SAND2021-XXXXR

LDRD PROJECT NUMBER: 223167

LDRD PROJECT TITLE: Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections

PROJECT TEAM MEMBERS: William Hart (PI), Michael Bynum, Carl Laird, John Siirola

ABSTRACT

This project explored large-scale nonlinear optimization approaches for simultaneous estimation of epidemiological model parameters, unobserved SARS-CoV-2 infections, and impact of nonpharmaceutical interventions (NPIs) from case data. Effective control of emerging infectious diseases requires decision-making tools that can estimate epidemiological parameters and quantify the impacts of intervention strategies. NPIs are a key tool for controlling the ongoing SARS-CoV-2 pandemic, even with the development of vaccines for this disease. Well-informed policies for NPIs are critical to mitigate COVID outbreaks and plan effective responses for COVID variants with new infectious characteristics. Previous research supported an analysis of the efficacy of NPIs in the U.S. prior to vaccine deployments, where unobserved infections were reconstructed from historical data. Here, we consider large-scale optimization models that simultaneously estimate transmission parameters and unobserved infections. Three models for the transmission parameters are considered: (1) models with temporally varying transmission parameters for all counties, (2) linear models capturing dependence of interventions on transmission parameters, and models with transmission parameters predicted by an ML model informed by county-specific NPI policies (as inputs). This research employed the Pyomo software to model these parameter estimation problems, and the scale of these problems motivated enhancements to Pyomo for deep learning models.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

INTRODUCTION

This work focuses on estimation of unknown states and parameters in a discrete-time, stochastic, SEIR model using reported case counts and mortality data. An SEIR model is based on classifying individuals with respect to their status in regards to the progression of the disease, where S is the number individuals who remain susceptible to the disease, E is the number of individuals who have been exposed to the disease but not yet infectious, I is the number of individuals who are currently infectious, and R is the number of recovered individuals [5, 1]. For convenience, we include in our notation the number of infections or transmissions, T, that represents the number of individuals transitioning from compartment S to compartment E over a particular interval. Similarly, we use C to represent the number of reported cases.

In compartment models, the rate of new transmissions is often characterized with relationships like the following:

$$T = \beta I \frac{S}{H}$$

where H is the population size. The transmission parameter β represents the average number of adequate contacts of an infected person during the infectious period, where an adequate contact is one that is sufficient to spread infection [5]. The transmission parameter β depends on the characteristics of the disease as well as the environment in which it spreads. Thus, the transmission parameters may depend on several factors, including population behavior, intervention and mitigation strategies, and population age structure. This parameter is related to the reproductive number R, which is often approximated as $R = \beta/\gamma$ where γ is the mean infectious period.

This report describes the research performed in the Sandia LDRD Project "Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections". The main goal of this project is to estimate changes in β over time and space as a decision-making aid for policy makers. We are motivated by SEIR patch models like the COVIDScenarioPipeline, which model all counties in the United States and account for interactions between them [3, 6]. Analysis of these national-scale SEIR models require the solution of large-scale parameter estimation problems, which is the central goal of this project.

Our preliminary research was performed as part of Sandia's COVID-19 Inference Project, and two publications provide detailed discussions of these early results [4, 8]. Additionally, Addendum A includes slides presented at a Sandia review of these results.

OVERVIEW

This project accomplished a variety of technical objectives, which are summarized here. Our preliminary research was performed as part of Sandia's COVID-19 Inference Project, which formed the basis for the extensions developed in this project. In this project, we revised and updated the documentation of these methods in a SAND report [4]:

1. Reconstruction of Unobserved Transmissions: We estimate the time-profile counts of each of the compartments (S-E-I^1-I^2-I^3-R) from reported cases or mortality data. We consider both deterministic and stochastic procedures for estimating the population counts within these compartments. These reconstructions provide initial conditions that can be used for models like those in the COVIDScenarioPipeline.

2. Transmission Parameter Inference: We estimate changes in transmission parameters per county over time throughout the United States, given inter-county mobility information and estimates of time profiles for SEIR compartments over time. The transmission parameters in these models represent the average number of contacts that are sufficient for transmission. They depend on the disease and on the setting in which it spreads, and changes in transmission parameters typically reflect, for example, changes in population behaviors, such as those related to intervention strategies like social distancing policies. A key differentiator of this approach is that the transmission parameters are estimated across all counties simultaneously using a fully-coupled model that includes the impact of mobility between counties.

In this project, the reconstruction methods were extended and used to generate estimates for the progression of COVID-19 in the United States in 2020. This data was used by academic collaborators to assess the efficacy of nonpharmaceutical interventions, which was published in Nature Communications [8].

3. Assessing the Effects of Nonpharmacological Intervention (NPI) Policies: In collaboration with others, we estimated the transmissibility of SARS-CoV-2 in 3,036 (out of 3,142, 97%) US counties using a mechanistic meta-population model that incorporates spatial coupling of transmission between counties to estimate weekly effective basic reproductive numbers (i.e., R_{eff}, the reproductive number adjusted for changes due to factors other than population susceptibility, such as social distancing) from confirmed cases and deaths from January 21 to July 5, 2020. We associated these R_{eff} estimates with NPIs and county level demographics, while accounting for, temporal variation, autocorrelation, and uncertainty in our estimates. This analysis provides a concrete basis for recommending NPIs in future policy implementations.

Most of this project focused on extensions of this prior work, which are documented in this report. We explored the application of large-scale optimization methods to two additional large-scale formulations:

4. Simultaneous Estimation of Compartment Counts and Transmission Parameters: Simultaneous estimation of compartment and transmission parameters results in a nonlinear parameter estimation that is significantly more challenging than the transmission parameter estimation problem we considered previously [4]. However, the simultaneous estimation of parameters allows for more explicit management of error estimates across different parts of the model, and provides significantly more flexibility to explore richer models for reporting, vaccination, etc. We describe a nonlinear, fully-coupled parameter estimation model that includes estimation of temporal profiles for model noise terms (e.g., external transmissions, stochasticity) and transmission parameters.

5. Simultaneous Estimation of Parameters and NPI Efficacy: We extend the previous model and estimate the impact of NPIs on the transmission parameter profiles. In this formulation, we estimate a linear model describing the impact of interventions on the transmission parameters. Although there may be nonlinear dependence on the interventions, this model provides a base case for investigating the performance and effectiveness of more complex interactions.

These parameter estimation problems require the solution of large-scale nonlinear optimization problems. Our research focused on two aspects of these problems that reflect performance bottlenecks:

6. Data-Parallel Nonlinear Parameter Estimation: These inference formulations are large-scale models at a national (county-level) scale, however, they are loosely coupled between counties. Therefore, they are highly appropriate for decomposition-based optimization strategies that can exploit parallel computing on HPC. We demonstrate effective scale-up of parameter estimation using Parapint, which is a software package based on PyNumero and Pyomo that implements a parallel interior-point method. Structured optimization problems inherently induce structure in the linear systems solved at each iteration of the optimization problem, and Parapint exploits this structure to support parallel solution.

