Adapting Spatiotemporal Gaussian Process Regression for Multinomial Data

Hayley Tymeson

A thesis
submitted in partial fulfillment of the
requirements for the degree of

Master of Public Health

University of Washington
2019

Committee:
Robert Reiner
Christopher Adolph

Program authorized to offer degree:
Department of Global Health

University of Washington

Abstract

Adapting Spatiotemporal Gaussian Process Regression for Multinomial Data

Hayley Tymeson

Chair of the Supervisory Committee:
Robert Reiner
Department of Global Health

Estimating risk factor exposures over time is crucial to measuring and evaluating progress on behavioral, environmental, and occupational risks. Risk factor levels and trends are also a critical input in calculating the burden of diseases of health outcomes caused by exposure to health risks. Many health risk factors are multinomial in nature or in practice, and require more nuanced statistical treatment than traditional statistical methods provide. Spatiotemporal Gaussian Process Regression (ST-GPR) is a time-series model used primarily to estimate risk factor exposure within the Global Burden of Disease Project. We expanded the existing model to more accurately account for the unique requirements of ordinal and nominal multinomial time-series data. The expanded multinomial model was evaluated on two risk factor datasets of 1) occupational categories and 2) vaccination coverage, and compared based on out-of-sample fit and data coverage. We found that multinomial adaptations to the existing model led to slightly worse out-of-sample fit, yet major improvements in the validity of uncertainty and significant reductions in computational time.

1 Introduction

It is estimated that over 60% of all deaths worldwide can be attributed to health risk factors [2]. The World Health Organization defines risk factors as 'any attribute, characteristic or exposure of an individual that increases the likelihood of developing a disease or injury' [7]. Many common risk factors, such as low physical activity, smoking behavior, environmental or occupational hazards are entirely avoidable, yet persist at high levels throughout the population [2]. With so much disease and disability attributable to risky health behaviors and environmental or occupational exposures, it is important to understand and quantify the level and trends of risk factor exposure.

But risk factor exposures can be extremely difficult to estimate. First, unlike many questions in health and medicine, risk factors are often not biological but behavioral or environmental in nature. They often depend on cultural factors that vary enormously across and within countries, age-groups, and sexes. Second, risk factor exposures frequently change over time, as human behavior changes, policies striving to reduce harmful risk factors succeed or backfire, or population distributions fluctuate. Exposures to risks often change differentially across locations and demographics. For example, smoking prevalence varies strongly worldwide, by sex, and by age. In Greenland, over 40% of males and females smoke daily. That is a strong contrast to Burundi, where it is estimated that less than 1% of women and roughly 10% of males smoke daily, or Indonesia, where about 4% of women and 38% of men are thought to be daily smokers. The levels of smoking already vary widely by location and sex, but the time trends also diverge across locations and sexes. Many high-income countries showed large age-standardized reductions in smoking prevalence in males between 1990-2016, but smaller reductions in female smoking. Some countries, like Kuwait, were moving in the opposite direction over the same time period, recording large *increases* in smoking prevalence from 1990-2016 for both men and women. This does not even cover differences in age patterns or relationships between smoking and other demographic variables such as poverty [3].

Many risk factors also do not fit the requisite assumptions for common statistical modeling tools. Linear regression, for example, depends upon the independence of observations as a starting premise. But many risk factors are multinomial in nature or in practice, and violate this assumption. Data are multinomially distributed when composed of independent trials where the outcome variable can be one, and only one, of many unique categories. For example, rolling a six-sided dice many times produces a multinomial distribution. You can count the number of times a 1, 2, 3, 4, 5, or 6 is rolled. Each roll is an independent trial. If the outcome of a given roll is a 6, it is certain that the roll did not produce a 1, 2, 3, 4, or 5.

When a risk factor is grouped into categories, an increase in one category naturally requires a decrease in one or many other categories. For example, models of sanitation quality are often grouped into 'unimproved', 'improved', or 'optimal' quality. Imagine the proportion of households with optimal sanitation in a given country rises from 50% in 2000 to 90% in 2010, while the propor-

tion with improved sanitation stays at 5% through the same period. Then we know the proportion with unimproved sanitation fell from 45% to 5%, without even needing to observe this category. So, datapoints across categories for the same location are in fact highly *dependent*, not independent as linear regression requires. Multinomial risk factors then require alternative methods of analysis, while retaining all the difficulties associated with estimating risk factors in general.

Spatiotemporal Gaussian Process Regression (ST-GPR) is a fast and flexible time-series modeling strategy for behavioral, environmental, and metabolic risk factors. The model takes in location-,year-, age-, and sex-specific datapoints and discerns spatial and temporal trends amidst the noise, based on available data and model-specific parameters. ST-GPR is frequently used to model risk factors within the Global Burden of Disease Project due to its speed, flexibility, and ability to borrow information across space, time, and age to best approximate areas with missing or limited data [2]. However, ST-GPR currently relies upon a linear model to set the trends for the entire modeling process. Currently, multinomial models with N categories are split into either N or N-1 separate ST-GPR models, and estimates are re-scaled and combined at the end of the process to fit their multinomial constraints. Modeling these categories as separate entities risks overfitting, by assuming each trend has no external dependencies, and violates the assumptions of linear regression.

