# STAT 8320 Spring 2015 Assignment 2

Peng Shao 14221765

April 8, 2015

```
libname da2 'C:\Users\psy6b\Desktop\8320 datasets';
ods graphics on;
options ls=70 ps=35;

data da2.h5q2;
infile 'C:\Users\psy6b\Desktop\8320 datasets\growthdata.dat';
input t @;
do i=1 to 6;
input y @;
output;
end;
run;

data aver;
set da2.h5q2;
by t;
retain count yave;
if first.t then do;
count=0;
yave=0;
end;
count+1;
yave+y;
if last.t then do;
y=yave/count;
```

```
i=7;
output;
end;
keep t i y;
run;

data da2.h5q2plot;
set da2.h5q2 aver;
run;

proc sort data=da2.h5q2plot;
by i;
run;

symbol interpol=join;
proc gplot data=da2.h5q2plot;
plot y*t=i;
run;
quit;


proc nlmixed data=da2.h5q2;
parameters beta1=200
beta2=850
beta3=350
resvar=40
varu=900;
e=exp(-(t-beta2)/beta3);
model y ~ normal((beta1+u)/(1+e), resvar);
random u ~ normal(0,varu) subject=i out=EBlups;
predict beta1/(1+e) out=pred;
predict (beta1+u)/(1+e) out=predB;
estimate 'Beta_3=350?' beta3-350;
ods output ParameterEstimates=estimates;
run;

proc sort data=pred;
by i t;
```

```
run;
proc sort data=predB;
by i t;
run;
data panelplot;
merge predB(rename=(pred=PredB)) pred;
by i t;
length type $20;
keep i t type resp;
type='measurement';
resp=y;
output;
type='cluster-specific';
resp=predb;
output;
type='population-average';
resp=pred;
output;
run;

proc sgpanel data=panelplot;
panelby i/spacing=5 rows=2 columns=3 novarname;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;

proc print data=eblups;
title 'Estimation of Random Effect';
var i Estimate tValue Probt;
run;

title;
```

```
data da2.h5q31;
infile 'C:\Users\psy6b\Desktop\8320 datasets\ssttornado532001.dat';
retain ss1-ss49;
array ss{49} ss1-ss49;
if _N_=1 then do; input ss1-ss49;end;
loc+1;
drop ss1-ss49;
do t=1 to 49;
sst=ss{t};
input torn @;
output;
end;
run;

data da2.h5q32;
infile 'C:\Users\psy6b\Desktop\8320 datasets\MOtornlatlon.dat';
loc+1;
input lat lon;
run;

proc sql;
create table da2.h5q3
as select * from da2.h5q31 as a, da2.h5q32 as b
where a.loc=b.loc;
run;
quit;

proc glimmix data=da2.h5q3 noitprint;
class loc;
model torn = sst sst*loc / dist=poisson link=log ddfm=betwithin solution;
random intercept / subject=loc type=sp(exp)(lon lat);
nloptions tech=newrap;
covtest 's' zerog;
output out=h5q3out pred(ilink)=predicted lcl(ilink)=lower ucl(ilink)=upper pred
run;
proc glimmix data=da2.h5q3 noitprint;
class loc;
model torn = sst sst*loc / dist=poisson link=log ddfm=betwithin solution;
```

```
random intercept sst/ subject=loc type=sp(exp)(lon lat);
nloptions tech=newrap;
run;


data panelplot2;
set h5q3out;
length type $20;
keep loc t type resp;
t=t+1952;
type='measurement';
resp=torn;
output;
type='cluster-specific';
resp=predicted;
output;
type='lower bound';
resp=lower;
output;
type='upper bound';
resp=upper;
output;
run;

proc sgpanel data=panelplot2;
where loc le 4 and loc ge 1;
panelby loc/rows=2 columns=2 spacing=5;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;
proc sgpanel data=panelplot2;
where loc le 8 and loc ge 5;
panelby loc/rows=2 columns=2 spacing=5;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;
```

```
proc sgpanel data=panelplot2;
where loc le 12 and loc ge 9;
panelby loc/rows=2 columns=2 spacing=5;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;
proc sgpanel data=panelplot2;
where loc le 16 and loc ge 13;
panelby loc/rows=2 columns=2 spacing=5;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;
proc sgpanel data=panelplot2;
where loc le 20 and loc ge 17;
panelby loc/rows=2 columns=2 spacing=5;
vline t/response=resp group=type;
colaxis fitpolicy=thin alternate;
rowaxis alternate;
run;


proc sort data=h5q3out;
by loc;
run;
data h5q3eval;
set h5q3out;
by loc;
drop sst t;
retain sumtorn summar sumpred sumu suml;
if first.loc then do;
sumtorn=0;
summar=0;
sumpred=0;
sumu=0;
suml=0;
end;
```

```
sumtorn+torn;
summar+margpred;
sumpred+predicted;
sumu+upper;
suml+lower;
if last.loc then do;
torn=sumtorn;
margpred=summar;
predicted=sumpred;
upper=sumu;
lower=suml;
output;
end;
run;
proc sort data=h5q3eval;
by torn;
run;

proc sgplot data=h5q3eval noautolegend;
scatter x=torn y=predicted/ datalabel=loc;
series x=torn y=torn;
xaxis label='Actual Measurements';
yaxis label='Predicted';
KEYLEGEND "Observations" "Reference Line";
run;
proc sgplot data=h5q3eval noautolegend;
scatter x=lat y=lon/ datalabel=loc;
run;
```

Figure 1: Regression Analysis

The NLMIXED Procedure

Specifications

| | |
|---|---|
| Data Set | DA2.H5Q2 |
| Dependent Variable | y |
| Distribution for Dependent Variable | Normal |
| Random Effects | u |
| Distribution for Random Effects | Normal |
| Subject Variable | i |
| Optimization Technique | Dual Quasi-Newton |
| Integration Method | Adaptive Gaussian Quadrature |

Dimensions

| | |
|---|---|
| Observations Used | 60 |
| Observations Not Used | 0 |
| Total Observations | 60 |
| Subjects | 6 |
| Max Obs per Subject | 10 |
| Parameters | 5 |
| Quadrature Points | 1 |

