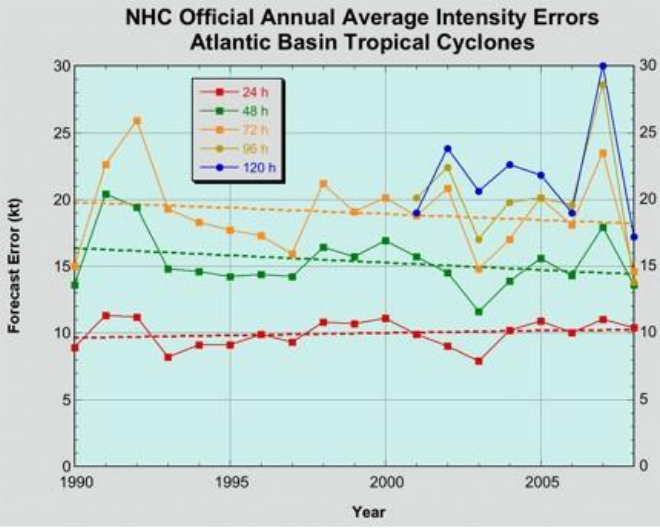
Statistical Approaches to Hurricane Intensity Prediction: A Historical Overview of the SHIPS Model

Greg Herman

The Problem

- Accurate short and medium-range prediction of hurricanes and other high-impact meteorological phenomena is critical to the protection of lives and property
- Three approaches to forecasting:
  - Dynamical Models: Use estimate of the current state of the atmosphere and equations governing the atmosphere's evolution to predict future states of the atmosphere
  - Statistical Models: Use current observations in conjunction with a record of past observations and their corresponding outcomes to generate predictions about the future state of the atmosphere
  - Combined Dynamical/Statistical: Use both current observations *and* output from a dynamical model or models to in conjunction with a record of past observations and/or model predictions to generate the current forecast



Dynamical Models

- Synonymous with NWP Models; examples include GFS, NAM, ECMWF, GEM, WRF, RAMS, NOGAPS, HWRF, etc.
- Using estimate of current state of atmosphere and physical laws governing the atmosphere's evolution, simulate the atmospheric state in the near-term future
- Dynamical models introduce inherent bias
  - Insufficient resolution leads to features not being (adequately) resolved
  - Parameterizations and other model details also contribute to systemic bias
- Also generally quite computationally expensive

Statistical Models

- Unlike dynamical models, statistical models do not model the state of the atmosphere at all
- Instead, use current and recent analyses of atmospheric state, along with climatology, to predict future states of atmospheric variables based on historical relationships
- Much faster/cheaper
- Does not require as significant an understanding of the phenomenon being predicted
- ‘Non-physical’

Combined Dynamical/Statistical Approaches: Post-Processing

- Take output from dynamical model and apply to statistical model
- Looks for historical relationships between model fields and observed atmospheric variables of interest in addition to traditional statistical model predictors
- Correct bias of dynamical models
- Can provide downscaling/resolution for relatively coarse model
- Still ‘non-physical’

Calibration Methods

- Bias Correction (Shift & Stretch)
- **Multivariate Regression**
- Terrain Mapping
- Quantile Mapping
- Bayesian Model Averaging
- Analogs
- Machine Learning
- Neural Networks
- Many more…

Multivariate Regression: Introduction

- Idea: Want to know some variable
- $u(x_i, y_j, t_z) = u_{ijz} = y_{u_{ijz}}$
- Instances: m records of form $< y_{u_{ijz}}, [1, x_1, \ldots, x_n] >$
- m = number of records (length of time)
- n = number of predictors
- Combine to form:
- $\boldsymbol{y}_{\boldsymbol{u}_{ijz}} = \boldsymbol{\beta X} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\beta}$ is a weights vector of length n, and ε is the error term vector of length m
- In *multivariate* linear regression, you have *k* instances of these *m* records corresponding to *k* predictands

Please note: Vector notation throughout largely dropped for simplicity. Unless otherwise noted, all vector products denote the *inner* product

Multivariate Regression: Computing $\boldsymbol{\beta}$

- Many possible methods to estimate optimal coefficients for $\boldsymbol{\beta}$
- Simplest and most popular method is Ordinary Least Squares (OLS)
- Idea: Minimize total (summed) squared error (residuals) of predictive linear model:
- $\mathrm{SSR}(\beta) = \sum_{q=1}^{M}(y_q - \boldsymbol{x_q}\boldsymbol{\beta})^2 = \sum_{q=1}^{M} {\varepsilon_q}^2$
- $\widehat{\boldsymbol{\beta}} = argmin\big(SSR(\boldsymbol{\beta})\big)$

To find $\widehat{\boldsymbol{\beta}}$, differentiate $SSR(\boldsymbol{\beta})$ with respect to $\boldsymbol{\beta}$ and set to zero.