7. Accelerating Machine Learning Models in Pyomo: We considered an extension of our parameter estimation model including a general machine learning model to capture the efficacy of NPIs. However, this significantly increases the computational challenge of both representing and solving the associated parameter estimation problem. We tailored Pyomo's problem representation to more efficiently express the dense, regular

expressions that are commonly used in machine learning models. This tailored representation significantly reduced the time needed to generate parameter estimation models with machine learning equations.

METHODS, RESULTS, AND DISCUSSION

The following sections describe research accomplishments for the four technical objectives that were the focus of this project (objectives 4-7 summarized above).

Simultaneous Estimation of Compartment Counts and Transmission Parameters

We extended our previous research [4] to explore the application of optimization methods to large-scale formulations that simultaneously estimate transmission parameters, unobserved states (compartment populations), and the impacts of intervention strategies on transmission. In the following, we use the notation introduced in [4]. To simplify this research, we consider disease propagation models that only have one compartment *I*, and we consider models that do not include inter-county mobility. However, the methods described here can be generalized to problems with those features.

The following optimization formulation simultaneously estimates compartment population counts and transmission parameters over time, given data for the reported cases over time, $\hat{C}_t$:

$$
\begin{aligned}
\min \quad & \sum_{t \in \mathcal{D}} \left(C_t - \hat{C}_t \right)^2 + \alpha_\epsilon^T ||\epsilon^T||_1 + \alpha^\delta ||\delta||_1 \\
\text{s.t.} \quad & T_t = \frac{\beta_t I_t S_t}{H} + \epsilon_t^T \\
& \delta_t = \beta_{t+1} - \beta_t \\
& S_{t+1} = S_t - T_t \\
& E_{t+1} = E_t + T_t - \sigma E_t \\
& I_{t+1} = I_t + \sigma E_t - \gamma I_t \\
& R_{t+1} = R_t + \gamma I_t \\
& C_{t+1} = \rho \sigma E_t \\
& \beta_t \geq 0 \\
& S_0 = H \\
& E_0, I_0, R_0, C_0 = 0
\end{aligned}
\tag{1}
$$

Formulation (1) includes L_1 regularization terms in the objective for noise or errors in the transmission expression (e.g., due to externally seeded transmissions or stochasticity) and temporal changes in the transmission parameters β_t. The regularization terms allow the parameter values to adapt to rapidly changing dynamics. For example, we would expect discontinuous changes in transmission parameters at points in time where public policies change that impact disease transmission.

We performed simulation studies to evaluate this formulation with known disease dynamics. Our results demonstrate the ability of this formulation to robustly recover discrete changes in the temporal profiles for both these terms in the model. Here, we illustrate the results from one of the case studies. Simulated data was generated using disease parameters consistent with literature values for Sars-CoV-2 ($\sigma = \frac{1}{5.2}, \gamma = \frac{1}{4}$), a seed of 5 external transmissions at day 10 ($\epsilon_{10}^T = 5$) and a change of the transmission parameter from 0.75 to 0.55 for the days 70-90 in the simulation. Figure 1 below shows the estimation results for ϵ^T and β for this case study:

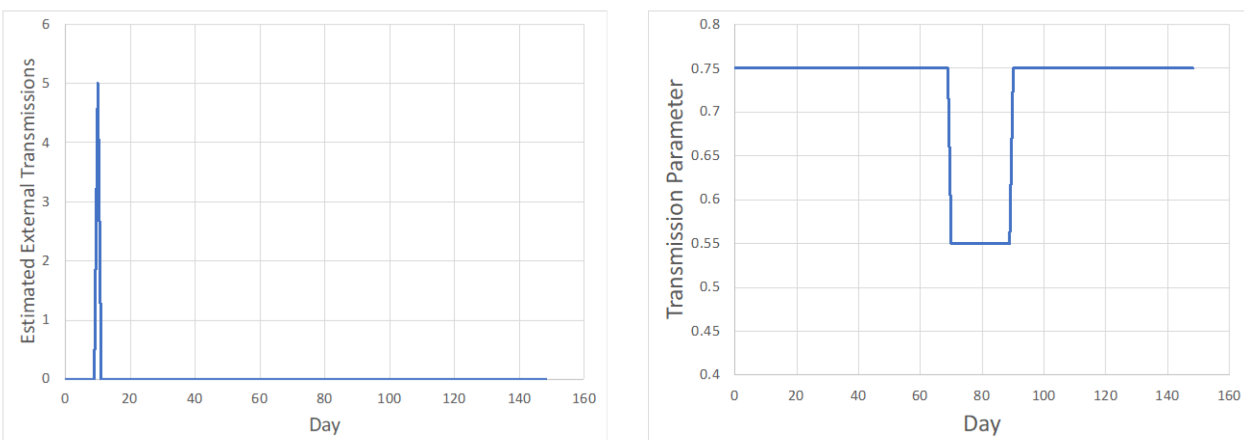


Figure 1: Demonstration of the L1-regularized formulation to recover discrete changes in profiles of estimated transmissions and transmission parameters. Simulation was performed using a seed of 5 external transmissions on day 10 and a change in the transmission parameter from 0.75 to 0.55 for days 70-90.

Simultaneous Estimation of Parameters and NPI Efficacy

In general, we wish to characterize the impact of non-pharmaceutical interventions (NPIs) on disease transmission. In Yang et al. [8], we provide a detailed retrospective assessment of the impact of NPIs to manage the COVID-19 pandemic. This research involved the estimation of transmission parameter profiles, and the correlation of those profiles with interventions. In this work, we explore simultaneous estimation of the impact of interventions along with model parameters. This provides increased flexibility in the inference tools that allow more complex models to be explored

We generalized Formulation (1) to integrate a linear model to predict the efficacy of NPIs:

$$
\begin{aligned}
\min \quad & \sum_{j \in J} \sum_{t \in \mathcal{D}} \left(C_{jt} - \hat{C}_{jt} \right)^2 + \alpha_\epsilon^T ||\epsilon^T||_1 \\
\text{s.t.} \quad & T_{jt} = \frac{\beta_{jt} I_{jt} S_{jt}}{H_j} + \epsilon_{jt}^T \\
& \beta_{jt} = \beta_j^0 + \sum_{i \in NPI} \beta_i^\Delta \lambda_{j,i,t} \\
& S_{j,t+1} = S_{jt} - T_{jt} \\
& E_{j,t+1} = E_{jt} + T_{jt} - \sigma E_{jt} \\
& I_{j,t+1} = I_{jt} + \sigma E_{jt} - \gamma I_{jt} \\
& R_{j,t+1} = R_{jt} + \gamma I_{jt} \\
& C_{j,t+1} = \rho \sigma E_{jt} \\
& \beta_{jt} \geq 0 \\
& S_{j0} = H_j \\
& E_{j0}, I_{j0}, R_{j0}, C_{j0} = 0
\end{aligned}
\tag{2}
$$

Here, λ_{jit} is binary data where a 1 indicates that the i-th NPI is active at time t for county j and the β_i^Δ variables represent the reduction in transmission as a function of the intervention. Formulation (2) simultaneously estimates compartment counts, model parameters and the model for the transmission rates. Although there may be nonlinear dependence on the interventions, this model provides a base case for investigating the performance and effectiveness of more complex interactions.

