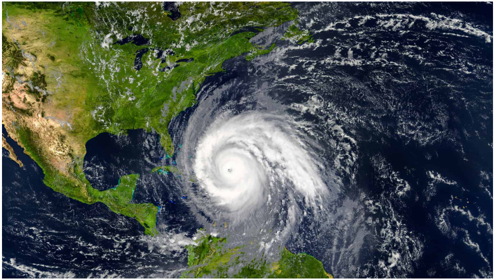
BDA Project: Hurricane forecasting in Stan

José Miguel Ramírez & Jonas Lindblad
Aalto University

Hurricane introduction

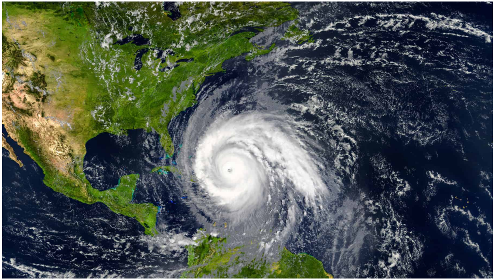
Hurricane introduction

- Destructive storms occuring in the late summer and fall in the northern hemisphere's tropical region.



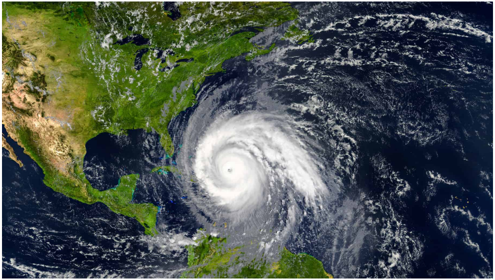
Hurricane introduction

- Destructive storms occuring in the late summer and fall in the northern hemisphere's tropical region.
- Classified by their wind intensity at the eye wall.



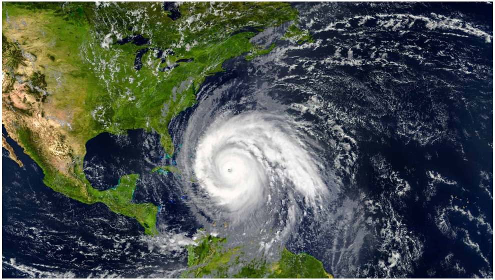
Hurricane introduction

- Destructive storms occuring in the late summer and fall in the northern hemisphere's tropical region.
- Classified by their wind intensity at the eye wall.
- They can cause extreme levels of flooding and destroy many buildings.



Hurricane introduction

- Destructive storms occuring in the late summer and fall in the northern hemisphere's tropical region.
- Classified by their wind intensity at the eye wall.
- They can cause extreme levels of flooding and destroy many buildings.
- Monetary damages and loss of lives increase with an almost exponential character as a function of storm intensity.



Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
    - Dynamical: simulate the laws of physics

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
    - Dynamical: simulate the laws of physics
    - Statistical: estimate based on historical data

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
  - Dynamical: simulate the laws of physics
  - Statistical: estimate based on historical data
- Dynamical vs. statistical: good at long- and short-range forecasts respectively

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
  - Dynamical: simulate the laws of physics
  - Statistical: estimate based on historical data
- Dynamical vs. statistical: good at long- and short-range forecasts respectively
- . . . but for hurricane forecasts short-range is usually more interesting

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
  - Dynamical: simulate the laws of physics
  - Statistical: estimate based on historical data
- Dynamical vs. statistical: good at long- and short-range forecasts respectively
- ...but for hurricane forecasts short-range is usually more interesting
- *Rapid intensification*: forecasted better by dynamical models

Hurricane forecasting basics

- Forecasters predict two quantities: track and intensity
- Two kinds of models: *dynamical* and *statistical*
  - Dynamical: simulate the laws of physics
  - Statistical: estimate based on historical data
- Dynamical vs. statistical: good at long- and short-range forecasts respectively
- . . . but for hurricane forecasts short-range is usually more interesting
- *Rapid intensification*: forecasted better by dynamical models
- This project: a *statistical* model for *intensity*

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!
- The NHC regression model: Statistical Hurricane Intensity Prediction Scheme (SHIPS)

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!
- The NHC regression model: Statistical Hurricane Intensity Prediction Scheme (SHIPS)
- SHIPS: $\sim$140 covariates, many calculated from data sources more easily available to the NHC

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!
- The NHC regression model: Statistical Hurricane Intensity Prediction Scheme (SHIPS)
- SHIPS: ~140 covariates, many calculated from data sources more easily available to the NHC
- SHIPS dataset: publically available with no restrictions (link: SHIPS Development)

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!
- The NHC regression model: Statistical Hurricane Intensity Prediction Scheme (SHIPS)
- SHIPS: ∼140 covariates, many calculated from data sources more easily available to the NHC
- SHIPS dataset: publically available with no restrictions (link: SHIPS Development)
- . . . but the documentation is terrible

Hurricane forecasting basics: the SHIPS data

The US government forecasting agency, the National Hurricane Center (NHC), uses a large number of models operationally. The models (together: the *model ensemble*) are used together with experienced meteorologists' judgment to provide the official forecast.

