

# **A Comparison of Three Procedures for Creating Bivariate Pairs to be Analyzed with Least Squares Regression**

Adam Craig  
Department of Mathematical Sciences  
Montana State University

June 4, 2004

A writing project submitted in partial fulfillment  
of the requirements for the degree

Master of Science in Statistics

In May of 1996, whirling disease was detected in Little Prickly Pear Creek near Helena, Montana. Since then, the disease has become widespread in the Missouri River tailwater below Holter Dam (Leathe et al. 2002). The outbreak of whirling disease is of high concern because this section of the Missouri River is a reputed rainbow trout fishery. Whirling disease has caused significant declines in other prized trout fisheries largely due to greatly diminished recruitment of juvenile and young of the year rainbow trout (Vincent 1996).

Rainbow trout in this section of the river primarily use three tributaries for spawning (Little Prickly Pear Creek, Sheep Creek, and the Dearborn River), each with varying risks of whirling disease infection (Munro 2004). Adult rainbow trout in the mainstem Missouri River are assumed to be the progeny of these tributary-spawning fish. The relative extent of each tributary's contribution to the mainstem population of rainbow trout is unknown, however (Munro 2004). The doctoral research of Andrew Munro at Montana State University attempted to quantify the contribution of each of these tributaries to the mainstem population. The information gained from this research was hoped to provide the means for targeting tributaries contributing higher numbers of trout. Knowledge of such could reveal where whirling disease is causing the most detriment to the Missouri River mainstem population of rainbow trout (Munro 2004).

Otolith microchemistry was used to quantify the proportions of fish born in the various tributaries currently living in the Missouri River. An otolith is a calcified structure in the inner ear of fish. The information storage capability of otoliths has been recognized for many years (Munro 2004). For example, the Fisheries Agency of Japan (1967) realized that otolith microchemistry could be used to trace the movements of

Pacific salmon. Approximately 2% of an otolith is composed of trace elements (Campana 1999). These trace elements, namely Sr:Ca ratios are the basis of otolith microchemistry. Trace elements in the aquatic environment in which a fish lives are incorporated onto the surface of the otolith as the fish grows (Campana 1999). Using Time of Flight Secondary Ion Mass Spectrometry, Dr. Munro was able to quantify the Sr:Ca ratio of an otolith and compare that to the Sr:Ca ratio in the water of the fish's natal tributary (Munro 2004).

The remainder of this manuscript elucidates the nature of the data collected by Dr. Munro and the problems with using conventional least-squares regression to analyze them. Besides data collected from the Missouri River, similar data were collected from three lakes (Heart, Lewis, Yellowstone) in Yellowstone National Park. As an alternative to conventional least-squares regression, three resampling procedures will be described and their results compared for both the Missouri River and Yellowstone National Park datasets.

In both datasets, the data were composed of Sr:Ca ratios for a given water body as well as Sr:Ca ratios for otoliths coming from that body of water. Initial analyses performed by Dr. Munro included a least-squares regression with Sr:Ca ratios of water samples as an explanatory variable and Sr:Ca ratios of otolith samples as the response. The relationship found therein was hoped to explain how the chemistry of a fish's environment is reflected by a fish's otolith chemistry. The regression performed resulted in an ostensibly well-fitting linear model ( $Sr:Ca_{oto} = 0.3234 * Sr:Ca_{water} - 0.1285$ ,  $r^2 = 0.9677$ ,  $p\text{-value} < 0.0001$ ) (Munro 2004). The results of this regression were not based

on actual bivariate data, however. The pairs were in fact pairs of mean water chemistry and mean otolith chemistry from each water body.

Table 1. Summary statistics for otolith and water Sr:Ca ratio measurements by location.

Otolith Sr:Ca Ratios			
Location	n	Mean	Std. Dev.
Dearborn	39	0.4304	0.1010
Little Prickly Pear	31	0.4410	0.0926
Sheep	40	0.6283	0.1146
Missouri	13	0.6971	0.2257
Lewis	29	0.3939	0.2602
Heart	29	0.5380	0.1103
Yellowstone	42	0.8966	0.1277

Water Sr:Ca Ratios			
Location	n	Mean	Std. Dev.
Dearborn	9	1.7428	0.6738
Little Prickly Pear	30	2.0667	0.6800
Sheep	9	2.9674	0.2350
Missouri	12	3.0107	0.2312
Lewis	4	1.5345	0.7096
Heart	4	2.3974	0.2268
Yellowstone	8	3.9728	0.1017

A number of additional problems exist with these data. To begin, the data are not truly paired measurements on an individual fish. That is to say the chemistry measurement of the water body is not paired with each otolith chemistry measurement. Typically water samples were collected and analyzed up to a year prior to the otoliths being collected and analyzed. Unequal sample sizes of water chemistry measurements and otolith chemistry measurements are present among the different water bodies. Finally, the variability within water and otolith chemistry measurements is not homogenous among the water bodies.

Although, the data do not conform to the assumptions of least-squares regression, we cannot ignore the presence of a relationship between water and otolith Sr:Ca ratios. As a result, the goal of this exercise was to explore and compare several means of creating bivariate pairs of water and otolith chemistry measurements to be used in least-squares regression.

The procedures performed were centered on use of the delete-1 jackknife. Typically the delete-1 jackknife is a technique for estimating the bias and standard error of an estimate (Efron and Tibshirani, 1993). We felt use of this technique was an initial attempt toward finding a solution for modeling the data collected by Dr. Munro. Support for using jackknifing comes from the fact that jackknife datasets are more similar on average to their original datasets and no information is lost when the estimator used is a linear statistic (Efron and Tibshirani, 1993).

The first procedure involved creating delete-1 jackknife mean datasets of water Sr:Ca ratios for each water body. As each individual observation was deleted from a water body's Sr:Ca ratio measurements, a mean of the remaining  $n-1$  observations was

calculated and stored. The resulting new dataset was composed of n delete-1 jackknife means for all water bodies. Next, each otolith measurement from a particular water body was iteratively paired with all of the delete-1 jackknife mean water measurements from the same water body. Separate least-squares lines and regression estimates were calculated for each iteration. The Matlab code used to calculate these jackknife datasets, a plot of the least-squares lines, relative frequency histograms of the regression estimates, scatterplot of the slope versus intercept estimates, and scatterplot of the jackknife dataset are contained in Appendix I. Figures 1A-1E are in regards to the Missouri River data and Figures 2A-2E regard the Yellowstone Natl. Park lakes. Summary statistics of the analyses contained in Appendix I are presented in Table 2.

Table 2. Summaries of regression estimates and coefficients of determination when responses were paired with each delete-1 jackknife mean from its respective water body.

#### Missouri River & Tributaries

Estimate	Mean	Variance
Intercept	0.088	9.0e-05
Slope	0.1857	1.3e-04
$r^2$	0.41	6.0e-05

#### Yellowstone National Park Lakes

Estimate	Mean	Variance
Intercept	0.0555	0.0031
Slope	0.2102	2.35e-04
$r^2$	0.6190	3.25e-05

The second procedure was similar to the first. Delete-1 jackknife mean datasets of otolith Sr:Ca ratios for each water body were created, however delete-1 jackknife mean datasets of water Sr:Ca ratios were also created. From these datasets, 10,000 random pairs of mean otolith and mean water chemistry measurements were iteratively selected. Again, separate least-squares lines and regression estimates were calculated for

each iteration. Appendix II contains the Matlab code used to calculate these jackknife datasets, a plot of the least-squares lines, relative frequency histograms of the regression estimates, scatterplot of the slope versus intercept estimates, and scatterplot of the jackknife dataset. Figures 3A-3E are in regards to the Missouri River data and Figures 2A-2E regard the Yellowstone Natl. Park lakes. Summary statistics of the analyses contained in Appendix II are presented in Table 3.

Table3. Summaries of regression estimates and coefficients of determination when delete-1 jackknife mean otolith measurement was paired with delete-1 jackknife mean water measurement from the same water body.

#### Missouri River & Tributaries

Estimate	Mean	Variance
Intercept	0.053	0.0012
Slope	0.2029	2.0e-04
$r^2$	0.93	5.0e-04

#### Yellowstone National Park Lakes

Estimate	Mean	Variance
Intercept	0.0565	0.0034
Slope	0.2093	2.74e-04
$r^2$	0.9904	1.26e-04

The final procedure did not require creating any delete-1 jackknife mean data sets. Using the same original mean pairs that Dr. Munro used, we iteratively deleted a single pair and calculated the least-squares line and regression estimates based on the remaining n-1 mean pairs. See Appendix III for Matlab code and plots of the least squares lines. Figure E regards the Missouri River data and Figure F regards the Yellowstone Natl. Park lakes. Summary statistics of the analyses contained in Appendix III are presented in Table 4.

Table 4. Summaries of regression estimates and coefficients of determination when an original mean response and mean water measurement pair was deleted.

Missouri River & Tributaries		
Estimate	Mean	Variance
Intercept	0.042	0.0067
Slope	0.2065	0.0011
$r^2$	0.95	6.0e-06

Yellowstone National Park Lakes		
Estimate	Mean	Variance
Intercept	0.0692	0.0053
Slope	0.2002	9.46e-04
$r^2$	1.0	0.0

#### Discussion:

From the three resampling procedures, the most salient result is the similarity among the slope estimates. All estimates for both the Missouri River and the Yellowstone Natl. Park lakes were approximately 0.20. This result supports Dr. Munro's findings of a significant positive association between water Sr:Ca ratios and otolith Sr:Ca ratios. Examination of the plots of the least-squares lines also supports this because all lines had positive slope.

We can see that the  $r^2$  values increased dramatically when pairs of mean water chemistry and mean otolith chemistry were used. This result is not surprising because by regressing pairs of means we have eliminated a large proportion of the variability within the otolith and water chemistry measurements. We should note the  $r^2$  value of 1.0 for the Yellowstone Natl. Park lakes when deleting one of the original pairs of means. Because only 3 lakes were in the dataset, the deletion of one pair of means left only two pairs for regression rendering an  $r^2$  value of 1.0.



This work has not provided a sure solution as how to analyze data of this nature. Rather, the results lend support to the nature of how environmental Sr:Ca ratios are reflected by otolith Sr:Ca ratios. Future research into this type of analysis may require simulations of unpaired data with significant linear association. Such work may reveal the distribution of regression parameters under these circumstances and provide the means for estimation.

## References

- Campana, S. E. 1999. Chemistry and composition of fish otoliths: pathways, mechanisms and applications. *Marine Ecology Progress Series* 188:263-297.
- Efron, B. and R. J. Tibshirani. 1993. *An introduction to the bootstrap*. Chapman & Hall, New York.
- Fisheries Agency of Japan. 1967. Study on identification of the Pacific salmon by neutron activation analysis. Survey and Research Division, Fisheries Agency of Japan. (Translated from Japanese by the Fisheries Research Board of Canada Translation Series No. 1655).
- Leathe, S. A., T. Horton, T. E. McMahon, A. R. Munro, and A. V. Zale. 2002. Interactions of life history, fish size, and infection risk on population-level effects of whirling disease in wild rainbow trout. Final Report to the Whirling Disease Initiative of the National Partnership on Management of Wild and Native Coldwater Fisheries, Bozeman, Montana.
- Munro, A. R. 2004. Identification of life history variation in salmonids using otolith microchemistry and scale growth patterns: implications for illegal introductions and for whirling disease in Missouri River rainbow trout. Doctoral Dissertation. Montana State University, Bozeman, Montana.
- Vincent, E. R. 1996. Whirling disease and wild trout: the Montana experience. *Fisheries* 21(6):32-33.

## Appendix I.

Matlab source code and outputted figures when regressing otolith Sr:Ca ratios versus delete-1 jackknife mean water Sr:Ca ratios.

```

load chemistry.dat
X=[chemistry(:,1), chemistry(:,2), chemistry(:,3)];
location=X(:,1);
otolith=X(:,2);
water=X(:,3);
length(X);
n=223;

% Designate vectors for jackknifing
DBx = []; HLx=[]; LLx=[]; LPx=[]; MRx=[]; SCx=[]; YLx=[];

g=isnan(water);% Makes index of 0s & 1s where NaN=1

% Create counter indices for each jackknife vector space
k1=1; k2=1; k3=1; k4=1; k5=1; k6=1; k7=1;

% Loop to parse numeric water chemistries into vectors
% for each location

for i=1:n;
    if g(i)==0
        if location(i)==1
            DBx(k1,1)=X(i,3);
            k1=k1+1;
        end
        if location(i) == 2
            HLx(k2,1)=X(i,3);
            k2=k2+1;
        end
        if location(i) == 3
            LLx(k3,1)=X(i,3);
            k3=k3+1;
        end
        if location(i)==4
            LPx(k4,1)=X(i,3);
            k4=k4+1;
        end
        if location(i) == 5
            MRx(k5,1)=X(i,3);
            k5=k5+1;
        end
        if location(i) == 6
            SCx(k6,1)=X(i,3);
            k6=k6+1;
        end
        if location(i) == 7
            YLx(k7,1)=X(i,3);
            k7=k7+1;
        end
    end
end

end

% Sizes of the parsed vectors for water chemistry by location %
nDBx=9; nHLx=4; nLLx=4; nLPx=30; nMRx=12; nSCx=9; nYLx=8;

% Loop to create the M.River tribs jackknife mean vectors %

for i=1:nDBx;
    DBxtmp=DBx;
    DBxtmp([i],:)=[]; % Remove row i from DBx and store
    DBxmn(i,1)=mean(DBxtmp); % Delete 1 means in vector DBxmn
end;

% Repeat loop for all Missouri River locations

for i=1:nLPx;

```

```

LPxtmp=LPx;
LPxtmp([i],:)=[];
LPxmn(i,1)=mean(LPxtmp);
end;

for i=1:nMRx;
MRxtmp=MRx;
MRxtmp([i],:)=[];
MRxmn(i,1)=mean(MRxtmp);
end;

for i=1:nSCx;
SCxtmp=SCx;
SCxtmp([i],:)=[];
SCxmn(i,1)=mean(SCxtmp);
end;

% Use same loop to create jackknife mean vectors for Yellowstone
% Lakes

for i=1:nHLx;
HLxtmp=HLx;
HLxtmp([i],:)=[];
HLxmn(i,1)=mean(HLxtmp);
end;

for i=1:nLLx;
LLxtmp=LLx;
LLxtmp([i],:)=[];
LLxmn(i,1)=mean(LLxtmp);
end;

for i=1:nYLx;
YLxtmp=YLx;
YLxtmp([i],:)=[];
YLxmn(i,1)=mean(YLxtmp);
end;

jackmatm=zeros(4,30); % Create matrix of jackknife mean vectors
% 4 rows for locations and 30 cols b/c
% LP has 30 x 1 vector of delete-1
% jackknife means
jackmatm(1,1:nDBx)=DBxmn'; % Row 1 cols 1-9 DB jackknife means
jackmatm(2,1:nLPx)=LPxmn'; % Row 2 cols 1-30 LP jackknife means
jackmatm(3,1:nMRx)=MRxmn'; % Row 3 cols 1-12 MR jackknife means
jackmatm(4,1:nSCx)=SCxmn'; % Row 4 cols 1-9 SC jackknife means

nDBy=sum(location==1); % Initialize space with size equal to number
nLPy=sum(location==4); % of otolith measurements
nMRY=sum(location==5);
nSCy=sum(location==6);

tot=nDBx*nLPx*nMRx*nSCx; % Initialize spaces for number of rows and
TOT=nDBy+nLPy+nMRY+nSCy; % number of columns in ratmat

ratmatm=zeros(TOT,tot); % Matrix of all combos of jackknife means

it=0;

for i= 1:nDBx;
for j= 1:nLPx;
for k= 1:nMRx;
for m= 1:nSCx;

it=it+1;

