# Human-in-the-loop Scenario Forecasting for Infrastructure Capacity Planning

Jiaping Zhang, Ana Bertran, Lauren Valdivia

Jun 2020

# Infrastructure Analytics & Data Science

*Our team aims to create a world-class analytics platform that enables infrastructure teams to intelligently model and manage their resources with respect to capacity, usage, performance, growth, cost and incidents. Through continuous investment in the data science workflow, we provide access to standard datasets, implement algorithms, and automate data analyses that reduce time to insight and enable data-driven infrastructure applications.*

Ana Bertran

Lauren Valdivia

Jiaping(JP) Zhang

# Agenda

**Capacity Planning Business Problem**
- Capacity planning overview
- Why is forecasting important?
- Forecasting challenges

**Human-in-the-Loop Scenario Forecasting Approach**
- Incorporating human inputs
- Impact inference of future events in time series
- Forecasting as a service
- Visualization & UX design

**Lessons Learned & Future Work**

# Capacity Planning Overview

Capacity planners ensure **efficient utilization** of Salesforce infrastructure while maintaining **customer success** by optimizing capacity using their **supply and demand** knowledge.



Cost Management

Remediation Tracking

Customer Success

Capacity Planners

Risk Assessment

Supply chain Management

# Why is forecasting important for infrastructure?

*Forecasting enables data-driven capacity planning and infrastructure operation at scale*

- Salesforce infrastructure is multi-substrate and multi-tenant and complex with constant changes and expanding global footprint

- Effective capacity planning is required to ensure trust and customer success

- As Salesforce continuously rolls out new products and features for customers, proper infrastructure must be accurately planned for, given the scale of our product usage

# Challenges: The scale and variety of our time series metrics
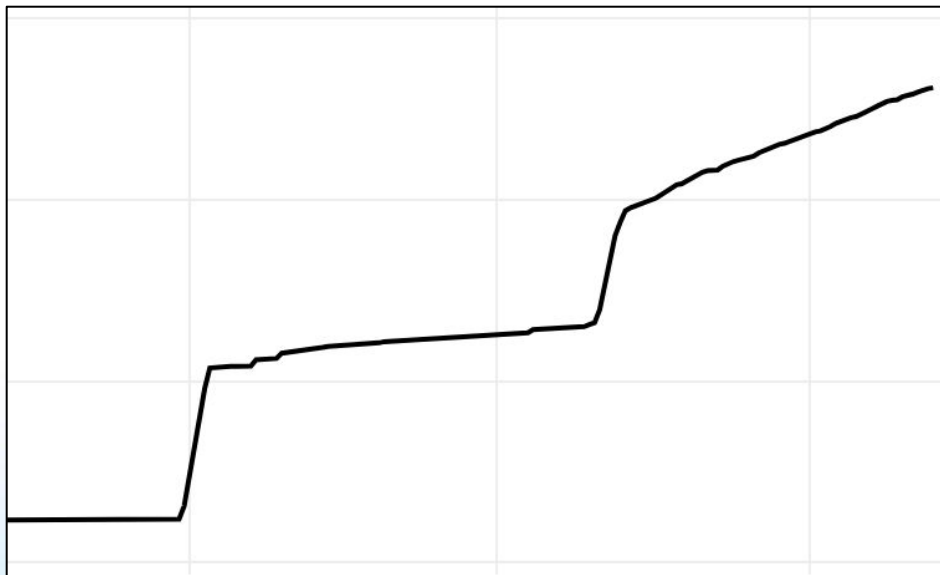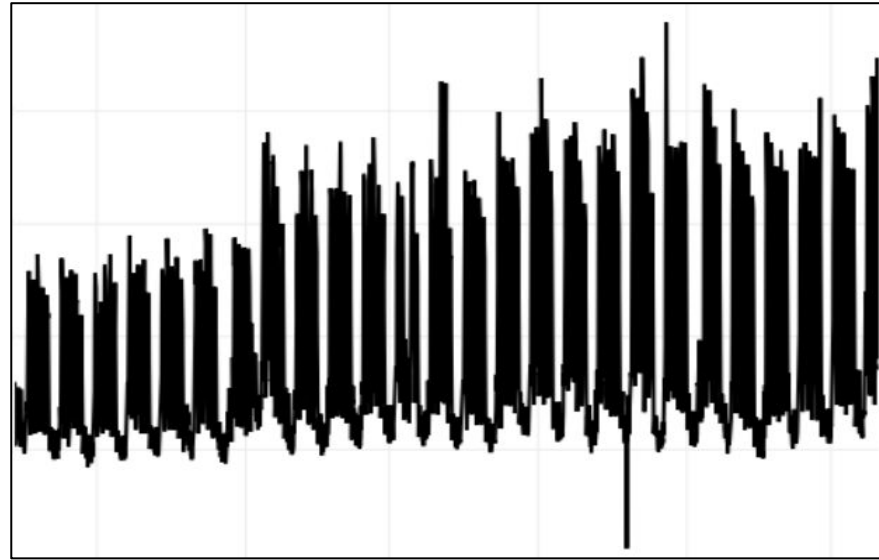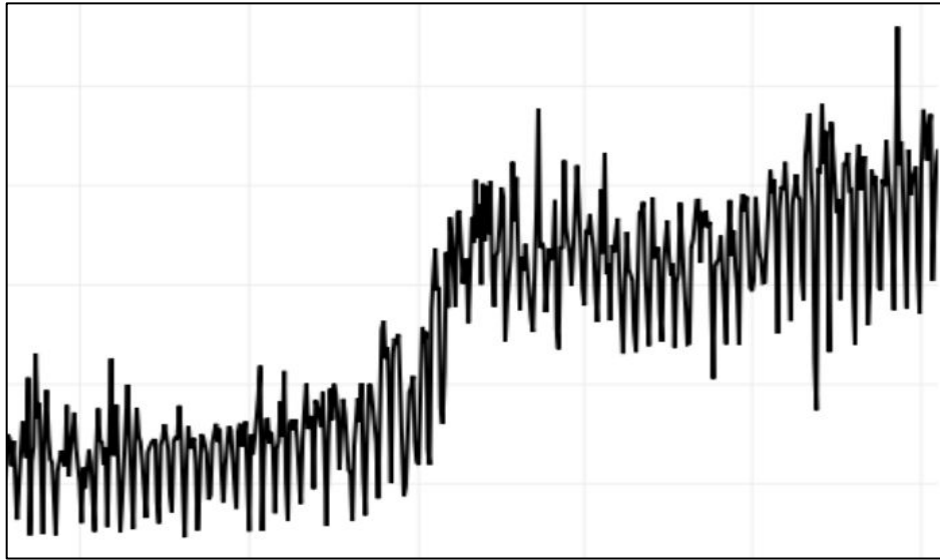
# Challenges: Planning for uncertainty
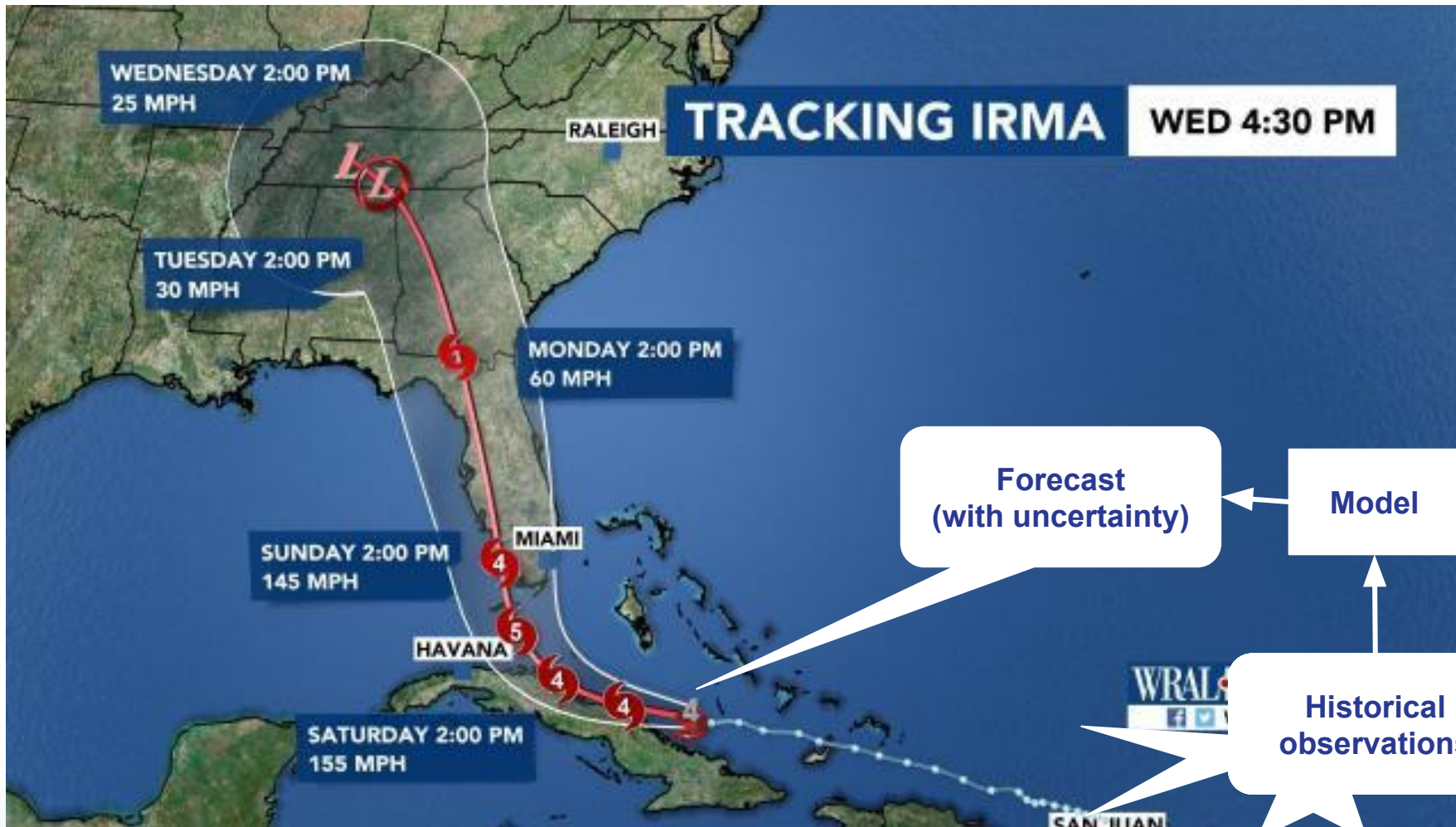