- Surprisingly, the best *single* (short-range) model is a multiple linear regression!
- The NHC regression model: Statistical Hurricane Intensity Prediction Scheme (SHIPS)
- SHIPS: ∼140 covariates, many calculated from data sources more easily available to the NHC
- SHIPS dataset: publically available with no restrictions (link: SHIPS Development)
- ... but the documentation is terrible
- SHIPS: only a point estimate; our project: a predictive distribution

Hurricane forecasting basics: the SHIPS data

The SHIPS developmental data is confusing!

```
ALBE 820602 12   20   21.7   87.1 1005 AL011982                                                          HEAD
 -12   -6    0    6   12   18   24   30   36   42   48   54   60   66   72   78   84   90   96  102  108  114  120 TIME
9999 9999   20   25   30   40   50   75   65   55   45   40   30   25   25   25   25   25   20 9999 9999 9999 9999 VMAX
9999 9999 1005 1004 1003 1001  995  985  992  998 1002 1005 1007 1008 1009 1010 1010 1010 1010 9999 9999 9999 9999 MSLP
9999 9999    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1 9999 9999 9999 9999 TYPE
             1    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0 HIST
9999 9999    0    5   10   20   30   55   45   35   25   20   10    5    5    5    5    5    0 9999 9999 9999 9999 DELV
9999 9999 9999    5    5   10   10   25  -10  -10  -10   -5  -10   -5    0    0    0    0   -5 9999 9999 9999 9999 INCV
9999 9999  217  222  226  228  232  240  248  249  249  250  251  252  253  254  255  255  255 9999 9999 9999 9999 LAT
9999 9999  871  865  858  850  842  836  834  841  848  842  841  840  839  836  833  830  826 9999 9999 9999 9999 LON
9999 9999  274  278  281  283  281  278  278  278  278  275  275  274  274  274  274  274  274 9999 9999 9999 9999 CSST
9999 9999  156  161  177  228  237  205  122  132  173  132  132  132  103   95   95   92   94 9999 9999 9999 9999 CD20
9999 9999   56   56   66  102  110   88   40   45   64   45   45   45   31   28   28   28   29 9999 9999 9999 9999 CD26
9999 9999   31   33   43   72   81   65   25   28   39   28   28   28   17   16   16   16   18 9999 9999 9999 9999 COHC
9999 9999   21   93  136   84   69  132  215  241  267  255  258  244  230  198  166  139  105 9999 9999 9999 9999 DTL
9999 9999    0   46   65   67   63   55   60   94   95  181  227  266  299  297  279  267  220 9999 9999 9999 9999 OAGE
9999 9999    0   10   17   22   27   33   41   60   54   98  113  120  123  115  101   91   68 9999 9999 9999 9999 NAGE
9999 9999  280  280  278  276  274  270  269  272  274  272  271  271  270  269  268  267  266 9999 9999 9999 9999 RSST    0
9999 9999  267  279  280  272  267  267  270  273  270  273  273  272  272  272  272  272  274 9999 9999 9999 9999 DSST    1
9999 9999  271  278  279  272  267  268  270  272  270  272  272  272  271  271  272  272  274 9999 9999 9999 9999 DSTA    1
           224  187  169  179  209  221  234  236  221  202  229  204  178  182  195  165  172 9999 9999 9999 9999 U200
           238  162  153  143  139  170  206  158  108   94  125  102  132  149  159  120  212 9999 9999 9999 9999 U20C
            99   96  163  101  134  164  126  -29  -39    3  -19  -68  -88  -54  -77 -149 -189 9999 9999 9999 9999 V20C
          3528 3528 3535 3523 3488 3506 3513 3509 3512 3525 3524 3515 3505 3514 3523 3483 3477 9999 9999 9999 9999 E000
           113  114  114  107   86  104  109  109  118  125  118  116  120  130  141  111  115 9999 9999 9999 9999 EPOS
             8    9    8    9   11   10   10   10   10   10    9   10   11   11   10   14   14 9999 9999 9999 9999 ENEG
            51   48   52   48   22   42   42   42   50   55   46   46   48   47   53   22   26 9999 9999 9999 9999 EPSS
            16   20   18   18   21   23   21   19   18   22   20   19   19   25   23   28   26 9999 9999 9999 9999 ENSS
            70   68   70   70   66   69   67   67   68   68   69   70   68   62   59   56   55 9999 9999 9999 9999 RHLO
            57   53   56   59   53   55   51   51   48   48   46   46   43   37   34   33   32 9999 9999 9999 9999 RHMD
            50   51   54   59   54   49   48   43   35   35   34   31   26   23   24   25   24 9999 9999 9999 9999 RHHI
           548   31   23   24   13   28   16   40   23   64   82   77   89   92   85  208  200   79 9999 9999 9999 PSLV
             7   17    0  -14   -8   14   -9  -37  -60  -47  -32  -52  -29  -30    0   -4  -20 9999 9999 9999 9999 Z850
            64   65   84   70   51  114   42   27   -7    8  -17  -37  -48  -52  -25  -32  -41 9999 9999 9999 9999 D200
             4    6    3    6   12    4   10    0  -16   -9   -3   -3   -2    2    7    0  -17 9999 9999 9999 9999 REFC
            -1    0   -1   -1   -2   -1    0    1    1    1    1    2    3    2    3    4    6 9999 9999 9999 9999 PEFC
           259  267  266  261  253  265  264  258  255  268  265  258  254  269  269  257  252 9999 9999 9999 9999 T000
            84   79   81   83   83   77   79   82   85   78   80   83   85   76   77   79   81 9999 9999 9999 9999 R000
           -27  -24  -31  -20  -20   -8  -16  -10   -8   -2  -18   -8   -8   -3   -7    4    8 9999 9999 9999 9999 Z000
           207  218  217  227  232  239  242  248  249  250  249  251  248  248  251  251  263 9999 9999 9999 9999 TLAT
           873  867  862  849  840  844  839  839  843  840  839  839  841  840  837  835  839 9999 9999 9999 9999 TLON
            67   76   69   63   70   83   67   60   65   62   60   51   41   32   22   18   16 9999 9999 9999 9999 TWAC
           100  104  106  110  110  138  114  102  109  101   98   84   73   62   42   32   34 9999 9999 9999 9999 TWXC
```

