Using Fourier Analysis in Evaluating Program Office Estimates

UPAS

Stephen Cox, USA PEO Aviation

Matthew Latham, USA AMCOM

Angela Lemke, USA AMCOM

June 2010

Special Thanks To:

Wallace Willard, NASA

Penny Pride, USA PEO Aviation

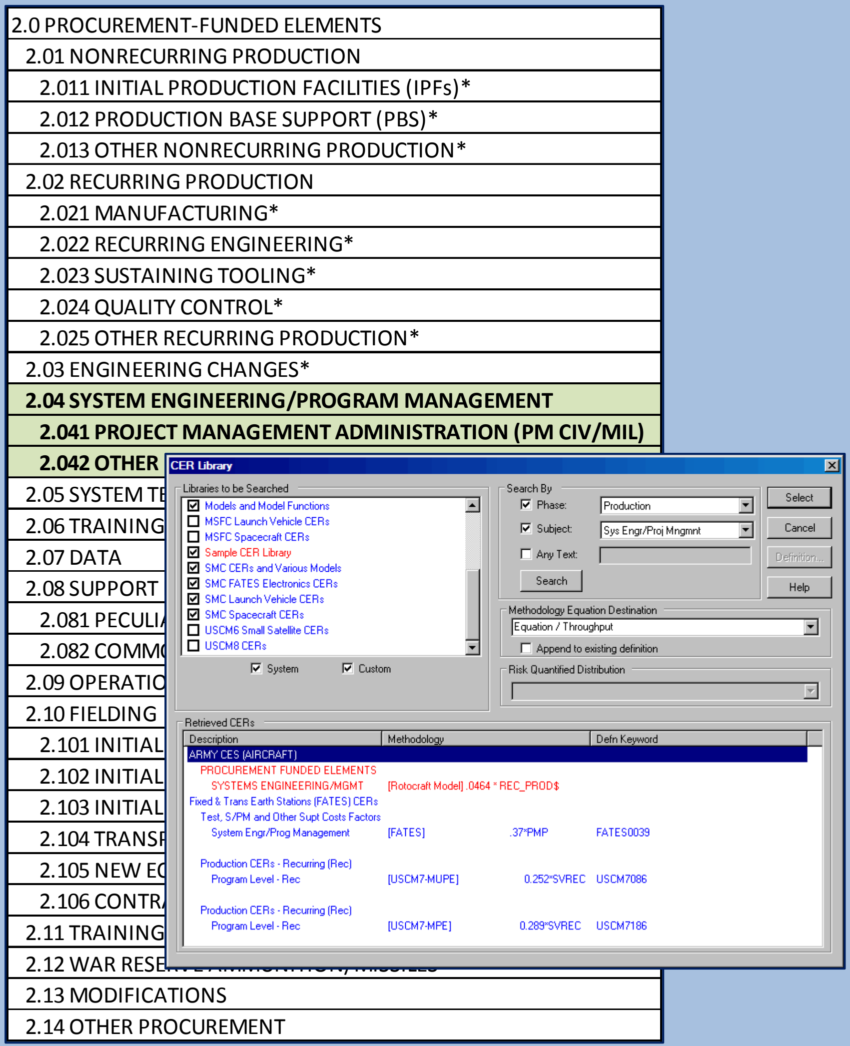
Distribution Statement A. Approved for public release; distribution is unlimited. This determination was made on March 31, 2010. U.S. Army PEO Aviation, Aviation Systems, SFAE-AV-AS, RSA, AL 35898.

How Do We Estimate SEPM?

- Historical data is available for analogous programs.
- We need to estimate Systems Engineering / Program Management (SEPM) for a new system.
- SEPM is not broken out in the historical data.
- We desire a good way to extract SEPM costs from the available historical data.

