Use and Misuse of Probabilistic Forecasting for Aviation Turbulence

John K. Williams

The Weather Company, an IBM Business

john.williams@us.ibm.com

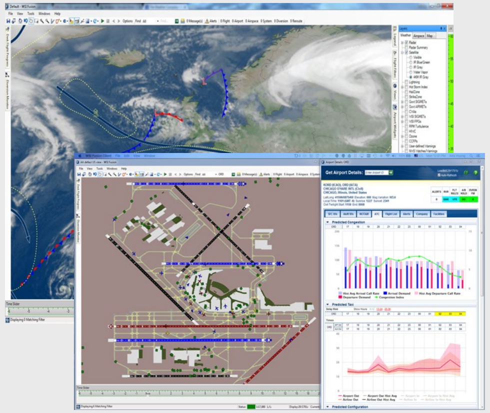
NCAR-MITRE-FAA Turbulence Mitigation Workshop IV

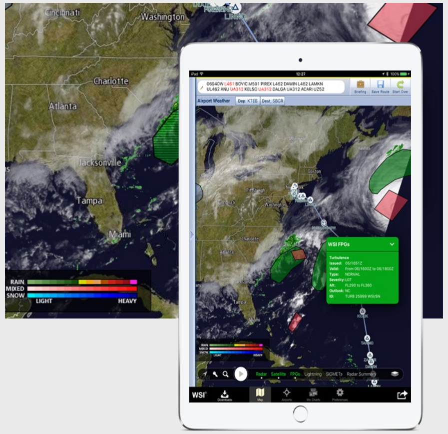
November 9, 2021

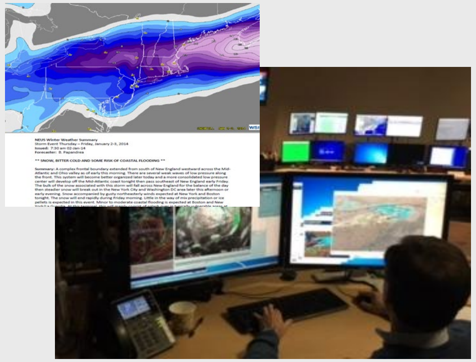
The Weather Company
An IBM Business

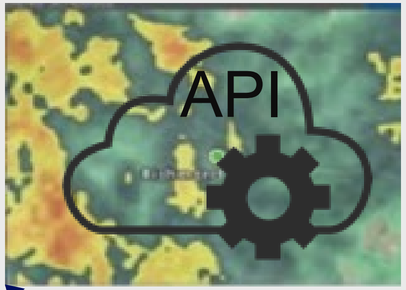
TWC/IBM Serves Airlines Around the World

Fusion	Pilotbrief	Hubcast	Enhanced Weather	Weather APIs
AOC planning and flight operations management.	Pilots preflight and inflight weather decisions.	Airport personnel weather conditions monitoring.	Operational decisions guidance.	Deliver data to customer platforms & applications.









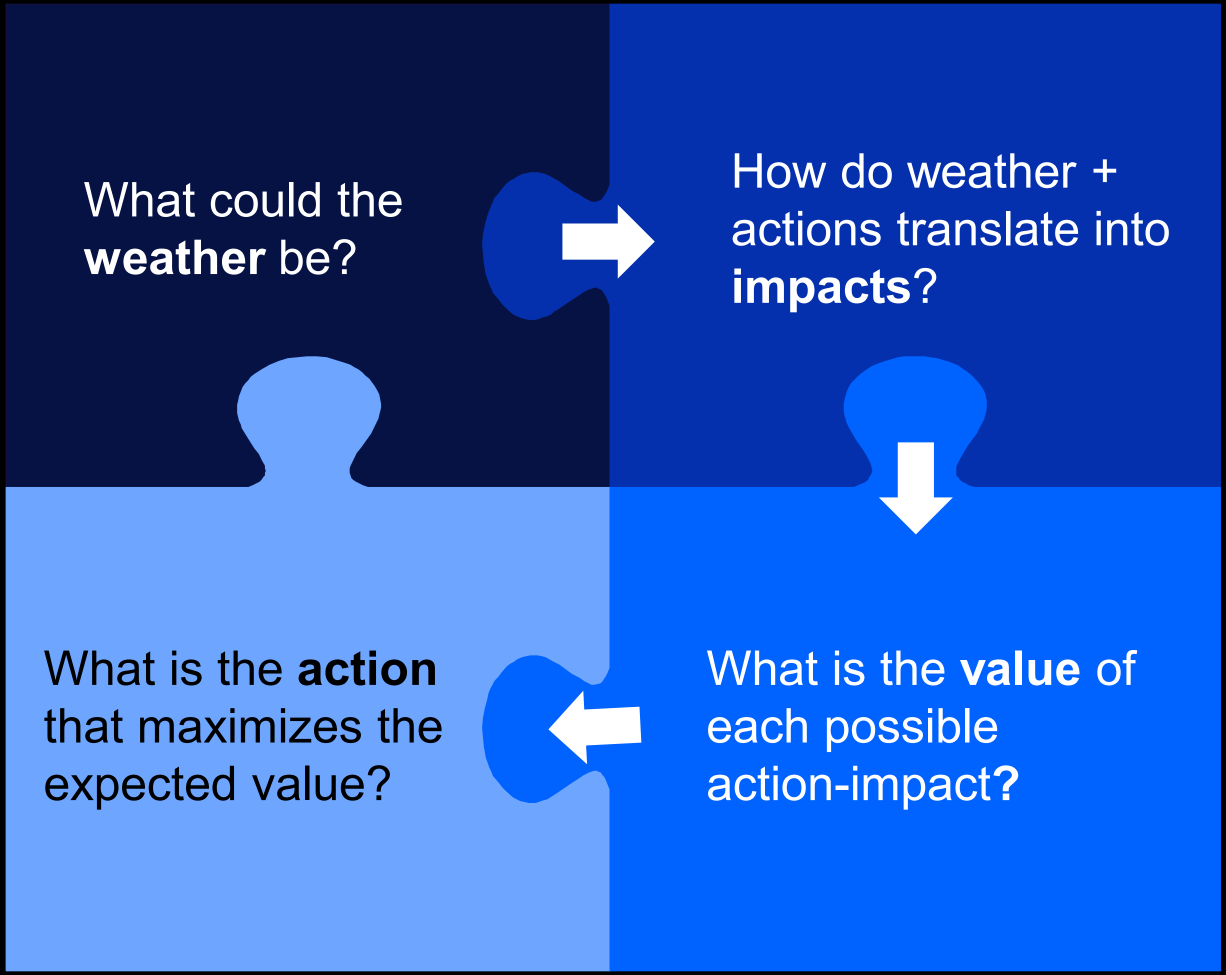
IBM's Turbulence Forecast

MPAS 15-km global NWP model

Our Goal: Better Weather-Impacted Decisions

"First, it should be understood that forecasts possess no intrinsic value. They acquire value through their ability to influence the decisions made by users of the forecasts."