```

```

ratmatm(1:nDBy, it)=jackmatm(1,i);
ratmatm((nDBy+1):(nDBy+nLPy), it)=jackmatm(2,j);
ratmatm((nLPy+1+nDBy):(nDBy+nLPy+nMRy), it)=jackmatm(3,k);
ratmatm((nLPy+1+nDBy+nMRy):(nDBy+nLPy+nMRy+nSCy), it)=jackmatm(4,m);

    end
end
end
end

```

```

Ym=zeros(223,1); % Create vector of otolith chemistries
Ym=Ym+(location==1); % to be paired with each column of ratmatm
Ym=Ym+(location==4);
Ym=Ym+(location==5);
Ym=Ym+(location==6);
Ym=Ym.*otolith;
Ym(find(Ym==0))=[];

```

```

jackmaty=zeros(3,8); % Create matrix of jackknife mean vectors
% 3 rows for locations and 8 cols b/c
% YL has 8 x 1 vector of delete-1
% jackknife means

```

```

jackmaty(1,1:nHLx)=HLxmn'; % Row 1 cols 1-4 HL jackknife means
jackmaty(2,1:nLLx)=LLxmn'; % Row 2 cols 1-4 LL jackknife means
jackmaty(3,1:nYLx)=YLxmn'; % Row 3 cols 1-8 YL jackknife means

```

```

nHLy=sum(location==2); % Initialize space with size equal to number
nLLy=sum(location==3); % of otolith measurements
nYLy=sum(location==7);

```

```

tot=nHLx*nLLx*nYLx; % Initialize spaces for number of rows and
TOT=nHLy+nLLy+nYLy; % number of columns in ratmat

```

```

ratmaty=zeros(TOT,tot); % Matrix of all combos of jackknife means

```

```

it=0;

```

```

for i= 1:nHLx;
    for j= 1:nLLx;
        for k= 1:nYLx;

```

```

it=it+1;

```

```

ratmaty(1:nHLy, it)=jackmaty(1,i);
ratmaty((nHLy+1):(nHLy+nLLy), it)=jackmaty(2,j);
ratmaty((nHLy+1+nLLy):(nHLy+nLLy+nYLy), it)=jackmaty(3,k);

```

```

    end
end
end

```

```

YY=zeros(223,1); % Create vector of otolith chemistries
YY=YY+(location==2); % to be paired with each column of ratmaty
YY=YY+(location==3);
YY=YY+(location==7);
YY=YY.*otolith;
YY(find(YY==0))=[];

```

```

% Run regression of each column of ratmatm on Y
% This is a regression of otolith on every delete-1 jackknife
% mean for water by location

```

```

X1=ones(123,1);

for i= 1:29160;

X=[X1, ratmatm(:,i)];
Betatm(:,i)=pinv(X'*X)*X'*Ym;

% Calculate vector of coefficients of determination
b1=Betatm(2,i);
Sx=sqrt(var(X(:,2)));
Sy=sqrt(var(Ym));
Rsqm(i)=(b1*(Sx/Sy))^2);

end;

% Run regression of each column of ratmaty on YY
% This is a regression of otolith on every delete-1 jackknife
% mean for water by location

X1=ones(100,1);

for i= 1:128;

X=[X1, ratmaty(:,i)];
Betaty(:,i)=pinv(X'*X)*X'*YY;

% Calculate vector of coefficients of determination
b1=Betaty(2,i);
Sx=sqrt(var(X(:,2)));
Sy=sqrt(var(YY));
Rsqu(i)=(b1*(Sx/Sy))^2);

end;

% Create x,y pairs of jackknife mean water chem vs. oto chem by site for plotting

DBp=ones(351,2); % 39 responses times 9 jackknife means
Ymd=Ym(1:nDBY);
it=0;
for i=1:9
    for j=1:39
        it=it+1;
        DBp(it,:)=[DBxmn(i), Ymd(j)];
    end
end

LPP=ones(930,2); % 31 responses times 30 jackknife means
Yml=Ym(nDBY+1:nDBY+nLPY);
it=0;
for i=1:30
    for j=1:31
        it=it+1;
        LPP(it,:)=[LPxmn(i), Yml(j)];
    end
end

MRp=ones(156,2); % 13 responses times 12 jackknife means
Ymm=Ym(nDBY+nLPY+1:nDBY+nLPY+nMRY);
it=0;
for i=1:12
    for j=1:13
        it=it+1;
        MRp(it,:)=[MRxmn(i), Ymm(j)];
    end
end

```

```

SCp=ones(360,2); % 40 responses times 9 jackknife means
Yms=Ym(nDBy+nLPy+nMRY+1:nDBy+nLPy+nMRY+nSCy);
it=0;
for i=1:9
    for j=1:40
        it=it+1;
        SCp(it,:)=[SCxmn(i), Yms(j)];
    end
end

```

```

HLp=ones(116,2); % 29 responses times 4 jackknife means
YYh=YY(1:nHLY);
it=0;
for i=1:4
    for j=1:29
        it=it+1;
        HLp(it,:)=[HLxmn(i), YYh(j)];
    end
end

```

```

LLp=ones(156,2); % 29 responses times 4 jackknife means
YYl=YY(nHLY+1:nHLY+nLLy);
it=0;
for i=1:4
    for j=1:29
        it=it+1;
        LLp(it,:)=[LLxmn(i), YYl(j)];
    end
end

```

```

YLp=ones(336,2); % 42 responses times 8 jackknife means
YYy=YY(nHLY+nLLy+1:nHLY+nLLy+nYLy);
it=0;
for i=1:8
    for j=1:42
        it=it+1;
        YLp(it,:)=[YLxmn(i), YYy(j)];
    end
end

```

% Plots for Missouri tribes and Yellowstone lakes

```

figure(1)
hold on;
plot(DBp(:,1),DBp(:,2),'o', LPp(:,1),LPp(:,2),'x', MRp(:,1),MRp(:,2),'^', ...
      SCp(:,1),SCp(:,2),'s')
title('Figure 1A. Delete-1 Jackknife Means for Water Chemistries vs. Otolith Chemistry')
xlabel('Delete-1 Jackknife Means for Water Sr:Ca')
ylabel('Otolith Sr:Ca')
legend('Dearborn', 'Prickly Pear', 'Missouri', 'Sheep', 2)
text(1.65,0.75, 'Dearborn')
text(2,0.64, 'Prickly')
text(2.75,0.81, 'Sheep')
text(2.92,1.35, 'Missouri')

```

```

figure(2)
hist(Betatm(1,:))
title('Figure 1C. Intercept Estimate')
xlabel('\beta_0 hat')
ylabel('Frequency')

```

```

figure(3)
hist(Betatm(2,:))
title('Figure 1D. Slope Estimate')

```



```
xlabel('\beta_1 hat')
ylabel('Frequency')
```

```
figure(4)
plot(Betatm(1,:), Betatm(2,:), '.')
title('Figure 1E. Intercept vs. Slope')
xlabel('\beta_0 hat')
ylabel('\beta_1 hat')
```

```
figure(5)
hold on;
x=[0 1];
for i=1:29160
    plot(x, [Betatm(1,i) Betatm(1,i) + Betatm(2,i)])
end
title('Figure 1B. Least Squares Lines')
xlabel('Delete-1 Mean Water Sr:Ca')
ylabel('Otolith Sr:Ca')
```

```
figure(6)
hold on;
plot(HLp(:,1),HLp(:,2),'o', LLp(:,1),LLp(:,2),'x', YLp(:,1),YLp(:,2),'^')
title('Figure 2A. Delete-1 Jackknife Means for Water Chemistries vs. Otolith Chemistry')
xlabel('Delete-1 Jackknife Means for Water Sr:Ca')
ylabel('Otolith Sr:Ca')
legend('Heart', 'Lewis', 'Yellowstone', 2)
text(1.35,1.25, 'Lewis')
text(2.35,0.85, 'Heart')
text(3.45,1.2, 'Yellowstone')
```

```
figure(7)
hist(Betaty(1,:))
title('Figure 2C. Intercept Estimate')
xlabel('\beta_0 hat')
ylabel('Frequency')
```

```
figure(8)
hist(Betaty(2,:))
title('Figure 2D. Slope Estimate')
xlabel('\beta_1 hat')
ylabel('Frequency')
```

```
figure(9)
plot(Betaty(1,:), Betaty(2,:), '.')
title('Figure 2E. Intercept vs. Slope')
xlabel('\beta_0 hat')
ylabel('\beta_1 hat')
```

```
figure(10)
hold on;
x=[0 1];
for i=1:128
    plot(x, [Betaty(1,i) Betaty(1,i) + Betaty(2,i)])
end
title('Figure 2B. Least Squares Lines')
xlabel('Delete-1 Mean Water Sr:Ca')
ylabel('Otolith Sr:Ca')
```