This paper expands ST-GPR to more accurately accommodate the unique requirements of multinomial models, by creating options for ordinal or nominal multinomial datasets. In order to evaluate the multinomial methods against existing, or default, methods, two example multinomial datasets were selected. We will describe how these models are estimated in the existing framework, and how those methods were adapted for multinomial datasets. Finally, we will compare in-sample and out-of-sample fit statistics and data coverage to evaluate each treatments of multinomial data within ST-GPR.

2 Data

Two multinomial datasets, one ordinal and one nominal, were selected to test the modeling changes implemented for each type of multinomial method. Ordinal multinomial models have a theoretical ordering method; for example, sanitation quality is sometimes categorized into 'unimproved', 'improved', and 'optimal' groupings to condense the information into actionable and measurable categories. The categories clearly have some intuitive ordering that may be relevant to modeling. Nominal multinomial data, on the other hand, have no relevant ordering whatsoever. Modeling disease strains, such as the strain of tuberculosis measured in a population, is a common example of a nominal multinomial dataset in the medical field, as strains are distinct and unordered categories.

To test the ordinal version of ST-GPR, we used data on Diphtheria-Pertussis-Tetanus (DPT) vaccination coverage. The vaccine requires multiple doses for optimal effectiveness, yet some children receive no doses or fewer than the op-

timal number of doses of the vaccine. Determining the proportion of the population that 1) abstains from vaccination, 2) receives the first or second dose of the DPT vaccine, or 3) receive three or more doses of the vaccine provides information to health officials on if, how, and where children are getting lost in the vaccination schedule. The categories are also clearly ordered, making this a clear and policy-relevant example of an ordinal model.

The second test case was a multinomial nominal dataset estimating the proportion of the population employed within ten occupational categories. Unlike the vaccines dataset, occupational categories have no theoretical ordering. The categories were defined according to the major groups in the International Labour Organization's International Standard Classification of Occupations (ISCO-08) [6]. Overall, the dataset included occupational data from roughly 1400 sources spanning 379 GBD locations.

3 Methods

3.1 Existing model methods

ST-GPR currently takes in a dataset and a configuration file of model parameters. The dataset is prepared and consolidated by the modeler, and should include datapoints with corresponding data variance, indexed by GBD location, year, GBD five-year age groups, and sex. More information on the GBD location hierarchy can be found at `http://ghdx.healthdata.org/record/ihme-data/gbd-2017-cause-rei-and-location-hierarchies`. The configuration file contains model parameters that dictate how much smoothing should be applied over geography, ages, and sexes. The data pass through three separate modeling stages. Before modeling begins, the data is offset and transformed within ST-GPR. Under the existing methods, occupational and vaccine coverage data are logit-transformed for modeling, as both datasets are prevalences that must be within [0,1] bounds. Datapoints that are equal to 0 or 1 are offset by a small amount, δ, before applying the logit transformation.

The first stage is a linear fixed-effects or mixed-effects regression, which creates a first-pass estimate to fill in the time-series and incorporate covariate information. The linear regression is run using the *lme4* package in R 3.6.0 [1].

The second stage runs a local smoother to capture any residual spatial, temporal, and age-pattern signals that were missed in the first-stage model. This information is then 'added' back into the model estimates from the fully-specified first stage.

An intermediary stage estimates two uncertainty components that feed into the final Gaussian Process stage. First, we estimate a sex-specific, but otherwise global, model parameter *amplitude* (α_s), as a function of the absolute differences between the spatiotemporally-smoothed estimates and the data. Amplitude acts as a scalar on the variance-covariance matrix of the Gaussian Process. Second, we attempt to capture non-sampling variance (NSV) by location and sex ($\delta_{l,s}$), or variance between the datapoints that cannot be explained by sampling

variance alone. We use non-sampling variance to increase datapoint uncertainty where data are extremely heterogeneous.

The third and final modeling stage runs a separate Gaussian Process with a Mátern kernel over time for each location-age-sex group, combining the space-time-age smoothed second-stage estimate and the data into a single jointly multivariate final estimate. The Gaussian Process is run using the *pymc* package in Python [8]. Lastly, countries with nested sub-national locations are either re-scaled or aggregated to make nested estimates internally consistent in a post-modeling 'raking' stage. A more detailed description of the methods behind ST-GPR can be found in the Appendix of the GBD2017 Risk Factors capstone [2].

3.2 Multinomial expansion

The multinomial expansion of ST-GPR involved changes to three main aspects of the model: the first stage modeling form, the dimensions of output predictions, and the estimation of non-sampling variance and amplitude.

For multinomial nominal cases, the model first fits a multinomial regression using the *multinom* function in the *nnet* package [9]. Spatiotemporal smoothing and the Gaussian Process stages were run separately on each category for the rest of the model using the logit-transformed data and first stage as starting inputs. At the end of modeling, estimates were transformed out of logit space and re-scaled such that for $[m_{l,y,a,s,1}, m_{l,y,a,s,2}, ..., m_{l,y,a,s,N}]$, where m is the proportion for a given location l, year y, age-group a, and sex s falling in category $c = [1, 2, ..., N]$,

$$\sum_{c=1}^{N} m_{l,y,a,s,c} = 1 \tag{1}$$

Ordinal models required entirely different adaptations. Instead of changing the first stage regression and running the model essentially as-is, the datapoints themselves were adjusted to reflect their ordinal nature via the continuation ratio method [5]. The entire process was then run on these ordinally-transformed data.

