Review of Seasonal models for ln(CRP)

Ed Stanek

Introduction

We review methods for estimating seasonality of LN(crp). The analyses are based on sine and cosine models similar to those used in the Seasons Main Outcome paper. Since the models fit to ln(crp) were fit in STATA, and we used SAS for the main paper, our review fits some simple seasonal models using SAS (taking advantage of earlier results from model development) to estimate amplitude of the seasonal effect. This advantage of using the simulation study is that we were able to investigate the correspondence of the estimated SE of the amplitude from the fitted models, relative to the standard error of repeated estimates of the amplitude from a simulation (checking the estimation of the SE). An additional advantage of the simulation is that the SAS program code (validated by the simulation) can be directly used to estimate the seasonal CRP models.

Data reported for the analyses of CRP was based on different measures of CRP than those contained in the original data. As an initial step, we compared these two data sources. Our comparison was based on adding the new CRP values to the original data. Finally, we fit seasonal models to Ln CRP and evaluate confidence intervals for the amplitude when transformed back to the natural scale.

Description of Data

We first combine the new CRP data with the original data from the Season's study and compare the CRP results. Table 1 illustrates the number of subjects, and the number of records per subject in the two data files.

Table 1. Comparison of Number of Measures per Subject in New and Old Data

new(In new Crp data file: NEW)
old(In original Seasons Data: OLD)

Frequency	1	2	4	5	Total
0	4	2	52	57	115
2	0	0	12	74	86
3	0	0	7	101	108
4	0	0	2	133	135
5	0	0	0	205	205
Total	4	2	73	570	649

Source: yz05p02.sas Seasons study 7/03/2005 Ying Zhou

Table 2 presents similar results where measures of CRP were given. Notice that measures were made on CRP for 113 subjects in the original Seasons data file. Using the newer

CRP file, 534 subjects had CRP measures. Also notice that on 111 subjects, measures were reported on CRP in both data sets.

Table 2. Comparison of Number of CRP Measures per Subject in New and Old Data

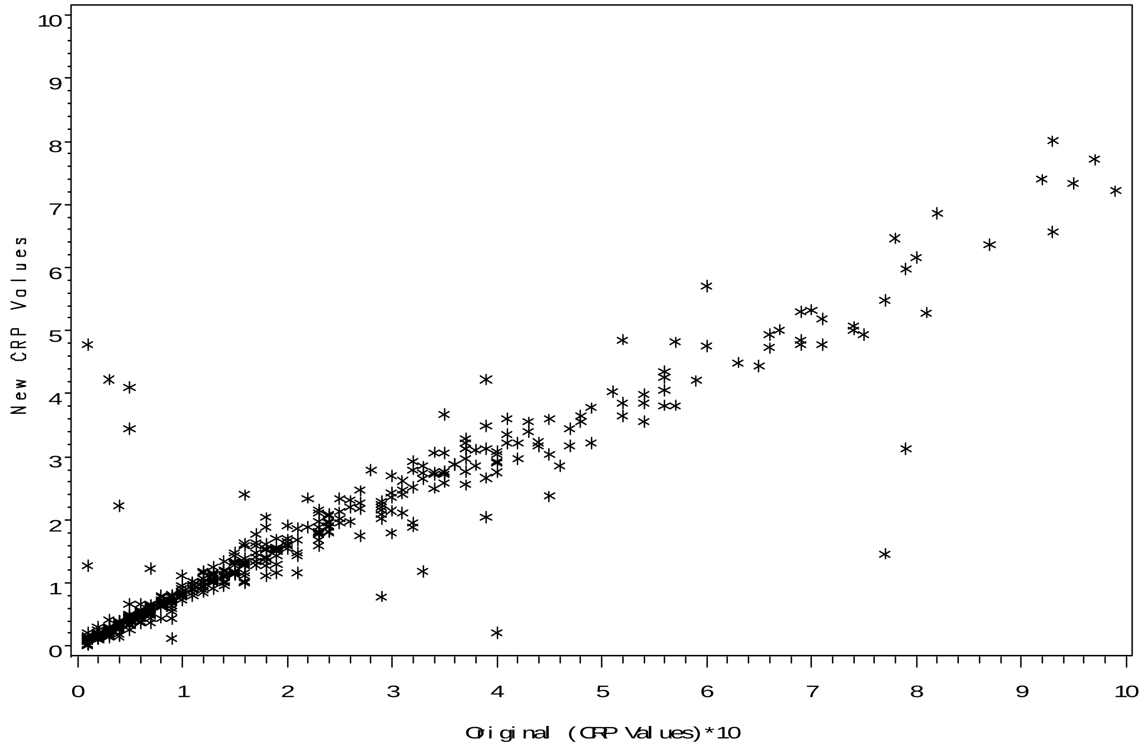
logcrp(Natural Log CRP (new): LOGCRP)
crp(C-reactive protein (units?)*: crp)