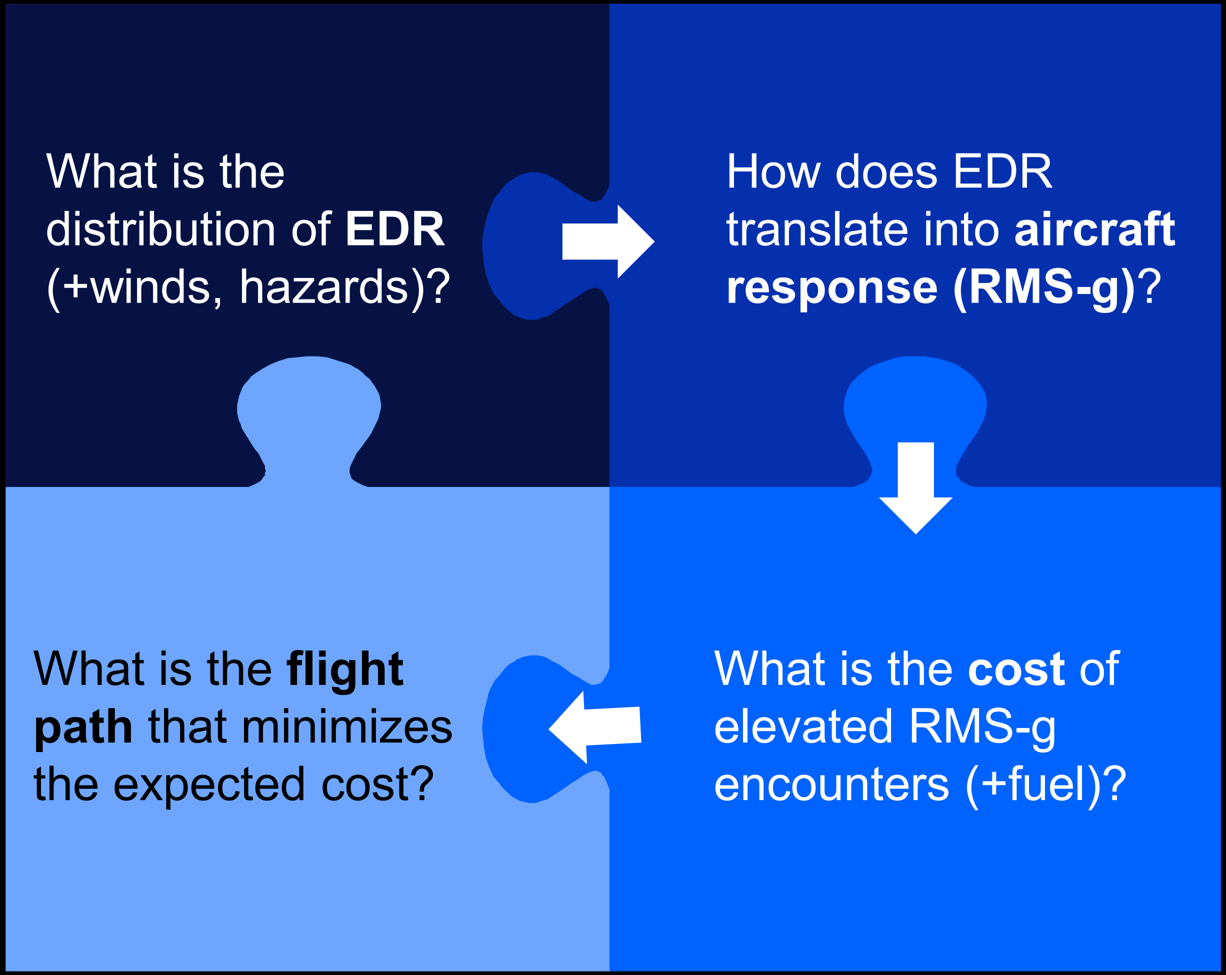
Murphy, A. H., 1993: What is a Good Forecast? An essay on the Nature of Goodness in Weather Forecasting, *Wea. Forecasting*, **8**, 281-293.



Our Goal: Better Turbulence Avoidance Decisions

"First, it should be understood that forecasts possess no intrinsic value. They acquire value through their ability to influence the decisions made by users of the forecasts."

Murphy, A. H., 1993: What is a Good Forecast? An essay on the Nature of Goodness in Weather Forecasting, *Wea. Forecasting*, **8**, 281-293.

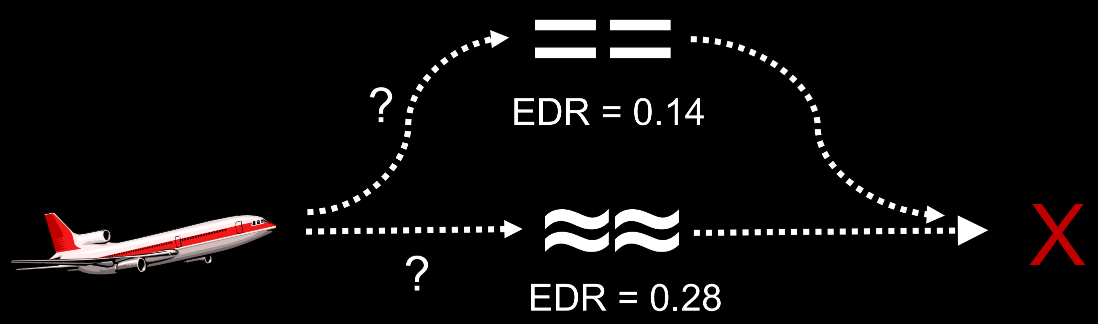


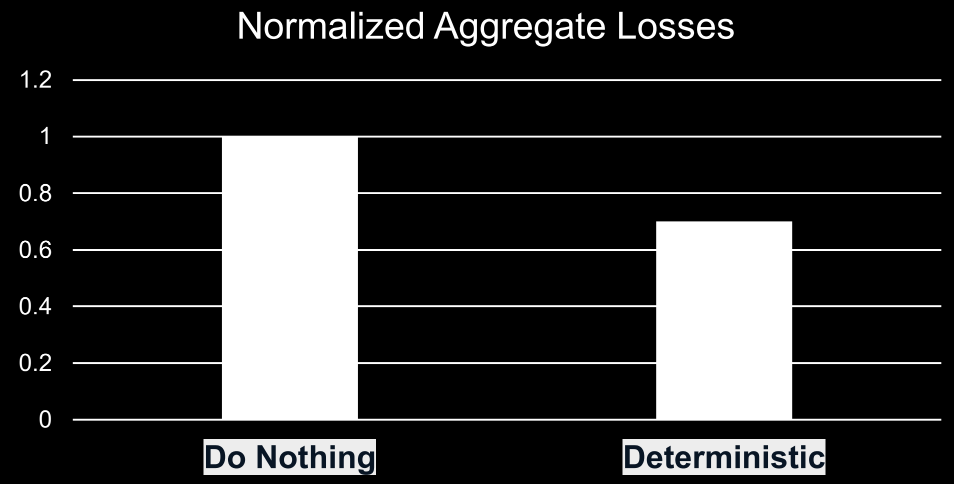
Probabilistic Forecasts Enable Better Decisions

Reroute around potential turbulence?

- Deviation is expensive in fuel and time
- *But* a severe turbulence accident could be much more costly.

Using a deterministic forecast helps …





Using deterministic forecast mitigates losses by 30%*

* Notional, not meant as a rigorous claim..

Probabilistic Forecasts Enable Better Decisions

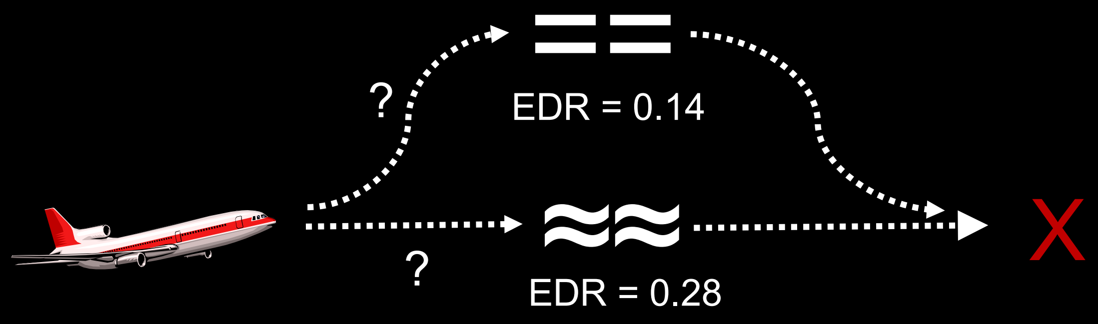
Reroute around potential turbulence?

