PREDICTING HOSPITAL ADMISSIONS USING CLIMATE DATA

SCHOOL OF SCIENCE & TECHNOLOGY

AN AI-POWERED PUBLIC HEALTH MODELING APPROACH

Group 3

CLARA SOBEJANO - LOUIS-ESMEL KODO - LUCAS IHNEN

MASSIMO TASSINARI - NITIN JANGIR - YOUSEF MOHAMMAD

Table of Contents

Turn gridMET Data into Business Insights 03
From Climate to Care 04
Influenza-Related Hospital admissions 05
Data Sources: GridMET + FluView 06
Modular & Reusable Extraction Pipeline 07
Exploratory Data Analysis 08
Feature Engineering 10
Model Evolution 11
Model Selection 12
Impact & Market Potential 13
Scalability 14
Pricing & ROI 15
Limitations & Future Improvements 16
Next Steps 17

THE CHALLENGE

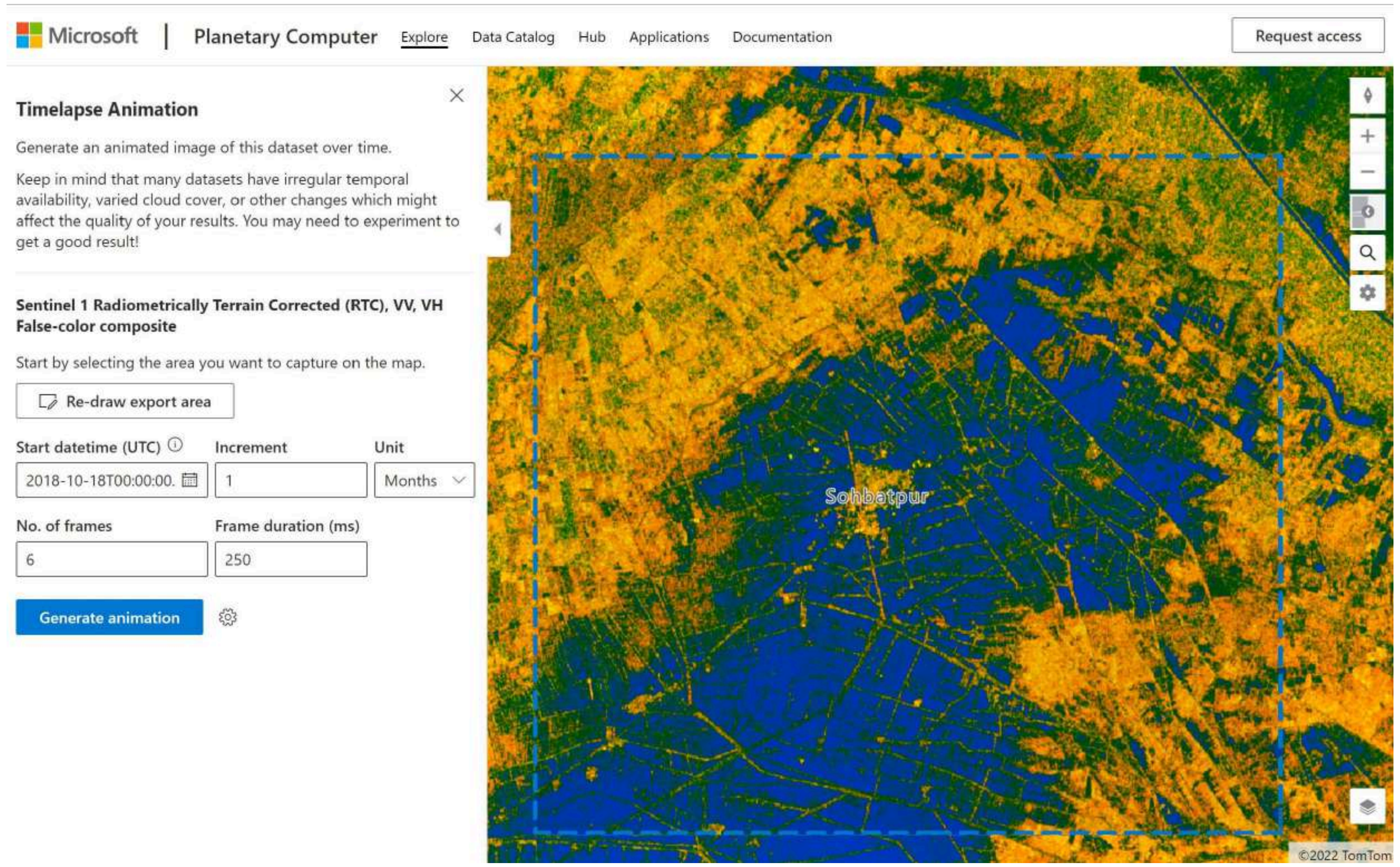
Turn gridMET Data into Business Insights