In brief, for an ordinal model with N categories, we model $N-1$ categories and treat the category with the least data as a residual category, θ. The residual category θ is then solved for upon completion of modeling. So for estimates $[m_{l,y,a,s,1}, m_{l,y,a,s,2}, ..., m_{l,y,a,s,N}]$, where m is the proportion for a given location l, year y, age-group a, and sex s falling in category $c = [1, 2, ..., N]$, such that $\sum_{c=1}^{N} m_{l,y,a,s,c} = 1$, we convert data via the following formula:

$$\tilde{m}_{l,y,a,s,c} = \left\{ \begin{array}{ll} c = 1, & \theta \\ 1 < c < N, & \frac{m_{l,y,a,s,c}}{1 - \sum_{j>c}^{N} m_{l,y,a,s,j}} \\ c = N, & m_{l,y,a,s,N} \end{array} \right\}$$

We then take these adjusted proportions within each category for each location, year, age, and sex, and model them separately through ST-GPR in logit

space. Unlike nominal multinomial datasets, the first stage remained a linear mixed-effects or fixed effects regression on logit-transformed data for each non-residual category $\tilde{m}_{l,y,a,s,c}$ in the model. At the end, we reverse the ordinal transformations and solve for the residual category θ.

$$\hat{m}_{l,y,a,s,c} = \left\{ \begin{array}{ll} c = N, & \tilde{m}_{l,y,a,s,c} \\ N > c > 1, & \tilde{m}_{l,y,a,s,c} \times (1 - \sum_{j>c}^{N} \tilde{m}_{l,y,a,s,c}) \\ c = 1, & 1 - \sum_{j>c}^{N} \tilde{m}_{l,y,a,s,j} \end{array} \right\}$$

To test how best to estimate uncertainty, we ran out-of-sample 5-fold cross-validation on 10 sets of resampled data for all test and comparison models. Data were held out by matching patterns of missingness in random locations in the model until 20% of the data was held out, following methods used in a similar cross-validation model [4].

The 'default' version of each model replicated the current implementation of ST-GPR for categorical models. For vaccination coverage models, the default method is to run *two* models with data transformed before input via the continuation ratio method, then use the ST-GPR final outputs to solve for the true proportions for all three categories of interest. The first model contains data on the unadjusted proportion of children receiving three doses of the DPT vaccine. The second model contains the adjusted ratio of all those receiving 1 or 2 doses of the DPT vaccine. The residual category is the proportion receiving no doses of the DPT vaccine. Each model was run with out-of-sample cross-validation separately, and outputs were combined post-modeling to make up the comparison ('default') set.

For the nominal occupational categories dataset, the default method is to launch 10 separate models for each of the 10 potential occupational categories, and re-scale the outputs to fit their categorical constraints upon completion. Each of the 10 models was run with 10 holdouts and likewise combined and re-scaled upon completion to make up the comparison 'default' estimates.

Identical model settings and covariates were used for both the 'default' version and the multinomial version of all models being tested.

For both ordinal and multinomial datasets, we explored three different ways to estimate non-sampling variance and amplitude. ST-GPR estimates a single value of amplitude for each sex in the model, while non-sampling variance values are estimated for a given location and sex. Since the default method of running multinomial data involves running separate models for each category, this essentially means amplitude is estimated for each sex and category ($\alpha_{s,c}$), while NSV is estimated for each location, sex, and category ($\delta_{l,s,c}$). There was accumulating evidence that uncertainty is being underestimated in the model based on in-sample coverage. One potential explanation is that estimating categories of multinomial datasets via separate models leads to overfitting, as each category has its own amplitude and NSV calculated without including the residuals from other categories in the multinomial dataset. The three methods involved 1) estimating amplitude and non-sampling variance using residuals from all categories as α_s, 2) estimating amplitude and non-sampling variance separately by cate-

gory (similarly to the default models), and 3) estimating non-sampling variance by category but using all categories' residuals for amplitude estimation. Each of these methods were evaluated out-of-sample.

4 Results

The multinomial test cases involved running one multinomial model for both vaccination coverage and occupational categories each, and comparing to the combined 'default' models for each risk factor. Models were compared via in-sample and out-of-sample root-mean-squared-error (RMSE) and mean-absolute-error (MAE). The adequacy of uncertainty was measured via data coverage, or the proportion of in-sample or out-of-sample datapoints in the model that fell within the estimated uncertainty interval. Uncertainty intervals were estimated in 5% increments, from 5% confidence intervals to 95% confidence intervals.

4.1 Nominal case study: Occupational categories

The multinomial nominal models were compared for fit statistics at the first and final stage. Final outputs were further compared across all three methods of uncertainty estimation, as detailed in Table 1. Though models were evaluated on both root mean squared error and mean absolute error, only RMSE plots will be shown, as the trends and conclusions from MAE were identical to RMSE.

Using a multinomial regression as the first stage of ST-GPR actually performed worse than running separate linear regressions and re-scaling estimates, according to out-of-sample RMSE and MAE. Though the overall RMSEs were extremely close and out-of-sample RMSE was actually lower in 7/10 categories, the multinomial regression performed far worse when estimating the agricultural workers and armed forces occupational categories across holdouts when looking at mean RMSE by location.