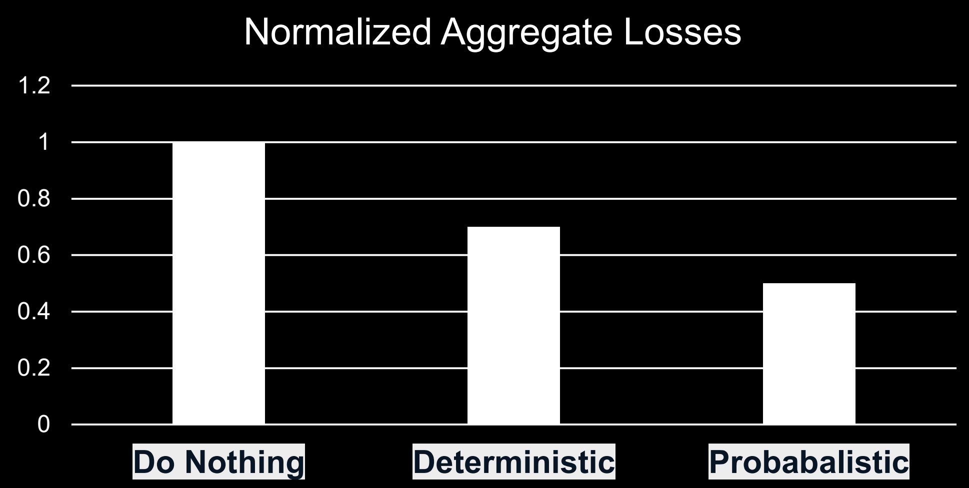
- Deviation is expensive in fuel and time
- *But* a severe turbulence accident could be much more costly.

Using a deterministic forecast helps …

… and using a probabilistic forecast in conjunction with cost-loss optimization further improves the net outcome.

- ***Probabilistic turbulence forecasts have 25% greater value than deterministic***
  Buchanan (2016), in Sharman & Lane (eds), *Aviation Turbulence*

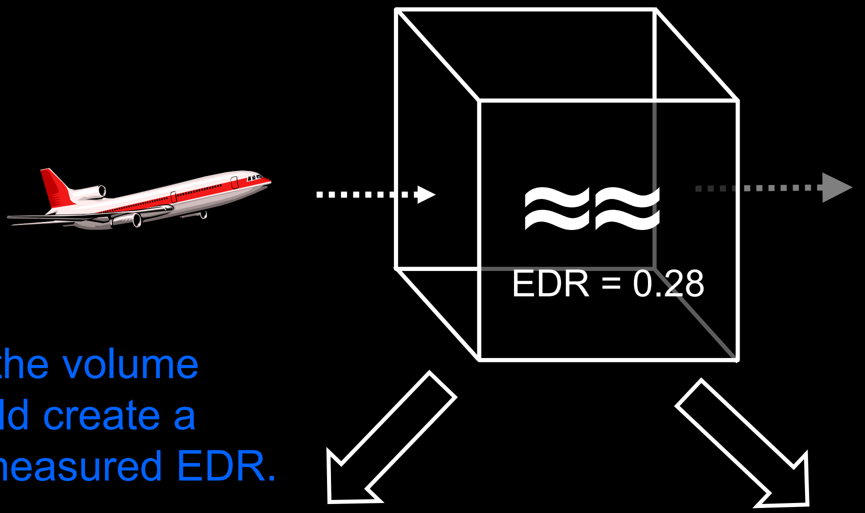




Using probabilistic forecast mitigates losses by 50%*

* Notional, not meant as a rigorous claim…

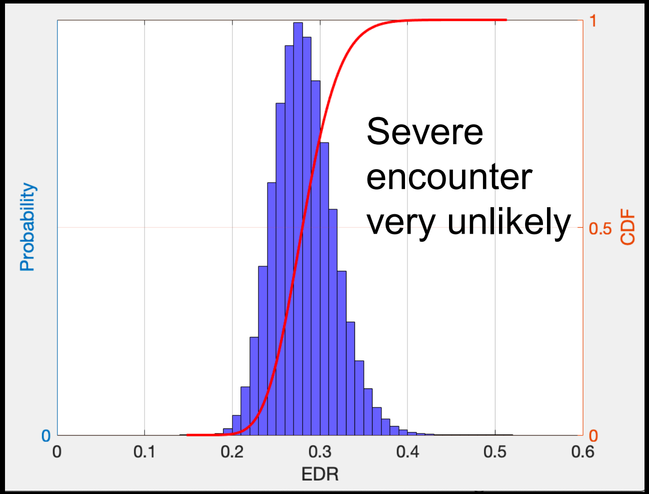
Importance of Forecast Uncertainty



Flying through the volume repeatedly would create a *distribution* of measured EDR.

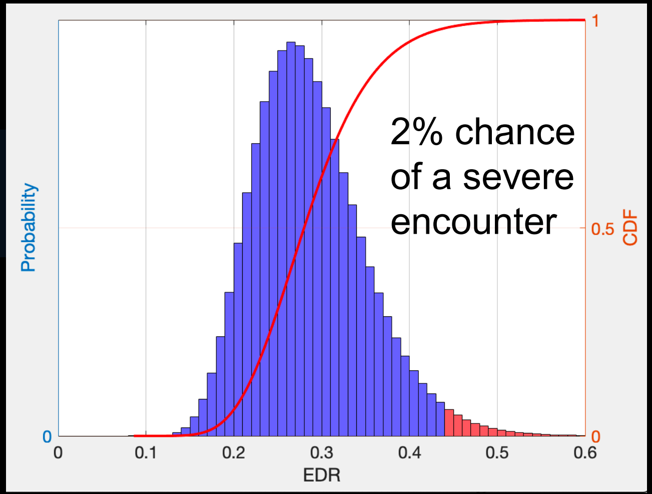
Sources of uncertainty:

- NWP model forecast errors
- Limited NWP model resolution
- Diagnostic linkage to aircraft scales
- Inhomogeneity
- Random nature of turbulence
- …



E[EDR] = 0.28; P(EDR > 0.44) < 0.01%

OR



E[EDR] = 0.28; P(EDR > 0.44) < 2.0%

?

Cost-Loss

Available Actions

Possible Weather

	Deviate	Don't Deviate
Severe Turb P(Y) = 2% = 0.02	**No Encounter (-$500)**	**Severe Encounter (-$50,000)**
Benign Turb. P(N) = 98% = 0.98	**No Encounter (-$500)**	**No Encounter ($0)**
Expected value	**-$500**	**-$1,000**

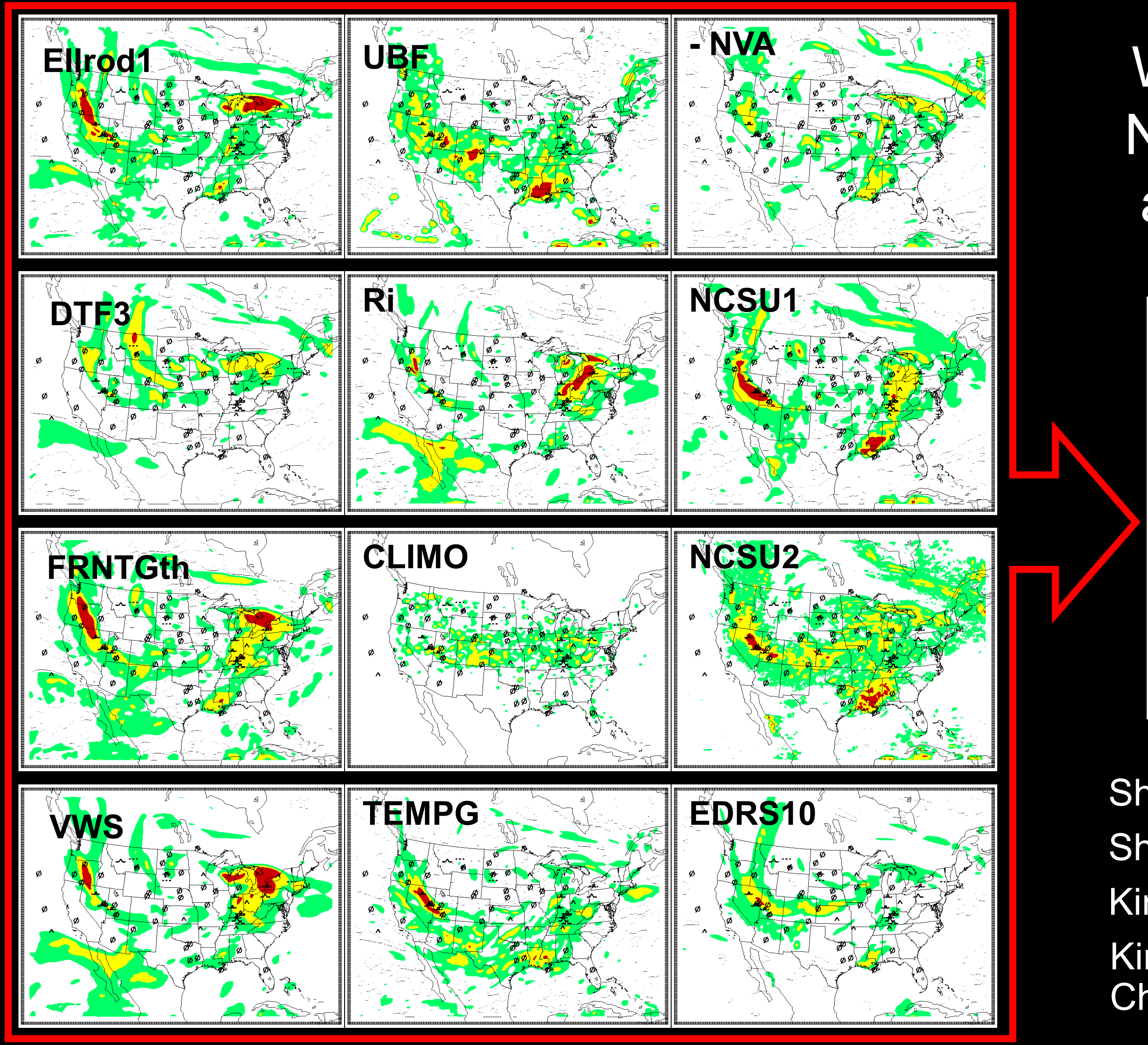
Optimal decision:
Action with highest expected value
⇒ Deviate!

