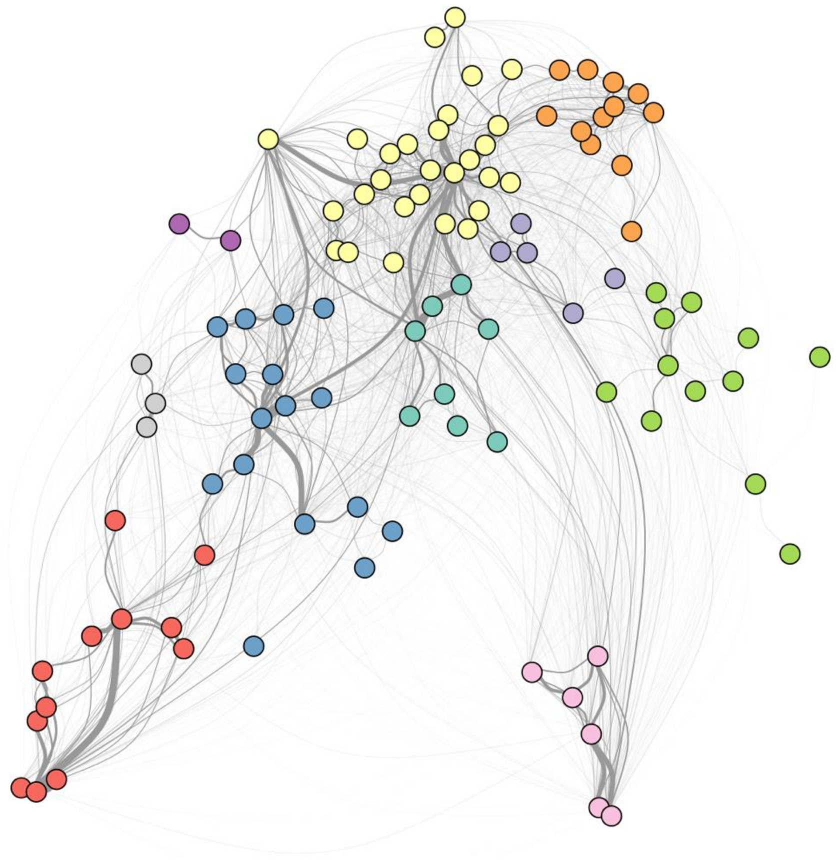
Locations of WWTPs are usually known and even better – their catchments

Approach: Partition catchments based on distance matrix – stipulate 10 ‘communities’ – choose the highest population in each.

Mobility driven selection

Rational: Captures how individuals move between catchments therefore indicating which are likely to be highly correlated based due to transmission and shedding into the Wastewater.

Mobile phone data from the Swiss Sunrise network (35% market share). **'Trips'** defined as a person travelling from **one place to another** and remaining there for **> 30 mins.** We **filtered for regular travellers** and **averaged the number of trips** made between catchments over the analysis period**.**



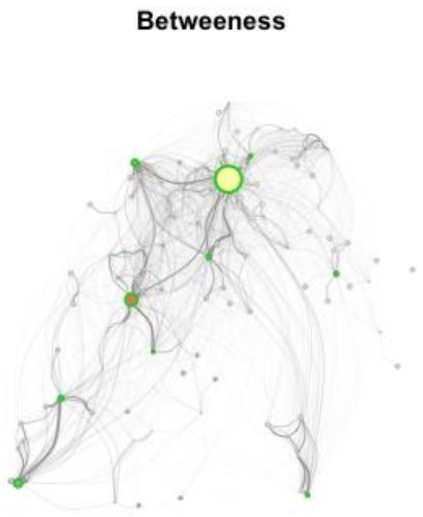
Mobility driven selection

Rational: Captures how individuals move between catchments therefore indicating which are likely to be highly correlated based due to transmission and shedding into the Wastewater.

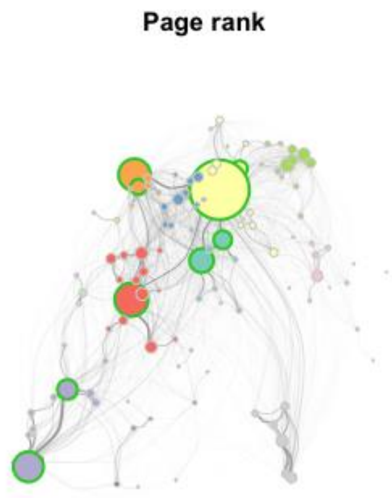
Approaches:

10 highest degree

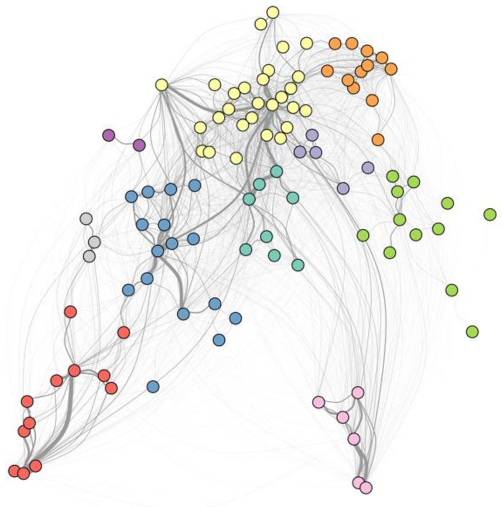
10 highest betweenness



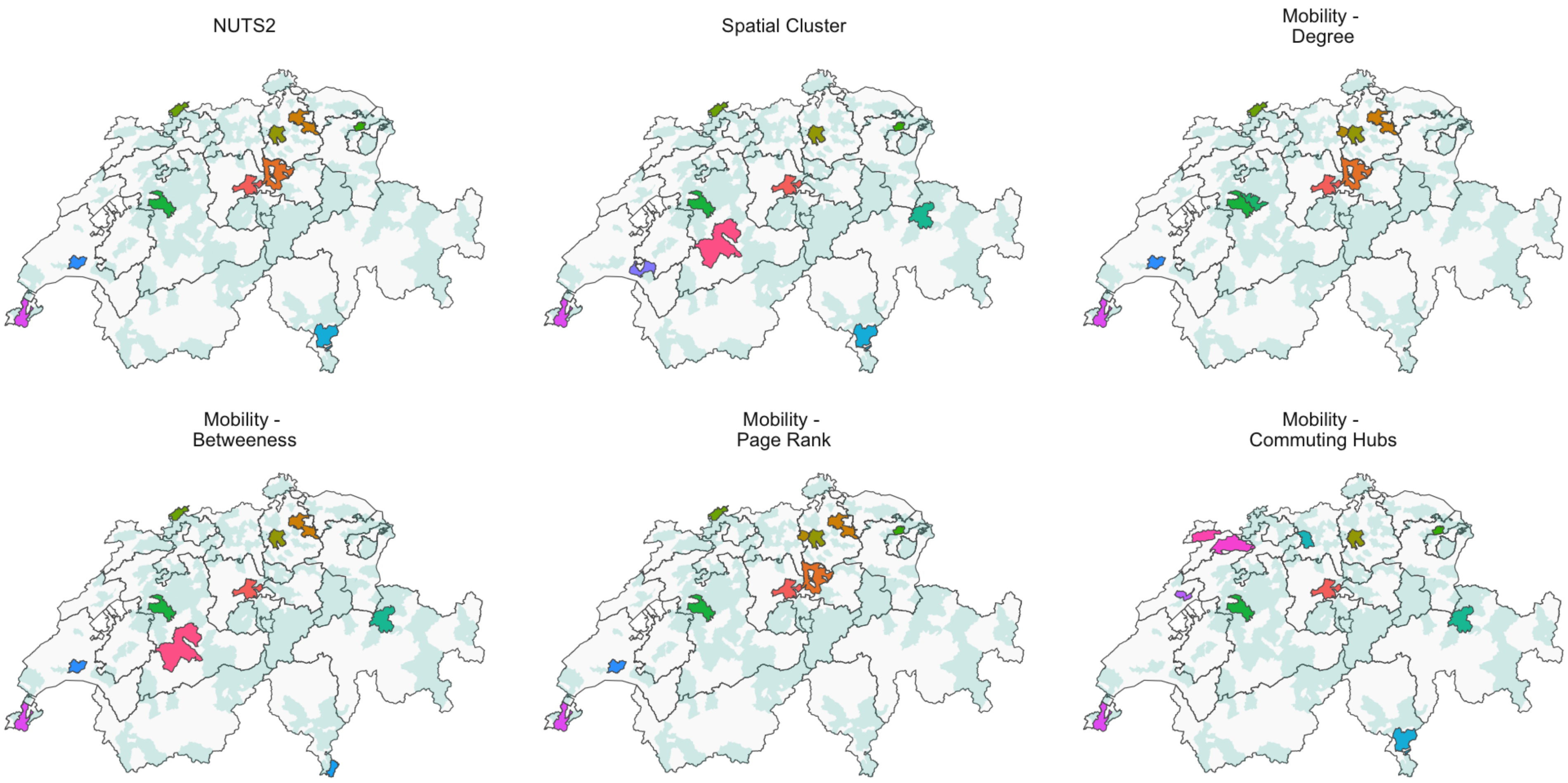
10 highest page rank



Identify ‘commuting hubs’ (info map) – choose the highest degree node from each.