We demonstrate this formulation on a case study using simulated data. For this study, we assumed three different non-pharmaceutical interventions, each with a different reduction in the transmission parameter. We determine the base transmission parameter for each county β_j^0 used in each simulation randomly (uniformly selected from 0.25-0.75). The timing of seeding of external transmissions was determined to occur randomly between day 0 and day 30 for each county. The number of external transmissions was set to 5 for each of the counties. The time profiles for different interventions across the counties were also specified randomly. The duration of the intervention was uniformly drawn from 10-75 days, and the starting day of the intervention was drawn uniformly from 10-140 days (minus the duration).

These specifications were used to generate simulation data for 100 counties. Noise (10% standard deviation) was added to the reported cases from the simulation before passing this data to the inference implementation. Table 1 below shows the simulated and estimated values for the β_i^{Δ} parameters indicating the estimated reduction observed by the interventions. Additional studies with less noise and more counties resulted in increased accuracy of the estimated parameters.

Intervention	True Impact on β (Simulation)	Estimated Impact on β (Inference)
NPI-1	**-0.119**	**-0.115**
NPI-2	**-0.197**	**-0.229**
NPI-3	**-0.241**	**-0.232**

Table 1. Simulated and estimated impact of different intervention strategies

Figure 2 shows the reported cases, estimated infections, and estimated transmission parameter profiles for this case study. Our results indicate that NPI efficacy can be estimated simultaneously with other parameters in an infectious disease model. Estimates of the impact of interventions is improved with data from multiple regions (e.g. counties). Our results demonstrate robustness to data errors, though there may be challenges with correlated errors.

Data-Parallel Nonlinear Parameter Estimation

Scalable methods for solving Formulation (2) is important for this large-scale problem as we increase the size (time span and geographical discretization considered), and as we increase model complexity (e.g., extending the model to track multiple variants, additional compartments for vaccination status and reporting processes). Thus, we considered strategies for parallelizing parameter estimation. In particular, we observed that Formulation (2) is a weakly-coupled model. If the impact of interventions are specified (the β_i^{Δ} variables), then the remaining parameters can be independently determined for each county. Consequently, this model is well-suited for parallel solution with advanced decomposition-based optimization techniques.

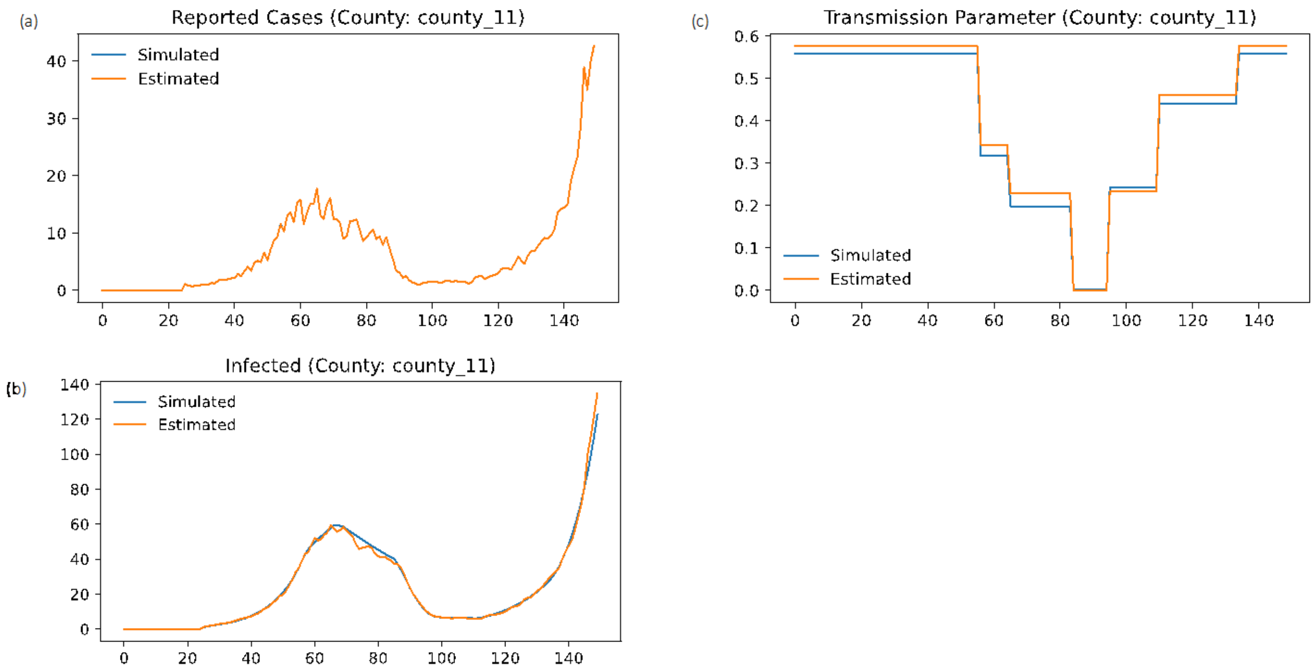


Figure 2. Estimation results from simulated data for 100 counties. Results shown for county #11. Figure (a) shows the reported cases, figure (b) shows the simulated and estimated values of infections over time, and figure (c) shows the simulated and estimated transmission parameter profiles.

We demonstrate the scalability of this inference problem using the parallel nonlinear optimization package, Parapint [7]. Parapint implements an interior-point method, where the dominant computational cost is the solution of a large linear system to compute the step direction at each iteration of the algorithm. Parapint is built on the principle that a large-scale optimization problem is inherently structured, and the structure of the problem will inherently induce structure in the large linear system describing optimality conditions. Parapint is built on Pyomo and PyNumero, which provides interfaces for building complex, performant optimization algorithms in Python.

We implemented Formulation (2) with Pyomo and applied Parapint to perform data-parallel parameter estimation on a test case with 1000 counties. Figure 3 below shows strong scaling results on this problem, where Parapint is compared against the runtime performance of a serial interior point solver. On 128 cores, this approach was able to achieve 94X speedup, obtaining a solution in only 11 seconds. This fast runtime enables epidemiologists to interact with these models and explore the impact of changes in model parameters.