	Method	Amplitude	NSV
1	(Default) Amplitude and NSV by category	$\alpha_{s,c}$	$\delta_{l,s,c}$
2	Amplitude and NSV across categories	α_s	$\delta_{l,s}$
3	Amplitude across categories, NSV by category	α_s	$\delta_{l,s,c}$

Table 1: The three explored methods for estimating uncertainty parameters

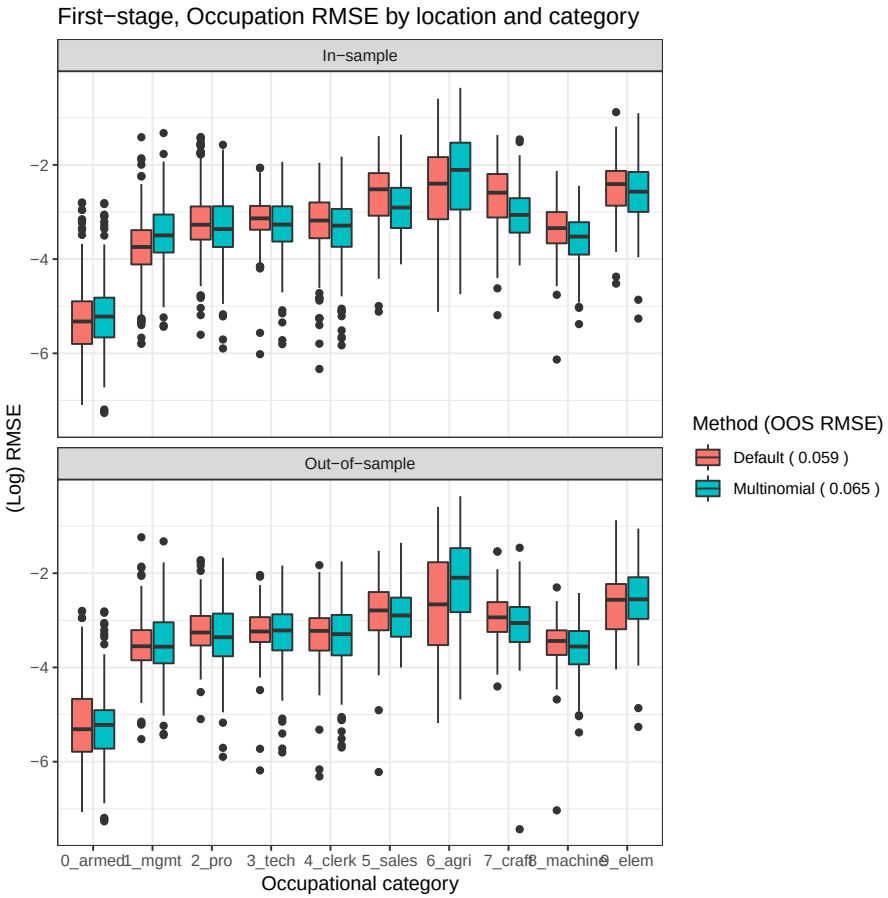

Figure 1: In-sample and Out-of-sample RMSE on first-stage estimates calculated for each location in the model by category. Legend shows mean out-of-sample RMSE across all first stage values for both the default and multinomial models.

Final estimates also had higher RMSE and MAE for the full model, and were more uniformly deficient across categories.

The uncertainty method that led to the lowest out-of-sample error for final estimates was method (3), as described in Table 1. This method also led to the best out-of-sample uncertainty based on coverage statistics. This method was expected a priori to perform best, as non-sampling variance depends on heterogeneity in datapoints on the same scale, while amplitude more fully depends on the size of the residuals from all categories, as a change in any category affects other categories as well. Despite lingering systematic bias in in-sample coverage, out-of-sample coverage is roughly in line with predicted coverage at all measured confidence intervals.

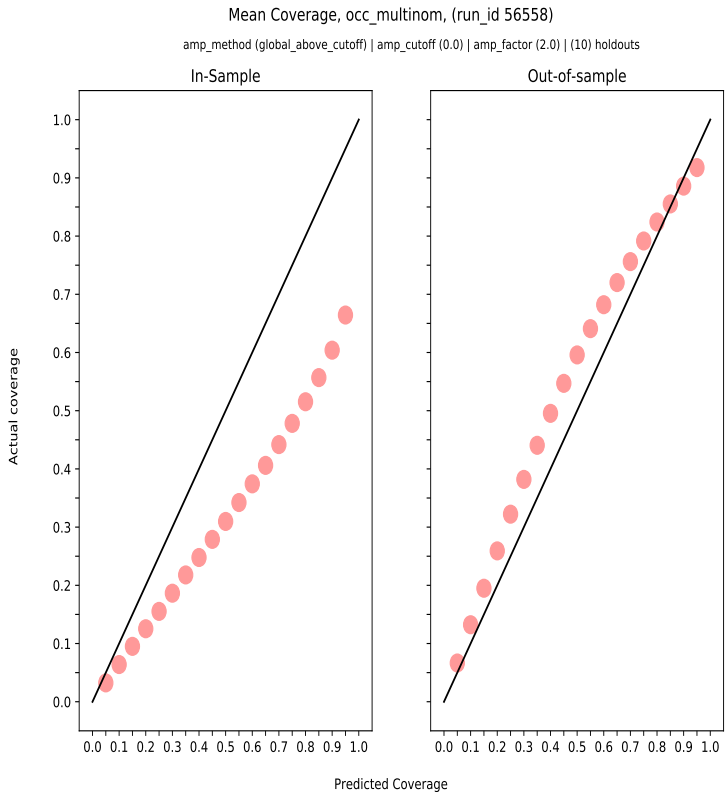

Figure 2: Data coverage for the best multinomial occupational model, estimating amplitude across categories and non-sampling variance by category.