Parameters

| beta1 | beta2 | beta3 | resvar | varu | NegLogLike |
|---|---|---|---|---|---|
| 200 | 850 | 350 | 40 | 900 | 229.066515 |

Figure 1: *continued*

Iteration History

| Iter | Calls | NegLogLike | Diff | MaxGrad | Slope |
|---|---|---|---|---|---|
| 1 | 8 | 222.858676 | 6.207839 | 0.135255 | -3.44269 |
| 2 | 12 | 221.325222 | 1.533454 | 0.108492 | -1.10246 |
| 3 | 16 | 220.4229 | 0.902322 | 0.104243 | -0.27509 |
| 4 | 21 | 217.621537 | 2.801363 | 0.067182 | -0.77932 |
| 5 | 23 | 217.520935 | 0.100602 | 0.080529 | -0.72008 |
| 6 | 25 | 217.347603 | 0.173332 | 0.058901 | -0.43968 |
| 7 | 27 | 217.190674 | 0.156929 | 0.039087 | -0.15741 |
| 8 | 29 | 217.100941 | 0.089734 | 0.032728 | -0.10549 |
| 9 | 31 | 217.01773 | 0.08321 | 0.020822 | -0.16928 |
| 10 | 34 | 217.000377 | 0.017353 | 0.012076 | -0.02334 |
| 11 | 37 | 216.988644 | 0.011733 | 0.013059 | -0.00631 |
| 12 | 39 | 216.98306 | 0.005584 | 0.006165 | -0.01167 |
| 13 | 41 | 216.980417 | 0.002642 | 0.003475 | -0.00838 |
| 14 | 44 | 216.978694 | 0.001723 | 0.001598 | -0.00346 |
| 15 | 48 | 216.977806 | 0.000888 | 0.003132 | -0.00003 |
| 16 | 54 | 216.842805 | 0.135001 | 0.020194 | -0.00172 |
| 17 | 56 | 216.712242 | 0.130563 | 0.007343 | -0.11317 |
| 18 | 59 | 216.708022 | 0.00422 | 0.000293 | -0.00773 |
| 19 | 62 | 216.708009 | 0.000013 | 0.000037 | -0.00002 |
| 20 | 65 | 216.708009 | 1.783E-7 | 3.067E-6 | -2.54E-7 |

NOTE: GCONV convergence criterion satisfied.

Fit Statistics

| | |
|---|---|
| -2 Log Likelihood | 433.4 |
| AIC (smaller is better) | 443.4 |
| AICC (smaller is better) | 444.5 |
| BIC (smaller is better) | 442.4 |

Figure 1: *continued*

Parameter Estimates

| Parameter | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha |
|---|---|---|---|---|---|---|
| beta1 | 199.41 | 15.2372 | 5 | 13.09 | <.0001 | 0.05 |
| beta2 | 797.42 | 14.6250 | 5 | 54.52 | <.0001 | 0.05 |
| beta3 | 298.48 | 11.4146 | 5 | 26.15 | <.0001 | 0.05 |
| resvar | 49.8315 | 9.5902 | 5 | 5.20 | 0.0035 | 0.05 |
| varu | 1346.95 | 784.96 | 5 | 1.72 | 0.1468 | 0.05 |

Parameter Estimates

| Parameter | Lower | Upper | Gradient |
|---|---|---|---|
| beta1 | 160.24 | 238.58 | 4.83E-7 |
| beta2 | 759.82 | 835.01 | -3.07E-6 |
| beta3 | 269.14 | 327.82 | 2.884E-6 |
| resvar | 25.1791 | 74.4838 | 2.602E-6 |
| varu | -670.87 | 3364.76 | -1.06E-8 |

Additional Estimates

| Label | Estimate | Standard Error | DF | t Value | Pr > \|t\| | Alpha |
|---|---|---|---|---|---|---|
| Beta_3=350? | -51.5207 | 11.4146 | 5 | -4.51 | 0.0063 | 0.05 |

Additional Estimates

| Label | Lower | Upper |
|---|---|---|
| Beta_3=350? | -80.8627 | -22.1786 |

Figure 1: *continued*

```
                  Estimation of Random Effect

           Obs    i    Estimate     tValue      Probt

            1     1      9.0186     0.58954    0.58113
            2     2     48.3441     3.15760    0.02516
            3     3    -61.3289    -4.00406    0.01028
            4     4     28.0234     1.83138    0.12654
            5     5      7.4494     0.48697    0.64687
            6     6    -31.5072    -2.05886    0.09457
                      The GLIMMIX Procedure

                        Model Information

           Data Set                        DA2.H5Q3
           Response Variable               torn
           Response Distribution           Poisson
           Link Function                   Log
           Variance Function               Default
           Variance Matrix Blocked By      loc
           Estimation Technique            Residual PL
           Degrees of Freedom Method       Between-Within
                     Class Level Information

Class    Levels    Values

loc          20    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
                   19 20
             Number of Observations Read         980
             Number of Observations Used         980
```

Figure 1: *continued*

Dimensions

| | |
|---|---|
| G-side Cov. Parameters | 2 |
| Columns in X | 22 |
| Columns in Z per Subject | 1 |
| Subjects (Blocks in V) | 20 |
| Max Obs per Subject | 49 |

Optimization Information

| | |
|---|---|
| Optimization Technique | Newton-Raphson |
| Parameters in Optimization | 2 |
| Lower Boundaries | 2 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

Convergence criterion (PCONV=1.11022E-8) satisfied.