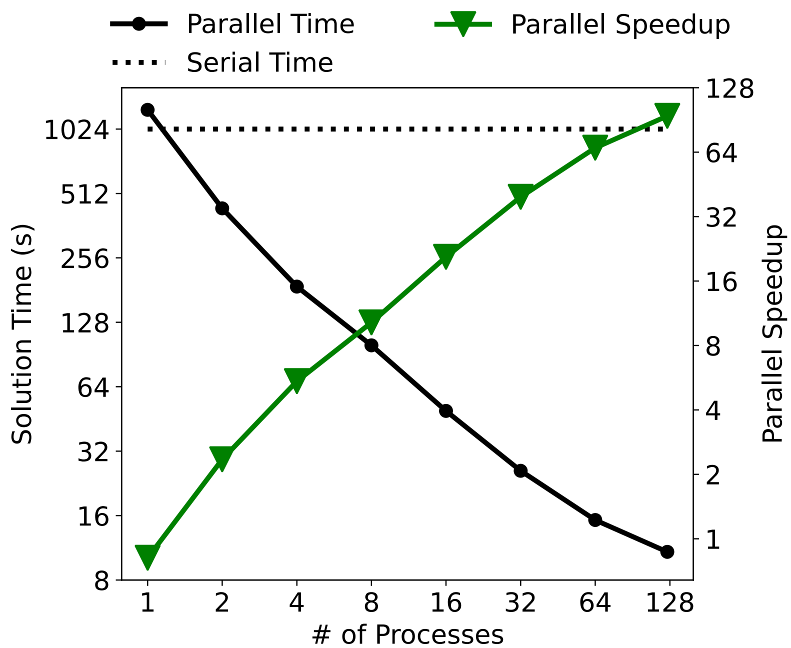


Figure 3. Strong scaling results with Parapint for a parameter estimation problem with 1000 counties.

Accelerating Machine Learning Models in Pyomo

Complex, nonlinear models for the impact of interventions on β_{jt} may be necessary to capture nonlinear interactions between NPIs. For example, the transmission parameter could be represented by a multilayer neural network model, for which the county baseline transmission parameter (i.e., variables β_j^0) and the intervention timing (i.e., the known values λ_{jit}) are inputs to the network while the output layer is a single prediction of β_{jt}. Standard neural network representations can be added to the Pyomo model directly as continuous, differentiable expressions. Consequently, the optimization methods used for Formulations (1) and (2) should be applicable to these new formulations.

Formulations (1) and (2) were implemented using the Pyomo optimization modeling software [2]. Preliminary research representing machine learning models with Pyomo has highlighted performance bottlenecks that limit the size of a neural network model that can be practically expressed. This motivated this project to address performance bottleneckes in Pyomo to accelerate the expression of machine learning models.

Previous LDRD research has identified several representations of neural networks for embedding in optimization problems in Pyomo models. In the *full-space representation*, all intermediate variables from each layer are seen by the optimizer. Pyomo provides fast, sparse construction of this representation, but these problems can be difficult to solve. In the *reduced-space representation*, the parametric model is expressed in terms of inputs and outputs only, with deep expression trees that substitute out the intermediate variables. These deep expression trees are computationally expensive for Pyomo to generate, because its design is tailored for large, sparse models.

This project focused on improvements to the Pyomo framework for translating dense expressions from Pyomo to the NL format that used by many nonlinear optimization solvers. We used a dense 3-layer neural network model with 100 nodes per layer to test Pyomo's performance. Prior to our improvements, the reduced-space representation of the neural network model took approximately 42 seconds for translation to the NL format.

Pyomo was improved with three developments: (1) profiling the NL writer led to removal of bottlenecks associated with deep expression trees, (2) the implementation was modified to require only a single pass through the expression tree, and (3) intermediate, repeated expressions are cached. After these improvements, this same neural network required only ~1.5 seconds for translation, which is a 27 times reduction in computational time. These improvements are implemented as part of the new NL writer in Pyomo

ANTICIPATED OUTCOMES AND IMPACTS

The main goal of this project is to estimate changes in infectious transmission over time and space as a decision-making aid for policy makers. As part of early work on this project, we refined and published earlier research, which led to the following publications:

- D. Cummings, W.E. Hart, B. García-Carreras, C.D. Laird, E.C. Lee, J. Lessler, A. Staid, *Spatio-temporal Estimates of Disease Transmission Parameters for COVID-19 with a Fully-Coupled, County-Level Model of the United States*, Sandia National Laboratories, September, 2021. (to appear)
- Yang, B., Huang, A. T., Garcia-Carreras, B., Hart, W. E., Staid, A., Hitchings, M. D., ... & Cummings, D. A. (2021). Effect of specific non-pharmaceutical intervention policies on SARS-CoV-2 transmission in the counties of the United States. Nature communications, 12(1), 1-10.

This research demonstrated capabilities for estimating the unobserved state of a disease progression, which could be used to initialize simulation models like COVIDScenarioPipeline. Similarly, this capability was used to inform our assessment of the impacts of NPIs [8].

Subsequent research in this project has focused on extending these demonstrations with more sophisticated models that (1) simultaneously estimate compartment and transmission parameters, and (2) simultaneously estimate compartment parameters and a predictive model for transmission rates. These extensions provide a more flexible, integrated analysis of disease progression, but they lead to significantly more complex, nonlinear parameter estimation problems. We expect this research to inform ongoing collaborations academic collaborators at CMU and U Florida, providing new insights into the relative utility of NPIs for the U.S. response to the COVID-19 pandemic.

This research is an exemplar for the integration of data-informed models with algebraic models that are commonly used in operations research (OR). The computational bottlenecks addressed here reflect a broader challenge for the integration of machine learning and OR methods, relating to both model generation and model optimization. This challenge has been recognized by a variety of Sandia researchers, and the "Data-Informed Operations Research" LDRD proposal has been submitted to address this challenge (NSIST, FY22-24).

CONCLUSION

The Sandia LDRD Project "Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections" has developed new strategies for modeling the spread of infectious diseases to demonstrate nonlinear programming methods for simultaneous estimation of unobserved states in the population, parameters describing disease dynamics, and the impact of nonpharmaceutical interventions. With these tools it is possible to simultaneous estimate these states and parameters within a flexible framework that supports straightforward improvements and extensions of the compartment model. Furthermore, we demonstrated that there is opportunity for performance improvements of these large-scale inference problems with parallel nonlinear tools (i.e., Parapint).

The capabilities demonstrated in this project represent will catalyze future work in the analysis of large-scale coupled infectious disease models. Given ongoing improvements of optimization tools like Pyomo, there is also opportunity to directly estimate machine learning models hybridized with the existing structure of the compartment model. Additionally, scalable optimization solvers like Parapint will enable rapid interaction with national-scale models. However, future research is needed to demonstrate these capabilities on real-world data and inter-county interactions. Additionally, analysis of NPIs will require detailed information about per-county policy executions. Collection of this information is currently prohibitively difficult in most cases, but this capability could demonstrate the utility of this information and motivate its collection by government agencies.