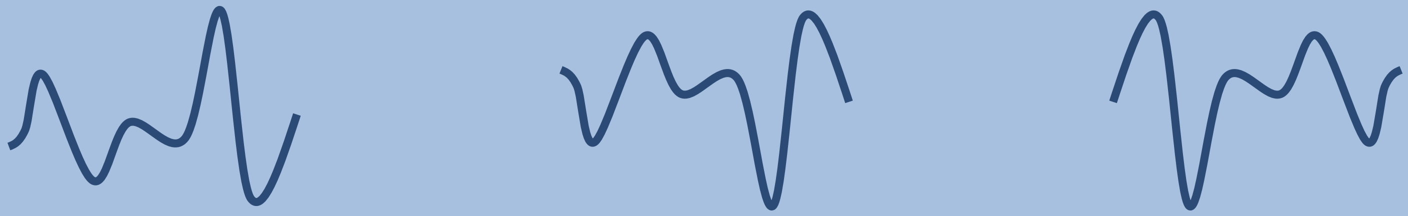
SEPM As A Factor

- In many cases, SEPM is estimated as a factor of another cost element.
  - 7-35% of Prime Mission Product
  - 4.6% of Recurring Production
- What if SEPM is a function and not just a factor?



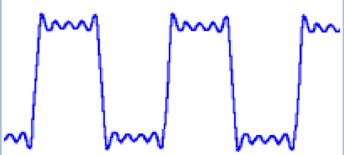
SEPM As A Function

- If a Program Office Estimate (POE) is a summation of functions, then SEPM might be a function.
- If SEPM is a function, it could be extracted from the data set using Fourier Transform (FT).
- If SEPM is a constant then it would only move the series up or down, it would not be a term identified by the transform.



What is a Fourier Series?

- The Fourier Series is used in Fourier Analysis of periodic functions (ex. current, voltage, harmonics).



A periodic function

- Fourier Analysis uses trigonometric functions for various waveforms and the summation of those functions.
- The sum of the trigonometric functions is the **Fourier Series.**

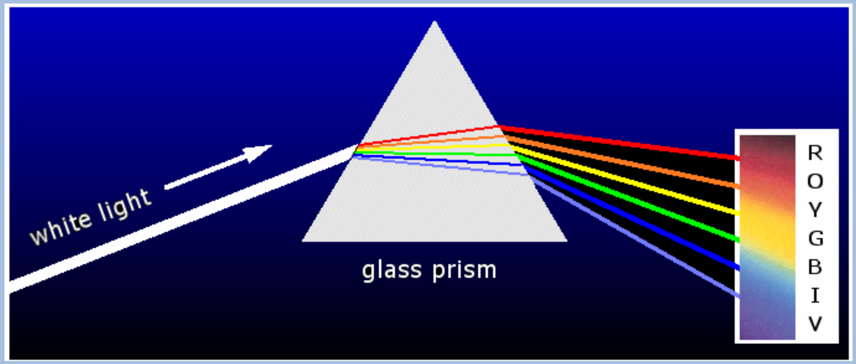
- The Fourier Series is named after Jean Baptiste Joseph Fourier (1768-1830).
- Fourier was a French mathematician as well as a friend of Napoleon and scientific adviser to Napoleon's army.

Jean Fourier

http://www.intmath.com/Fourier-series/Fourier-intro.php, 3/1/2010

Fourier Transform Description

- In Fourier Analysis, a Fourier Transform (FT) is the process that decomposes a function into basic pieces.
- A FT is like a prism that splits white light into its constituents.
- A FT takes a function (typically of time) and transforms it into a function (typically of frequency).
- In essence, a FT describes frequencies that are present in the original function.