Figure 1A. Delete-1 Jackknife Means for Water Chemistries vs. Otolith Chemistry

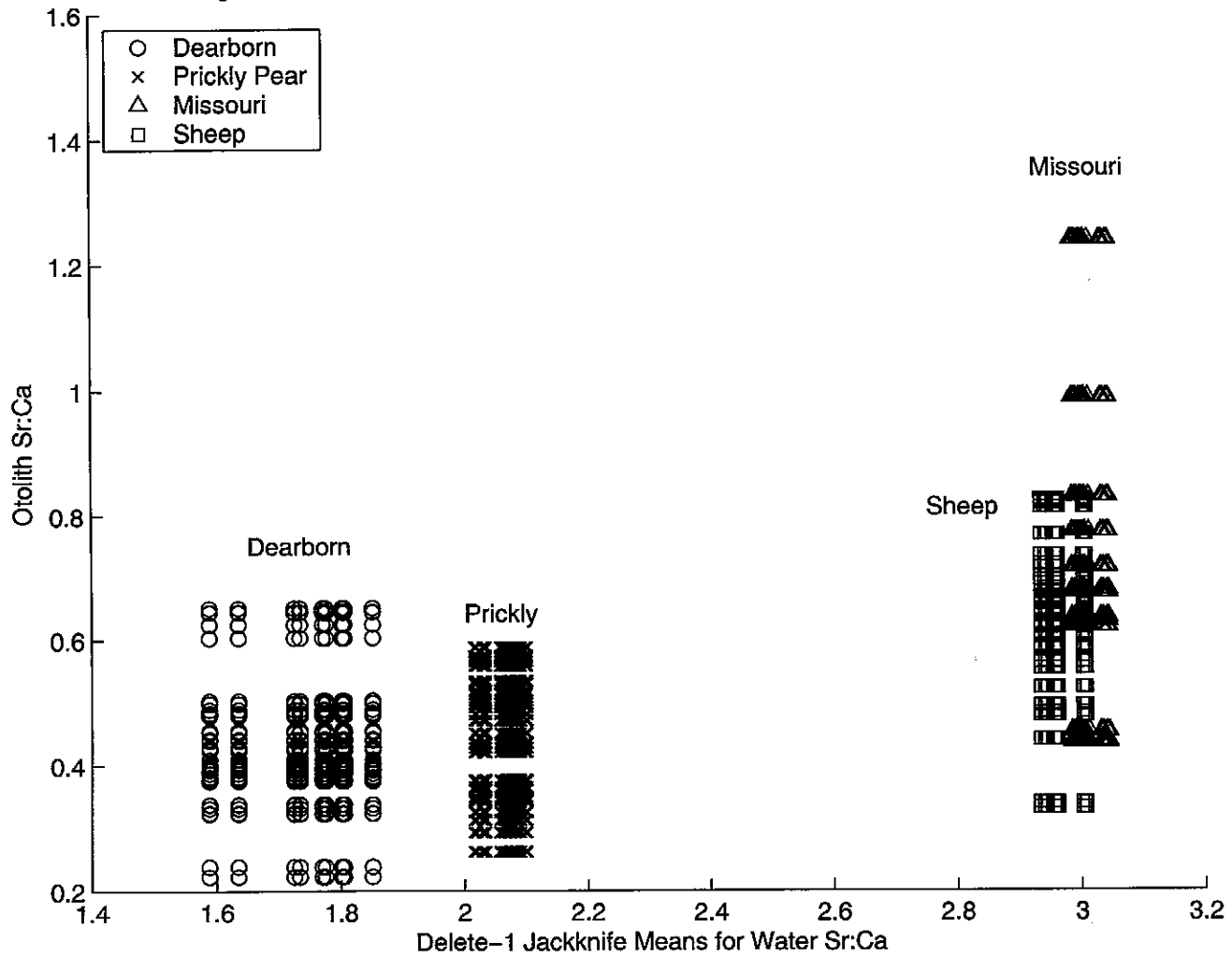


Figure 1B. Least Squares Lines

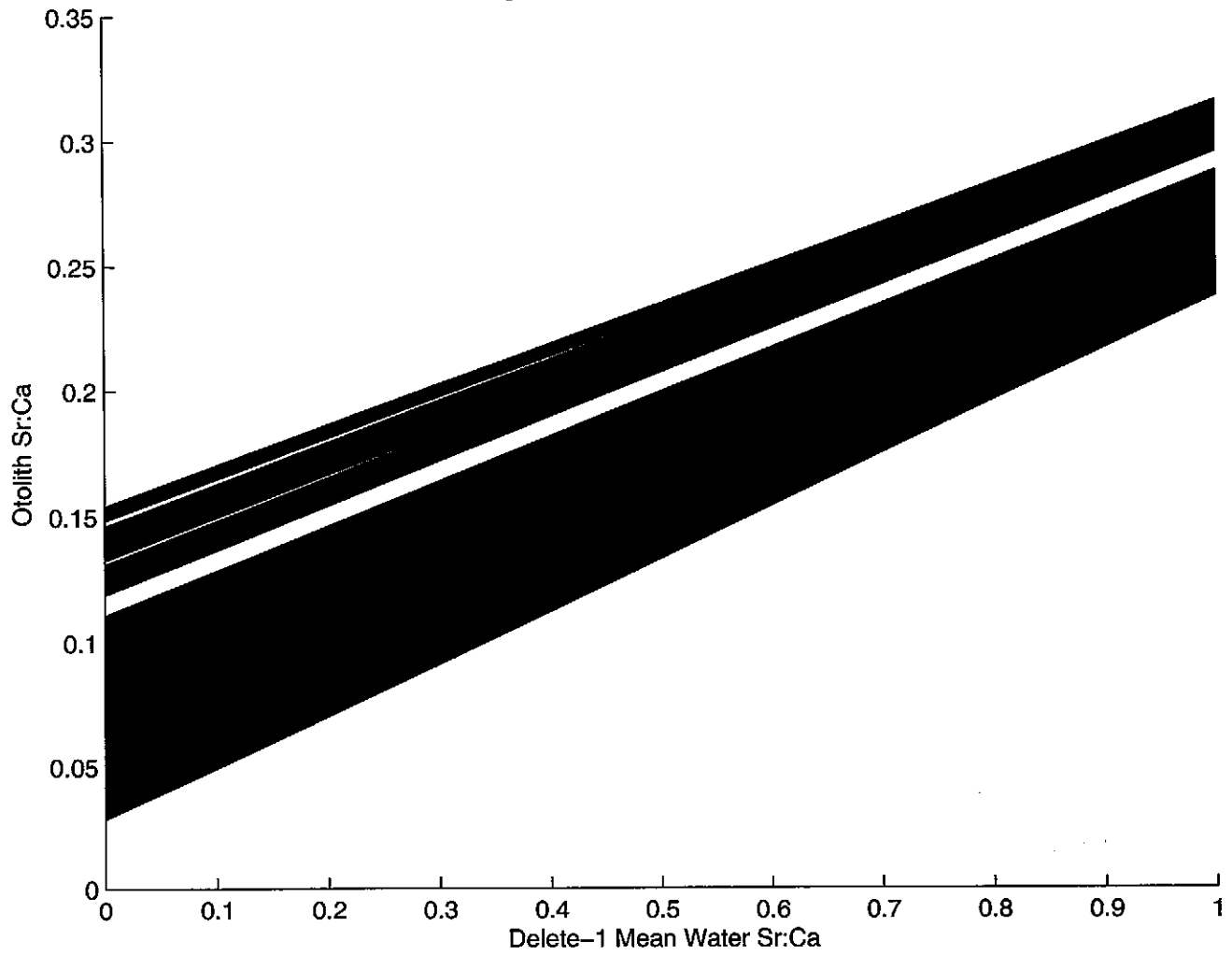


Figure 1C. Intercept Estimate

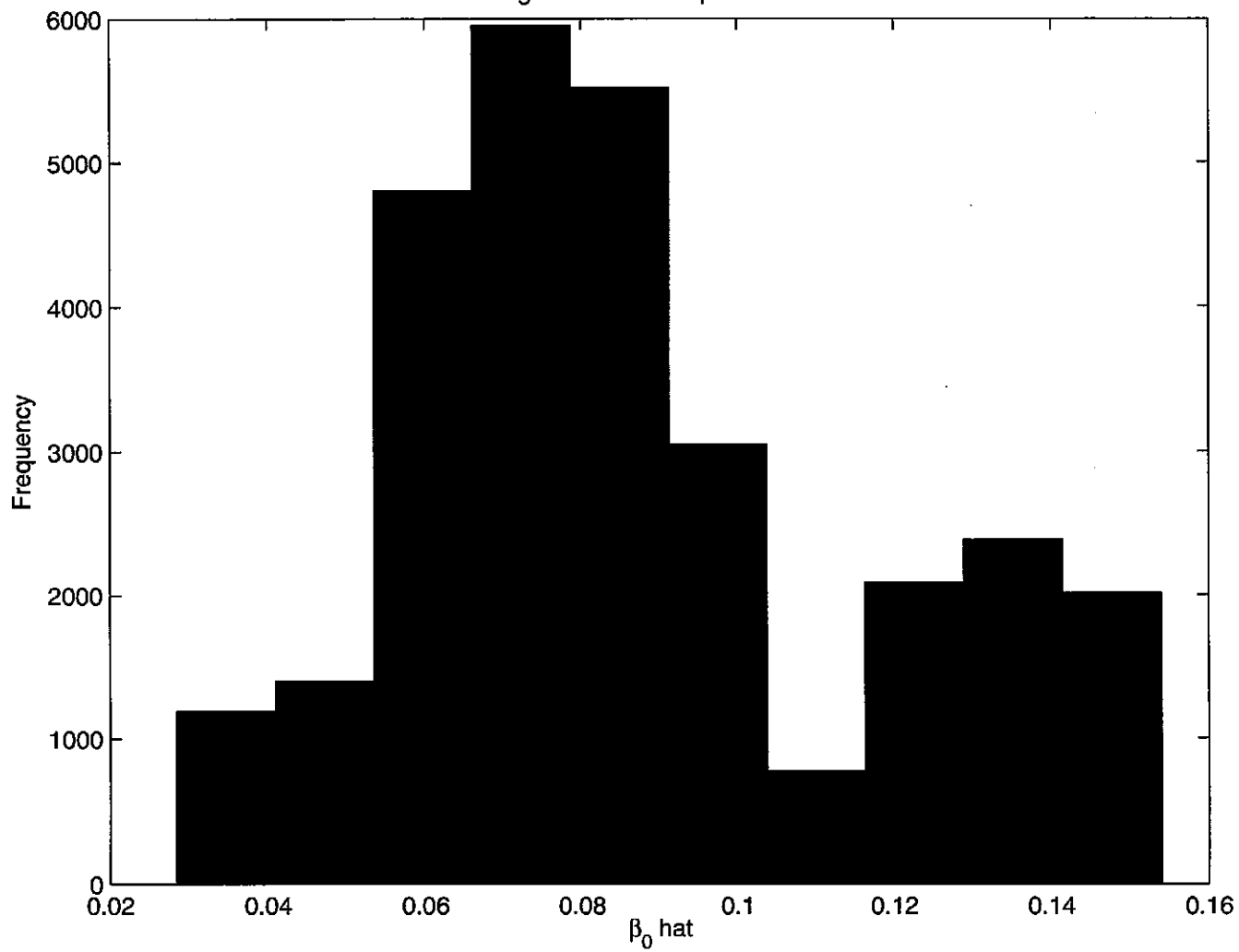


Figure 1D. Slope Estimate

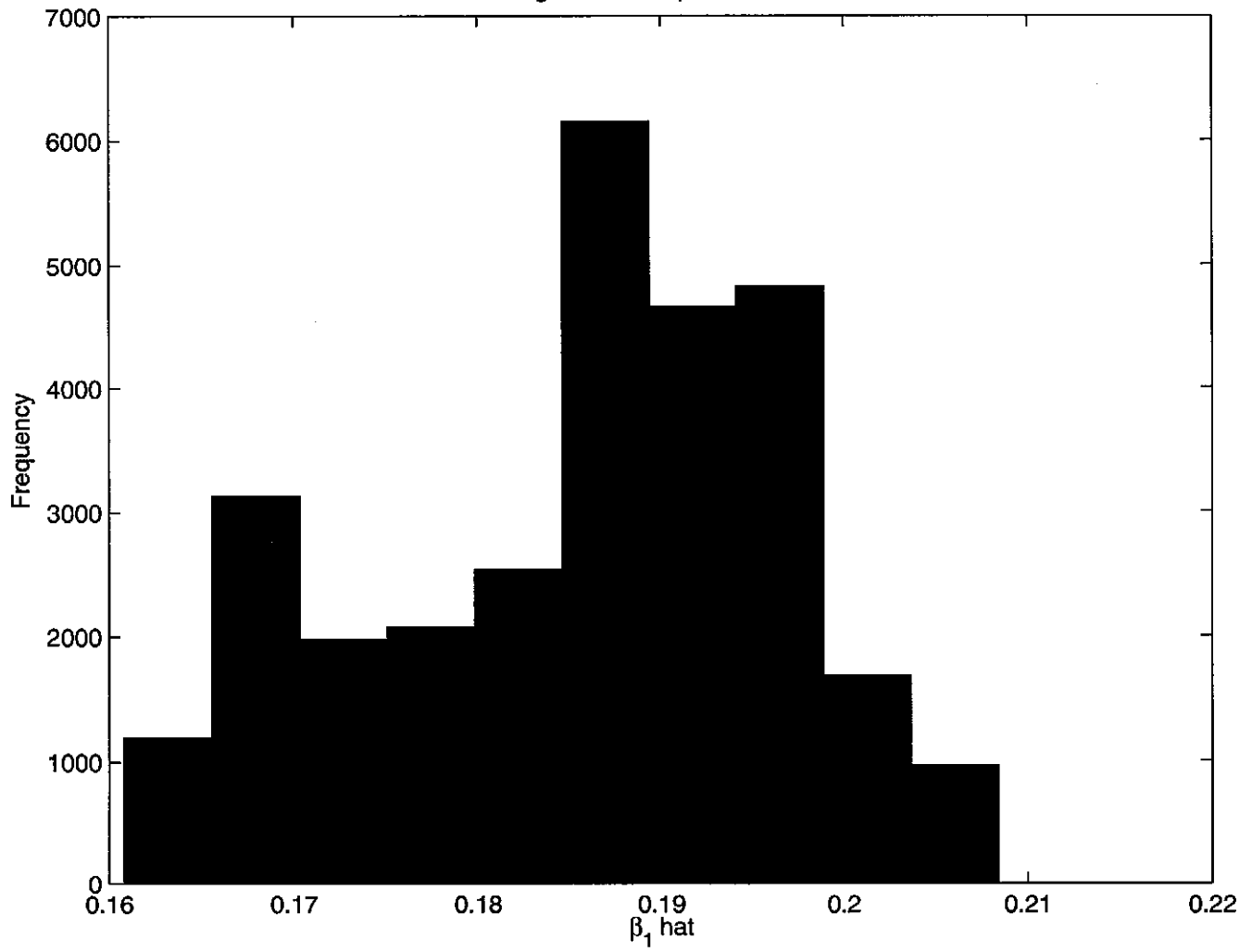


Figure 1E. Intercept vs. Slope

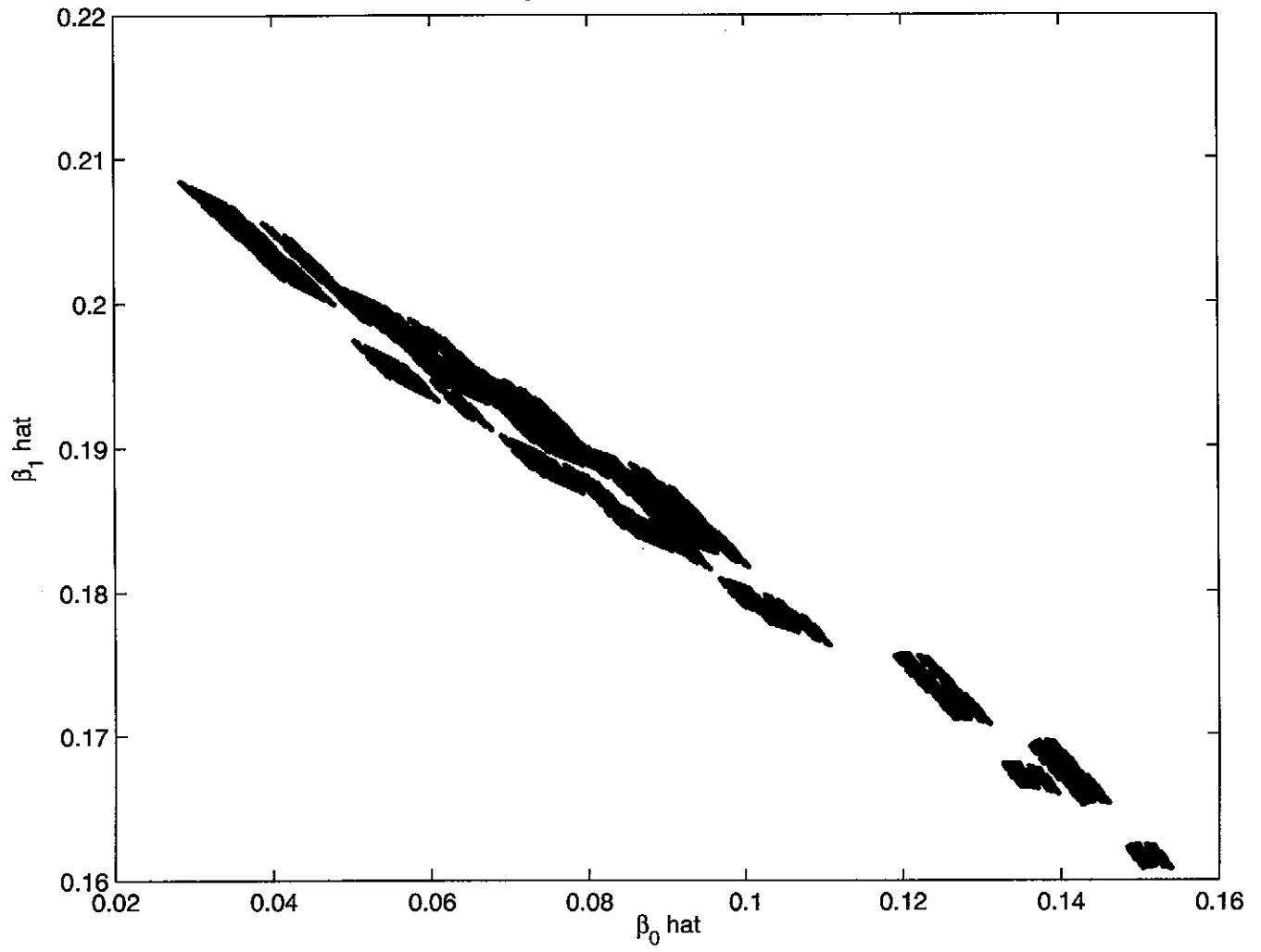