- Uncertainty of forecasts
- Contingency planning
- Future change points
- Domain knowledge

# Challenges: Planning for different scenarios

**Why It's So Freaking Hard To Make A Good COVID-19 Model**

By Maggie Koerth, Laura Bronner and Jasmine Mithani

Filed under Coronavirus

- X factors
- Different scenarios under certain constraints
- Long-term strategic planning
- Iterative decision-making

from FiveThirtyEight
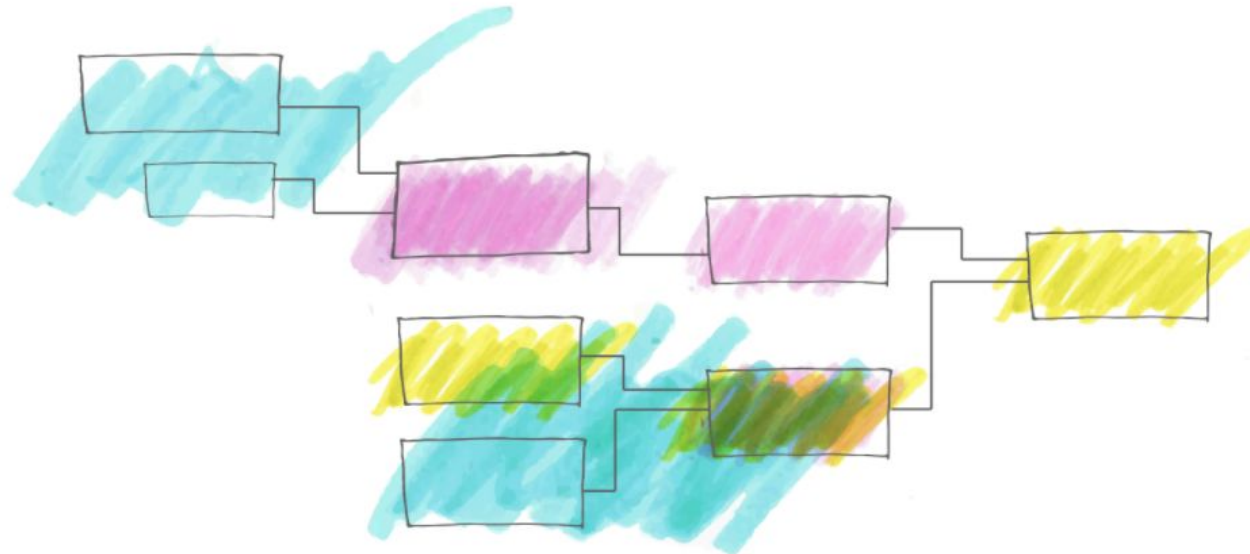
# Planning = Forecasting **+ Adjustments**

