

# A Multiple-Model Ensemble Examination of the Probabilistic Prediction of Tropical Cyclones

Christopher Melhauser

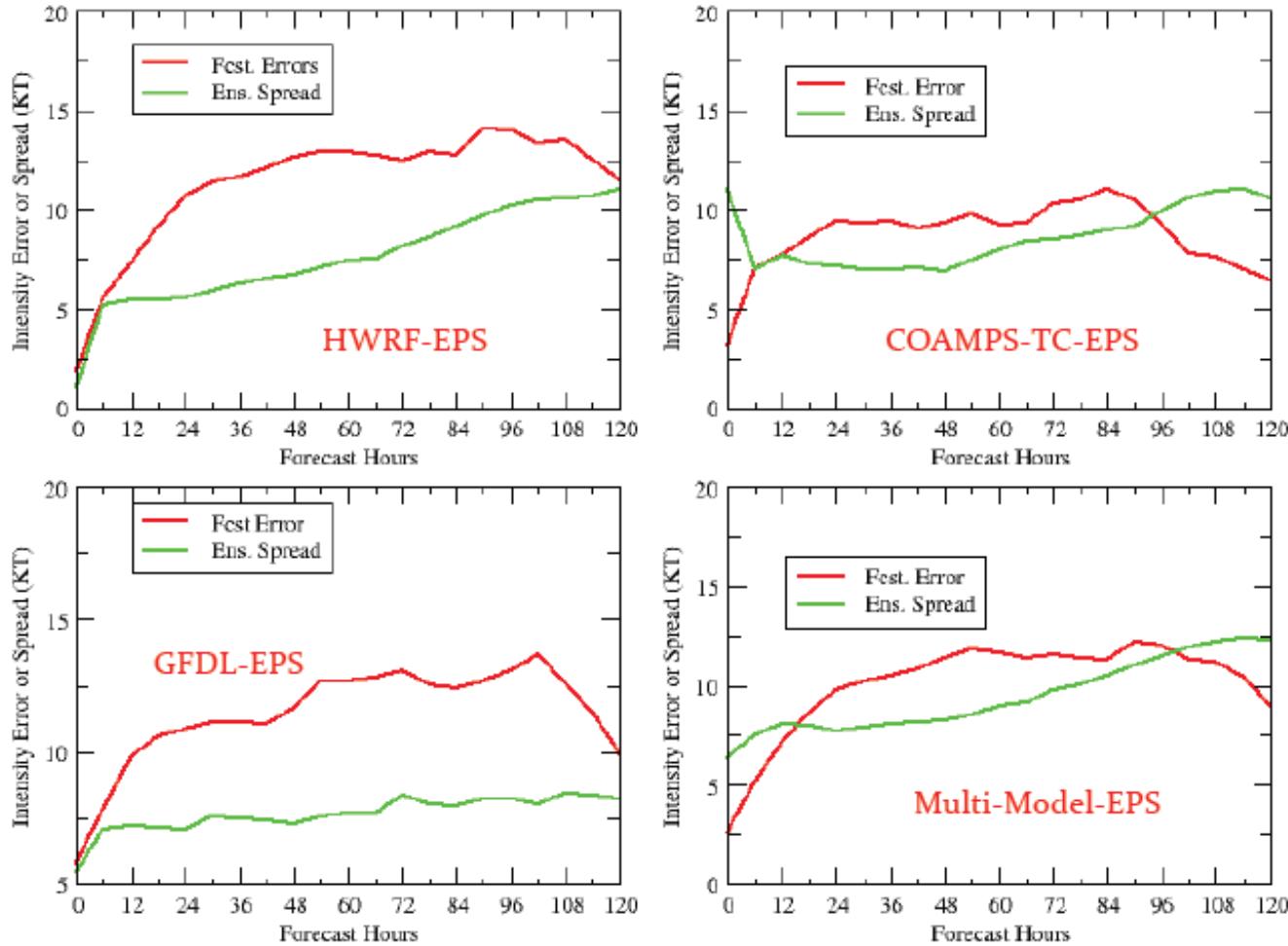
End-of-Semester Group Meeting  
December 11, 2015

Sponsored by ONR, NOAA, and NASA

Given pseudo-operational configurations of three TC-tuned regional models ...

- 1) How do the mean and spread of an ensemble with the same initial perturbations evolve using “multiple-models”?
  
- 2) Are single-core ensembles sufficient for representing model uncertainties in TC prediction or do we need multi-core ensembles?

# Why do we want more spread?



2014 HFIP Annual Meeting (Zhang and Peng)

An ensemble system is performing “well” when:  
**ensemble spread  $\approx$  forecast error**

- **Pseudo-operational configurations of TC tuned regional models**
  - Same initial conditions for all models
  - Fixed SST from operational GFS
  - Atmospheric component only
- **Model Versions**
  - WRF-ARW v3.6
  - HWRF v3.5b
  - COAMPS-TC v1.0 (Feb 2015)
- **Model Configurations**

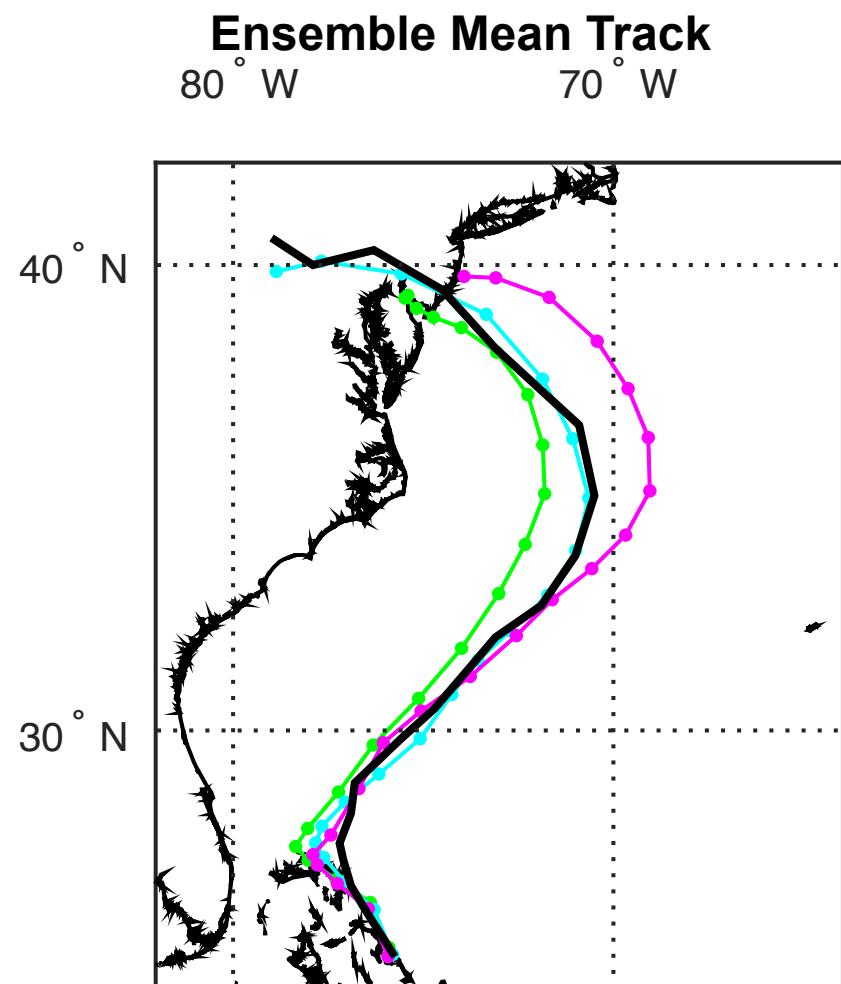
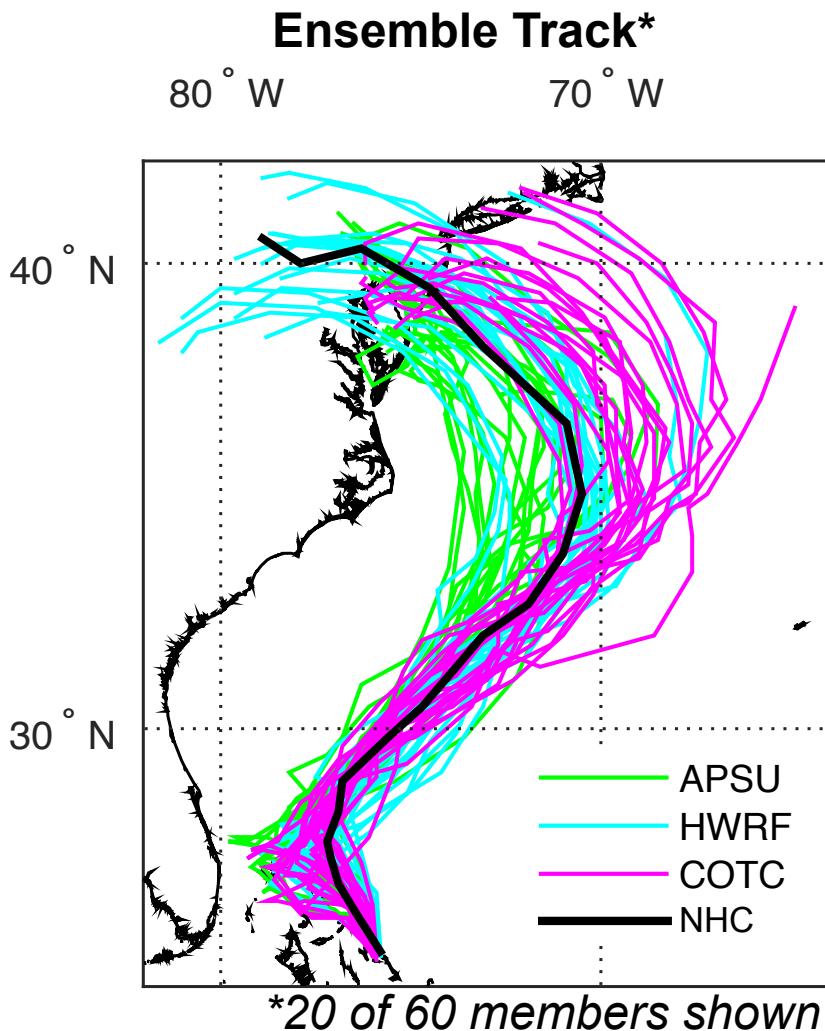
 APSU	PSU WRF-ARW w/ 2014 real-time configuration
 HWRF	HWRF w/ 2013 HWRF configuration, but no ocean
 COTC	COAMPS-TC w/ 2014 HFIP configuration

# Hurricane Sandy (2012) Ensemble Track

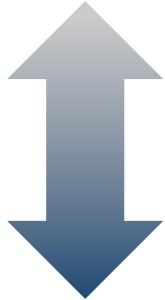
5 day forecast initialized 2012-10-26 00 UTC

Ensemble track mean and spread reproduced by each model-core

- Systematic mean difference at longer lead times



# WRF-ARW Multi-Physics Ensemble Configuration



## “APSU-Like” Physics

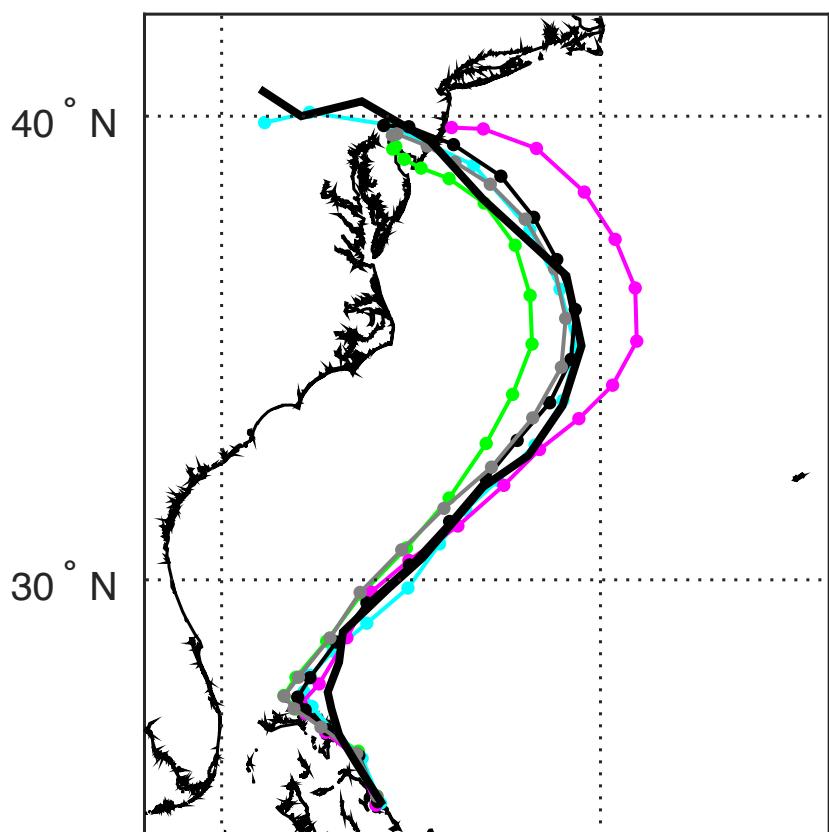
*Modify microphysics, radiation,  
PBL, surface drag, cumulus*

## “HWRF-Like” Physics

# Hurricane Sandy (2012) Multi-Model Ensembles

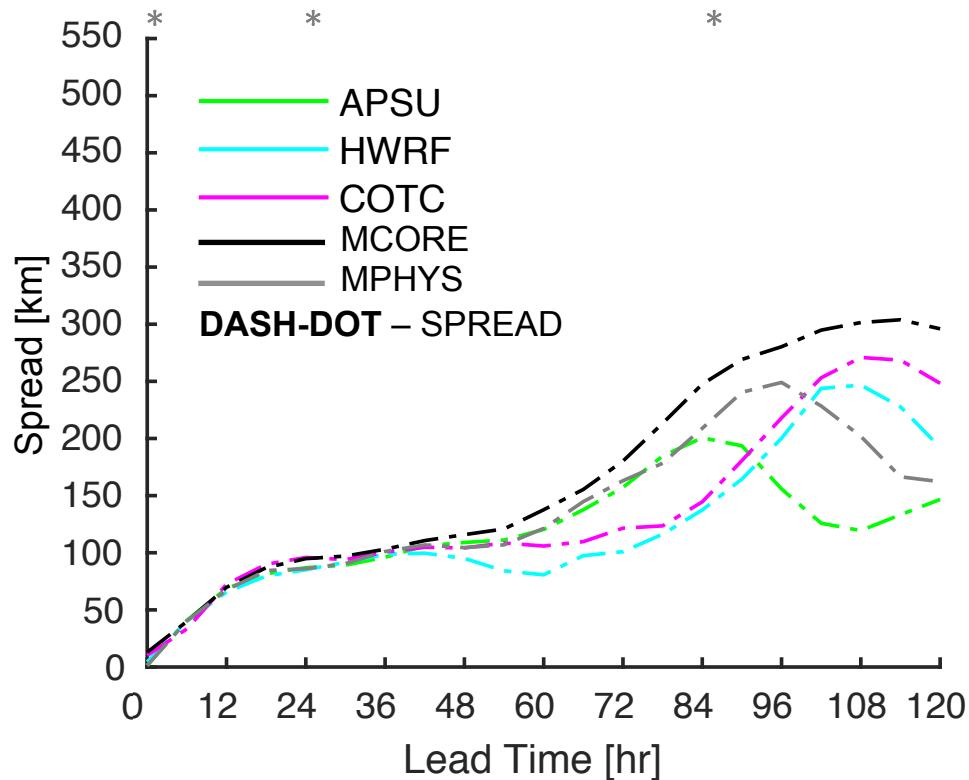
Multi-core and multi-physics ensemble track error distributions cannot be proven statistically different\*\*