$$0 = \frac{d}{d\boldsymbol{\beta}}\sum_{q=1}^{M}(y_q - \boldsymbol{x_q}\boldsymbol{\beta})^2 = \frac{d}{d\boldsymbol{\beta}}\sum_{q=1}^{M}{y_q}^2 - 2y_q\boldsymbol{x_q}\boldsymbol{\beta} + \boldsymbol{x_q}^{\mathbf{2}}\boldsymbol{\beta}^{\mathbf{2}}$$

$$= \sum_{q=1}^{M}\frac{d}{d\boldsymbol{\beta}}({y_q}^2 - 2y_q\boldsymbol{x_q}\boldsymbol{\beta} + \boldsymbol{x_q}^{\mathbf{2}}\boldsymbol{\beta}^{\mathbf{2}}) = 2\big(\boldsymbol{x_q}^{\mathbf{2}}\widehat{\boldsymbol{\beta}} - y_q\boldsymbol{x_q}\big)$$

$\boldsymbol{x}^{\mathbf{2}}\widehat{\boldsymbol{\beta}} = \boldsymbol{y}\boldsymbol{x}$ (note: x is MxN, y is 1xM)

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{x})^{-1}\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{y}$$

Multivariate Regression: Assumptions

- Predictand is a linear combination of the predictors (i.e. the predictand/predictor relationship is both linear and additive)
  - Not as significant of an assumption as one may initially think; variables may transformed to preserve a linear relationship.
    - Example: Temperature T is measured (or predicted in a model) at some location. That one 'observation' may serve as many predictors in the linear model:
      - $x_1 = T$
      - $x_2 = T^2$
      - $x_3 = T^4$
      - $x_4 = \ln T$
      - …

Multivariate Regression: Assumptions

- Exogeneity: $E[\varepsilon|\boldsymbol{X}] = 0$
  - Residuals are uncorrelated with the predictors
  - Mean (expected) error is zero
- No perfect multicollinearity
  - Recall in my derivation of the OLS estimator for $\widehat{\boldsymbol{\beta}}$, matrix inversion was required to recover the estimate for $\widehat{\boldsymbol{\beta}}$
  - Recall from linear algebra that the following statements are equivalent for an NxN matrix A:
    - Matrix A is invertible
    - Matrix A's determinant is non-zero
    - Matrix A is non-singular
    - Matrix A has rank N
    - The columns of A are linearly independent
  - Long story short: If two predictors are perfectly correlated, then two columns of the predictor matrix will be linearly dependent and the predictor matrix will be singular and thus non-invertible
- Homoscedasticity
  - Variance in the residuals (errors) is *independent* of:
    - Time
    - Predictand Value
    - Any Predictor Value
- Independence of Residuals
  - Predictand is uncorrelated with residuals
  - Weaker than 'pure' statistical independence
- Residuals are normally distributed

Early History of Statistical Hurricane Modeling: Pre-SHIPS

- 1979: Jarvinen and Neumann developed the Statistical Hurricane Intensity Forecast (SHIFOR) model
- Used only climatology and persistence variables with a linear regression technique to predict tropical cyclone intensity changes
- Still used as a baseline today

Statistical Hurricane Intensity Prediction Scheme (SHIPS)

- Arose from work by John Kaplan, Mark DeMaria, and others during the late 1980s
- First forecasts in 1990- out to 48h, predictions at 00Z and 12Z
- Applies multiple linear regression to predict hurricane intensity change

SHIPS: Early Work

- Initial SHIPS used only climatological, persistence, and synoptic variables as predictors (regressors)
- Storm Intensity: Maximum 1-minute sustained surface wind
  - Determined by NHC's "best track" data (uses aircraft reconnaissance where available, otherwise primarily satellite techniques)
- Training Data: All named TCs (incl. depression stage) from 1989-1992 over the Atlantic Basin that are entirely over water, plus sampling of TC cases from '82-'89.
  - 510 analyses, 49 TCs total
  - Reduced to 300 analyses by forecast hour 72 due to land interaction and/or dissipation
- Was pure statistical model until 1997

SHIPS

- Predictand: Intensity change from initialization
  - Different regression for reach time interval from 0 to 72 hour forecast
- Predictor (independent variable) selection
  - Developers deemed it optimal to subjectively select small subset of potential predictors that the developers considered to have a plausible physical relationship with the predictand (intensity change)
    - Concerned about overfitting
- Predictor considered significant if ≠ 0 at 95% confidence, determined via F-statistic
- Same predictor set at all intervals, even when set of significant predictors changes
- After subjectively selecting set of “reasonable candidate predictors” from set of “total candidate predictors”, backward iterative scheme applied to remove least-significant predictor from set of reasonable predictors until all remaining predictors statistically-significantly non-zero at at least one time interval

Greg's Aside (read: Rant)

