



# Establishing, Growing Use of, Trusting, and Increasing Value from AI/ML-based Weather Decision Support Services and Solutions

John Celenza, Weather Team Lead

31 October 2024 🤖



# Zipline: 2 Platforms

- P1 - Fixed wing aircraft, 3 meter wingspan
- P2 - a 5 prop e-vtol “docking” drone with a droid

# Platform 1



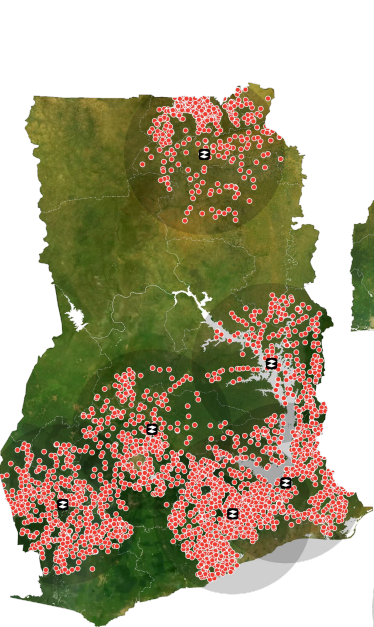
# CI-1 Zipline Daloa at Nightfall.



# Zipline launched with Walmart in 2021

Pea Ridge, Arkansas  
AR-1 (Colocated DC)  
Launched July 2021

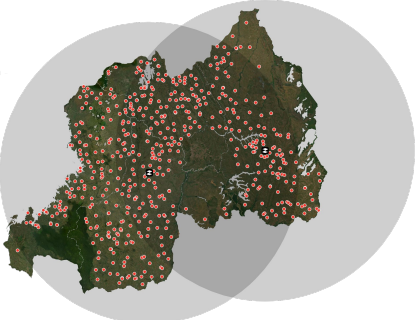
# Where We Fly



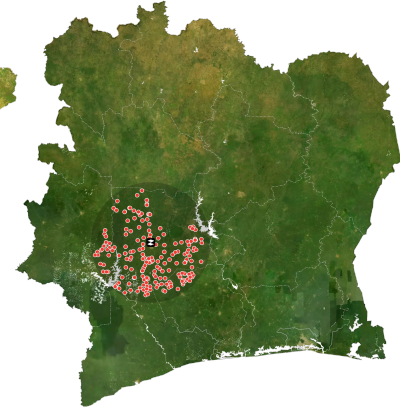
Ghana



Nigeria



Rwanda



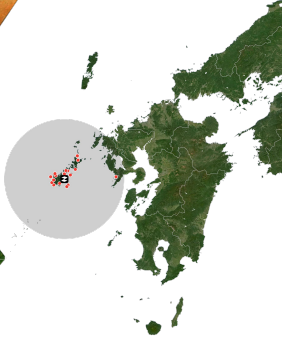
Côte d'Ivoire



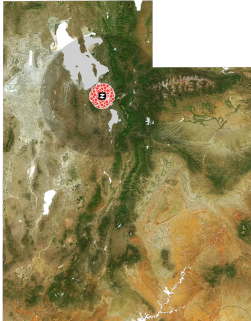
North Carolina  
Remote Ops Control Center



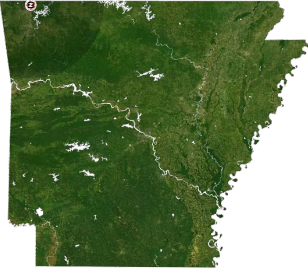
Kenya



Japan



Utah



Arkansas

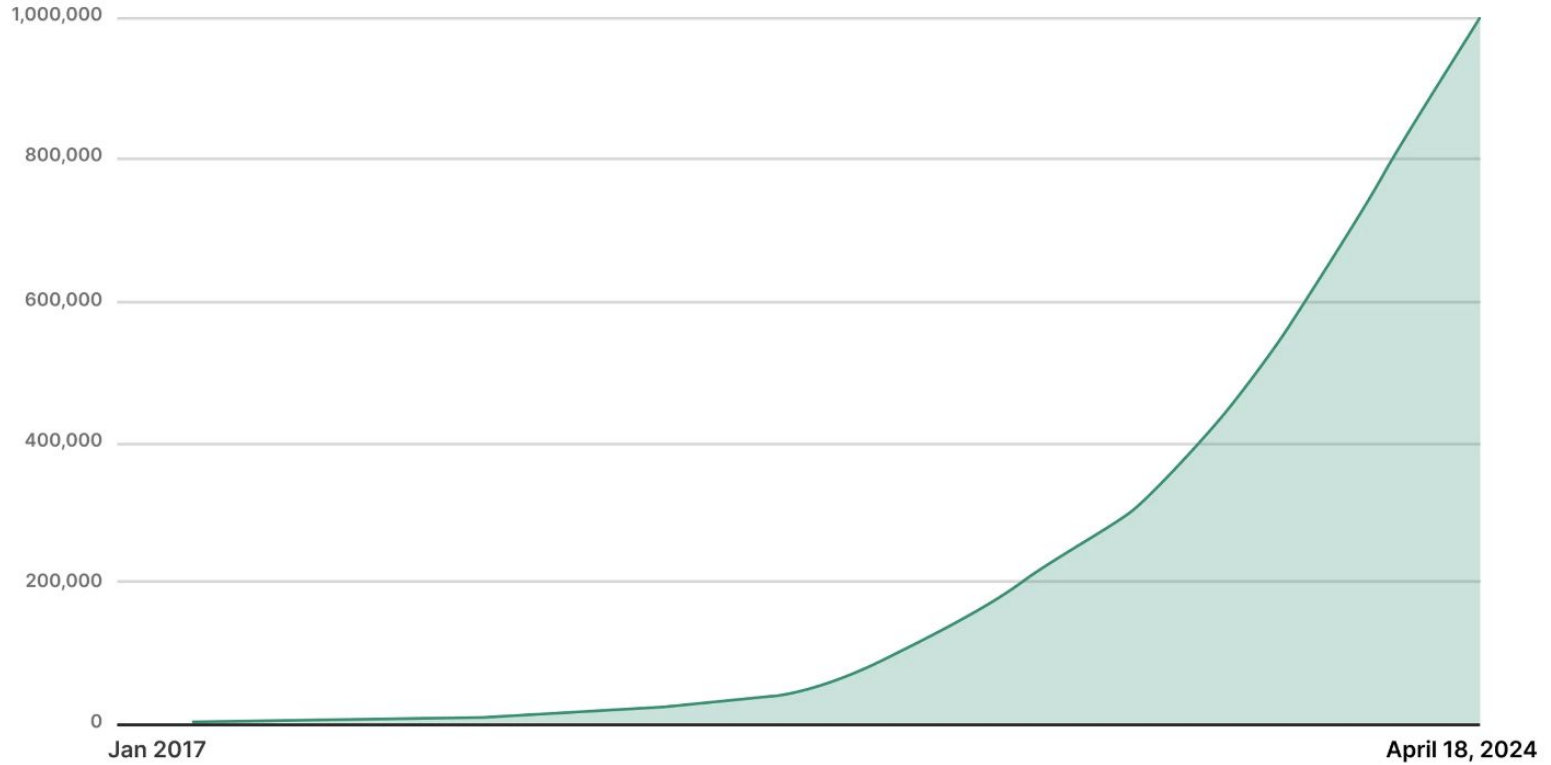


Zips fly the circumference of Earth at least **every 8 hours**

NASA's Artemis flew 2M km in 25 days  
**Zipline flew 2.8M km** in the same 25 days

Artemis burned 755,000 gallons of fuel;  
Zipline burned **none**

# Zipline's commercial drone deliveries over time





# Platform 2



# Docking stations come in different shapes and sizes

## Integrated Station(s)

- Orders loaded from inside a building
- Better integrated into workflows
- Expandable high-volume versions available (if needed)
- ~2+ orders per minute



## Instamount Station

- Orders loaded outside in a parking lot
- Faster to get started
- ~1 order per minute



# Delivery that's thoughtfully unnoticeable

Instant Delivery



Zip quietly hovers above...