**Ensemble Mean Track**



Increased track spread at longer lead times  
for multi-core or multi-physics

**Track Spread\***



\*NHC best track data as verification

\*\* Bootstrapped Kolmogorov–Smirnov test  
(10,000 samples;  $\alpha=0.05$ )

# Hurricane Edouard (2014) Ensemble Track and Intensity

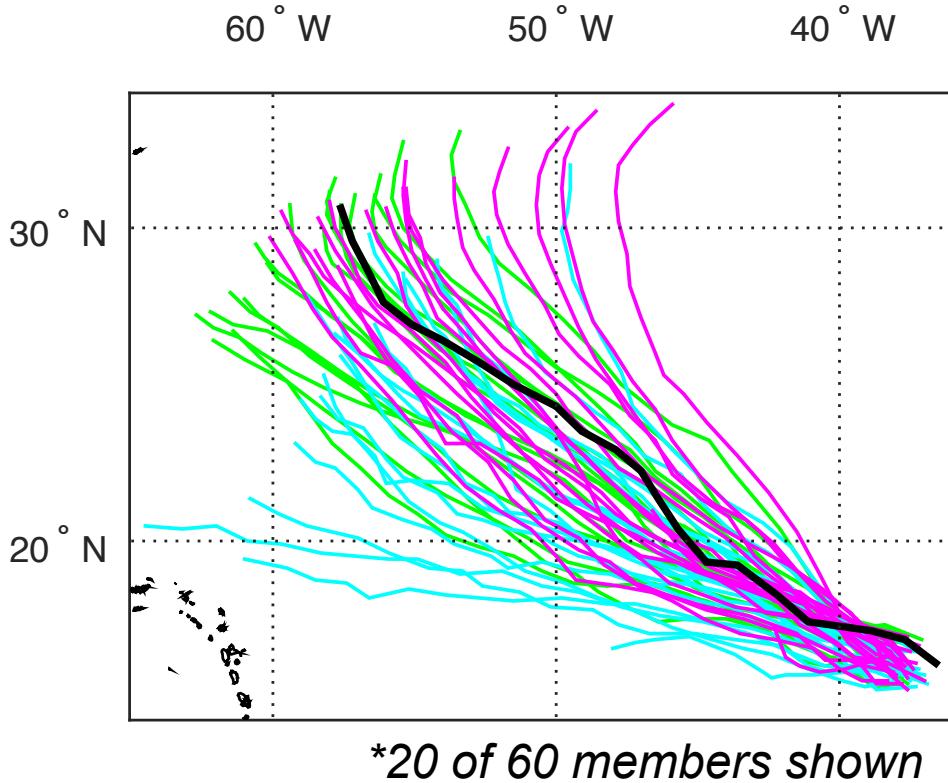
5 day forecast initialized 2014-09-11  
12 UTC

Multi-core ensemble spread  
larger than any individual single-  
core

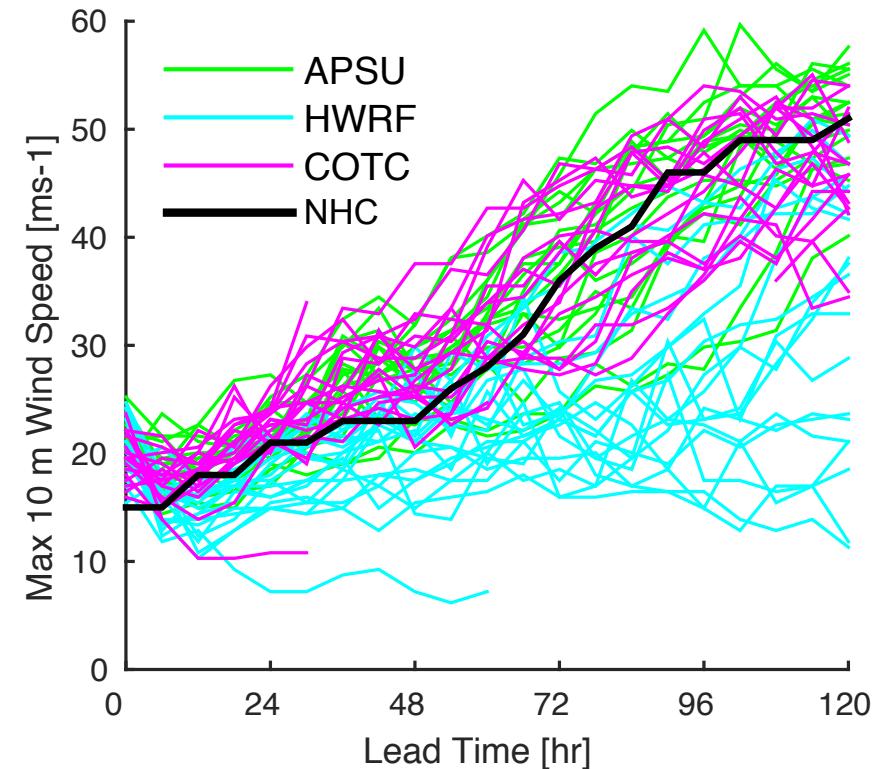
Spread:

APSU ~ COTC !~ HWRF

## Ensemble Track\*



## Ensemble Intensity\*



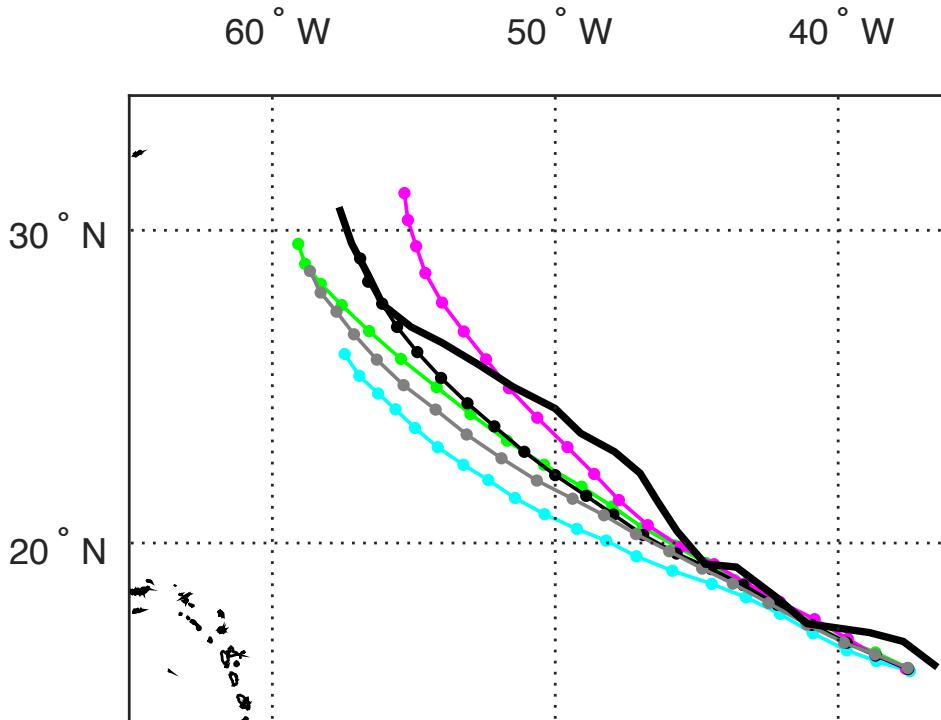
# Hurricane Edouard (2014) Mean Track and Intensity

5 day forecast initialized 2014-09-11  
12 UTC

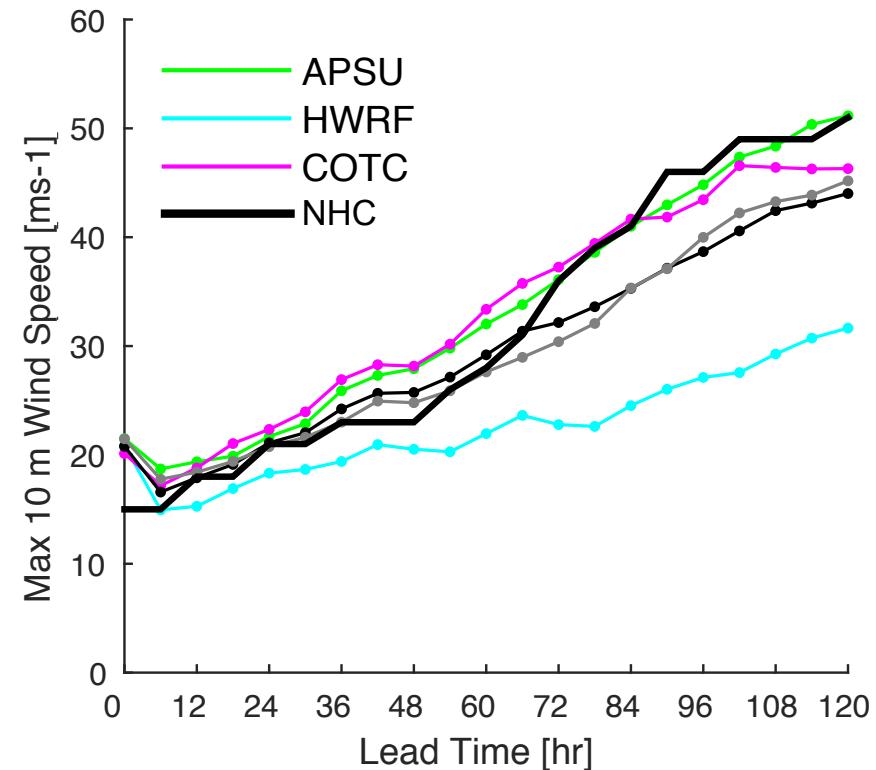
HWRF impacting MCOR  
model performance

In practice, cannot remove the  
worst performing model

**Ensemble Mean Track**



**Ensemble Mean Intensity**



# Hurricane Edouard (2014) Intensity Verification

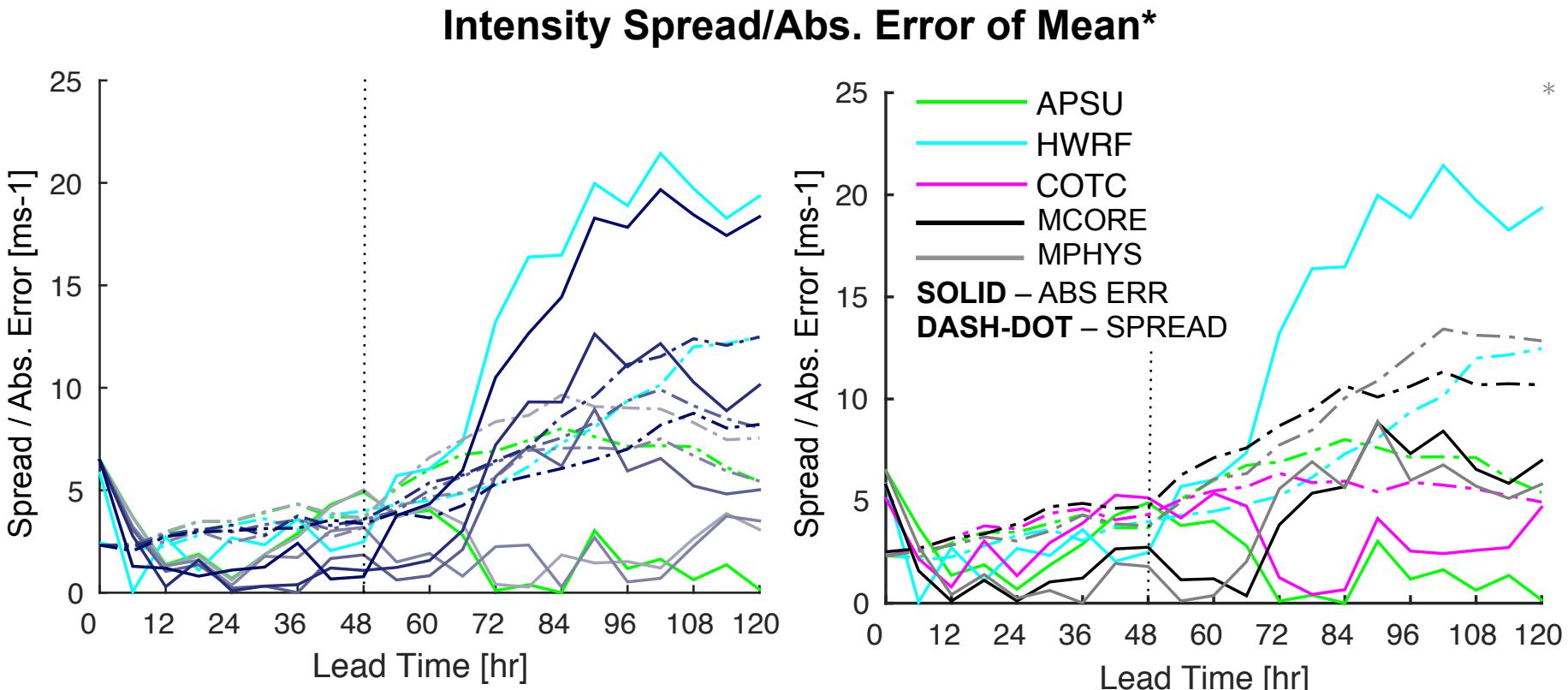
“APSU-Like” Physics

*Modify microphysics, radiation, PBL, surface drag, cumulus*

“HWRF-Like” Physics

Similar spread evolution for multi-core and multi-physics ensembles

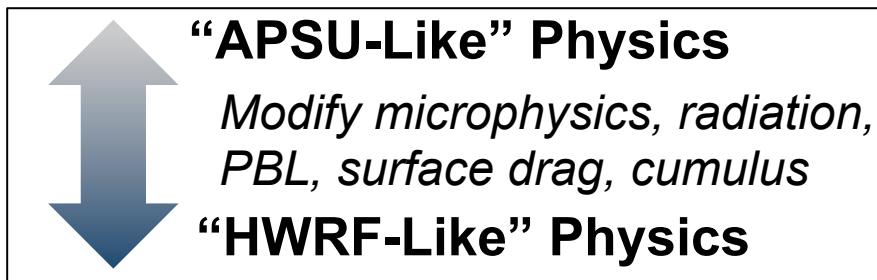
Both multi-core and multi-physics improve mean error within 3 days, but detriment after RI