- Developer's concern about overfitting with too many predictors is very real
- However, it is rarely optimal to subjectively select predictors in the manner conducted in the development of the SHIPS model; you may be throwing away valuable information contained within the discarded candidate predictors. There are typically better ways to avoid overfitting without discarding too many variables. Typically, the fact that adding more predictors yields strong overfitting indicates the algorithm is not *robust*.
- There are several alternative approaches which, in most contexts, are more reasonable:
  - Dimensionality Reduction (feature extraction): Reduce a high-dimensional dataset to a lower dimensionality dataset by generating a small subset of new features which explain the vast majority of the variance of the original dataset.
    - Many possible methods: Principal Component/Empirical Orthogonal Functions Analysis (PCA/EOF Analysis) is probably most common
    - Partial Least Squares Regression (PLSR) is often a good alternative when the predictand is known and the dimensionality of the data is larger the number of records
  - Penalize regression models based on their complexity- this also prevents overfitting
    - Examples:
    - Ridge Regression: Penalizes large weights
    - LASSO Regression: Penalizes non-zero weights
  - There are also Feature Selection algorithms, though the former two approaches tend to be favored
- It is *never* correct to draw conclusions about the skill or capability of a model by limiting analysis/validation to data on which the model was trained
- It is *essential* that, in the development of any statistical model, step one is the separation of *training data* and *validation data*
  - Model is trained exclusively on the training data; many different model configurations may be tested, and evaluated either through in-sample cross-validation, or on a separate group of *testing data* determined at step one
  - Once a model is deemed 'optimally tuned' on the training data, apply the model to the validation data. Once you do this *you cannot go back and revise the model*
- Maybe this wasn't well established circa 1990…

SHIPS (1993): Climatology & Persistence Predictors

TABLE 1. Climatological and persistence predictors.

1)	Absolute value of Julian date - 253.	(JDATE)
2)	Initial storm intensity.	(VMX)
3)	Intensity change during previous 12 h.	(DVMX)
4)	Initial storm latitude (°N).	(LAT)
5)	Initial storm longitude (°W).	(LONG)
6)	Eastward component of storm motion vector.	(USM)
7)	Northward component of storm motion vector.	(VSM)
8)	Magnitude of storm motion vector.	(CSM)

TABLE 2. Normalized regression coefficients for the climatological and persistence predictors. The coefficients that are significant at the 95% level are underlined; r^2 is the percent of the total variance explained by the regression.

Variable	Forecast interval 12	24	36	48	60	72
1) VMX	−0.24	−0.36	−0.44	−0.51	−0.57	−0.61
2) DVMX	+0.46	+0.34	+0.25	+0.22	+0.23	+0.20
3) LONG	+0.19	+0.27	+0.27	+0.26	+0.24	+0.24
4) JDATE	−0.12	−0.18	−0.20	−0.23	−0.24	−0.24
5) LAT	−0.08	−0.11	−0.14	−0.18	−0.16	−0.13
6) VSM	−0.04	−0.07	−0.11	−0.11	−0.14	−0.14
7) CSM	−0.03	+0.00	+0.05	+0.06	+0.12	+0.18
8) USM	−0.09	−0.11	−0.09	−0.05	−0.04	−0.05
r^2 (%)	29.8	28.6	30.2	33.6	38.4	41.3

SHIPS (1993): Climatology & Persistence Predictors

- VMX: More intense storm implies already peaked (intensity favored to decrease)
- DVMX: Persistence (increase in past 12 hours favors continued intensity increase)
- LONG: Positive coeff <-> SST increase to the west
- JDATE: Stronger storms at peak of season
- VSM/USM: North/East storm motion tends not to favor intensification

TABLE 3. Synoptic predictors.

1) Maximum possible intensity–initial intensity.	(POT)
2) Magnitude of 850–200-mb vertical shear.	(SHR)
3) Time tendency of vertical shear magnitude.	(DSHR)
4) The 200-mb relative eddy angular momentum flux convergence.	(REFC)
5) The 200-mb planetary eddy angular momentum flux convergence.	(PEFC)
6) The 850-mb relative angular momentum.	(SIZE)
7) Distance to nearest major landmass.	(DTL)

$$POT = \left(\frac{12}{FHOUR+12} \sum_{f=FHOUR} 66.5 + 108.5 * e^{0.1813(SST_f - 30.0)} \right) - INT_{FHOUR}$$

$$SHR = \begin{cases} \frac{12}{FHOUR+12} \sum_{f=FHOUR} \overline{\boldsymbol{V}_{200_{t=f}}} - \overline{\boldsymbol{V}_{850_{t=f}}}, FHOUR < 36 \\ \frac{1}{4} \sum_{f=0}^{36} \overline{\boldsymbol{V}_{200_{t=f}}} - \overline{\boldsymbol{V}_{850_{t=f}}}, FHOUR \geq 36 \end{cases}$$, averaged over 600km radius relative to storm center at t=FHOUR

$$DSHR = \frac{SHR(t=24) - SHR(t=0)}{24}$$

DTL uses storm center *at forecast time* in calculation

The integrated relative angular momentum (SIZE) is included as a measure of the extent of the outer circulation of the tropical cyclone. This variable is defined by

$$\mathrm{SIZE} = \int_{r_1}^{r_2} (rV) r dr, \qquad (4)$$

where V is the 850-mb tangential wind, $r_1 = 400$ km, and $r_2 = 800$ km. Although there is usually not enough

REFC and PEFC are evaluated at 200 mb using

$$\mathrm{REFC} = -r^{-2} \partial / \partial r (r^2 \overline{U'_L V'_L}) \qquad (2)$$

$$\mathrm{PEFC} = -\overline{f'U'}, \qquad (3)$$

where r is the radius from the storm center, U is the radial wind, V is the tangential wind, f is the Coriolis parameter, the overbar denotes an azimuthal average, the prime represents a deviation from the azimuthal average, and the subscript L indicates that these quantities (U or V) are evaluated in a coordinate system moving with the storm. Qualitatively, REFC tends to

REFC, PEFC, SIZE all use analysis t=0 values

TABLE 4. Normalized regression coefficients for the combined climatological, persistence, and synoptic predictors. The coefficients that are significant at the 95% level are underlined; r^2 is the percent of the total variance explained by the regression.