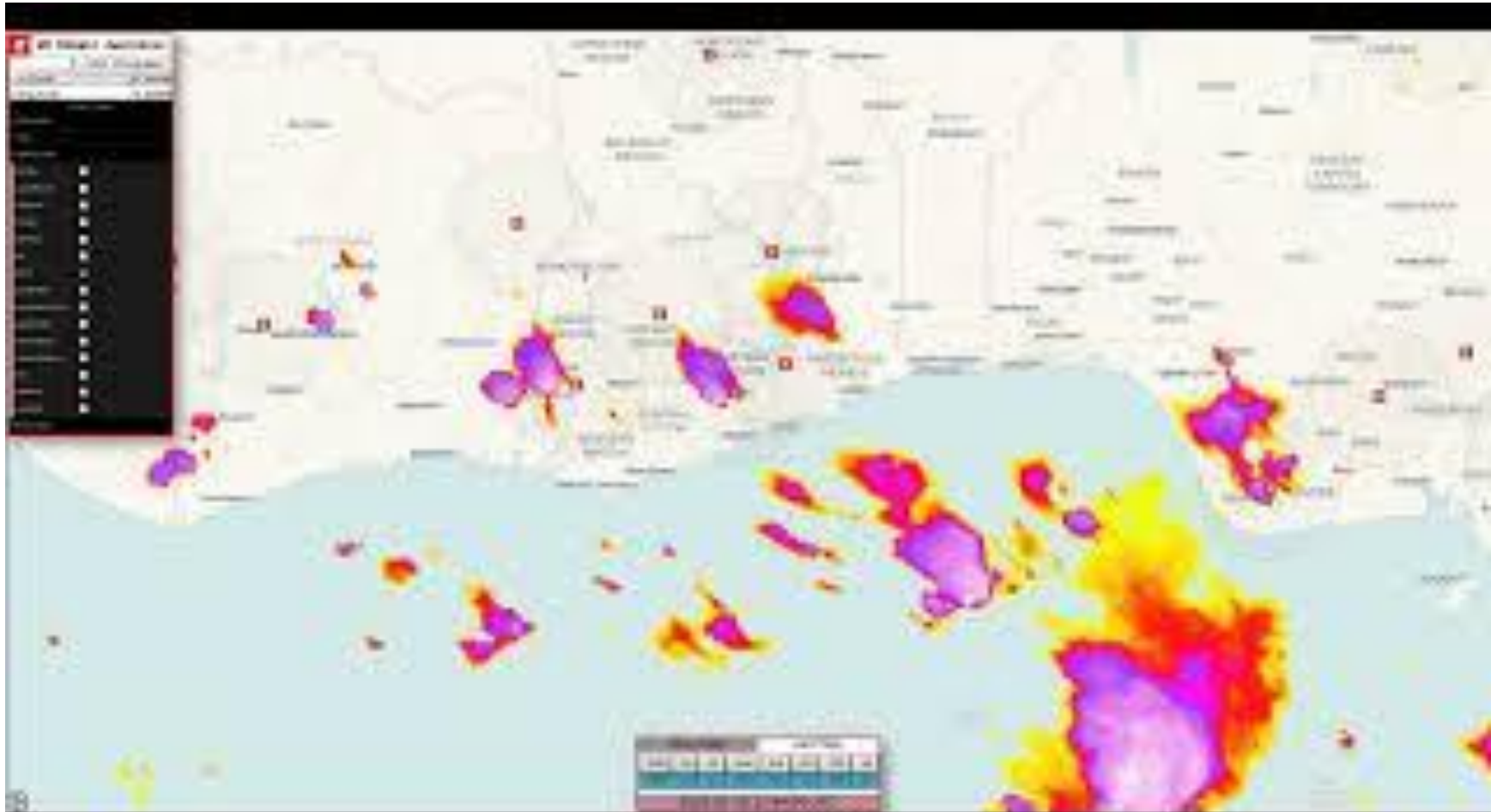
...while the delivery droid gently places deliveries in a matter of seconds



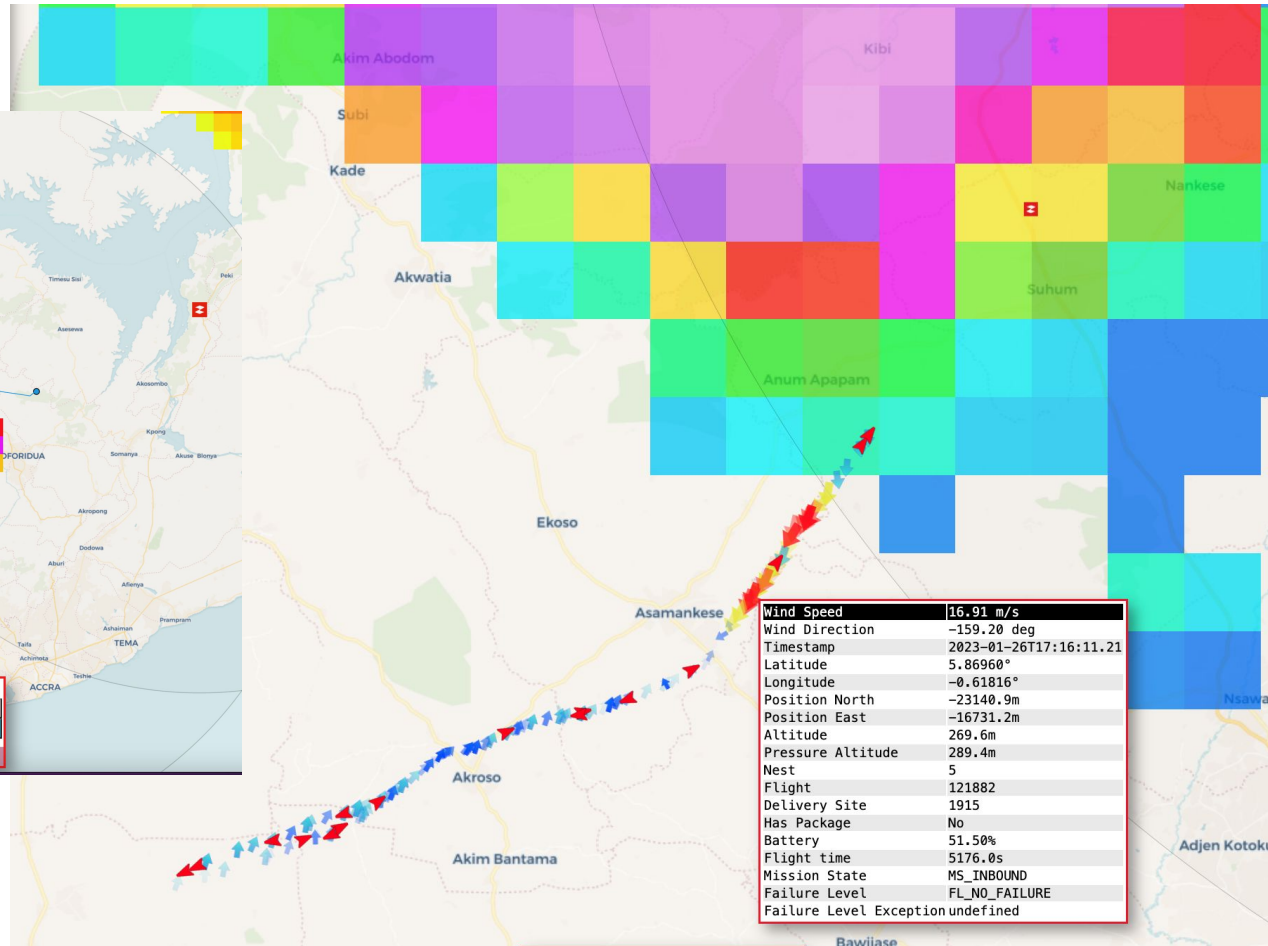
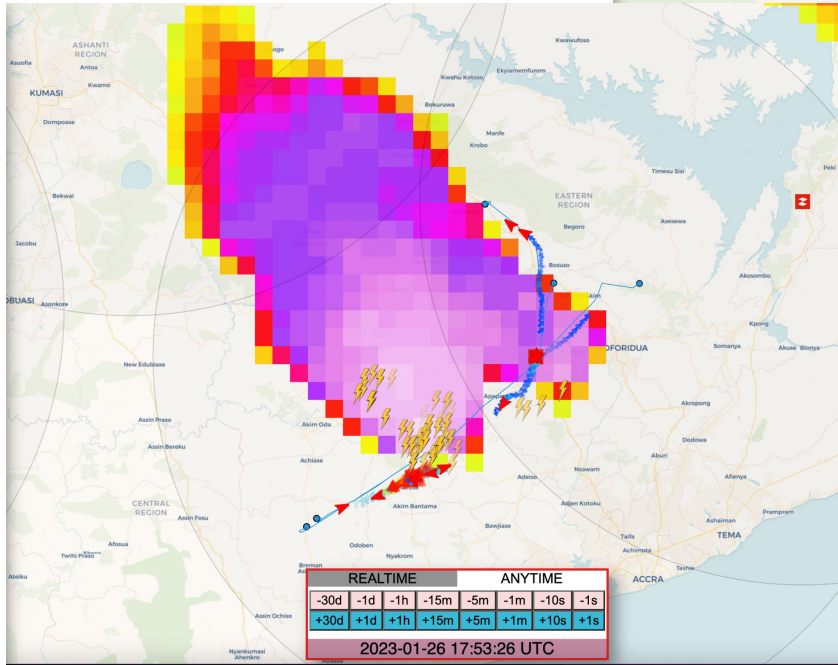