\*NHC best track data as verification

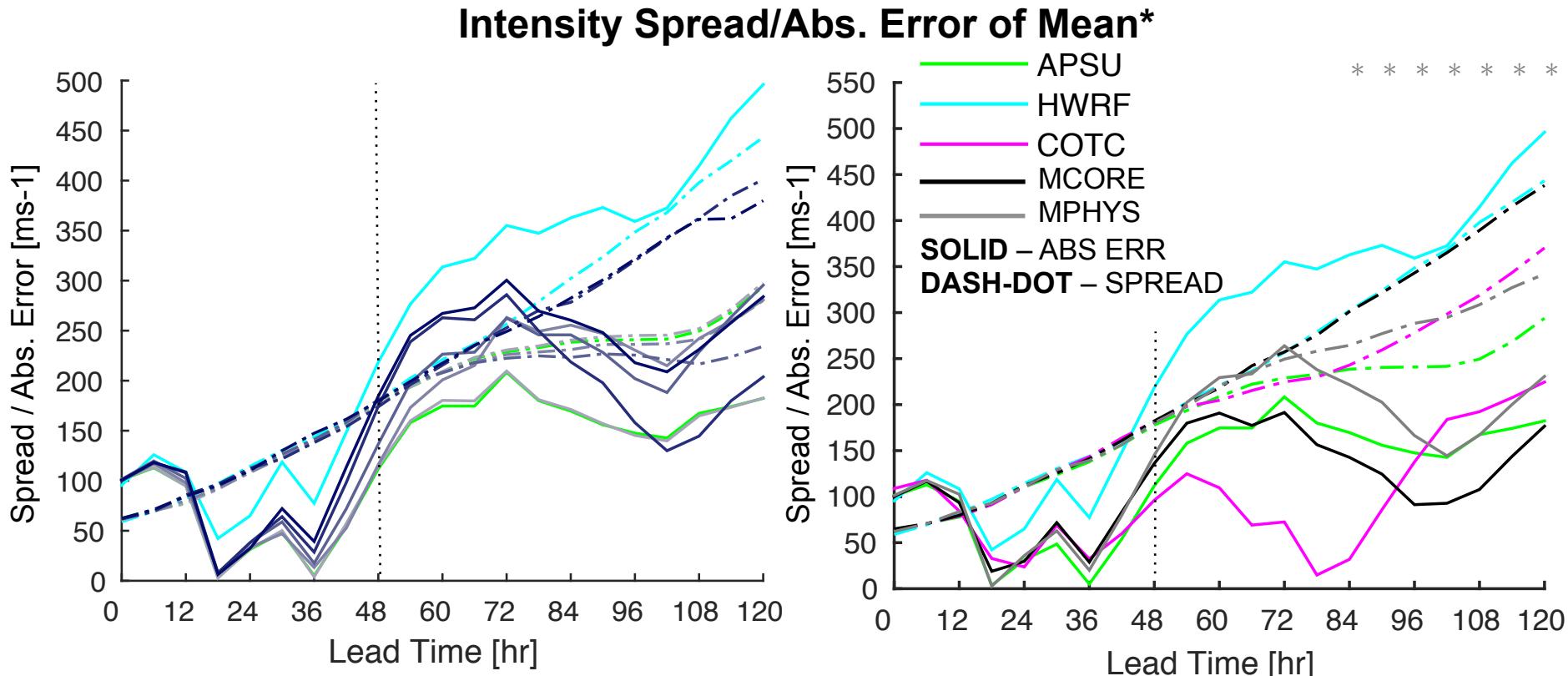
\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

# Hurricane Edouard (2014) Track Verification



Missing solutions similar to COTC  
(further east track displacement)  
in multi-physics ensemble

Lacking broad set of physics  
combinations



\*NHC best track data as verification

\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

# Stochastic Physics Options in WRF-ARW

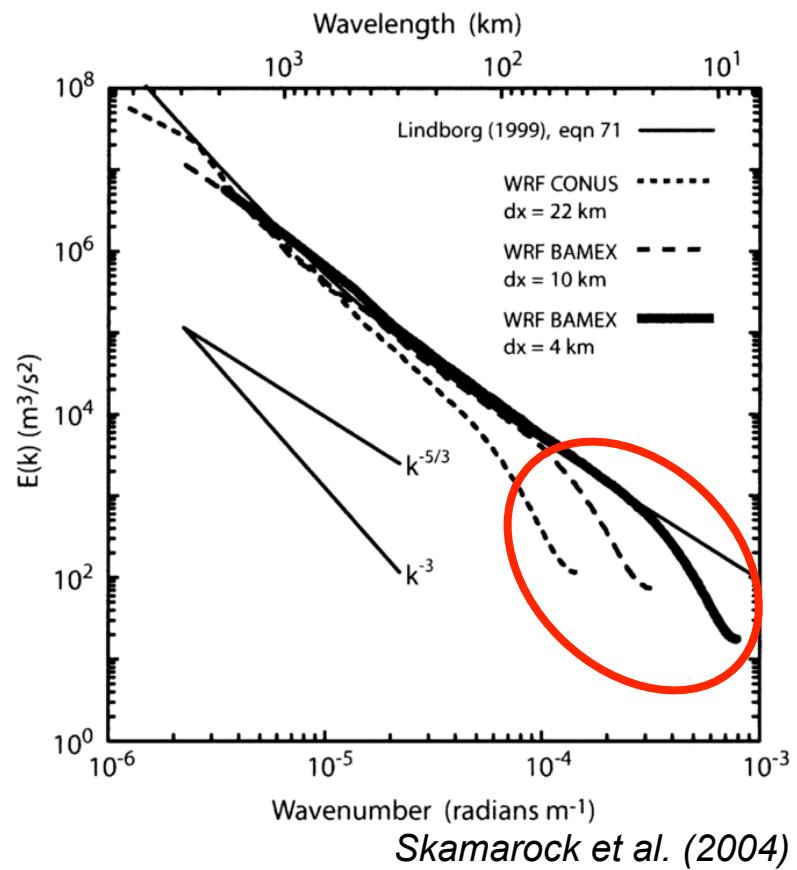
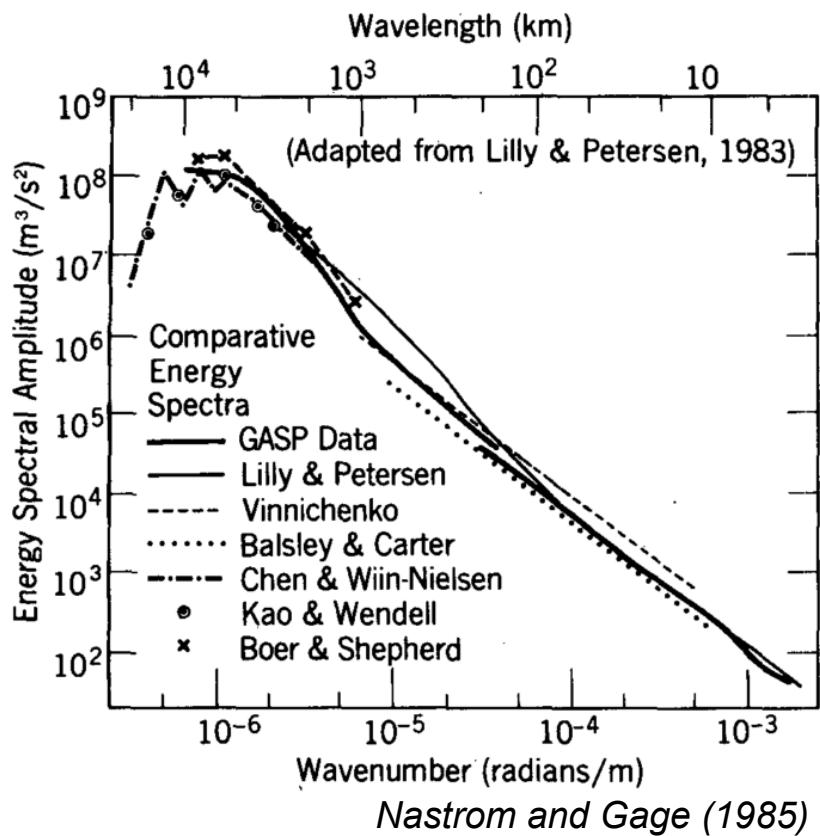
## (1) Stochastic Kinetic Energy Backscatter Scheme (SKEBS; Shutts (2005), Berner et al. 2009)

- Simulate upscale-propagating errors to account for shortcomings in un-resolved subgrid-scale processes
- Perturb rotational u- and v-wind components and potential temperature (additive)
- Spatially and temporally correlated perturbations

## (2) Stochastically Perturbed Parameterization Tendencies (SPPT; Palmer et al. 2009, Berner et al. 2015)

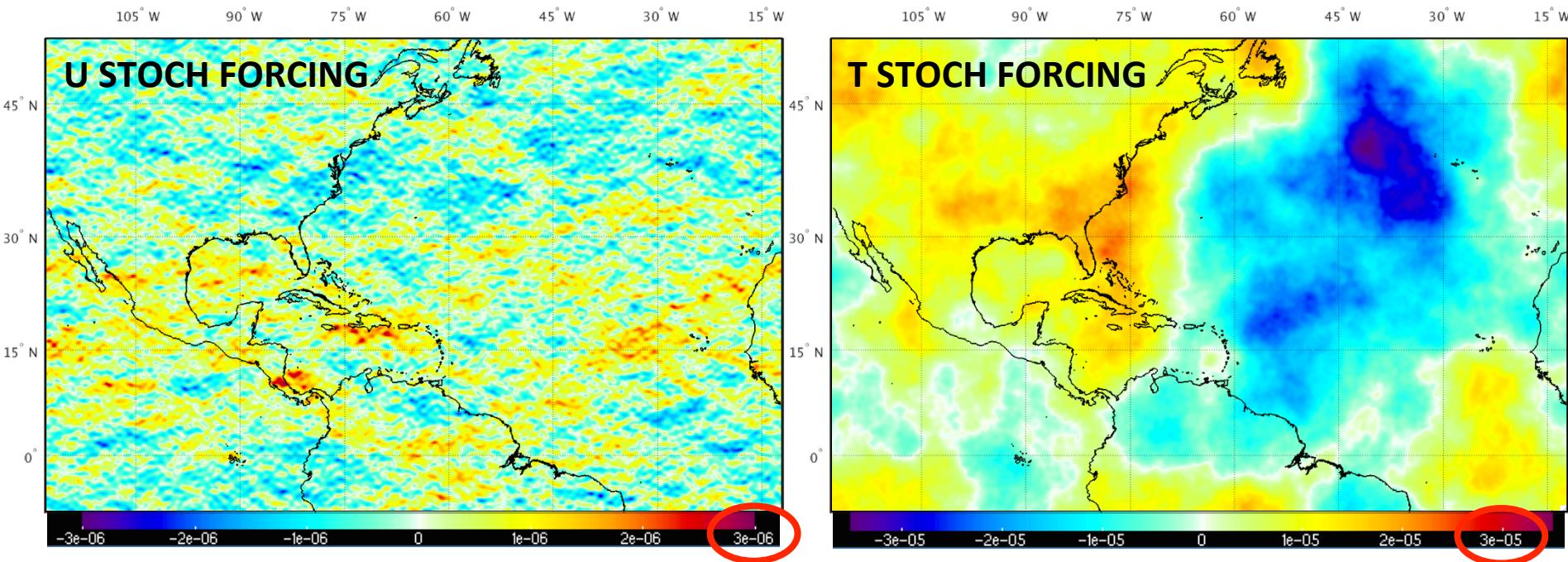
- Account for uncertainties in existing parameterization schemes
- Perturb parameterization tendencies (multiplicative)
- Similar spatial and temporally correlated perturbations to SKEBS

# Observed and WRF Modeled Kinetic Energy Spectra



- Model deficiencies can cause “spectral tails” in kinetic energy power spectrum

# Single Time Step Stochastic Forcing Example



- Can modify spatial decorrelation distance, temporal decorrelation time, and vertical structure of forcing

# Stochastic Physics Options in WRF-ARW

## (1) Stochastic Kinetic Energy Backscatter Scheme (SKEBS; Shutts (2005), Berner et al. 2009)

- Simulate upscale-propagating errors to account for shortcomings in un-resolved subgrid-scale processes
- Perturb rotational u- and v-wind components and potential temperature (additive)
- Spatially and temporally correlated perturbations

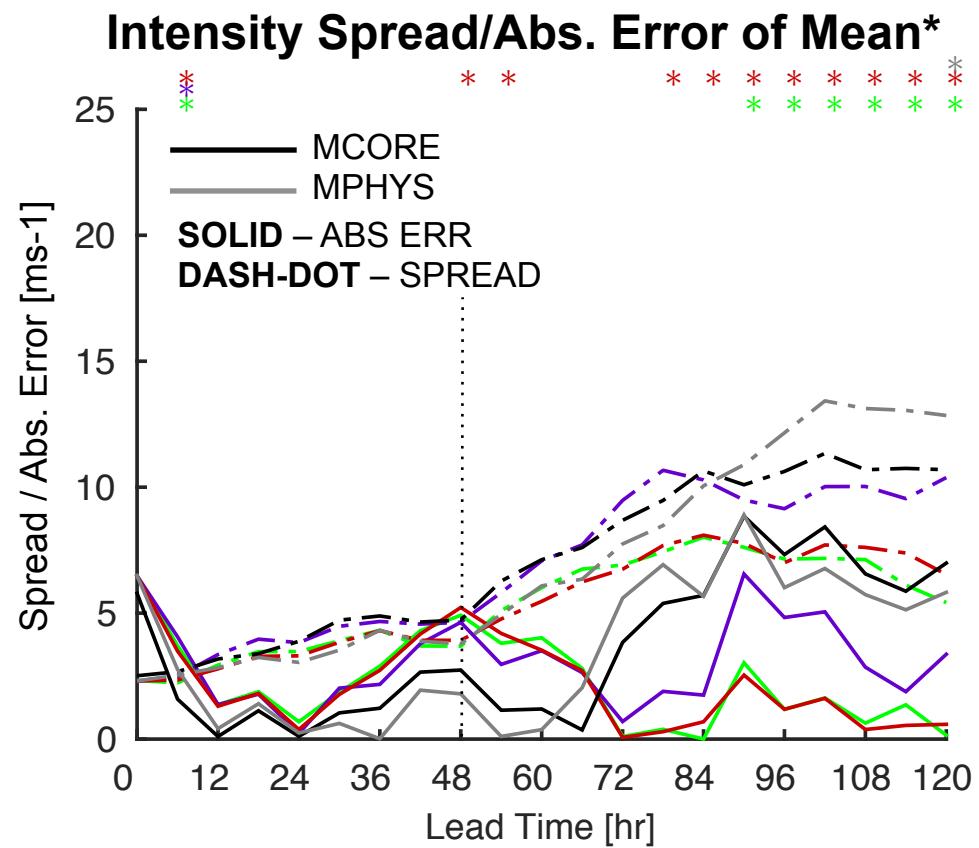
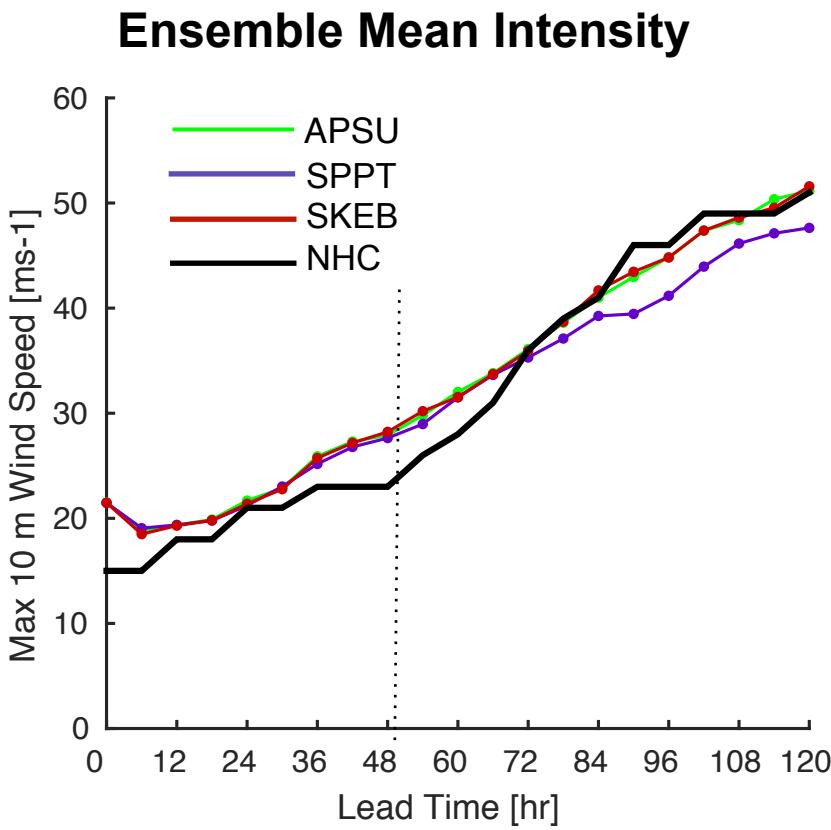
## (2) Stochastically Perturbed Parameterization Tendencies (SPPT; Palmer et al. 2009, Berner et al. 2015)

- Account for uncertainties in existing parameterization schemes
- Perturb parameterization tendencies (multiplicative)
- Similar spatial and temporally correlated perturbations to SKEBS

# Hurricane Edouard (2014) Intensity – Stochastic Physics

SKEBS minimally impacts ensemble mean intensity and intensity spread.