Variable	Forecast interval 12	24	36	48	60	72
1) POT	<u>+0.32</u>	<u>+0.46</u>	<u>+0.56</u>	<u>+0.63</u>	<u>+0.68</u>	<u>+0.70</u>
2) SHR	<u>−0.20</u>	<u>−0.26</u>	<u>−0.31</u>	<u>−0.36</u>	<u>−0.31</u>	<u>−0.25</u>
3) DVMX	<u>+0.40</u>	<u>+0.28</u>	<u>+0.18</u>	<u>+0.16</u>	<u>+0.18</u>	<u>+0.16</u>
4) REFC	+0.03	<u>+0.08</u>	<u>+0.16</u>	<u>+0.22</u>	<u>+0.19</u>	<u>+0.17</u>
5) PEFC	+0.08	<u>+0.12</u>	<u>+0.12</u>	<u>+0.10</u>	<u>+0.10</u>	+0.10
6) JDATE	−0.04	−0.06	−0.07	−0.09	<u>−0.12</u>	<u>−0.13</u>
7) LONG	<u>+0.14</u>	<u>+0.14</u>	+0.08	+0.03	−0.02	−0.09
8) DTL	<u>+0.12</u>	<u>+0.12</u>	+0.09	+0.05	+0.00	−0.09
9) SIZE	<u>+0.11</u>	<u>+0.11</u>	<u>+0.09</u>	+0.05	+0.05	+0.03
10) DSHR	−0.01	+0.06	<u>+0.13</u>	+0.07	−0.03	<u>−0.11</u>
r^2 (%)	35.7	39.4	44.4	50.4	52.0	53.6

- Larger departure from MPI indicates more potential for hurricane to intensify and correlates with future hurricane intensification
- Large vertical shear is inhibitive of hurricane intensification
- REFC/PEFC: Eddy momentum flux convergence corresponds to interactions of storm with large-scale, synoptic flow to make upper-level hurricane circulation more cyclonic; this can work to intensify storm
  - I don’t really understand this…see Holland and Merrill (1984) or DeMaria, Baik, and Kaplan (1993) for theoretical argument for this

SHIPS (1993-1996)

TABLE 1. Predictors used in the DK94 (first 11) and later versions of SHIPS.

1) POT	Maximum possible intensity-initial intensity
2) SHR	Magnitude of 850–200-mb vertical shear
3) DVMX	Intensity change during previous 12 h
4) REFC	200-mb relative eddy angular momentum flux convergence
5) PEFC (removed 1995)	200-mb planetary eddy angular momentum flux convergence
6) JDAY	Absolute value of Julian day—253
7) LONG (removed 1994)	Initial storm longitude
8) DTL (removed 1994)	Distance to nearest major landmass
9) SIZE (removed 1997)	850-mb relative angular momentum
10) DSHR (removed 1996)	Time tendency of vertical shear magnitude
11) POT2	POT^2
12) T200 (added 1995)	Average 200-mb temperature within 1000 km of storm center
13) U200 (added 1995)	Average 200-mb zonal wind within 1000 km of storm center
14) Z850 (added 1997)	Average 850-mb vorticity within 1000 km of storm center
15) LSHR (added 1997)	SHR times the sine of the initial storm latitude
16) D200 (added 1998)	Average 200-mb divergence within 1000 km of storm center
17) SPDX (added 1998)	Zonal component of initial storm motion vector
18) VMX (added 1998)	Initial storm maximum wind

TABLE 2. Normalized SHIPS errors (%) for Atlantic intensity forecasts 1993–97. A negative number indicates skill relative to SHIFOR (improvement over climatology and persistence). The sample size at each forecast interval is indicated in parentheses.

	Forecast interval (h)				
Year	12	24	36	48	72
1993	0 (133)	1 (112)	7 (94)	2 (80)	−23 (56)
1994	−3 (110)	−8 (91)	−5 (75)	−5 (59)	43 (32)
1995	1 (468)	1 (428)	1 (389)	5 (347)	14 (285)
1996	0 (341)	−1 (309)	−5 (279)	−5 (251)	−7 (210)
1997	−3 (104)	−7 (86)	−18 (70)	−28 (59)	−31 (44)

SHIPS (1997)

- Use of only synoptic predictors at analysis time thought to be major limitation to SHIPS
- Dynamical models at the time very poor; tropical cyclones often initialized hundreds of km from storm center
- Forecast variables thus became wildly inaccurate; using model variables for shear produced degraded results relative to using just values at time of analysis
- Used AVN model (NCEP's global model- predecessor to GFS)
- Needed to remove TC from model fields before use