Selected WWTPs



Spatio-temporal model

Modelling framework

$$\log(vl_{it}) \sim Gamma(\mu_{it}, \sigma_{vl}^2)$$

Viral loads per person (10bn’s of gene copies per person)

Intercept (Baseline VL)

Coefficients

Local spatio-temporal component

$$\mu_{it} \sim \alpha + \mathbf{X}_i\beta + \eta_t + z_{it}$$

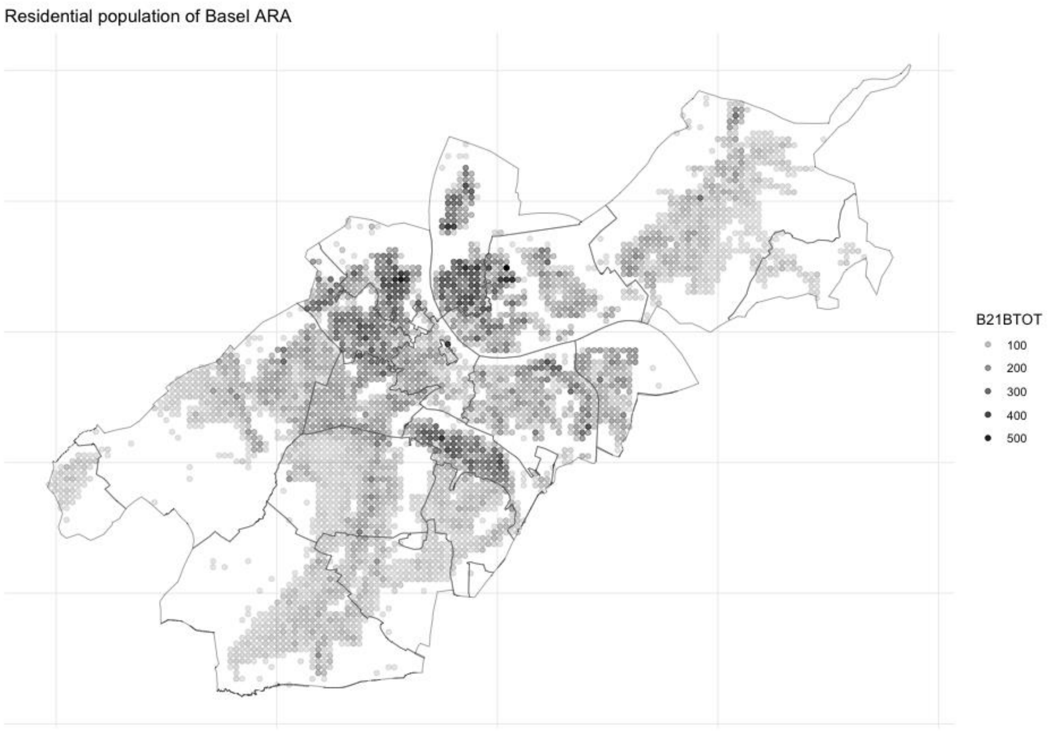
Popn. covariates

Global temporal component (1st order random walk)

Model covariates

1. Population density
2. Proportion under 20yo
3. Proportion over 65yo
4. Proportion born non CH/EU
5. WW processing Lab

$$\mu_{it} \sim \alpha + \mathbf{X}_i\beta + \eta_t + z_{it}$$



Spatiotemporal component

$$\mu_{it} \sim \alpha + \mathbf{X}_i\beta + \eta_t + z_{it}$$

$$t = 0 \qquad \boldsymbol{z_t} \sim Normal(0, \Sigma_z)$$

$$t > 1 \qquad \boldsymbol{z_t} \sim Normal(\rho z_{t-1}, \Sigma_z)$$

Matern covariance function:

(Covariance between the i-th and j-th site)

Scales the overall variance

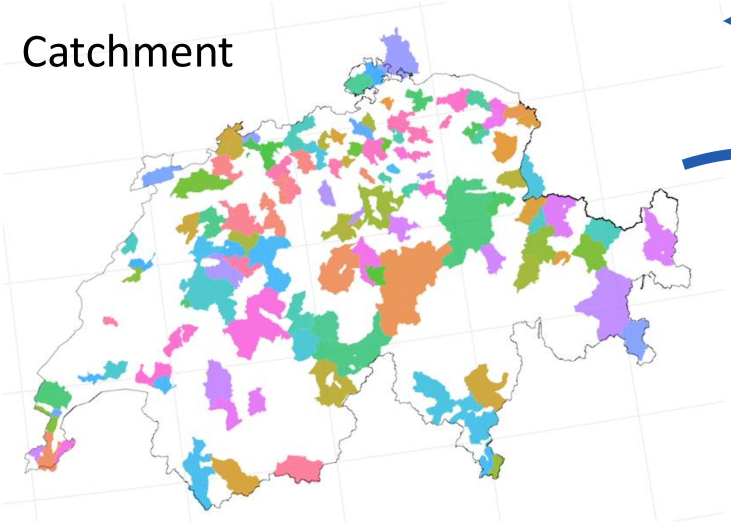
$$\Sigma(z_i, z_j) = \frac{\sigma_z^2}{\Gamma(\lambda)2^{\lambda-1}}\left(kd_{ij}\right)^{\lambda} K_\lambda\left(kd_{ij}\right)$$

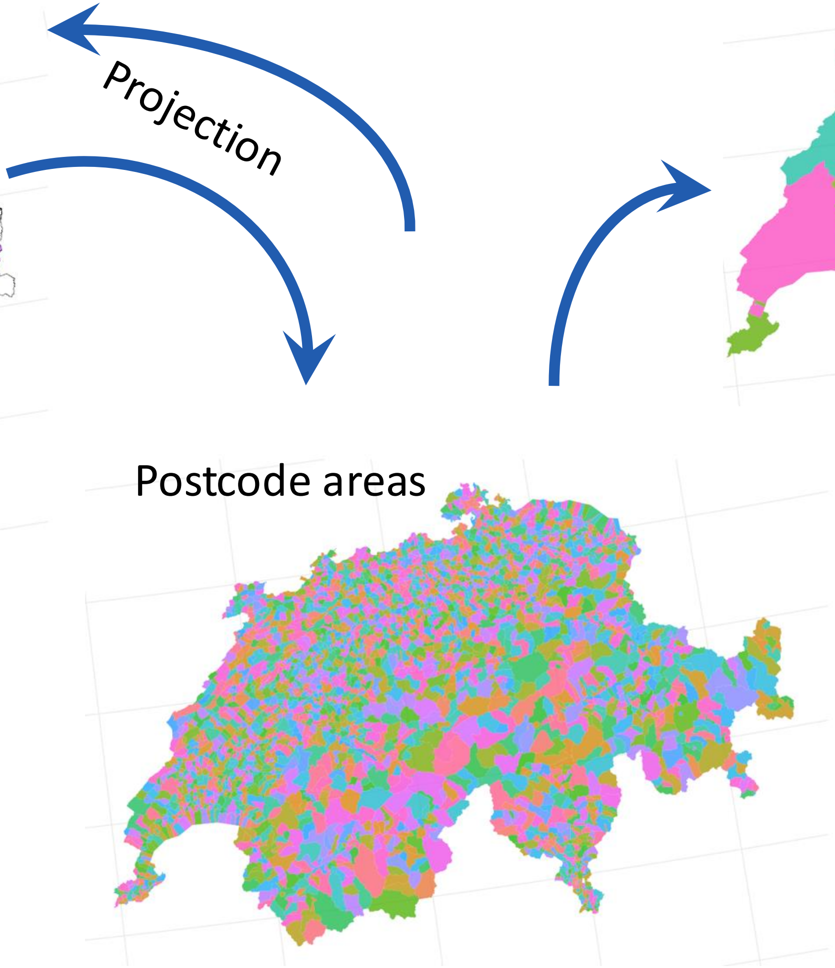
Gamma function

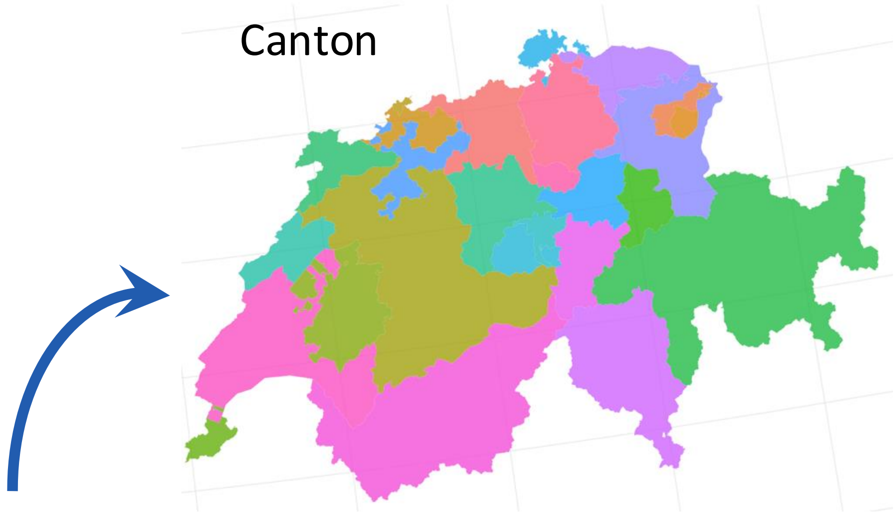
Type 2 Bessel function (order λ)