*in a business context*
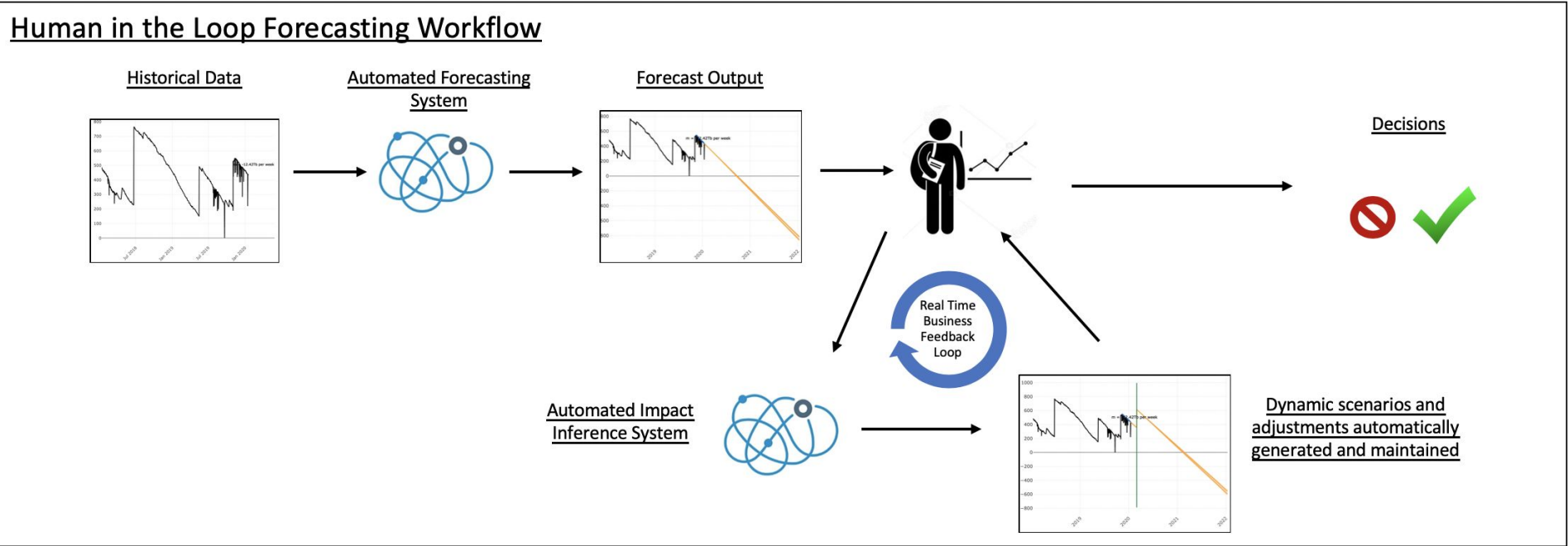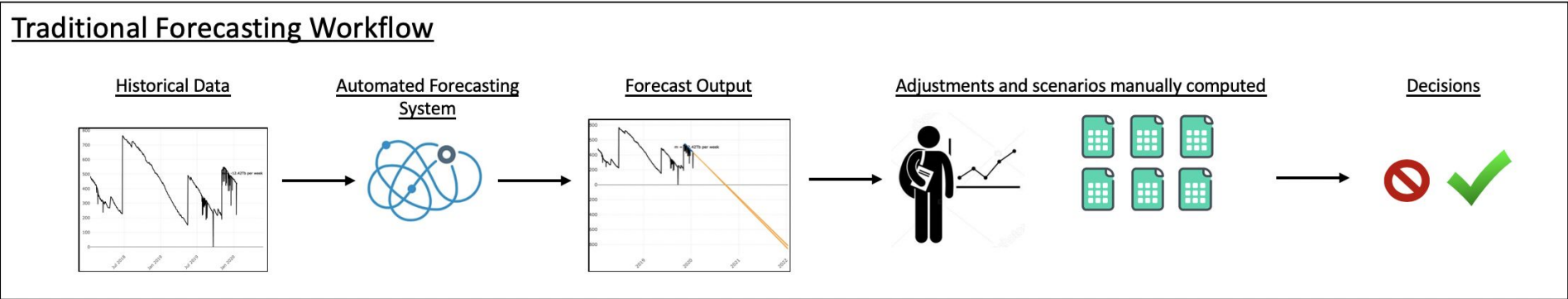
# Human-in-the-loop scenario forecasting framework
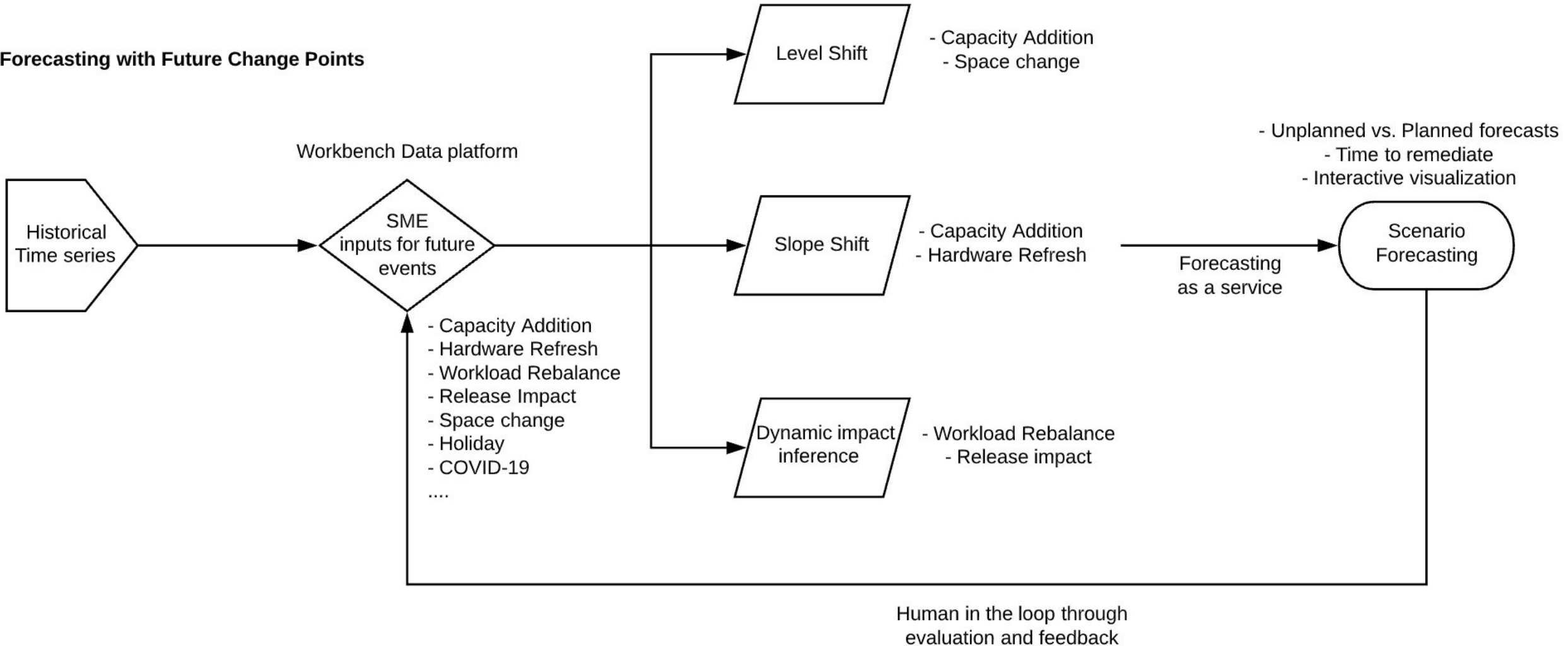
# Human Inputs