In comparison, the data coverage for the default method, where amplitude and non-sampling variance are both estimated individually for each category, is vastly underestimated using the same model parameters. To put the results in context, the 95% confidence interval for the default model only contains 40% of the data for both in-sample and out-of-sample models.

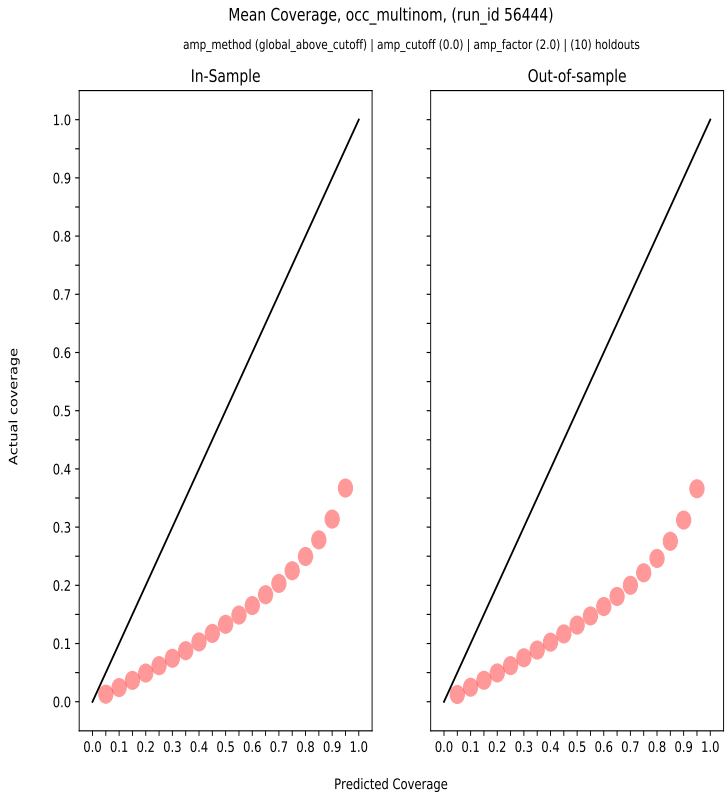

Figure 3: Data coverage for the default occupational model, estimating amplitude and non-sampling variance by category.

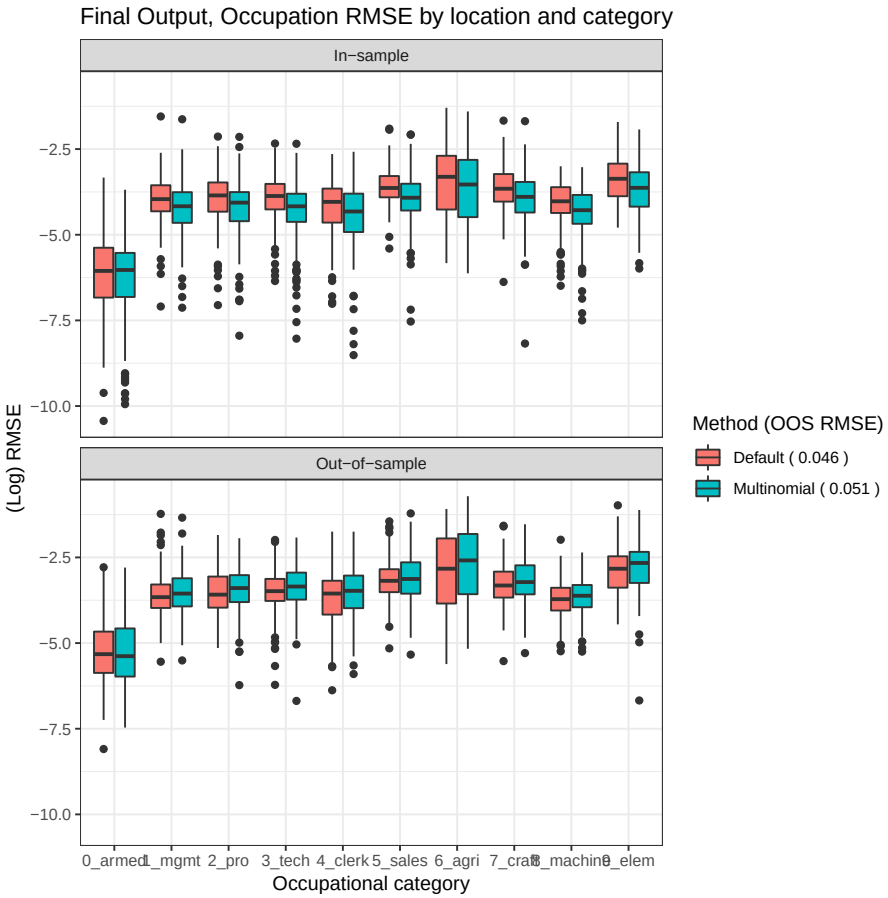

Figure 4: In-sample and Out-of-sample RMSE on first-stage estimates calculated for each location in the model by category.

The final estimates from the Gaussian Process stage show an interesting pattern. The in-sample fit is uniformly better for the multinomial model across categories, while the out-of-sample fit is almost uniformly worse compared to the default method.