d – Distance
λ – Smoothness
k - Range parameter

Estimating viral loads at differnect geographic aggregations



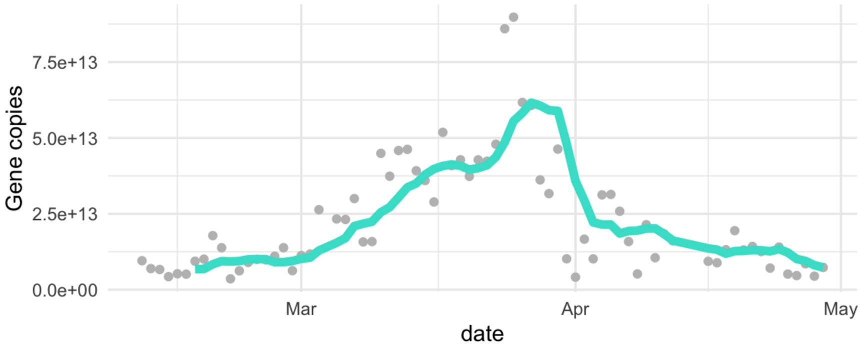




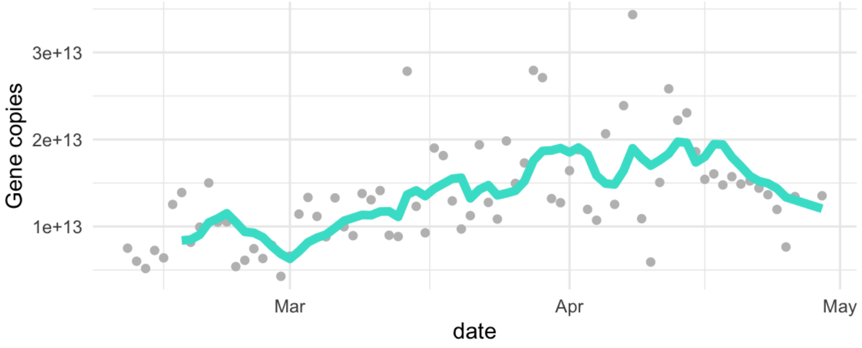
Wastewater data

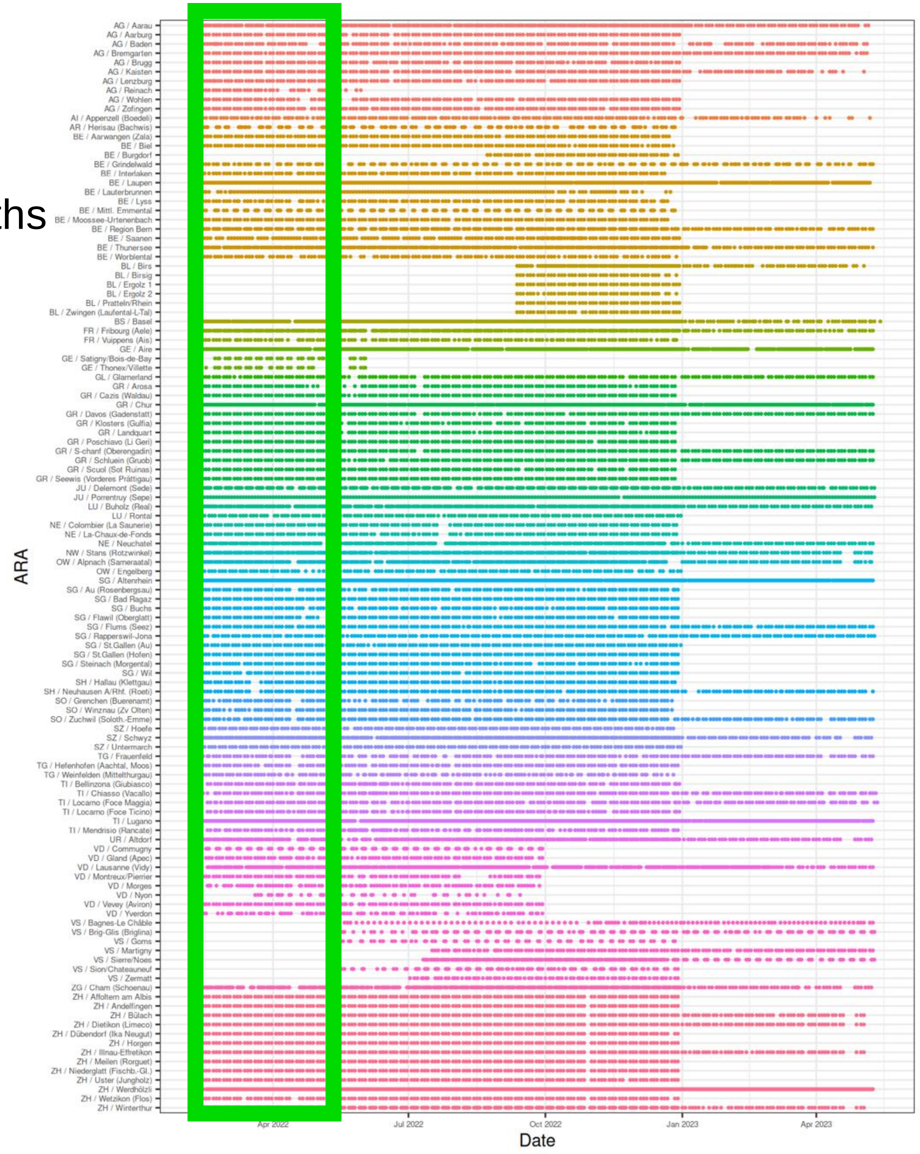
In this work we use concentrations from the very **early comprehensive phase (11th February – 29th April)** - where there were **104 WWTPs** consistently sampled for several months – with no changes in lab protocol.

Daily viral load per 100k population in Basel

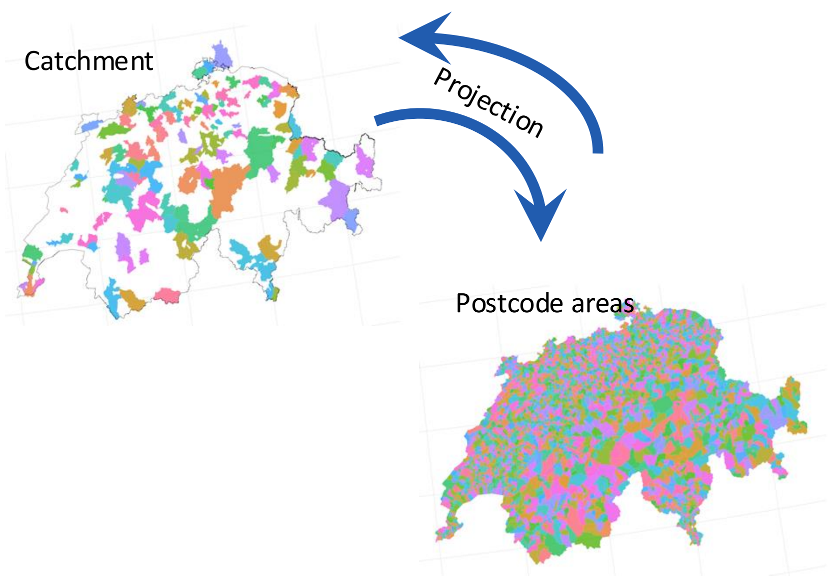


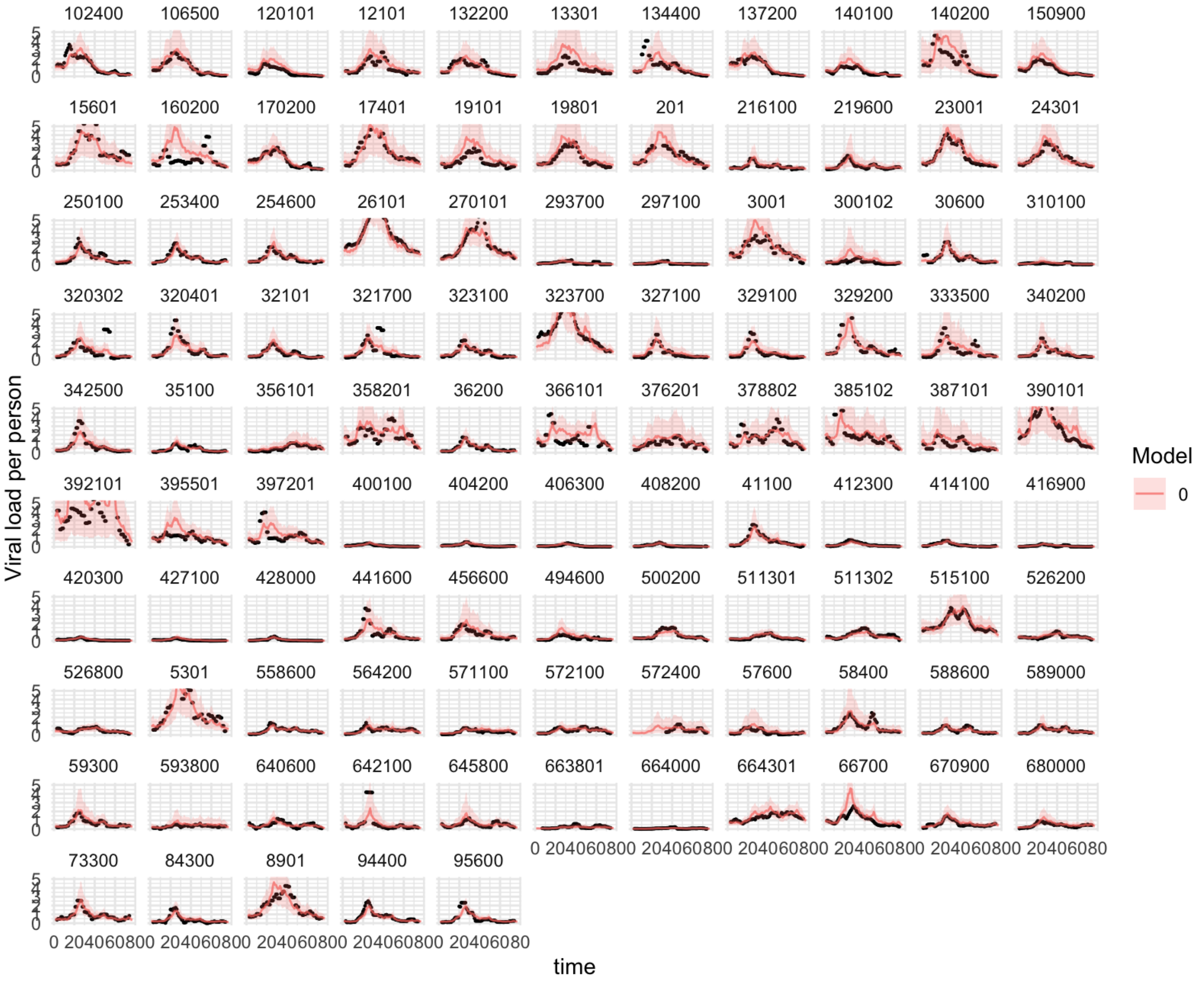
Daily viral load per 100k population in Geneva (Aire)





