The Umbrella Concept – a simple probability approach.

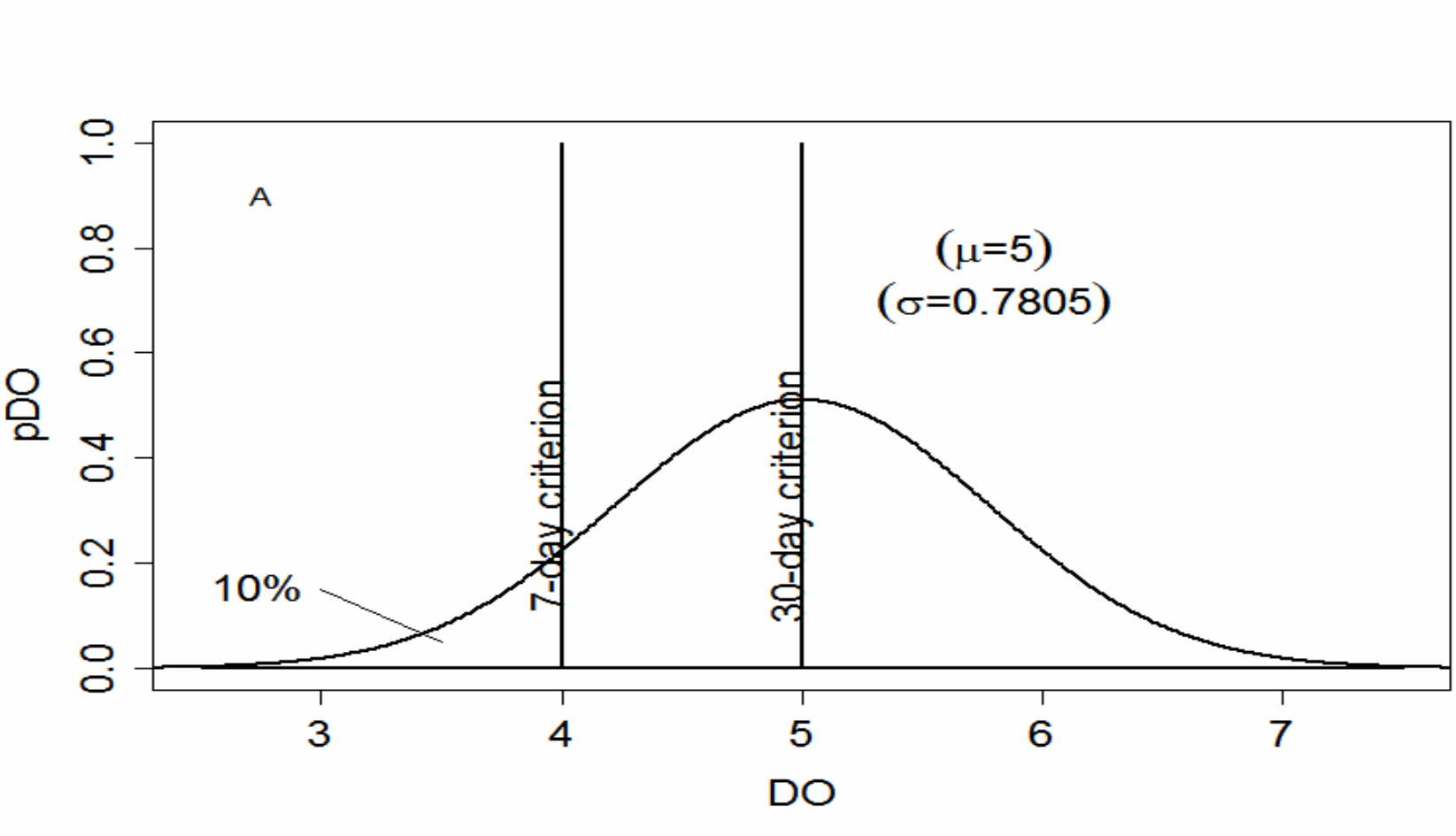
Elgin Perry
CBP – TMAW

Umbrella Workshop
Mar. 16-17, 2011

Two Part Talk

- Risk approach to 30-day mean and umbrella criterion for 7-day mean.
- Risk approach to 30 or 7-day mean as umbrella for instantaneous minimum.