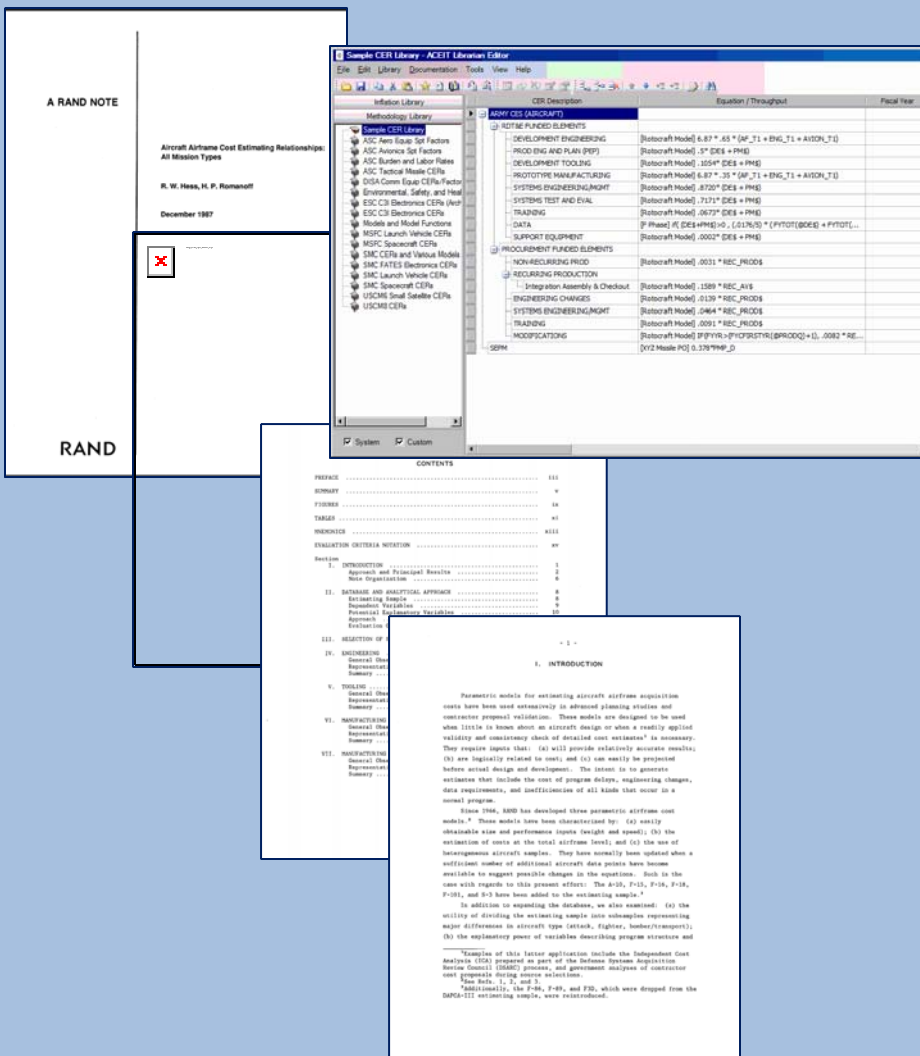
White light consists of all the ***frequencies*** of the color spectrum and the prism splits them apart and allows us to see the separate ***frequencies***.

http://www.intmath.com/Fourier-series/Fourier-intro.php (3/1/2010)
Beckwith, T. et al. *Mechanical Measurements, 6th ed*. Pearson Education, 2007. p. 121-124.

UPAS Program Office Estimate

- A short suspense cost estimate is needed for a new aircraft system, the Unmanned Paper Airplane System (UPAS).
- Specifically, leadership is interested in procurement phase costs.
- A cost estimate for key cost drivers of the procurement phase was created for the UPAS using cost estimating relationships (CER) for analogous systems.
- CERs were used to calculate non-recurring production, recurring production, systems engineering / program management, and fielding costs.