SPPT increases spread, but degrades mean at longer lead times relative to control.



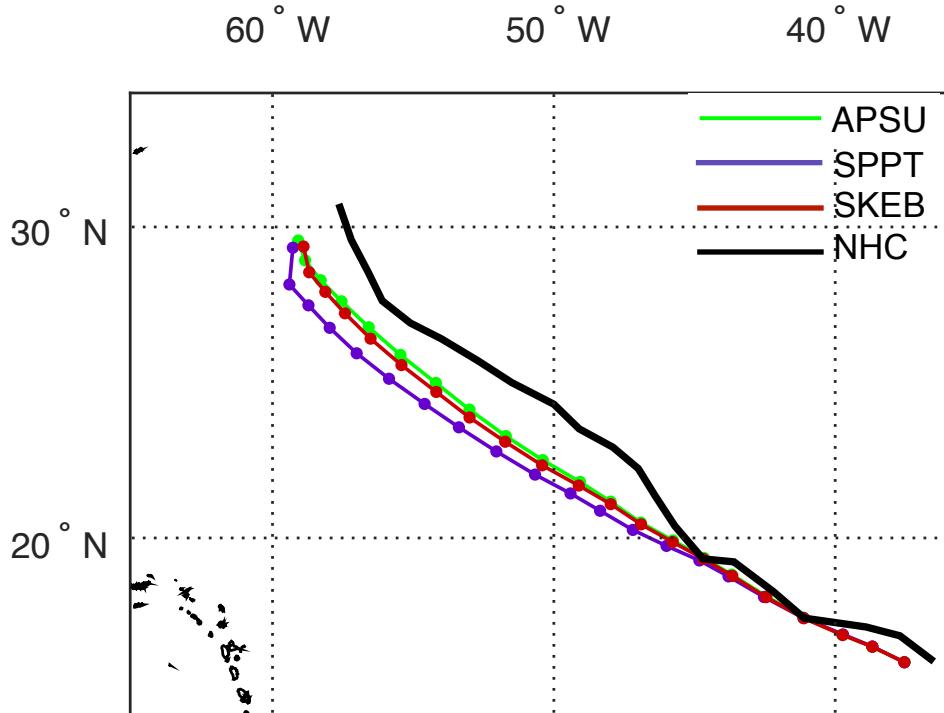
\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

# Hurricane Edouard (2014) Track – Stochastic Physics

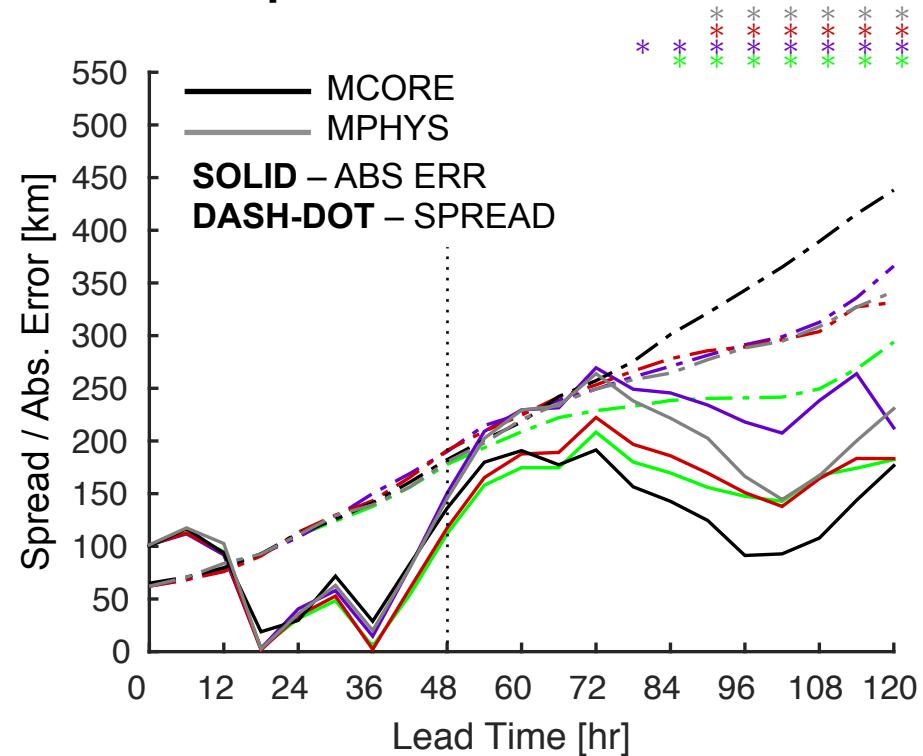
Both SKEB and SPPT can impact track spread  $\sim > 24\text{-}36\text{ h}$

SPPT detrimental to ensemble mean forecast

Ensemble Mean Track



Track Spread/Abs. Error of Mean\*



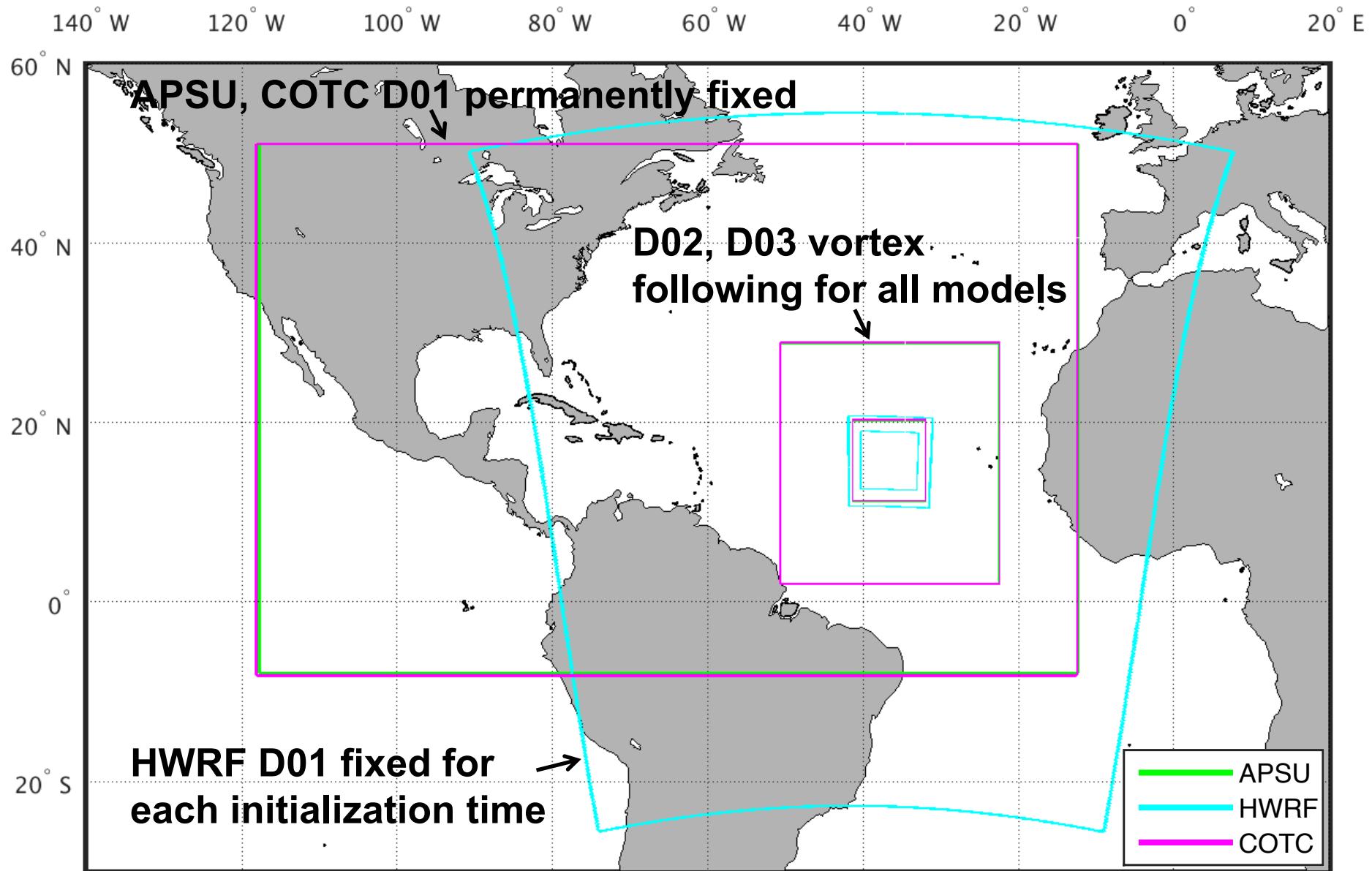
\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

# Discussion

- Recall: initializing three state-of-the-art regional TC-tuned models with the same global resolution IC/BC
- Ensemble track mean and spread generally reproduced by each model-core track and intensity solutions using identical initial perturbations
  - Systematic differences in mean track and intensity evident between model-cores
- Modifying single-core physics can alter mean and spread. Track and intensity error distributions for single-core multi-physics ensemble generally cannot be statistically proven different from multi-core ensemble
  - Stochastic physics can have similar effect
- Single-core multi-physics may be sufficient for TC prediction
- How much spread is sufficient?? Extremely hard to determine!

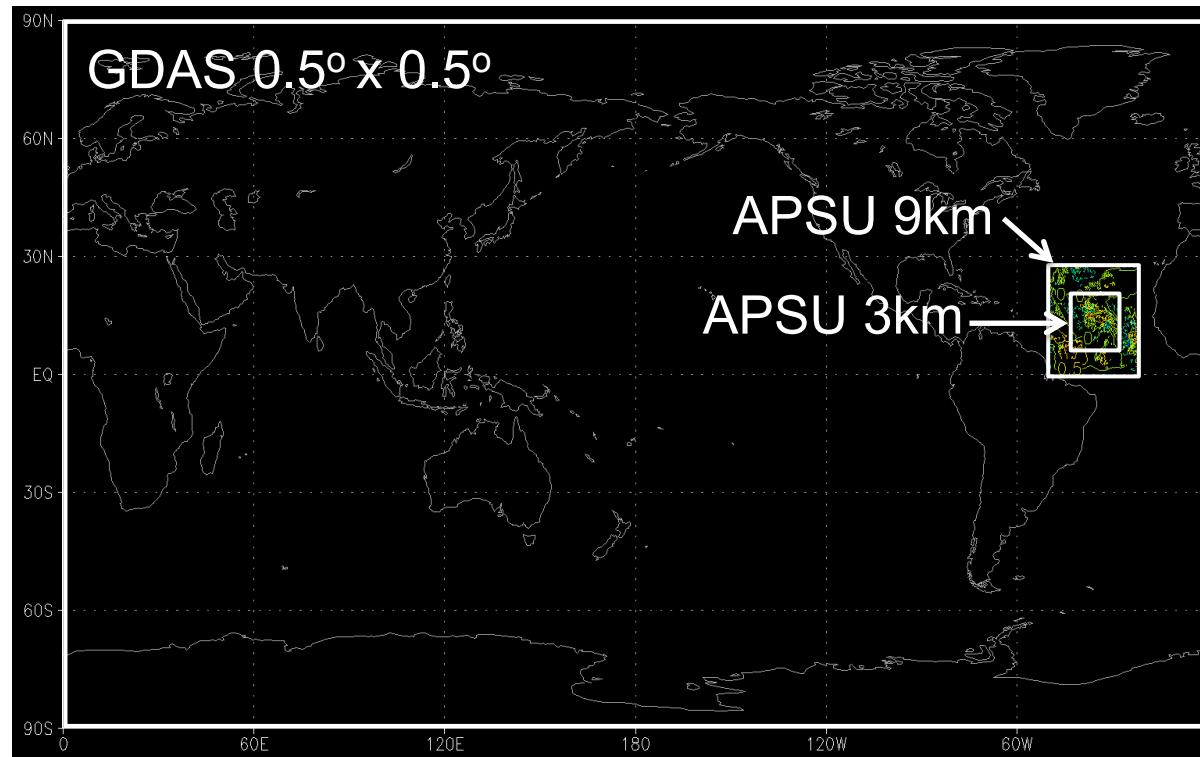
# **SUPPLEMENTARY SLIDES**

# Domain Setup



# Initial and Boundary Conditions

- Initial Conditions:
  - PSU WRF-EnKF (APSU) real-time 60-member ensemble analysis and perturbations
  - Operational GDAS analysis
  - APSU + GDAS interpolated and merged onto  $0.1^\circ \times 0.1^\circ$  grid on standard pressure levels

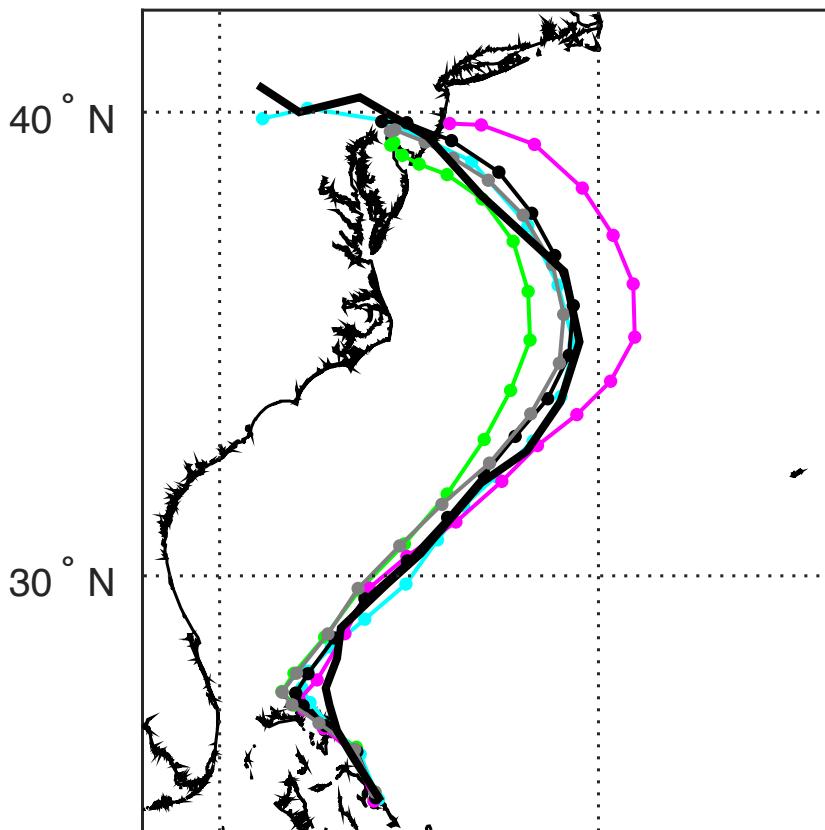


- Boundary Conditions: GFS forecast 6-120 hr (not perturbed)