Action-Impact

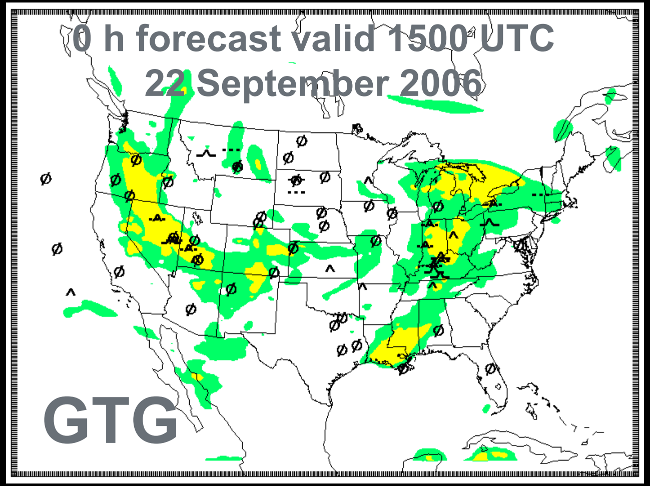
Impact Value

Expected Value = Sum of (Impact Values x Probabilities)

GTG Methodology Produces Rich Information



Weighted consensus of rescaled NWP-derived “diagnostics” tuned against turbulence observations



Sharman et al. (2006), *WAF*

Sharman and Pearson (2017), *JAMC*

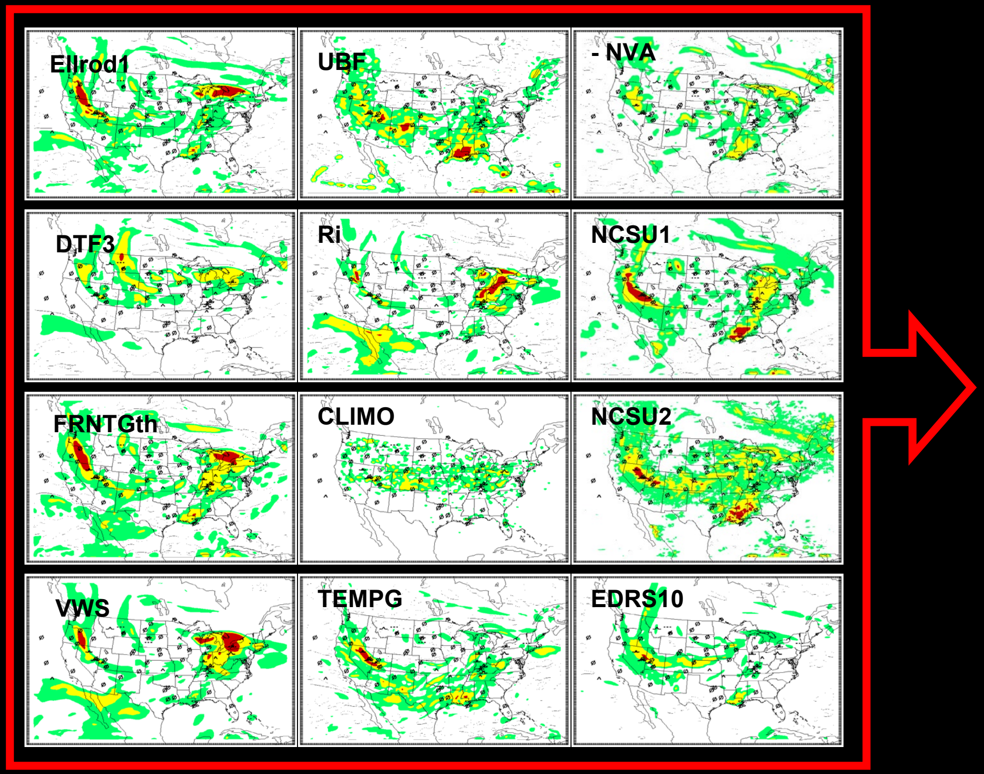
Kim, Chun, Sharman & Keller, (2011), *JAMC*

Kim, Sharman, Strahan, Scheck, Bartholomew, Cheung, Buchanan & Gait (2018), *BAMS*