Why GridmET

- High resolution & historical depth
- Nationwide coverage for wide-scale applications
- Rich meteorological signals
- Open & replicable

Potential Application Domains

- Climate Risk Modeling
- Wildfire Prediction
- Agriculture Advisory
- Renewable Energy Planning
- Hospital Admissions Forecasting



CONTEXT

FROM CLIMATE TO **CARE**

Climate-sensitive health events are rising: heatwaves, wildfires, poor air quality.

Hospitals face **unpredictable** spikes in demand.

Accurate, early forecasting = better staffing, resource allocation, reduced costs.

We aim to leverage U.S. climate data to **forecast flu-related hospitalizations.**

BUSINESS USE CASE

Influenza-Related **Hospital admissions**

Why Hospital Admissions ?

- Unanticipated patient surges
- Emergency overcrowding
- Staffing & supply disruptions
- Can cost large hospitals **$2-4M annually in efficiency loss**

Why Influenza ?

- ***High impact and frequency:*** Up to 710,000 hospitalizations/year in the U.S.
- ***Strong weather link:*** Flu activity correlates with temperature, humidity, and seasonality.
- ***Rich & reliable data:*** Weekly CDC reports available at national & state levels.
- ***Operational relevance:*** Enables resource planning and cost reduction in hospitals.
- ***Scalability:*** Framework can later be adapted to asthma, heat strokes, pollution-related illnesses...

Target Users

- Public health agencies (CDC, state health departments)
- Hospital networks, insurers
- Climate & risk forecasting startups

Return on Investment

- Operational efficiency during outbreaks
- Reduced costs from avoidable over/under-staffing
- Scalable architecture adaptable across states