SHIPS (1997)

- Complex procedure to eliminate TC from model

$$F(x, y) = \begin{cases} f(x, y) & \text{for } r \geq R \\ g(x, y) & \text{for } r < R, \end{cases} \qquad (3.8)$$

where

$$\nabla^2 g = 0 \qquad (3.9)$$

$$g(x, y) = f(x, y) \quad \text{at } r = R \qquad (3.10)$$

$$r = [(x - x_0)^2 + (y - y_0)^2]^{1/2} \qquad (3.11)$$

R = 500km (100mb)-1000km (1000mb)

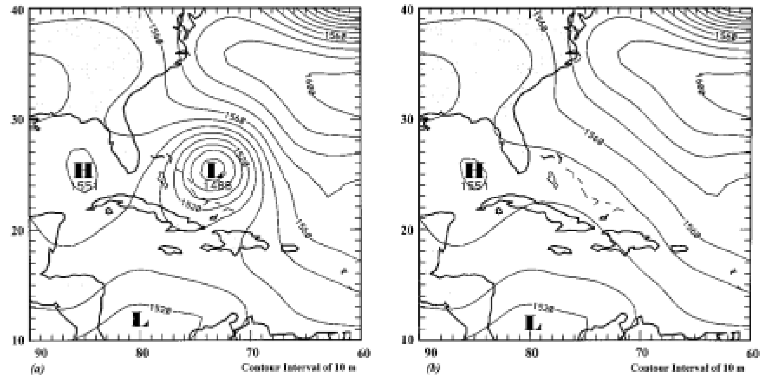


FIG. 4. The 850-mb geopotential height from the aviation model at 0000 UTC 4 Sep 1996 (a) before and (b) after the application of the Laplacian filter. Hurricane Fran was located very close to the height minimum near the center of (a).

- Laplacian filtered AVN analysis used to initialize very simple limited area model (generalized version of LBAR model from Horsfall et al. (1997)) . For details on this step, see DeMaria and Kaplan (1999)
  - 10 vertical levels, ~150km horizontal resolution
  - ‘Dry’ primitive equations only, no parameterizations
  - ‘Smoothed’ transition between LBAR, AVN fields
  - AVN BCs
  - Run to 72h for all cases from 1989-1996 (85 total storms, 1025 cases)

SHIPS (1997)

- T200: Cold upper levels more favored for intensification
- U200: More easterly upper-level favored for intensification (likely related to low shear)
- Z850: More cyclonic lower-levels favors intensifications
- POT2: Correction for +POT coefficients; attributed to weak storms lacking organization to rapidly intensify

TABLE 3. Normalized regression coefficients for the 1997 version of SHIPS for the Atlantic basin. Coefficients that are significant at the 95% level are underlined, r^2 is the percent of the total variance explained by the regression, and N is the developmental sample size.

	Forecast interval (h)					
Variable	12	24	36	48	60	72
POT	+0.62	+0.69	+0.73	+0.79	+0.84	+0.96
SHR	−0.35	−0.43	−0.43	−0.43	−0.44	−0.42
DVMX	+0.40	+0.30	+0.23	+0.18	+0.13	+0.08
T200	−0.08	−0.13	−0.15	−0.18	−0.20	−0.22
U200	−0.08	−0.11	−0.15	−0.19	−0.20	−0.21
Z850	+0.09	+0.12	+0.13	+0.13	+0.14	+0.14
REFC	+0.07	+0.07	+0.07	+0.09	+0.12	+0.14
JDAY	−0.03	−0.03	−0.04	−0.05	−0.08	−0.10
POT2	−0.30	−0.24	−0.21	−0.22	−0.24	−0.34
LSHR	+0.23	+0.27	+0.26	+0.25	+0.24	+0.24
r^2	36	40	45	50	53	54
N	1025	929	836	752	676	605

SHIPS (1998-2003): Quick changes

- Forecasts extended to 5 days (120h)
- Updates to training data
  - Coefficients re-trained including previous season’s data
  - Tropical depressions added in 2001
  - Subtropical storms added in 2003
- Use NHC track forecasts instead of LBAR
- Switched to NCAR re-analysis in place of operational
- Explored use of more data sources, including satellite obs and aircraft
- Scrapped LBAR in 2001 in favor of GFS fields
  - These now include representation of TCs

SHIPS (1998-2003): Predictor Updates

TABLE 1. Predictors used in SHIPS 1997–2003. The predictors that are evaluated at the beginning of the forecast period are static (S), and predictors that are evaluated along the forecasted track of the storm are time dependent (T). An X indicates that the predictor was used in that year and a dash (—) indicates it was not used that year.