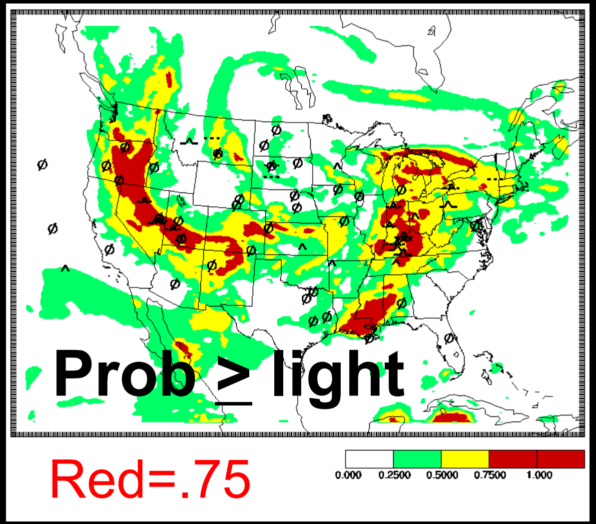
Images courtesy of Bob Sharman

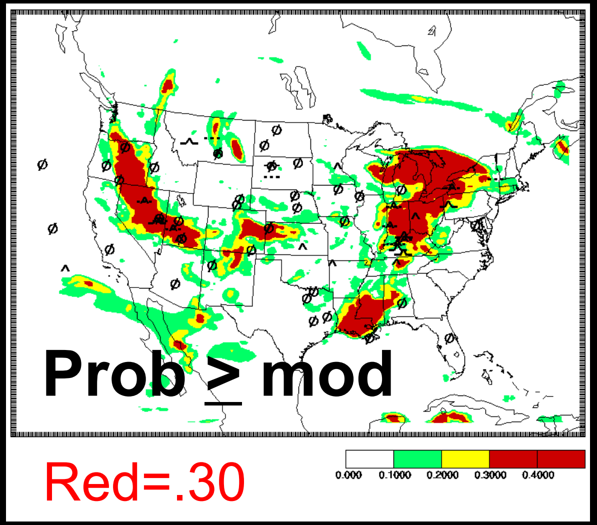
Forecasting Exceedance Probabilities

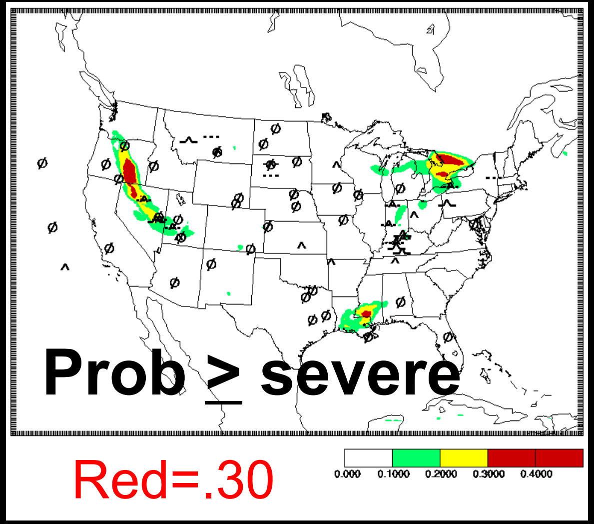
One approach: Probability ≈ fraction of diagnostics exceeding category threshold



Images courtesy of Bob Sharman







0 h forecast valid at 22 Sep 2006 15Z

Approach also applied to multi-model and time-lagged ensembles:

Kim, Chan, Sridhar, and Sharman (2015), *BAMS*

Buchanan (2016)

Shin, Deierling, Muñoz-Esparza, Sharman (2021), ARAM

Advanced Machine Learning Approaches

Training against turbulence observations, e.g., *in situ* EDR, Sharman et al. (2014), *JAMC*

Clear air turbulence forecasts

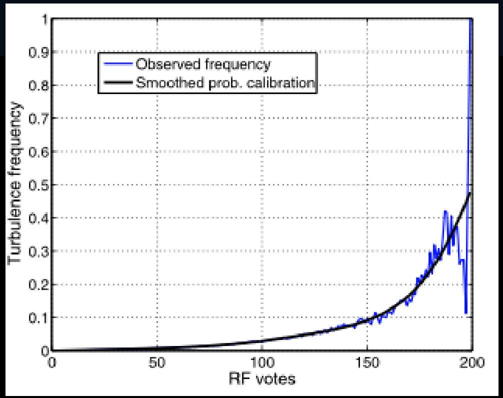
Table 1. High-altitude turbulence forecast skill scores for years 2010 and 2011, evaluated using pilot reports and automated aircraft turbulence reports. AI methods were trained on 40,000 random samples from 2011 with 30% turbulence cases and evaluated on all of 2010, and vice versa. The *k*-nearest neighbors method used 100 analogs. TSS is the true skill score.

Method	Year	ROC AUC	Max CSI	Max TSS
GTG weighted mean	2010	0.791	0.137	0.443
	2011	0.775	0.132	0.418
Logistic regression	2010	0.822	0.162	0.496
	2011	0.805	0.149	0.461
k-nearest neighbors	2010	0.832	0.167	0.514
	2011	0.818	0.163	0.482
Random forest	2010	0.849	0.179	0.541
	2011	0.830	0.169	0.499

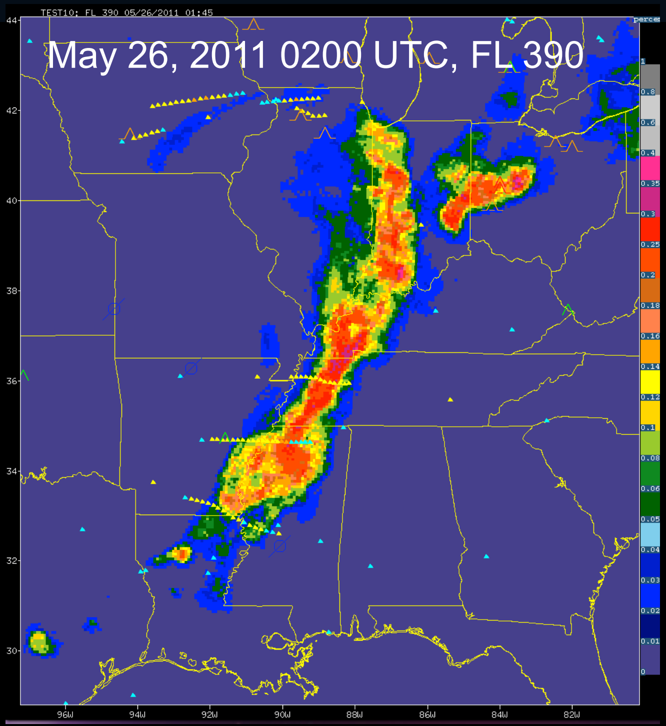
ML and GTG skill comparison, forecasting EDR ≥ 0.3. From McGovern, …, **Williams** (2017), *BAMS*

Muñoz-Esparza, **Sharman**, and Deierling (2020), *JAMC* use RFs and Gradient Boosted Regression Trees.

Convectively-induced turbulence nowcast



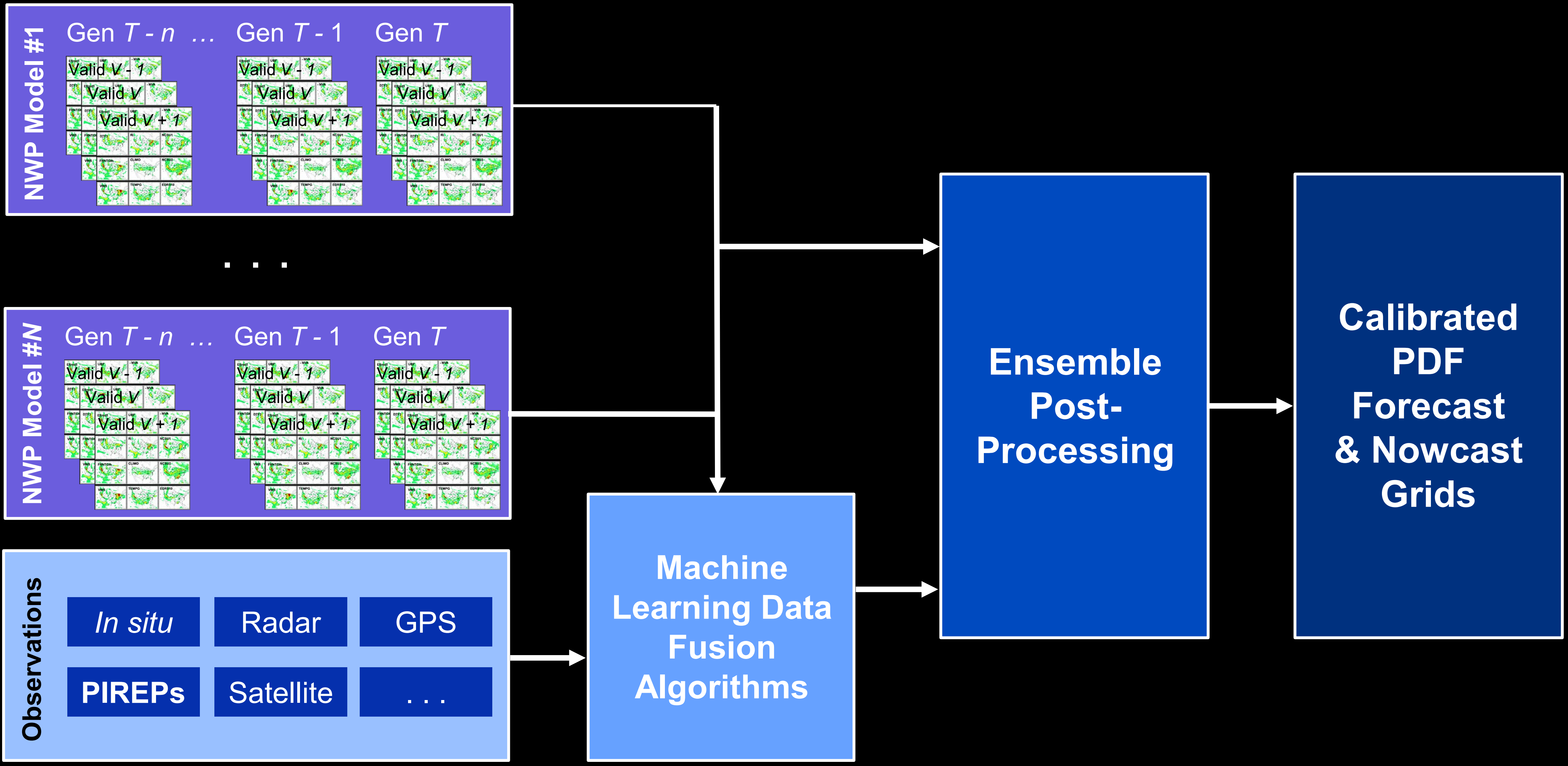
Random forest votes calibrated to probability