4.2 Ordinal case study: Vaccination coverage

Finding a comparison case for the ordinal implementation of the data proved difficult, as the continuation ratio method was already being implemented for many multinomial ordinal outcomes. Modelers would transform the data via the continuation ratio method and then run $N-1$ separate models for an N-category ordinal response variable. Implementation of this method was, until the uncertainty estimation, more of a technical and computational improvement

than a methodological improvement. However, like the occupational models shown above, data coverage showed severe and systematic underestimation of uncertainty in the default ordinal method. The main remaining decision was to determine how to calculate non-sampling variance and amplitude within the multinomial version to give the best model fit and uncertainty.

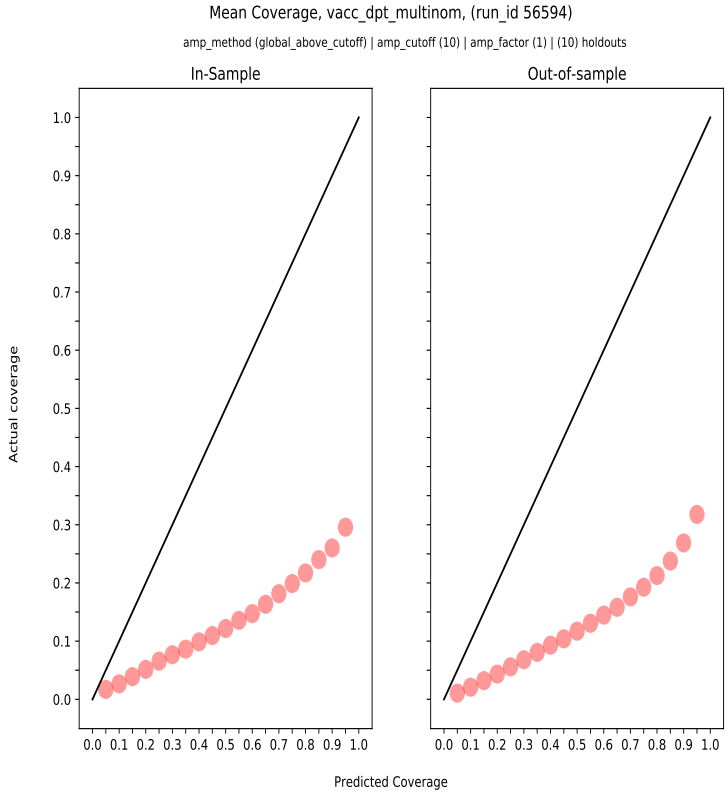

Figure 5: Data coverage for the default vaccination model, estimating amplitude and non-sampling variance by category.

The ordinal model followed many of the same patterns as the multinomial model. Once again, the best of the three uncertainty estimation methods was method (3), estimating amplitude across categories and NSV by category. Estimating amplitude across categories led to moderate improvements in data coverage metrics, but slight declines in out-of-sample model fit compared to the default method. Uncertainty is still significantly underestimated, but the coverage at the 95% confidence interval rose from roughly 32% in the default model to 45% in the multinomial model, and shows some out-of-sample improvement across all confidence intervals measured.

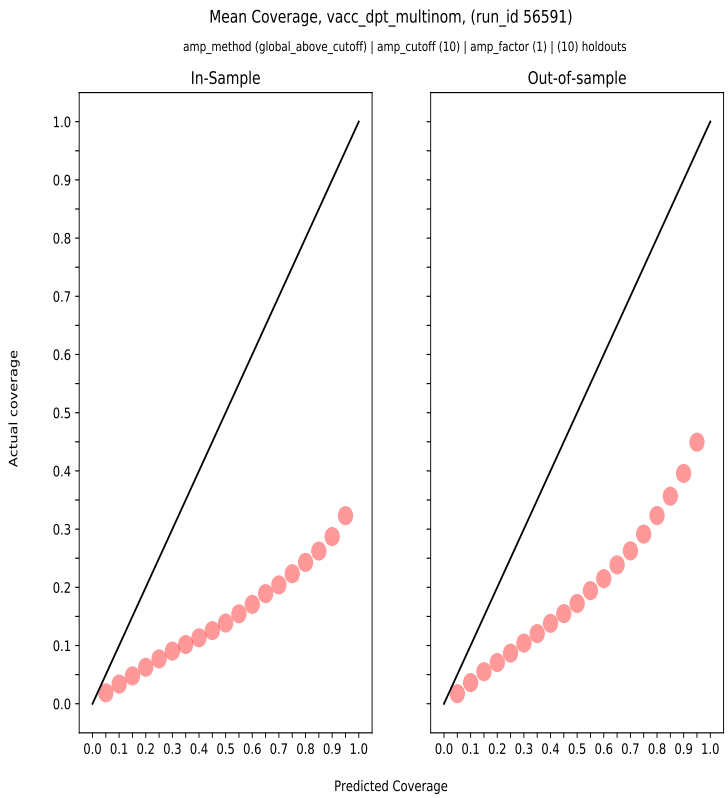

Figure 6: Data coverage for the best ordinal vaccination model, estimating amplitude across categories and non-sampling variance by category.

This improvement in uncertainty estimation seems to have come at a cost of model fit for the final estimates. Since nothing but the uncertainty estimation changed in this comparison, we can be more causally certain that the changes in estimating amplitude led to the worse performance in model fit. Root-mean-squared-error, calculated for each location in the model, was consistently slightly worse across holdouts, which reduces the likelihood of mere coincidence driving the reductions in model performance due to increased uncertainty. Interestingly, though RMSE by location was slightly higher in general in the multinomial version, there was consistently less variation in RMSE in the multinomial version.