Figure 2A. Delete-1 Jackknife Means for Water Chemistries vs. Otolith Chemistry

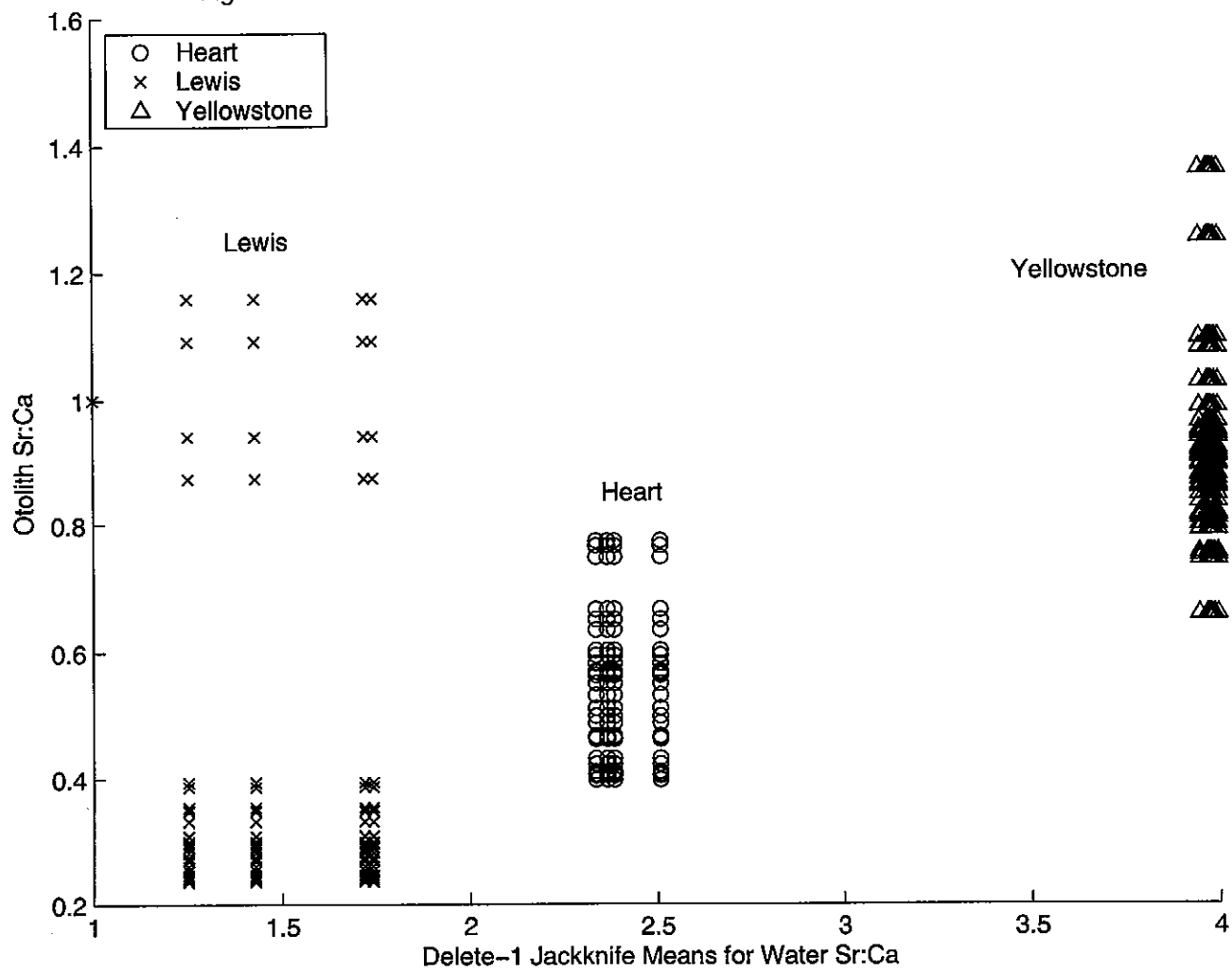


Figure 2B. Least Squares Lines

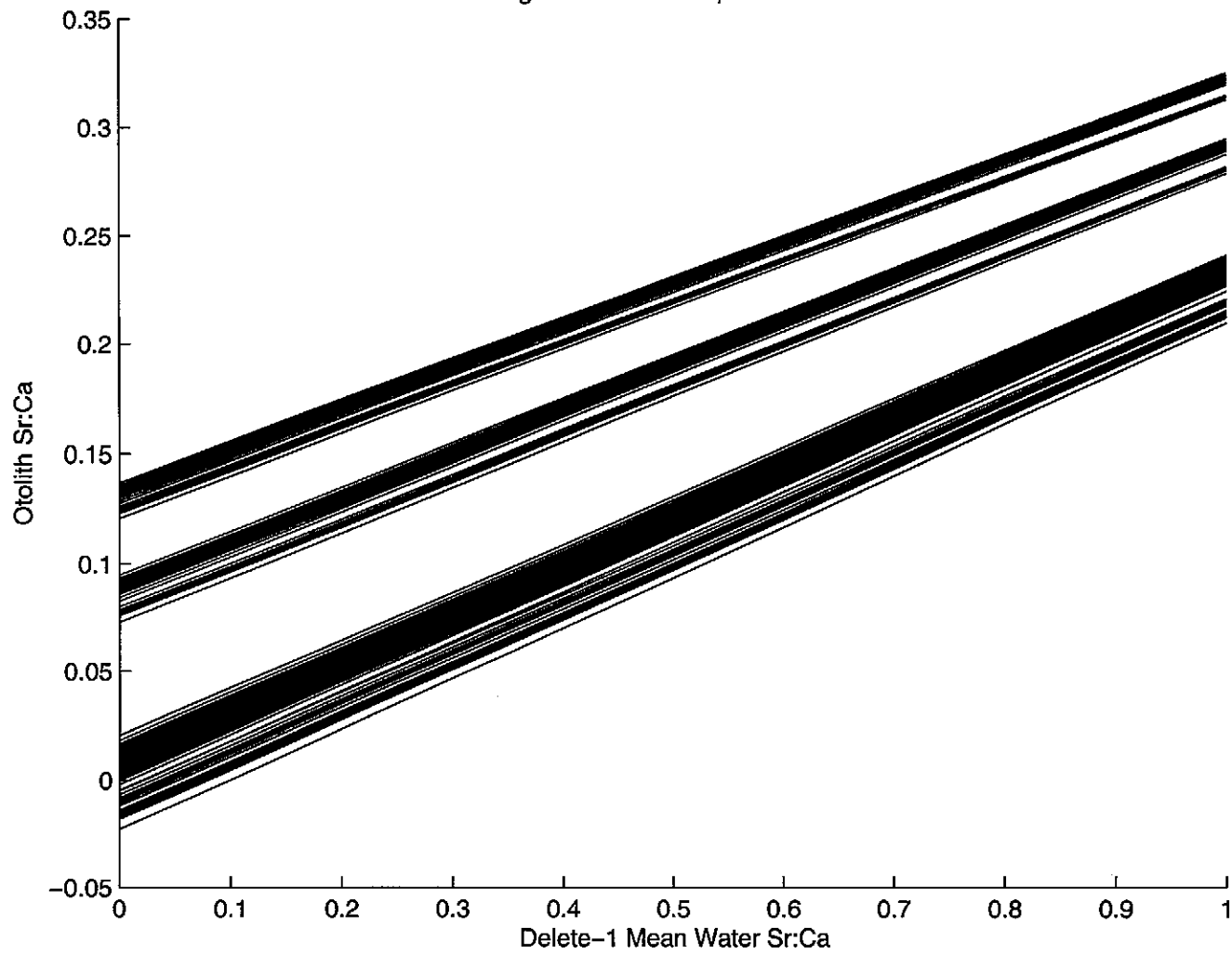




Figure 2C. Intercept Estimate

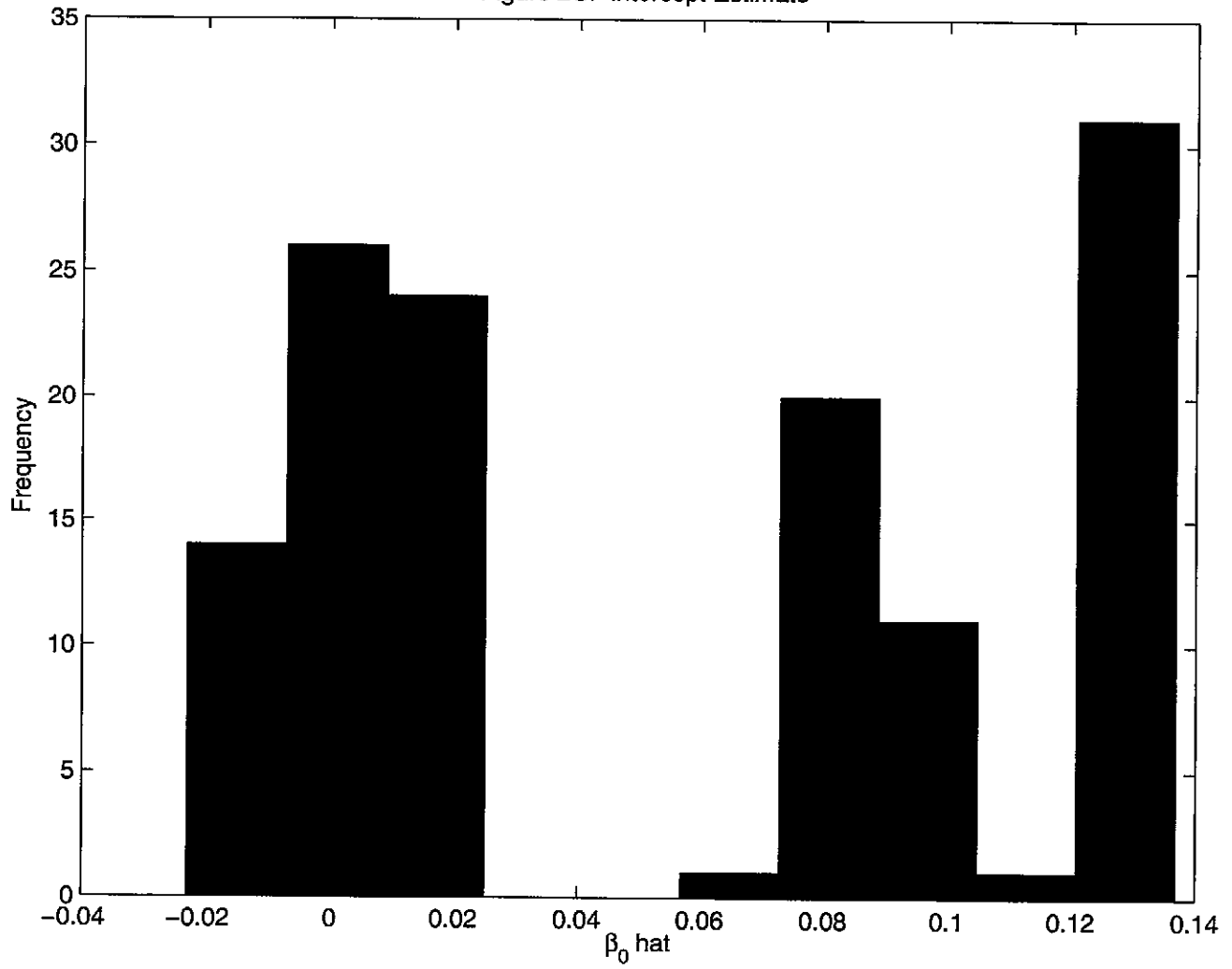


Figure 2D. Slope Estimate

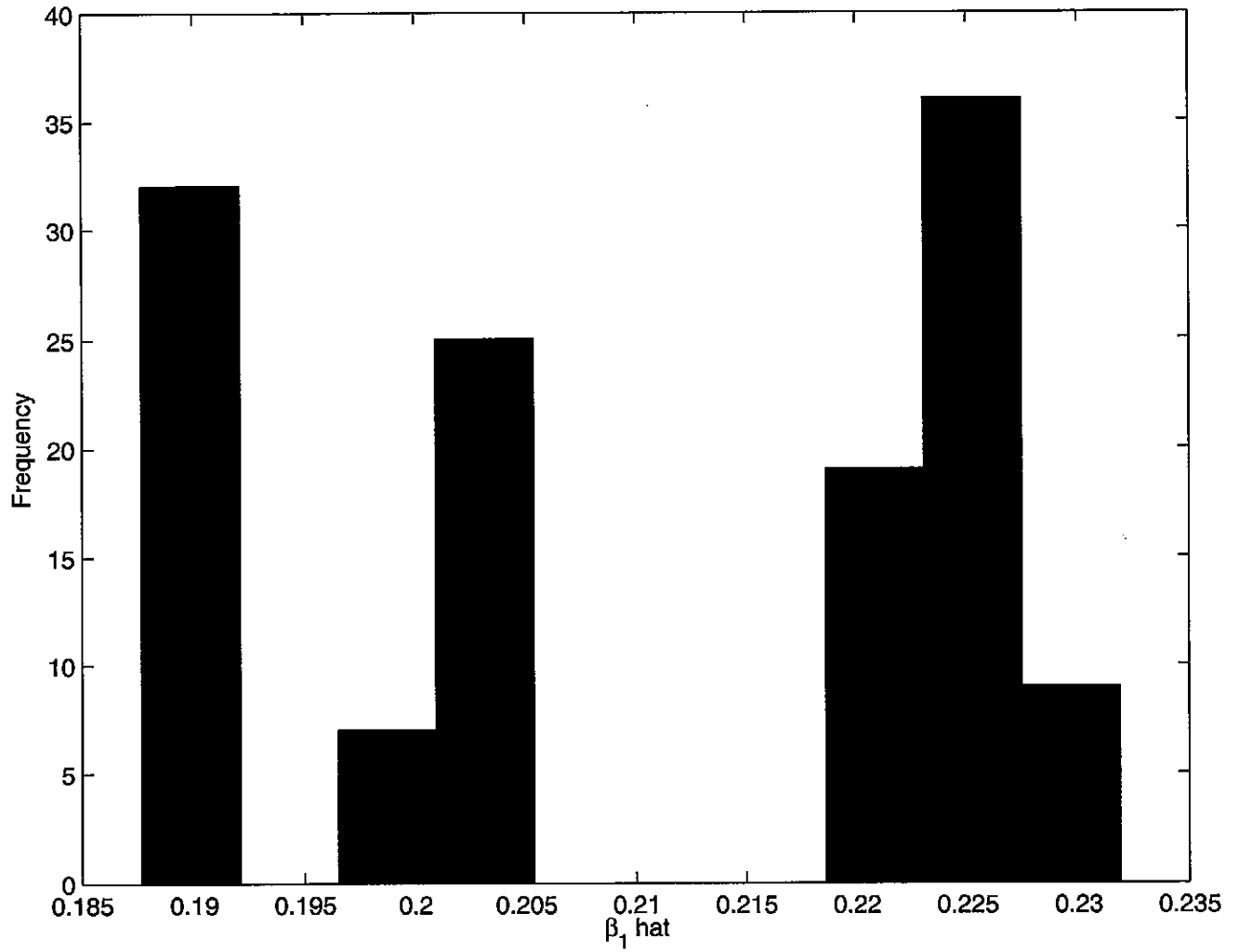
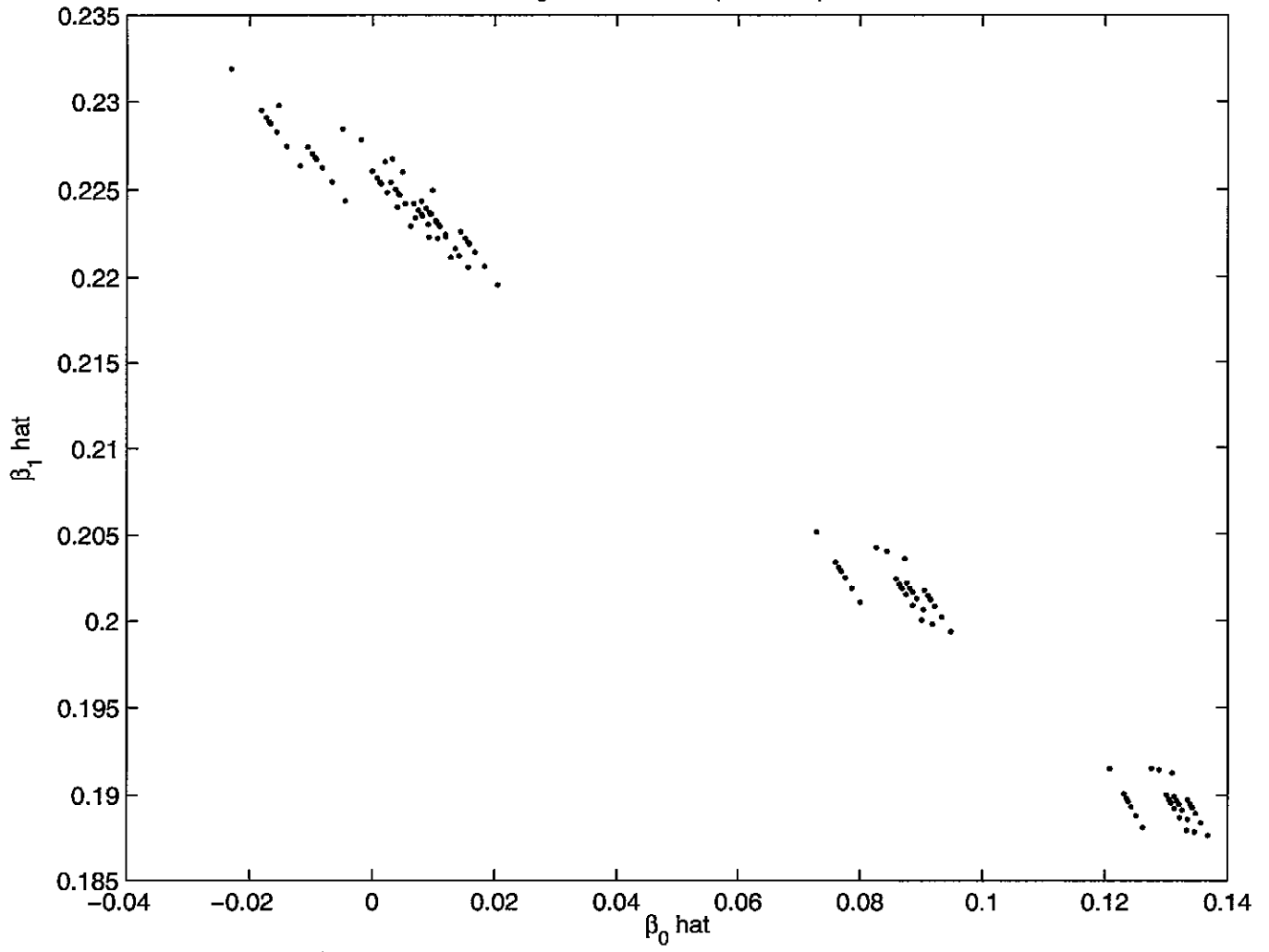