Predictor	Static (S) or time dependent (T)	Year						
		1997	1998	1999	2000	2001	2002	2003
1) Absolute value of (Julian day − peak season value)	S	X	X	X	X	X	X	—
2) Gaussian function of (Julian day − peak value)	S	—	—	—	—	—	—	X
3) Initial maximum winds	S	—	X	X	X	X	X	X
4) Max wind change during the past 12 h	S	X	X	X	X	X	X	X
5) Initial max winds times previous 12-h change	S	—	—	—	—	—	—	X
6) Zonal component of storm motion	S	—	X	X	X	X	X	X
7) Pressure level of storm steering	S	—	—	—	—	X	X	X
8) 200-hPa divergence	S	—	X	X	X	—	X	X
9) 200-hPa eddy momentum flux convergence	S	—	—	—	—	—	X	—
10) 200-hPa eddy momentum flux convergence	T	X	X	X	X	X	—	—
11) Max potential intensity − current intensity	T	X	X	X	X	X	X	X
12) 850–200-hPa vertical shear	T	X	X	X	X	X	X	X
13) 200-hPa zonal wind	T	X	X	X	X	X	X	—
14) 200-hPa temperature	T	X	X	X	X	X	X	X
15) 850–700-hPa relative humidity	T	—	—	—	—	X	X	—
16) 500–300-hPa relative humidity	T	—	—	—	—	—	—	X
17) 850-hPa relative vorticity	T	X	X	X	X	X	X	X
18) Surface − 200-hPa θ_e deviation of lifted parcel	T	—	—	—	—	—	—	X
19) Vertical shear times sine of storm latitude	T	X	X	X	X	X	X	X
20) Square of potential − current intensity	T	X	X	X	X	X	X	X
21) Initial intensity time shear	T	—	—	—	—	—	X	X

- Most significant predictors:
  - Difference between current intensity and estimated Maximum Potential Intensity (MPI)
  - Vertical Wind Shear
  - Persistence
  - Upper-Tropospheric Temperature

SHIPS (1998-2003): Land Decay (DSHIPS)

- 12-hour forecast track/intensity forecasts interpolated to 1-hour forecast values
- Portion of track over land identified
- Empirical decay model applied for land segments
- After re-entering water, forecast re-adjusted so that change consistent with original forecast

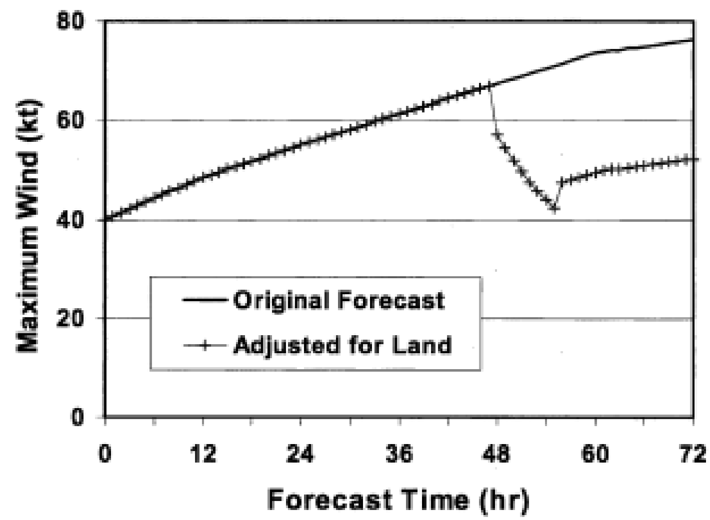
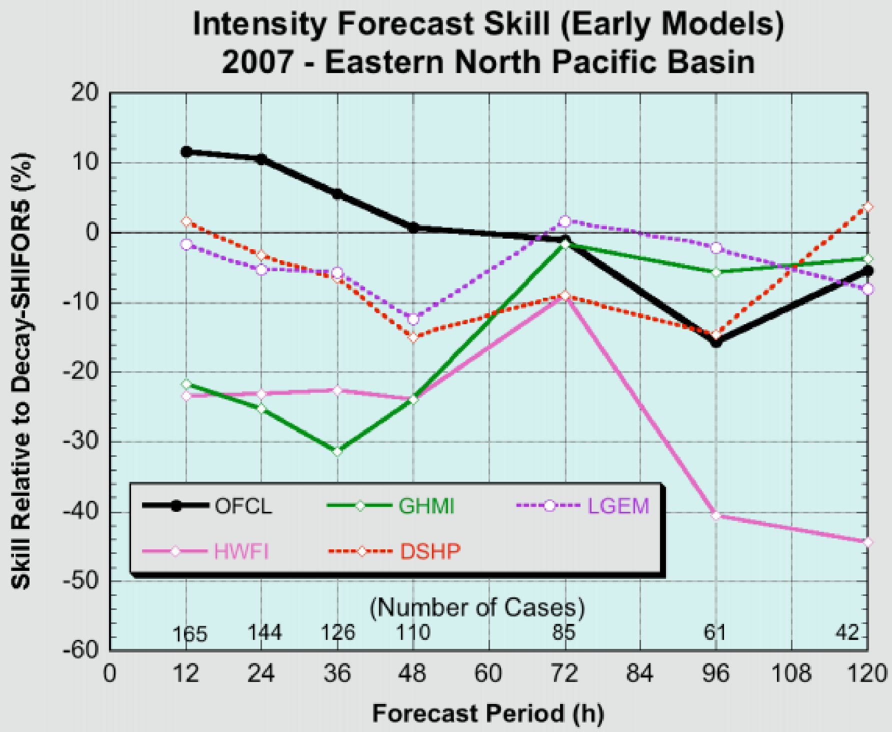
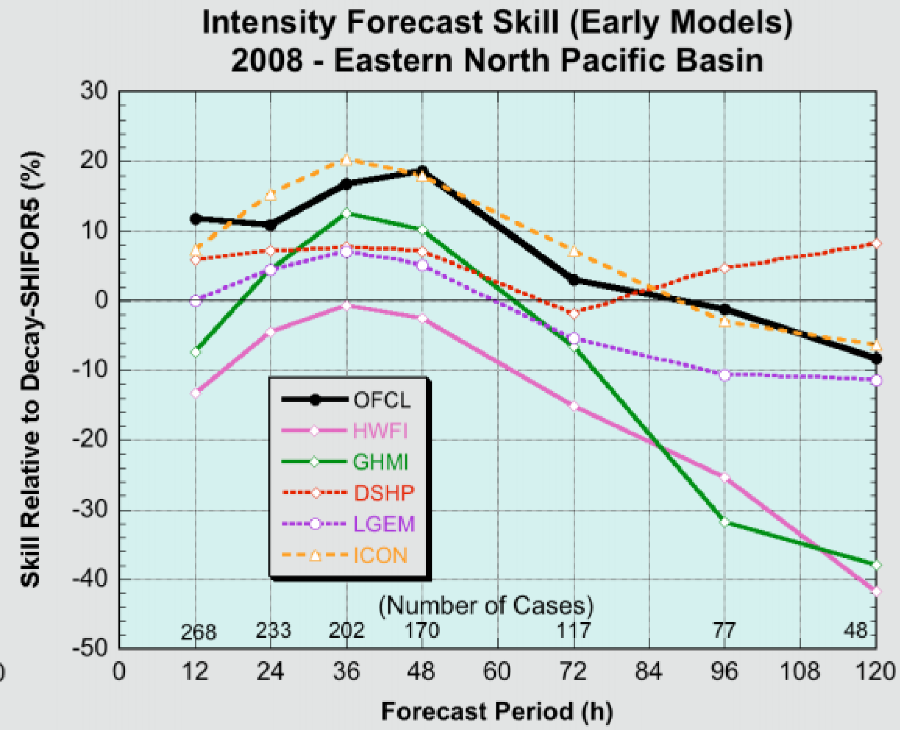