Estimating the standard deviation that creates an umbrella.



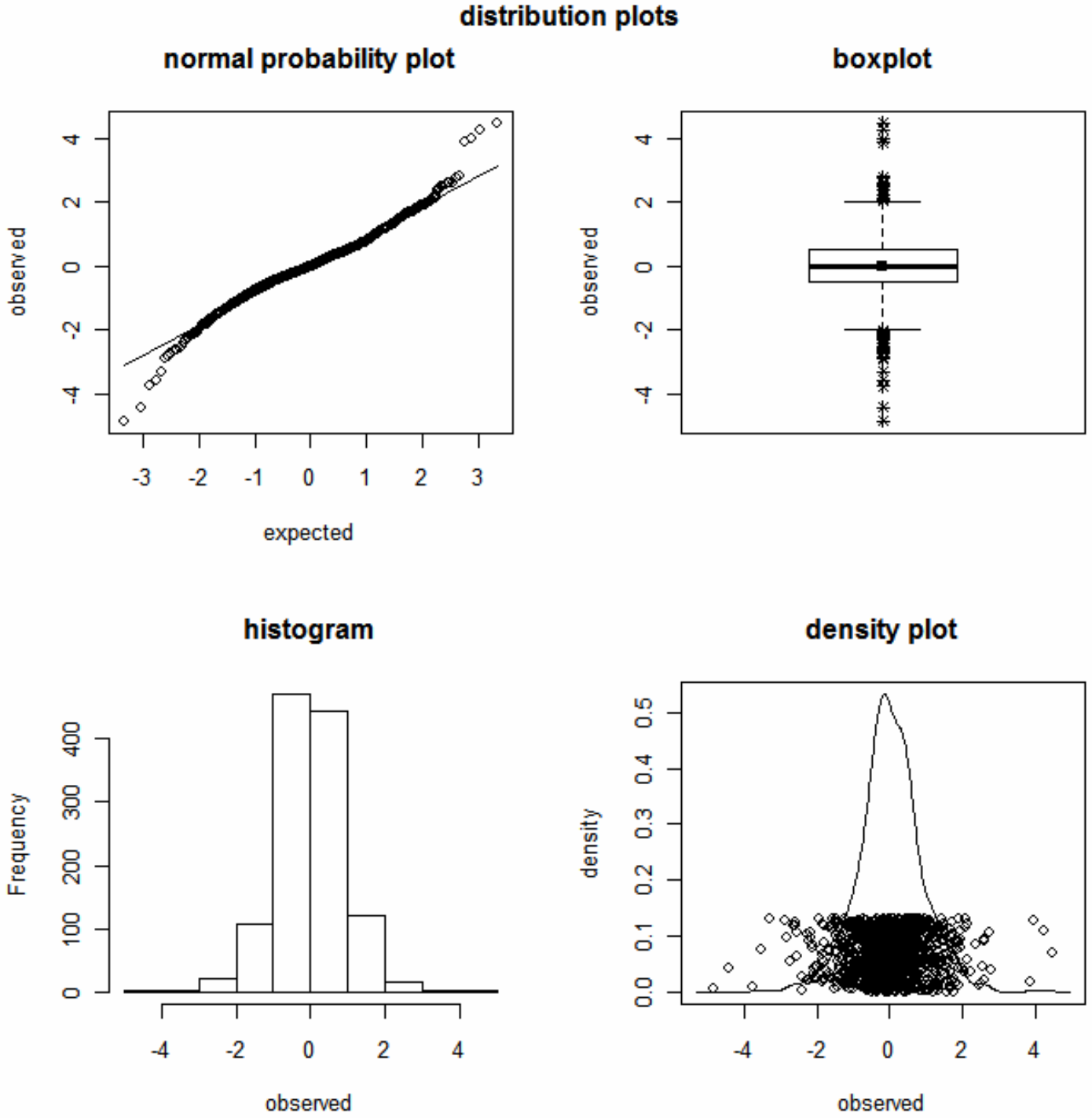