LABORATORY DIRECTED RESEARCH & DEVELOPMENT
WHERE INNOVATION BEGINS

REFERENCES

[1] R.M Anderson, M.M Robert. *Infectious diseases of humans: Dynamics and control*. Oxford Science Publications, 36:118, 1991.

[2] M.L. Bynum, G.A. Hackebeil, W.E. Hart, C.D. Laird, B.L. Nicholson, J.D. Siirola, J.-P. Watson, D.L. Woodruff. *Pyomo - Optimization modeling in python, 3rd Edition*, Volume 67. Springer International Publishing, 2021.

[3] COVIDScenarioPipeline, 2020. https://github.com/HopkinsIDD/COVIDScenarioPipeline

[4] D. Cummings, W.E. Hart, B. García-Carreras, C.D. Laird, E.C. Lee, J.Lessler, A. Staid, *Spatio-temporal Estimates of Disease Transmission Parameters for COVID-19 with a Fully-Coupled, County-Level Model of the United States*, Sandia National Laboratories, September, 2021. (to appear)

[5] H.W. Hethcote. *The mathematics of infectious diseases*. SIAM review, 42(4):599–653, 2000.

[6] J.C. Lemaitre, K.H. Grantz, J. Kaminsky, H.R. Meredith, S.A. Truelove, S.A. Lauer, L.T. Keegan, S. Shah, J. Wills, K. Kaminsky, J. Perez-Saez, J. Lessler, E.C. Lee. *A scenario modeling pipeline for COVID-19 emergency planning.* medRxiv, 2020. doi: 10.1101/2020.06.11.20127894. URL https://www.medrxiv.org/content/early/2020/09/11/2020.06.11.20127894

[7] Parapint, 2021. https://github.com/Parapint/parapint

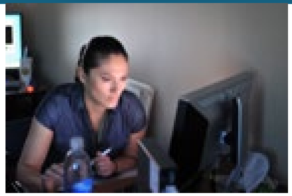
[8] B. Yang, et al. (2021). *Effect of specific non-pharmaceutical intervention policies on SARS-CoV-2 transmission in the counties of the United States*. Nature communications, 12(1), 1-10.

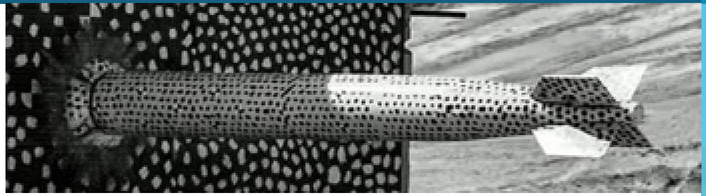
ADDENDUM A

CIS LDRD REVIEW PRESENTATION

LDRD Ending Project Review

Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections (Project 223167)





PI: William E. Hart (01464), PM: John Feddema (01460)

Team: M. Bynum, C.D. Laird, J. Siirola, A. Staid

FY21, $58k

Summary

This project developed large-scale nonlinear optimization approaches for simultaneous estimation of model parameters and unobserved SARS-CoV-2 infections

Motivation

- Well-informed policies for non-pharmaceutical interventions (NPIs) are critical to mitigate COVID outbreaks and plan effective responses for variants with new infectious characteristics.
- No estimation approaches exist that can simultaneously estimate time-varying transmission parameters, reporting rates, and unobserved states with a fully-coupled all-county model of the United States.

Goals

- Extend Sandia's previous estimation approaches, which separately estimate unobserved infections and transmission parameters
- Integrate a predictive model for the utility of NPIs to estimate the effect of NPI policies on transmission rates
- Demonstrate scalable solution strategies on large, fully-coupled nonlinear formulations

Overview and Accomplishments

Description	Details	Outcomes
Reconstruction of unobserved transmissions	Finalized FY20 activities. Prepared new results for NPI journal article.	SAND Report
Transmission parameter inference	Finalized FY20 activities.	SAND Report
Assessing the effects of NPI policies	Revised estimates and submitted journal article	Journal Article (Nature Communications)
Simultaneous estimation of compartment and transmission parameters	Developed new formulations. Demonstrated ability to track fast changes in transmission rates.	LDRD Report
Simultaneously estimation with prediction of NPI efficacy	Developed new formulations. Demonstrated ability to identify linear NPI relationships. Scalability demonstration using Parapint.	LDRD Report
Accelerating machine learning models in Pyomo	Optimized Pyomo model generation for machine learning models with large, repeated subexpressions	LDRD Report, Pyomo software contributions

Inference model – Overview

$$
\begin{aligned}
\min \quad & \sum_{j\in J}\sum_{t\in\mathcal{D}} \left(C_{jt} - \hat{C}_{jt}\right)^2 + \alpha_\epsilon^T ||\epsilon^T||_1 \\
\text{s.t.} \quad & T_{jt} = \frac{\beta_{jt} I_{jt} S_{jt}}{H_j} + \epsilon_{jt}^T \\
& \beta_{jt} = \beta_j^0 + \sum_{i\in NPI} \beta_i^\Delta \lambda_{j,i,t} \\
& S_{j,t+1} = S_{jt} - T_{jt} \\
& E_{j,t+1} = E_{jt} + T_{jt} - \sigma E_{jt} \\
& I_{j,t+1} = I_{jt} + \sigma E_{jt} - \gamma I_{jt} \\
& R_{j,t+1} = R_{jt} + \gamma I_{jt} \\
& C_{j,t+1} = \rho\sigma E_{jt} \\
& \beta_{jt} \geq 0 \\
& S_{j0} = H_j \\
& E_{j0}, I_{j0}, R_{j0}, C_{j0} = 0
\end{aligned}
$$

Objective: MLE fit for reported cases with 1-norm regularization of the transmission noise

Transmissions follow common SEIR model with time-varying transmission parameter and model noise

Spatially and temporally varying transmission parameters

- Base transmission parameter value is county dependent
- Transmission parameter adjusted by interventions
- County-specific intervention timing
- Multiple counties needed to estimate intervention impact

Reported cases subject to under-reporting

Large-scale nonlinear programming model (# days x # counties)

Inference model – Comparison with Prior SNL Models

$$
\begin{aligned}
\min \quad & \sum_{j \in J} \sum_{t \in T} \left(C_{jt} - \hat{C}_{jt} \right)^2 + \boxed{\alpha_\epsilon^T ||\epsilon^T||_1} \\
\text{s.t.} \quad & T_{jt} = \boxed{\frac{\beta_{jt} I_{jt} S_{jt}}{H_j}} + \boxed{\epsilon_{jt}^T} \\
& \beta_{jt} = \boxed{\beta_j^0 + \sum_{i \in NPI} \beta_i^\Delta \lambda_{j,i,t}} \\
& S_{j,t+1} = S_{jt} - T_{jt} \\
& E_{j,t+1} = E_{jt} + T_{jt} - \sigma E_{jt} \\
& I_{j,t+1} = I_{jt} + \sigma E_{jt} - \gamma I_{jt} \\
& R_{j,t+1} = R_{jt} + \gamma I_{jt} \\
& C_{j,t+1} = \rho \sigma E_{jt} \\
& \beta_{jt} \geq 0 \\
& S_{j0} = H_j \\
& E_{j0}, I_{j0}, R_{j0}, C_{j0} = 0
\end{aligned}
$$