Frequency	0	3	4	5	Total
0	113	2	0	0	115
2	86	0	0	0	86
3	96	6	4	2	108
4	115	3	6	11	135
5	126	1	16	62	205
Total	536	12	26	75	649

Source: yz05p02.sas Seasons study 7/03/2005 Ying Zhou

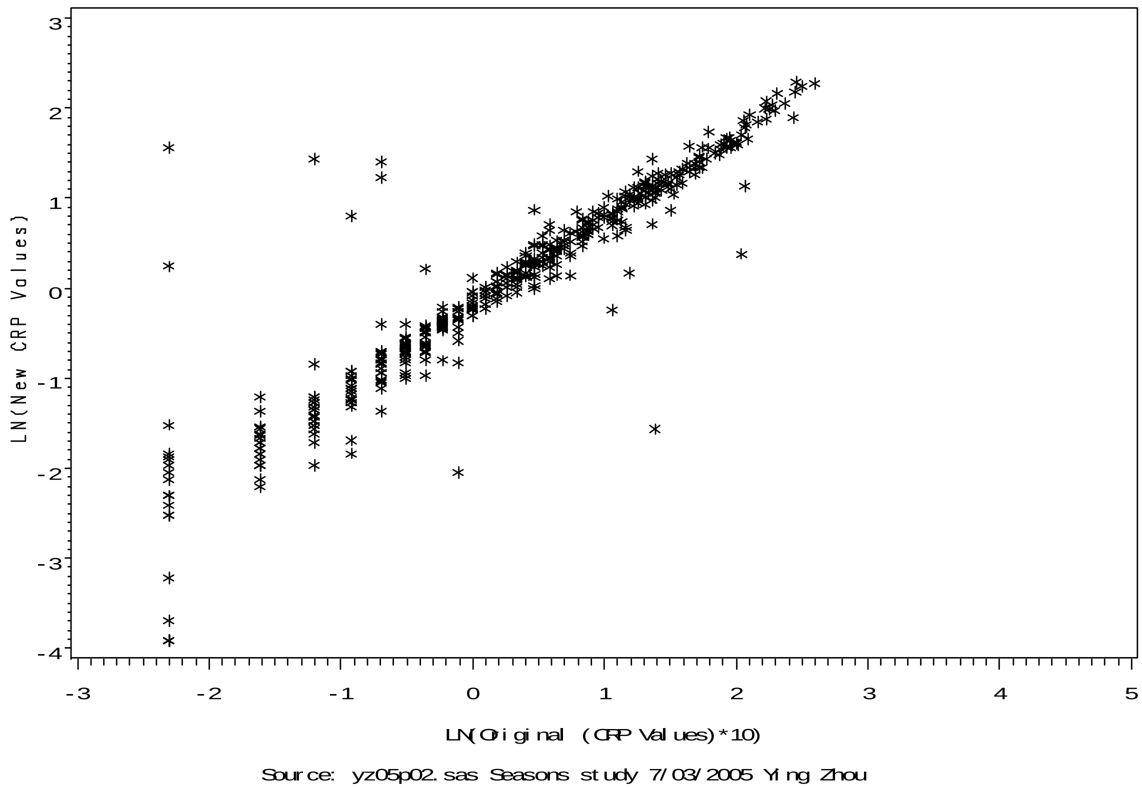
We next compared the CRP measures among the 111 subjects who had been measured in both data sets. We constructed this comparison using a ln scale, and using a natural scale, using a scatter plot to illustrate the correspondence of CRP values. Preliminary analysis indicated that the original CRP values needed to be multiplied by 10 to facilitate the comparison.