- Custom inputs from capacity planners

# Event Impact Inference Modeling

- Various events can result in heterogeneous impacts on time series in terms of level shift, slope change and dynamic state switch
- We need to build supplement models to simulate/predict the event impact, i.e hw gain, release impact, etc.
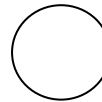
Hardware Gain
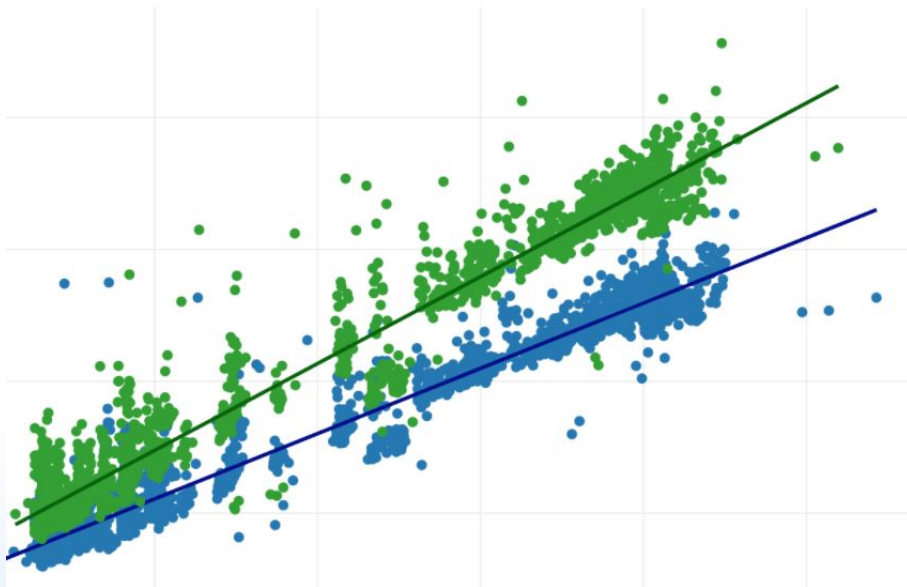
Release Impact Prediction
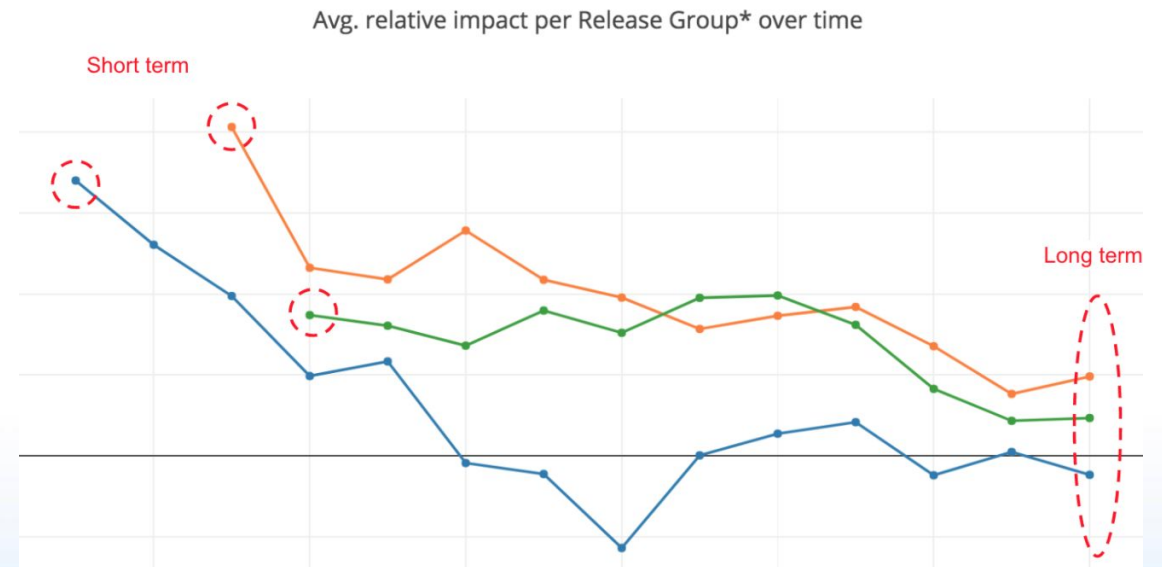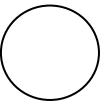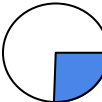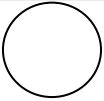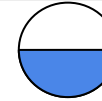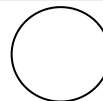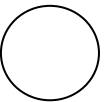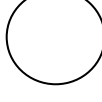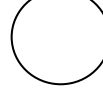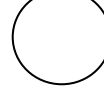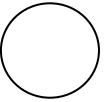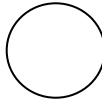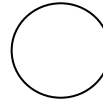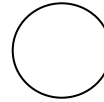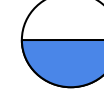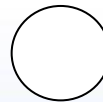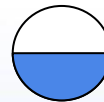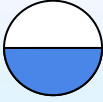
# Shepherd: Forecasting as a Service