From Figs. 1 and 2 of **Williams** (2014), *Machine Learning*

RF Prob(EDR ≥ 0.2) based on NWP, satellite and radar data

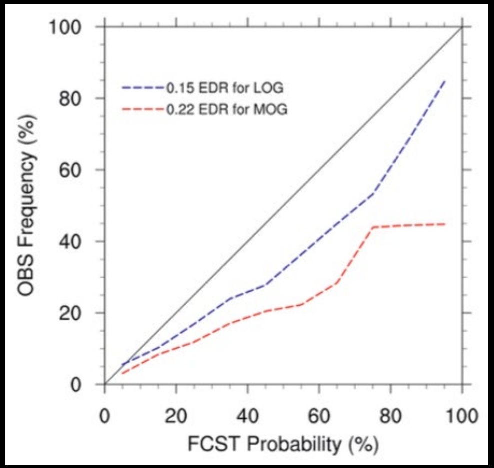
Probabilistic Forecasting Approach



Ideal Probabilistic Forecast Attributes

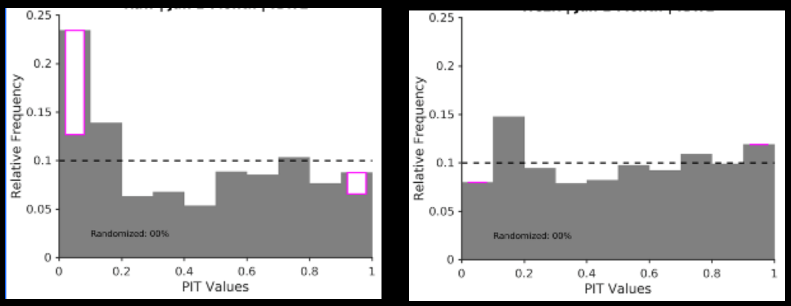
- Clearly defined
  - E.g., likelihood per nautical mile or per minute of flight
  - Otherwise, can't account for turbulence volume size / duration
- Calibrated
  - Probabilities must be accurate (reliable; flat PIT diagrams)
  - Otherwise, derived cost/loss estimates are flawed
- As "sharp" as possible
  - More specific forecasts preferable (low Brier Score, CRPS)
  - Otherwise, can be calibrated but not useful
- Forecast the entire Probability Distribution Function
  - Provides probability for every EDR the aircraft may encounter
  - Otherwise, challenging to translate to encounter costs for a variety of aircraft types, weights, flight conditions

Reliability Diagram



Kim et al., *BAMS* (2018)

Probability Integral Transform Diagram

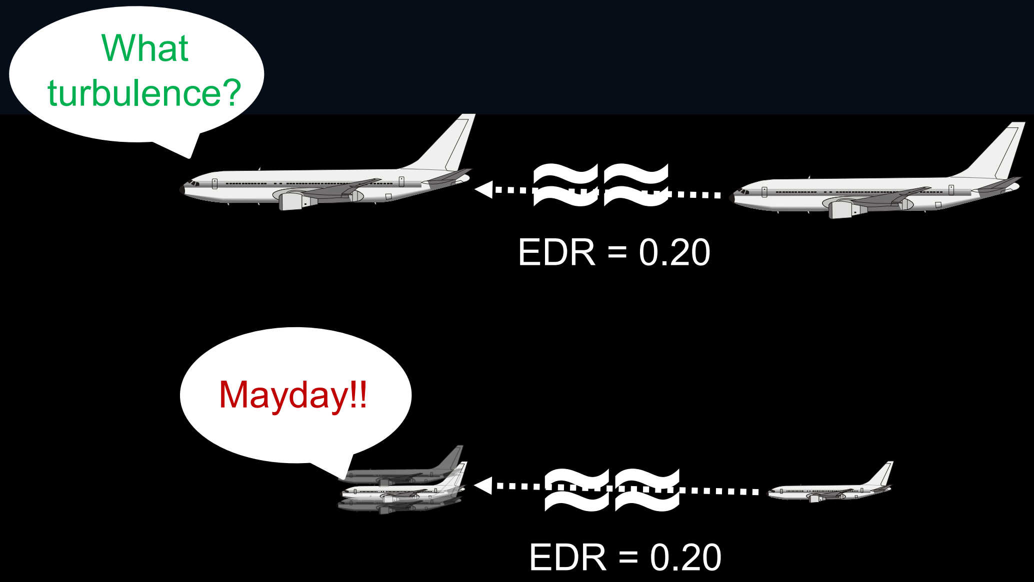


Uncalibrated

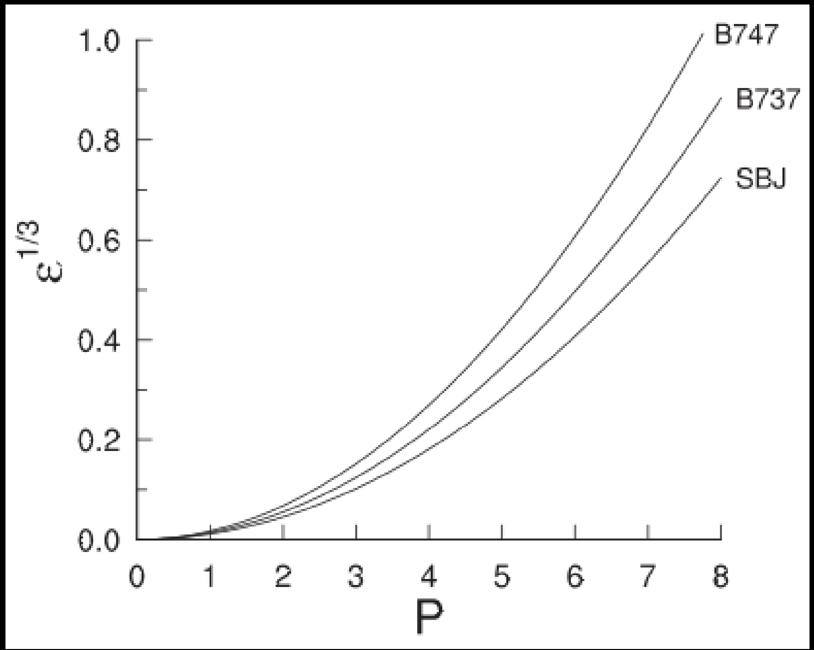
Calibrated

Categorical Probabilities Aren’t Enough

Turbulence impact depends on aircraft type, weight, airspeed, cabin status, …



$$\sigma_g = F(\mathrm{ac}, V_T, z, W)\varepsilon^{1/3}, \qquad (20)$$



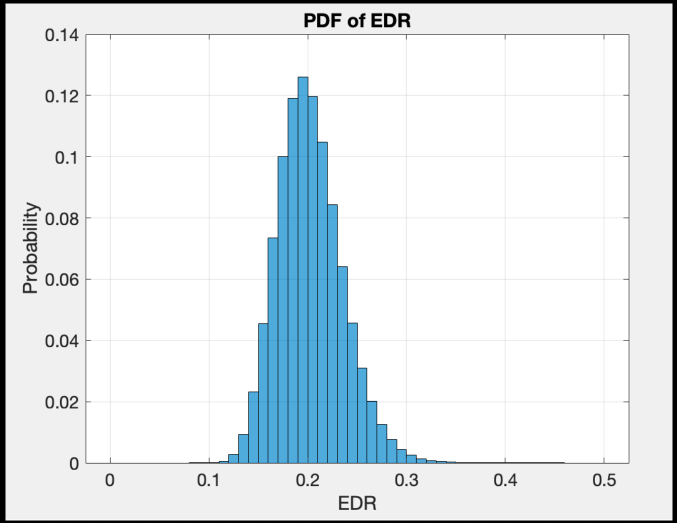
Eq. (20) and Fig. 7 from Sharman et al. (2014), *JAMC*