Figure 2E. Intercept vs. Slope



## Appendix II.

Matlab source code and outputted figures when regressing delete-1 jackknife mean otolith Sr:Ca ratios versus delete-1 jackknife mean water Sr:Ca ratios.

```

load chemistry.dat
X=[chemistry(:,1), chemistry(:,2), chemistry(:,3)];
location=X(:,1);
otolith=X(:,2);
water=X(:,3);
length(X);
n=223;

% Designate vectors for jackknifing
DBx = []; HLx=[]; LLx=[]; LPx=[]; MRx=[]; SCx=[]; YLx=[];

g=isnan(water);% Makes index of 0s & 1s where NaN=1

% Create counter indices for each jackknife vector space
k1=1; k2=1; k3=1; k4=1; k5=1; k6=1; k7=1;

% Loop to parse numeric water chemistries into vectors
% for each location

for i=1:n;
  if g(i)==0
    if location(i)==1
      DBx(k1,1)=X(i,3);
      k1=k1+1;
    end
    if location(i) == 2
      HLx(k2,1)=X(i,3);
      k2=k2+1;
    end
    if location(i) == 3
      LLx(k3,1)=X(i,3);
      k3=k3+1;
    end
    if location(i)==4
      LPx(k4,1)=X(i,3);
      k4=k4+1;
    end
    if location(i) == 5
      MRx(k5,1)=X(i,3);
      k5=k5+1;
    end
    if location(i) == 6
      SCx(k6,1)=X(i,3);
      k6=k6+1;
    end
    if location(i) == 7
      YLx(k7,1)=X(i,3);
      k7=k7+1;
    end
  end
end

end

% Sizes of the parsed vectors for water chemistry by location %
nDBx=9; nHLx=4; nLLx=4; nLPx=30; nMRx=12; nSCx=9; nYLx=8;

% Loop to create the M.River tribs jackknife mean vectors %

for i=1:nDBx;
  DBxtmp=DBx;
  DBxtmp([i],:)=[]; % Remove row i from DBx and store
  DBxmn(i,1)=mean(DBxtmp); % Delete 1 means in vector DBxmn
end;

% Repeat loop for all Missouri River locations

for i=1:nLPx;

```

```

LPxtmp=LPx;
LPxtmp([i],:)=[];
LPxmn(i,1)=mean(LPxtmp);
end;

for i=1:nMRx;
MRxtmp=MRx;
MRxtmp([i],:)=[];
MRxmn(i,1)=mean(MRxtmp);
end;

for i=1:nSCx;
SCxtmp=SCx;
SCxtmp([i],:)=[];
SCxmn(i,1)=mean(SCxtmp);
end;

% Use same loop to create jackknife mean vectors for Yellowstone
% Lakes

for i=1:nHLx;
HLxtmp=HLx;
HLxtmp([i],:)=[];
HLxmn(i,1)=mean(HLxtmp);
end;

for i=1:nLLx;
LLxtmp=LLx;
LLxtmp([i],:)=[];
LLxmn(i,1)=mean(LLxtmp);
end;

for i=1:nYLx;
YLxtmp=YLx;
YLxtmp([i],:)=[];
YLxmn(i,1)=mean(YLxtmp);
end;

% Create vector of otolith chemistries
% for each Missouri R. trib

DBy=zeros(223,1);
DBy=DBy+(location==1);
DBy=DBy .* otolith;
DBy(find(DBy==0))=[];

LPy=zeros(223,1);
LPy=LPy+(location==4);
LPy=LPy .* otolith;
LPy(find(LPy==0))=[];

MRy=zeros(223,1);
MRy=MRy+(location==5);
MRy=MRy .* otolith;
MRy(find(MRy==0))=[];

SCy=zeros(223,1);
SCy=SCy+(location==6);
SCy=SCy .* otolith;
SCy(find(SCy==0))=[];

% Create vector of otolith chemistries
% for each Yellowstone Lake

HLy=zeros(223,1);
HLy=HLy+(location==2);
HLy=HLy .* otolith;

```

```
HLy(find(HLy==0))=[];
```

```
LLy=zeros(223,1);
LLy=LLy+(location==3);
LLy=LLy .* otolith;
LLy(find(LLy==0))=[];
```

```
YLy=zeros(223,1);
YLy=YLy+(location==7);
YLy=YLy .* otolith;
YLy(find(YLy==0))=[];
```

```
% Sizes of the parsed vectors for otolith chemistry by location %
nDBy=39; nHLy=29; nLLy=29; nLPy=31; nMRy=13; nSCy=40; nYLy=42;
```

```
% Loop to create delete-1 jackknife mean otolith vectors
% for Missouri R. tribs
```

```
for i=1:nDBy;
    DBytmp=DBy;
    DBytmp([i],:)=[]; % Remove row i from DBy and store
    DBymn(i,1)=mean(DBytmp); % Delete 1 means in vector DBymn
end;
```

```
% Repeat loop for all Missouri River locations
```

```
for i=1:nLPy;
    LPytmp=LPy;
    LPytmp([i],:)=[];
    LPymn(i,1)=mean(LPytmp);
end;
```

```
for i=1:nMRy;
    MRytmp=MRy;
    MRytmp([i],:)=[];
    MRymn(i,1)=mean(MRytmp);
end;
```

```
for i=1:nSCy;
    SCytmp=SCy;
    SCytmp([i],:)=[];
    SCymn(i,1)=mean(SCytmp);
end;
```

```
% Loop to create delete-1 jackknife mean otolith vectors
% for Yellowstone Lakes
```

```
for i=1:nHLy;
    HLytmp=HLy;
    HLytmp([i],:)=[];
    HLyymn(i,1)=mean(HLytmp);
end;
```

```
for i=1:nLLy;
    LLytmp=LLy;
    LLytmp([i],:)=[];
    LLyymn(i,1)=mean(LLytmp);
end;
```

```
for i=1:nYLy;
    YLytmp=YLy;
    YLytmp([i],:)=[];
    YLyymn(i,1)=mean(YLytmp);
end;
```

```

Betatm=zeros(2,10000); % Initialize space for regression parameters
X1=ones(4,1); % Initialize first column of X matrix
s11=zeros(10000,1); s12=s11; s13=s11; s14=s11; % storage vectors for
sw1=s11; sw2=s11; sw3=s11; sw4=s11; % plotting

for i= 1:10000; % Loop to create random pairs of delete-1 jackknife
% means of otolith & water chemistry for regression

loc1= ceil(nDBy * rand); % Create random number within length of nDBy
loc2= ceil(nLPy * rand);
loc3= ceil(nMRy * rand);
loc4= ceil(nSCy * rand);

w1=ceil(nDBx * rand); % Create random number within length nDBx
w2=ceil(nLPx * rand);
w3=ceil(nMRx * rand);
w4=ceil(nSCx * rand);

% Create vector of random delete-1 jackknife mean otoliths
Y=[DBymn(loc1); LPymn(loc2); MRymn(loc3); SCymn(loc4)];

% Create design matrix for regression. Column of ones and column
% of random delete-1 jackknife mean chemistries
X=[X1 [DBxmn(w1); LPxmn(w2); MRxmn(w3); SCxmn(w4)]];

Betatm(:,i)=pinv(X'*X)*X'*Y; % Create 2-row matrix of regression
% coefficients. Beta0 then Beta1

% Calculate vector of coefficients of determination
b1=Betatm(2,i);
Sx=sqrt(var(X(:,2)));
Sy=sqrt(var(Y));
Rsqm(i)=((b1*(Sx/Sy))^2);

% Initialize vectors for all random delete-1
% jackknife mean chemistries and otoliths
s11(i)=Y(1); s12(i)=Y(2); s13(i)=Y(3); s14(i)=Y(4); % Otoliths
sw1(i)=X(1,2); sw2(i)=X(2,2); sw3(i)=X(3,2); sw4(i)=X(4,2); % Chemistries

end;

figure(1)
hist(Betatm(1,:))
title('Figure 3C. Intercept Estimate')
xlabel('\beta_0 hat')
ylabel('Frequency')

figure(2)
hist(Betatm(2,:))
title('Figure 3D. Slope Estimate')
xlabel('\beta_1 hat')
ylabel('Frequency')

figure(3)
plot(Betatm(1,:), Betatm(2,:), '.')
title('Figure 3E. Intercept vs. Slope')
xlabel('\beta_0 hat')
ylabel('\beta_1 hat')

figure(4)
plot(sw1,s11,'o', sw2,s12,'x', sw3,s13,'^', sw4,s14,'s')
title('Figure 3A. Delete-1 Jackknife Means for Otolith & Water Chemistries')
xlabel('Delete-1 Jackknife Means for Water Sr:Ca')
ylabel('Delete-1 Jackknife Means for Otolith Sr:Ca')
legend('Dearborn', 'Prickly Pear', 'Missouri', 'Sheep', 4)

```



```

text(1.65,0.45, 'Dearborn')
text(2,0.47, 'Prickly')
text(2.75,0.63, 'Sheep')
text(2.73, 0.7, 'Missouri')

figure(5)
hold on;
x=[0 1];
for i=1:100
    plot(x,[Betatm(1,i) Betatm(1,i) + Betatm(2,i)])
end
title('Least Squares Lines')
xlabel('Delete-1 Mean Water Sr:Ca')
ylabel('Delete-1 Mean Otolith Sr:Ca')

figure(6)
hold on;
x=[0 1];
for i=1:10000
    plot(x,[Betatm(1,i) Betatm(1,i) + Betatm(2,i)])
end
title('Figure 3B. Least Squares Lines')
xlabel('Delete-1 Mean Water Sr:Ca')
ylabel('Delete-1 Mean Otolith Sr:Ca')

% Repeat same code for Yellowstone Lakes %

Betaty=zeros(2,10000);
X1=ones(3,1);
s11=zeros(10000,1); s12=s11; s13=s11;
sw1=s11; sw2=s11; sw3=s11;

for i= 1:10000;

loc1= ceil(nHLy * rand);
loc2= ceil(nLLy * rand);
loc3= ceil(nYLy * rand);

w1=ceil(nHLx * rand);
w2=ceil(nLLx * rand);
w3=ceil(nYLx * rand);

Y=[HLymn(loc1); LLymn(loc2); YLymn(loc3)];

X=[X1 [HLxmn(w1); LLxmn(w2); YLxmn(w3)]];

Betaty(:,i)=pinv(X'*X)*X'*Y;

% Calculate vector of coefficients of determination
b1=Betaty(2,i);
Sx=sqrt(var(X(:,2)));
Sy=sqrt(var(Y));
Rsqy(i)=((b1*(Sx/Sy))^2);

s11(i)=Y(1); s12(i)=Y(2); s13(i)=Y(3);
sw1(i)=X(1,2); sw2(i)=X(2,2); sw3(i)=X(3,2);

end;

figure(7)
hist(Betaty(1,:))
title('Figure 4C. Intercept Estimate')
xlabel('\beta_0 hat')
ylabel('Frequency')

figure(8)

```

```
hist(Betaty(2,:))
title('Figure 4D. Slope Estimate')
xlabel('\beta_1 hat')
ylabel('Frequency')
```

```
figure(9)
plot(Betaty(1,:), Betaty(2,:), '.')
title('Figure 4E. Intercept vs. Slope')
xlabel('\beta_0 hat')
ylabel('\beta_1 hat')
```

```
figure(10)
plot(sw1,sl1,'o', sw2,sl2,'x', sw3,sl3,'^')
title('Figure 4A. Delete-1 Jackknife Means for Otolith & Water Chemistries')
xlabel('Delete-1 Jackknife Means for Water Sr:Ca')
ylabel('Delete-1 Jackknife Means for Otolith Sr:Ca')
legend('Heart', 'Lewis', 'Yellowstone', 4)
text(1.4,0.45, 'Lewis')
text(2.35,0.6, 'Heart')
text(3.5,0.9, 'Yellowstone')
```

```
figure(11)
hold on;
x=[0 1];
for i=1:10000
plot(x,[Betaty(1,i) Betaty(1,i) + Betaty(2,i)])
end
title('Figure 4B. Least Squares Lines')
xlabel('Delete-1 Mean Water Sr:Ca')
ylabel('Delete-1 Mean Otolith Sr:Ca')
```