Figure 1. Scatter plot of New CRP values vs (Original CRP Values)*10 (n= 111 Subject)



Source: yz05p02.sas Seasons study 7/03/2005 Ying Zhou

Figure 2. Scatter plot of LN(New CRP) vs LN((Original CRP)*10) (n= 111 Subject)



The correlation of these measures of CRP are given in Table 3. On the natural scale, the correlation is 0.95, whereas on the natural log scale, this correlation is 0.93. However, notice that these correlations lump together repeated measures and subject, and do not separate the correlation of CRP measures for subjects, from the correlation of CRP measures per day.

Table 3. Simple Correlations of New and Old CRP on natural and LN scales

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ncrp	2061	1.72013	1.83304	3545	0.02000	9.96000
crp1	515	2.86485	5.92215	1475	0.10000	76.90000
logcrp	2061	0.01524	1.08644	31.40756	-3.91202	2.29858
nlcrp1	515	0.23967	1.23943	123.42946	-2.30259	4.34251

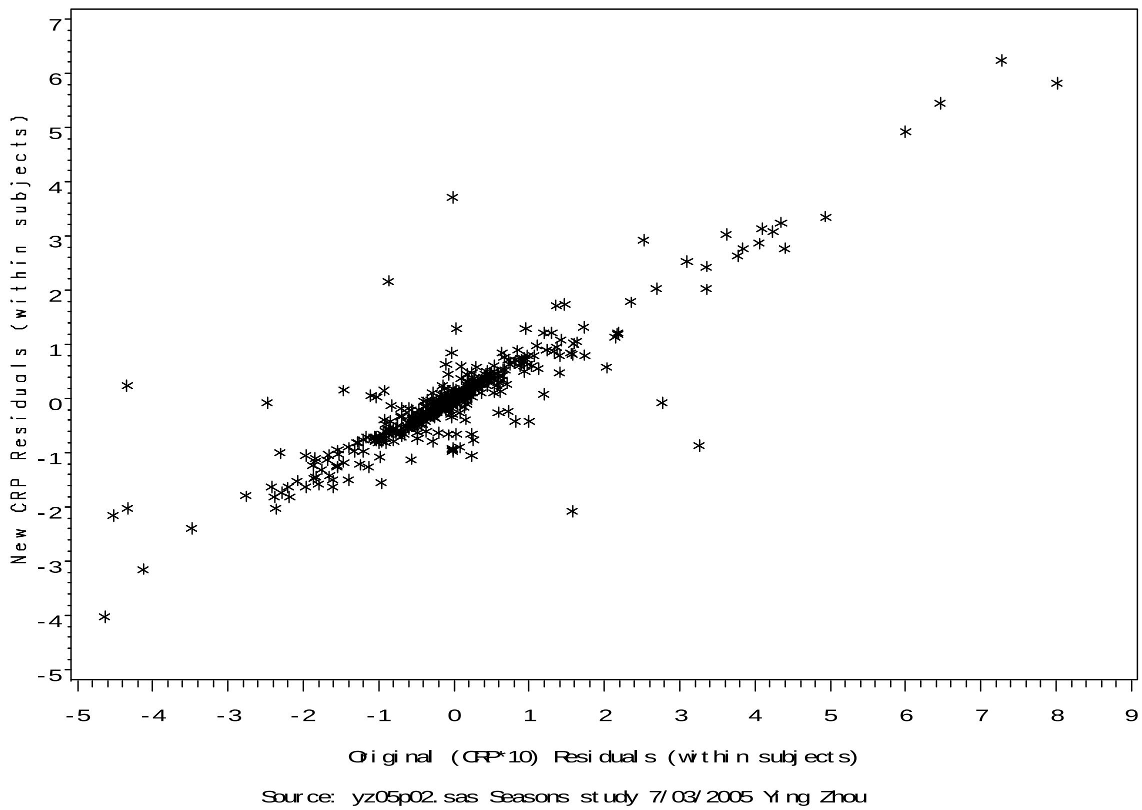
Pearson Correlation Coefficients
Prob > |r| under H0: Rho=0
Number of Observations