**Selected Models** (Quantile regression, Facebook Prophet)

| | Linear regression | ARIMA | TBATS | Quantile regression | Facebook Prophet | **Shepherd** 🐕 |
|---|---|---|---|---|---|---|
| Deals with missing values and data preprocessing | full | empty | quarter | quarter | quarter | full |
| Robust to outliers, automatic anomaly detection | empty | small quarter | quarter | full | quarter | full |
| Fit segmented varying trend and seasonality | empty | half | empty | quarter | three-quarter | quarter |
| Auto detects change points | empty | empty | empty | empty | half | quarter |
| Adjusts for historical & **future change points** | empty | empty | empty | empty | half | full |
| Can leverage external features | half | full | empty | half | full | full |
| Mature packages | full | full | full | full | quarter | half |

# Interactive visualization & UX design

- Interactive visualization
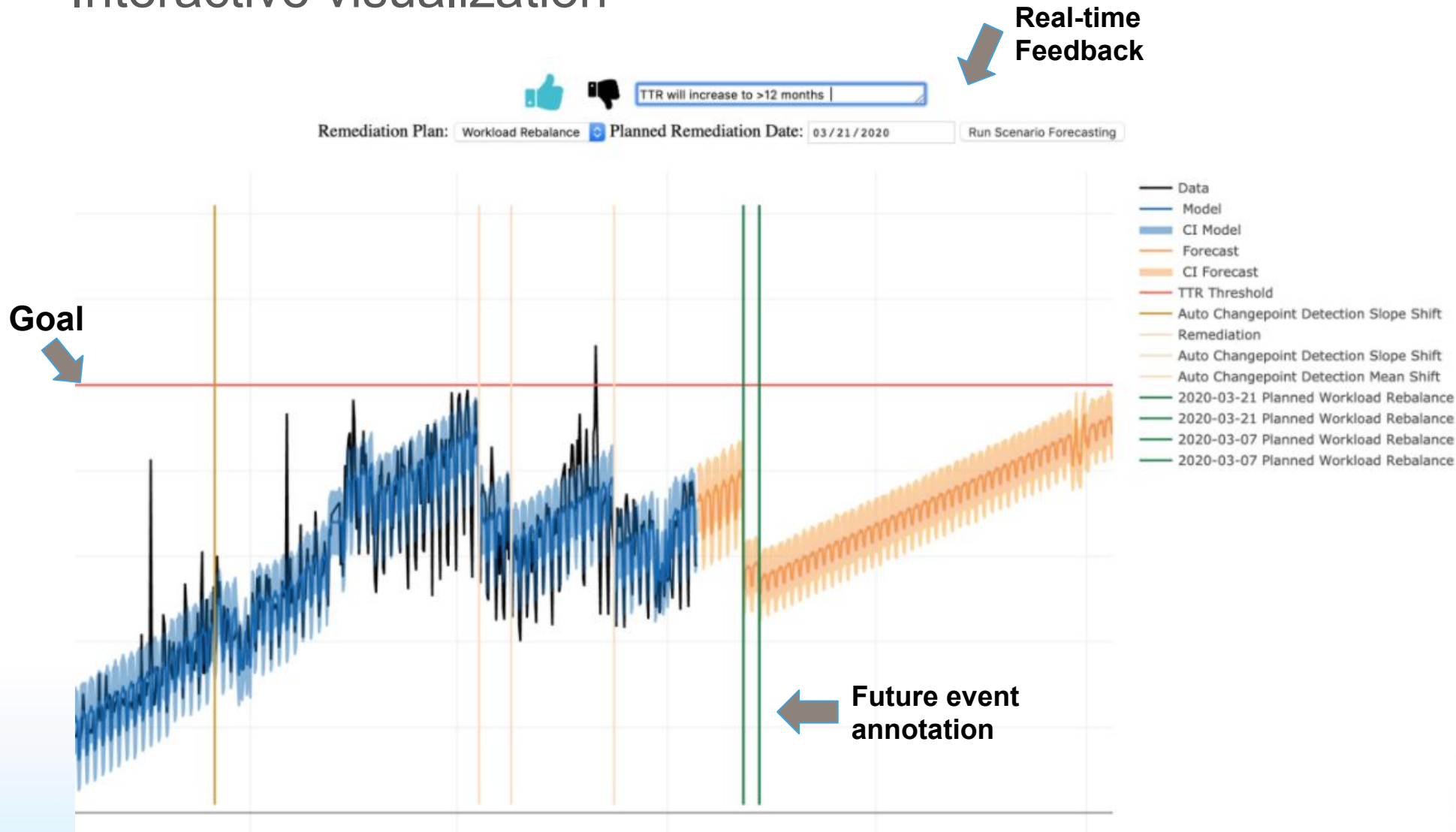
# Is it safe to make a certain remediation?