Hurricane forecasting basics: the SHIPS data

We are making *synoptic* models.

T = 0

```
 ALBE 820602 12   20   21.7   87.1 1005 AL011982                                                        HEAD
  -12   -6    0    6   12   18   24   30   36   42   48   54   60   66   72   78   84   90   96  102  108  114  120 TIME
 9999 9999   20   25   30   40   50   75   65   55   45   40   30   25   25   25   25   25   20 9999 9999 9999 9999 VMAX
 9999 9999 1005 1004 1003 1001  995  985  992  998 1002 1005 1007 1008 1009 1010 1010 1010 1010 9999 9999 9999 9999 MSLP
 9999 9999    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1    1 9999 9999 9999 9999 TYPE
              1    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0    0 HIST
 9999 9999    0    5   10   20   30   55   45   35   25   20   10    5    5    5    5    5    0 9999 9999 9999 9999 DELV
 9999 9999 9999    5    5   10   10   25  -10  -10  -10   -5  -10   -5    0    0    0    0   -5 9999 9999 9999 9999 INCV
 9999 9999  217  222  226  228  232  240  248  249  249  250  251  252  253  254  255  255  255 9999 9999 9999 9999 LAT
 9999 9999  871  865  858  850  842  836  834  841  848  842  841  840  839  836  833  830  826 9999 9999 9999 9999 LON
 9999 9999  274  278  281  283  281  278  278  278  278  275  275  274  274  274  274  274  274 9999 9999 9999 9999 CSST
 9999 9999  156  161  177  228  237  205  122  132  173  132  132  132  103   95   95   92   94 9999 9999 9999 9999 CD20
 9999 9999   56   56   66  102  110   88   40   45   64   45   45   45   31   28   28   28   29 9999 9999 9999 9999 CD26
 9999 9999   31   33   43   72   81   65   25   28   39   28   28   28   17   16   16   16   18 9999 9999 9999 9999 COHC
 9999 9999   21   93  136   84   69  132  215  241  267  255  258  244  230  198  166  139  105 9999 9999 9999 9999 DTL
 9999 9999    0   46   65   67   63   55   60   94   95  181  227  266  299  297  279  267  220 9999 9999 9999 9999 OAGE
 9999 9999    0   10   17   22   27   33   41   60   54   98  113  120  123  115  101   91   68 9999 9999 9999 9999 NAGE
 9999 9999  280  280  278  276  274  270  269  272  274  272  271  271  270  269  268  267  266 9999 9999 9999 9999 RSST     0
 9999 9999  267  279  280  272  267  267  270  273  270  273  273  272  272  272  272  272  274 9999 9999 9999 9999 DSST     1
 9999 9999  271  278  279  272  267  268  270  272  270  272  272  272  271  271  272  272  274 9999 9999 9999 9999 DSTA     1
            224  187  169  179  209  221  234  236  221  202  229  204  178  182  195  165  172 9999 9999 9999 9999 U200
            238  162  153  143  139  170  206  158  108   94  125  102  132  149  159  120  212 9999 9999 9999 9999 U20C
             99   96  163  101  134  164  126  -29  -39    3  -19  -68  -88  -54  -77 -149 -189 9999 9999 9999 9999 V20C
           3528 3528 3535 3523 3488 3506 3513 3509 3512 3525 3524 3515 3505 3514 3523 3483 3477 9999 9999 9999 9999 E000
            113  114  114  107   86  104  109  109  118  125  118  116  120  130  141  111  115 9999 9999 9999 9999 EPOS
              8    9    8    9   11   10   10   10   10   10    9   10   11   11   10   14   14 9999 9999 9999 9999 ENEG
             51   48   52   48   22   42   42   42   50   55   46   46   48   47   53   22   26 9999 9999 9999 9999 EPSS
             16   20   18   18   21   23   21   19   18   22   20   19   19   25   23   28   26 9999 9999 9999 9999 ENSS
             70   68   70   70   66   69   67   67   68   68   69   70   68   62   59   56   55 9999 9999 9999 9999 RHLO
             57   53   56   59   53   55   51   51   48   48   46   46   43   37   34   33   32 9999 9999 9999 9999 RHMD
             50   51   54   59   54   49   48   43   35   35   34   31   26   23   24   25   24 9999 9999 9999 9999 RHHI
            548   31   23   24   13   28   16   40   23   64   82   77   89   92   85  208  200   79 9999 9999 9999 PSLV
              7   17    0  -14   -8   14   -9  -37  -60  -47  -32  -52  -29  -30    0   -4  -20 9999 9999 9999 9999 Z850
             64   65   84   70   51  114   42   27   -7    8  -17  -37  -48  -52  -25  -32  -41 9999 9999 9999 9999 D200
              4    6    3    6   12    4   10    0  -16   -9   -3   -3   -2    2    7    0  -17 9999 9999 9999 9999 REFC
             -1    0   -1   -1   -2   -1    0    1    1    1    1    2    3    2    3    4    6 9999 9999 9999 9999 PEFC
            259  267  266  261  253  265  264  258  255  268  265  258  254  269  269  257  252 9999 9999 9999 9999 T000
             84   79   81   83   83   77   79   82   85   78   80   83   85   76   77   79   81 9999 9999 9999 9999 R000
            -27  -24  -31  -20  -20   -8  -16  -10   -8   -2  -18   -8   -8   -3   -7    4    8 9999 9999 9999 9999 Z000
            207  218  217  227  232  239  242  248  249  250  249  251  248  248  251  251  263 9999 9999 9999 9999 TLAT
            873  867  862  849  840  844  839  839  843  840  839  839  841  840  837  835  839 9999 9999 9999 9999 TLON
             67   76   69   63   70   83   67   60   65   62   60   51   41   32   22   18   16 9999 9999 9999 9999 TWAC
            100  104  106  110  110  138  114  102  109  101   98   84   73   62   42   32   34 9999 9999 9999 9999 TWXC
```