# P1 Africa: It's stormy!

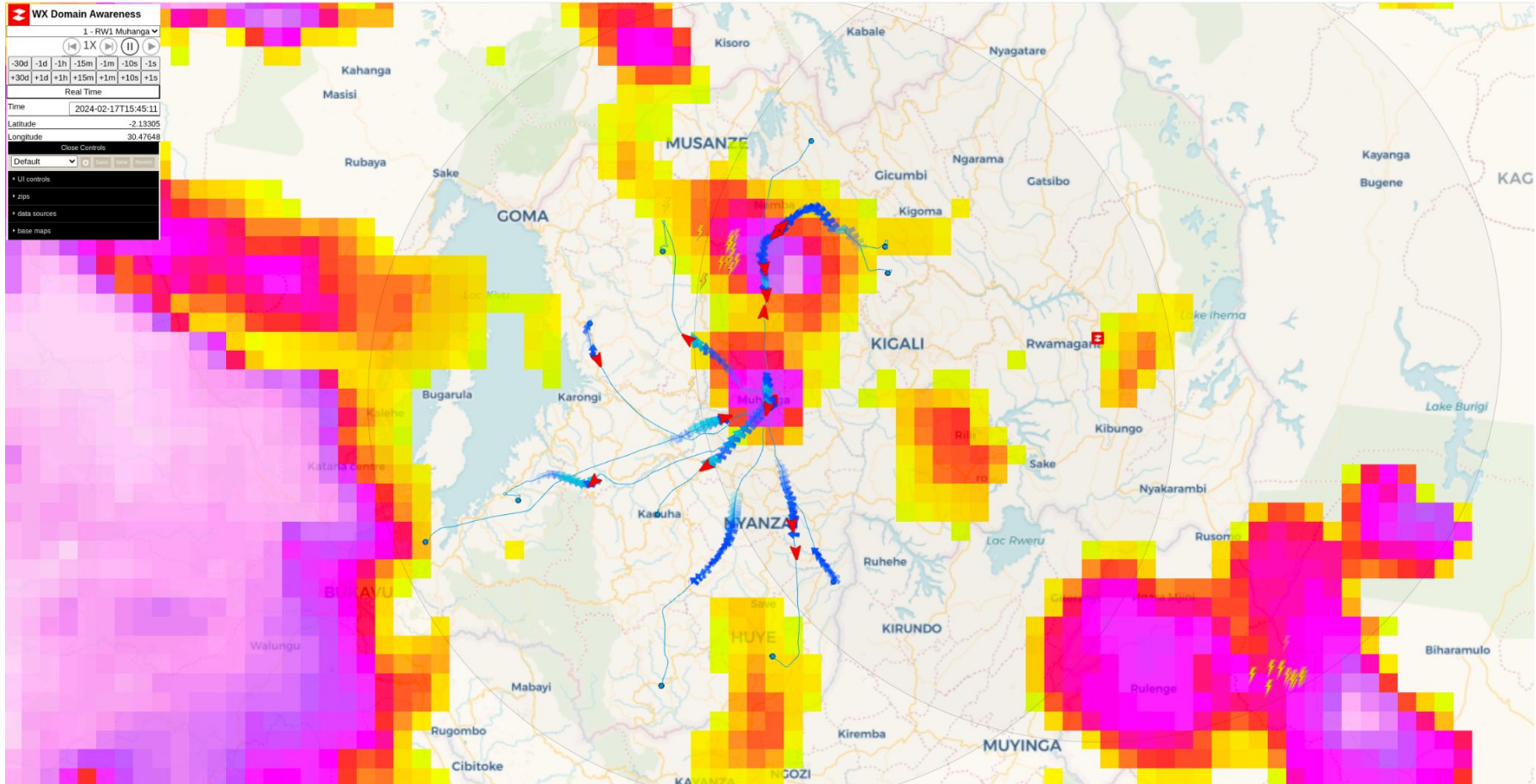
# Western Africa Severe Weather



# 2023-01-26 Ghana



# Typical Snapshot





# Avoiding Severe Wind: Let's build a Model



# What Zips Measure

**Observational Frequency** - 50 Hz

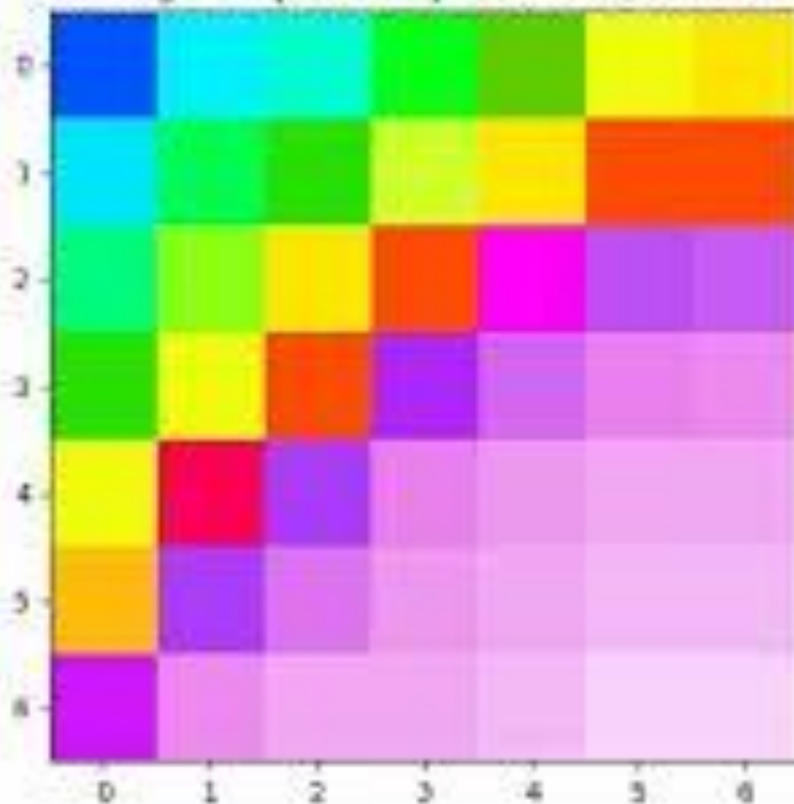
**Average Flight Duration** - <1 Hour Globally