Fit Statistics

| | |
|---|---|
| -2 Res Log Pseudo-Likelihood | 5209.11 |
| Generalized Chi-Square | 1158.20 |
| Gener. Chi-Square / DF | 1.21 |

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error |
|---|---|---|---|
| Variance | loc | 0.1122 | 0.09420 |
| SP(EXP) | loc | 1.0000 | . |

Figure 1: *continued*

Solutions for Fixed Effects

| Effect | loc | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | -2.3014 | 0.1299 | 19 | -17.71 | <.0001 |
| sst | | 0.5251 | 0.7078 | 940 | 0.74 | 0.4583 |
| sst*loc | 1 | -2.3186 | 0.9117 | 940 | -2.54 | 0.0111 |
| sst*loc | 2 | -1.0147 | 1.0165 | 940 | -1.00 | 0.3184 |
| sst*loc | 3 | 0.04973 | 0.9467 | 940 | 0.05 | 0.9581 |
| sst*loc | 4 | -1.5813 | 0.9489 | 940 | -1.67 | 0.0960 |
| sst*loc | 5 | -2.3085 | 0.9207 | 940 | -2.51 | 0.0123 |
| sst*loc | 6 | 0.1876 | 0.9982 | 940 | 0.19 | 0.8510 |
| sst*loc | 7 | 0.6270 | 0.8985 | 940 | 0.70 | 0.4855 |
| sst*loc | 8 | -0.02175 | 0.9252 | 940 | -0.02 | 0.9812 |
| sst*loc | 9 | 0.4048 | 0.9063 | 940 | 0.45 | 0.6552 |
| sst*loc | 10 | -1.1604 | 1.0120 | 940 | -1.15 | 0.2518 |
| sst*loc | 11 | 0.6323 | 0.9086 | 940 | 0.70 | 0.4866 |
| sst*loc | 12 | -0.1940 | 0.9969 | 940 | -0.19 | 0.8457 |
| sst*loc | 13 | -0.4288 | 0.9597 | 940 | -0.45 | 0.6551 |
| sst*loc | 14 | -0.2062 | 1.0594 | 940 | -0.19 | 0.8457 |
| sst*loc | 15 | -0.3897 | 0.9737 | 940 | -0.40 | 0.6891 |
| sst*loc | 16 | -1.6004 | 0.9606 | 940 | -1.67 | 0.0960 |
| sst*loc | 17 | 0.4378 | 0.8840 | 940 | 0.50 | 0.6205 |
| sst*loc | 18 | 0.1002 | 0.9572 | 940 | 0.10 | 0.9167 |
| sst*loc | 19 | -0.4643 | 0.8798 | 940 | -0.53 | 0.5978 |
| sst*loc | 20 | 0 | . | . | . | . |

Type III Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| sst | 1 | 940 | 0.19 | 0.6607 |
| sst*loc | 19 | 940 | 2.04 | 0.0053 |

Figure 1: *continued*

```
                  Tests of Covariance Parameters
               Based on the Residual Pseudo-Likelihood

 Label      DF    -2 Res Log P-Like      ChiSq    Pr > ChiSq    Note

 s           2              5212.01       2.89        0.2352    --

--: Standard test with unadjusted p-values.
                        The GLIMMIX Procedure

                          Model Information

             Data Set                      DA2.H5Q3
             Response Variable             torn
             Response Distribution         Poisson
             Link Function                 Log
             Variance Function             Default
             Variance Matrix Blocked By    loc
             Estimation Technique          Residual PL
             Degrees of Freedom Method     Between-Within
                       Class Level Information

   Class    Levels    Values

   loc          20    1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
                      19 20
                Number of Observations Read         980
                Number of Observations Used         980
                              Dimensions

                   G-side Cov. Parameters           2
                   Columns in X                    22
                   Columns in Z per Subject         2
                   Subjects (Blocks in V)          20
                   Max Obs per Subject             49
```

Figure 1: *continued*

Optimization Information

| | |
|---|---|
| Optimization Technique | Newton-Raphson |
| Parameters in Optimization | 2 |
| Lower Boundaries | 2 |
| Upper Boundaries | 0 |
| Fixed Effects | Profiled |
| Starting From | Data |

Convergence criterion (PCONV=1.11022E-8) satisfied.

Estimated G matrix is not positive definite.

Fit Statistics

| | |
|---|---|
| -2 Res Log Pseudo-Likelihood | 5209.11 |
| Generalized Chi-Square | 1158.20 |
| Gener. Chi-Square / DF | 1.21 |

Covariance Parameter Estimates

| Cov Parm | Subject | Estimate | Standard Error |
|---|---|---|---|
| Variance | loc | 0.1122 | 0.09420 |
| SP(EXP) | loc | 1.0000 | . |

Figure 1: *continued*

Solutions for Fixed Effects

| Effect | loc | Estimate | Standard Error | DF | t Value | Pr > \|t\| |
|---|---|---|---|---|---|---|
| Intercept | | -2.3014 | 0.1299 | 19 | -17.71 | <.0001 |
| sst | | 0.6842 | 0.8048 | 940 | 0.85 | 0.3955 |
| sst*loc | 1 | -2.3722 | 0.9425 | 940 | -2.52 | 0.0120 |
| sst*loc | 2 | -1.0899 | 1.0884 | 940 | -1.00 | 0.3169 |
| sst*loc | 3 | -0.1626 | 1.0903 | 940 | -0.15 | 0.8815 |
| sst*loc | 4 | -1.7982 | 0.9983 | 940 | -1.80 | 0.0720 |
| sst*loc | 5 | -2.3047 | 0.9505 | 940 | -2.42 | 0.0155 |
| sst*loc | 6 | 0.2787 | 1.1485 | 940 | 0.24 | 0.8083 |
| sst*loc | 7 | 0.4961 | 1.0736 | 940 | 0.46 | 0.6442 |
| sst*loc | 8 | -0.3755 | 1.0646 | 940 | -0.35 | 0.7244 |
| sst*loc | 9 | 0.1688 | 1.0704 | 940 | 0.16 | 0.8747 |
| sst*loc | 10 | -1.2240 | 1.0758 | 940 | -1.14 | 0.2555 |
| sst*loc | 11 | 0.5692 | 1.0841 | 940 | 0.53 | 0.5997 |
| sst*loc | 12 | -0.2797 | 1.1234 | 940 | -0.25 | 0.8034 |
| sst*loc | 13 | -0.7348 | 1.0723 | 940 | -0.69 | 0.4934 |
| sst*loc | 14 | -0.06291 | 1.1820 | 940 | -0.05 | 0.9576 |
| sst*loc | 15 | -0.6257 | 1.0882 | 940 | -0.58 | 0.5654 |
| sst*loc | 16 | -1.7484 | 1.0079 | 940 | -1.73 | 0.0831 |
| sst*loc | 17 | 0.07983 | 1.0491 | 940 | 0.08 | 0.9394 |
| sst*loc | 18 | -0.03744 | 1.1037 | 940 | -0.03 | 0.9729 |
| sst*loc | 19 | -1.2069 | 0.9933 | 940 | -1.22 | 0.2246 |
| sst*loc | 20 | 0 | . | . | . | . |