Chebyshev's inequality

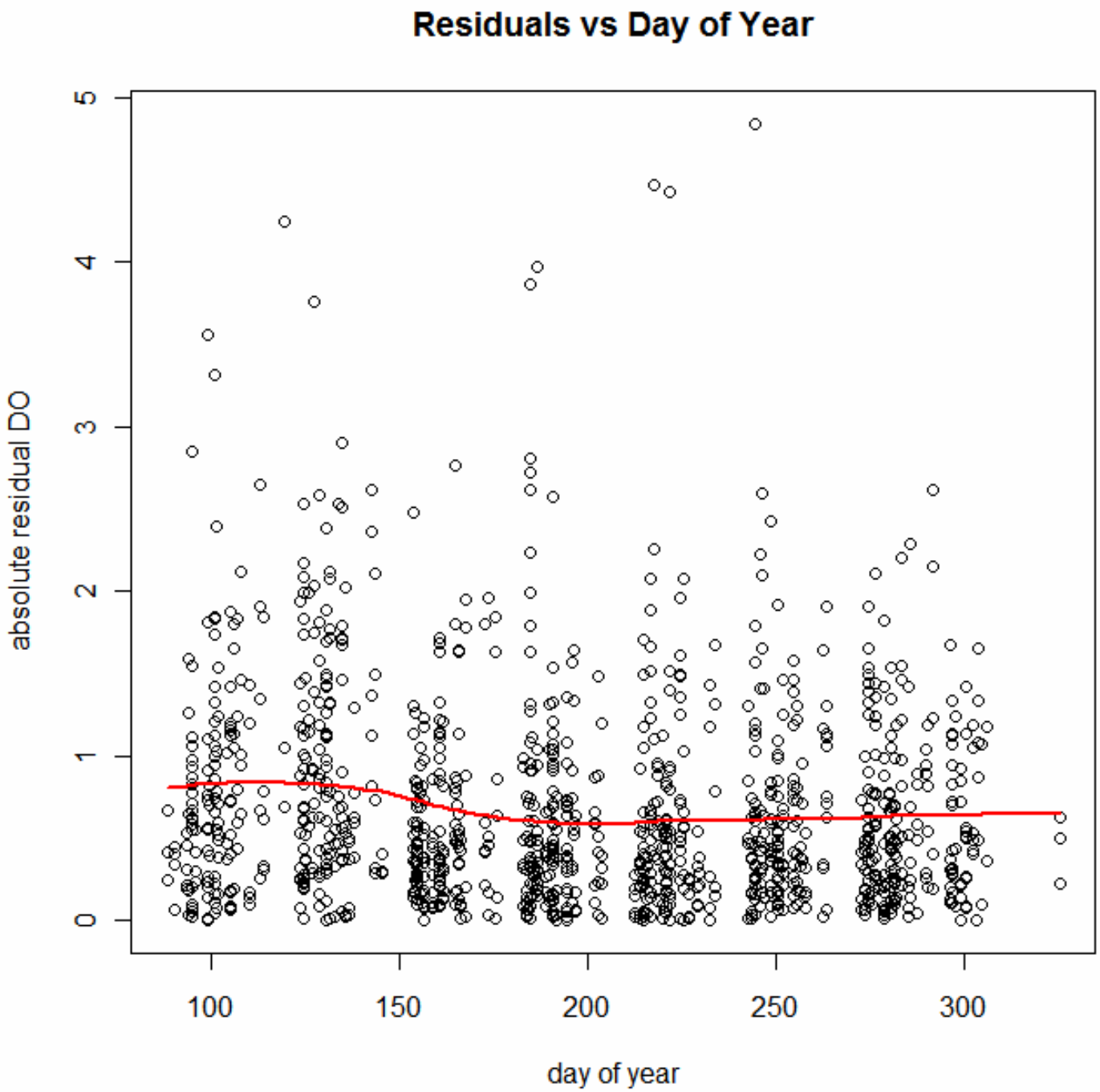
- A one-tailed variant with $k > 0$, is