# Hurricane Sandy (2012) Multi-Model Ensembles

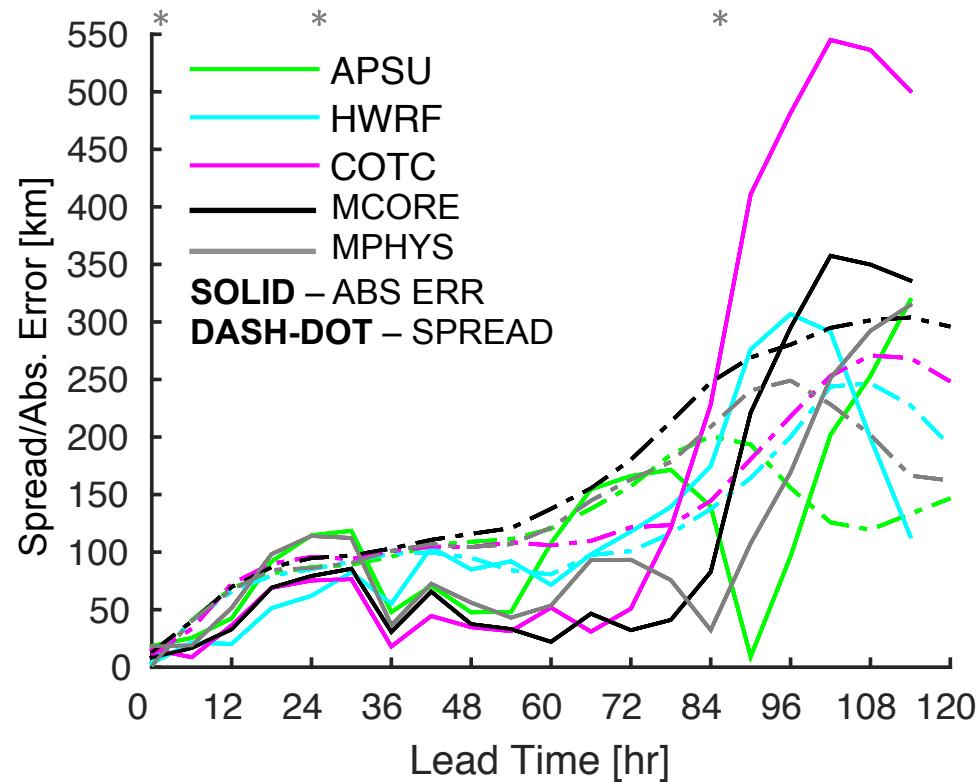
Multi-core and multi-physics ensemble track error distributions cannot be proven statistically different\*\*

Ensemble Mean Track  
80° W      70° W



Increased track spread at longer lead times  
for multi-core or multi-physics

Track Spread/Abs. Error\*

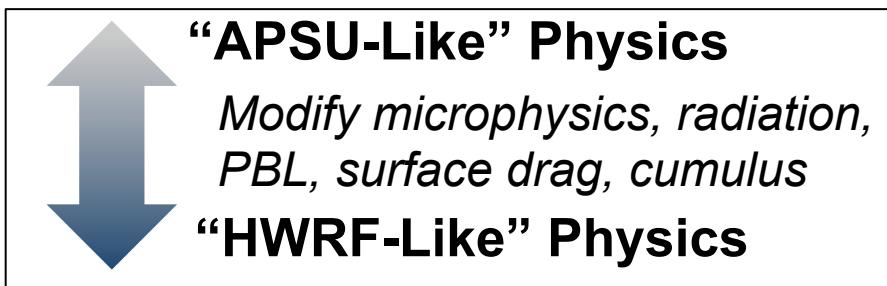


\*NHC best track data as verification

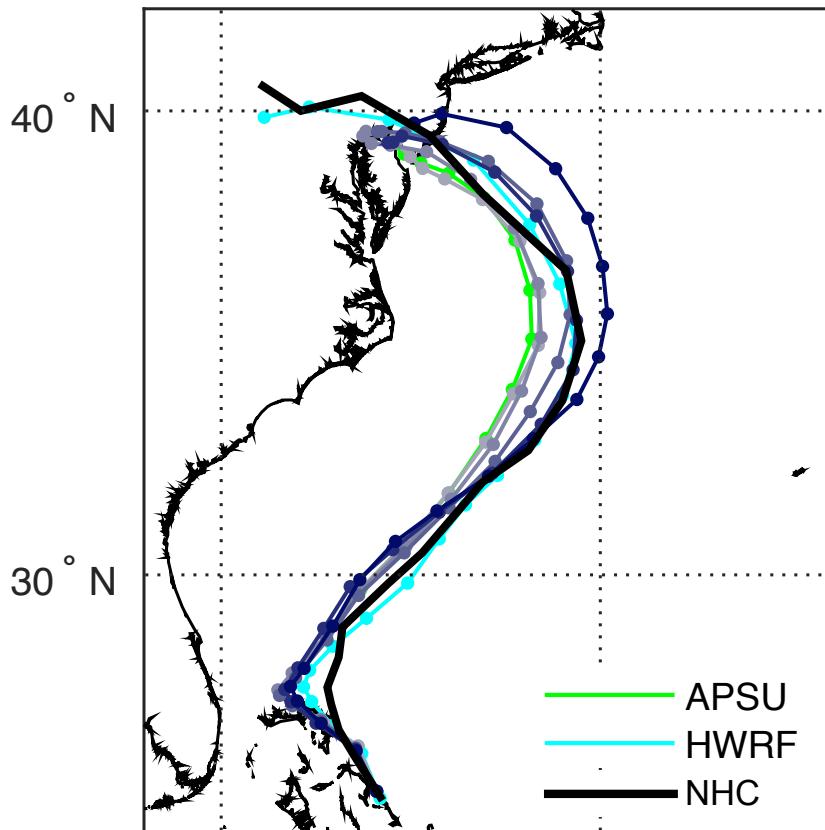
\*\* Bootstrapped Kolmogorov–Smirnov test  
(10,000 samples;  $\alpha=0.05$ )

# Hurricane Sandy (2012)

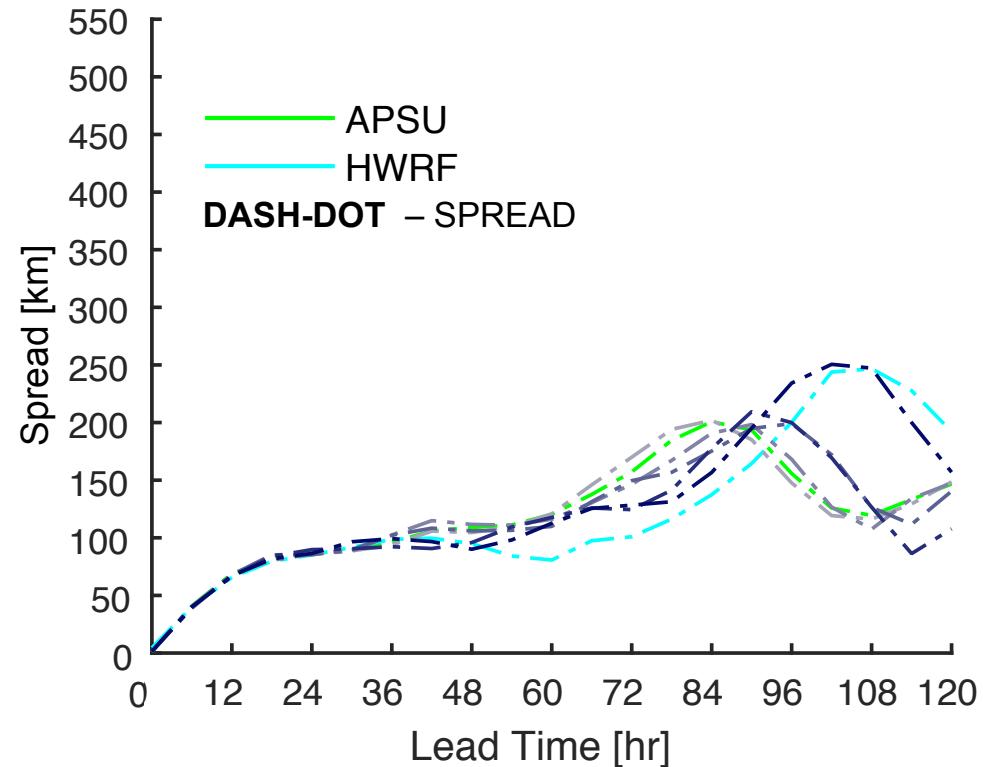
## Multiple Physics Ensembles



Ensemble Mean Track



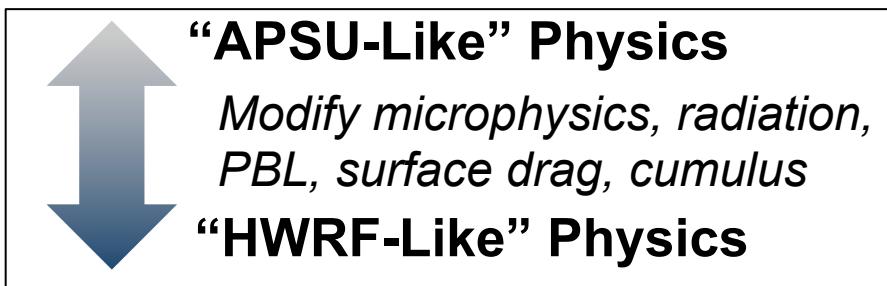
Track Spread\*



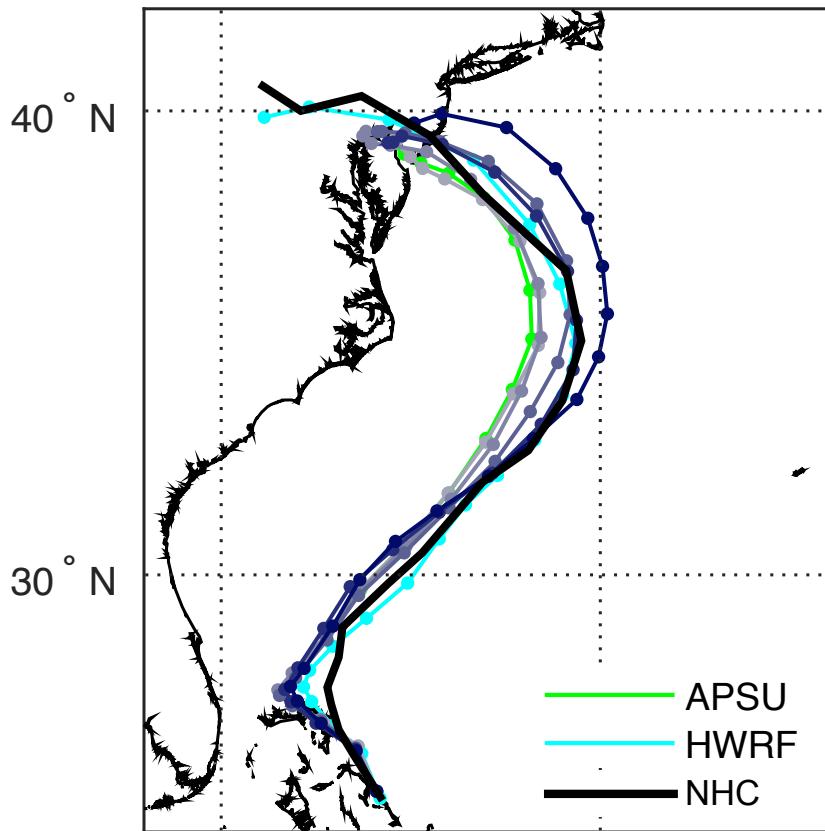
\*NHC best track data as verification

# Hurricane Sandy (2012)

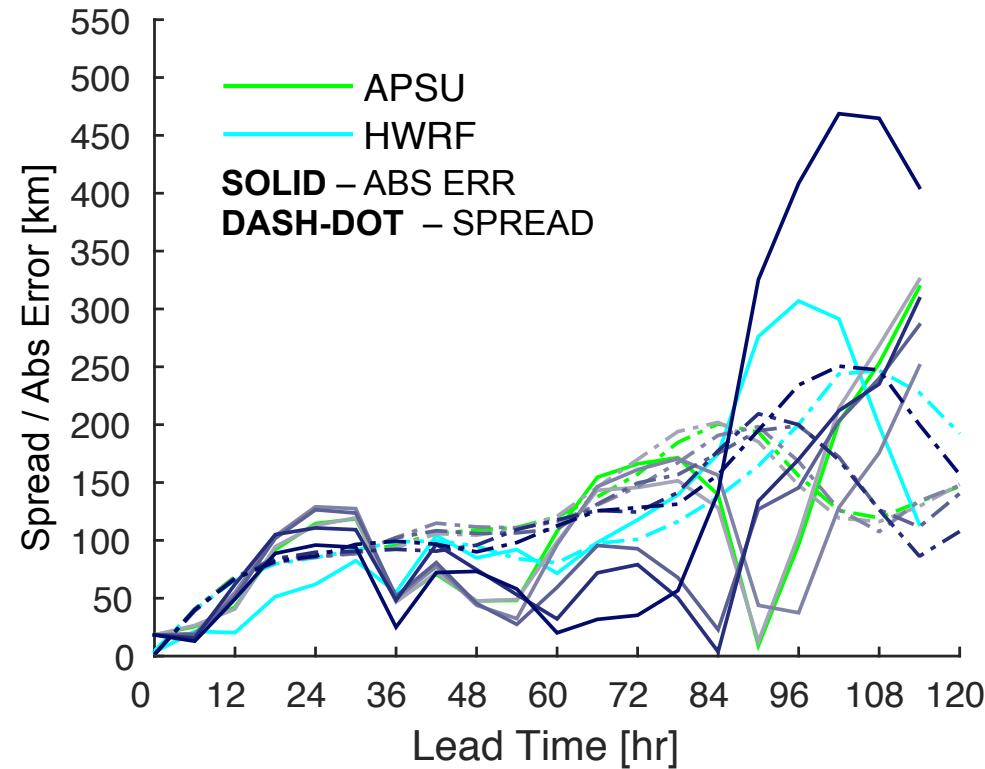
## Multiple Physics Ensembles



Ensemble Mean Track

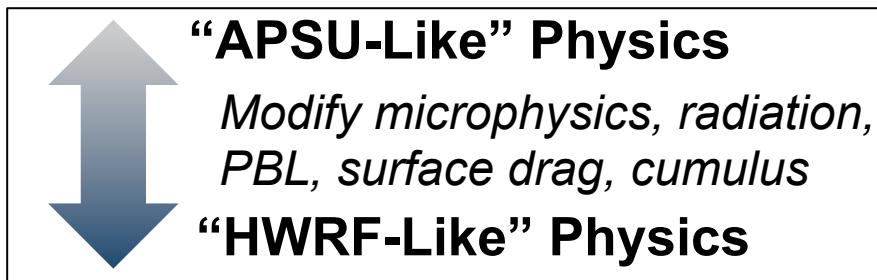


Track Spread/Abs. Error\*



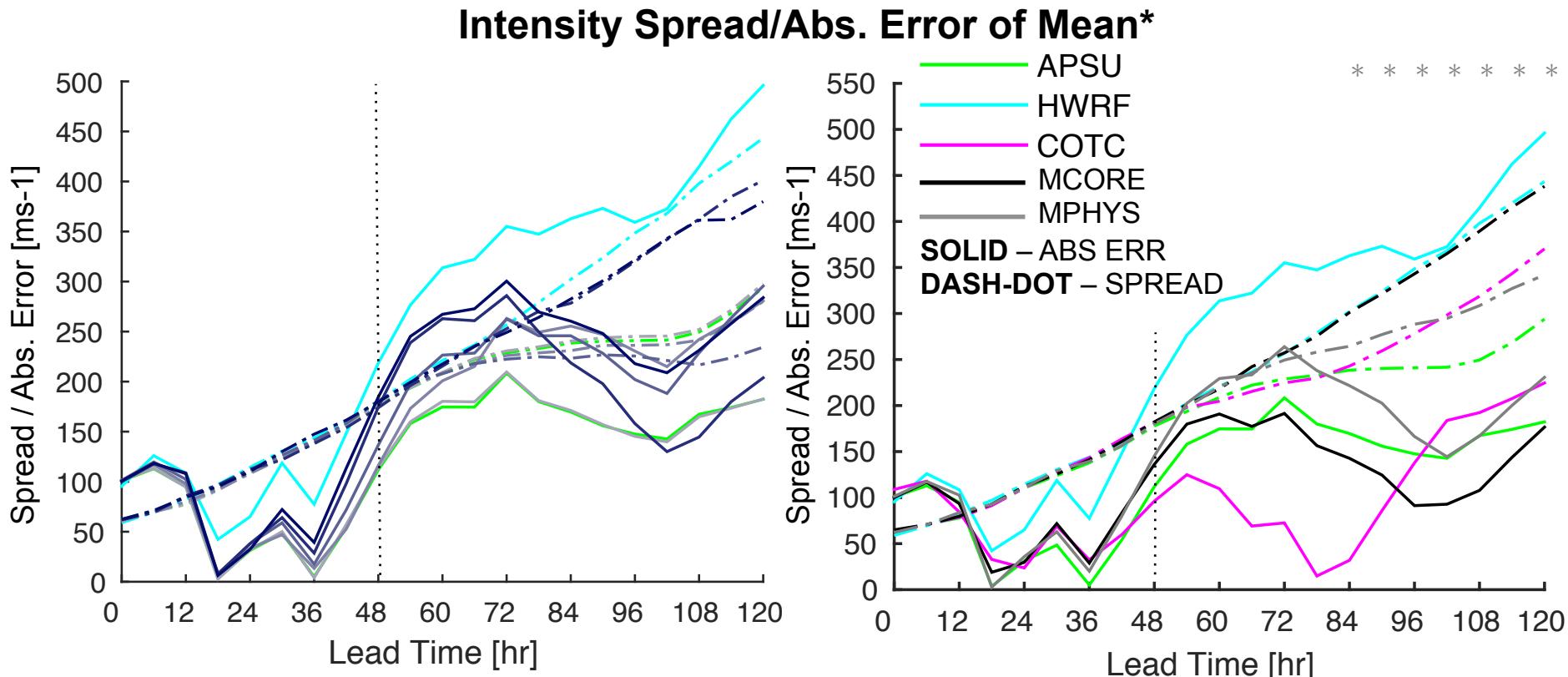
\*NHC best track data as verification

# Hurricane Edouard (2014) Track Verification



Missing solutions similar to COTC  
(further east track displacement)  
in multi-physics ensemble