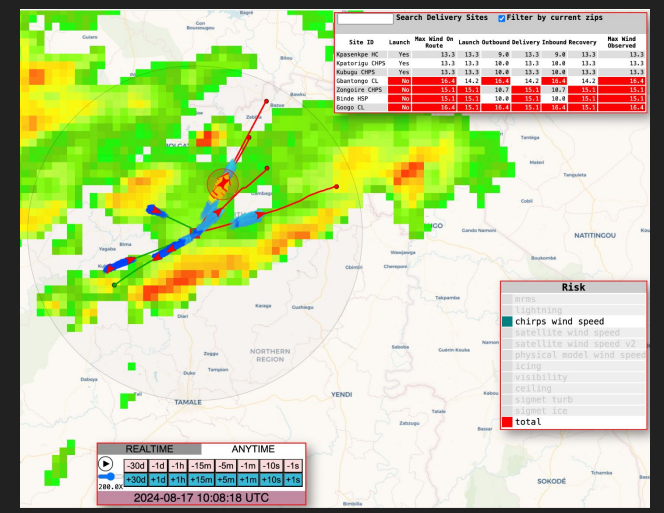
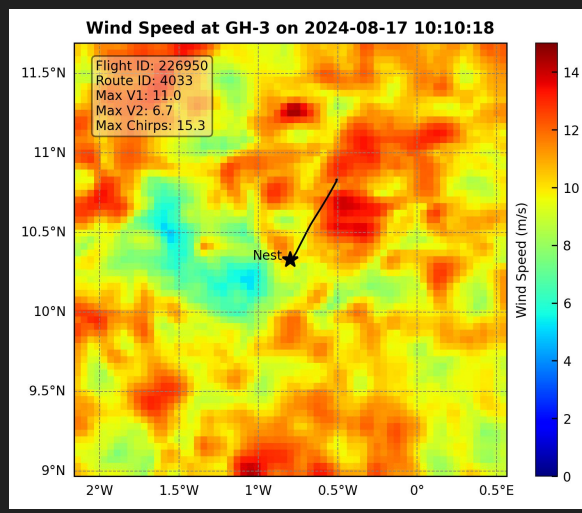
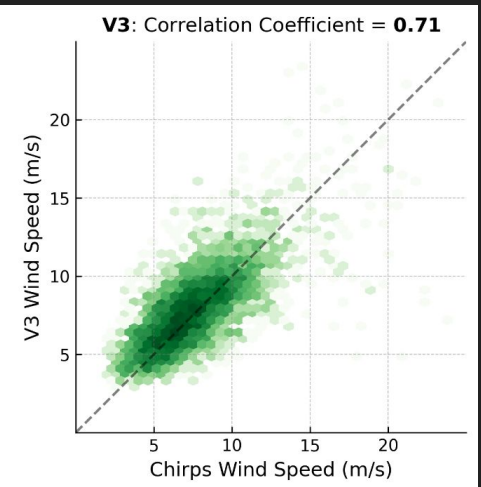
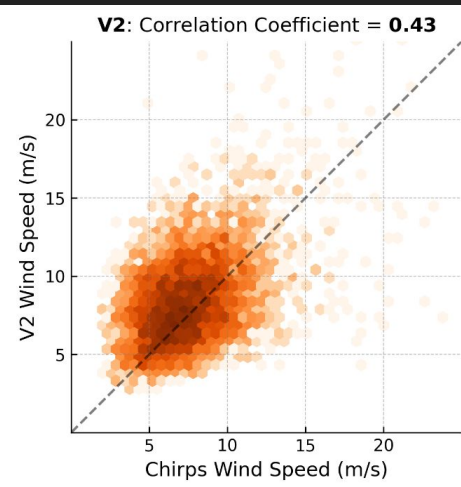
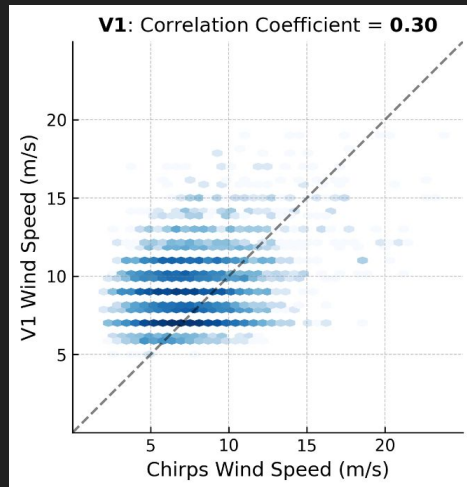
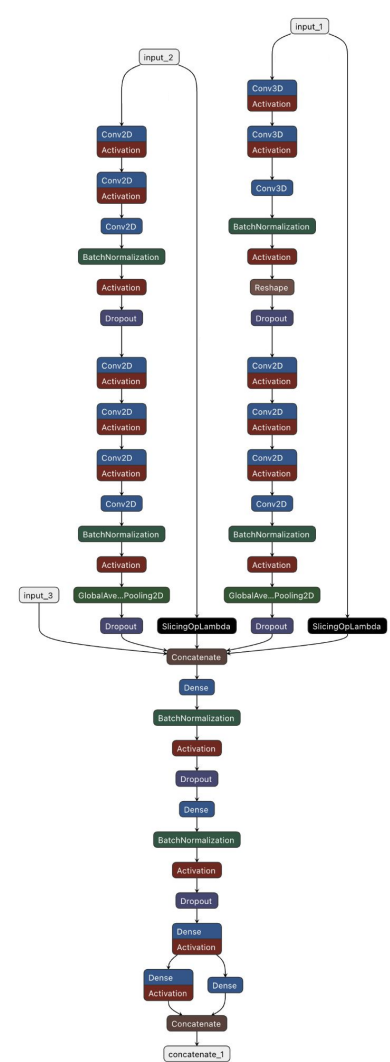
## Variables

- **Wind Component (NED)** -  $\sim \pm 1.5$  m/s
- **Temperature** -  $\sim \pm 2.5$  °C
- **Relative Humidity**
- **Static Pressure** - High accuracy, used for pressure altitude

# Let's train a model!

training example: i:54 y:2.323618173599243







# Trusting AI: How do we use this model responsibly?

# Trusting AI

## It's all about safety

- What are the lines of defense to ensure safety?
- How do we quantify risk and business impact?

# Trusting AI

## It's all about safety

### Zip Safety Through Engineering

- **Bully** the weather - be nimble and thrustful
- **Avoid** the weather - through forecasting and nowcasting
- **Safely** end flight - parachute deployment

**Being the meteorologist, I am going to talk about avoiding the weather.**

# Trusting AI

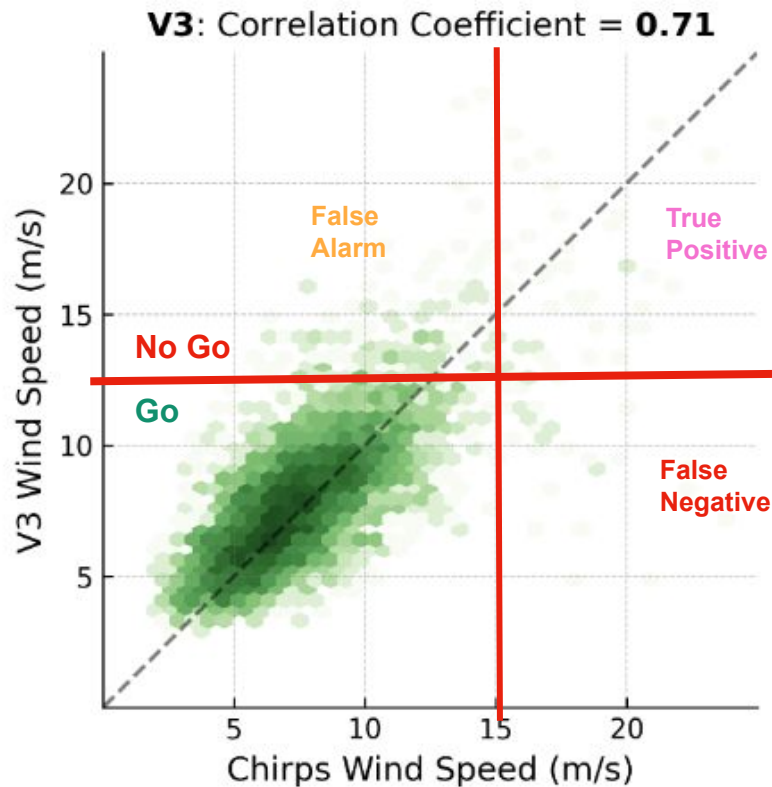
## It's all about safety - Probability of Parachute vs Delay

### The Forecasting Trade:

When we have a predictor, we need to determine a threshold that separates the good from the bad;

the nominal flights from the parachute deploying flights.

We trade safety for delay when choosing a threshold.



# Trusting AI

## It's all about safety - Probability of Parachute vs Delay

To compute **probability of parachute deployment** over a set of forecasting thresholds, we find the following probabilities.