Inputs: Workload rebalance + future remediation date

Unplanned forecasts - Before
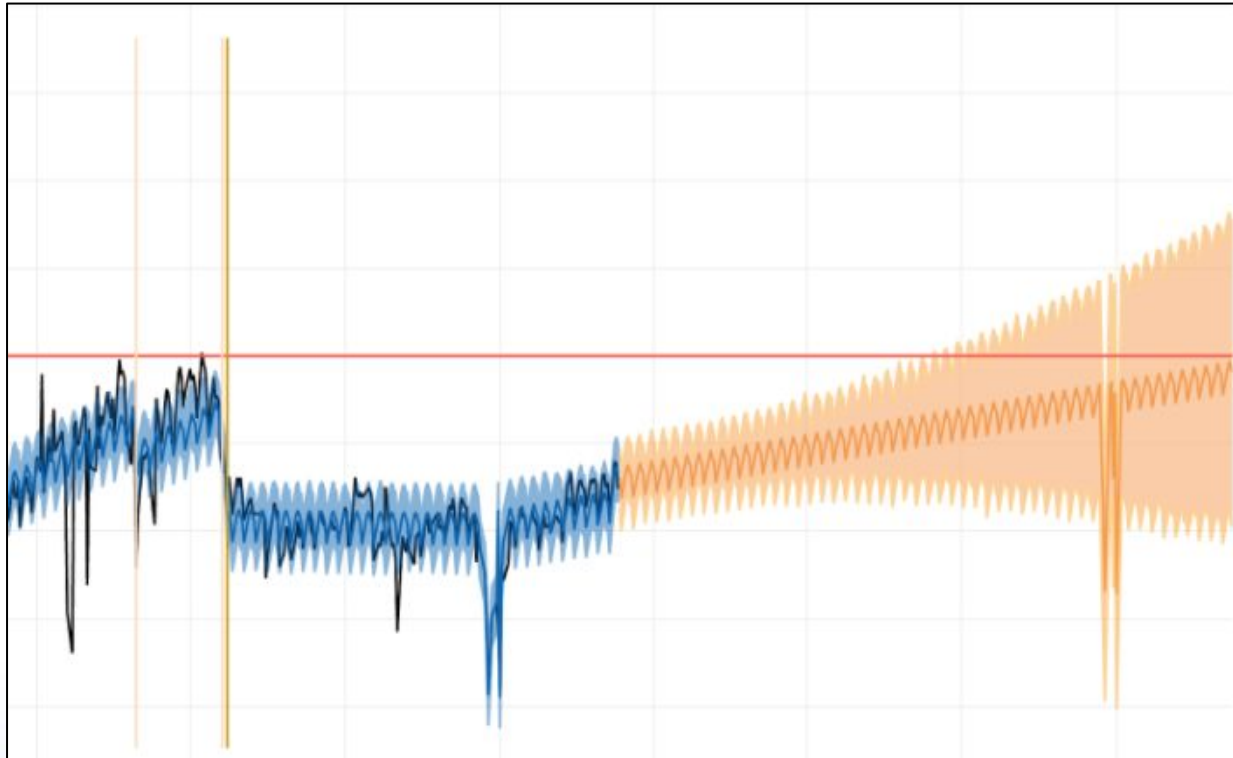
Planned forecasts
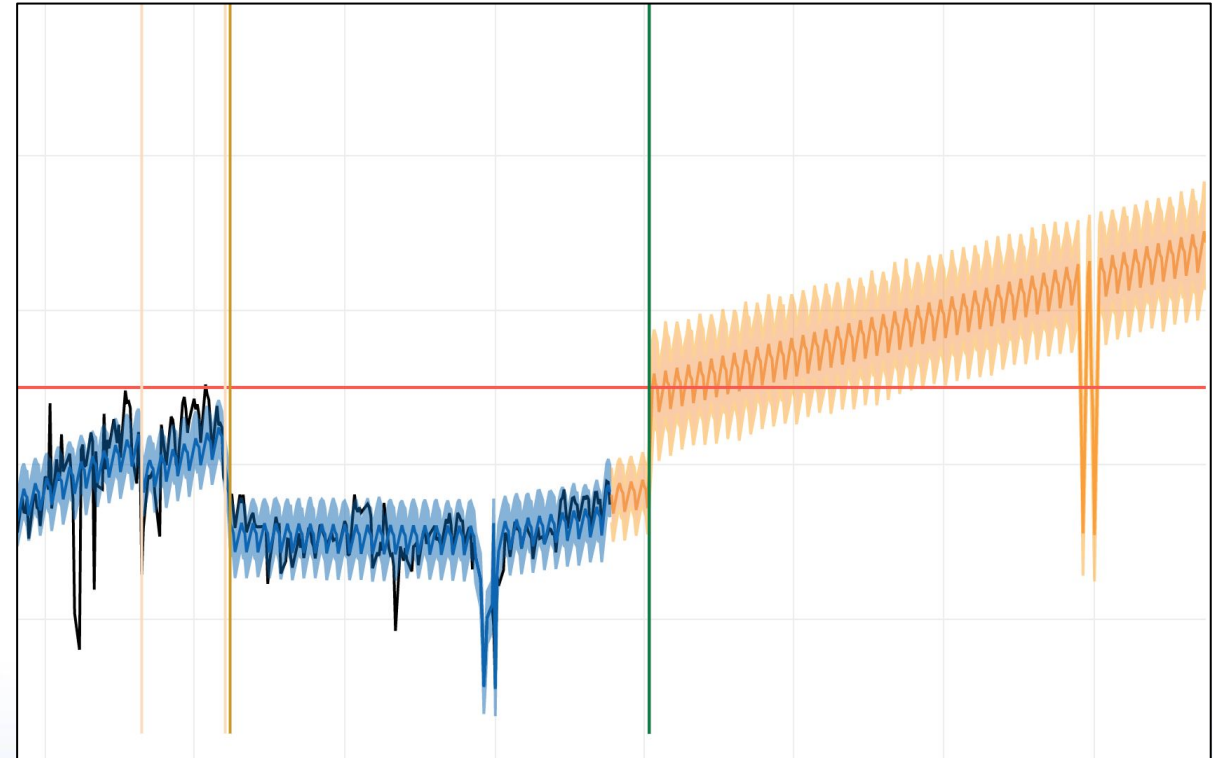
# How can we account for various degree of release regression impact for capacity planning?