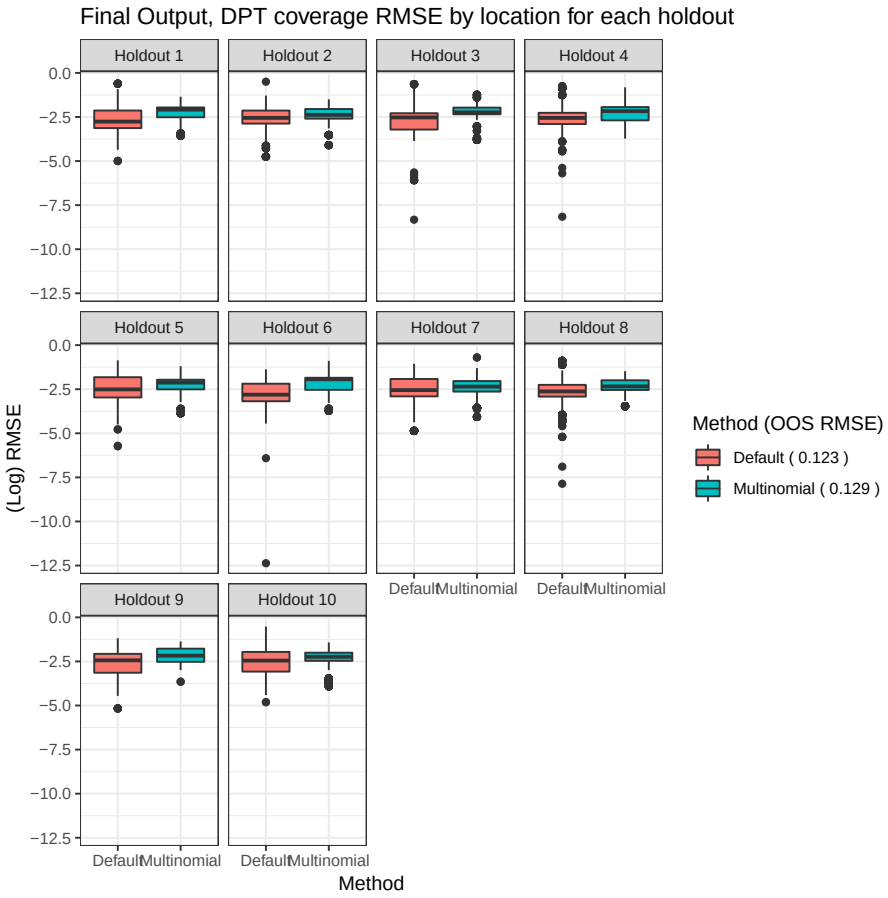

Figure 7: Vaccination models, Mean RMSE by location for each holdout.

When looking at the out-of-sample RMSE by category, the multinomial models show lower in-sample RMSE but higher out-of-sample RMSE relative to the default models, similar to the final outputs in the occupational models.

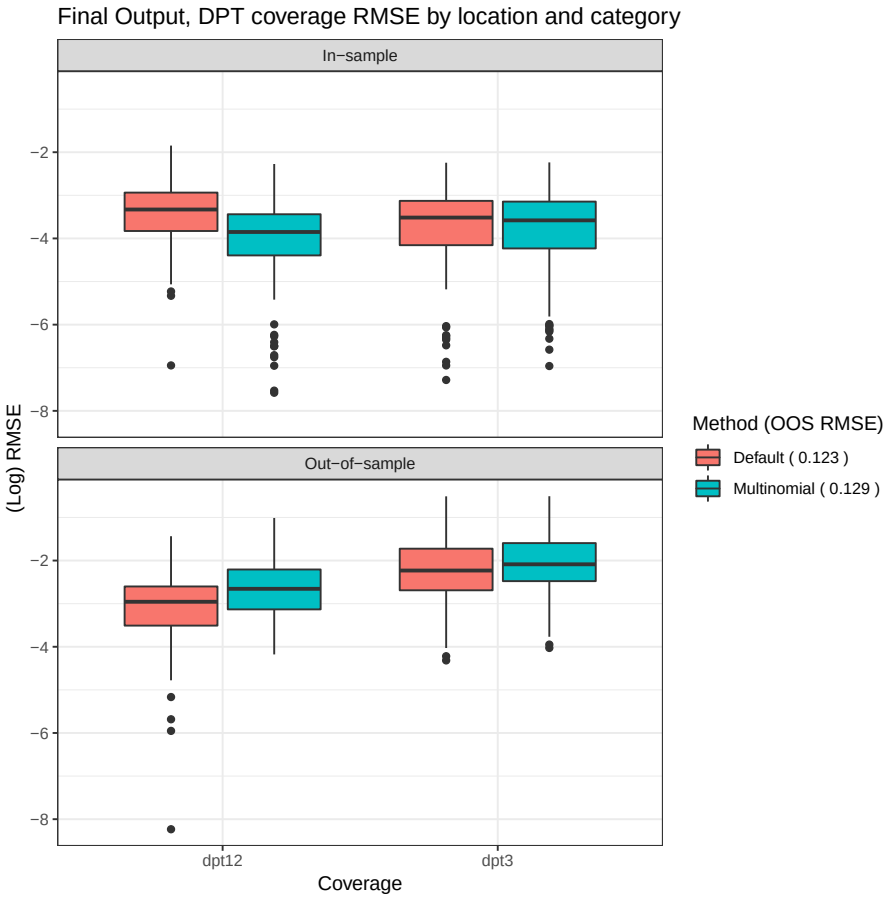

Figure 8: Vaccination models, Mean RMSE by location for each category.