The algorithm:

For each “forecasting threshold” - eg, 0 to 30 m/s

For every flight sample - these samples can be from real flights, or a proxy

$i$ th  $P(\text{parachute deployment}) =$

$P(\text{parachute deployment} \mid \text{wind speed in flight}) *$

$(1 - P(\text{forecast model predicted above forecasting threshold} \mid \text{flight sample time/location}))$

$P(\text{parachute deployment} \mid \text{forecasting threshold}) =$

$\text{mean}(\text{list of } i\text{th } P(\text{parachute deployment}))$



# Trusting AI

## It's all about safety - Probability of Parachute vs Delay

### Examples:

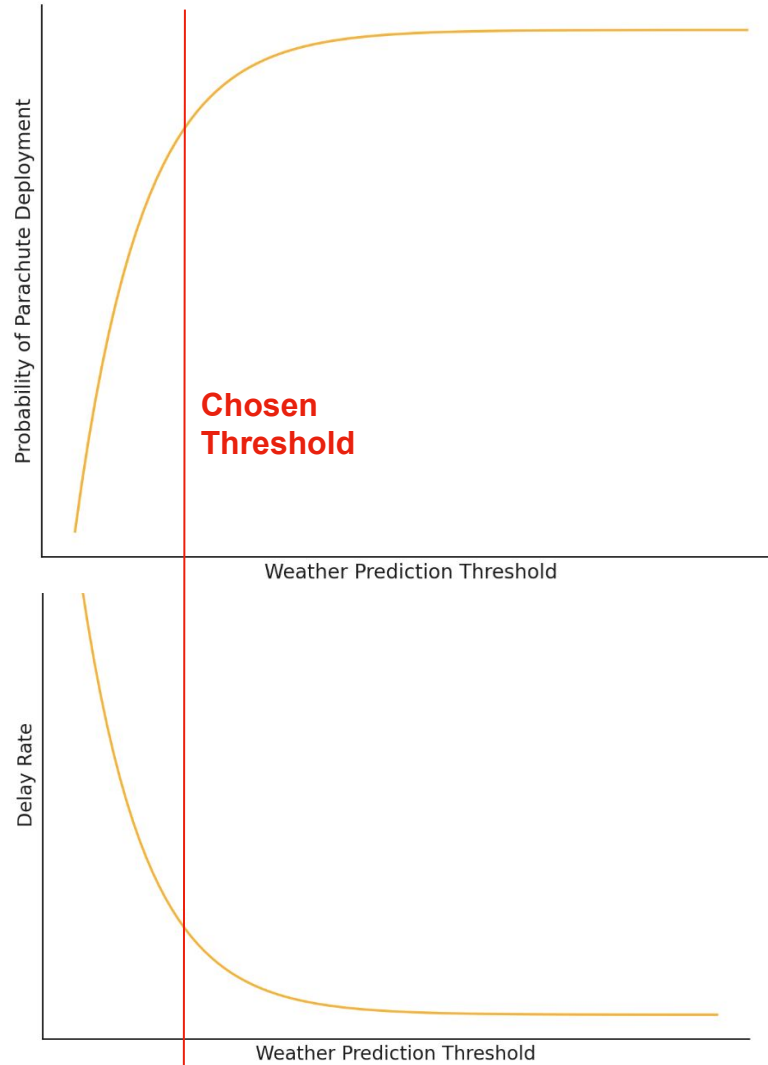
- if we threshold our wind model to **0 m/s** (no go if the wind exceeds 0 m/s), then we have 100% safety, but 100% delay (non operation)
- if we threshold to **100 m/s**, then we have the safety of **non-mitigation**, but 0% delay

The right answer could be the 2nd bullet if we've bullied the weather.  
Most of the time, the right answer is somewhere in between.

# Trusting AI

## From ML to Operations

- **Operations Team -**
  - **Maximum Uptime**
  - **Minimize Delay to customer**
  - **Operate Safely**
- **Weather Team -**
  - **Maximize Safety**
  - **Minimize Frustration for Operator**
  - **Maximize Confidence with Regulator**



# Trusting AI

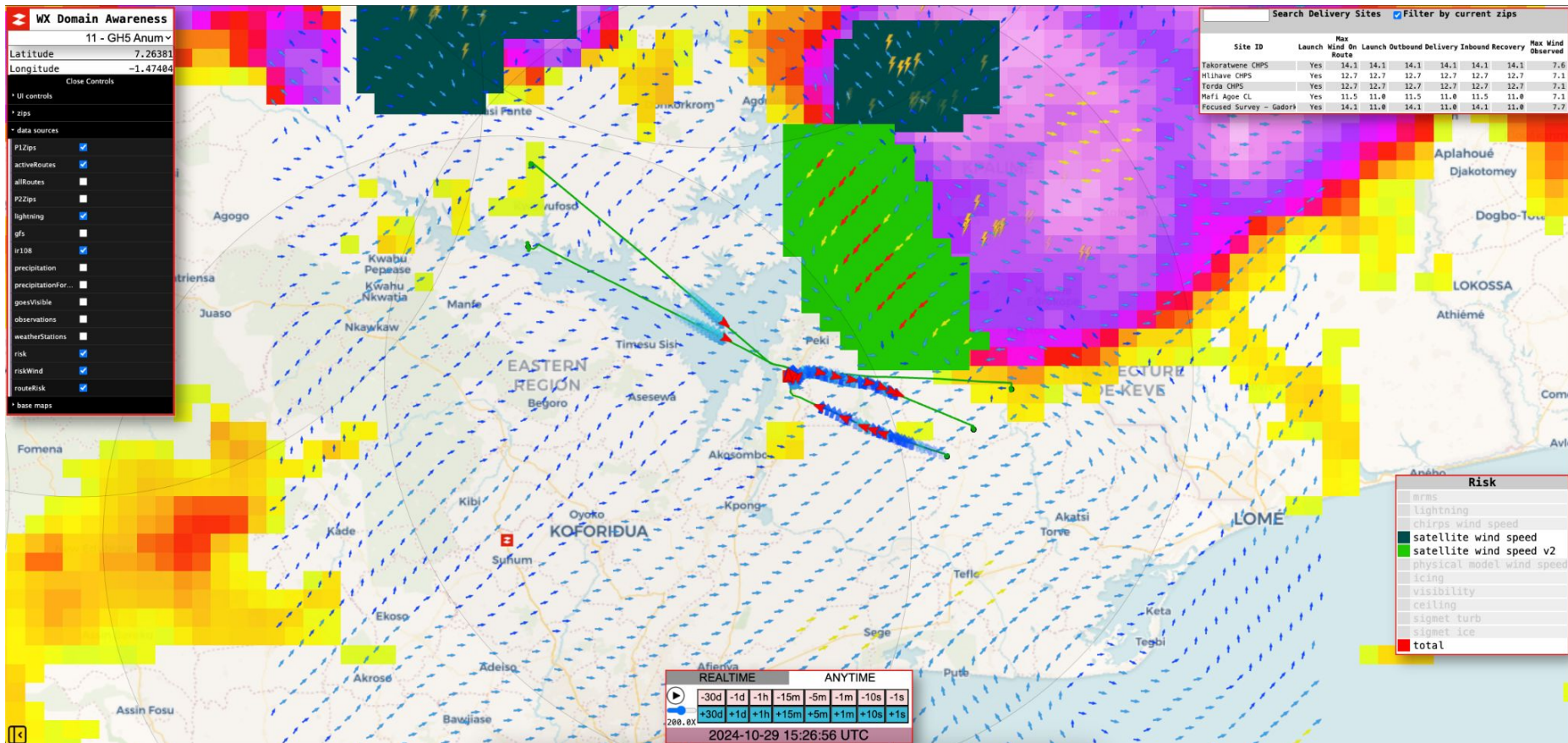
## It's all about safety - The Tried and True Way?

But do we totally trust the AI? No.