Hurricane forecasting basics: the SHIPS data

We are making *synoptic* models and choosing variables.

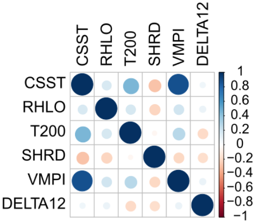
T = 0

ALBE	820602	12	20	21.7	87.1	1005	AL011982																HEAD	
-12	-6	0	6	12	18	24	30	36	42	48	54	60	66	72	78	84	90	96	102	108	114	120	TIME	
9999	9999	20	25	30	40	50	75	65	55	45	40	30	25	25	25	25	25	20	9999	9999	9999	9999	VMAX	
9999	9999	1005	1004	1003	1001	995	985	992	998	1002	1005	1007	1008	1009	1010	1010	1010	1010	9999	9999	9999	9999	MSLP	
9999	9999	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	9999	9999	9999	9999	TYPE	
		1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	HIST	
9999	9999	0	5	10	20	30	55	45	35	25	20	10	5	5	5	5	5	0	9999	9999	9999	9999	DELV	
9999	9999	9999	5	5	10	10	25	-10	-10	-10	-5	-10	-5	0	0	0	0	-5	9999	9999	9999	9999	INCV	
9999	9999	217	222	226	228	232	240	248	249	249	250	251	252	253	254	255	255	255	9999	9999	9999	9999	LAT	
9999	9999	871	865	858	850	842	836	834	841	848	842	841	840	839	836	833	830	826	9999	9999	9999	9999	LON	
9999	9999	274	278	281	283	281	278	278	278	278	275	275	274	274	274	274	274	274	9999	9999	9999	9999	CSST	
9999	9999	156	161	177	228	237	205	122	132	173	132	132	132	103	95	95	92	94	9999	9999	9999	9999	CD20	
9999	9999	56	56	66	102	110	88	40	45	64	45	45	45	31	28	28	28	29	9999	9999	9999	9999	CD26	
9999	9999	31	33	43	72	81	65	25	28	39	28	28	28	17	16	16	16	18	9999	9999	9999	9999	COHC	
9999	9999	21	93	136	84	69	132	215	241	267	255	258	244	230	198	166	139	105	9999	9999	9999	9999	DTL	
9999	9999	0	46	65	67	63	55	60	94	95	181	227	266	299	297	279	267	220	9999	9999	9999	9999	OAGE	
9999	9999	0	10	17	22	27	33	41	60	54	98	113	120	123	115	101	91	68	9999	9999	9999	9999	NAGE	
9999	9999	280	280	278	276	274	270	269	272	274	272	271	271	270	269	268	267	266	9999	9999	9999	9999	RSST	0
9999	9999	267	279	280	272	267	267	270	273	270	273	273	272	272	272	272	272	274	9999	9999	9999	9999	DSST	1
9999	9999	271	278	279	272	267	268	270	272	270	272	272	272	271	271	272	272	274	9999	9999	9999	9999	DSTA	1
		224	187	169	179	209	221	234	236	221	202	229	204	178	182	195	165	172	9999	9999	9999	9999	U200	
		238	162	153	143	139	170	206	158	108	94	125	102	132	149	159	120	212	9999	9999	9999	9999	U20C	
		99	96	163	101	134	164	126	-29	-39	3	-19	-68	-88	-54	-77	-149	-189	9999	9999	9999	9999	V20C	