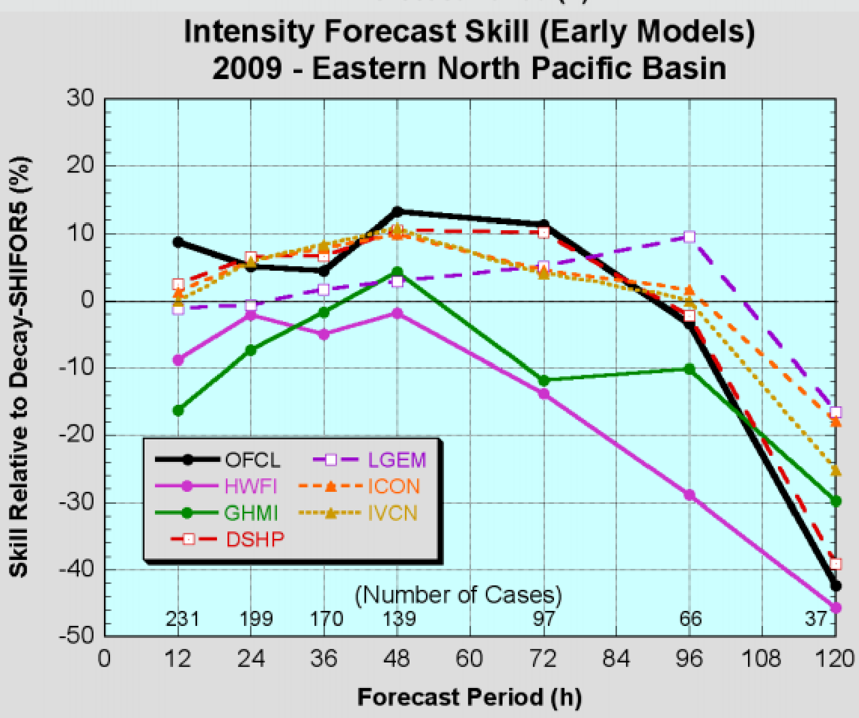
FIG. 1. The original SHIPS forecast and the forecast adjusted for movement over land for Hurricane Isidore beginning at 1200 UTC on 18 Sep 2002. Isidore crossed western Cuba between 48 and 55 h.