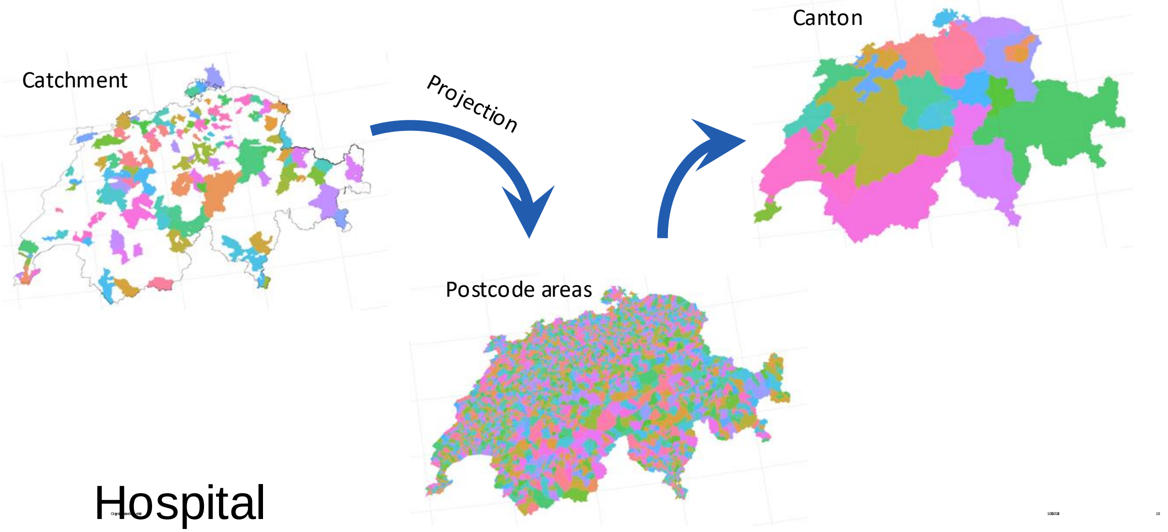
Full model test… Back-validation



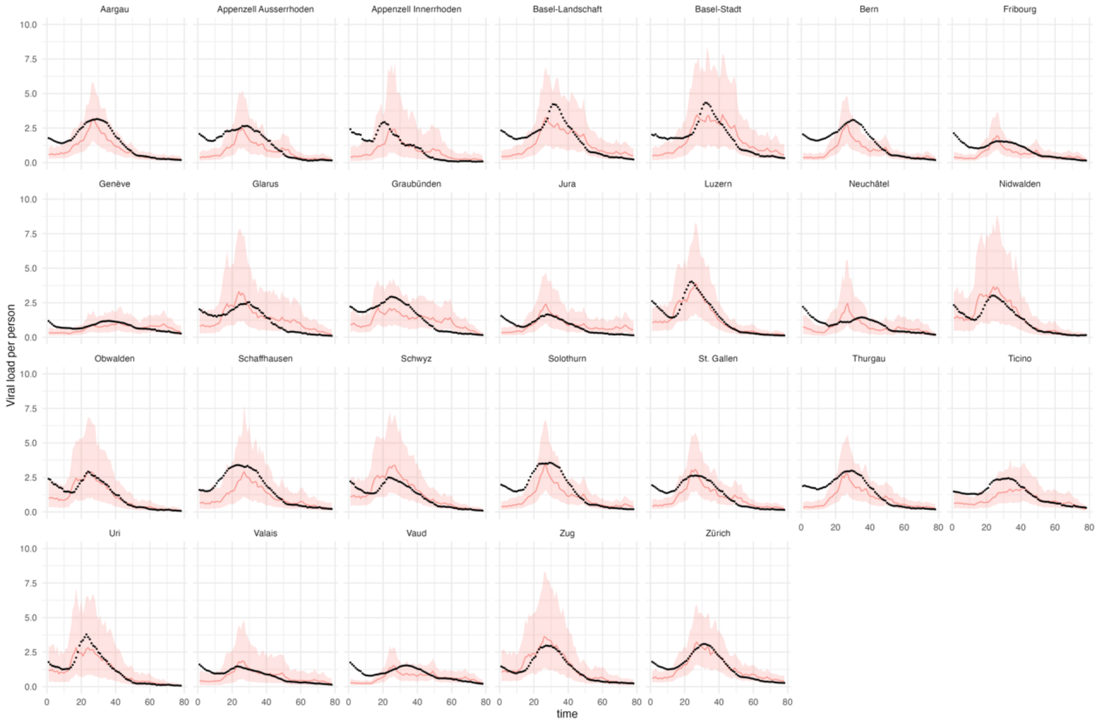


Projections to cantons…

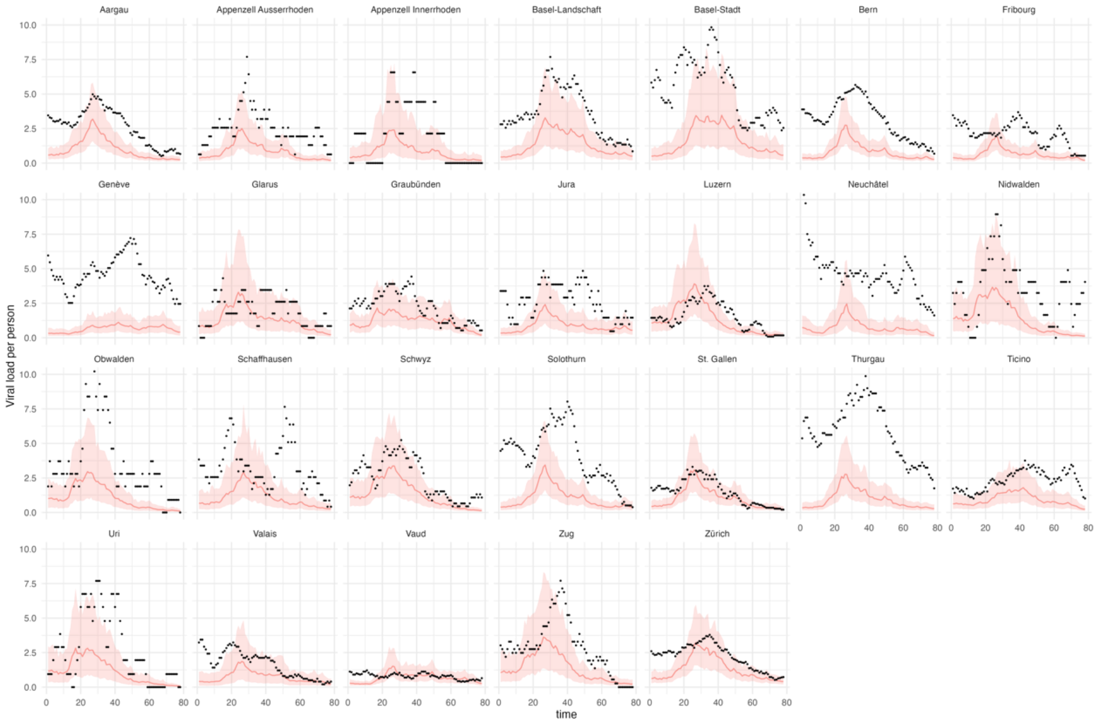
Allows us to compare directly to clinical data- which is not available at catchment level