		3528	3528	3535	3523	3488	3506	3513	3509	3512	3525	3524	3515	3505	3514	3523	3483	3477	9999	9999	9999	9999	E000	
		113	114	114	107	86	104	109	109	118	125	118	116	120	130	141	111	115	9999	9999	9999	9999	EPOS	
		8	9	8	9	11	10	10	10	10	10	9	10	11	11	10	14	14	9999	9999	9999	9999	ENEG	
		51	48	52	48	22	42	42	42	50	55	46	46	48	47	53	22	26	9999	9999	9999	9999	EPSS	
		16	20	18	18	21	23	21	19	18	22	20	19	19	25	23	28	26	9999	9999	9999	9999	ENSS	
		70	68	70	70	66	69	67	67	68	68	69	70	68	62	59	56	55	9999	9999	9999	9999	RHLO	
		57	53	56	59	53	55	51	51	48	48	46	46	43	37	34	33	32	9999	9999	9999	9999	RHMD	
		50	51	54	59	54	49	48	43	35	35	34	31	26	23	24	25	24	9999	9999	9999	9999	RHHI	
		548	31	23	24	13	28	16	40	23	64	82	77	89	92	85	208	200	79	9999	9999	9999	PSLV	
		7	17	0	-14	-8	14	-9	-37	-60	-47	-32	-52	-29	-30	0	-4	-20	9999	9999	9999	9999	Z850	
		64	65	84	70	51	114	42	27	-7	8	-17	-37	-48	-52	-25	-32	-41	9999	9999	9999	9999	D200	
		4	6	3	6	12	4	10	0	-16	-9	-3	-3	-2	2	7	0	-17	9999	9999	9999	9999	REFC	
		-1	0	-1	-1	-2	-1	0	1	1	1	1	2	3	2	3	4	6	9999	9999	9999	9999	PEFC	
		259	267	266	261	253	265	264	258	255	268	265	258	254	269	269	257	252	9999	9999	9999	9999	T000	
		84	79	81	83	83	77	79	82	85	78	80	83	85	76	77	79	81	9999	9999	9999	9999	R000	
		-27	-24	-31	-20	-20	-8	-16	-10	-8	-2	-18	-8	-8	-3	-7	4	8	9999	9999	9999	9999	Z000	
		207	218	217	227	232	239	242	248	249	250	249	251	248	248	251	251	263	9999	9999	9999	9999	TLAT	
		873	867	862	849	840	844	839	839	843	840	839	839	841	840	837	835	839	9999	9999	9999	9999	TLON	
		67	76	69	63	70	83	67	60	65	62	60	51	41	32	22	18	16	9999	9999	9999	9999	TWAC	
		100	104	106	110	110	138	114	102	109	101	98	84	73	62	42	32	34	9999	9999	9999	9999	TWXC	

Hurricane forecasting basics: our selection

We have not done statistical variable selection. Choice of variable subset is based on theory.

- **CSST**: (climatological) sea surface temperature
- **RHLO**: low-altitude relative humidity
- **T200**: air temperature at 200 mb (very high altitude)
- **SHRD**: wind shear between 850 and 200 mb
- **VMPI**: maximum potential intensity

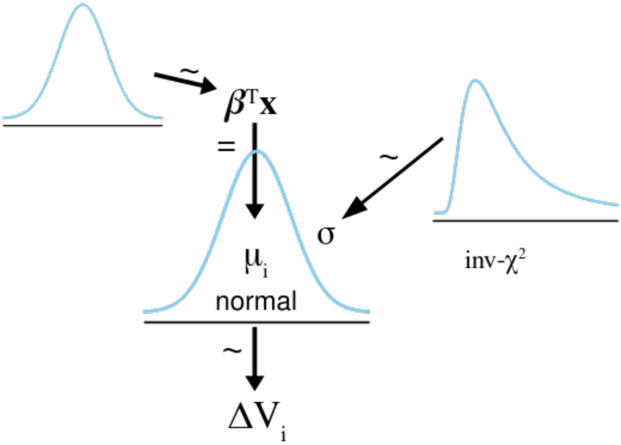


- for testing, we have variable sets *A*, *B*, *C*
- *A*: LAT/LON, VMAX, CSST, SHRD
- *B*: LAT/LON, VMAX, CSST, SHRD, VMPI
- *C*: LAT/LON, VMAX, CSST, SHRD, VMPI, RHLO, T200