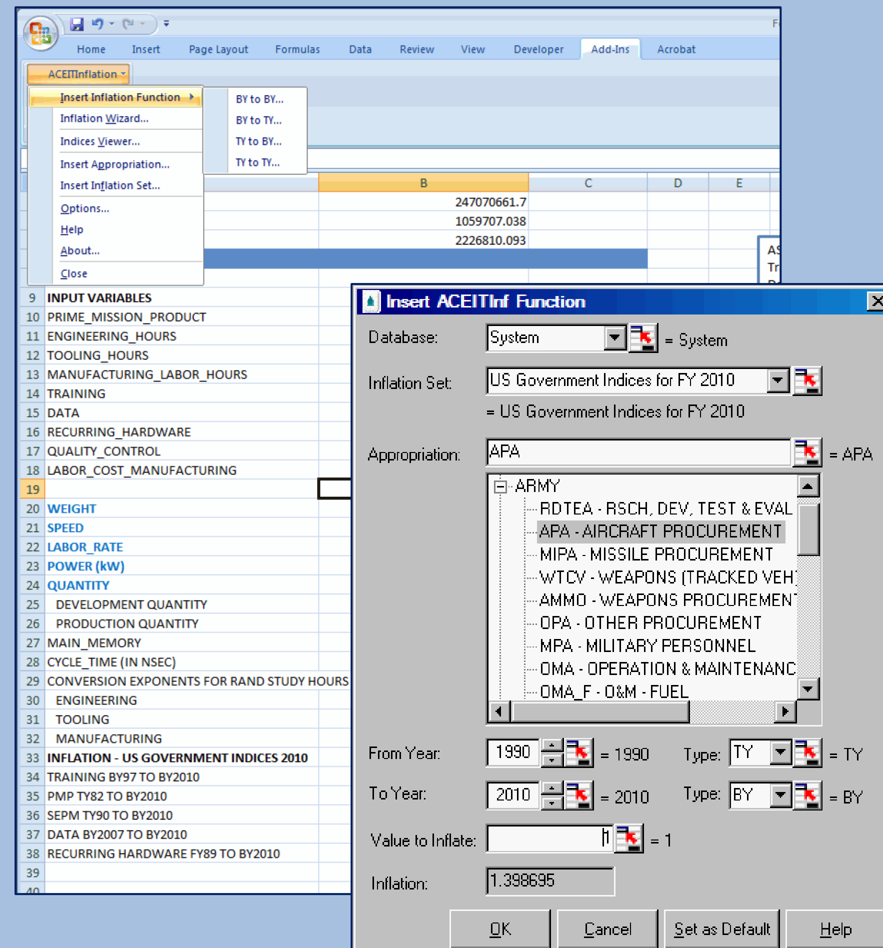
Cost Estimating Relationships



- Sources for CERs included
  - ACEIT Librarian
  - RAND study, “Aircraft Airframe Cost Estimating Relationships: All Mission Types” by R. W. Hess and H. P. Romanoff, December 1987.
- CERs were selected to be functions of system characteristics such as weight and power.

Normalizing the CERs

- CERs were normalized to Base Year 2010 (BY2010).
- OSD Inflation Indices for 2010 were applied.
- ACEIT Inflation Utility 7.2 was used to generate inflation factors.



Labor Hour Adjustment

- Engineering, tooling, and manufacturing labor hours were adjusted to reflect production only.
- The CERs included both development and production.
- A conversion equation was provided in the RAND study.

E36

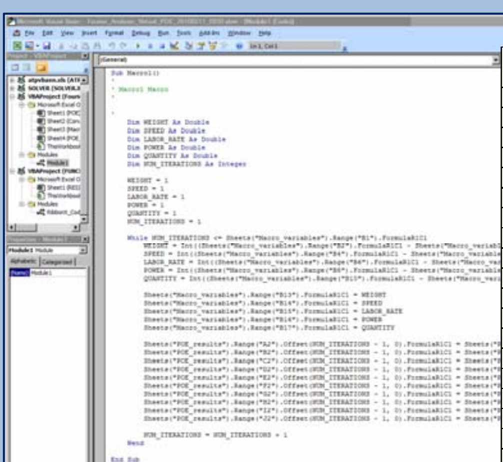
	A	B	C	E	F	G	I	J	K
2	Characteristic	Database Range		Characteristic	Database Range		Characteristic	Database Range	
3	Empty Weight (lb)	9,753	320,085	Empty Weight (lb)	9,753	320,085	Empty Weight (lb)	9,753	320,0
4	Maximum Speed (kn)	389	1,250	Maximum Speed (kn)	389	1,250	Maximum Speed (kn)	389	1,25
5	Number of Test AC	10	33	Number of Test AC	10	33	Number of Test AC	10	33
6									
7	Conversion Exponents			Conversion Exponents			Conversion Exponents		
8	**Engineering**	0.163		**Tooling**	0.263		**Mnf. Labor Hours**	0.641	
9									
10									
11									
12	Maximum Speed (SP)	600		Maximum Speed (SP)	600		Maximum Speed (SP)	600	
13	Empty Weight (EW)	10000		Empty Weight (EW)	10000		Empty Weight (EW)	10000	
14	Production Qty	116.1		Production Qty	116.1		Production Qty	116.1	
15	Development Qty	12.9		Development Qty	12.9		Development Qty	12.9	
16	Total Qty	129		Total Qty	129		Total Qty	129	
17									
18	Engineering			Engineering			Engineering		
19	*Hours			*Hours			*Hours		
20	ENGR(100) in thousands	4022.577708		TOOL(100) in thousands	2212.021418		LABR(100) in thousands	5940.7916	
21	ENGR(Cum) in thousands	4193.055061		TOOL(Cum) in thousands	2365.235678		LABR(Cum) in thousands	6994.107806	
22	ENGR(Dev) in thousands	2880.915799		TOOL(Dev) in thousands	1290.845965		LABR(Dev) in thousands	1598.572443	
23	ENGR(Prod) in thousands	1312.139261		TOOL(Prod) in thousands	1074.389714		LABR(Prod) in thousands	5395.535362	
24	*Dollars			*Dollars			*Dollars		
25	Engineering Labor Rate (1986$)	$59.10		Tooling Labor Rate (1986$)	$60.70		Tooling Labor Rate (1986$)	$50.10	
26	Engineering Cost ($K-1986)	77547.43034		Tooling Cost ($K-1986)	$65,215.46		Tooling Cost ($K-1986)	$270,316.32	
27									