<u>Regularization</u> – Better prediction of fast changes

<u>Nonlinearities</u> – Added to simultaneously fit compartment and transmission parameters

<u>Mobility</u> - Omitted to simplify assessment of impact of nonlinearities

Accounting for stochasticity in transmission

<u>NPI Effects</u> – Added to account for the impact of NPI policies

Simulation-Estimation Studies Were Used to Test Estimation Methods

$$
\begin{aligned}
\min \quad & \sum_{j \in J} \sum_{t \in \mathcal{D}} \left(C_{jt} - \hat{C}_{jt} \right)^2 + \alpha_\epsilon^T ||\epsilon^T||_1 \\
\text{s.t.} \quad & T_{jt} = \frac{\beta_{jt} I_{jt} S_{jt}}{H_j} + \epsilon_{jt}^T \\
& \beta_{jt} = \beta_j^0 + \sum_{i \in NPI} \beta_i^\Delta \lambda_{j,i,t} \\
& S_{j,t+1} = S_{jt} - T_{jt} \\
& E_{j,t+1} = E_{jt} + T_{jt} - \sigma E_{jt} \\
& I_{j,t+1} = I_{jt} + \sigma E_{jt} - \gamma I_{jt} \\
& R_{j,t+1} = R_{jt} + \gamma I_{jt} \\
& C_{j,t+1} = \rho \sigma E_{jt} \\
& \beta_{jt} \geq 0 \\
& S_{j0} = H_j \\
& E_{j0}, I_{j0}, R_{j0}, C_{j0} = 0
\end{aligned}
$$

Simulations with 3-1000 counties for 150 days

Disease model parameters from COVID-19 literature (e.g., latent, infectious)

Base transmission parameters (β_j^0) for each county uniformly distributed between 0.25-0.75

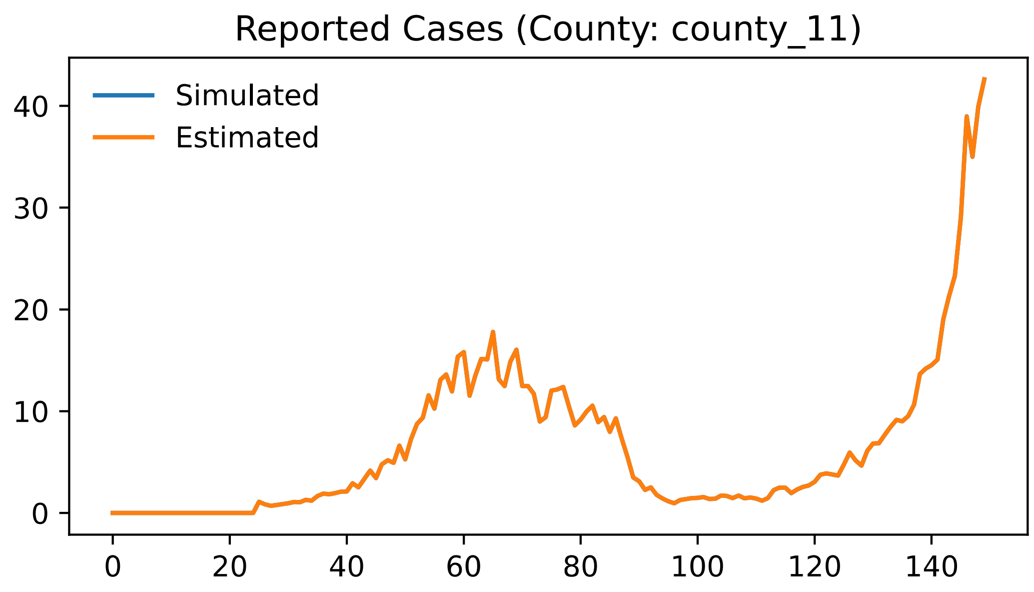
- R_eff: ~1-3

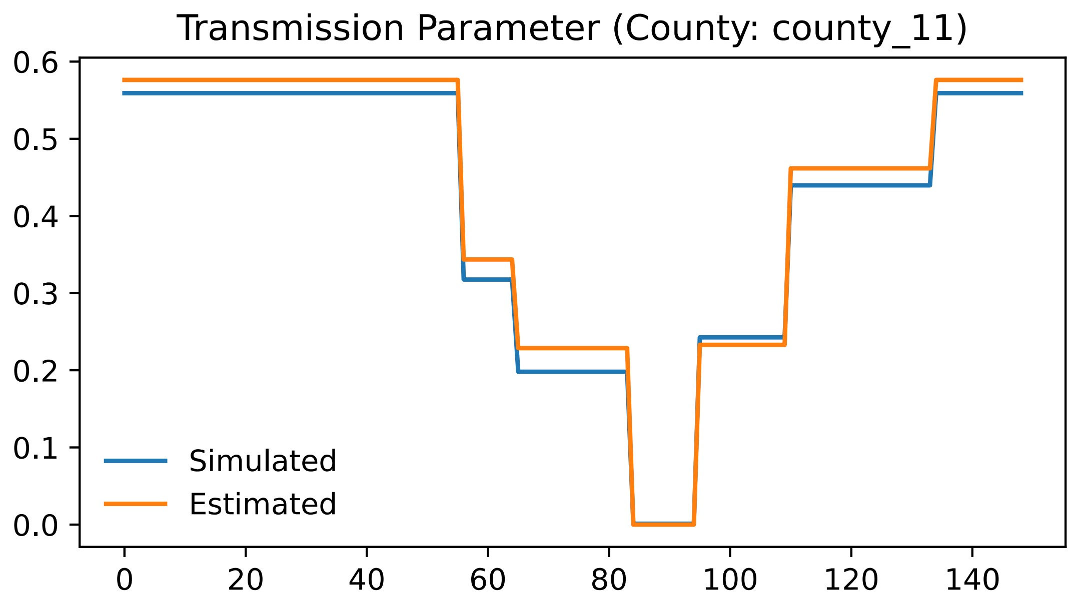
Intervention timing start and duration determined randomly for each county