Intensity change predictive model

Intensity change predictive model

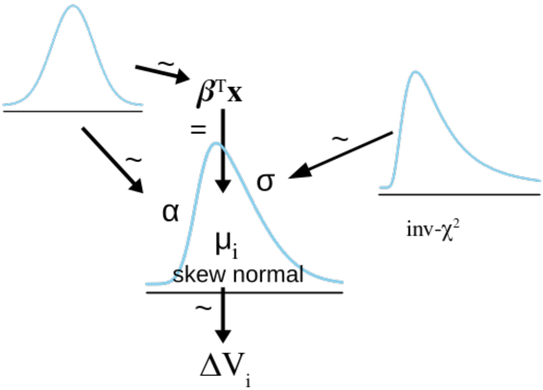
The SHIPS Blunder: a simple linear regression



Intensity change predictive model

Model 2: regression with skewness

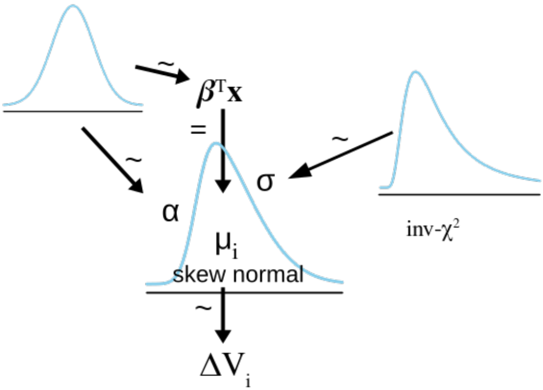
- errors not symmetric around the mean prediction!



Intensity change predictive model

Model 2: regression with skewness

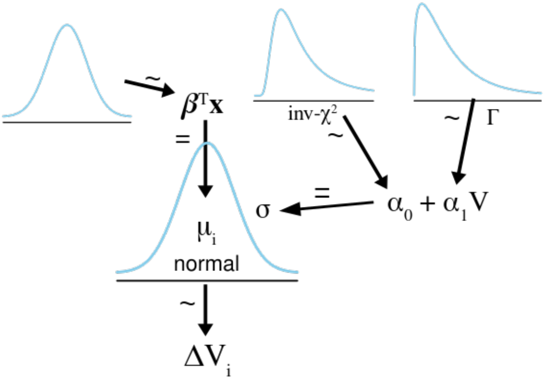
- errors not symmetric around the mean prediction!
- rapid intensification!



Intensity change predictive model

Model 3: regression with a linear model for standard deviation

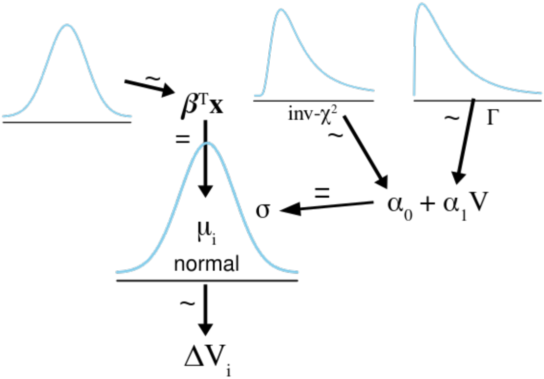
- fewer storms reach higher values of `VMAX`



Intensity change predictive model

Model 3: regression with a linear model for standard deviation

- fewer storms reach higher values of `VMAX`
- allow for higher variance to account for larger historical uncertainty



Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}

Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development

Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development
- reminder: we do 12-hour forecasts

Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development
- reminder: we do 12-hour forecasts
- we standardized all of our data; priors chosen to be weak in standardized scale

Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development
- reminder: we do 12-hour forecasts
- we standardized all of our data; priors chosen to be weak in standardized scale
- SHIPS data: 1982-2019; our restriction: 2017-2019

Models: some remarks

- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development
- reminder: we do 12-hour forecasts
- we standardized all of our data; priors chosen to be weak in standardized scale
- SHIPS data: 1982-2019; our restriction: 2017-2019
- poor problem setup! True model is the laws of physics, but we are fitting a regression