	ncrp	crp1	logcrp	nlcrp1
ncrp	1.00000	0.95393	0.84808	0.80312
CRP (new): NCRP		<.0001	<.0001	<.0001
	2061	488	2061	488
crp1	0.95393	1.00000	0.79191	0.65317
Original CRP*10: CRP1	<.0001		<.0001	<.0001
	488	515	488	515
logcrp	0.84808	0.79191	1.00000	0.93364
Natural Log CRP (new): LOGCRP	<.0001	<.0001		<.0001
	2061	488	2061	488
nlcrp1	0.80312	0.65317	0.93364	1.00000
Natural Log CRP*10 (original): NLCRP1	<.0001	<.0001	<.0001	
	488	515	488	515

Source: yz05p02.sas Seasons study 7/03/2005 Ying Zhou

We evaluated one more scatter plot to compare CRP measures within subject. This plot was formed by first subtracting the average CRP value per subject (using only days with both old and new CRP measures), and then plotting the residuals (after subtracting the mean). This plot was only done on the natural scale.

Figure 3. Scatter plot of within Subject residuals of New CRP values vs Original CRP*10 (n=[illegible] Subjects)



There was a fairly high correlation of within subject residual CRP values (see Table 4).

Table 4. Simple Correlations of Residual New and Old CRP on natural scales

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
ncrp_r	488	0	0.96768	0	-4.01167	6.24000
crp1_r	488	0	1.27824	0	-4.63333	8.02000

Pearson Correlation Coefficients, N = 488
Prob > |r| under H0: Rho=0

	ncrp_r	crp1_r
ncrp_r	1.00000	0.88781
		<.0001
crp1_r	0.88781	1.00000
	<.0001	

Source: yz05p02.sas Seasons study 7/03/2005 Ying Zhou

We can estimate the measurement error in CRP, assuming that the two CRP measures are replicates. This can be done by fitting a mixed model to these data, treating the two measures as replicates. It results in variance components for day to day variance equal to 1.0984, and for response error variance given by 0.1866. We use this estimate in subsequent simulations of seasonal effects.

We fit simple mixed models to the new CRP measures both on a natural and a logn scale. These models had a random subject effect. The result of the model fitting are given in Table 5.

Table 5. Simple mixed model for new CRP measures

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z	649
Subjects	1
Max Obs Per Subject	3150

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
id	2.3249
Residual	1.2569

Fit Statistics

-2 Res Log Likelihood	7423.0
AIC (smaller is better)	7427.0
AICC (smaller is better)	7427.0
BIC (smaller is better)	7436.0

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.7820	0.07085	533	25.15	<.0001

Source: yz05p03.sas Seasons study 7/03/2005 Ying Zhou

Table 6. Simple mixed model for new LN(CRP) measures

Dimensions

Covariance Parameters	2
Columns in X	1
Columns in Z	649
Subjects	1
Max Obs Per Subject	3150

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
id	0.8550
Residual	0.3375

Fit Statistics

-2 Res Log Likelihood	4862.6
AIC (smaller is better)	4866.6
AICC (smaller is better)	4866.6
BIC (smaller is better)	4875.5

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.04372	0.04220	533	1.04	0.3006