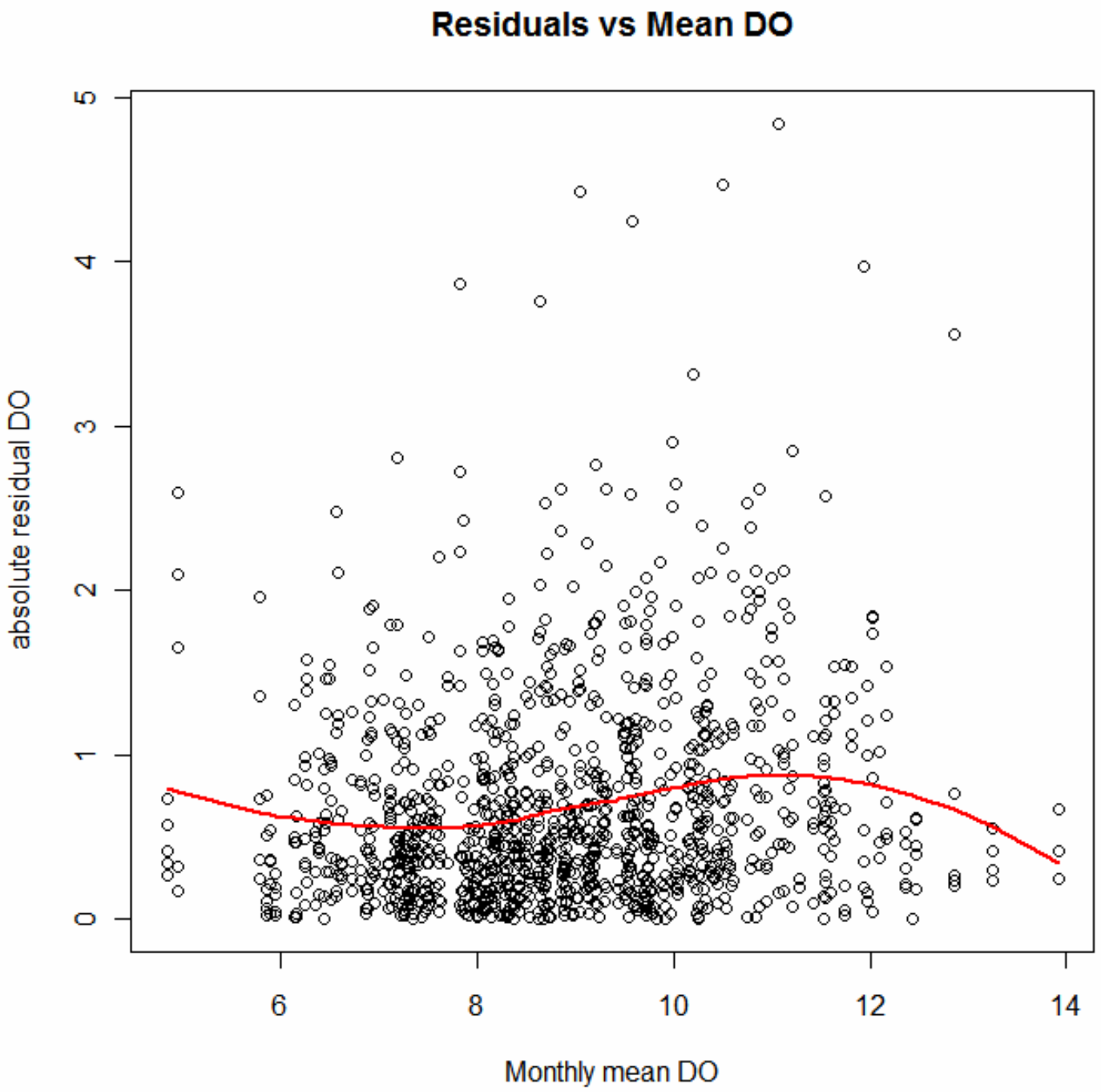
$$\Pr(X - \mu \geq k\sigma) \leq \frac{1}{1+k^2}$$