Type III Tests of Fixed Effects

| Effect | Num DF | Den DF | F Value | Pr > F |
|---|---|---|---|---|
| sst | 1 | 940 | 0.15 | 0.6975 |
| sst*loc | 19 | 940 | 1.97 | 0.0077 |

Figure 2: Graphs for Regression Analysis

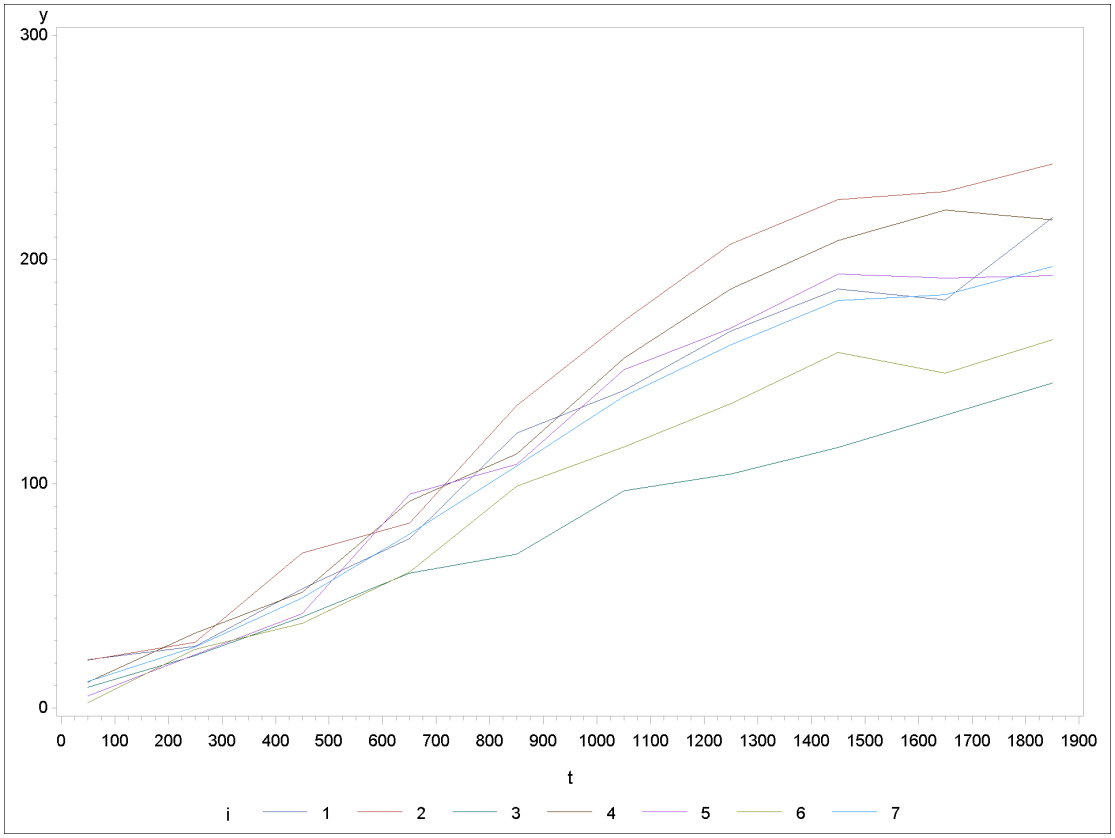

Figure 2: *continued*