DATA SOURCES

GridMET + FluView

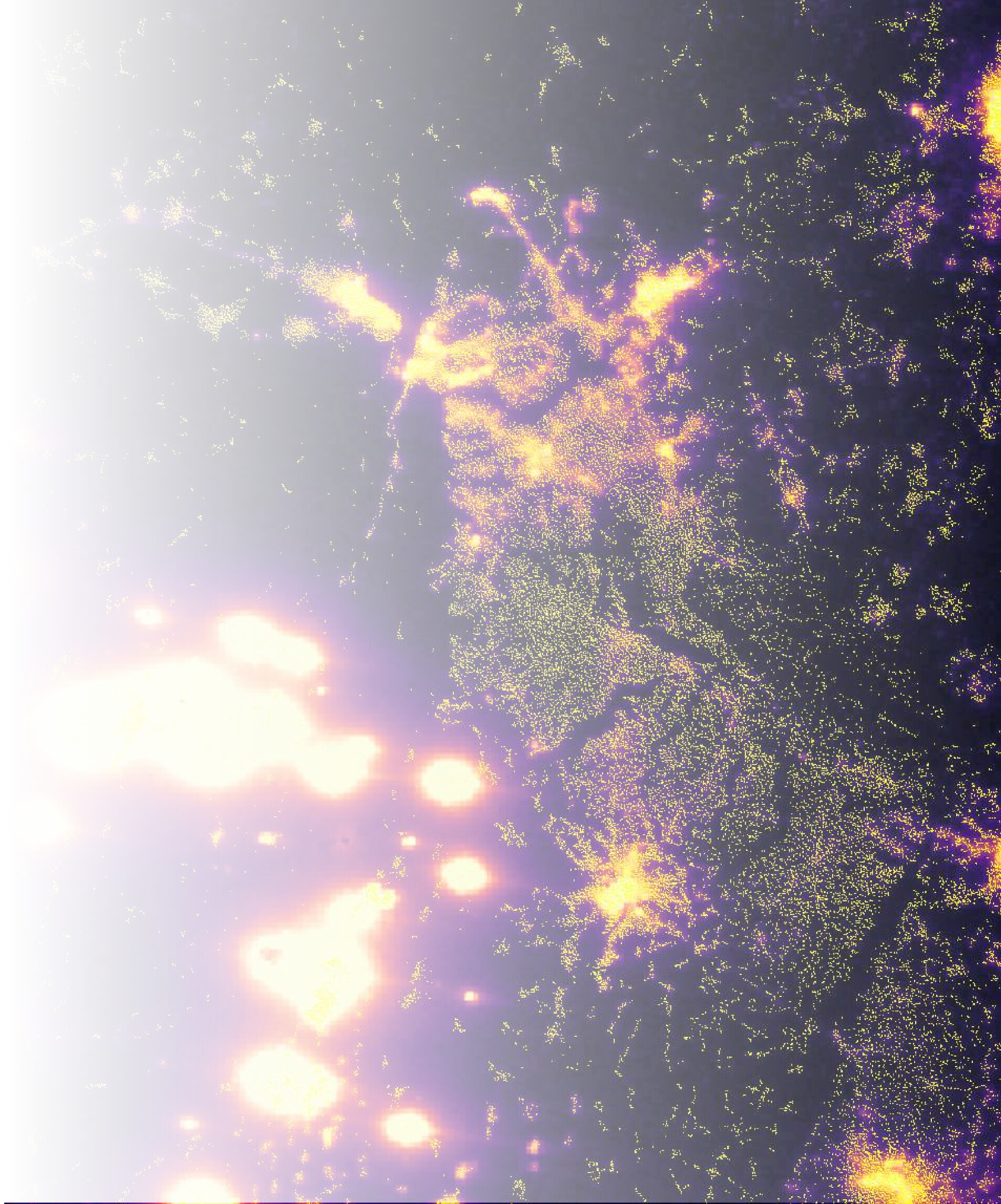
GridMET Data

- 40 years of daily U.S. weather, 4 km grid
- 10+ variables (temp, humidity, precip, wind, fire danger...)
- > 12 billion records in daily granularity and (lat, long) pairs

