Compare VBGFs

Preliminaries

Load Necessary Packages

```
> library(FSA)        # for filterD(), headtail(), col2rgbt(), vbFuns(), vbStart()
> library(dplyr)      # for mutate()
> library(nlstools)   # for nlsBoot()
> library(AICcmodavg) # for aictab()
```

Load Data and Make Some Preparations

```
> # Set your working directory to where your external data files (and scripts) are located.
> setwd("C:/aaaWork/Web/GitHub/RcourseNunavut2016/Handouts")
> dSC <- read.csv("SawyerCo_reduced.csv")
> wae <- filterD(dSC,waterbody %in% c("LAKE CHIPPEWA","SAND LAKE"),
                 species=="Walleye",!is.na(len),!is.na(age))
```

```
> xlbl <- "Age (yrs)"
> ylbl <- "Total Length (in)"
> clr1 <- c("black","blue")
> clr2 <- col2rgbt(clr1,1/5)
```

Quick Summaries

```
> ( sum <- Summarize(len~age+waterbody,data=wae,digits=1) )
Warning: First RHS variable was converted to a factor.
```

	age	waterbody	n	nvalid	mean	sd	min	Q1	median	Q3	max	percZero
1	2	LAKE CHIPPEWA	1	1	297.0	NA	297	297.0	297.0	297.0	297	0
2	3	LAKE CHIPPEWA	40	40	328.6	35.7	264	301.5	326.5	356.5	404	0
3	4	LAKE CHIPPEWA	17	17	387.9	18.4	361	376.0	386.0	396.0	429	0
4	5	LAKE CHIPPEWA	18	18	403.4	39.5	343	380.0	395.0	410.5	498	0
5	6	LAKE CHIPPEWA	21	21	460.9	50.4	358	424.0	472.0	485.0	546	0
6	7	LAKE CHIPPEWA	20	20	446.1	52.5	356	416.2	437.0	486.2	531	0
7	8	LAKE CHIPPEWA	11	11	476.7	33.4	419	456.0	475.0	492.5	533	0
8	9	LAKE CHIPPEWA	6	6	468.2	27.5	429	449.5	473.5	488.5	498	0
9	10	LAKE CHIPPEWA	1	1	470.0	NA	470	470.0	470.0	470.0	470	0
10	4	SAND LAKE	16	16	324.8	12.0	307	314.2	324.0	338.0	343	0
11	5	SAND LAKE	9	9	359.6	16.4	343	348.0	361.0	363.0	396	0
12	6	SAND LAKE	16	16	394.1	23.8	356	380.0	394.0	409.0	432	0
13	7	SAND LAKE	4	4	399.2	36.5	368	381.5	388.5	406.2	452	0
14	8	SAND LAKE	6	6	424.8	26.3	386	411.2	428.0	436.5	462	0
15	9	SAND LAKE	6	6	442.5	23.8	406	429.0	450.0	455.2	470	0

```
> plot(len~age,data=wae,pch=19,col=clr2[waterbody],xlab=xlbl,ylab=ylbl)
> lines(mean~fact2num(age),data=filterD(sum,waterbody=="LAKE CHIPPEWA"),col=clr1[1],lwd=2)
> lines(mean~fact2num(age),data=filterD(sum,waterbody=="SAND LAKE"),col=clr1[2],lwd=2)
```
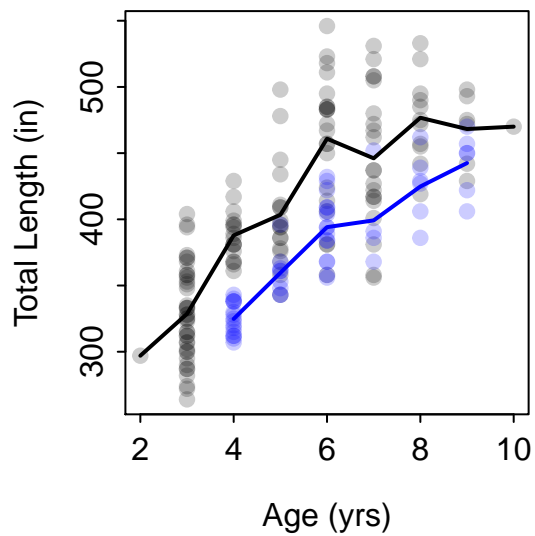