Cases

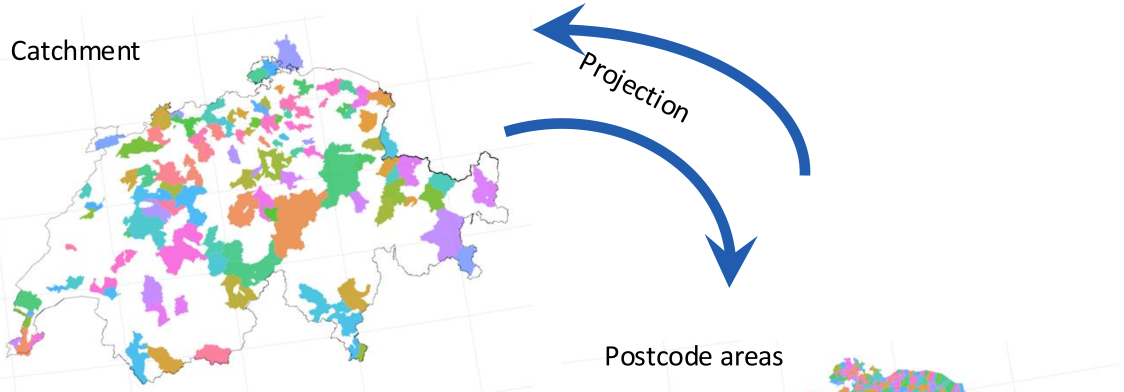


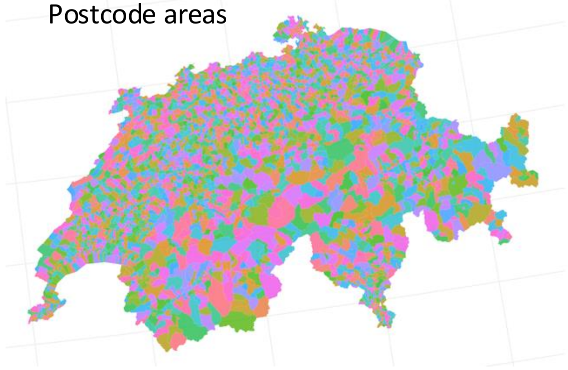
Hospital admissions

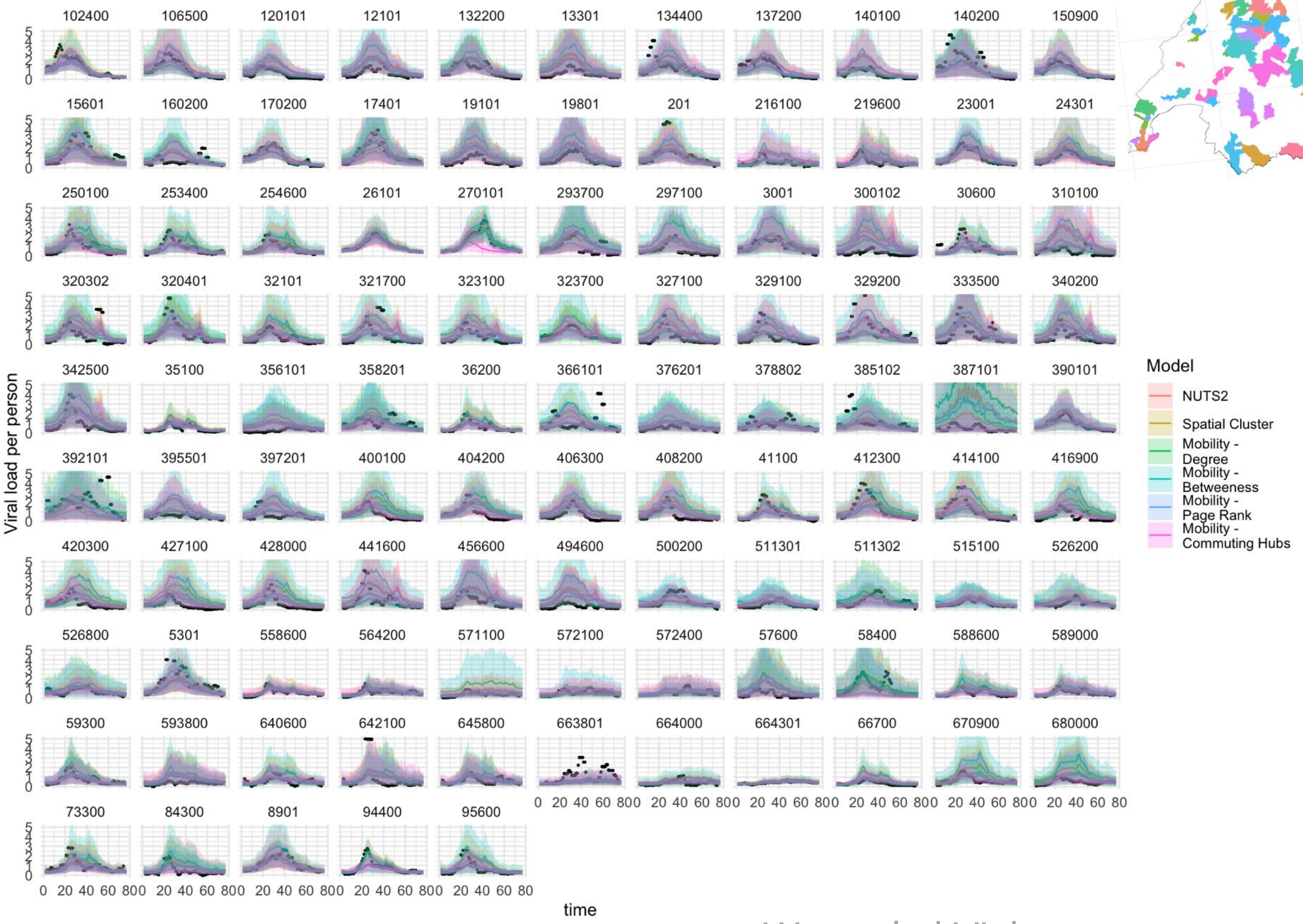


Evaluation of site selection

Re-prediction of WWTP measurements







- Continuous ranked probability score (**CRPS**)
- Absolute error of the median prediction (**AE_median**)
- **Bias** (average proportion over/under predicted)
- Median absoulute dispersion of the prediction (**mad**)

We scaled VL by

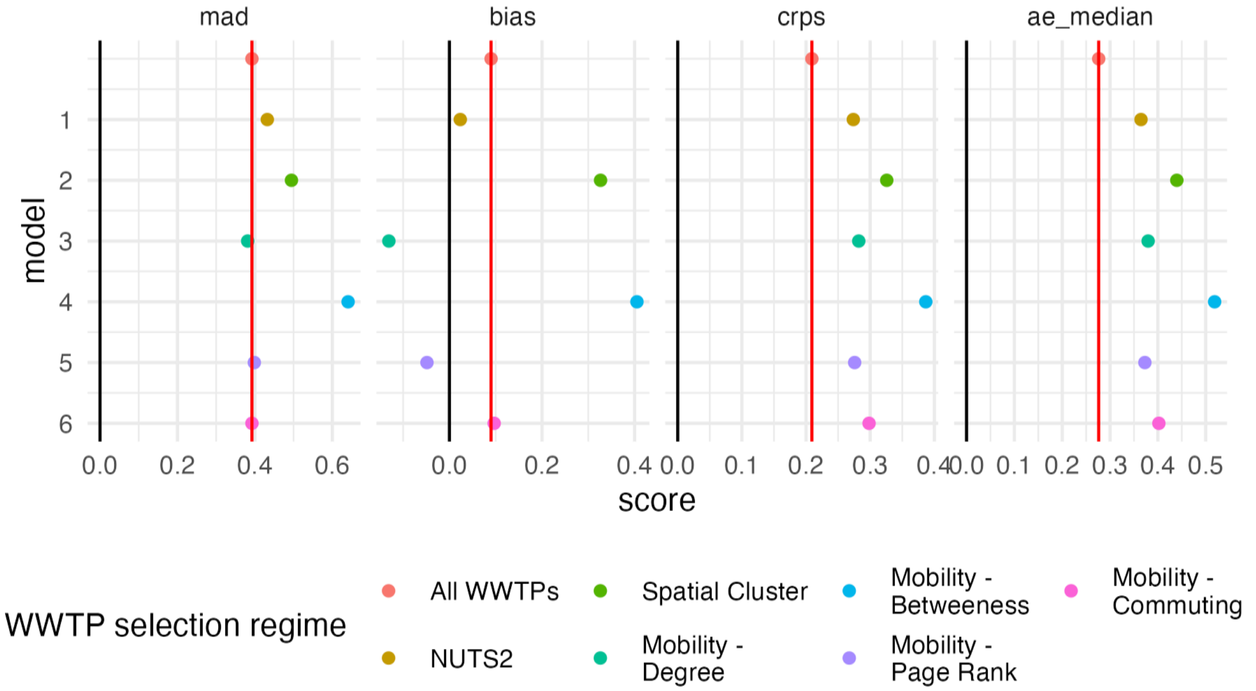
Overall scores

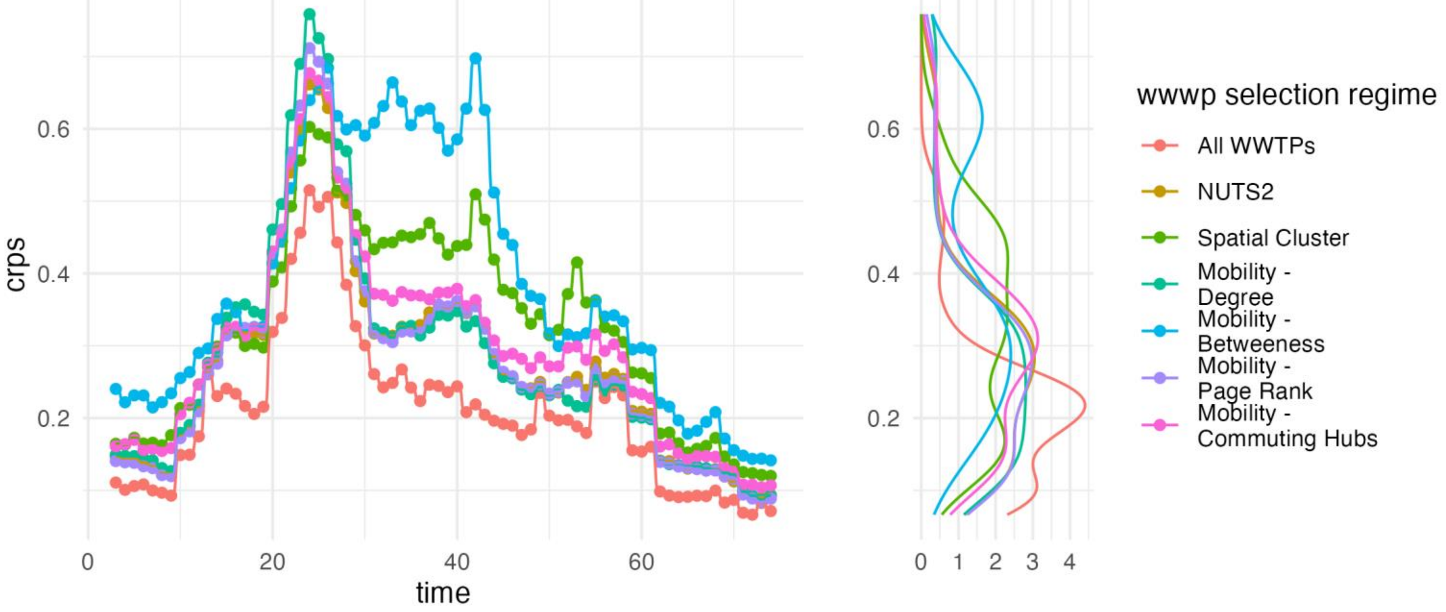
The **betweeness** based selection **performs comparatively poorly** in all measures