FluView Data

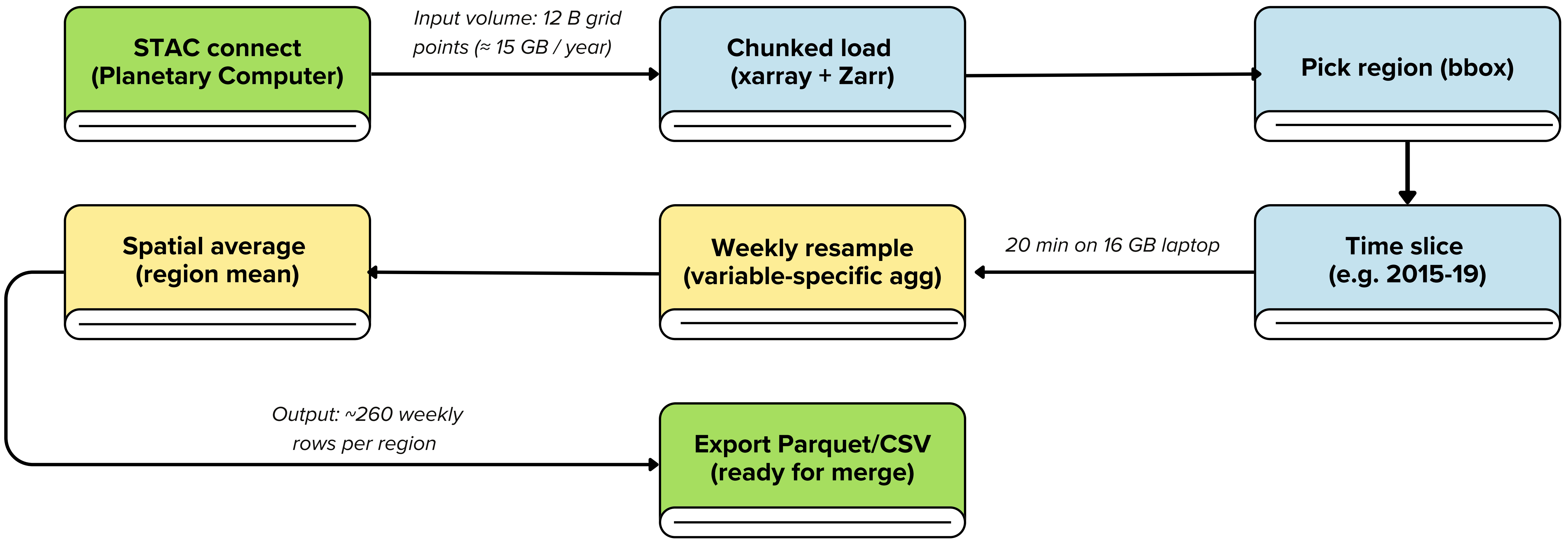
- Coverage back to 2009, refreshed in near-real time
- City & State level laboratory confirmed weekly influenza hospitalisations
- Age group, Virus Type, Race & Gender granularity
- Gold-standard source for national flu burden estimates

Filter gridMET to hospital catchment box
Resample daily → weekly
Join on Monday-aligned weeks



DATA EXTRACTION

Modular & Reusable Extraction Pipeline



Tool & Tech Stack

DEVELOPMENT

Exploratory Data Analysis

Seasonality

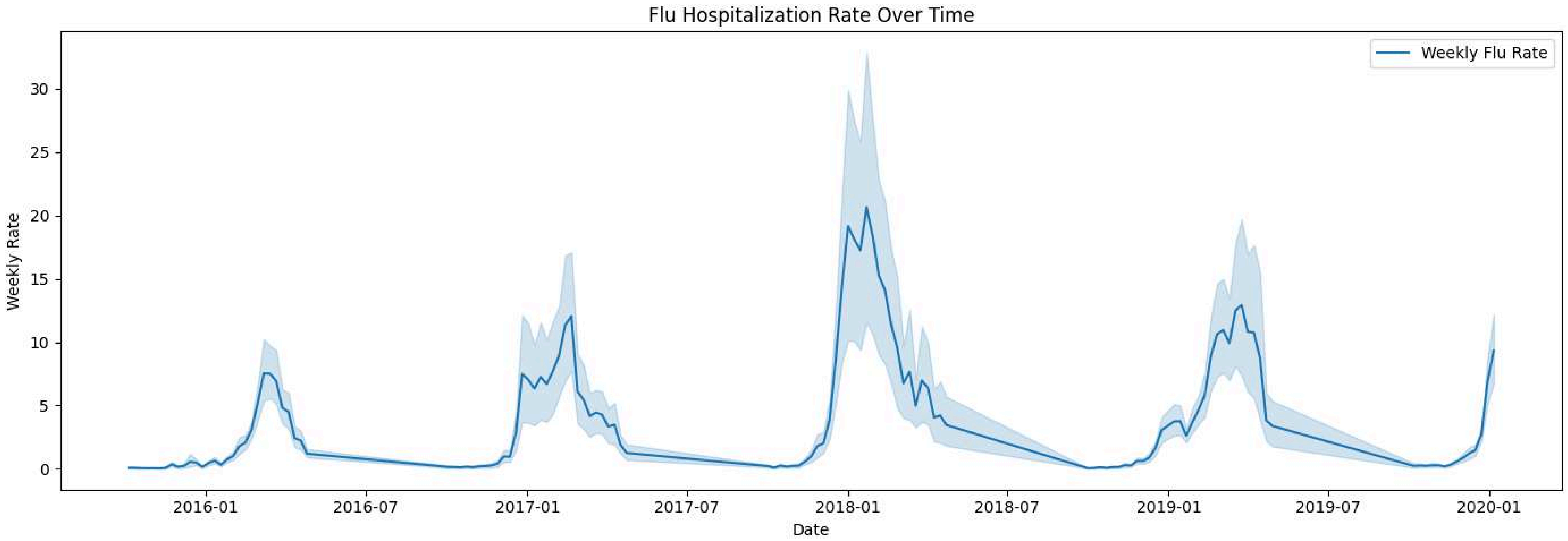
Flu hospitalizations peak Dec–Mar. Some years (e.g., 2018) spike much higher.