Inputs: Predicted impacts of release regression + impact duration

Small impact
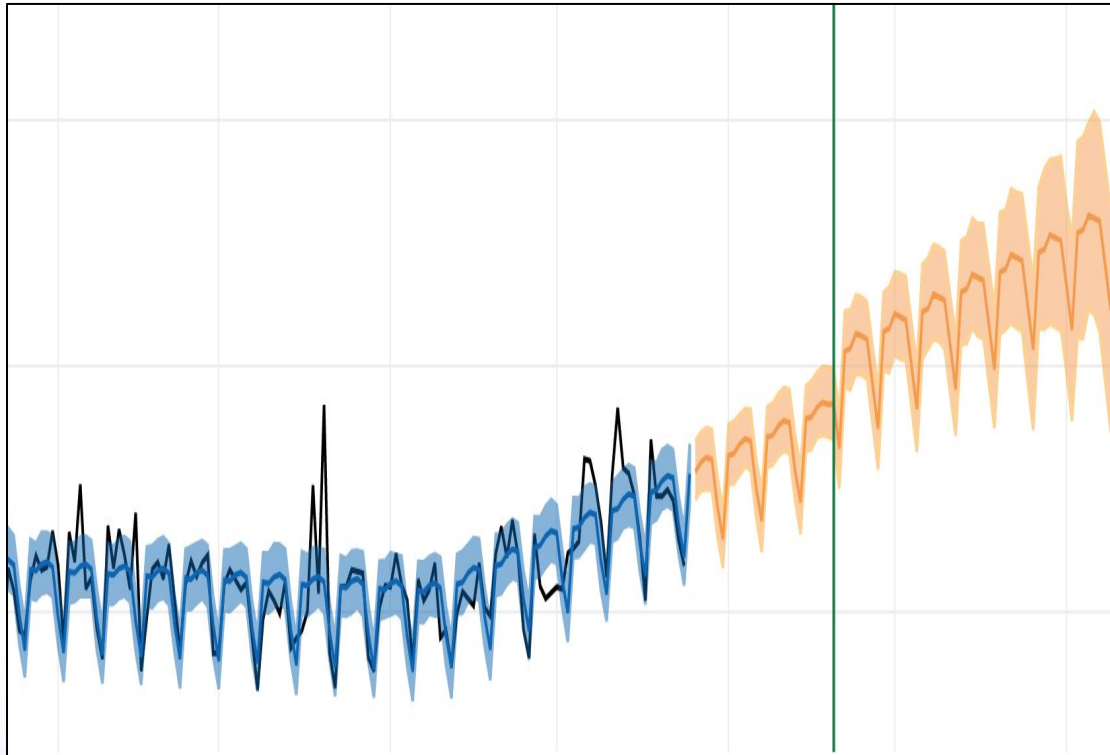
Significant impact
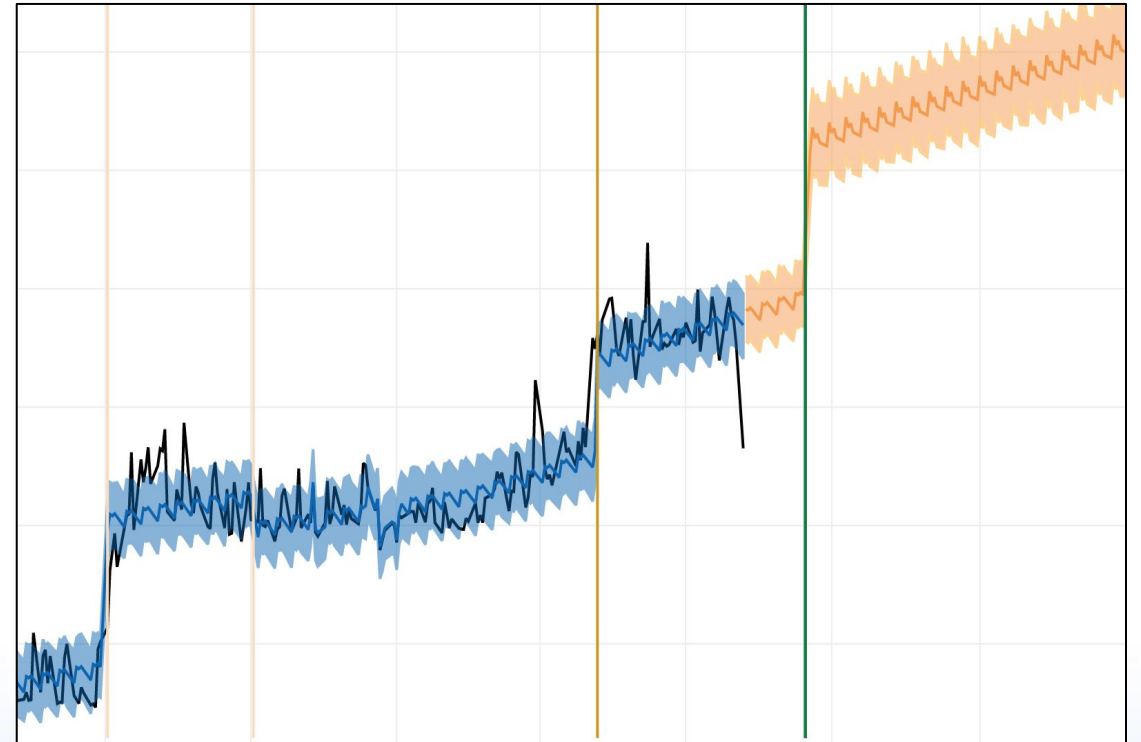
# How would different events compound the impacts?