Lacking broad set of physics  
combinations



\*NHC best track data as verification

\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

# Hurricane Edouard (2014) Intensity Verification

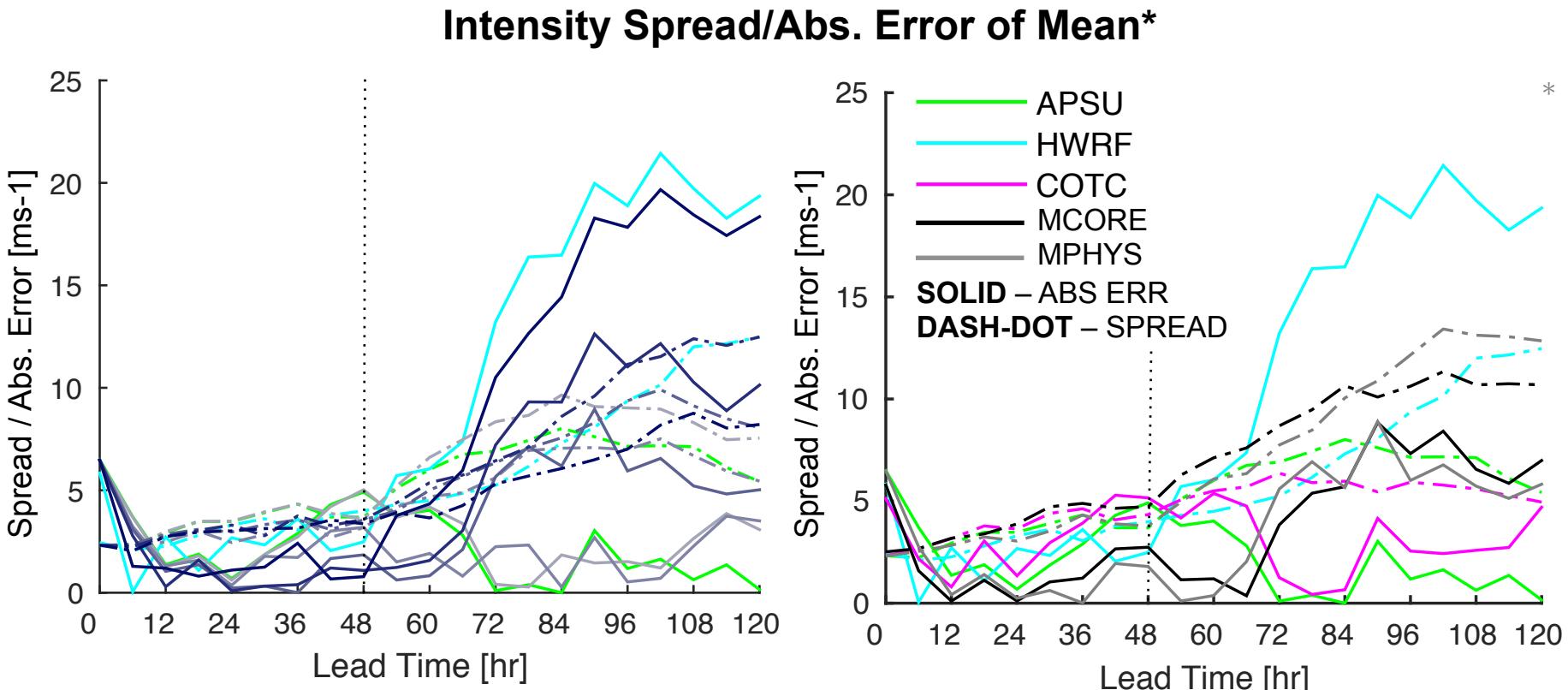
“APSU-Like” Physics

*Modify microphysics, radiation,  
PBL, surface drag, cumulus*

“HWRF-Like” Physics

Similar spread evolution for multi-core and multi-physics ensembles

Both multi-core and multi-physics improve mean error within 3 days, but detriment after RI

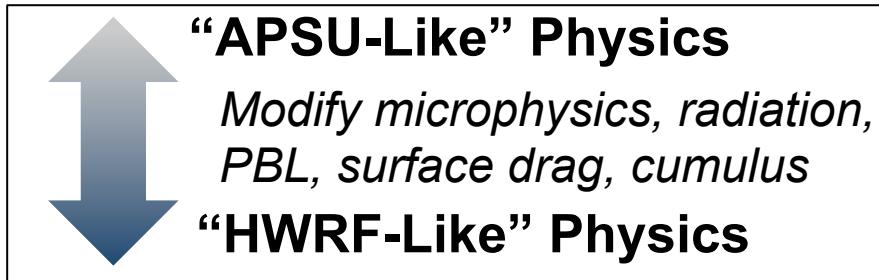


\*NHC best track data as verification

\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

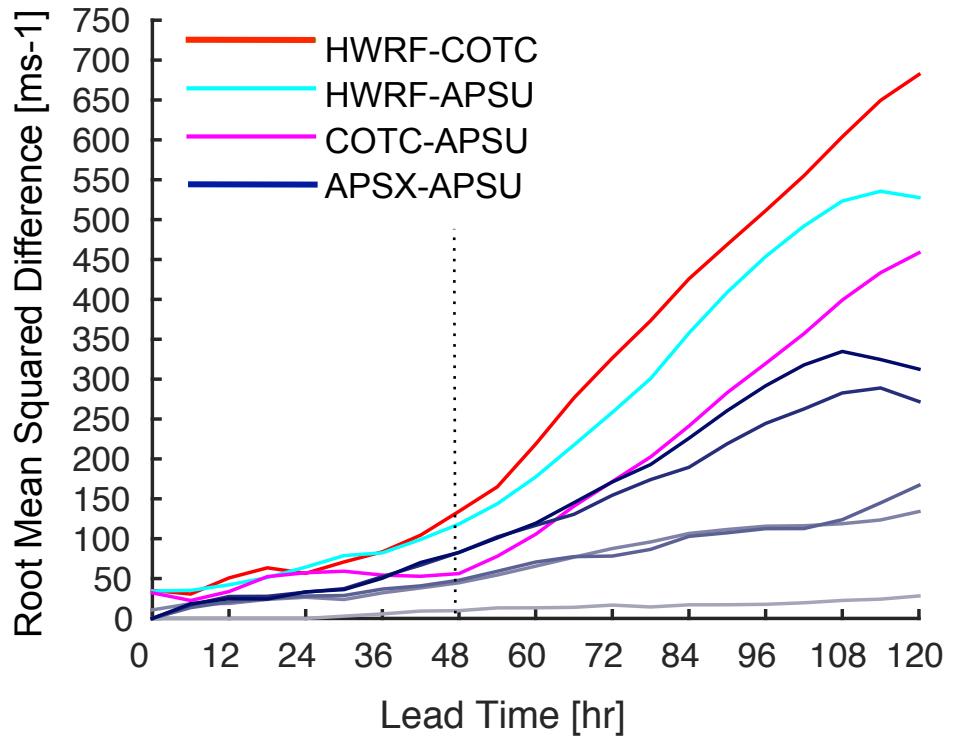
# Hurricane Edouard (2014)

## Model Error Growth

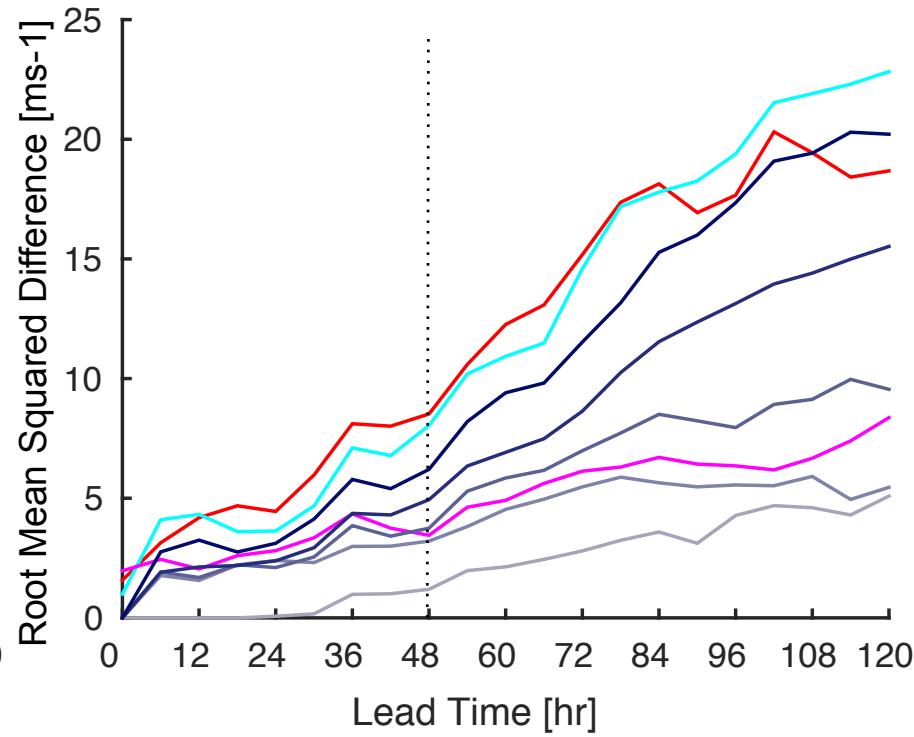


Error growth between models for the same ensemble member is highly dependent on model and/or physics configuration