Fitting Most Complex Model and Checking Assumptions

```
> ( svOm <- vbStarts(len~age,data=wae,plot=TRUE) )
```
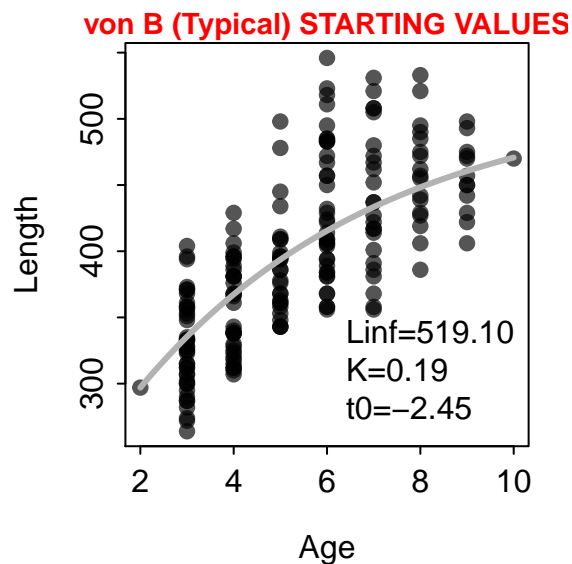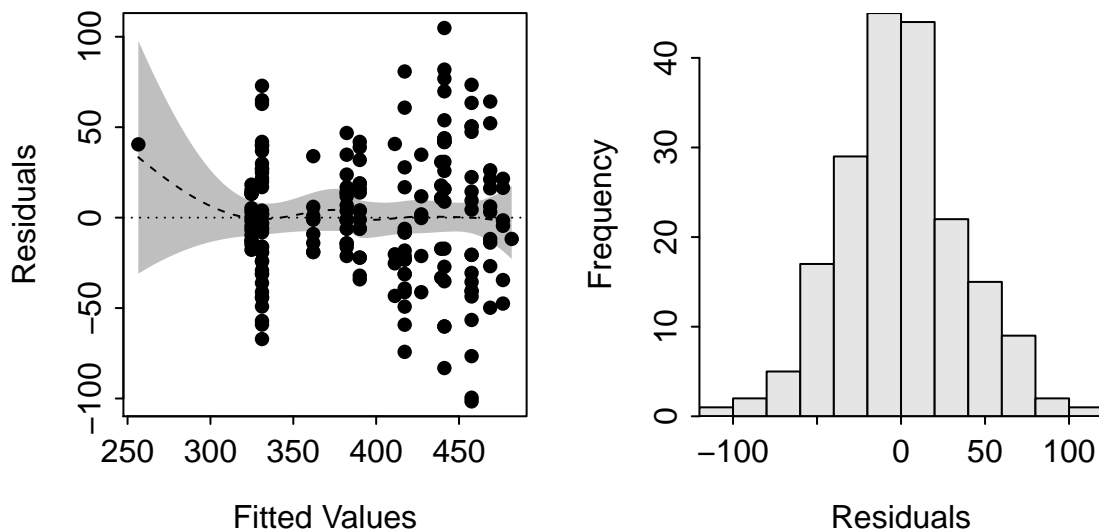


```
$Linf
[1] 519.1043

$K
[1] 0.1905709

$t0
[1] -2.454813
```

```
> ( svLKt <- Map(rep,svOm,c(2,2,2)) )
$Linf
[1] 519.1043 519.1043

$K
[1] 0.1905709 0.1905709

$t0
[1] -2.454813 -2.454813
```

```
> vbLKt <- len~Linf[waterbody]*(1-exp(-K[waterbody]*(age-t0[waterbody])))
> fitLKt <- nls(vbLKt,data=wae,start=svLKt)
> residPlot(fitLKt)
```