Climate Trends

Temperature, humidity, and wind show seasonal cycles but weak direct correlation to flu rates.

Lag Effects

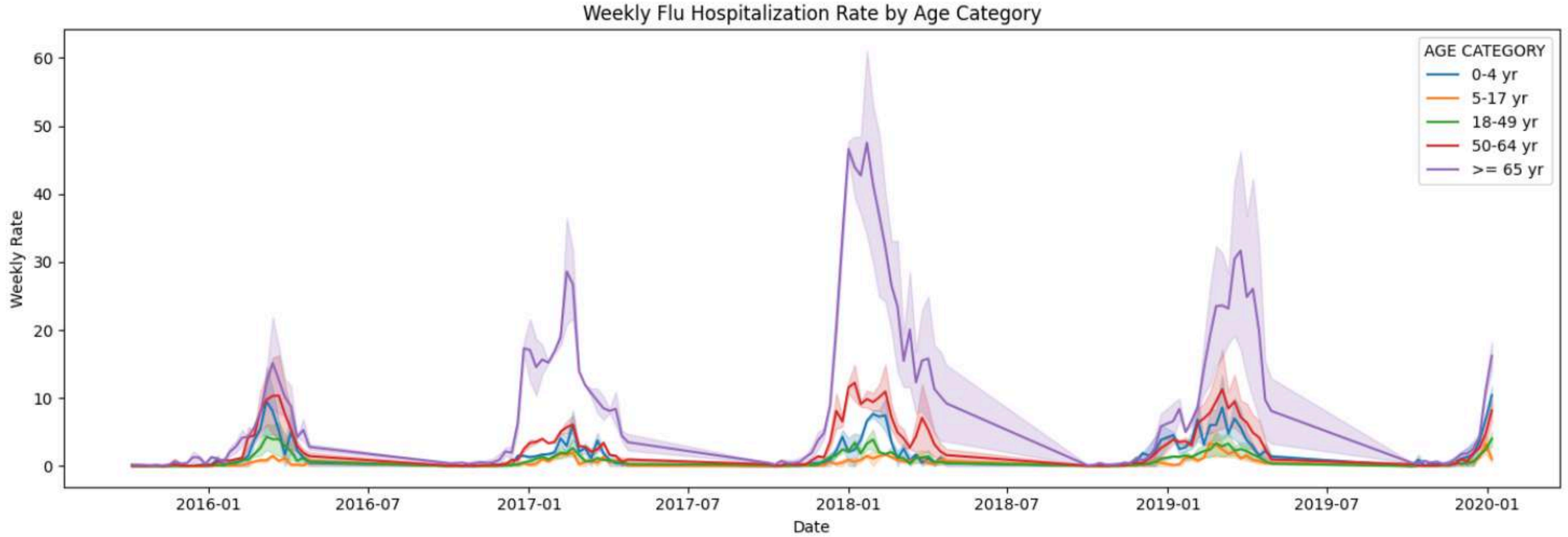
Delayed impact of weather → lag features needed for modeling.