- X-u = difference between 7-day and 30-day mean
- $1/(1+k^2) = 0.10 \Rightarrow k = 3$
- $k\sigma = 1 \Rightarrow \sigma = 1/3$


Fairfax Station
Clover
Alexandria
Camp Springs
Upper Marlboro
Churchton
McDaniel
Springfield
Burke
Clifton
Rose Hill Farms
Clinton
Tracys Landing
Deale
Wittman
Saint Micha
Manassas
Donovans Corner
Croom
Leitch
Sherwood
Bozman
Piscataway Creek
Newington
Friendship
Pohic
Fenwick
Cheltenham
Rose Haven
Neavitt
Royal
Lorton
Pohic
Dunkirk
North Beach
Tilghman
Occoquan
Brentsville
Hoadly
Hall
Piscataway
Brandywine
Bellevue
Oxfor
Chesapeake Beach
Woodbridge
Accokeek
Sunderland
Randle Cliff Beach
Horsehead
Dale City
Potomac River
Bryans Road
Waldorf
Lower Marlboro
Willows
Mattawoman Creek
Huntingtown
Indian Head
White Plains
Aquasco
Hudson
Pomfret
Hudson Creek
Brooks Creek
Dumfries
Marbury
Hunting Creek
Bryantown
Triangle
Pisgah
Prince Frederick
La Plata
Hughesville
Barstow
Quantico
Madison
Ruby
Port Tobacco
Port Republic
Benedict
Ironsides
Woolford
Garrisonville
Welcome
Charlotte Hall
Saint Leonard
Taylors Island
Nanjemoy
Bel Alton
Mechanicsville
Faulkner
Stafford
Broomes Island
Lusby
Popes Creek
Potomac Creek
Brooke
Helen
Newburg
Chaptico
Loveville
Hollywood
Dowell
Honga
Mount Victoria
Falmouth
Dahlgren
Clements
Solomons
Fredericksburg
Weedonville
Bushwood
Leonardtown
Mile Fork
Monroe Bay
Rosier Creek
King George
Compton
Lexington Park
Sealston
Colonial Beach
Abell
Maryton
Dogue
Callaway
Great Mills
Coltons Point
Jersey
Park Hall
Mattox Creek
Corbin
Valley Lee
Index
Nomini Bay
Saint Marys City
Popes Creek
Port Royal
Stratford
Drayden
Dameron
Guinea
Rollins Fork
Coles Point
Saint Inigoes
Zacata
Ridge
Woodford
Foneswood
Montross
Mount Holly
Scotland
Supply
Loretto
Hague
Yeocomico River
Chance
Bowling Green
Hustle
Kinsale
Milford
Champlain
Acorn
Ladysmith


distribution plots
normal probability plot
observed
expected
-4
-2
0
2
4
-3
-2
-1
0
1
2
3
boxplot
observed
histogram
Frequency
observed
0
100
200
300
400
density plot
density
observed
0.0
0.1
0.2
0.3
0.4
0.5


Residuals vs Day of Year
absolute residual DO
0
1
2
3
4
5
100
150
200
250
300
day of year


Residuals vs Mean DO
absolute residual DO
0
1
2
3
4
5
6
8
10
12
14
Monthly mean DO

Standard Deviation Estimates

Data Used	Estimated SD	Umbrella SD
all data	0.9649	0.7805
30-day mean DO < 8.0	0.8439	0.7805

Risk of 7-day failure | 30-day

30-day mean DO	Prob(|sd=0.9649)	Prob(|sd=0.8439)
5.0	0.1500	0.1180
5.1	0.1271	0.0962
5.2	0.1068	0.0775
5.3	0.0889	0.0617
5.4	0.0734	0.0486
5.5	0.0600	0.0377

Risk approach to Umbrella for Instantaneous Minimum Criterion

- Issue:
  - High degree of autocorrelation
  - Difficult to analytically obtain marginal EDF.
  - Use simulation

Simulation Set Up

- Parse Buoy data into 1-week (or less) time series.
- Fit time series model to each location/week (n=251)
- Estimate mean and V/C matrix for time-series parameter vectors.