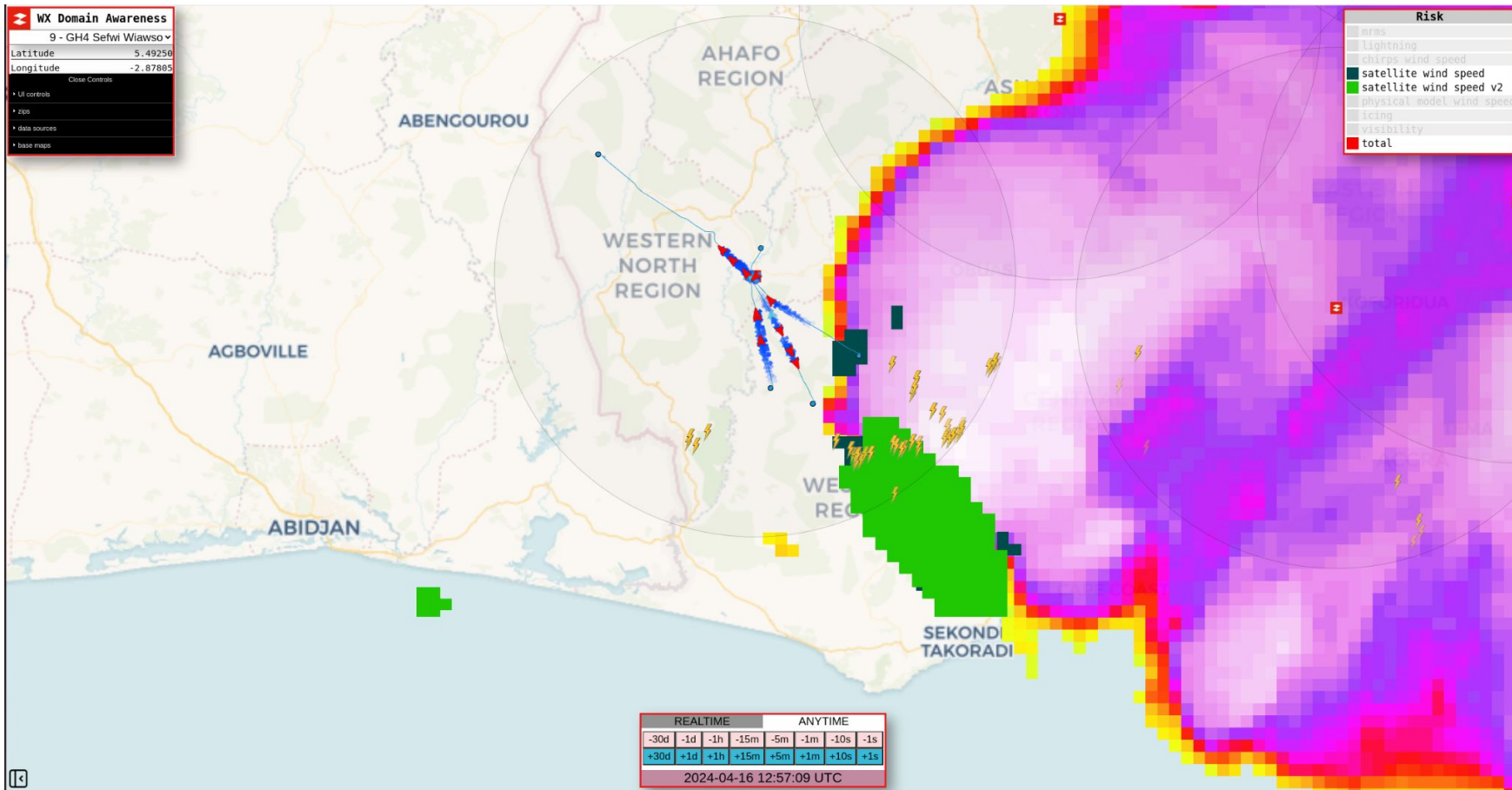
Zipline puts humans and nowcasting in the loop

- **Zips measure the wind** - true observation limits risk to other aircraft
- **Pilots in Command** - observers on the ground can make informed decisions
- **Meteorologist Expertise** - we staff meteorologists to observe the entire fleet

# 2024-10-28 Ghana 5

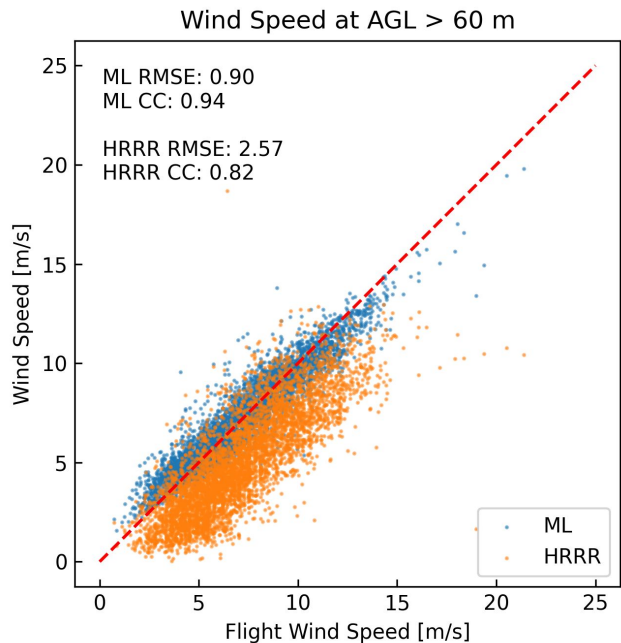


# 2024-04-16 Ghana

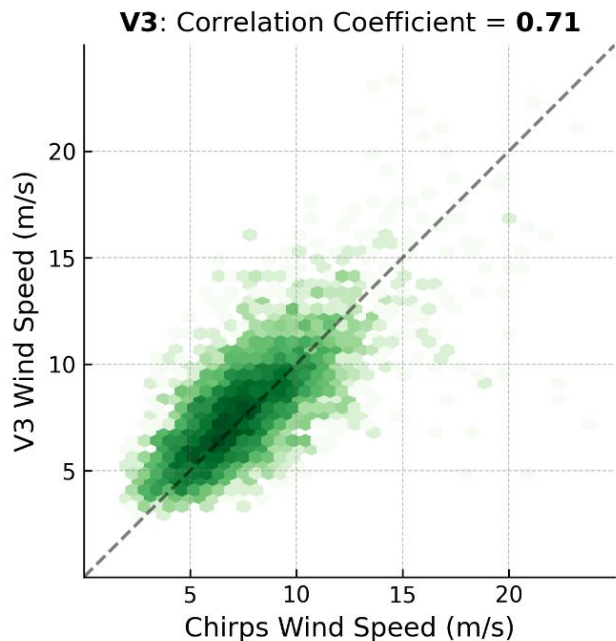


# Comparing Satellite to HRRR NWP Performance

## P1 AR-1 vs HRRR



## P1 Africa vs METEOSAT

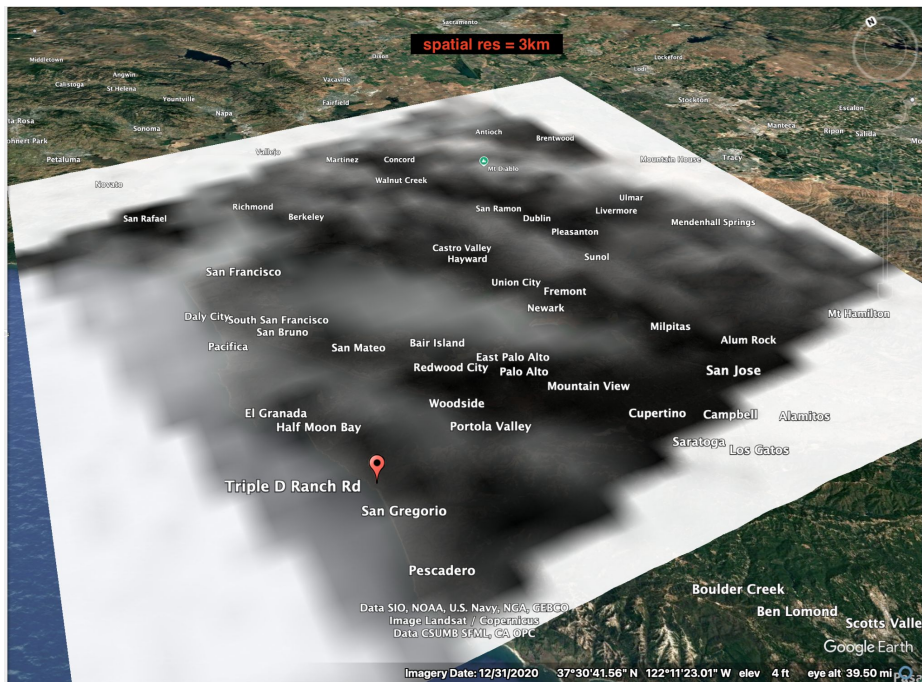


# With Terrain: HRRR Doesn't Cut It

- HRRR can't "see" San Bruno Mountain's Curved flow
- It's a pretty big feature
- I live this curved flow on my bike daily!