Exploratory Data Analysis

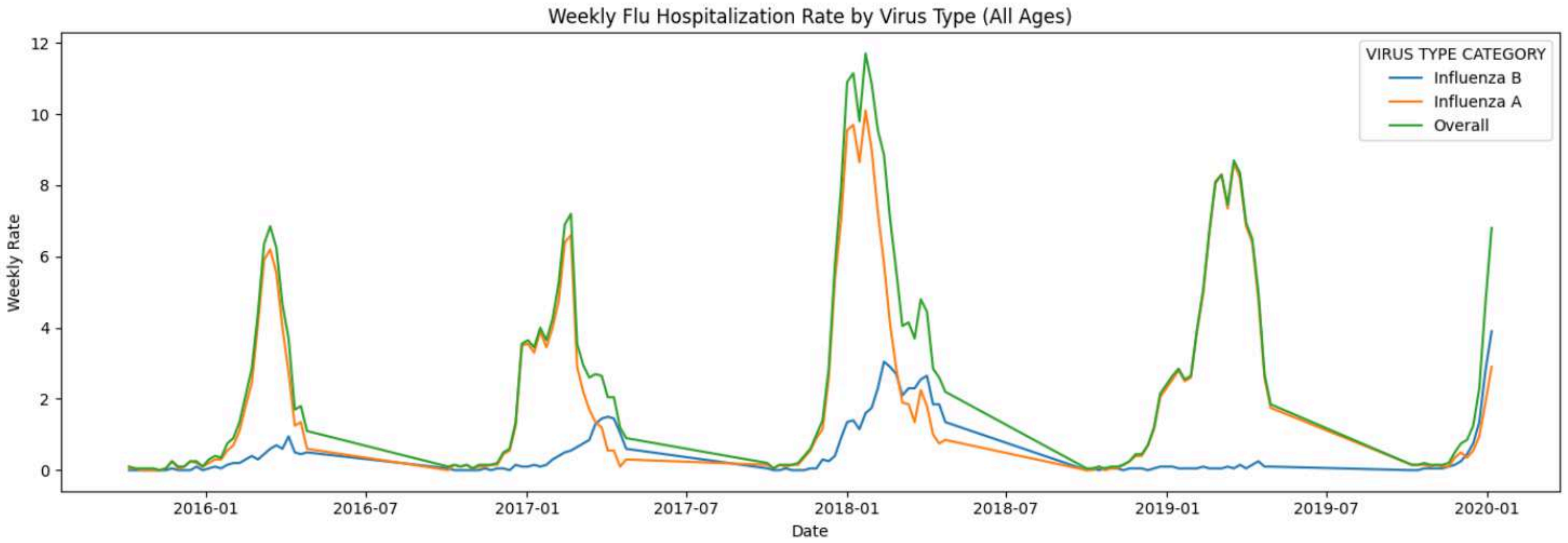
Demographics

Elderly (65+) and infants are most affected ➞ priority for forecasting and planning.



Virus Types

Influenza A dominates hospital burden: ~4x higher than B.



DEVELOPMENT

Feature Engineering

Lag Features

Past climate data (1–8 weeks) to capture delayed weather effects.

Rolling Statistics

2–8 week moving averages, variability, min/max values to capture short-term trends and volatility in weather patterns.

Interaction Terms

Combined variables (e.g., humidity × temperature) to reflect compound effects.

Seasonality Features

Engineered features to help the model recognize recurring seasonal patterns in flu activity.

Lagged Target Variables

Added flu rates from the previous 1–3 weeks to provide short-term historical context.

DEVELOPMENT

Model Evolution

Phase 1: Baseline Modeling (v1.0 - v2.0)	**Phase 2: Feature Selection + Tuning (v3.0 - v4.1)**	**Phase 3: Hybrid Residual Modeling (v.5.0 - v5.2)**	**Phase 4: Simplification & Optimization (v6.0 - v6.2)**
Goal: Establish baseline performance and test initial feature strategies	Goal: Improve accuracy and reduce overfitting	Goal: Enhance short-term prediction by correcting residual errors	Goal: Maximize performance with a simpler, more robust model
Key Changes: Used full feature set, then tested autoregressive and seasonal signals	Key Changes: Introduced hyperparameter tuning and selected top features	Key Changes: Modeled deviations from a lag-based baseline; added model explainability	Key Changes: Focused on top features, pruned redundancy, added reproducibility
Performance: RMSE from 2.43 → 1.73	Performance: RMSE improved to 1.55	Performance: RMSE improved to 1.37, R^2: 0.799	Performance: Final RMSE: 1.35, R^2: 0.821