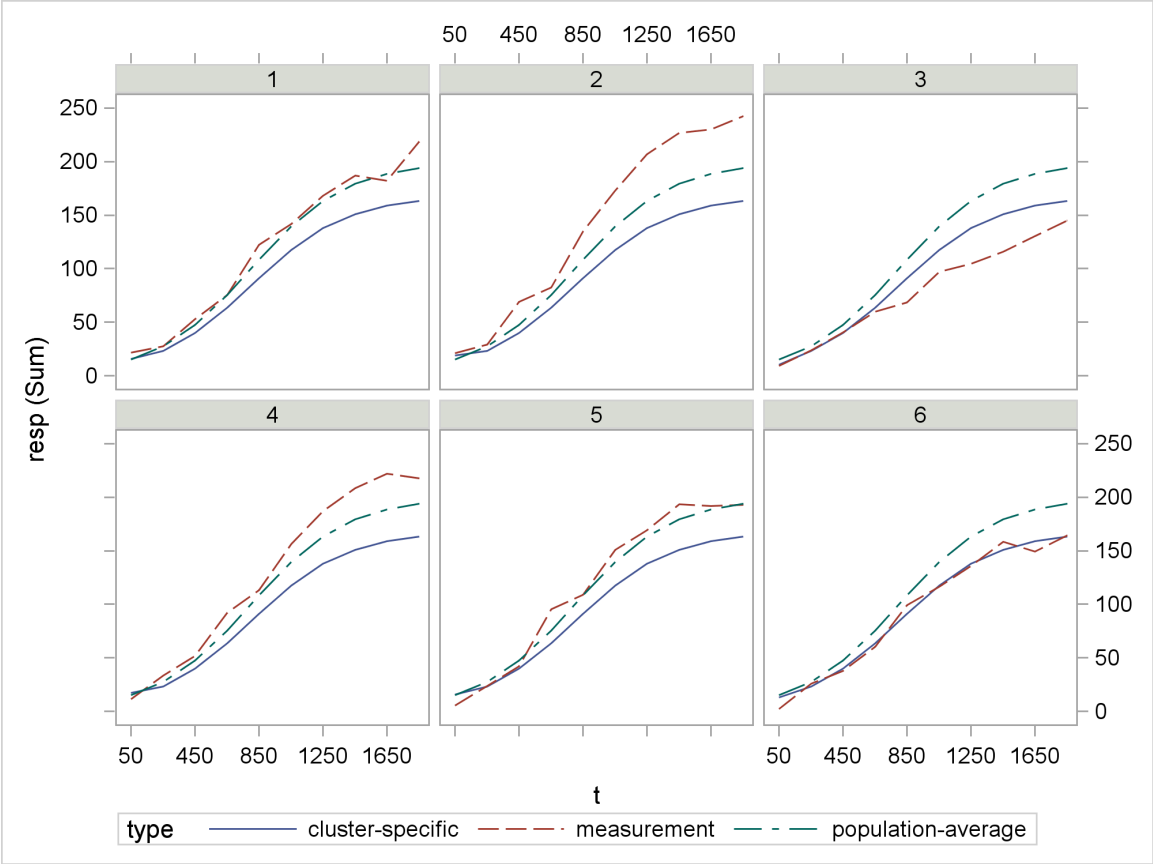

Figure 2: *continued*

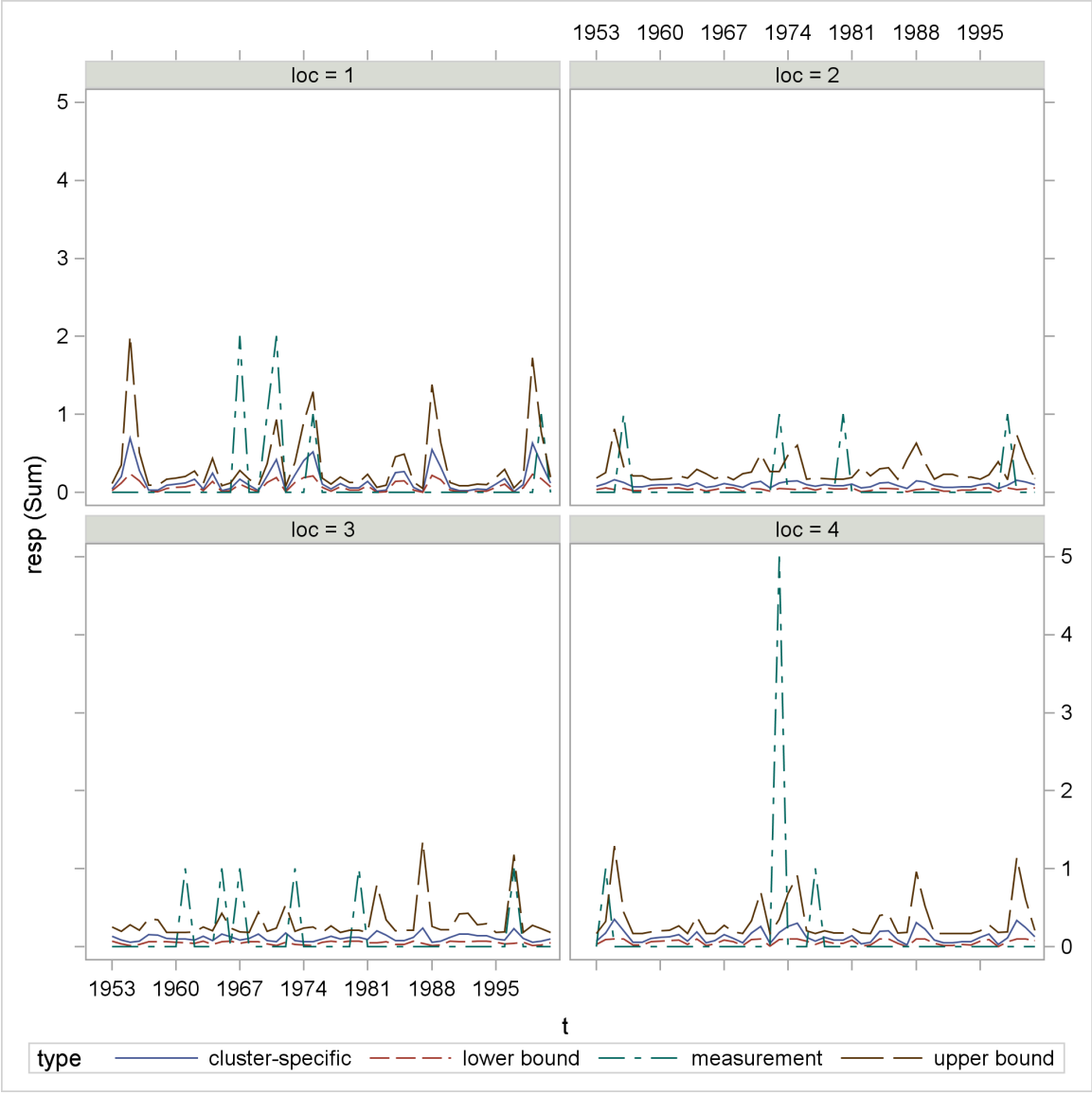

Figure 2: *continued*

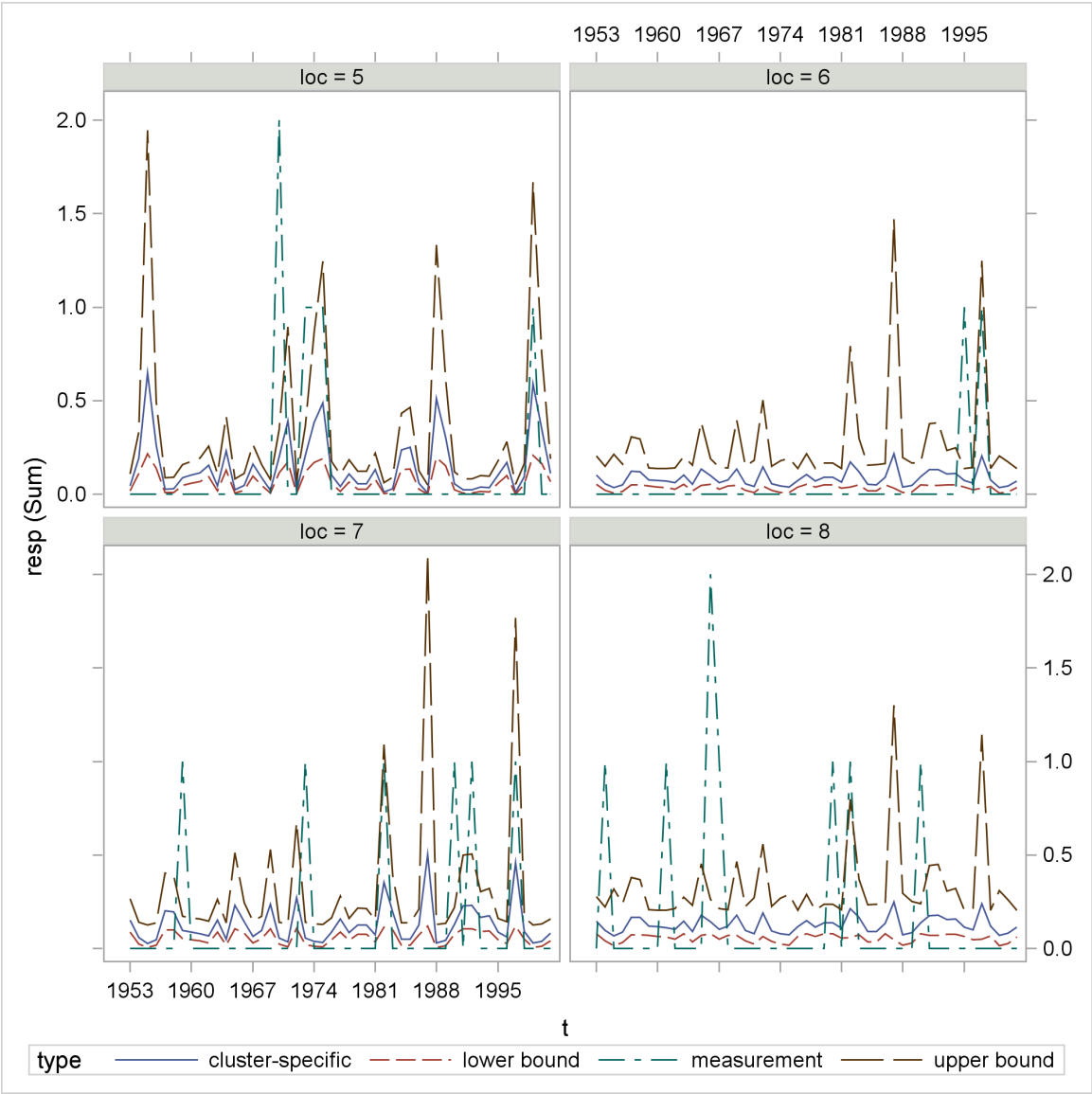