Are There Any Differences?

```
> vbOm <- len~Linf*(1-exp(-K*(age-t0)))
> fitOm <- nls(vbOm,data=wae,start=svOm)
```

```
> extraSS(fitOm,com=fitLKt,sim.name="{Omega}",com.name="{Linf,K,t0}")
Model 1: {Omega}
Model A: {Linf,K,t0}

    DfO    RSSO DfA    RSSA Df      SS      F    Pr(>F)
1vA 189 338461 186  241781  3  96679 24.791 1.535e-13
```

```
> lrt(fitOm,com=fitLKt,sim.name="{Omega}",com.name="{Linf,K,t0}")
Model 1: {Omega}
Model A: {Linf,K,t0}

    DfO  logLikO DfA  logLikA Df   logLik  Chisq Pr(>Chisq)
1vA 189 -990.004 186 -957.712  3  -32.292 64.584  6.157e-14
```

Is the Most Complex Model Warranted?

```
> vbLK <- len~Linf[waterbody]*(1-exp(-K[waterbody]*(age-t0)))
> ( svLK <- Map(rep,svOm,c(2,2,1)) )
$Linf
[1] 519.1043 519.1043

$K
[1] 0.1905709 0.1905709

$t0
[1] -2.454813
```

```
> fitLK <- nls(vbLK,data=wae,start=svLK)
> vbLt <- len~Linf[waterbody]*(1-exp(-K*(age-t0[waterbody])))
> svLt <- Map(rep,svOm,c(2,1,2))
> fitLt <- nls(vbLt,data=wae,start=svLt)
> vbKt <- len~Linf*(1-exp(-K[waterbody]*(age-t0[waterbody])))
> svKt <- Map(rep,svOm,c(1,2,2))
> fitKt <- nls(vbKt,data=wae,start=svKt)
> extraSS(fitLK,fitLt,fitKt,com=fitLKt,com.name="{Linf,K,t0}",
          sim.names=c("{Linf,K}","{Linf,t0}","{K,t0}"))
Model 1: {Linf,K}
Model 2: {Linf,t0}
Model 3: {K,t0}
Model A: {Linf,K,t0}

    Df0        RSS0 DfA        RSSA Df          SS      F Pr(>F)
1vA 187 241786.383 186 241781.303  1       5.081 0.0039 0.9502
2vA 187 242080.068 186 241781.303  1     298.765 0.2298 0.6322
3vA 187 241871.187 186 241781.303  1      89.884 0.0691 0.7929
```

Can the Model be Reduced to Only One Parameter that Differs?

```
> vbL <- len~Linf[waterbody]*(1-exp(-K*(age-t0)))
> ( svL <- Map(rep,svOm,c(2,1,1)) )
$Linf
[1] 519.1043 519.1043

$K
[1] 0.1905709

$t0
[1] -2.454813
```

```
> fitL <- nls(vbL,data=wae,start=svL)
> vbK <- len~Linf*(1-exp(-K[waterbody]*(age-t0)))
> svK <- Map(rep,svOm,c(1,2,1))
> fitK <- nls(vbK,data=wae,start=svK)
Error in nls(vbK, data = wae, start = svK): step factor 0.000488281 reduced below 'minFactor' of 0.000976562
```

```
> fitK <- nls(vbK,data=wae,start=svK,control=list(minFactor=1e-15,maxiter=500))
```

```
> extraSS(fitL,fitK,com=fitLK,com.name="{Linf,K}",sim.names=c("{Linf}","{K}"))
Model 1: {Linf}
Model 2: {K}
Model A: {Linf,K}

    DfO    RSSO DfA    RSSA Df      SS      F  Pr(>F)
1vA 188 246207.2 187 241786.4  1  4420.8 3.4191 0.06602
2vA 188 242717.7 187 241786.4  1   931.3 0.7203 0.39713
```