B39

	A	B	C	E
2	Characteristic	Database Range		Characte
3	Empty Weight (lb)	9753	320085	Empty Weight (lb)
4	Maximum Speed (kn)	389	1250	Maximum Speed (k
5	Number of Test AC	10	33	Number of Test AC
6				
7	Conversion Exponents			Conversion Expone
8	**Engineering**	0.163		**Tooling**
9				
10				
11				
12	Maximum Speed (SP)	600		Maximum Speed (S
13	Empty Weight (EW)	10000		Empty Weight (EW
14	Production Qty	=POE!$B$26		Production Qty
15	Development Qty	=POE!$B$25		Development Qty
16	Total Qty	=POE!$B$24		Total Qty
17				
18	Engineering			Engineering
19	*Hours			*Hours
20	ENGR(100) in thousands	=0.0103*(B13^0.777)*(B12^0.894)		TOOL(100) in thous
21	ENGR(Cum) in thousands	=B20*((B16/100)^B8)		TOOL(Cum) in thou
22	ENGR(Dev) in thousands	=B20*((B15/100)^B8)		TOOL(Dev) in thous
23	ENGR(Prod) in thousands	=B21-B22		TOOL(Prod) in thou
24	*Dollars			*Dollars
25	Engineering Labor Rate (1986$)	59.1		Tooling Labor Rate
26	Engineering Cost ($K-1986)	=B23*B25		Tooling Cost ($K-19
27				

Data Collection Tool

- A Visual Basic program was written in Microsoft (MS) Excel to generate a data set for Fourier Analysis.
- A random value was selected within each variable range of the UPAS POE.
- The estimate was calculated.
- Results were saved in a MS Excel worksheet.
- The process continued for a specified number of iterations.



PRODUCTION TOTAL	NONRECURRING	RECURRING	SEPM	FIELDING	WEIGHT	SPEED	LABOR_RATE	POWER (kW)	QUANTITY
373048427.1	150494079	219303257.6	1047675.711	2203414.859	228707	848	87		
671405772.3	190102321.4	477817555.1	1126938.819	2358956.96	250180	401	93		
612412531	172994443	438475675	166718.8605	775694.1104	23827	745	96		
204583977.9	123670956	76794630.76	1335377.508	2783013.613	308278	1140	69		
334468715.5	129847834.4	201986396.3	833598.8767	1800885.932	172636	1050	69		
438165231.6	151927669	284387322.3	544763.1602	1305477.072	102283	925	89		
540724952.8	187071795.2	349980846.4	1189101.529	2483209.668	267267	1099	87		
682352142.9	192902484.5	487850026	446862.2633	1152770.185	80157	988	100		
656265351	176249370.6	478847915.9	268169.6088	899894.8999	42763	1250	90		
296535159.4	122795600.1	172594550.2	258131.6612	886877.4048	40802	477	94		
677352849.6	168249943.1	507261030.3	541549.2463	1300326.955	101541	718	77		
500096107.7	146447762.8	351439162.4	679765.8952	1529416.627	134312	628	73		
632964072.8	181656768.4	449068026.8	690850.6644	1548426.899	137012	744	91		
572393442.5	163000279.4	407861877.4	419523.1662	1111762.413	74166	549	87		
404017467.6	159827024	240261620.5	1273620.501	2655202.536	290830	614	93		
229953128.2	126276489	99701765.13	1288676.086	2686197.94	295066	933	88		
229834095.2	121388384.1	105687308.1	877390.8332	1881012.213	183862	987	98		
426956509.6	156288669	267969972.8	856050.9057	1841816.884	178375	1178	82		
344947154.8	133924628.7	208426222.8	820018.9376	1776284.396	169182	788	79		
651729132.2	181622352.5	469125931.9	184575.9647	796271.8563	27005	599	100		
327463941.2	122869312.4	202511530.7	632954.275	1450143.809	123023	811	73		
423085045.1	146922868.8	273172485.5	958048.6069	2031642.152	204875	856	73		