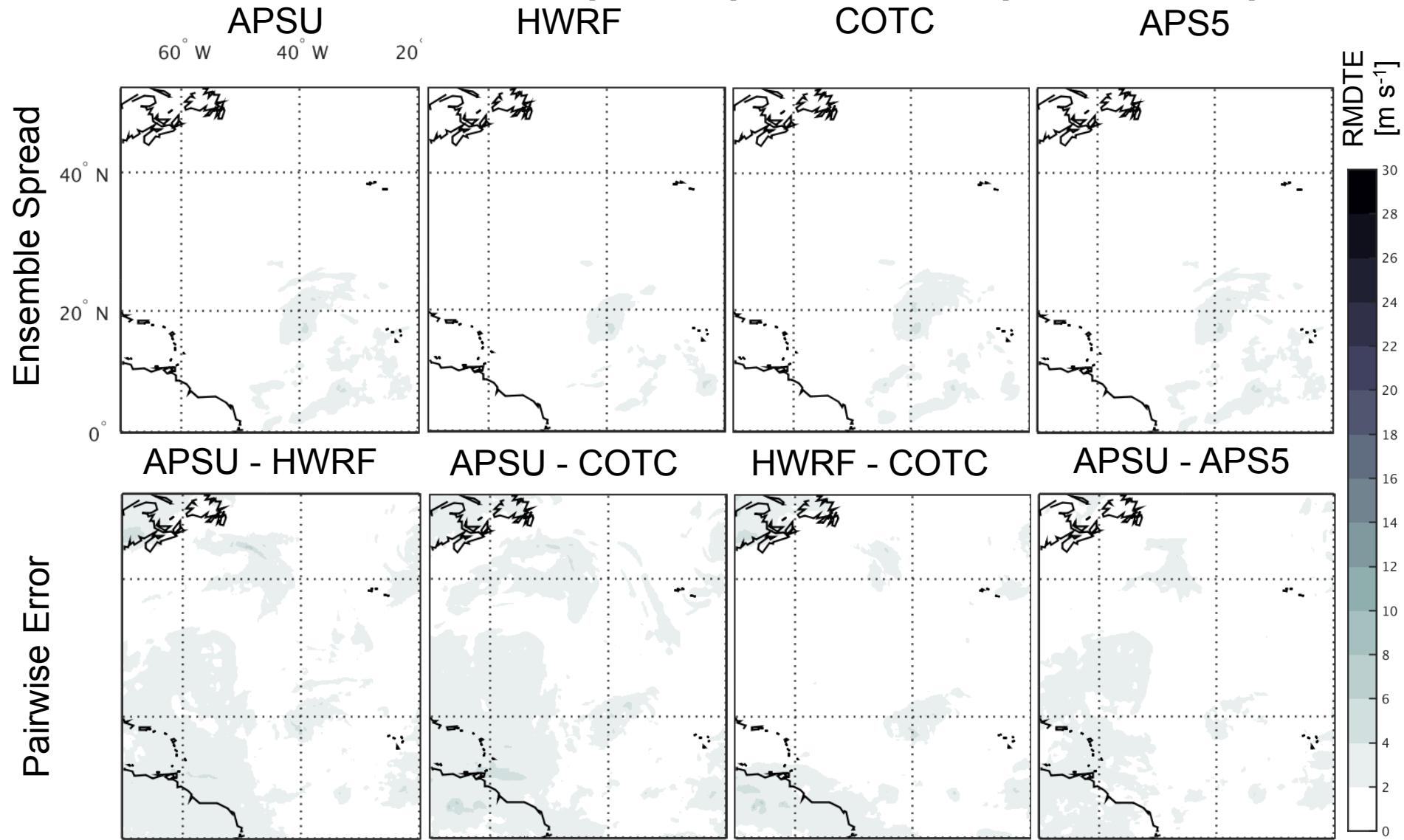
Pairwise Track RMS Difference



Pairwise Intensity RMS Difference



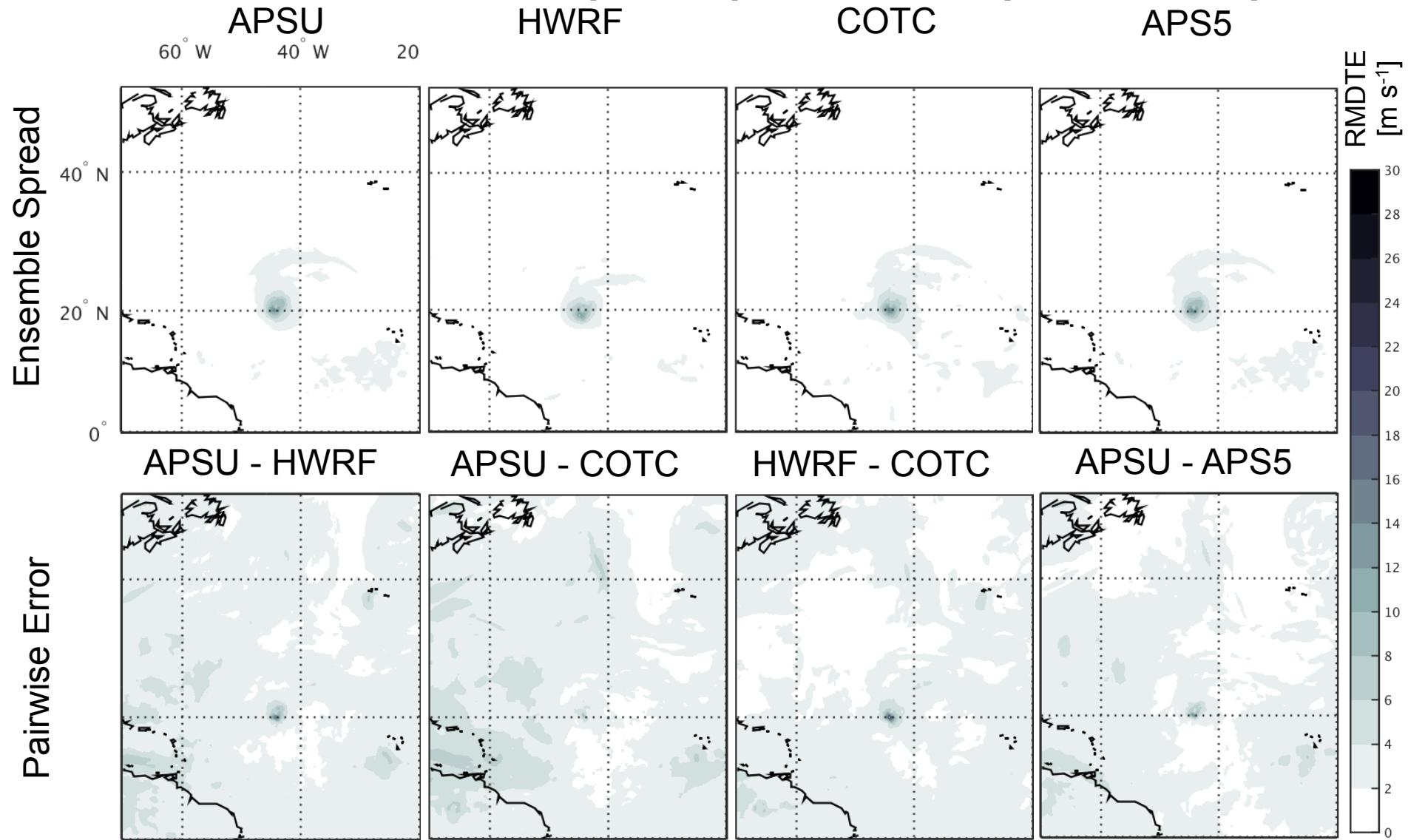
# Hurricane Edouard (2014) RMDTE (FHR 012)



$$RMDTE_{i,j} = \sqrt{\frac{1}{k} \sum_k \frac{1}{n} \sum_n U'_{i,j,k,n}^2 + V'_{i,j,k,n}^2 + \kappa T'_{i,j,k,n}^2}$$

where ' indicates difference  
between experiments and  $\kappa = \frac{c_p}{T_0}$

# Hurricane Edouard (2014) RMDTE (FHR 036)



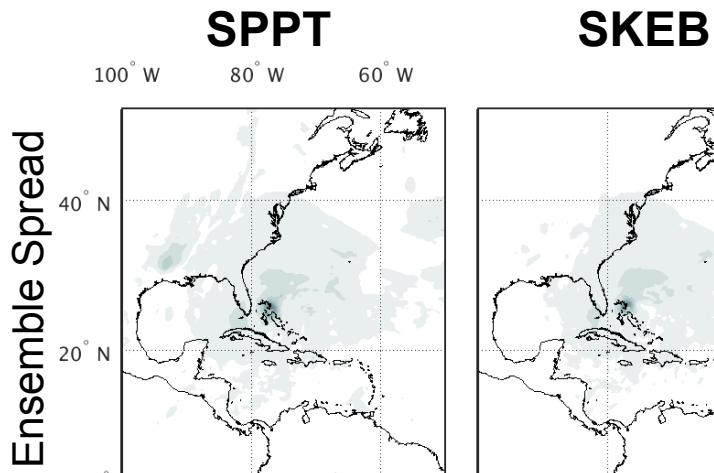
$$RMDTE_{i,j} = \sqrt{\frac{1}{k} \sum_k \frac{1}{n} \sum_n U'_{i,j,k,n}^2 + V'_{i,j,k,n}^2 + \kappa T'_{i,j,k,n}^2}$$

where ' indicates difference  
between experiments and  $\kappa = \frac{c_p}{T_0}$

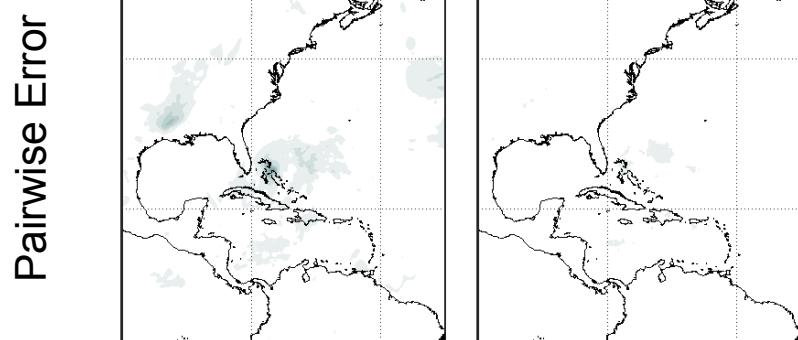
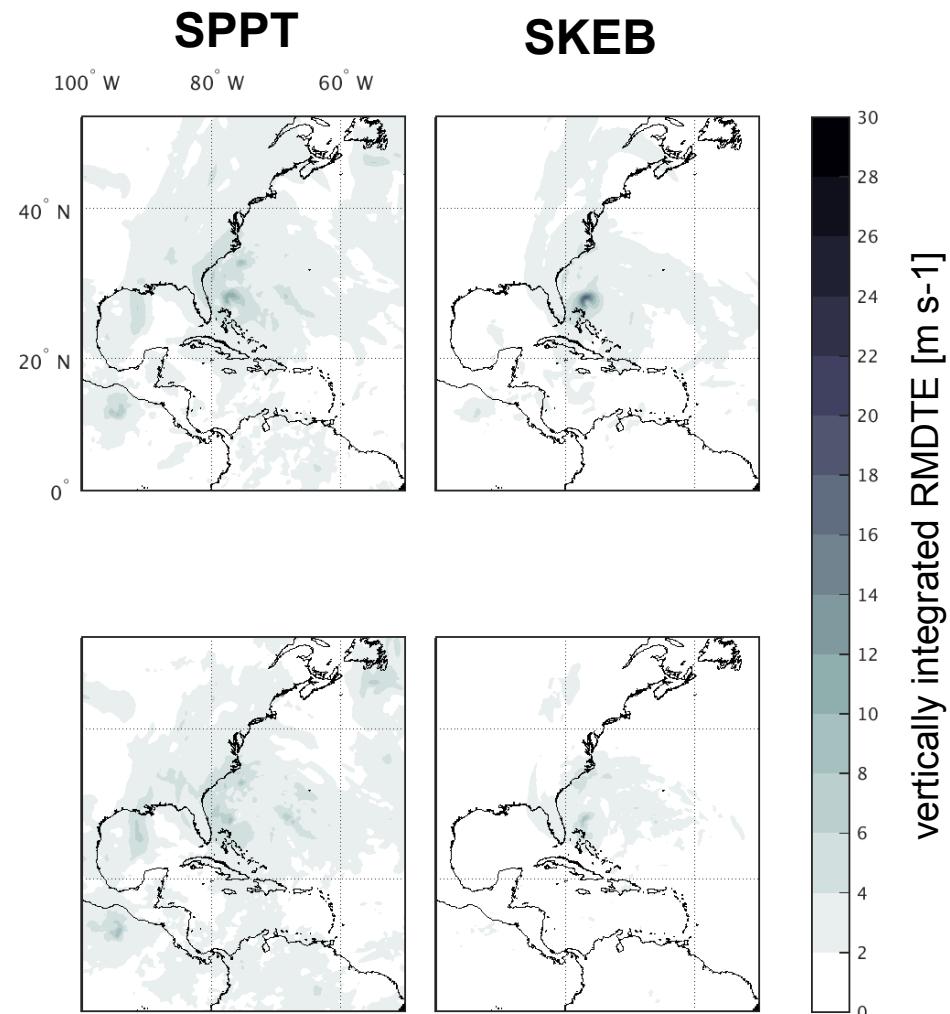
# Horizontal RMDTE (HU Sandy) – Stochastic

Initialized at 2012-10-26 00 UTC

FHR 012



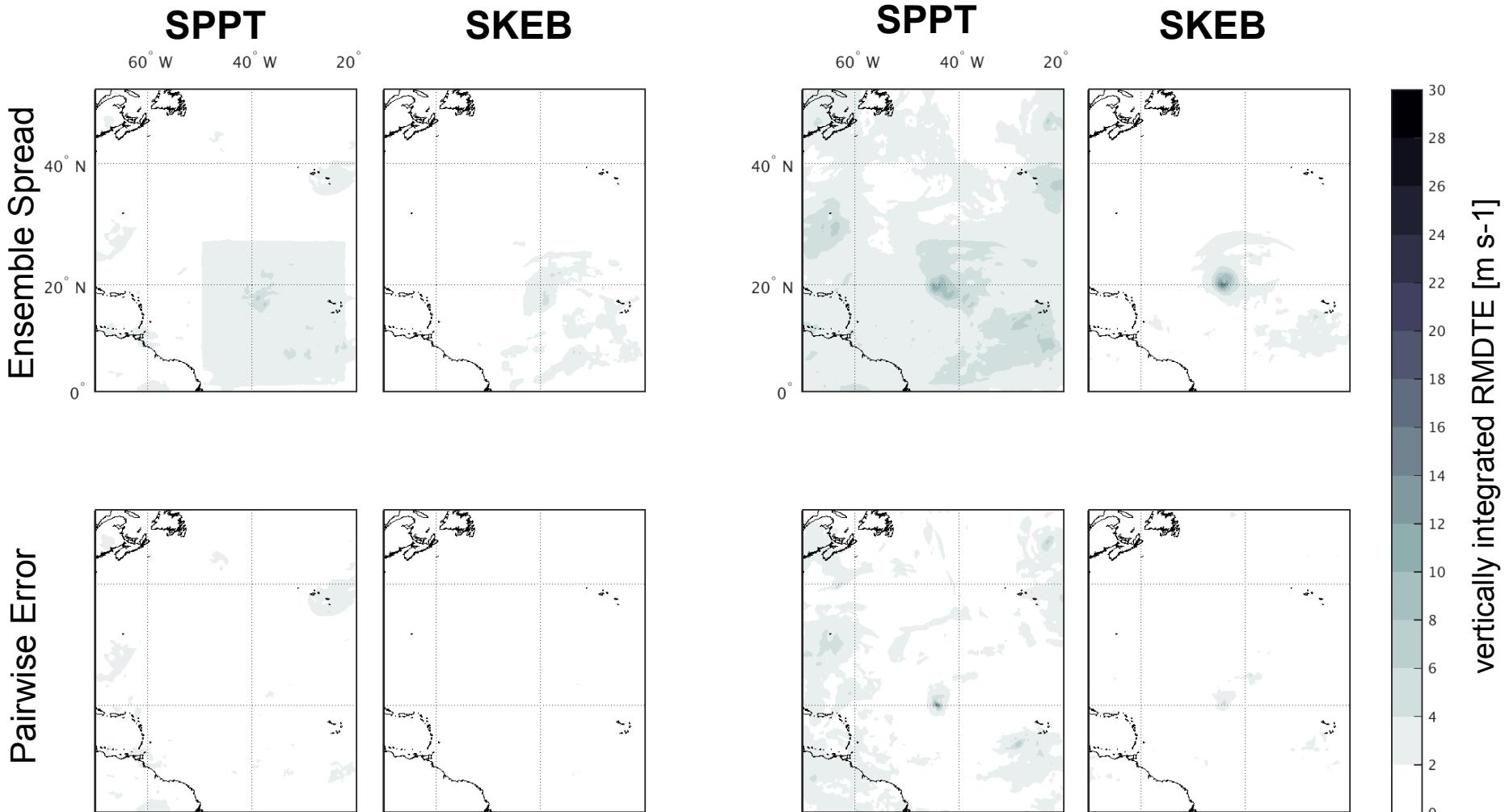
FHR 036



# Hurricane Edouard (2014) RMDTE

FHR 012

FHR 036

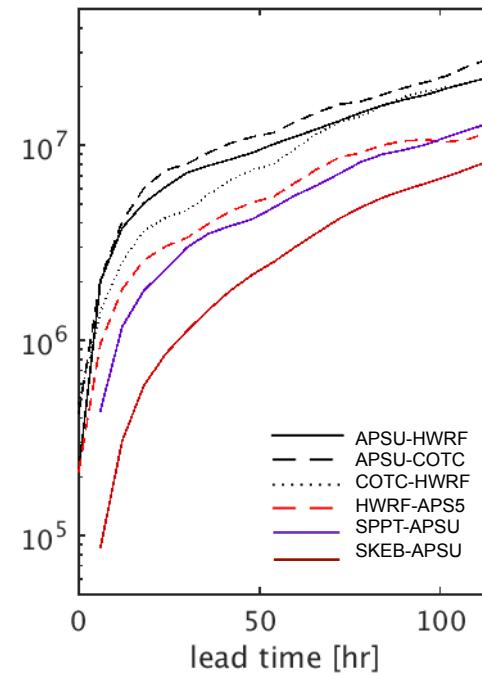
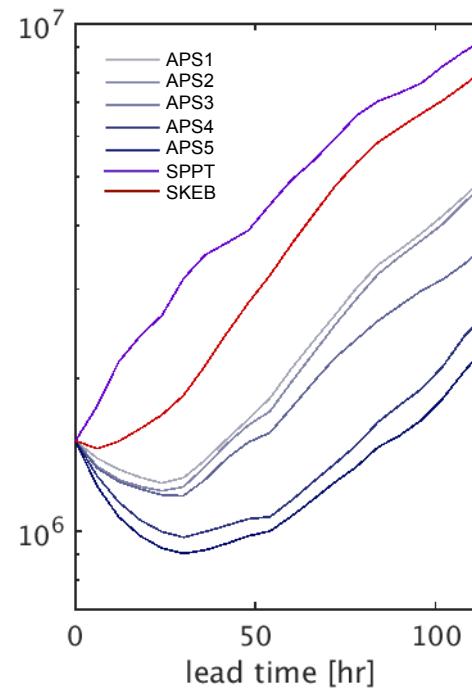
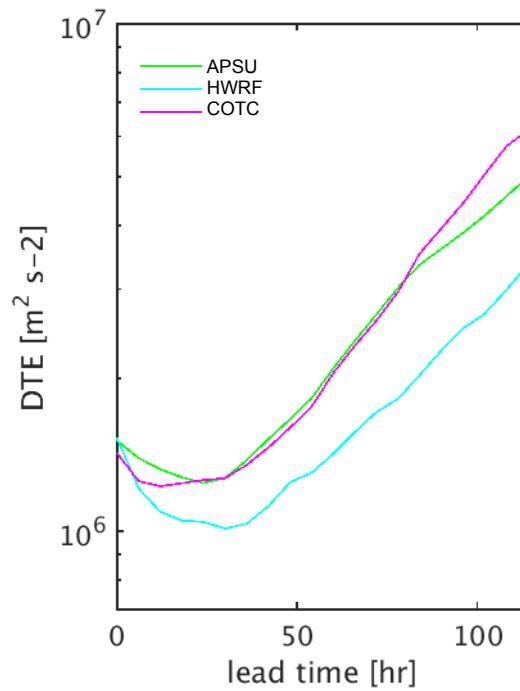


$$RMDTE_{i,j} = \sqrt{\frac{1}{k} \sum_k \frac{1}{n} \sum_n U'_{i,j,k,n}^2 + V'_{i,j,k,n}^2 + \kappa T'_{i,j,k,n}^2}$$

where ' indicates difference  
between experiments and

# Domain integrated DTE (HU Edouard)

6-hrly forecasts initialized at 2014-09-11 12 UTC



Lets keep it simple...