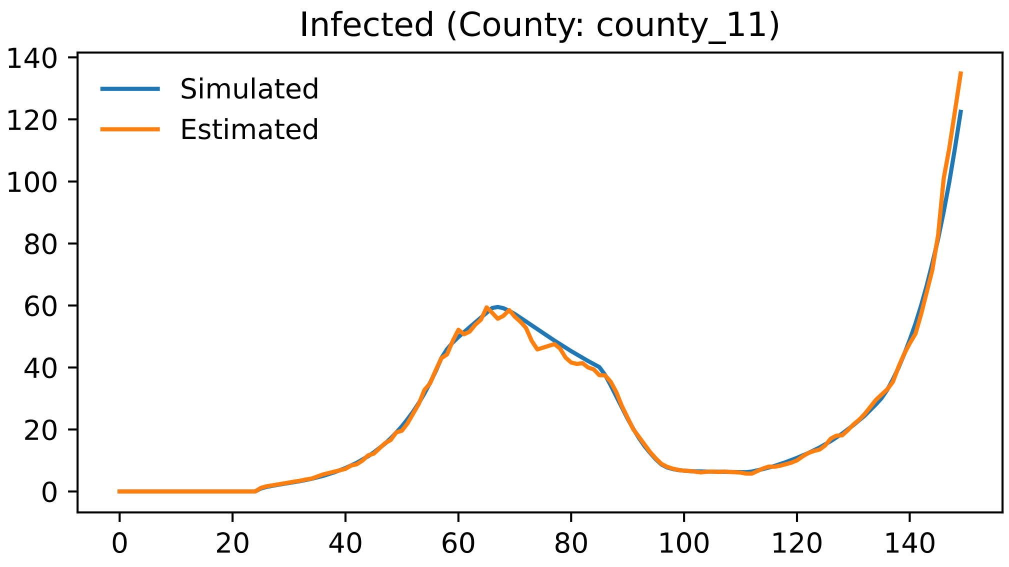
Simulations performed with significant stochastic noise on reported cases

- increased accuracy with reduced noise

Estimation Results with 100 Counties







Estimation of coupled parameters

Intervention	True Impact on β (Simulation)	Estimated Impact on β (Inference)
NPI-1	-0.119	-0.115
NPI-2	-0.197	-0.229
NPI-3	-0.241	-0.232

Increased accuracy with reduced noise and more counties

Next Steps – Scaling Up the Analysis

National-Scale Models
- Require data from ~3000 counties
- Data for ~1.5 years

Model Initialization
- Nonlinear optimizers are sensitive to the initial point used for optimization
- We have used a per-county optimization for initialization, but faster methods are often possible

Mobility
- This models inter-county interactions and exposures
- Significantly increases the difficulty of parameter estimation

Nonlinear NPI Models
- We expect NPIs to have correlated effects
- The linear model of NPI effects could be replaced with machine learning models
- These introduction additional nonlinearities and estimation parameters

Next Steps Completed
- Data-parallel parameter estimation with Parapint
- Accelerating Pyomo for machine learning models

Data-Parallel Parameter Estimation with Parapint

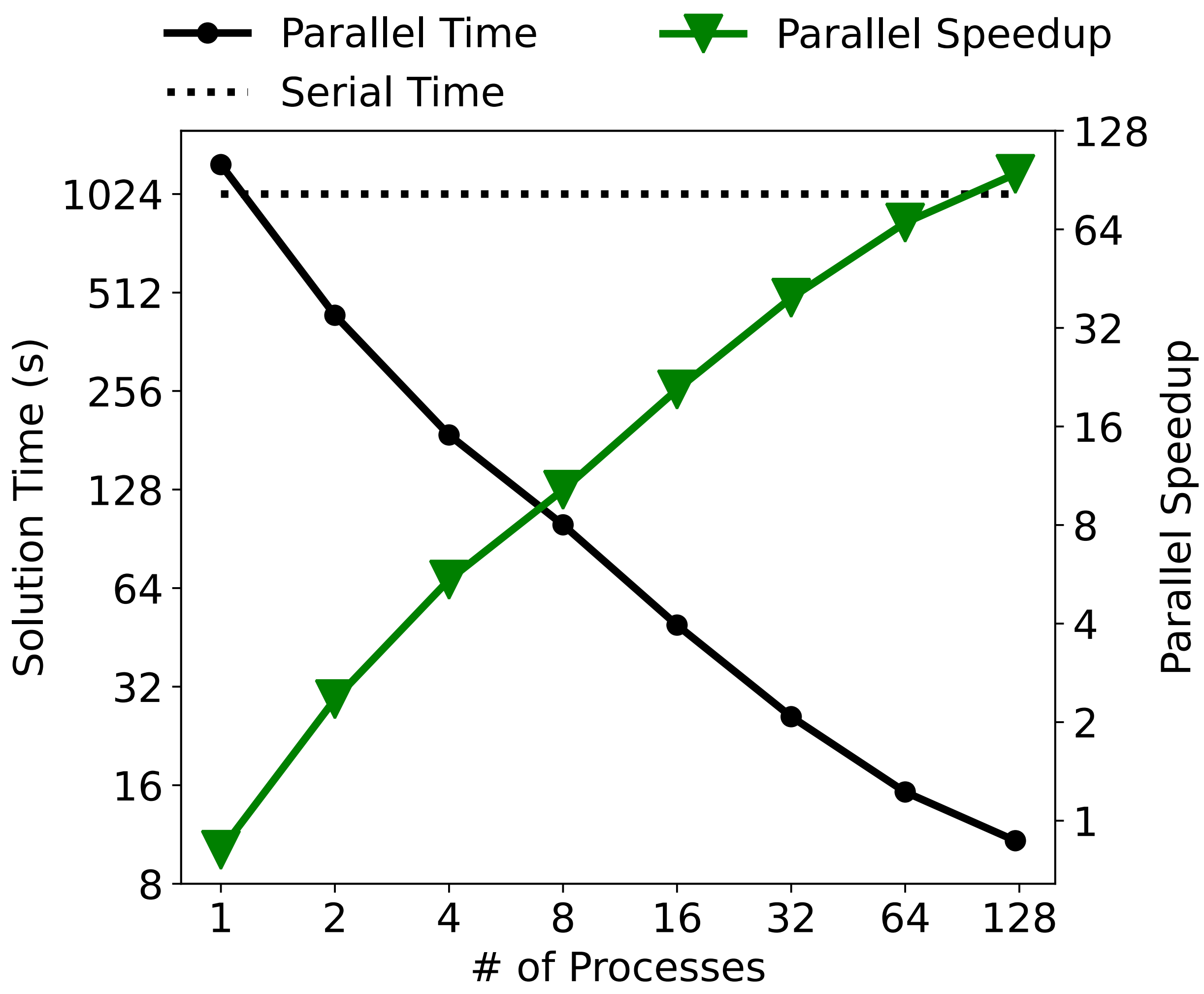
Goal: fast estimation to support exploratory analysis of model parameters

Idea: Parallelize estimation using Parapint

- Originally developed for parallel solution of dynamic optimization problems
- Structured decomposition can used for data-parallel estimation problems

Scalability study

- Compute end-to-end runtime, including model setup and initialization
- Computed solution time and speedup relative to serial optimizer
- **Solution on 128 cores in 11 seconds!**



Accelerating Pyomo for Machine Learning Models

Goal: Enable nonlinear representation of NPI effects

$$\beta_{jt} = \beta_j^0 + \sum_{i \in NPI} \beta_i^{\Delta} \lambda_{j,i,t} \quad \rightarrow \quad \beta_{jt} = NNet(\beta_j^0, \lambda_{j,t}; w, b)$$

Full-space NNet representation
- Includes all intermediate variables from each layer
- Large problems that are challenging to solve

Reduced-space NNet representation
- Consistent with tools like TensorFlow
- Express NNet as a single large expression tree

Pyomo improvements for NL model writer
- Reduce # of passes through the expression tree
- Caching of intermediate repeated expressions

	Full	Reduced
Old Time	0.1s	42.0s
New Time	0.1s	1.5s

Comparison of time for NL writer with 3-layer NNet model (100 nodes)

New NL writer is 27x faster!