Models: some remarks

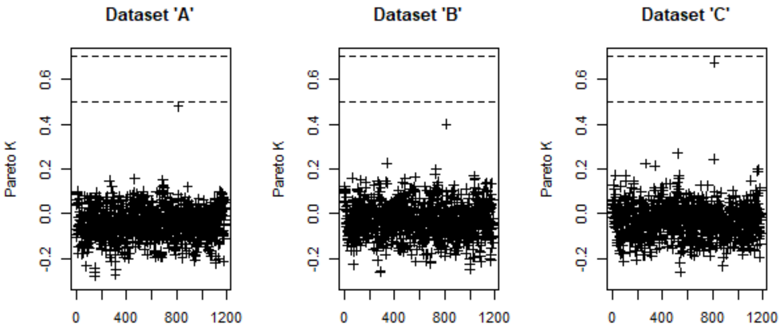
- SHIPS: predict V_{max}; our models: predict ΔV_{max}
- we also decided to include latitude and longitude as variables; geographic region is known to be important for storm development
- reminder: we do 12-hour forecasts
- we standardized all of our data; priors chosen to be weak in standardized scale
- SHIPS data: 1982-2019; our restriction: 2017-2019
- poor problem setup! True model is the laws of physics, but we are fitting a regression
- models were programmed in Stan; sampling with `rstan` resulted in no divergences or issues except for the skew model and the issue was solved by increasing max tree depth to 15

Forecasting: Model Comparison

Forecasting: Model Comparison

Dataset comparison for the **linear regression model** (LOOCV)

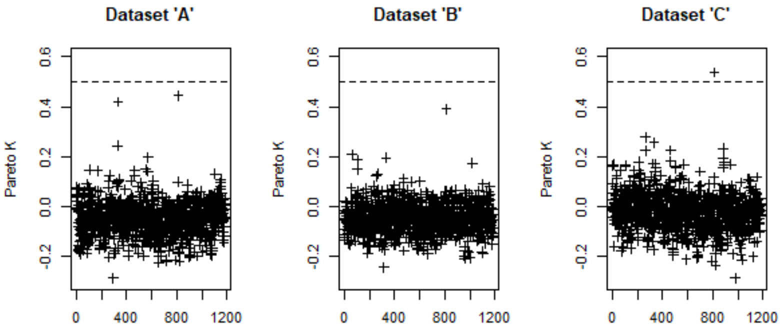
Dataset	**elpd_diff**	**se_diff**
C	0.0	0.0
B	-25.0	6.5
A	-27.4	6.3



Forecasting: Model Comparison (2)

Dataset comparison for the **skewed regression model** (LOOCV)

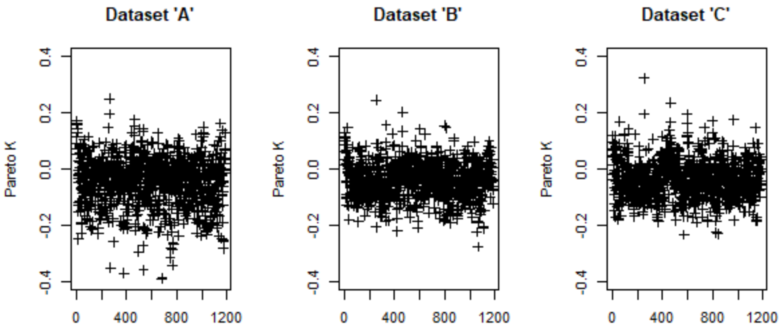
Dataset	**elpd_diff**	**se_diff**
C	0.0	0.0
B	-23.2	6.2
A	-28.7	6.2



Forecasting: Model Comparison (3)

Dataset comparison for the **Changing variance model** (LOOCV)

Dataset	**elpd_diff**	**se_diff**
C	0.0	0.0
B	-32.6	8.2
A	-37.1	8.2



Forecasting: Model Comparison (4)

Model comparison using the Dataset C (LOOCV)

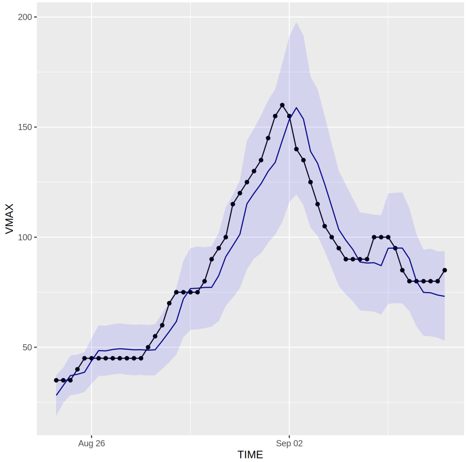
Model	**elpd_diff**	**se_diff**
Variance	0.0	0.0
Skew	-176.3	27.8
Linear	-205.5	34.9

Marginal posteriors

- θ_1: constant term
- θ_2, θ_3: latitude, longitude
- θ_4: sea surface temperature (CSST)
- θ_5: relative humidity (RHLO)
- θ_6: wind shear (SHRD)
- θ_7: maximum potential intensity (VMPI)
- θ_8: air temperature at 200 mb (T200)
- θ_9: intensity at storm core (VMAX)