DEVELOPMENT

Model Selection

We tested 10+ model versions to improve accuracy and efficiency:

- Started with basic tree models (v1.0)
- Gradually added tuning, feature selection, and hybrid methods
- Final model (v6.2) uses residual correction + 10 key features + automated tuning

Model Version	RMSE	R^2
v2.0 (Baseline)	2.43	0.467
v6.2 (Final)	1.35	0.821

Best predictive performance

Minimal feature set (efficient & scalable)

Tuning via Optuna

Interpretable hybrid design

45% error reduction, 75% variance explained

Impact & Market Potential

Hybrid model reduces flu-related hospital admissions by **44.9%, saving $125K per hospital annually.**

Results

Baseline Model
RMSE = 2.43
R^2 = 0.467

Hybrid Model
RMSE = 1.34
R^2 = 0.823

36-point R^2 gain & 45% fewer errors

Financial Impact

- 120 fewer admissions/year
- $12,500 avg. cost per admission
- 12 hospitals in Rochester → $1.5M saved
- $125K saved per hospital annually

Market Scalability

- **5,000+ hospitals** in U.S.
- Just **1% market** = **50 clients/year**
- Path to scale via **partnerships with health systems, public health agencies**, **insurers**.

Scalability

Scalable Deployment

Pipeline works at **city & state level**
(Rochester → California)

Generalizes well:
R^2 = 0.93 in CA using same architecture

Reusable across geographies
→ minimal adaptation needed

"Train once, deploy many" model with GRIDMET & health data

Product Strategy

Offered as a **Model-as-a-Service (MaaS)**

Delivered via **secure API** for hospitals, insurers, agencies

Designed for **data privacy & compliance**

Pricing & ROI

Product Tier	Product Offering	No. of Hospitals	Annual Price	Target Customer	Product ROI*
Pro	Forecast API & Dashboards	1	$100,000	Medium hospitals, regional clinics	**125%**
Enterprise	Multi-state forecasting, limited fine-tuning	5	$450,000	Large health systems, insurers, agencies	**139%**
Custom	Custom models and fine tuning for different diseases	Unlimited	$2,000,000	Local Government agencies, national insurers	**188%** (for 30 hospitals)

* Product ROI calculated using the potential savings explained previously

FUTURE

Limitations & Future improvements

1

Equal weighting overrepresents unpopulated areas.

Improvement:
Population-weighting boosts accuracy, with more compute.

2

Limited compute forces basic, rectangular regions.

Improvement:
Cloud boosts resolution, flexibility, and speed.

3

Aggregated, low-frequency health data limits precision.

Improvement:
Use ZIP-level, weekly data (e.g., HCUP) to improve predictions.

4

GRIDMET data stops at 2020, missing recent climate trends.

Improvement:
Update or integrate newer data to reflect current conditions.