A37

	A	B	C	D
1	NUM_ITERATIONS	20000		
2	WEIGHT MAX	320085		
3	WEIGHT MIN	9753		
4	SPEED MAX	1250		
5	SPEED MIN	389		
6	LABOR_RATE MAX	100		
7	LABOR_RATE MIN	68		
8	POWER MAX	360		
9	POWER MIN	3.7		
10	QUANTITY MAX	200		
11	QUANTITY MIN	10		
12				
13	WEIGHT	231943		
14	SPEED	1228		
15	LABOR_RATE	81		
16	POWER (kW)	242		
17	QUANTITY	129		

Fourier Series - Definition

$$f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{\infty} a_n \cos(n x) + \sum_{n=1}^{\infty} b_n \sin(n x).$$

- General definition of a Fourier Series
- Using sine and cosine functions accounts for phase shift

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)\, dx$$

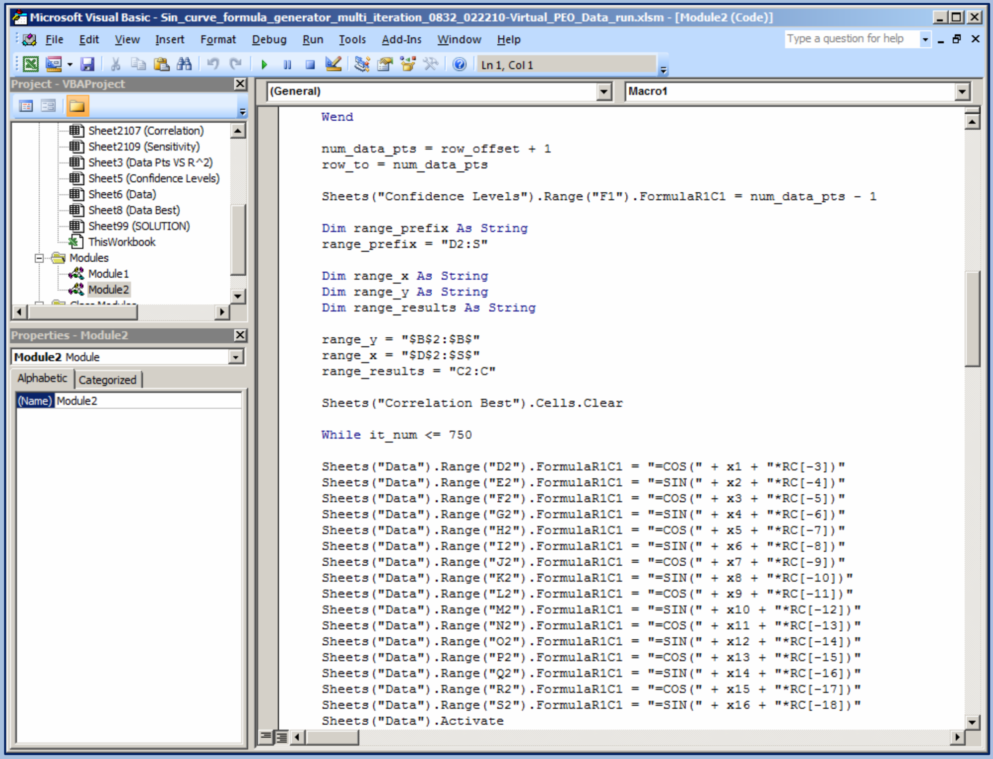
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(n x)\, dx$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(n x)\, dx$$

- Formal definition of coefficients

Fourier Analysis Program (FAP)

- The objective is to produce a Fourier series that will describe cost data.
  - Amplitude becomes cost
  - Time becomes a variable based on a system specification
- The code to perform the Fourier Analysis was written in Visual Basic.