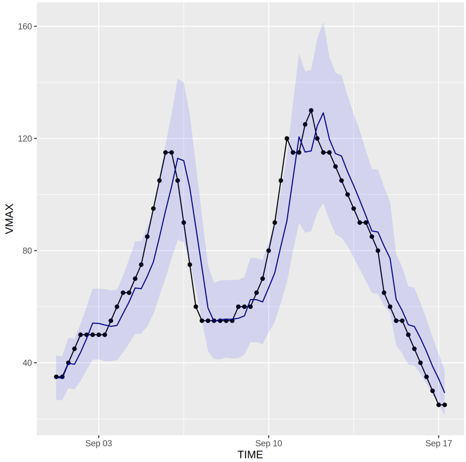
Forecasting: checking predictions

Hurricane Dorian 2019. The image shows a 90% credible interval.
Black dotted line: true VMAX



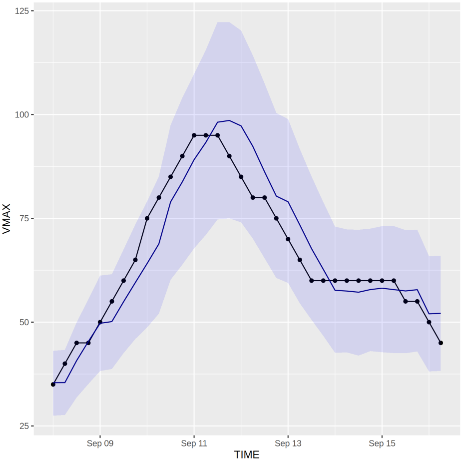
Forecasting: checking predictions

Hurricane Florence 2018. The image shows a 90% credible interval.
Black dotted line: true VMAX



Forecasting: checking predictions

Hurricane Helene 2018. The image shows a 90% credible interval.
Black dotted line: true VMAX



Concluding section

Concluding section

Further development ideas:

- variable selection in full SHIPS dataset

Concluding section

Further development ideas:

- variable selection in full SHIPS dataset
- more time series autoregressive components

Concluding section

Further development ideas:

- variable selection in full SHIPS dataset
- more time series autoregressive components
- use LGEM model (will explain)

A final takeaway

> *"The hurricane will be moving through an environment of low vertical wind shear, a moist mid-level atmosphere, and increasing upper-ocean heat content, and this is ideal for some additional intensification. However, given that the SHIPS models do not show any significant change in the intensity, the NHC forecast keeps Irma a powerful hurricane through five days."* — *NHC 5AM advisory Sep. 5th*

While the NHC published the advisory, Irma was undergoing rapid intensification from 130 to 175 knots over a span of only a few hours.

Conclusions & contact info

- Takeaway: the SHIPS model is terrible!
- Simple changes to the predictive distribution can improve the model
- Further development using Bayesian methods seems promising and there are several possible directions

More at our Github repo: https://github.com/jnlb/bda-hurricane-modeling

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

Contact info:

José Miguel Ramírez

Jonas Lindblad

- rocket-chat: `@jnlb`

Additional information

The SHIPS model:

$$y_i \sim \mathcal{N}(\alpha + X_i \cdot \beta_{N-1}, \sigma), \; i = 1, \dots, r,$$

where we let X_i denote the i:th row of the data, β_{N-1} is an $N-1$-dimensional parameter vector, and r is the number of observations (rows) in the data. Its priors were

$$\begin{bmatrix} \alpha_0 \\ \beta_{N-1,0} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}_N, 10 \cdot \mathbf{I}_N), \sigma_0 \sim \mathrm{Inv}\text{-}\chi^2(\tfrac{1}{10}).$$

The skew-normal regression model:

$$y_i \sim \mathrm{SkewNormal}(\alpha + X_i \cdot \beta_{N-1}, \sigma, \psi), \quad i = 1, \dots, r,$$

with priors

$$\begin{bmatrix} \alpha_0 \\ \beta_{N-1,0} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}_N, 10 \cdot \mathbf{I}_N), \sigma_0 \sim \mathrm{Inv}\text{-}\chi^2(\tfrac{1}{10}), \; \psi_0 \sim \mathcal{N}(0, 1).$$

Additional information

The variance model:

$$y_i \sim \mathcal{N}(\alpha + X_i \cdot \beta_{N-1}, \sigma + \gamma |V_{max,i}|), \ i = 1, \ldots, r,$$

where, again, we use the same notation as before and let $V_{max,i}$ denote the V_{max}-value of the i:th row. The priors that were fed into Stan were

$$\begin{bmatrix} \alpha_0 \\ \beta_{N-1,0} \end{bmatrix} \sim \mathcal{N}(\mathbf{0}_N, 10 \cdot \mathbf{I}_N), \ \sigma_0 \sim \text{Inv-}\chi^2(\tfrac{1}{10}), \ \gamma_0 \sim \Gamma(1, 1).$$

Additional information

The variance regression was run in `rstan` with the following options

```
variance_m <- rstan::stan_model(file = file.path(mod_path,
              "minimal3.stan"))
variance_model <- rstan::sampling(variance_m,
                data = stan_data,
                control = list(max_treedepth = 10),
                iter=4000, seed = SEED)
```