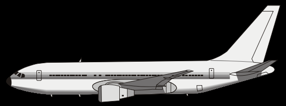
Prob(EDR ≥ 0.2) = 15% may have quite varied decision implications…

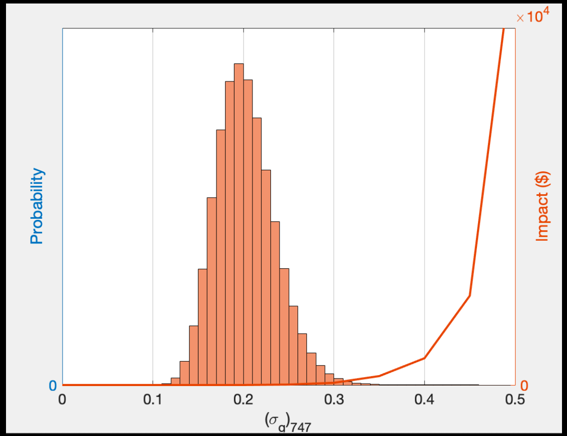
PDF forecasts provide relevant information for a variety of situations!

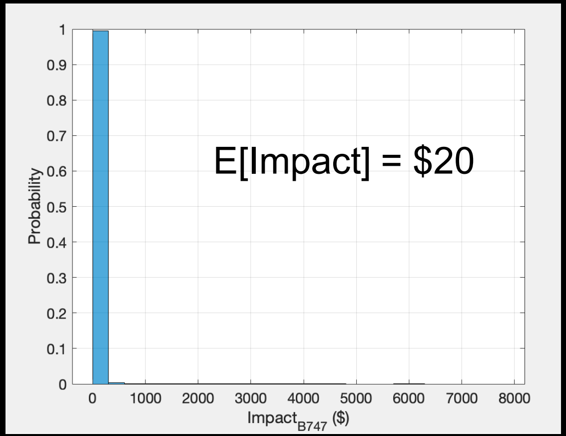
Forecast the EDR Probability Distribution Function!

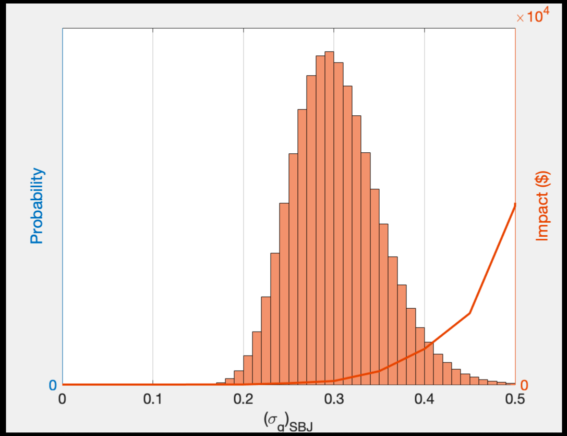
EDR PDF → σ_g PDFs → Impact PDFs

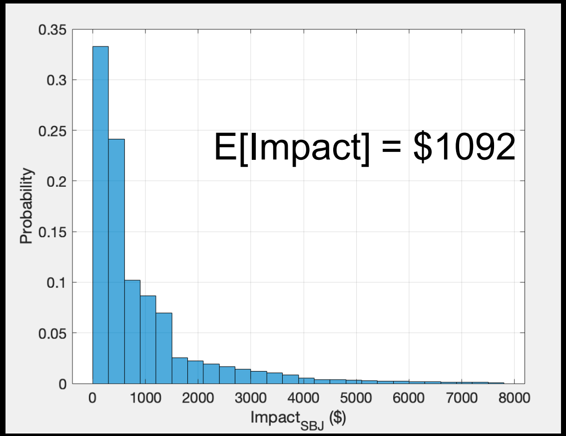












Using Hazard Forecasts to Find Safe Routes

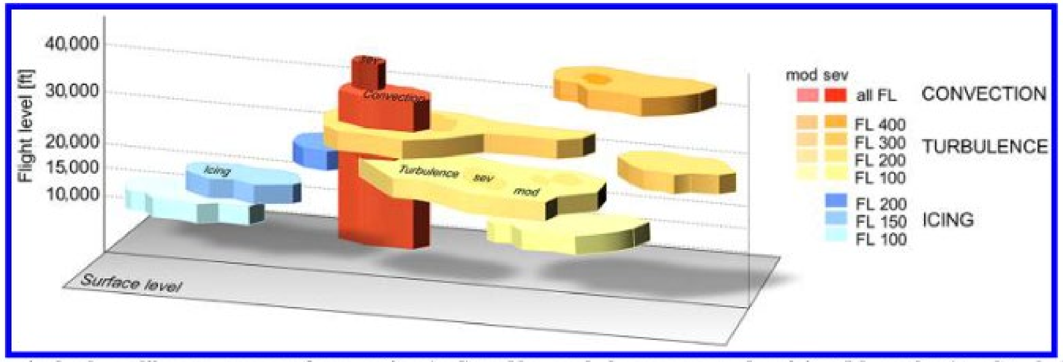
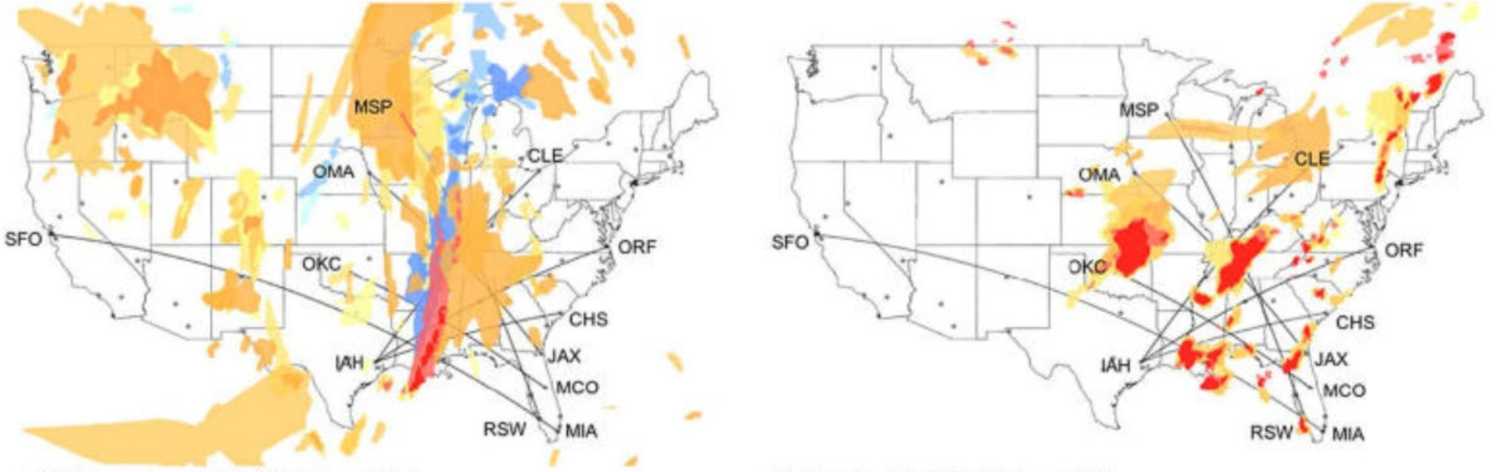


Fig. 1 Schematic of typical columnlike occurrence of convection (red) and layered phenomena such as icing (blue colors) and turbulence (yellow/orange).



a) 18 November 2015, 0600 hrs UTC

b) 14 July 2016, 1600 hrs UTC

Fig. 2 Sample weather situations, with hazard areas color coded according to the legend in Fig. 1, overlaid on great-circle routes connecting seven city pairs (San Francisco International Airport (SFO)-Miami International Airport (MIA), George Bush Intercontinental Airport Houston (IAH)-Cleveland Hopkins International Airport (CLE), IAH-Norfolk International Airport (ORF), IAH-Charleston International Airport (CHS), Will Rogers World Airport Oklahoma City (OKC)-Orlando International Airport (MCO), Eppley Airfield Omaha Airport (OMA)-Jacksonville International Airport (JAX), Minneapolis–Saint Paul International Airport (MSP)-Southwest Florida International Airport (RSW)).