Source: yz05p03.sas Seasons study 7/03/2005 Ying Zhou

Modelling Seasonal Effects using Sine and Cosine Curves

Seasonal effects in the mixed models were evaluated two ways. First, the date of blood draw was used to classify each lipid measure into a season of the year (Winter: Nov. 6 - Feb. 4; Spring: Feb. 5 - May 6; Summer: May 7 - August 5; Fall: August 6 - Nov 5). The season was used as a fixed effect in the mixed model analyses, and estimates of change in cholesterol between seasons were constructed. Second, the date of blood draw was used to define sine and cosine coefficients for a sine shaped seasonal model that assumed a period of 365 days (see Koopmans, L.H. (1974), and http://www-unix.oit.umass.edu/~seasons/se35.pdf). Estimates of fixed effect regression coefficients for the sine and cosine terms in the mixed model were transformed to estimate the amplitude and phase of the seasonal effects. A first order Taylor series expansion was used to construct estimates of the variance of the amplitude and phase from the variance estimates of the sine and cosine coefficients (see http://www-unix.oit.umass.edu/~seasons/se35.pdf])

Initial gender specific mixed models were fit solely with seasonal effects, using subjects as random effects. Subsequent models were fit that controlled for various time dependent covariates separately, including body mass index, percent calories from saturated fat, physical activity (total waking METs and sport METs), total light exposure, hemoglobin, relative plasma volume, and age. Dietary and activity variables used were summarized from 24-hour reports prior to the blood draw, as explained subsequently. We evaluated the impact of including the covariates by assessing the extent to which the estimated amplitude of the seasonal effect changed when the covariates were considered; percent change in seasonal amplitude was used for this purpose and was obtained by comparing amplitude from the model with covariates to that of the model without covariates. Similar analyses were also conducted for subjects in the upper or lower quartile distribution of blood lipids.

Prior to the blood draw on subjects in the SEASONS Study, up to three 24HR measures were made of dietary intake, physical activity, and light exposure in each quarter. Since only one lipid measure was made during the three month interval, we first estimated average covariate values intake (or activity) for each subject during the interval. Rather than using simple three-day averages as the estimates, we used best linear unbiased predictors (BLUP) from mixed models to estimate longer term (28 day) average intake (or activity) (Ref. Stanek et al 1999), assuming that the longer term average covariate values had stronger relationships with lipids. Thus, the BLUP estimates of dietary intake, physical activity, and light exposure were used in analyses of lipids.

Finally, % of the study population by gender with total cholesterol greater than or equal to 240mg/dl in both winter and summer was obtained using the subject average at each season. Percent of the study population by gender with LDL cholesterol greater than or equal to 160mg/dl was obtained by the same method.

Seasonal Models of CRP

We fit seasonal models to CRP using SAS code identical to the SAS code in sne05p06.sas (for the simulation program). The models were fit based on the natural scale, and on the natural log scale. The results based on the natural scale are given in Table 7.

Table 7. Results of Mixed models on new CRP measures (natural scale)

Dimensions

Covariance Parameters	2
Columns in X	3
Columns in Z	643
Subjects	1
Max Obs Per Subject	3150
Number of Observations Read	3150
Number of Observations Used	2061

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	8357.41145426	
1	2	7431.22979937	0.00167565
2	1	7427.82947375	0.00006388
3	1	7427.70991522	0.00000011
4	1	7427.70971304	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
id	2.3309
Residual	1.2537

Fit Statistics

-2 Res Log Likelihood	7427.7
AIC (smaller is better)	7431.7
AICC (smaller is better)	7431.7
BIC (smaller is better)	7440.6

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	1.7818	0.07094	533	25.12	<.0001
sin	-0.05041	0.03716	1525	-1.36	0.1752
cos	0.06343	0.03582	1525	1.77	0.0768

Covariance Matrix for Fixed Effects

Row	Effect	Col1	Col2	Col3
1	Intercept	0.005032	-0.00004	-0.00004
2	sin	-0.00004	0.001381	-0.00002
3	cos	-0.00004	-0.00002	0.001283