Region based selection (NUTS2), Degree and Page Rank appear to perform better than the clustering based approaches

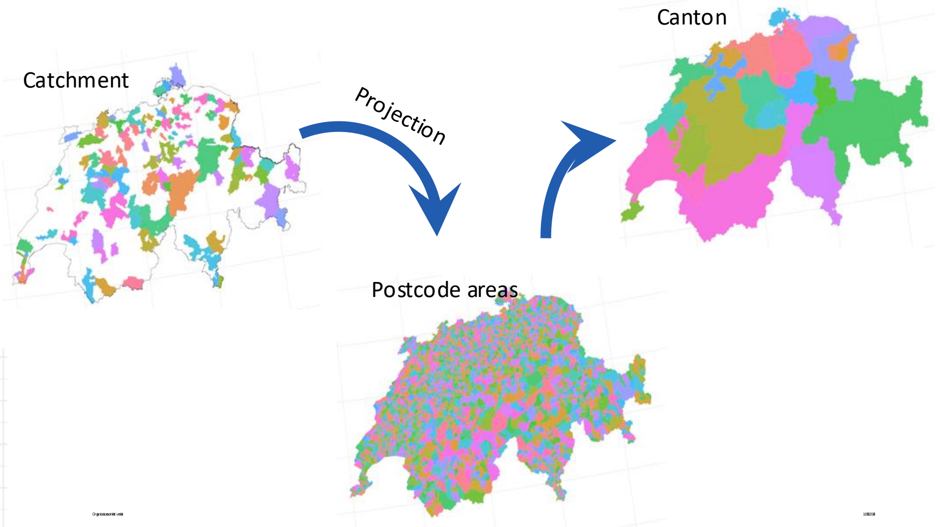
These are both **less biased overall and more accurate** (lower CRPS and absolute error)

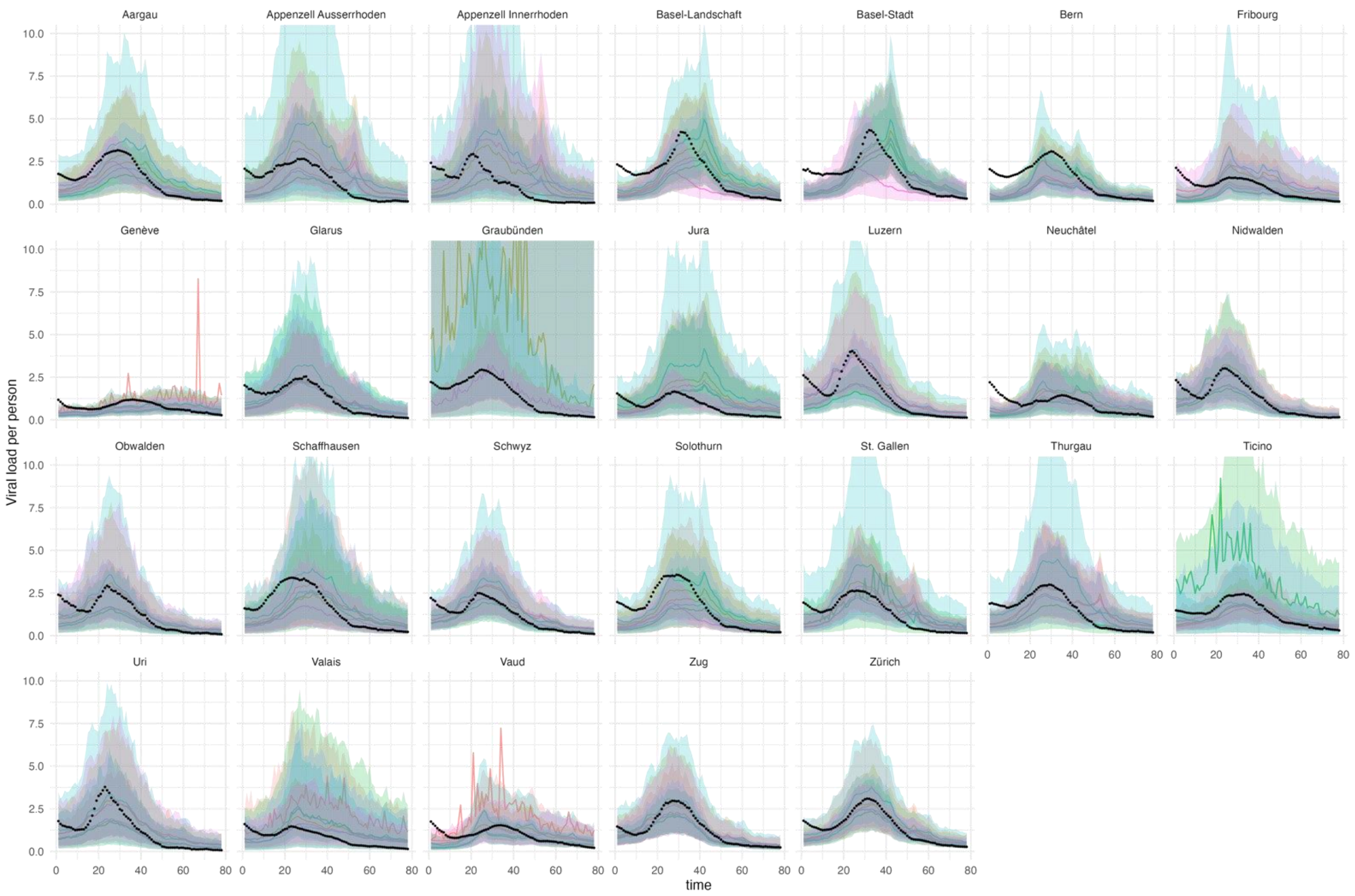
All but the betweeness based selection have **comparable dispersion** to the preditions with **all WWTPs**





Correlation with clinical data – Cases

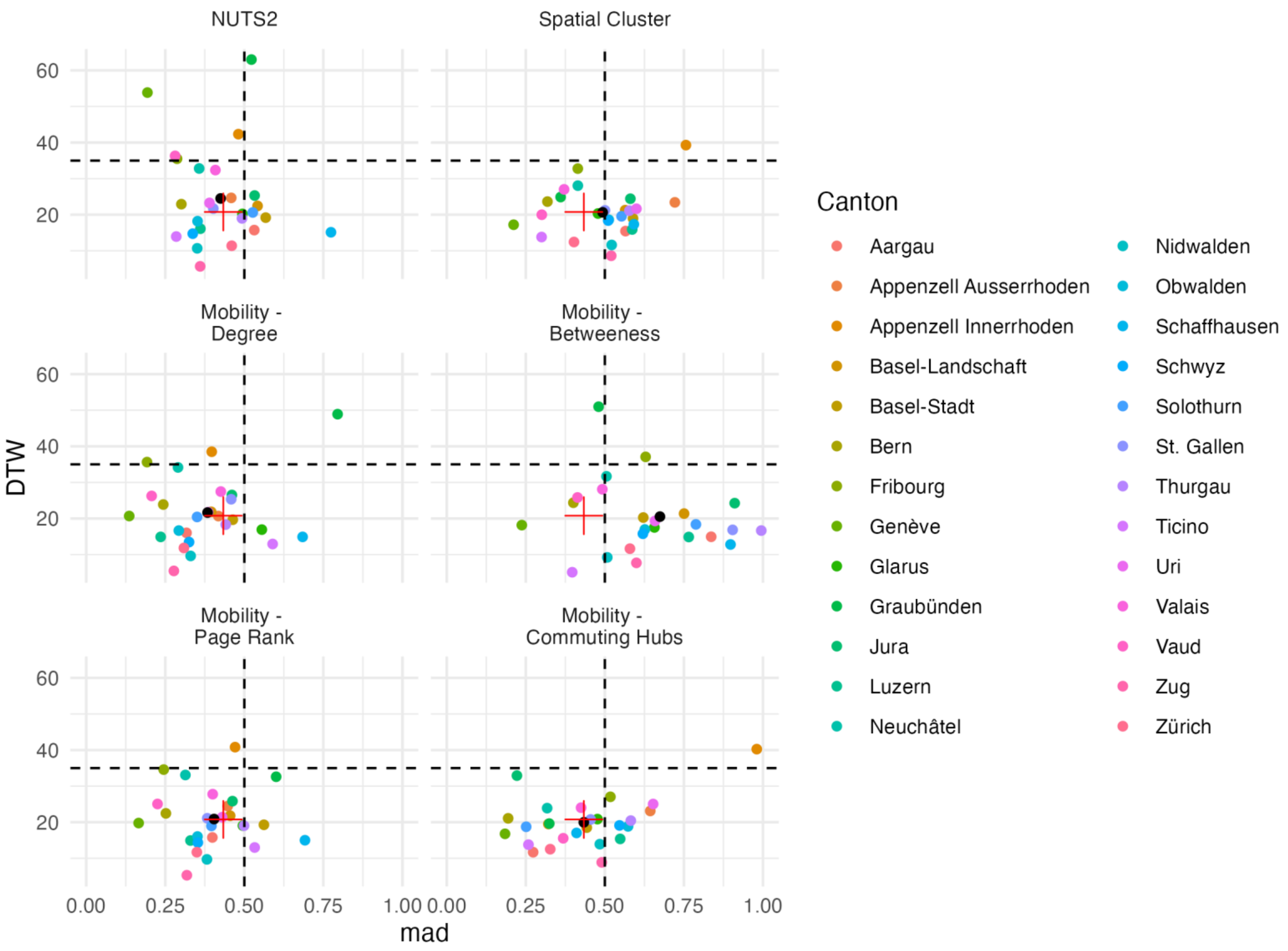




- Dynamic Time Warping (**DTW**) distance between standardised median and case incidence per 100k
- Median absoulute dispersion of the prediction (**mad**)