Inputs: A list of future events + event dates/periods
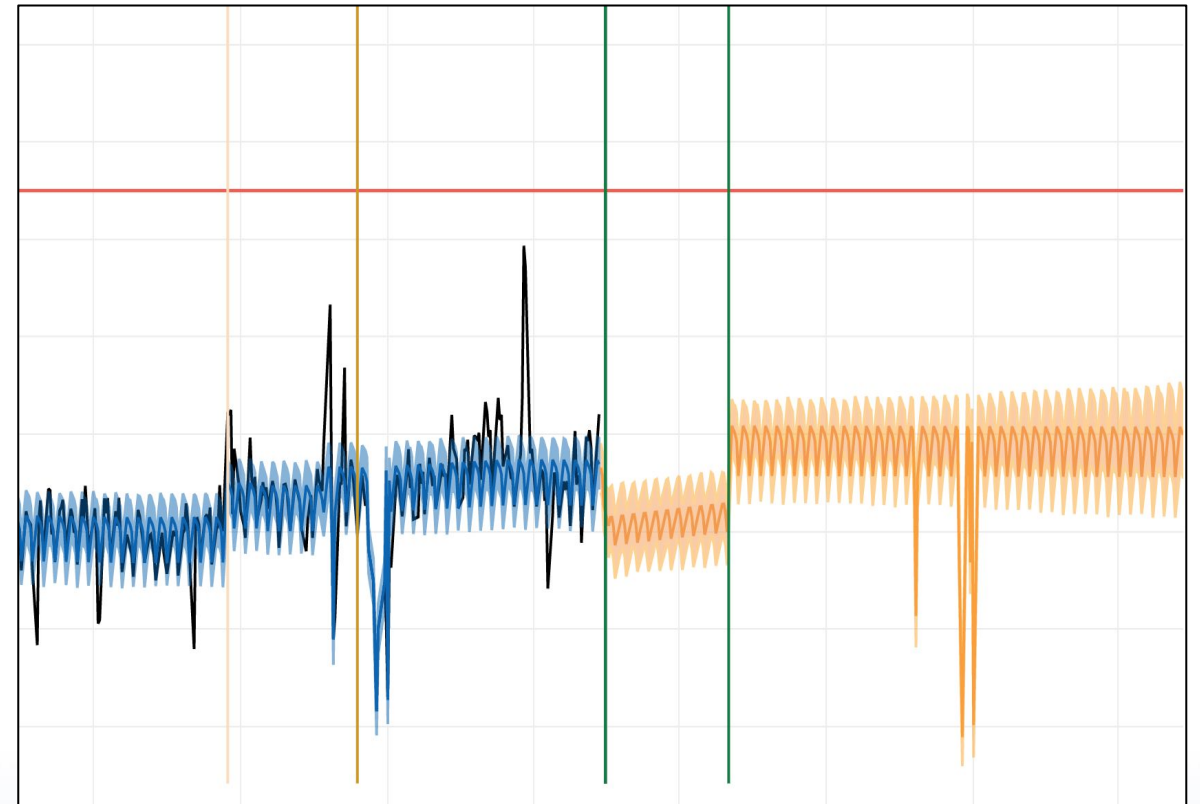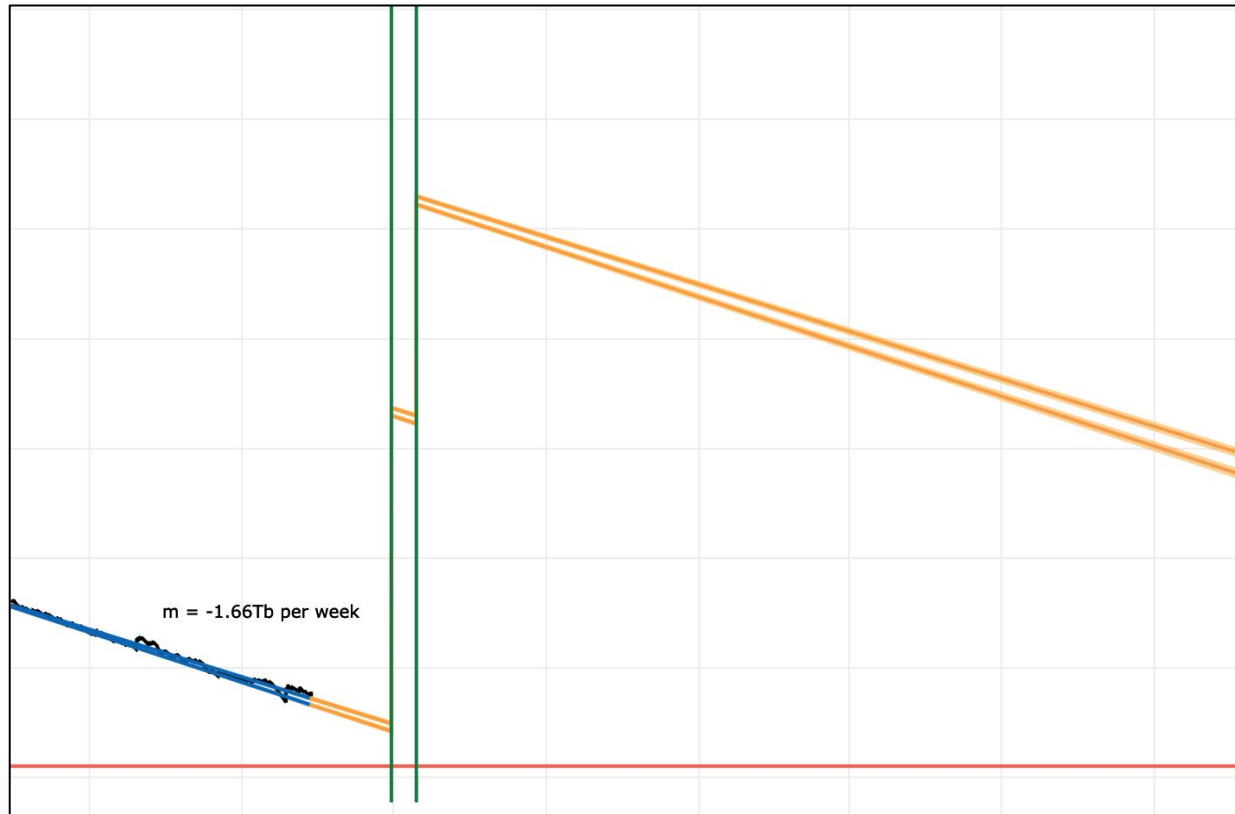
# Lessons learned - The role of human in the loop

**Goal**

what's the goal of stakeholders?

- communicate the uncertainty
- interaction to facilitate the feedback loop

- decision-making process
- collaboration model
- sources of errors

**Data & Algorithms**

How can we augment with human inputs?

**Uncertainty**

Where does the uncertainty come from?

- known unknowns vs. unknown unknown
- data collection process
- adjust for the known unknowns if possible

# Business Impacts

- It provides the **visibility** of how a critical capacity planning decision is being made as a **single source of truth**

- It provides **immediate feedback** on the effectiveness of a remediation and enables efficient communication and coordination with other business partners

- It boosts the **agility** in order to adapt to constant changes in the human-driven planning process based on more accurate and realistic forecasts

- It helps address the **scalability** challenge by synthesizing the most up-to-date but sometimes scattered information about future events

# Future Work

- Refine accuracy measurement methodology of planned forecasts

- Incorporate more causal inference/ML models for other future events with dynamic impacts on time series forecasting

- Enhance more real-time on-demand forecasting

## Acknowledgement

**Daisuke Kawamoto**
Product Manager

**Anil Bansal**
Eng Dir.

**Will Forrester**
EngLead Data Science

**Chun-Ju Lin**
FE Engineer

**Swanand Pethe**
BE Engineer

**Raghuram Gururajan**
BE Engineer

**Lucila Orengo**
UX

**Joanne Sun**
BE Engineer

**Anish Kanchan**
FE Engineer

THANK YOU