Figure 3A. Delete-1 Jackknife Means for Otolith &amp; Water Chemistries

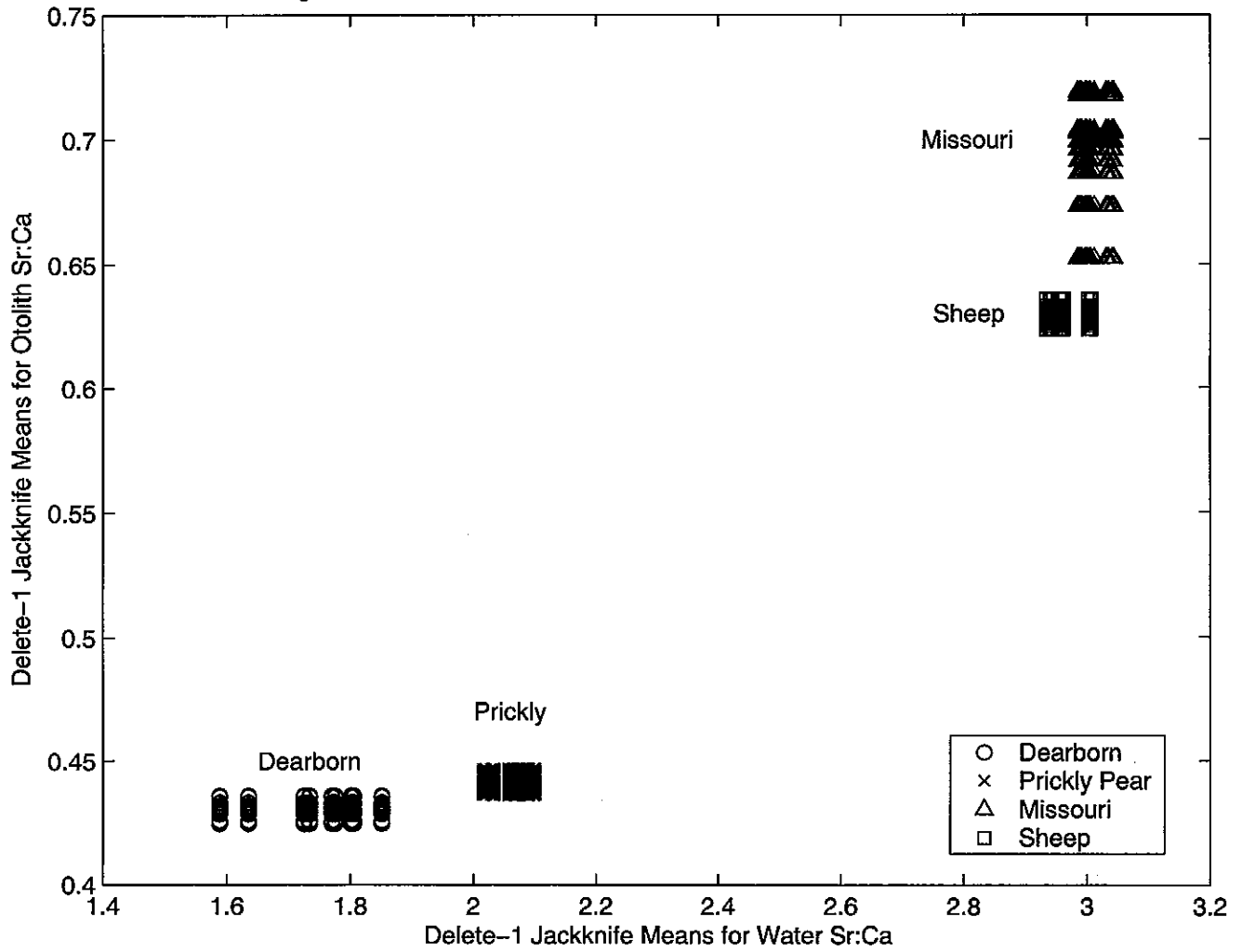


Figure 3B. Least Squares Lines

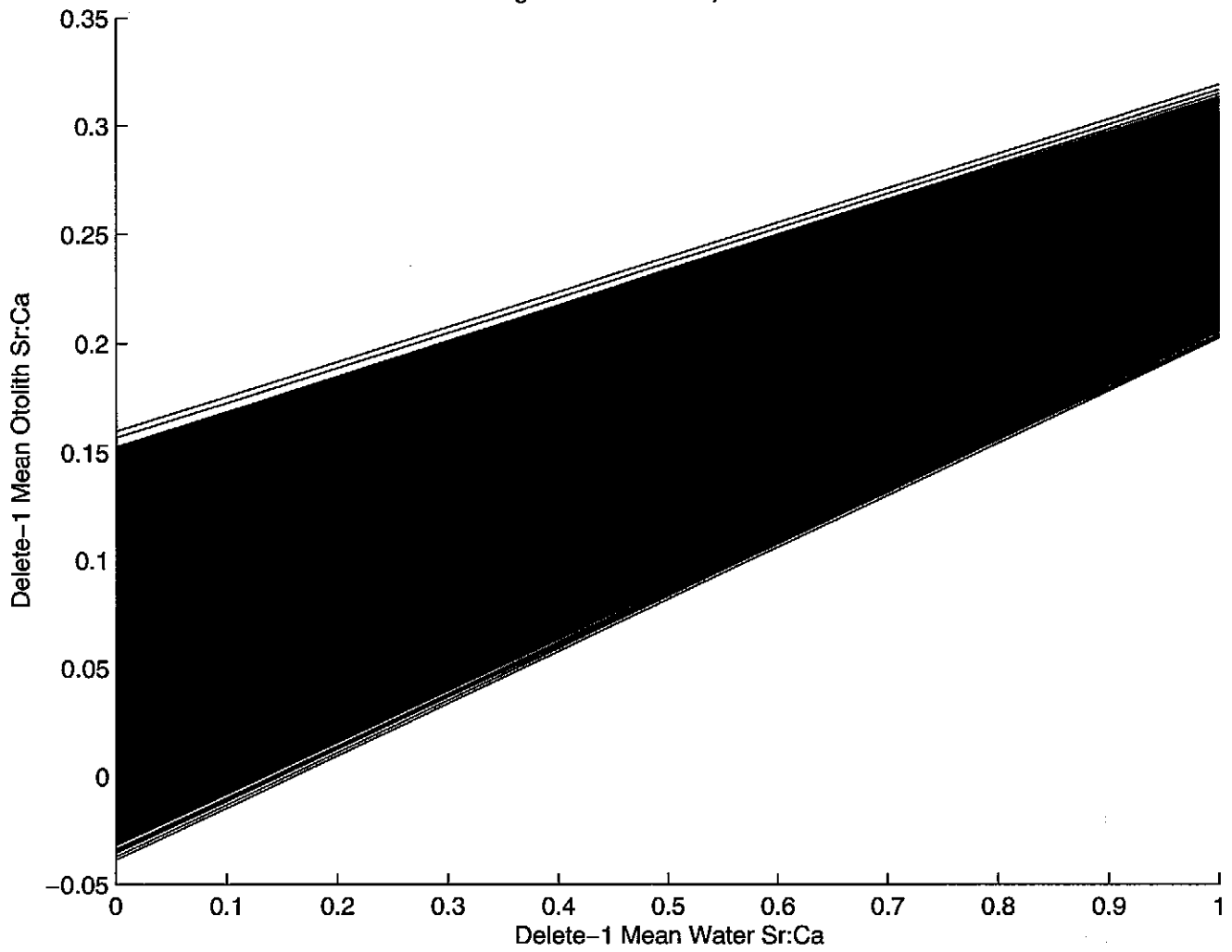


Figure 3C. Intercept Estimate

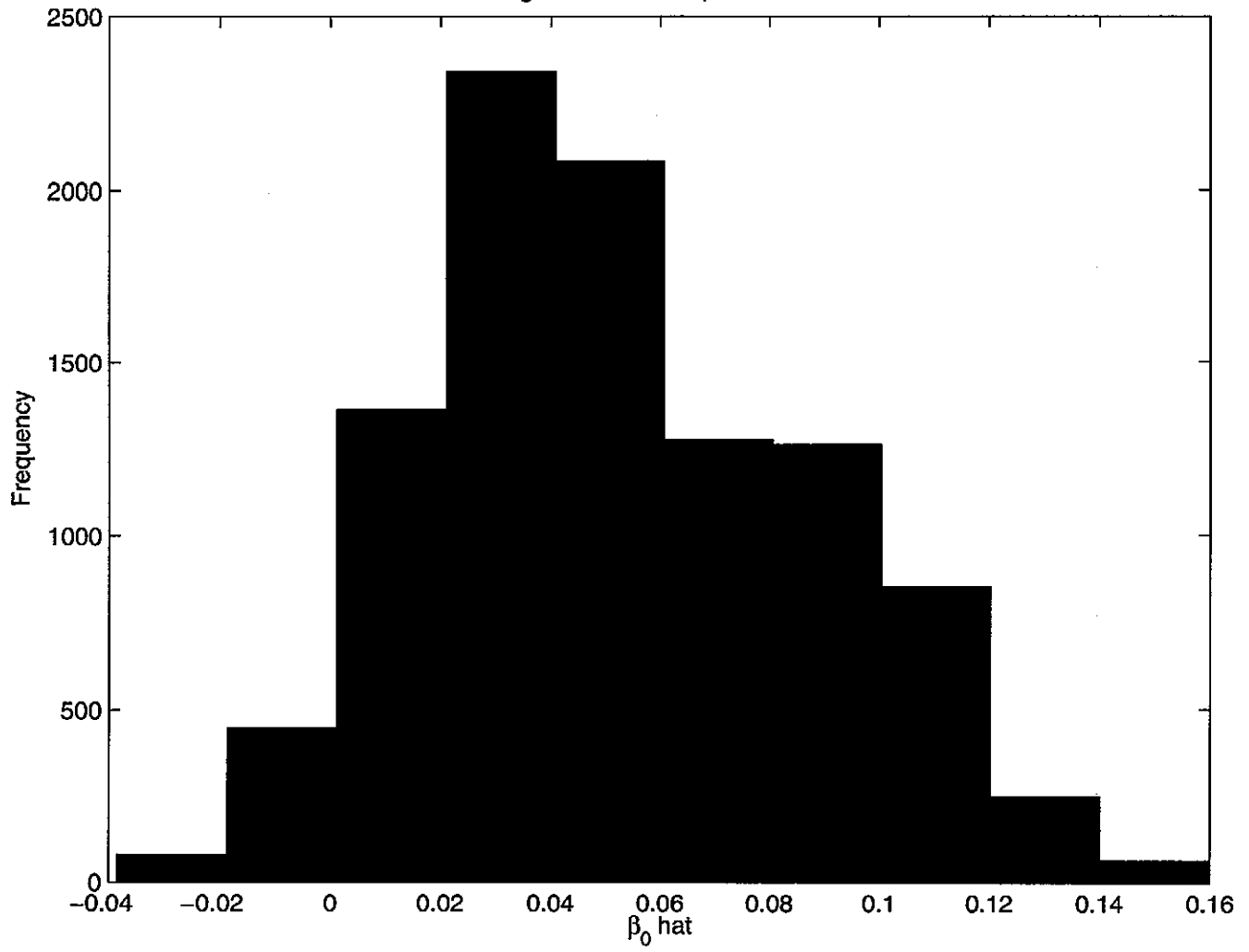


Figure 3D. Slope Estimate

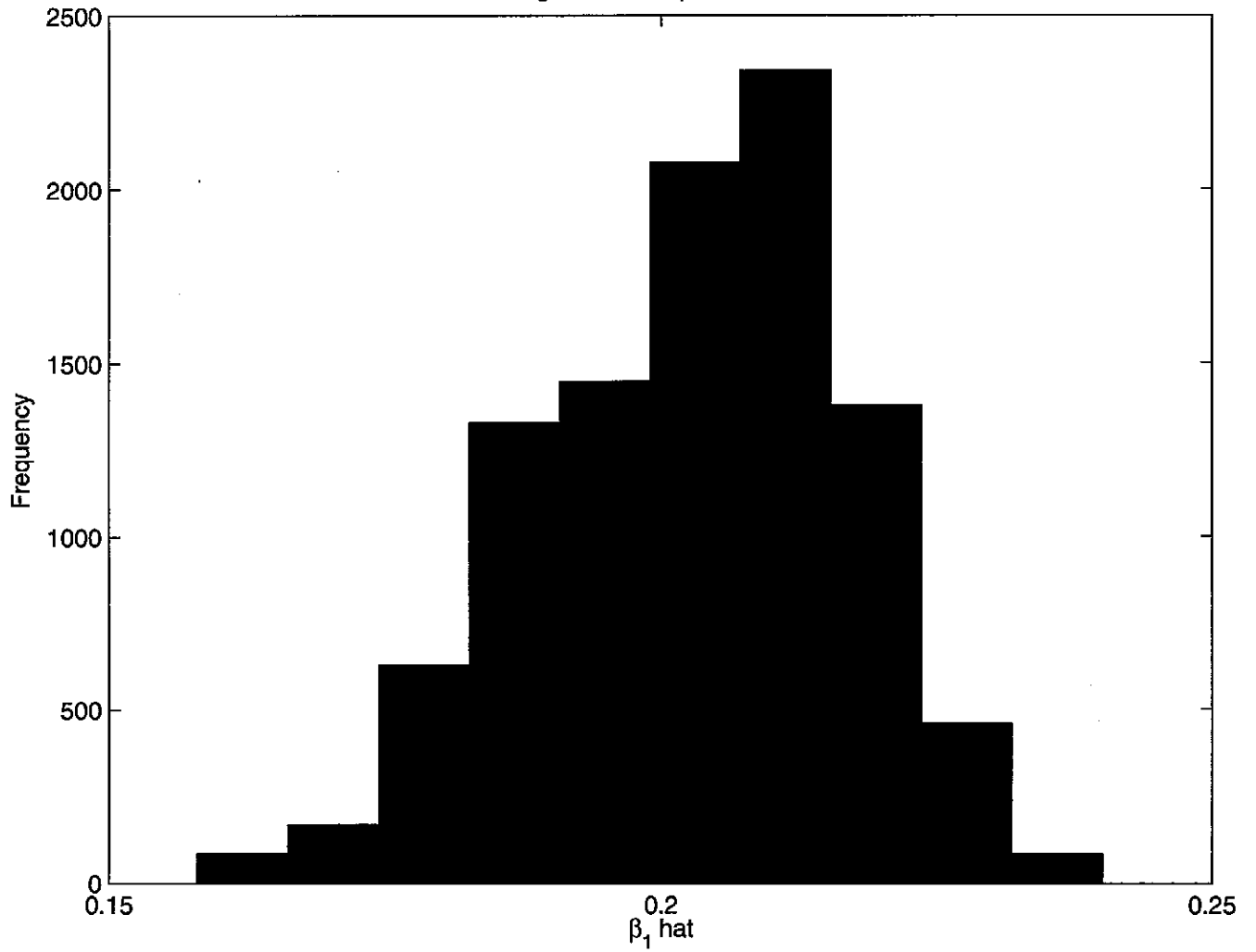


Figure 3E. Intercept vs. Slope

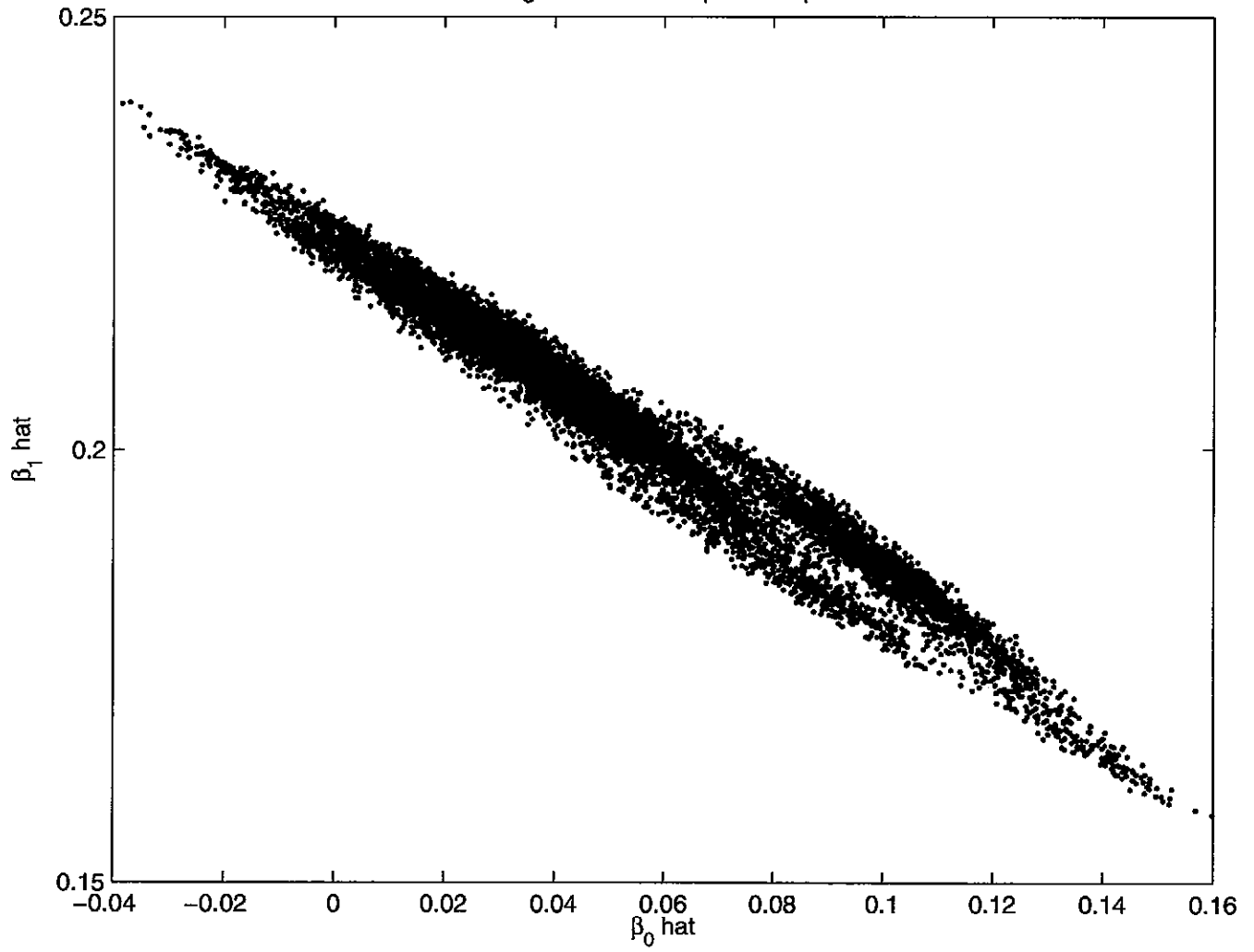


Figure 4A. Delete-1 Jackknife Means for Otolith &amp; Water Chemistries

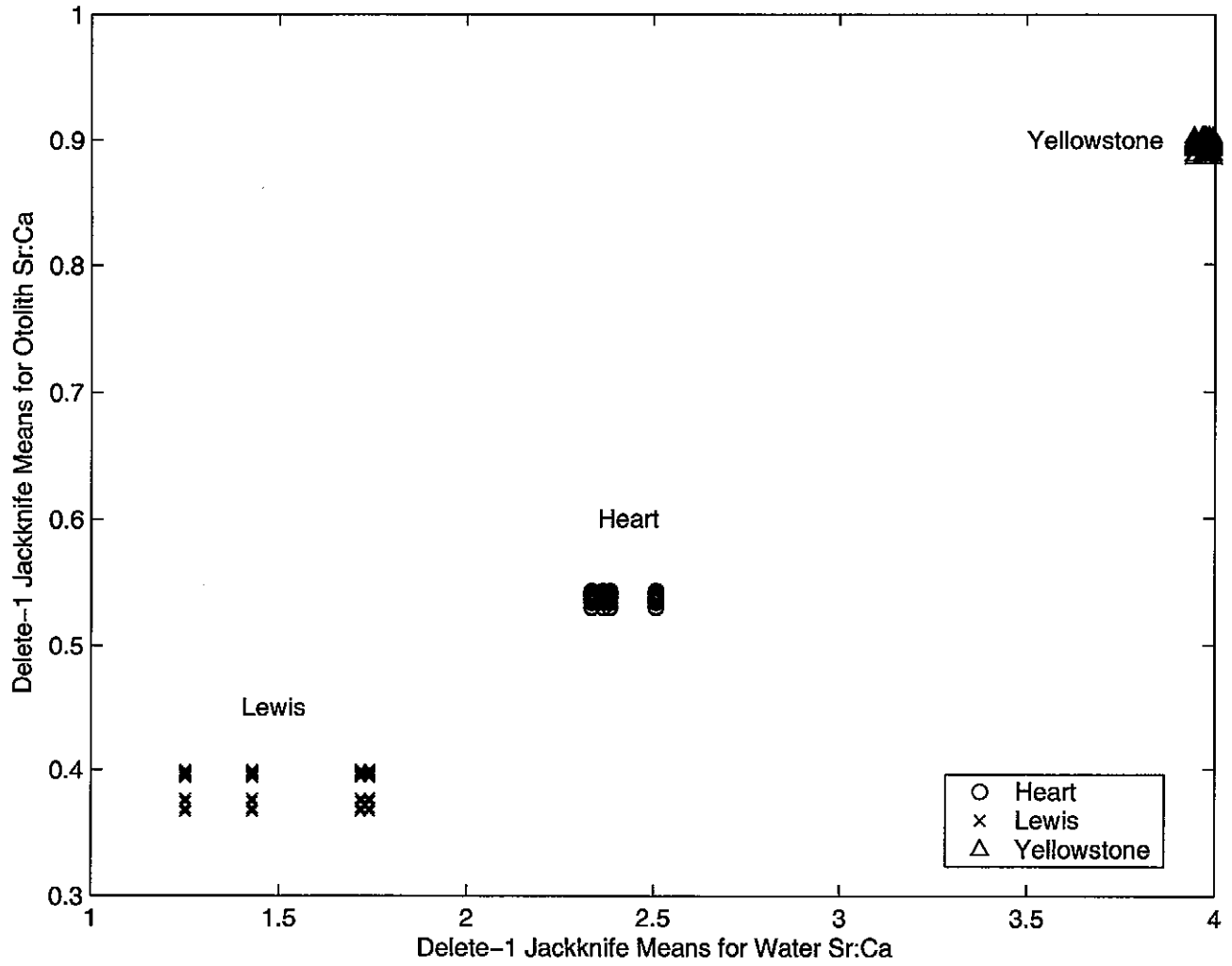




Figure 4B. Least Squares Lines

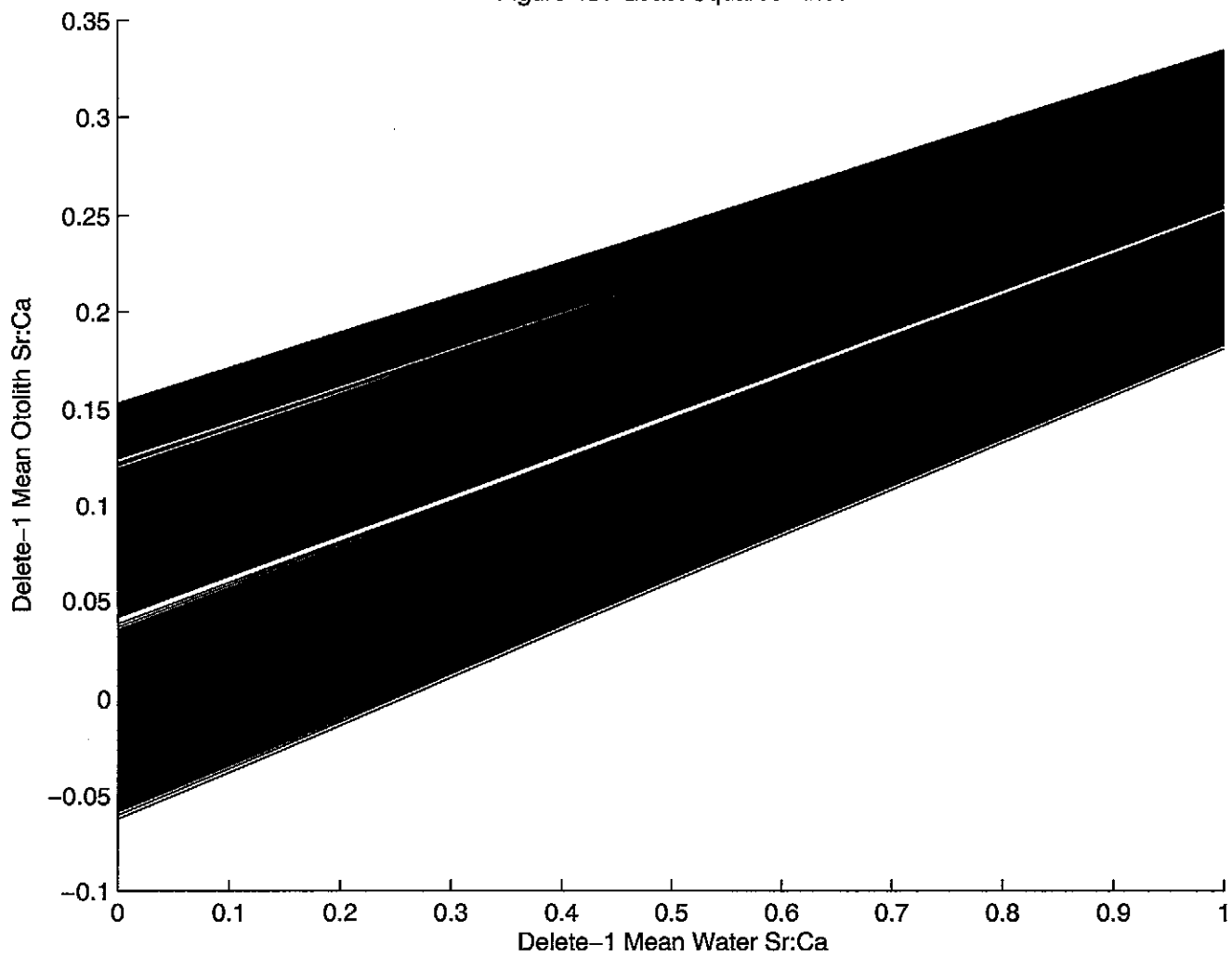


Figure 4C. Intercept Estimate

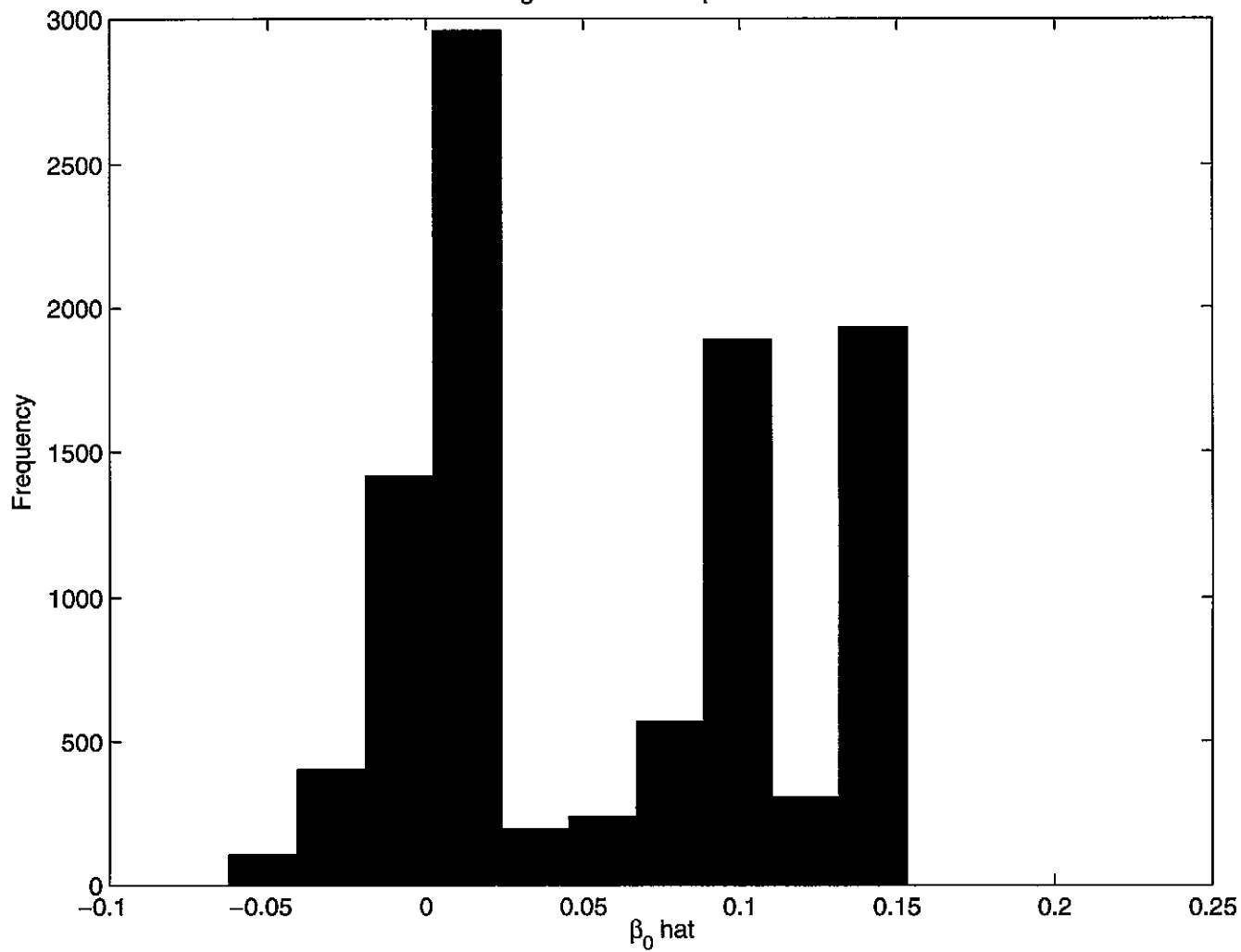


Figure 4D. Slope Estimate

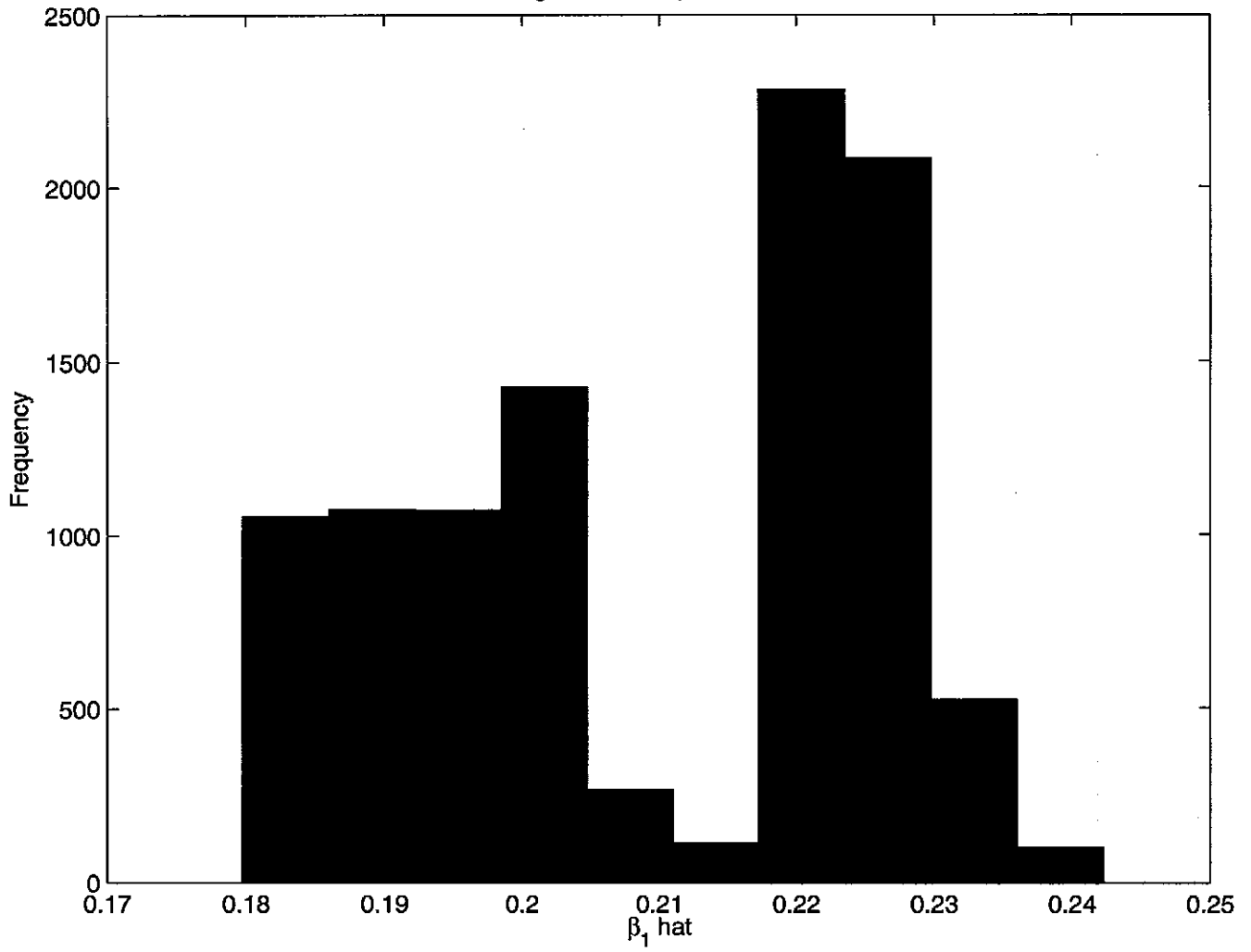
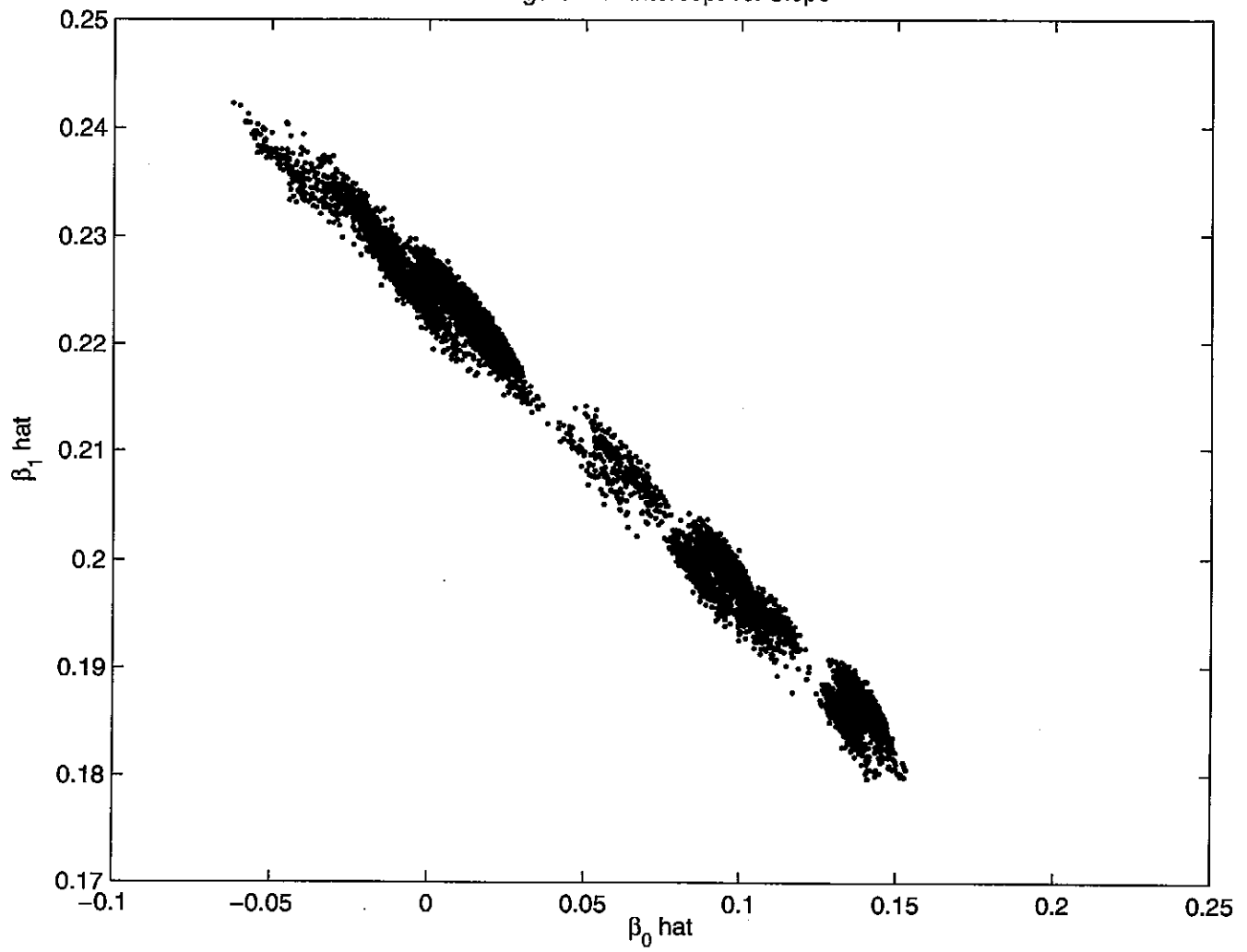


Figure 4E. Intercept vs. Slope



### Appendix III.

Matlab source code and outputted figures when deleting single pair of original mean otolith Sr:Ca ratio and mean water Sr:Ca ratio.

```

load chemistry.dat
X=[chemistry(:,1), chemistry(:,2), chemistry(:,3)];
location=X(:,1);
otolith=X(:,2);
water=X(:,3);
length(X);
n=223;

% Designate vectors for jackknifing
DBx = []; HLx=[]; LLx=[]; LPx=[]; MRx=[]; SCx=[]; YLx=[];

g=isnan(water);% Makes index of 0s & 1s where NaN=1

% Create counter indices for each jackknife vector space
k1=1; k2=1; k3=1; k4=1; k5=1; k6=1; k7=1;

% Loop to parse numeric water chemistries into vectors
% for each location

for i=1:n;
  if g(i)==0
    if location(i)==1
      DBx(k1,1)=X(i,3);
      k1=k1+1;
    end
    if location(i) == 2
      HLx(k2,1)=X(i,3);
      k2=k2+1;
    end
    if location(i) == 3
      LLx(k3,1)=X(i,3);
      k3=k3+1;
    end
    if location(i)==4
      LPx(k4,1)=X(i,3);
      k4=k4+1;
    end
    if location(i) == 5
      MRx(k5,1)=X(i,3);
      k5=k5+1;
    end