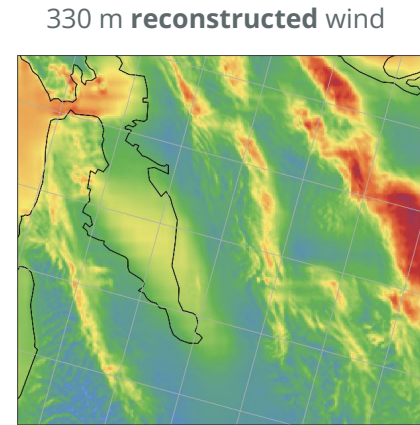
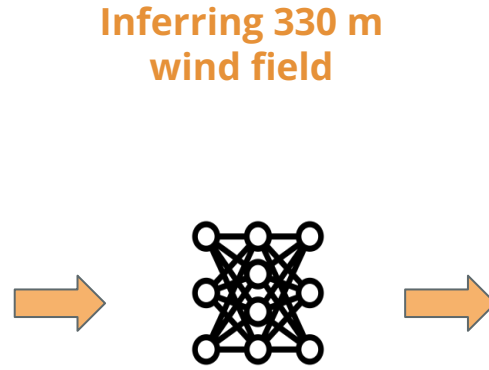
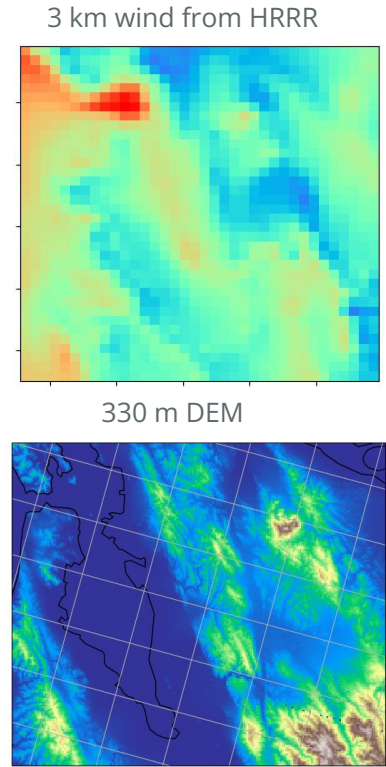