Correlation with clinical data – Cases

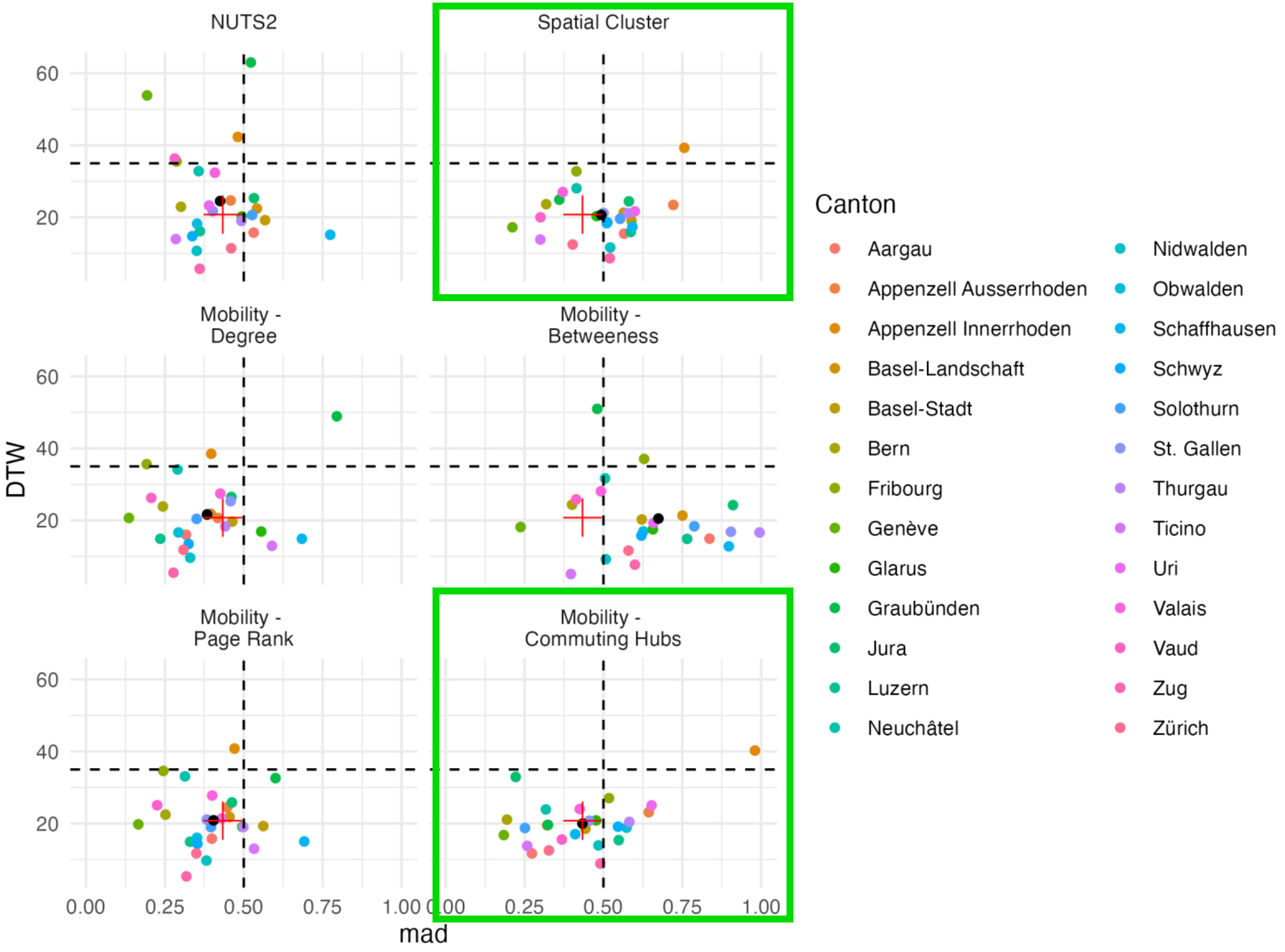
Converse to back-validation – cluster based selection perfoms marginally better than the others when evaluating correlation with case time-series.



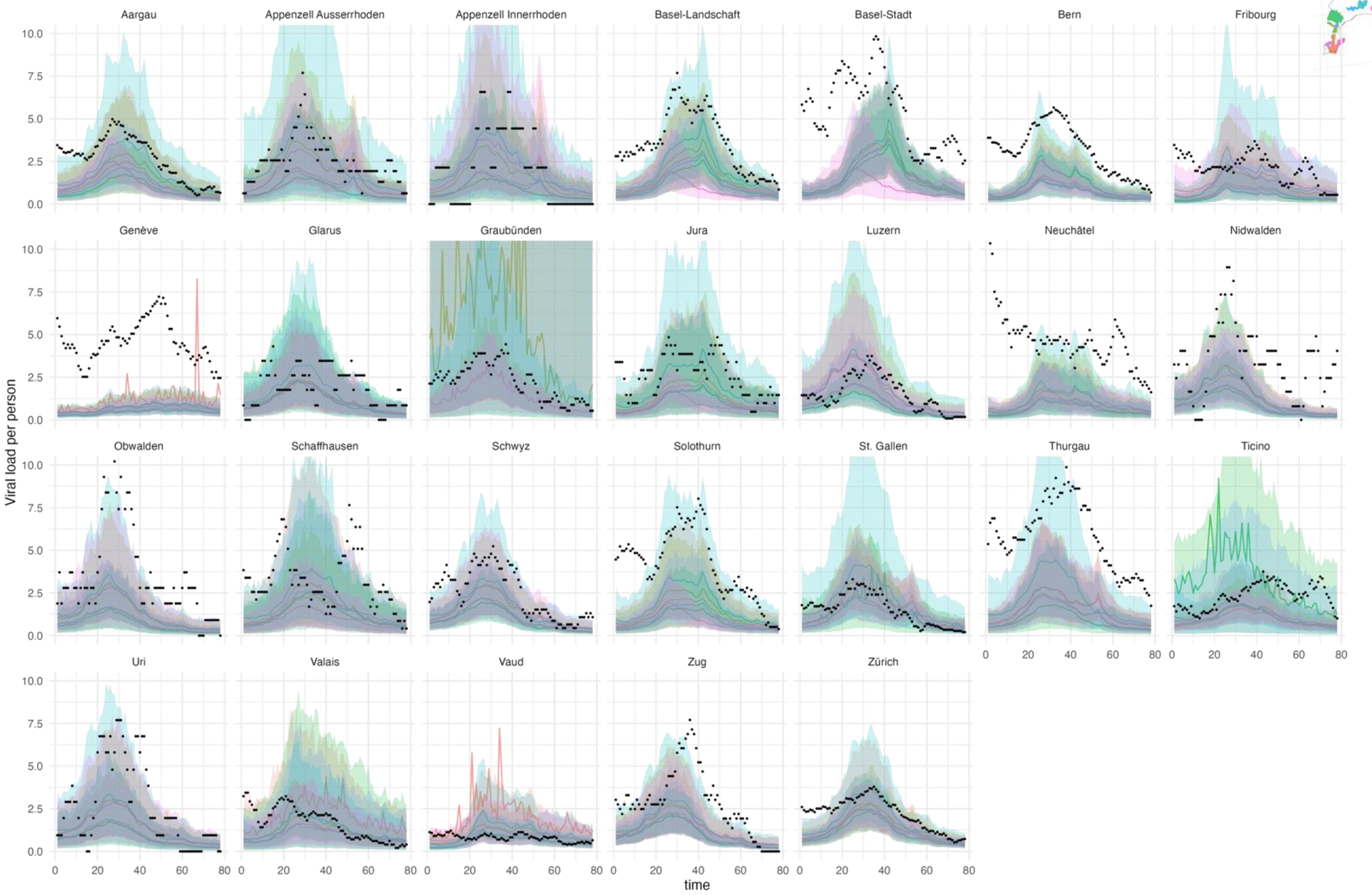
Correlation with clinical data – Cases

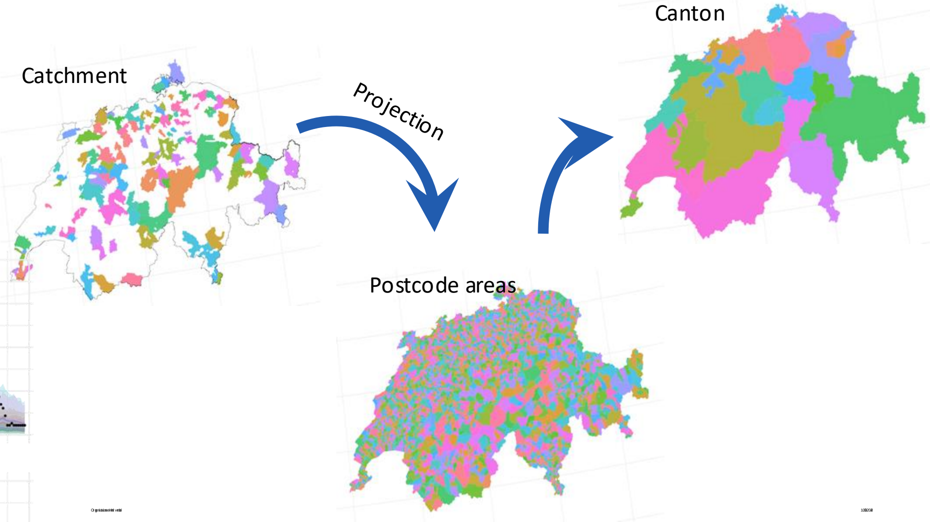
Converse to back-validation – cluster based selection perfoms marginally better than the others when evaluating correlation with case time-series.