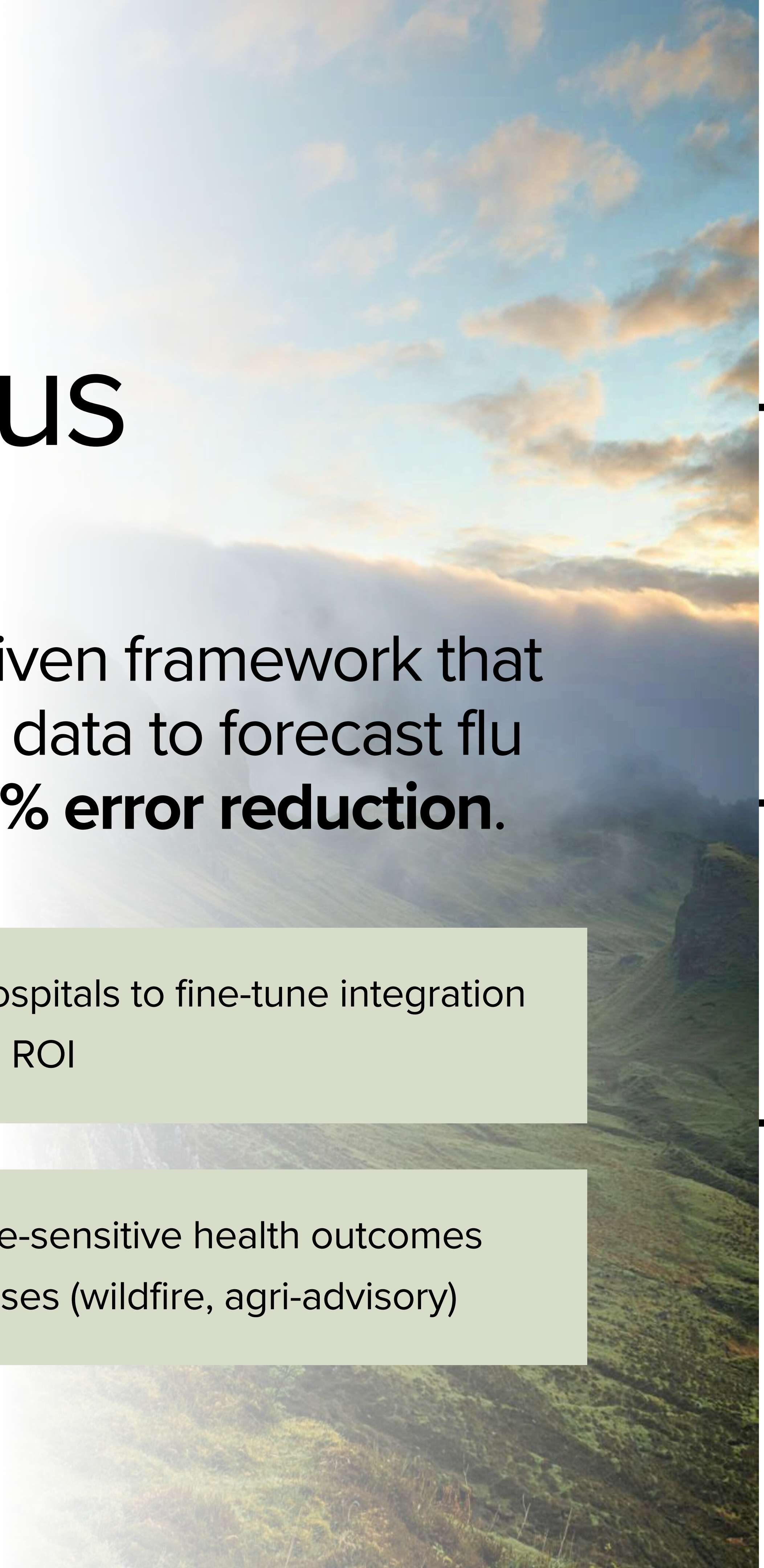
5

Simple models miss spikes and key factors.

Improvement:
Use advanced models and tuning for sharper forecasts.

CONCLUSION

Next Steps for us



We’ve built a scalable, data-driven framework that uses climate and surveillance data to forecast flu hospital admissions with **>44% error reduction**.

Pilot Deployment: Launch with 3–5 regional hospitals to fine-tune integration and demonstrate live ROI

Expand Applications: Extend to other climate-sensitive health outcomes (heatstroke, asthma) and non-health use cases (wildfire, agri-advisory)

Proven Impact

Hybrid model achieves RMSE 1.34 (vs. 2.43 baseline) and **R^2 0.82**, translating to ≈ 120 fewer admissions/year in a mid-sized city (~$1.5 M saved).

Scalability

Validated in Rochester and California, pipeline runs end-to-end on any region with minimal adaptation.

Comercial Readiness

Modular architecture ➔ **Model-as-a-Service**; subscription pricing delivers **125–188% ROI per hospital**.

Thank You !