PI's PROJECT LEGACY

Informing US COVID-19 Response

- The Nature Communications article informs future US policy on COVID-19 management
- Non-pharmaceutical interventions (NPIs) are a key tool for controlling the ongoing SARS-CoV-2 pandemic, even with the development of vaccines for this disease.
- Well-informed policies for NPIs are critical to mitigate COVID outbreaks and plan effective responses for COVID variants with new infectious characteristics.
- The modeling extensions demonstrated in this project provide a strategy for partially automating the assessment of NPI efficacy

Catalyzing future data-informed operations research (OR) applications

- The Pyomo enhancements will catalyze application of Pyomo in future machine learning (ML) applications
  - This will all support future efforts to develop hybrid OR-ML models using Pyomo
- The Parapint demonstration will catalyze future data-informed OR applications with this capability
  - Data-parallelism can be exploited to solve large parameter estimation problems
- Modeling disease dynamics is an application exemplar for the proposed NSIST DI-OR LDRD
  - Parameter estimation with NPI efficacy directly maps to one of the application exemplars in this project

PROJECT OUTPUTS

- Intellectual Property
  - None
- Publications
  - Yang, B., Huang, A.T., Garcia-Carreras, B. et al. *Effect of specific non-pharmaceutical intervention policies on SARS-CoV-2 transmission in the counties of the United States*. Nat Commun 12, 3560 (2021). https://doi.org/10.1038/s41467-021-23865-8
  - Hart, W.E., Laird, C.D., Staid, A. et al *Spatio-temporal estimates of disease transmission parameters for COVID-19 with a fully-coupled, county-level model of the United States*. Sandia National Laboratories, SAND 2021-XXXX, September, 2021.
  - Hart, W.E., Bynum, M., Laird, C.D., Siirola, J.D. *LDRD Final Report: Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections*, Sandia National Laboratories, SAND 2021-XXXXR, September, 2021.
- Presentations
  - None
- Awards, professional leadership/recognition
  - None

CAPABILITIES DEVELOPMENT

- Expected Impact of Capabilities on Future Work
  - The new formulations integrating predictive models of NPI efficacy are an exemplar of the type of data-informed operations research models that motivate the NSIST DI-OR LDRD proposal.
    - If funded, these models will be included in the library of application exemplars developed in that project.
  - The Parapint demonstration illustrates the scalable application of Parapint for parameter estimation
    - Analysts can perform national-scale computations in seconds
  - The new Pyomo model generation capabilities will be integrated into a future Pyomo release. This capability will facilitate the application of Pyomo to future data-driven applications.
- Career Development
  - None

TEAM BUILDING AND PARTNERSHIPS

- Carl Laird’s research team at CMU will continue this research with collaborators at U Florida and Sandia. We expect this research to continue to support COVID-19 analyses that support CDC/DHS policy decisions.
- The NSIST DI-OR LDRD proposal is a collaboration with Dr. Laird
- Joint journal article and SAND Report with collaborators at Hopkins and UF

IA/PM PROJECT LEGACY

How did this project contribute to IA strategic goals and objectives?

- This research project is strongly aligned with CIS data science research priorities
- This research demonstrated scalable methods that enable rapid national-scale analysis
  - Fast computations will empower epidemiologists to explore model parameterizations to understand model predictions
- It is unclear if this research will catalyze future COVID-centric funding
  - This reflects opportunities with external research programs rather than the innovative nature of this work

What are the key results from this research that will be useful to other current and future projects?

- The Nature Communications article is expected to inform future US policy on COVID-19 management
- The Pyomo enhancements and Parapint demonstration will catalyze future data-informed operations research applications

Technology insertion and follow-on funding for potential and realized ROI

- Pyomo is widely used in SPP and DOE projects
- The capabilities for data-informed operations research developed here will immediately support IDAES (J. Siirola), REDLY LDRD (W. Hart), DI-OR LDRD (W. Hart, proposed), Carrier SPP (R. Smith)

Large-scale Nonlinear Approaches for Inference of Reporting Dynamics and Unobserved SARS-CoV-2 Infections

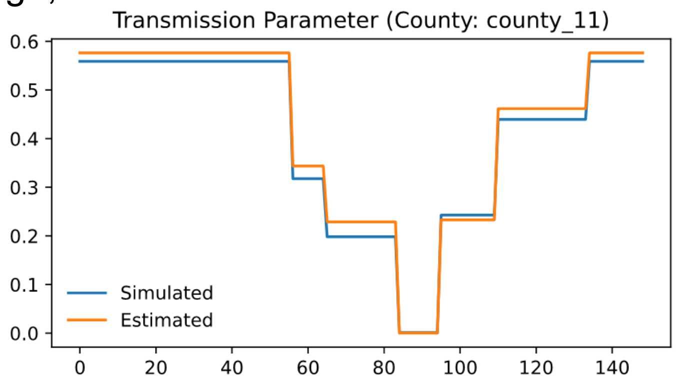
PI: William Hart, PM: John Feddema

Project goal(s)

Develop large-scale nonlinear optimization approaches for simultaneous estimation of model parameters and unobserved SARS-CoV-2 infections.

Predict the effect of nonpharmaceutical intervention (NPI) policies on transmission rates

Demonstrate scalable solution strategies for large, fully-coupled nonlinear formulations



Mission Impact

This research informs future US policy on COVID-19 management and response

The optimization capabilities demonstrate here can be leveraged in future data-informed operations research mission applications

Pyomo enhancements will resolve bottlenecks impacting any large, data-driven application

Key FY21 Accomplishments

Description
Assessing the effects of NPI policies (Journal Article)
Simultaneous estimation of compartment and transmission parameters
Simultaneously estimation with prediction of NPI efficacy
Accelerating machine learning models in Pyomo

Transition Plan

IDAES – Additional Pyomo enhancements and release of new Pyomo capabilities

NSIST DI-OR LDRD – Proposal to develop new capabilities for data-informed operations research