    if location(i) == 6
      SCx(k6,1)=X(i,3);
      k6=k6+1;
    end
    if location(i) == 7
      YLx(k7,1)=X(i,3);
      k7=k7+1;
    end
  end
end

end

% Create vector of otolith chemistries
% for each Missouri R. trib

DBy=zeros(223,1);
DBy=DBy+(location==1);
DBy=DBy .* otilith;
DBy(find(DBy==0))=[];

LPy=zeros(223,1);
LPy=LPy+(location==4);
LPy=LPy .* otilith;
LPy(find(LPy==0))=[];

MRy=zeros(223,1);

```

```

MRy=MRy+(location==5);
MRy=MRy .* otolith;
MRy(find(MRy==0))=[];

SCy=zeros(223,1);
SCy=SCy+(location==6);
SCy=SCy .* otolith;
SCy(find(SCy==0))=[];

% Create vector of otolith chemistries
% for each Yellowstone Lake

HLy=zeros(223,1);
HLy=HLy+(location==2);
HLy=HLy .* otolith;
HLy(find(HLy==0))=[];

LLy=zeros(223,1);
LLy=LLy+(location==3);
LLy=LLy .* otolith;
LLy(find(LLy==0))=[];

YLy=zeros(223,1);
YLy=YLy+(location==7);
YLy=YLy .* otolith;
YLy(find(YLy==0))=[];

% Initialize mean otolith & water chemistry by location

mDBx=mean(DBx); mDBy=mean(DBy);
mLPx=mean(LPx); mLPy=mean(LPy);
mMRx=mean(MRx); mMRy=mean(MRy);
mSCx=mean(SCx); mSCy=mean(SCy);
mHLx=mean(HLx); mHLy=mean(HLy);
mLLx=mean(LLx); mLLy=mean(LLy);
mYLx=mean(YLx); mYLy=mean(YLy);

% Create vector of mean water chemistry

Xm=[mDBx;mLPx;mMRx;mSCx];

% Create vector of mean otolith chemistry

Ym=[mDBy;mLPy;mMRy;mSCy];

% Loop to create delete-1 jackknife vectors
% of mean water chemistry & mean otolith chemistry

for i=1:4
    Xmtmp=Xm;
    Xmtmp([i],:)=[];

    Ymtmp=Ym;
    Ymtmp([i],:)=[];

X1=ones(3,1); % First column of design matrix for regression
Xmmat=[X1 Xmtmp]; % Design matrix for regression

Betatm(:,i)=pinv(Xmmat'*Xmmat)*Xmmat'*Ymtmp; % Regression parameters

% Calculate vector of coefficients of determination
b1=Betatm(2,i);
Sx=sqrt(var(Xmmat(:,2)));
Sy=sqrt(var(Ymtmp));
Rsqm(i)=(b1*(Sx/Sy))^2;

```

```

end

figure(1)