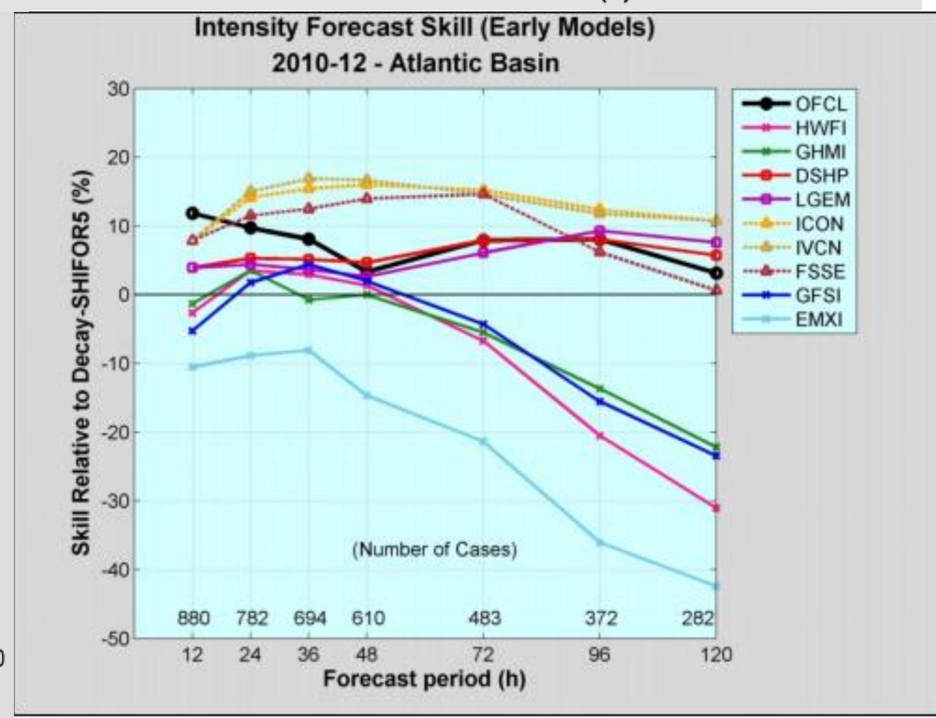
SHIPS (2003-Present)

- Applied SHIPS framework to many other basins
- Development of Rapid Intensification Index (RII)
- Logistic Growth Equation Model (LGEM) formulation added


Intensity Forecast Skill (Early Models)
2007 - Eastern North Pacific Basin
Skill Relative to Decay-SHIFOR5 (%)
OFCL
GHMI
LGEM
HWFI
DSHP
(Number of Cases)
165 144 126 110 85 61 42
Forecast Period (h)


Intensity Forecast Skill (Early Models)
2008 - Eastern North Pacific Basin
Skill Relative to Decay-SHIFOR5 (%)
OFCL
HWFI
GHMI
DSHP
LGEM
ICON
(Number of Cases)
268 233 202 170 117 77 48
Forecast Period (h)


Intensity Forecast Skill (Early Models)
2009 - Eastern North Pacific Basin
Skill Relative to Decay-SHIFOR5 (%)
OFCL
LGEM
HWFI
ICON
GHMI
IVCN
DSHP
(Number of Cases)
231 199 170 139 97 66 37
Forecast Period (h)


Intensity Forecast Skill (Early Models)
2010-12 - Atlantic Basin
Skill Relative to Decay-SHIFOR5 (%)
OFCL
HWFI
GHMI
DSHP
LGEM
ICON
IVCN
FSSE
GFSI
EMXI
(Number of Cases)
880 782 694 610 483 372 282
Forecast period (h)

Fortune Cookie Wisdom

- Simple techniques can be extremely powerful
- Post-processing can add great value to raw dynamical guidance
- Models can always be improved
- Long way to go to have satisfactory skill in forecasting of hurricane intensity
- *Always* separate training and validation data

References

DeMaria, M., and J. Kaplan, 1994: A statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic basin. Wea. Forecasting, **9**, 209 220

DeMaria, M., and J. Kaplan, 1999: An updated statistical hurricane intensity prediction scheme (SHIPS) for the Atlantic and eastern north Pacific basins. Wea. Forecasting, **14**, 326-337

DeMaria, M., M. Mainelli, L.K. Shay, J.A. Knaff and J. Kaplan, 2005: Further Improvements in the Statistical Hurricane Intensity Prediction Scheme (SHIPS). Wea. Forecasting, **20**, 531-543

Eckel, Tony. "Calibration." University of Washington. Seattle, WA. Spring 2011. ATMS 444/544 Lecture

Guesterin, Carlos.

"Linear Regression." University of Washington. Seattle, WA. 5 April, 2013. CSE 446 Lecture

"Overfitting." University of Washington. Seattle, WA. 8 April, 2013. CSE 446 Lecture

"Regularization." University of Washington. Seattle, WA. 10 April, 2013. CSE 446 Lecture

"Variable Selection LASSO: Sparse Regression." University of Washington. Seattle, WA. 12 April, 2013. CSE 446 Lecture

Jarvinen, Brian, and Neumann, Charles. "Statistical Forecasts of Tropical Cyclone Intensity for the North Atlantic Basin" . April 1979. http://www.nhc.noaa.gov/pdf/NWS-NHC-1979-10.pdf

Johnson, Riess, & Arnold. "Introduction to Linear Algebra." Fifth Edition. 2002

http://people.duke.edu/~rnau/testing.htm

http://www.nhc.noaa.gov/verification/verify3.shtml

http://rammb.cira.colostate.edu/research/tropical_cyclones/ships/index.asp

Bias/Variance Tradeoff