Avoidance regions based on forecast hazard intensity thresholds

Figs. from Sauer, Steiner, Sharman, Pinto and Deierling (2019), *J. Air Transportation*

Others: wind-optimal routes using severe turbulence probabilities

E.g., Williams (2009); Kim et al. (2016)

PDF Translation Enables Enhanced Decision Making

Individual Flight: Compare Routes

Option	Fuel Cost	Wx Hazard Impact	Delay Cost	Total
Route #1	$25k	$15k	$0k	$40k
Route #2	$30k	$4k	$4k	**$38k**
…				
Route #N	$35k	$2k	$8k	$45

Fleet Ops: Compare Playbooks

Option	Aircraft Ops.	Sched. Impact	Total
Playbook #1	$4.0M	$300k	$4.3M
Playbook #2	$3.8M	$200k	**$4.0M**
…			
Playbook #N	$4.5M	$50k	$5.0M

Or compare to climatological mean impact to characterize relative risk.

c.f. Fig. 3 in Lane, **Sharman**, …, **Williams** (2012), *BAMS*

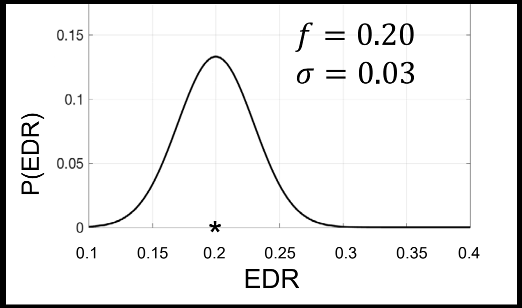
Methods for Creating PDF Forecasts

Goal: optimize sharpness (specificity) and calibration (probabilistic accuracy)

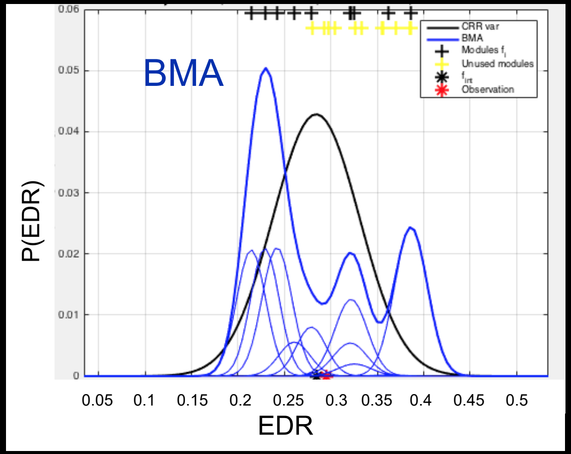
- Bayesian Model Averaging (Raftery et al. 2005, *MWR*)
- Heteroskedastic Extended Logistic Regression (Messner et al. 2014, *MWR*)
- Deep Learning (e.g., Grönquist et al. 2020; Bartholomew et al. 2021, ARAM 8.4)

Ex.: Bayesian Model Averaging

- Bias-correct / calibrate each forecast
- "Dress" each forecast EDR

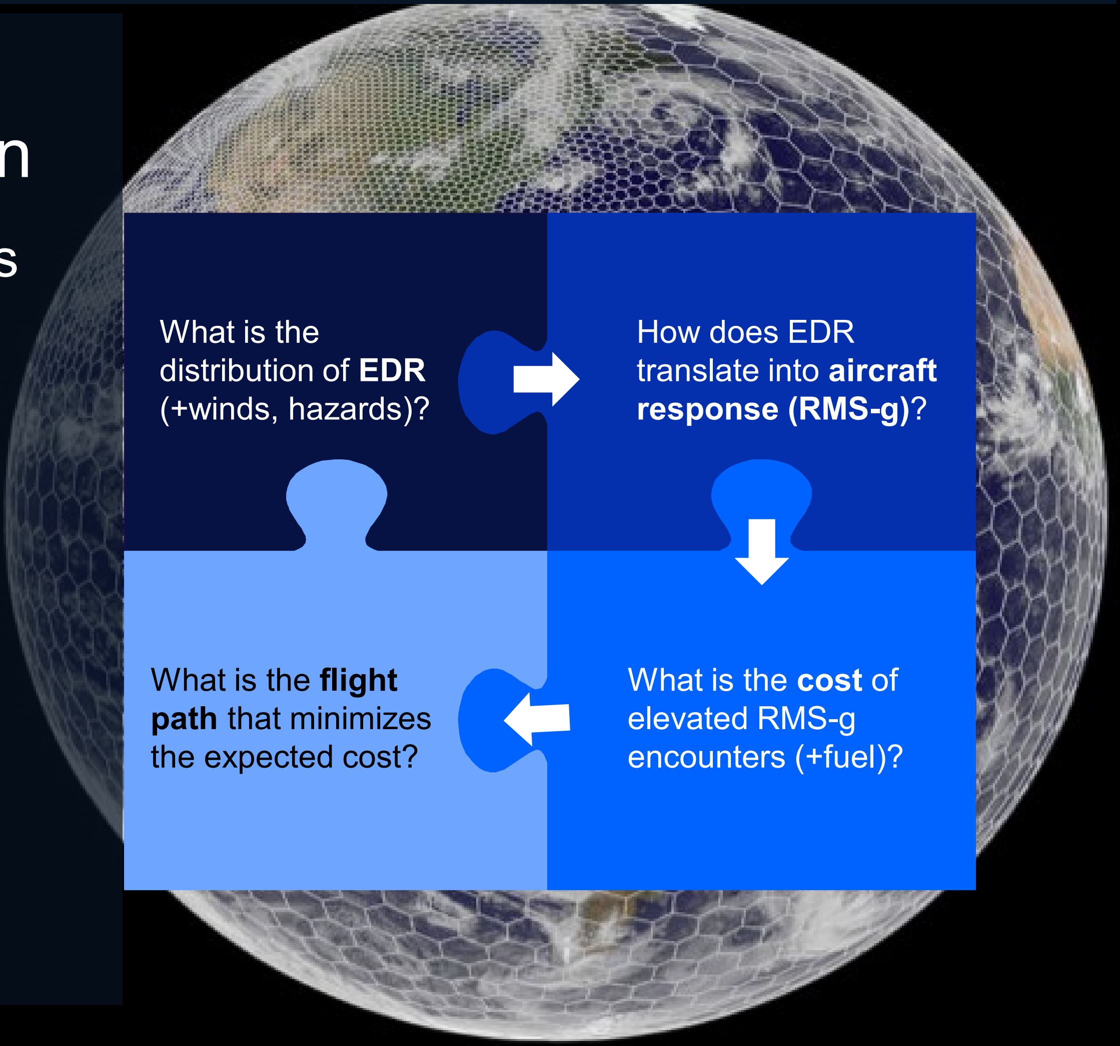


- Compute the weighted combination



Summary

- NWP model and turbulence diagnostic ensembles offer rich uncertainty information
  - Post-processing can produce calibrated probabilities
- Turbulence PDF forecasts support optimal decision making
  - Translation to aircraft response & encounter costs
  - Relative risk and cost maps
  - Route selection, deviation decisions, cabin management
- Uncalibrated, poorly defined or inaccurate probabilities can be misleading



John K. Williams

The Weather Company, an IBM Business

john.williams@us.ibm.com

The Weather Company

An IBM Business