2 step simulation

- First - simulate Vector of time-series parameters – based on mean and V/C matrix.

- Second – simulate a one-week time-series based on Vector from step 1.

Time-Series Model – 7 parameters

b_Int - the intercept,
b_cday - linear trend term,
b_sin, b_cos - diel trend coefficients
b_ar1,b_ar2 - autoregressive terms
mse - residual mean square error

MANOVA of Time-Series Vectors – Just CB4

(b_int,b_cday,b_sin,b_cos,b_AR1,b_AR2,mse)

Source	Pillai's Trace	Pr > F
month	1.9229	<.0001
TotDep	0.6465	<.0001
SampDep	0.5022	<.0001

Vector Mean for DO > 5.0

May	Jun	Jul	Aug	Sep	Oct	WaterDepth	SensorDepth
0	0	1	0	0	0	10	6

b_Int	b_cday	b_sin	b_cos	b_AR1	b_AR2	mse
5.0058	-0.0493	-0.4072	-0.0527	0.9333	-0.0319	0.3164

Adjusting mean DO

Sensor Depth	b_Int	b_cday	b_sin	b_cos	b_AR1	b_AR2	mse
6	5.0058	-0.0493	-0.4072	-0.0527	0.9333	-0.0319	0.3164
5	5.6733	-0.0476	-0.5114	0.0094	0.9328	-0.0294	0.4112
4	6.3408	-0.0460	-0.6156	0.0714	0.9324	-0.0268	0.5060
3	7.0082	-0.0443	-0.7198	0.1335	0.9320	-0.0243	0.6008

Partial Correlation Matrix

DF = 49	b_Int	b_cday	b_sin	b_cos	b_AR1	b_AR2	MSE
b_Int	1.00	-0.023 0.87	0.021 0.884	0.259 0.068	-0.166 0.248	0.267 0.0602	-0.066 0.6443
b_cday	-0.023 0.872	1.000	0.168 0.242	0.447 0.001	0.113 0.434	-0.087 0.5467	0.034 0.8144
b_sin	0.021 0.884	0.168 0.242	1.000	0.061 0.673	-0.228 0.109	0.120 0.4064	-0.137 0.3408
b_cos	0.259 0.068	0.447 0.001	0.061 0.673	1.000	-0.109 0.486	0.171 0.2330	-0.129 0.3694
b_AR1	-0.166 0.248	0.113 0.434	-0.228 0.109	-0.100 0.486	1.000	-0.90 <.0001	-0.073 0.6129
b_AR2	0.267 0.060	-0.087 0.546	0.120 0.406	0.171 0.233	-0.900 <.0001	1.000	-0.074 0.6084
MSE	-0.066 0.644	0.034 0.814	-0.137 0.3408	-0.129 0.369	-0.073 0.612	-0.074 0.6084	1.000

		7-day mean >= 4.0	7-day mean < 4.0	row marginal
Table 10a. Sensor Depth = 6 mean DO = 5.0058	failure Instantaneous minimum < 10%	542 65.86 %	9 5.08 %	551 55.1 %
	failure Instantaneous minimum > 10%	281 34.14 %	168 94.9 %	449 44.9 %
	column marginal	823 100%	177 100%	1000 100%

		7-day mean >= 4.0	7-day mean < 4.0	row marginal
Table 10b. Sensor Depth = 5 mean DO = 5.6733	failure Instantaneous minimum < 10%	656 69.42 %	1 1.8 %	657 65.7 %
	failure Instantaneous minimum > 10%	289 30.58 %	54 98.2 %	343 34.3 %
	column marginal	945 100%	55 100%	1000 100%

		7-day mean >= 4.0	7-day mean < 4.0	row marginal
Table 10d. Sensor Depth = 3 mean DO = 7.0082	failure Instantaneous minimum < 10%	834 83.6 %	0 0 %	834 83.4 %
	failure Instantaneous minimum > 10%	164 16.4 %	2 100 %	166 16.6 %
	column marginal	998 100%	2 100%	1000 100%