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
sin	1	1525	1.84	0.1752
cos	1	1525	3.14	0.0768

Estimates of Mean, Amplitude and Phase for CRP

Obs	mean	amp	v_amp	peakjul	v_phase
1	1.78176	0.16205	.005381304	326	678.850

Source: yz05p04.sas Seasons study 7/03/2005 Ying Zhou

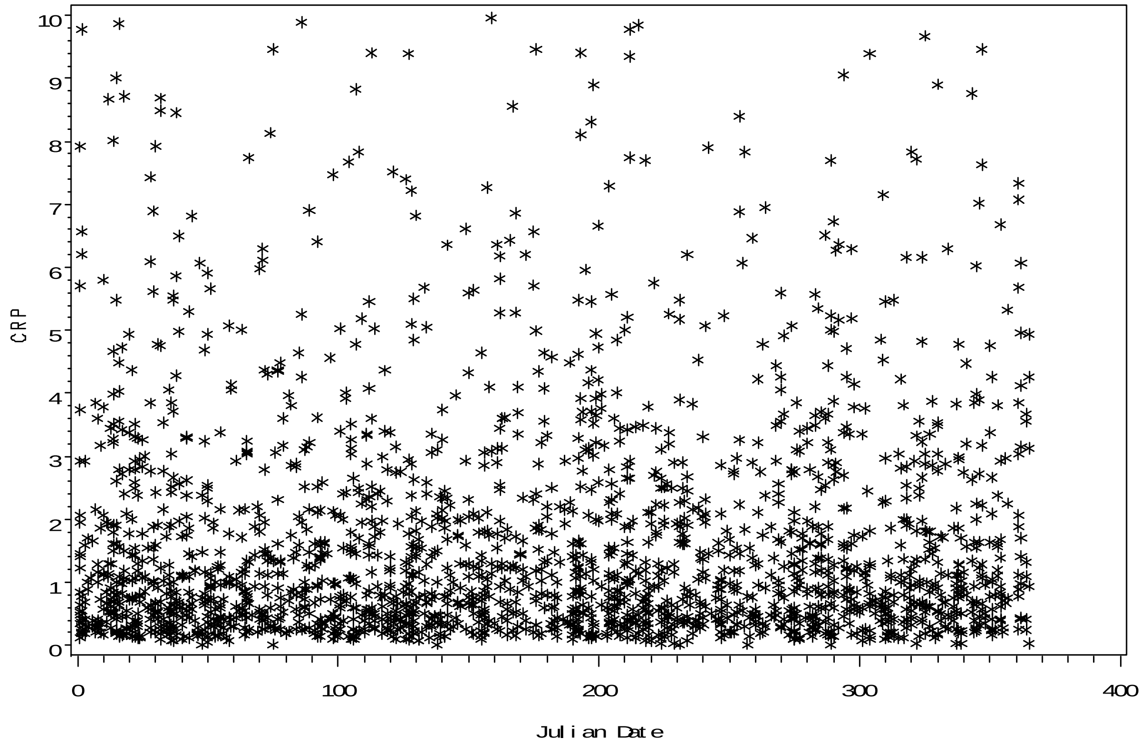
A list of the predicted CRP values by month of the year is given in Table 7b.

Table 7b List of Predicted CRP by 15th of month based on estimates from mixed model

Obs	mth	ncrp_p
1	1	1.82999
2	2	1.79097
3	3	1.74949
4	4	1.71665
5	5	1.70126
6	6	1.70744
7	7	1.73353
8	8	1.77255
9	9	1.81403
10	10	1.84687
11	11	1.86226
12	12	1.85608