5 Discussion

Overall, multinomial implementations performed slightly worse than their comparison models in out-of-sample fit statistics. However, multinomial models showed moderate to immense improvements in data coverage and uncertainty intervals for ordinal and multinomial models compared to treating categories as separate models. The analysis further revealed that the existing model is systematically underestimating uncertainty. This could be due in part to overfitting when estimating uncertainty separately for each category. Categories in a multinomial model do not fluctuate in isolation, but also in response to any changes in any alternate category. But the uncertainty on ordinal models remaining significantly underestimated despite inclusion of all but one category, suggesting there are structural biases in uncertainty estimation in the model.

The fact that out-of-sample coverage was actually *better* than in-sample uncertainty in the nominal test case also suggests some unknown biases are at play.

The occupational model showed higher out-of-sample root mean squared errors for both the final estimates and the first-stage outputs. Though the first stage model fit was *better* in-sample and only slightly worse out-of-sample, the statistical treatment was more accurate given the constraints of the data and the assumptions of each model. Since the first stage regression is also primarily useful for parsing out correlations and trends between covariates and data, a worse fit does not necessarily mean a worse model. The difficulty in comparing multinomial coefficients and linear regression coefficients makes it hard to make absolute statements about which treatment is best.

However, the categories that performed worst in the first stage regression in the multinomial version highlight a significant limitation in running multinomially compared to running separate models by category. There is currently no way to implement 'choice-specific' covariates within the multinomial model. For example, nine of the original 10 occupational models used the same covariates and linear regression formula for the first stage, but the category with the highest in-sample and out-of-sample error used a unique set of covariates. The model on the proportion of the population working in agriculture used covariates on absolute latitude and urbanicity, while the other category models used socio-demographic index and education level as covariates. Of course, for testing purposes, the same covariates (sdi and education) were used for all models, including the default agricultural model. Yet it is easy to imagine that occupational divisions across countries may have different drivers, though the multinomial model is only capable of running a single model for all categories. Future work should consider implementing a conditional logit model to allow for different covariate drivers for distinct categories falling within the same modeling group.

Lastly, there is the interesting finding that both the multinomial nominal and ordinal models performed consistently better in-sample, but consistently worse out-of-sample. This pattern is likely caused by the more accurate uncertainty intervals on the multinomial model, and believe the pattern would hold true for other modeling cases other than occupational and vaccination models. For context, the spatiotemporally-smoothed estimates act as the mean function in a Gaussian Process. This mean function is meshed with existing data, which have their own uncertainty as well. A higher amplitude allows the Gaussian Process to move away from the mean function more readily, if pulled away by data with more certainty. The pattern of lower in-sample RMSE but higher out-of-sample RMSE matches this scenario; in a model with more variance on the mean function, individual data points are too easily able to sway the model away from the main trend and towards confident datapoints, reducing the ability of the model to accurately predict datapoints held out from the model. This suggests a trade-off between valid uncertainty and model fit. Future work on this model should prioritize improvements in uncertainty propagation and evaluating the validity of datapoint uncertainty, so there is no need to sacrifice valid confidence

intervals for improved model fit.

In summary, the multinomial case rested on firmer statistical grounds and improved model uncertainty in both the occupational and vaccination test cases. On the technical side, the multinomial is vastly more computationally efficient and reduces the likelihood of human error, as a single model is easier to manage than ten. On the other hand, the multinomial model performed slightly worse on out-of-sample predictive validity. Looking at the pros and cons holistically, the incorporation of multinomial methods to ST-GPR likely merits further testing and investigation.

References

[1] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[2] GBD 2017 Risk Factor Collaborators. Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990-2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1923–1994, Nov 2018.

[3] Reitsma et al. Smoking prevalence and attributable disease burden in 195 countries and territories, 1990–2015: a systematic analysis from the global burden of disease study 2015. *The Lancet*, 389(10082):1885–1906, May 2017.

[4] Kyle J. Foreman, Rafael Lozano, Alan D. Lopez, and Christopher JL Murray. Modeling causes of death: an integrated approach using codem. *Population Health Metrics*, 10(1):1, Jan 2012.

[5] A. A. O'Connell. *Quantitative Applications in the Social Sciences: Logistic regression models for ordinal response variables.* SAGE Publications, Thousand Oaks, CA, fourth edition, 2006.

[6] International Labour Office. International Standard Classification of Occupations:ISCO-08. 1, 2012. `https://www.ilo.org/public/english/bureau/stat/isco/docs/publication08.pdf`.

[7] World Health Organization. Risk factors. `https://www.who.int/topics/risk_factors/en/`, 7 1993.

[8] Anand Patil, David Huard, and Christopher Fonnesbeck. Pymc: Bayesian stochastic modelling in python. *Journal of Statistical Software, Articles*, 35(4):1–81, 2010.

[9] W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S.* Springer, New York, fourth edition, 2002. ISBN 0-387-95457-0.