Figure 2: *continued*

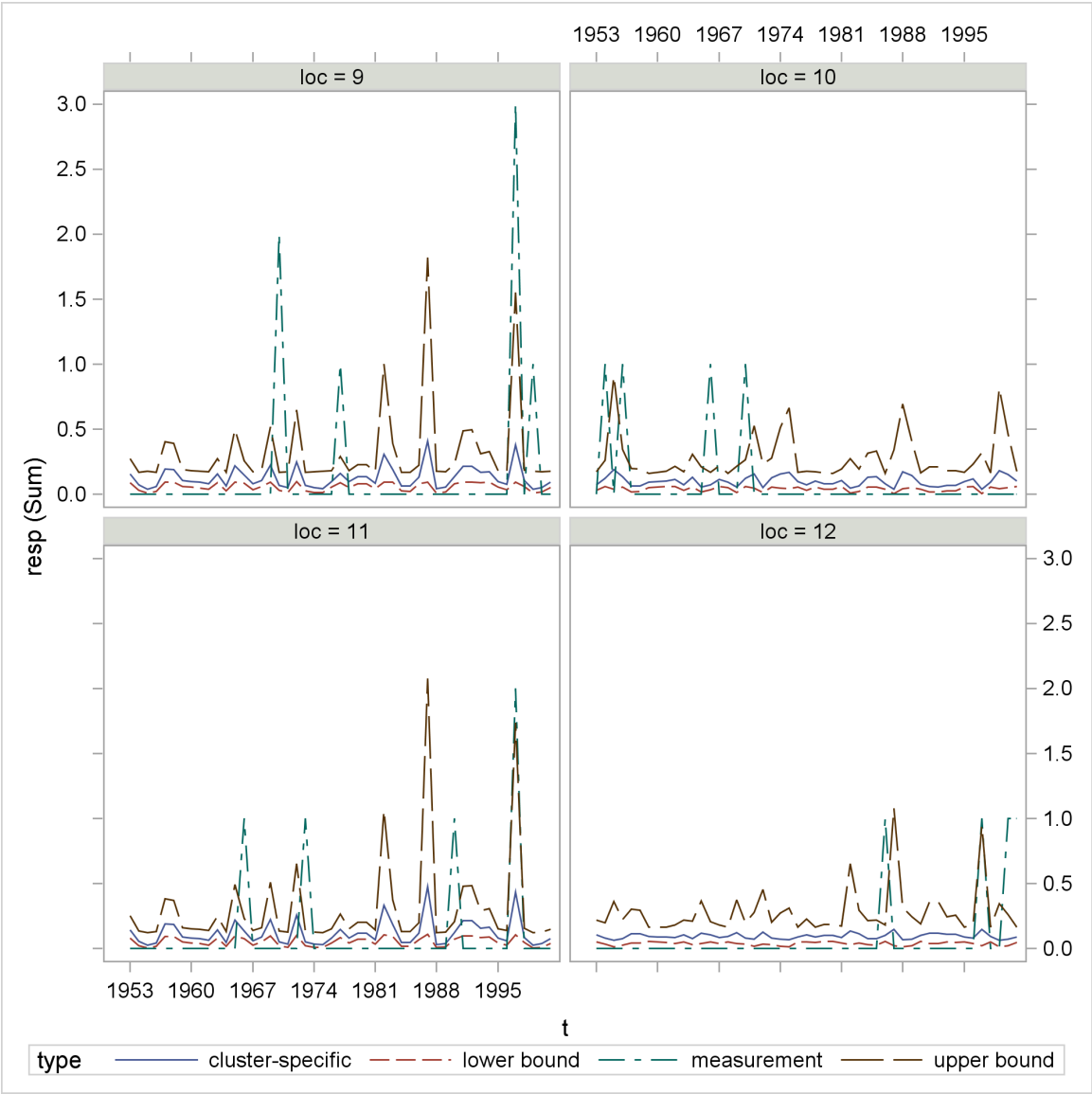

Figure 2: *continued*

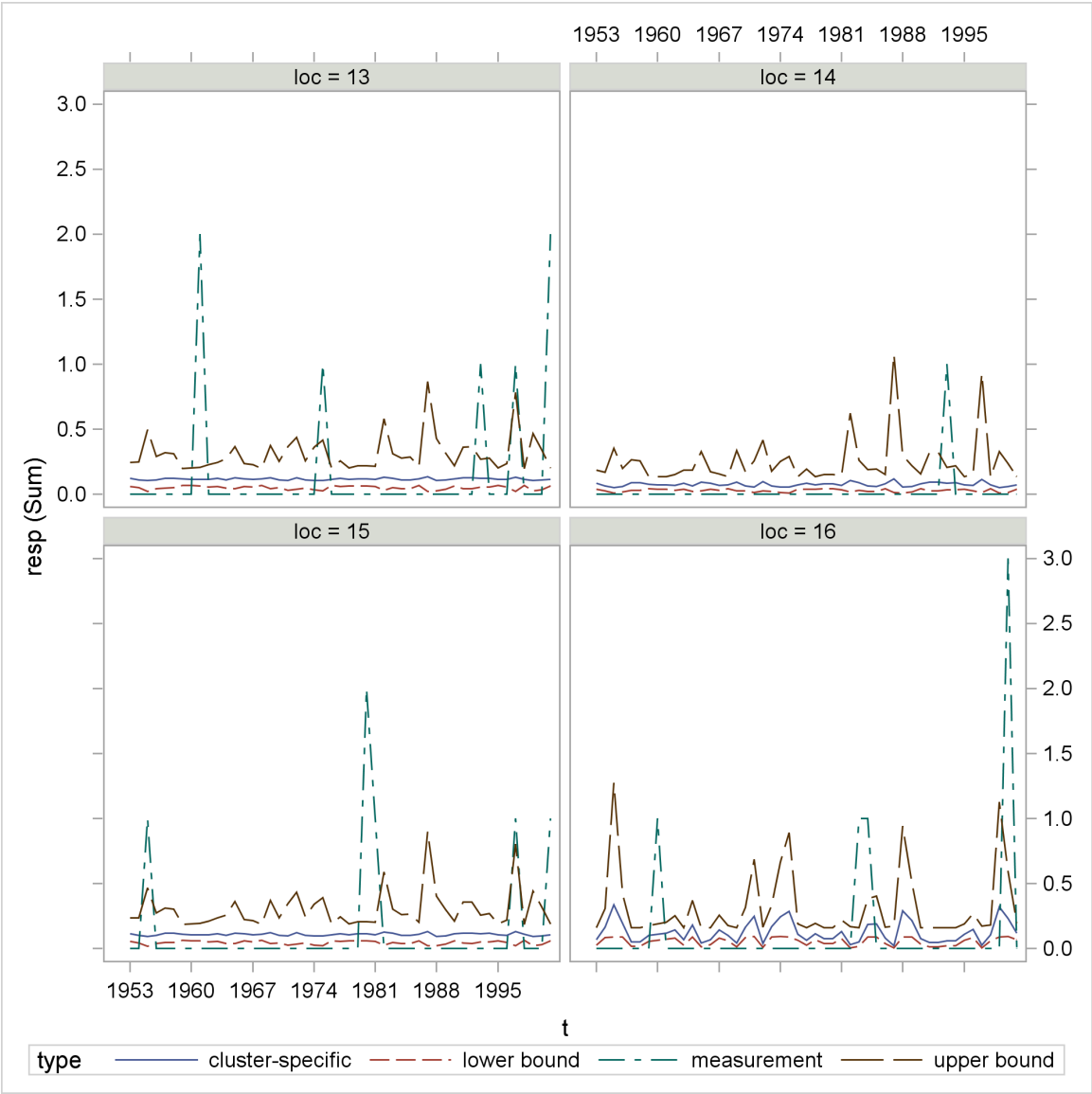