Source: yz05p04.sas Seasons study 7/03/2005 Ying Zhou

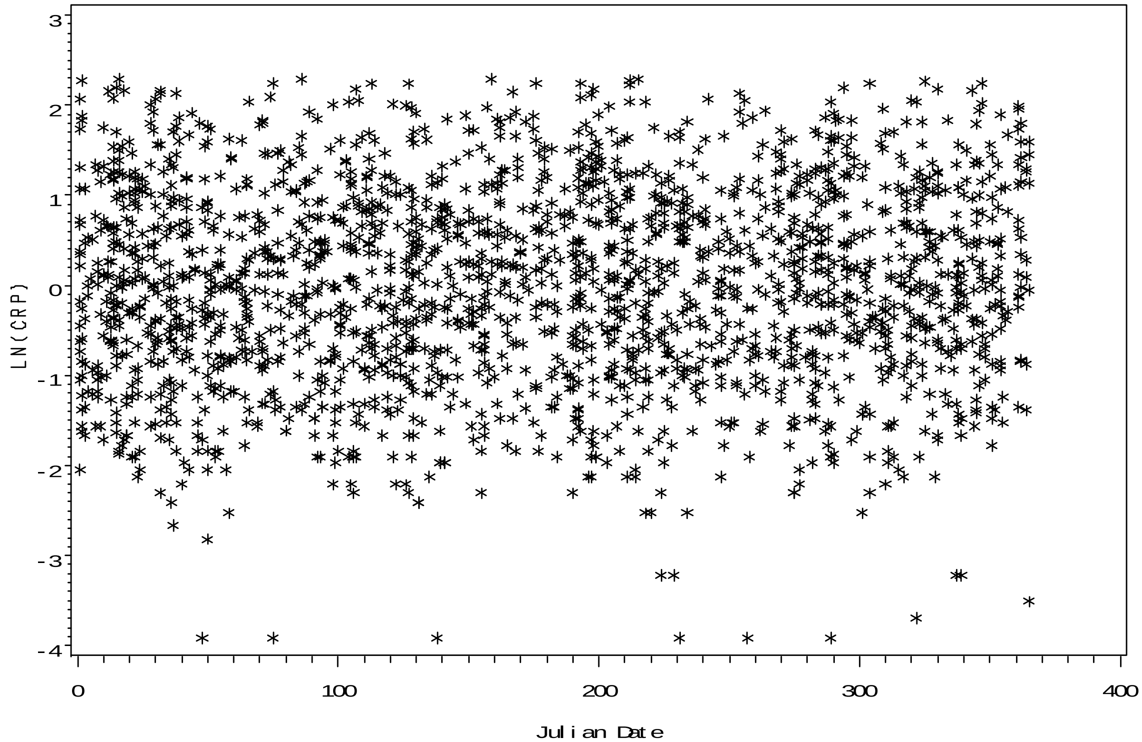
Figure 4. Scatter plot of CRP vs Julian Date



Source: yz05p04.sas Seasons study 7/03/2005 Ying Zhou

Next we fit a similar model but use LN(CRP) as the dependent variable.

Figure 5. Scatter plot of LN(CRP) vs Julian Date



Source: yz05p05.sas Seasons study 7/03/2005 Ying Zhou

Table 8. Results of Mixed models on new LN(CRP) measures (natural scale)

Dimensions

Covariance Parameters	2
Columns in X	3
Columns in Z	643
Subjects	1
Max Obs Per Subject	3150

Number of Observations Read	3150
Number of Observations Used	2061

Iteration History

Iteration	Evaluations	-2 Res Log Like	Criterion
0	1	6203.85340662	
1	2	4870.83466431	0.00000391
2	1	4870.83252809	0.00000000

Convergence criteria met.