The distribution of performance across all cantons is also imporved



Correlation with clinical data - Hospitalisations



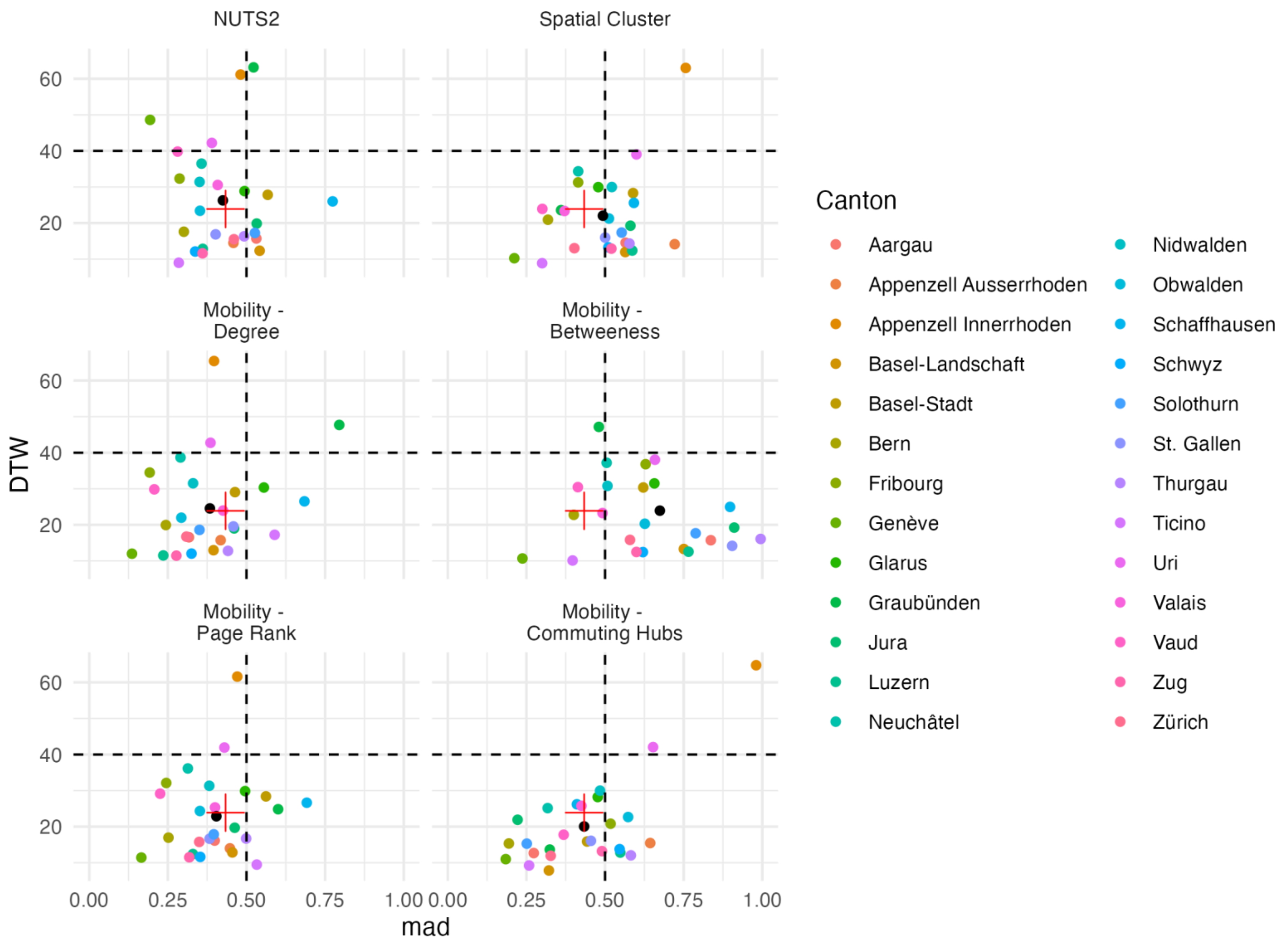


- Dynamic Time Warping (**DTW**) distance between standardised median and hospitalisation incidence per 100k
- Median absoulute dispersion of the prediction (**mad**)

Correlation with clinical data – Hospital admissions

Consistent with case – cluster based selection perfoms marginally better than the others when evaluating correlation with hospitalisation time-series.

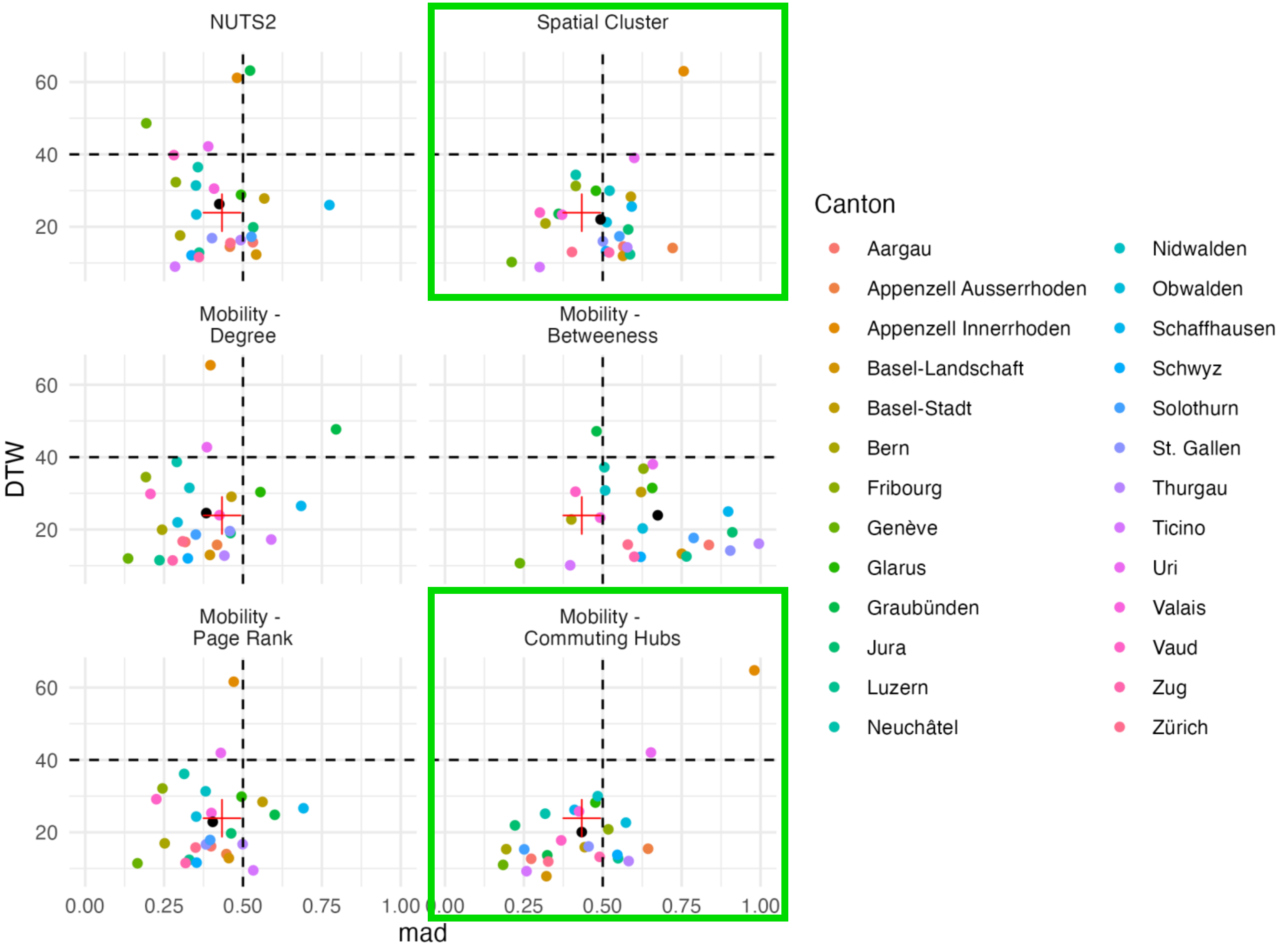
The distribution of performance across all cantons is again imporved as well



Correlation with clinical data – Hospital admissions

Consistent with case – cluster based selection perfoms marginally better than the others when evaluating correlation with hospitalisation time-series.

The distribution of performance across all cantons is again imporved as well



Summary

1. We defined six different WWTP site selection criteria
2. We evaluated these selection regimes by using a spatio temporal model to project expected viral loads in unmeasured catchments and also by comparing projections at Canton level with clinical data
3. We found that metrics highly correlated with population size performed better when making predictions at catchment level
4. However when exploring correlations with cases and hospitalisations at canton level, the regimes which clustered the catchments using spatial and mobility data performed better.
5. Overall all models performed comparably and were also relatively close in performace to the original model

Thanks!

Julien Riou,

Moritz Wagner,

Tim Julien,

Christoph Ort,

Tanja Stadler

Members of the WISE project. Especially

Charlie Gan and team

All the WWTP staff who do the sampling for us!

Swiss National Science Foundation

BAG OFSP UFSP

Computational Biology Group

cEvo

eawag aquatic research

unisanté

Centre universitaire de médecine générale et santé publique · Lausanne

james.munday@bsse.ethz.ch

@jdmunday

jdmunday.github.io

https://wise.ethz.ch/