hold on;
x=[0 1];
for i=1:4
plot(x,[Betatm(1,i) Betatm(1,i) + Betatm(2,i)])
end
title('Figure E. Delete-1 Jackknife Least Squares Lines')
xlabel('Mean Water Sr:Ca')
ylabel('Mean Otolith Sr:Ca')

% Repeat for Yellowstone Lakes

% Create vector of mean water chemistry
Xy=[mHLx;mLLx;mYLx];

% Create vector of mean otolith chemistry
Yy=[mHLY;mLLY;mYLY];

% Loop to create delete-1 jackknife vectors
% of mean water chemistry & mean otolith chemistry

for i=1:3
    Xytmp=Xy;
    Xytmp([i],:)=[];

    Yytmp=Yy;
    Yytmp([i],:)=[];

    X1=ones(2,1); % First column of design matrix for regression
    Xymat=[X1 Xytmp]; % Design matrix for regression

    Betaty(:,i)=pinv(Xymat'*Xymat)*Xymat'*Yytmp; % Regression parameters

% Calculate vector of coefficients of determination
b1=Betaty(2,i);
Sx=sqrt(var(Xymat(:,2)));
Sy=sqrt(var(Yytmp));
Rsqr(i)=(b1*(Sx/Sy))^2;

end

figure(2)
hold on;
x=[0 1];
for i=1:3
plot(x,[Betaty(1,i) Betaty(1,i) + Betaty(2,i)])
end
title('Figure F. Delete-1 Jackknife Least Squares Lines')
xlabel('Mean Water Sr:Ca')
ylabel('Mean Otolith Sr:Ca')

% Repeat, but combine Missouri and Yellowstone

X=[mDBx;mLPx;mMRx;mSCx;mHLx;mLLx;mYLx];
Y=[mDBy;mLPy;mMRy;mSCy;mHLY;mLLY;mYLY];

for i=1:7
Xtmp=X;
Xtmp([i],:)=[];
Ytmp=Y;
Ytmp([i],:)=[];

```



```
X1=ones(6,1);
Xmat=[X1 Xtmp];
Betat(:,i)=pinv(Xmat'*Xmat)*Xmat'*Ytmp;

% Calculate vector of coefficients of determination
b1=Betat(2,i);
Sx=sqrt(var(Xmat(:,2)));
Sy=sqrt(var(Ytmp));
Rsq(i)=(b1*(Sx/Sy))^2;

end

figure(3)
hold on;
x=[0 1];
for i=1:7
plot(x,[Betat(1,i) Betat(1,i) + Betat(2,i)])
end
title('Delete-1 Jackknife Least Squares Lines')
xlabel('Mean Water Sr:Ca')
ylabel('Mean Otolith Sr:Ca')
```

Figure E. Delete-1 Jackknife Least Squares Lines

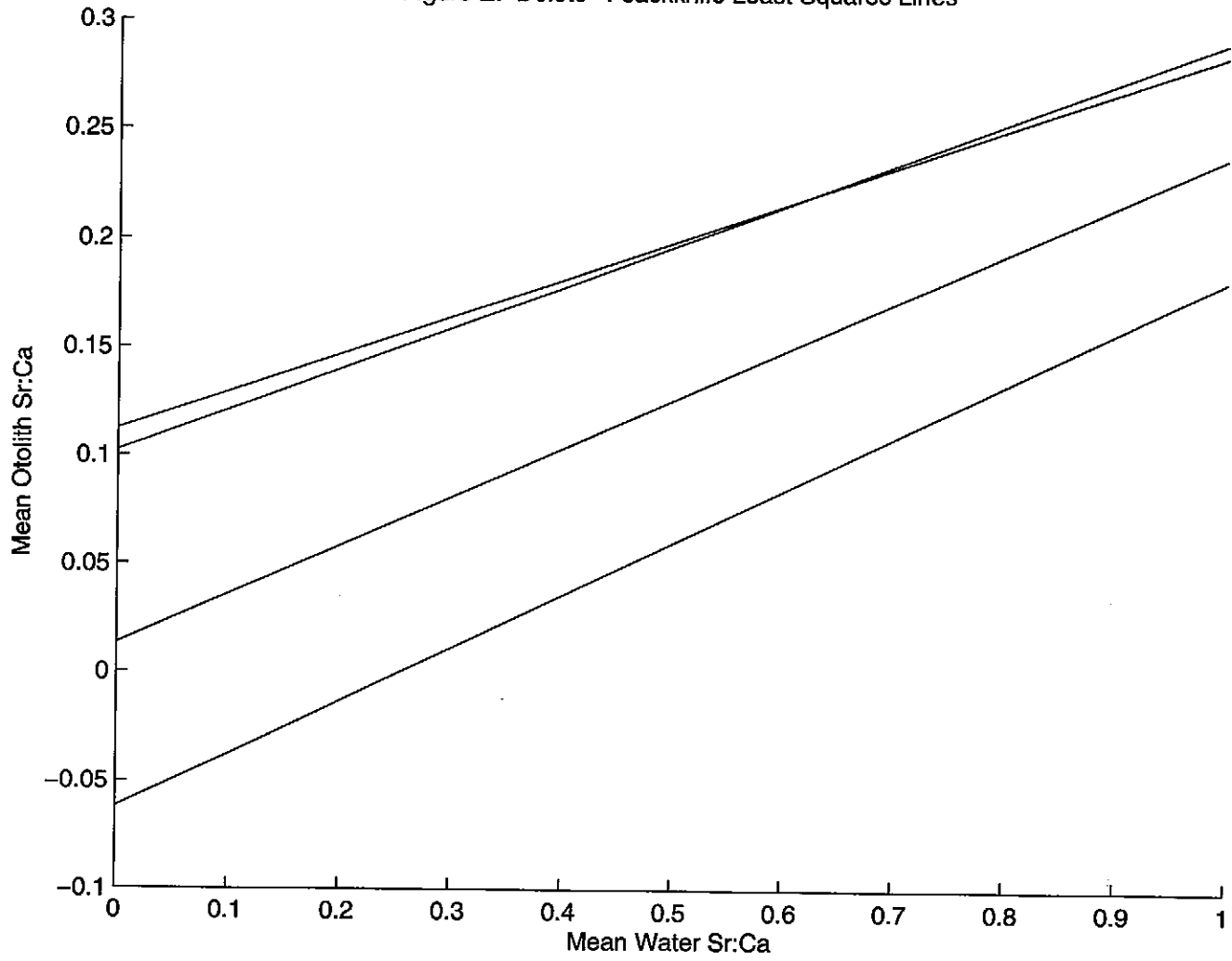


Figure F. Delete-1 Jackknife Least Squares Lines