Covariance Parameter Estimates

Cov Parm	Estimate
id	0.8559
Residual	0.3370

Fit Statistics

-2 Res Log Likelihood	4870.8
AIC (smaller is better)	4874.8
AICC (smaller is better)	4874.8
BIC (smaller is better)	4883.8

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	0.04415	0.04222	533	1.05	0.2963
sin	-0.03217	0.01931	1525	-1.67	0.0959
cos	0.02076	0.01860	1525	1.12	0.2646

Covariance Matrix for Fixed Effects

Row	Effect	Col1	Col2	Col3
1	Intercept	0.001783	-0.00001	-0.00001
2	sin	-0.00001	0.000373	-6.73E-6
3	cos	-0.00001	-6.73E-6	0.000346

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
sin	1	1525	2.78	0.0959
cos	1	1525	1.25	0.2646

Estimates of Mean, Amplitude and Phase for ln(CRP)

Obs	mean	amp	v_amp	peakjul	v_phase
1	0.044145	0.076573	.001484134	307	801.780

Source: yz05p05.sas Seasons study 7/03/2005 Ying Zhou

Table 8b. List of Predicted LN(CRP) by 15th of month based on estimates from Mixed model

Obs	mth	logcrp_p	Predicted Geometric Mean of CRP: cRP_P_GM
1	1	0.055871	1.05746
2	2	0.036077	1.03674
3	3	0.018445	1.01862
4	4	0.007698	1.00773
5	5	0.006718	1.00674
6	6	0.015766	1.01589
7	7	0.032418	1.03295
8	8	0.052213	1.05360
9	9	0.069845	1.07234
10	10	0.080592	1.08393
11	11	0.081572	1.08499
12	12	0.072524	1.07522

Source: yz05p05.sas Seasons study 7/03/2005 Ying Zhou

Table 9. Results of Amplitude and Phase Estimates for CRP by Sub-group
Estimates of Amplitude and Phase for CRP

Subgroup= Males

mean	amp	v_amp	peakjul	v_phase
1.78843	0.10068	0.011728	313	3667.41

Subgroup= Females

mean	amp	v_amp	peakjul	v_phase
1.77549	0.23214	.009470278	331	615.156

Subgroup= BMI LT 25

mean	amp	v_amp	peakjul	v_phase
1.12461	0.053102	0.012348	356	16096.37

Subgroup= BMI from 25 to 30

mean	amp	v_amp	peakjul	v_phase
1.76523	0.31433	0.014881	327	496.817

Subgroup= BMI from 25 to 30, males

mean	amp	v_amp	peakjul	v_phase
1.70066	0.44229	0.028615	344	479.323

Subgroup= BMI from 25 to 30, Females

mean	amp	v_amp	peakjul	v_phase
1.87400	0.26542	0.026621	270	1157.60

Subgroup= BMI GE 30

mean	amp	v_amp	peakjul	v_phase
2.88126	0.068099	0.025401	261	18145.63

Subgroup= Age Less than 50

mean	amp	v_amp	peakjul	v_phase
1.59355	0.17937	0.010227	318	1070.45

Subgroup= Age greater or equal to 50

mean	amp	v_amp	peakjul	v_phase
2.00545	0.15587	0.011372	335	1519.92

Subgroup= SBP GT 90

mean	amp	v_amp	peakjul	v_phase
1.79019	0.13032	.006276683	319	1205.08

Source: yz05p07.sas Seasons study 7/03/2005 Ying Zhou

Note: Individual records (quarters) were used to stratify data. This resulted in some subjects data being divided between groups (as for example with BMI and SBP)

We also calculated the average CRP in each season (using the “light season definitions”) and tabulated the percent of subjects with CRP values GT 3 in each season. The results are given in Table 10.

Table 10. Frequency of subjects whose mean crp is greater than 3 by season

The FREQ Procedure

Table of sea1 by crp_c

sea1(Light/Seasons/(SEA1))	crp_c(CRP Ave >3?: CRP_C)		
Frequency Row Pct	0	1	Total
Winter	439 84.26	82 15.74	521
Spring	445 85.41	76 14.59	521
Summer	416 84.04	79 15.96	495
Fall	453 88.48	59 11.52	512
Total	1753	296	2049

Frequency Missing = 2

Source: sne05p16.sas Seasons study 6/23/2005 ejs

References

Koopmans, L.H. (1974), and The spectral analysis of time series, Academic Press, New York 366p.