Figure 2: *continued*

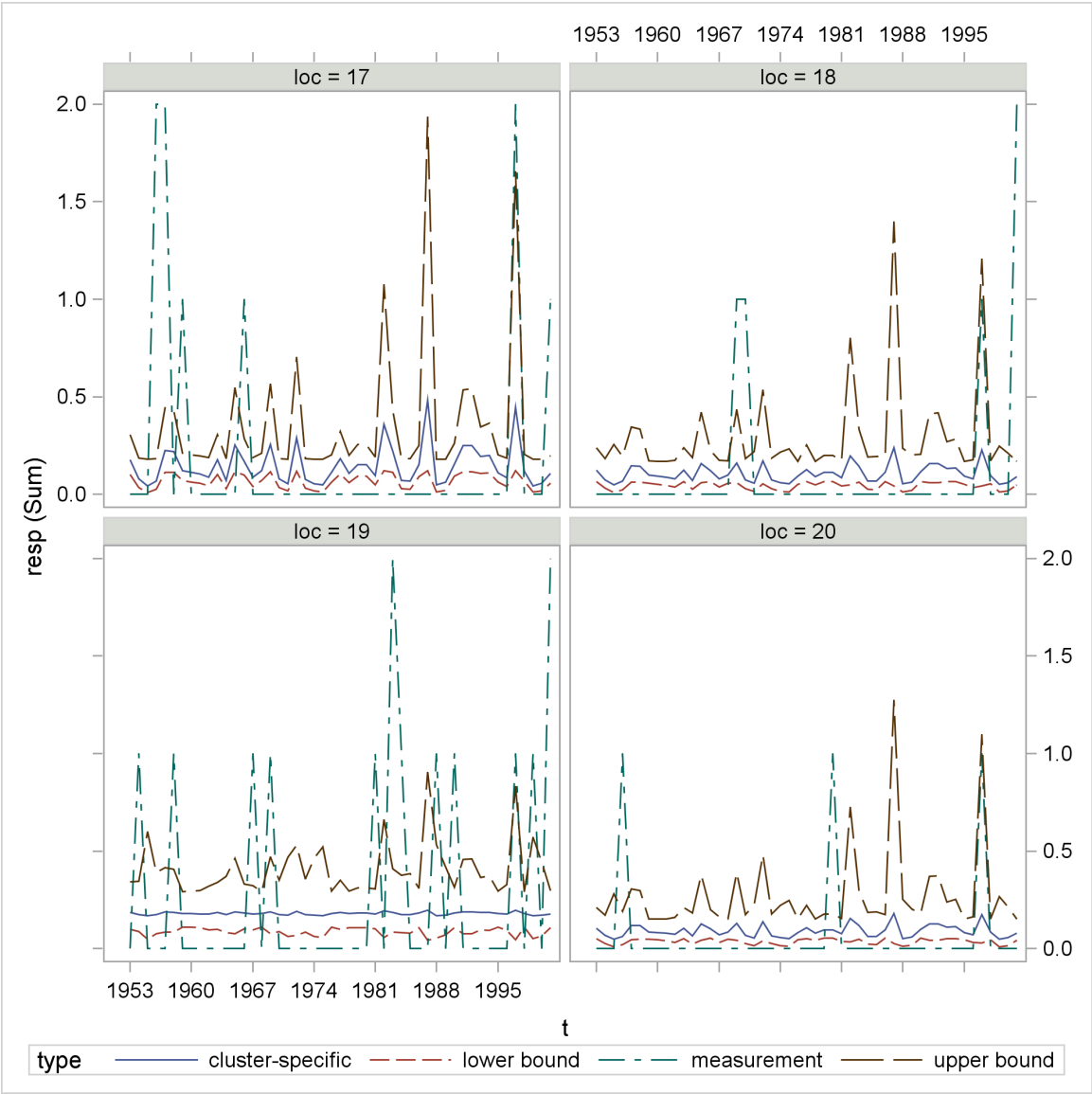

Figure 2: *continued*

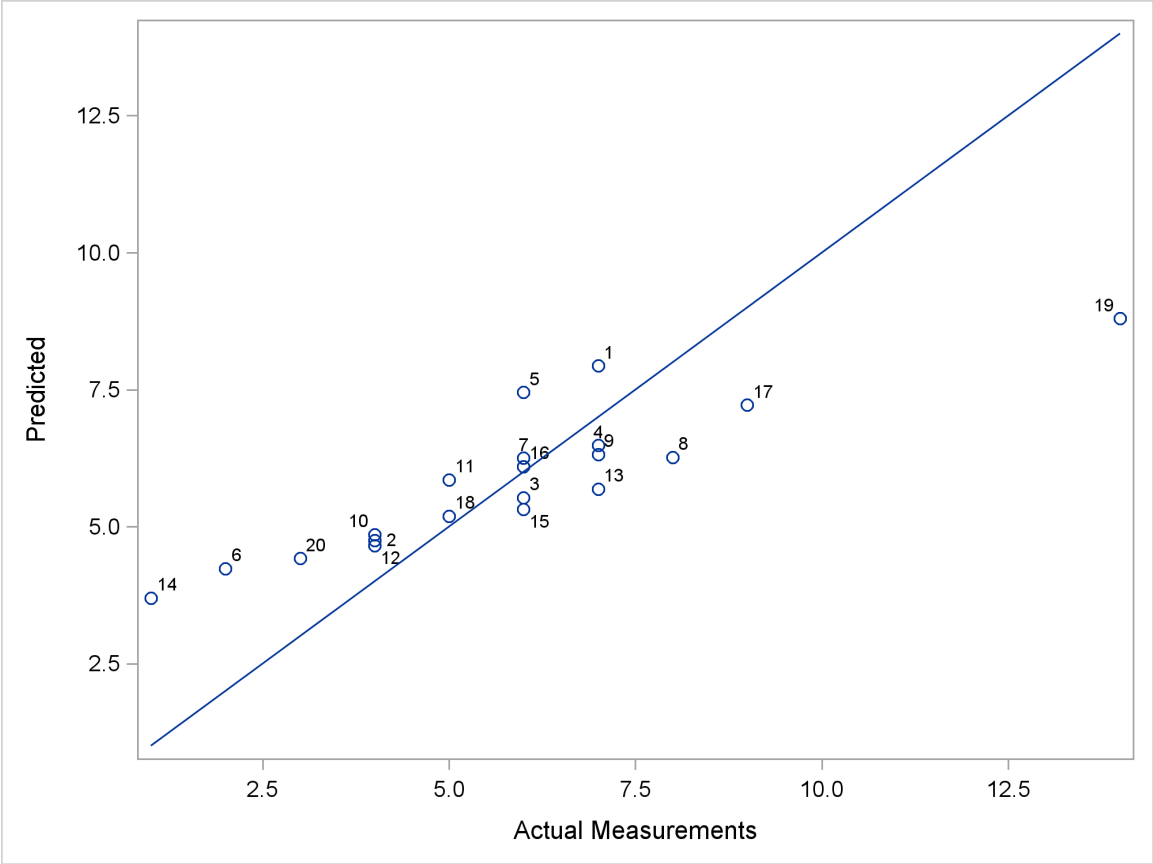

Figure 2: *continued*