# Downscaling Experiments - HRRR 3 km to 330 m





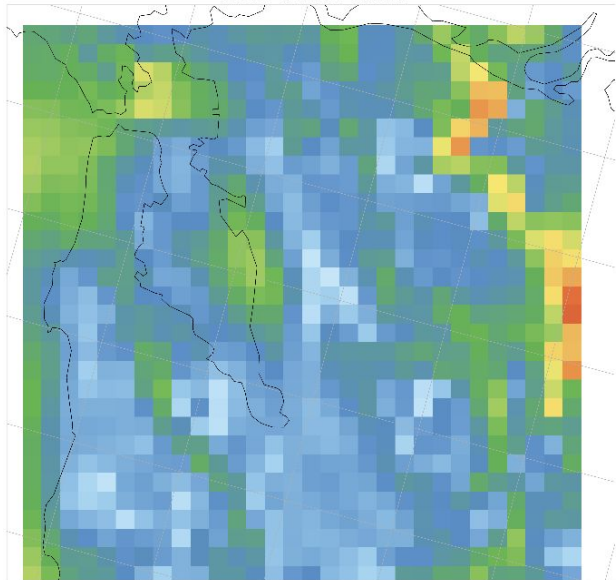
# Downscaling Model



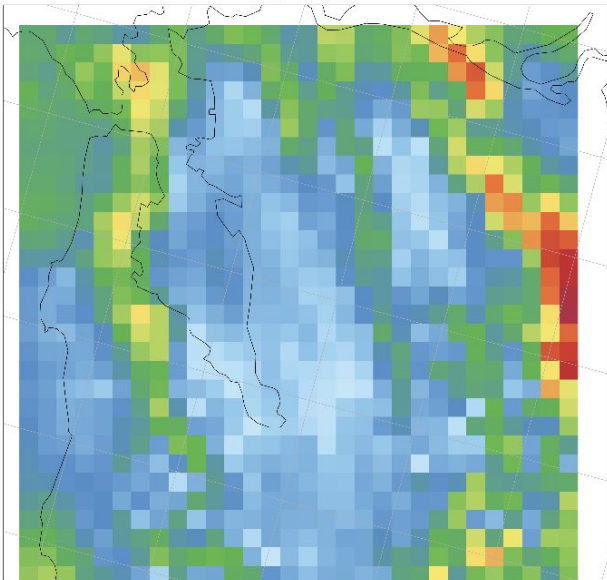
# Inference for CNN

2023-04-02T06:00:00.000000000

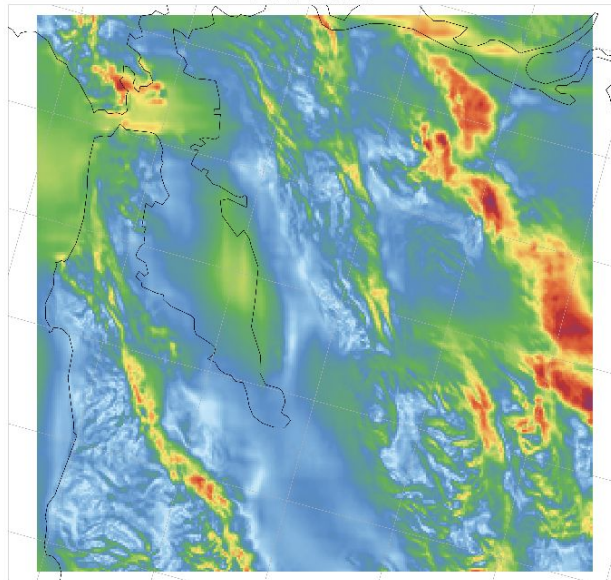
Observed 3km



ML 3km

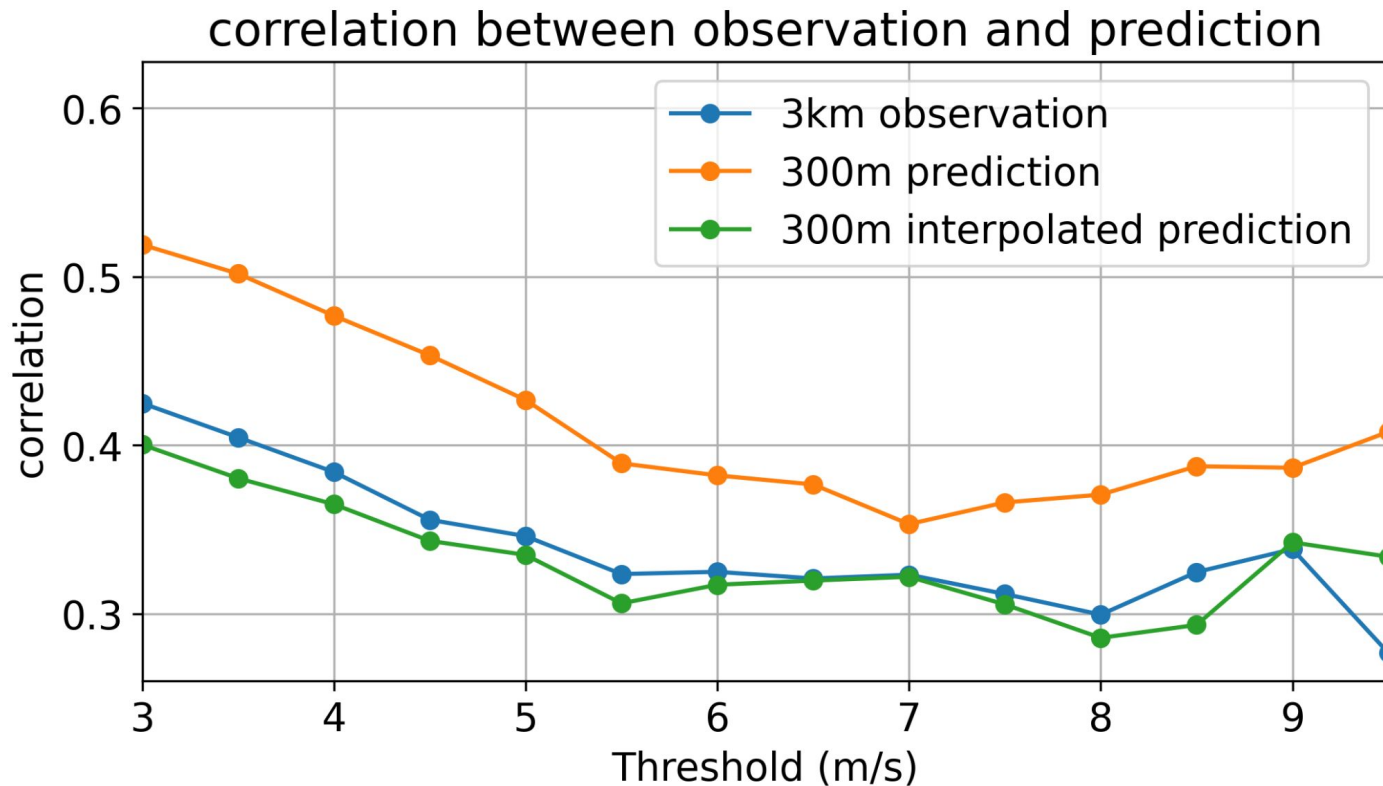


ML 330m



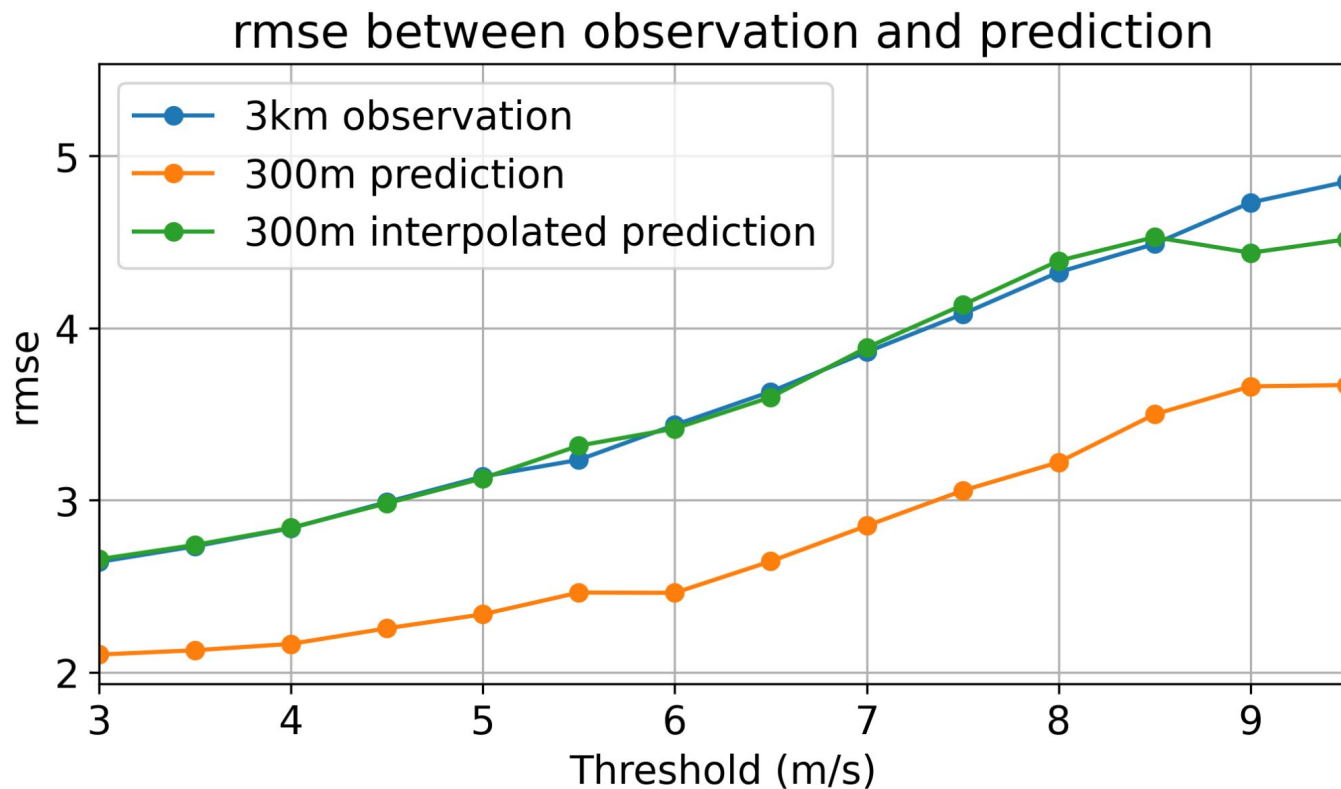
# Validation with madis at different thresholds

## Correlation



# Validation with madis at different thresholds

RMSE

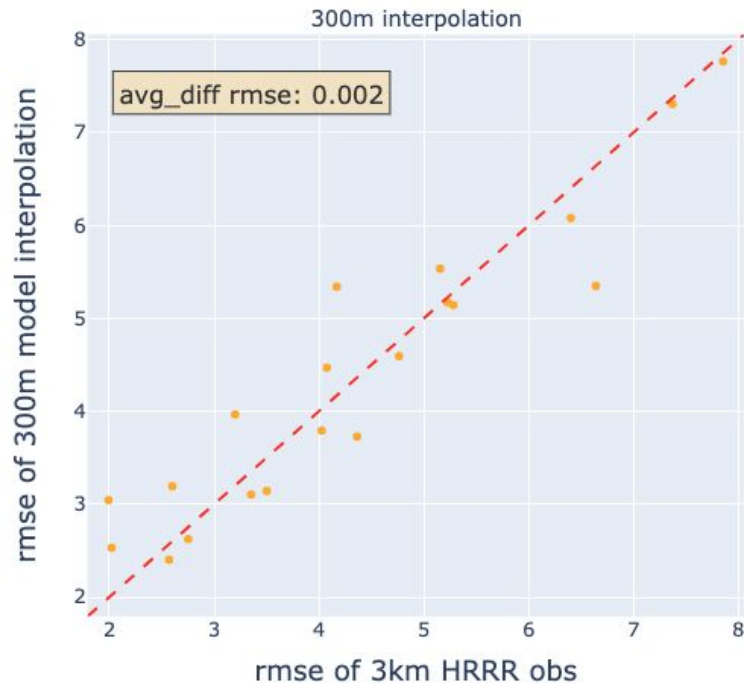
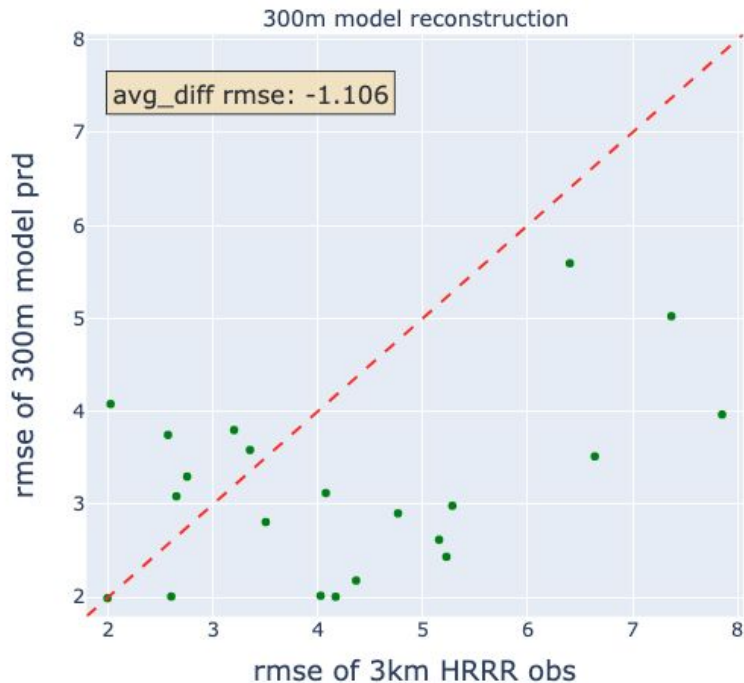


# Validating for wind above certain threshold

Comparison of rmse for wind higher than 7 m/s threshold

Features:

(Number of stations: 21)



# Thank you!

## Questions?

- **Contact: [john.celenza@flyzipline.com](mailto:john.celenza@flyzipline.com)**