Summarize Final Model

```
> summary(fitK,correlation=TRUE)

Formula: len ~ Linf * (1 - exp(-K[waterbody] * (age - t0)))

Parameters:
       Estimate Std. Error t value Pr(>|t|)
Linf  4.922e+02  1.711e+01  28.774  < 2e-16
K1    3.743e-01  8.318e-02   4.500 1.19e-05
K2    2.638e-01  4.688e-02   5.627 6.57e-08
t0   -9.452e-04  5.156e-01  -0.002    0.999

Residual standard error: 35.93 on 188 degrees of freedom

Correlation of Parameter Estimates:
   Linf  K1    K2
K1 -0.95
K2 -0.94  0.98
t0 -0.87  0.97  0.96

Number of iterations to convergence: 16
Achieved convergence tolerance: 1.431e-06
```

```
> round(cbind(coef(fitK),confint(fitK)),3)
Waiting for profiling to be done...
              2.5%   97.5%
Linf 492.234 467.551 539.594
K1     0.374   0.230   0.540
K2     0.264   0.175   0.352
t0    -0.001  -1.277   0.768
```

```
> vb <- vbFuns("typical")
> # Lake Chippewa
> crLC <- filterD(wae,waterbody=="LAKE CHIPPEWA")
> svLC <- list(Linf=492,K=0.37,t0=0)
> fitLC <- nls(len~vb(age,Linf,K,t0),data=crLC,start=svLC)
> # Sand Lake
> crSL <- filterD(wae,waterbody=="SAND LAKE")
> svSL <- list(Linf=492,K=0.26,t0=0)
> fitSL <- nls(len~vb(age,Linf,K,t0),data=crSL,start=svSL)
```

```
> offset <- 0.08
> # Lake Chippewa
> plot(len~I(age-offset),data=crLC,pch=19,col=clr2[1],ylim=c(200,550),xlab=xlbl,ylab=ylbl)
```

```
> curve(vb(x-offset,coef(fitLC)),from=2,to=10,col=clr1[1],lwd=2,add=TRUE)
> # Sand Lake
> points(len~I(age+offset),data=crSL,pch=19,col=clr2[2])
> curve(vb(x+offset,coef(fitSL)),from=4,to=9,col=clr1[2],lwd=2,add=TRUE)
> legend("bottomright",levels(wae$waterbody),pch=19,col=clr1,bty="n",cex=0.8)
```
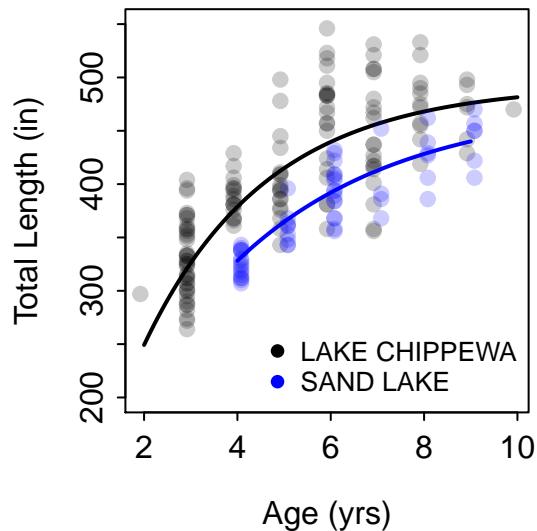

Using Information Criterion

Fit the Only Other Model not Fit Above

```
> vbt <- len~Linf*(1-exp(-K*(age-t0[waterbody])))
> svt <- Map(rep,svOm,c(1,1,2))
> fitt <- nls(vbt,data=wae,start=svt)
```

AICc Table

```
> ms <- list(fitOm,fitL,fitK,fitt,fitLK,fitLt,fitKt,fitLKt)
> mnames <- c("{Omega}","{Linf}","{K}","{t0}","{Linf,K}","{Linf,t0}","{K,t0}","{Linf,K,t0}")
> aictab(ms,mnames)

Model selection based on AICc:

              K    AICc Delta_AICc AICcWt Cum.Wt      LL
{K}           5 1926.49       0.00   0.35   0.35 -958.08
{Linf,K}      6 1927.88       1.39   0.17   0.52 -957.71
{K,t0}        6 1927.95       1.46   0.17   0.69 -957.75
{Linf,t0}     6 1928.12       1.63   0.16   0.85 -957.83
{Linf}        5 1929.23       2.74   0.09   0.94 -959.45
{Linf,K,t0}   7 1930.03       3.54   0.06   1.00 -957.71
{t0}          5 1935.54       9.05   0.00   1.00 -962.61
{Omega}       4 1988.22      61.73   0.00   1.00 -990.00
```