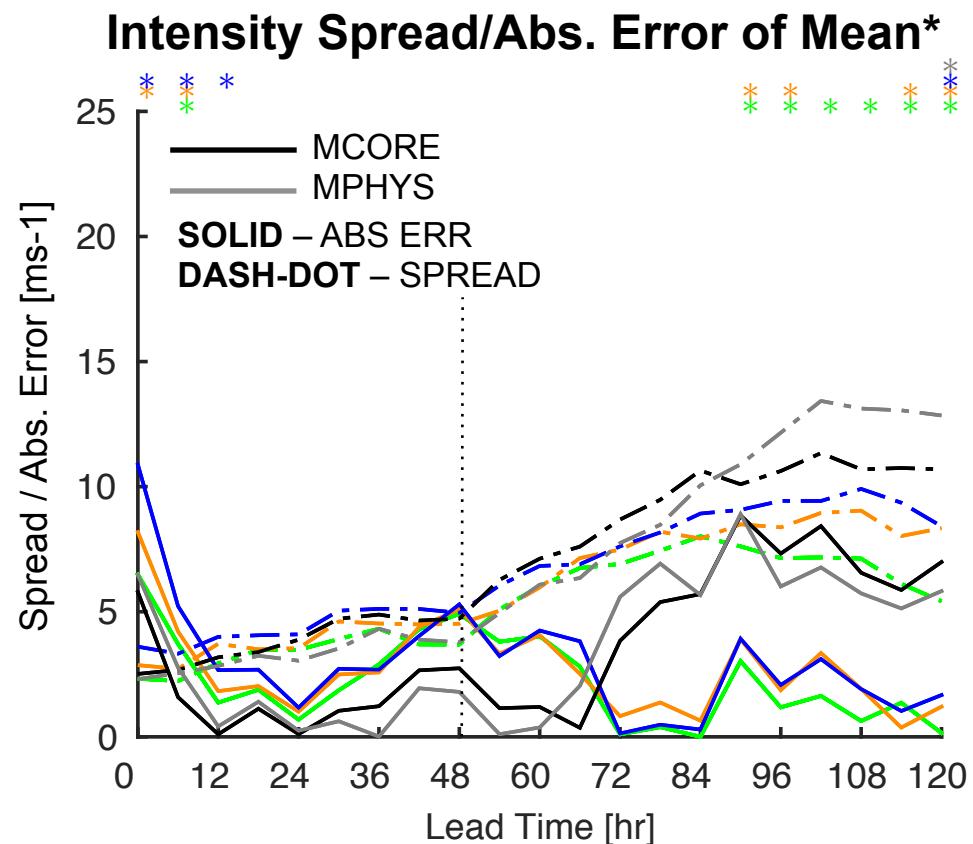
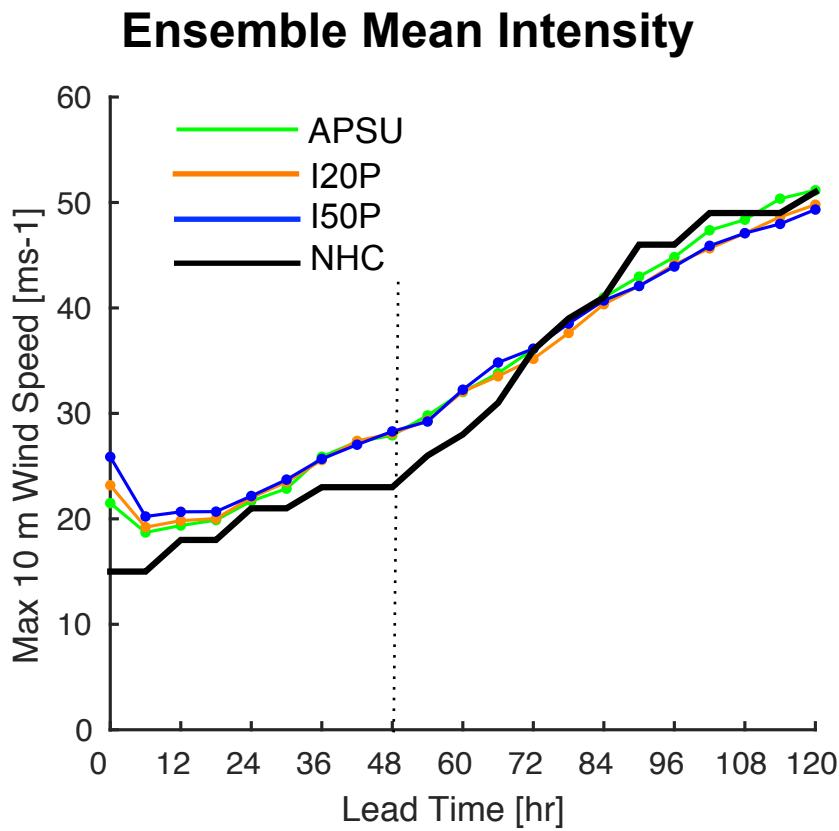
What about just inflating the initial conditions to get the desired spread? 20%? 50%?

# Hurricane Edouard (2014) Intensity – Increased Initial Perturbations

I20P: 20% inflation

I50P: 50% inflation

Inflating initial perturbations can cover the intensity spread of multi-core and multi-physics ensemble while providing lower error and longer lead times



\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

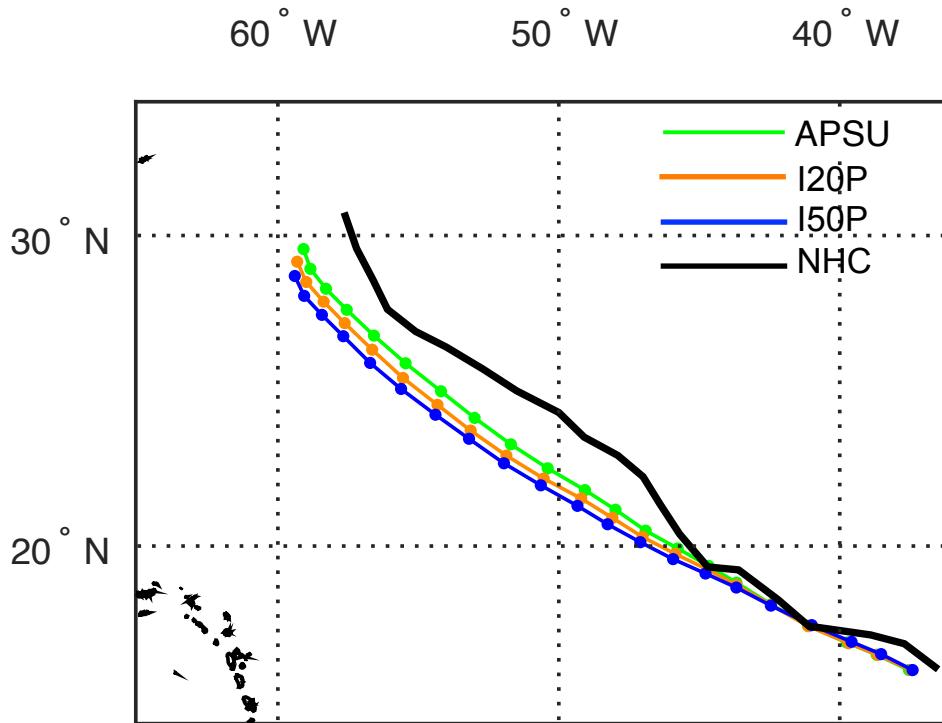
# Hurricane Edouard (2014)

## Track – Increased Initial Perturbations

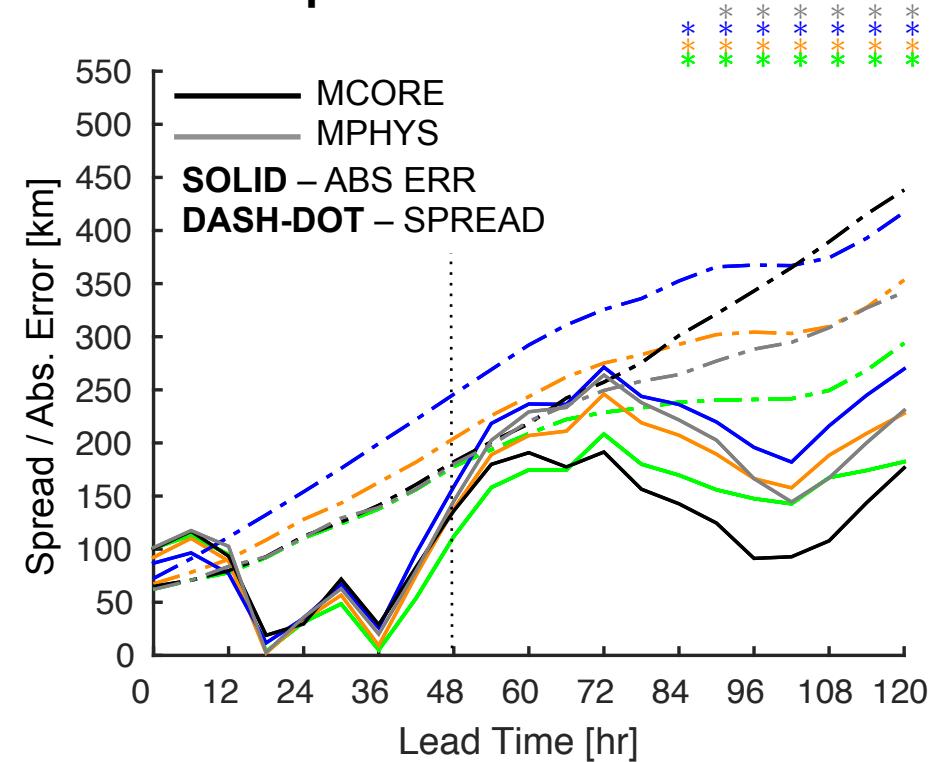
I20P: 20% inflation

I50P: 50% inflation

### Ensemble Mean Track



### Track Spread/Abs. Error of Mean\*



\*\* Bootstrapped K-S test (10,000 samples;  $\alpha=0.05$ )

Table 1

Model	DOMAIN	CP	MP	PBL	RAD	SFC
<b>APSU (2014)</b> (WRF-ARW)	D01: 27km (379x244) D02: 9km (304x304) D03: 3km (304x304)	D01 ONLY Grell-Freitas (Grell et al. 2013)	WSM-6 (Hong and Lim 1996)	YSU (Hong et al. 2006)	Dudhia shortwave (Dudhia 1989) RRTM longwave (Mlawer et al. 1997)	- Modified MM5 similarity (WRF option 91) - PSU formulation surface TC flux (Green and Zhang, 2013) - 5-layer thermal diffusion land surface
	Vertical Levels: 43					
	Model Top: 10 hPa					
<b>HWRF (2013)</b> (modified WRF-NMM)	D01: 0.18 deg (216x432) D02: 0.06 deg (88x170) D03: 0.02 deg (180x324)	D01 & D02 New SAS (HWRF) (Han and Pan 2011)	Tropical Ferrier (Ferrier 2005)	Modified GFS (Hong and Pan 1996); e.g. Gopalakrishnan et al. (2013); Zhang et al. (2013)	GFDL shortwave and longwave (Fels and Schwarzkopf 1981)	- HWRF surface physics - GFDL hurricane slab model land surface (Bob Tuleya 2011)
	Vertical Levels: 43					
	Model Top: 50 hPa					
<b>COTC (2015)</b> (COAMPS-TC)	D01: 27km (379x244) D02: 9km (304x304) D03: 3km (304x304)	D01 ONLY Kain-Fritsch scheme (Kain and Fritsch, 1983)	COAMPS v2 single-bulk (Rutledge and Hobbs, 1983) w/ drizzle	Mellor-Yamada 2.5 scheme (Mellor and Yamada 1982) w/ prognostic TKE	NOGAPS SW/LW (Harshvardhan et al., 1987)	COAMPS surface physics (Louis, 1979)
	Vertical Levels: 40					
	Model Top: ~ 12 hPa					

Table 2

Model	Time Discretization	Spatial Discretization	Prognostic Variables	Advection	Diffusion
<b>APSU (2014)</b> (WRF-ARW)  (see Skamarock et al. 2008 and references therein)	Runge-Kutta 3rd order predictor-corrector scheme (Wicker and Skamarock (2002)) with short time step time-splitting for high frequency acoustic modes	Horizontal: Arakawa C-grid  Vertical: mass + U,V and vertical velocity staggering	U, V, W, perturbation potential temperature, perturbation geopotential, perturbation surface pressure of dry air, TKE, $Q_v$ , $Q_r$ , $Q_s$ , $Q_g$ , $Q_i$ , $Q_c$	6th order accurate for momentum, scalars and geopotential	6 <sup>th</sup> order accurate
<b>HWRF (2013)</b> (modified WRF-NMM)  (see Janjic et al. (2010), Tallapragada et al. (2013, and references therein)	Forward-backward scheme with an implicit scheme for high frequency vertically propagating modes	Horizontal: Arakawa E-grid  Vertical: Lorenz staggering (mass + U,V on consistent levels)	U, V, T, non-hydrostatic pressure, hydrostatic surface pressure, $Q_v$ , $Q_r$ , $Q_i$ , $Q_{ci}$ , $Q_c$	Horizontal: modified Adams-Bashforth, for horizontal advection of u,v, and T, and Coriolis terms,  Vertical: Crank Nicholson for vertical advection of u,v, and T,  Scalars: upstream Lagrangian forward time differencing	2nd order accurate
<b>COTC (2015)</b> (COAMPS-TC)  (see Hack (1996), Chen et al. (2003), and references therein)	Centered-in-time (i.e. leap frog) 2 <sup>nd</sup> order scheme with short time step time-splitting for high frequency acoustic modes	Horizontal: Arakawa C-grid  Vertical: mass + U,V and vertical velocity staggering	U, V, W, $\theta$ , $\pi$ , TKE, $Q_v$ , $Q_r$ , $Q_i$ , $Q_g$ , $Q_s$ , $Q_c$	2 <sup>nd</sup> order accurate upstream, forward-in-time advection	4 <sup>th</sup> order accurate

Table 3

Experiment Name	CP	MP	PBL	RAD	SFC
<b>ALT1</b> APSU w/mod TC SFC flux	D01 ONLY Grell-Freitas	WSM-6	YSU	Dudhia shortwave RRTM longwave	Same to APSU w/ WRF TC surface flux (Garratt formulation)
<b>ALT2</b> APSU w/mod TC SFC flux, MP	D01 ONLY Grell-Freitas	Eta (Ferrier)	YSU	Dudhia shortwave RRTM longwave	Same as ALT2
<b>ALT3</b> APSU w/mod TC SFC flux, MP, RAD	D01 ONLY Grell-Freitas	Eta (Ferrier)	YSU	GFDL shortwave/ longwave	Same as ALT2
<b>ALT4</b> APSU w/mod TC SFC flux, MP, RAD, CP	D01 & D02 New SAS (HWRF)	Eta (Ferrier)	YSU	GFDL shortwave/ longwave	Same as ALT2
<b>ALT5</b> “HWRF-LIKE”	D01 & D02 New SAS (HWRF) (Han and Pan 2011)	Eta (Ferrier) (Rogers, Black, Ferrier, Lin, Parrish and DiMego 2001)	GFS (Hong and Pan 1996)	GFDL shortwave/ longwave (Fels and Schwarzkopf 1981)	Same as ALT2