Regression Analysis

- A table was created containing
  - User defined system variable
  - Cost
  - 8 sine functions
  - 8 cosine functions
- Regression analysis was performed on the virtual POE data to determine the coefficients in the Fourier Series.

SUMMARY OUTPUT

=Regression!$B$17+Regression!$B$18*COS(0.8*Data!A2)+Regression!$B$19*SIN(0.3*Data!A2)+Regression
COS(1.9*Data!A2)+Regression!$B$25*SIN(5.9*Data!A2)+Regression!$B$26*COS(0.7*Data!A2)+Regression!
SIN(6*Data!A2)+Regression!$B$32*COS(1.7*Data!A2)+Regression!$B$33*SIN(0.5*Data!A2)

ariables / (10T)	PRODUCTION TOTAL	Results	Coef 1	Coef 2	Coef 3	Coef 4	Coef 5	Coef 6	Coef 7	Coe
0.059870625	193338350.6	313634293.2	0.999122	0.219714	0.972864	0.208017	0.986963	0.260394	0.986963	0.
0.339695895	651728884.5	375053862.4	0.971862	0.95113	0.243509	0.927973	0.608062	0.997103	0.608062	0.1
0.345771734	296535035.7	377615350.7	0.970851	0.957831	0.220461	0.935687	0.594957	0.99878	0.594957	0.1
0.515240696	656265159	445395180.7	0.935661	0.944215	-0.42471	0.973083	0.178682	0.767244	0.178682	0.1
0.732664472	371010044.7	477261681.7	0.871342	0.417538	-0.95989	0.545736	-0.39622	-0.08204	-0.39622	0.2
0.928700055	308313536.4	449191618.3	0.796029	-0.29035	-0.88684	-0.10864	-0.8056	-0.81031	-0.8056	0.
1.043905572	356510172.8	431080778.9	0.744685	-0.66003	-0.59812	-0.48999	-0.94827	-0.9929	-0.94827	0.3
1.09955806	572393111.6	427160737	0.718126	-0.79969	-0.41151	-0.64945	-0.98511	-0.99211	-0.98511	0.3
1.337802486	317125508.6	462439068.3	0.592642	-0.97193	0.483843	-0.99955	-0.89135	-0.38652	-0.89135	0.3

	Coefficient	Standard Err	t Stat	P-value	Lower 95%	Upper 95%
Intercept	4.66E+08	29944399	15.57663	4.82E-17	4.06E+08	5.27E+08
Coef 1	42874586	36298687	1.181161	0.245735	-3.1E+07	1.17E+08
Coef 2	-6E+07	46567934	-1.29046	0.2056	-1.5E+08	34543590
Coef 3	-1E+08	37015884	-2.71476	0.010343	-1.8E+08	-2.5E+07
Coef 4	-1E+08	39849586	-2.54411	0.015668	-1.8E+08	-2E+07
Coef 5	-3272399	37615936	-0.087	0.931186	-8E+07	73172380
Coef 6	-1.5E+07	37123650	-0.40838	0.685554	-9.1E+07	60283704
Coef 7	-8.1E+07	40166435	-2.0262	0.050648	-1.6E+08	242956.6
Coef 8	78740447	47506594	1.657464	0.106626	-1.8E+07	1.75E+08
Coef 9	24952405	35937392	0.69433	0.492193	-4.8E+07	97985972
Coef 10	28932133	45414756	0.637065	0.528349	-6.3E+07	1.21E+08
Coef 11	0	0	65535	#NUM!	0	0
Coef 12	-3E+07	40096083	-0.75322	0.4565	-1.1E+08	51283728
Coef 13	34618769	43863315	0.789242	0.43544	-5.5E+07	1.24E+08
Coef 14	-6.8E+07	44676459	-1.53214	0.134742	-1.6E+08	22343045
Coef 15	-6E+07	47756510	-1.26275	0.215272	-1.6E+08	36748472
Coef 16	-1.5E+07	44216926	-0.34242	0.734144	-1.1E+08	74718910

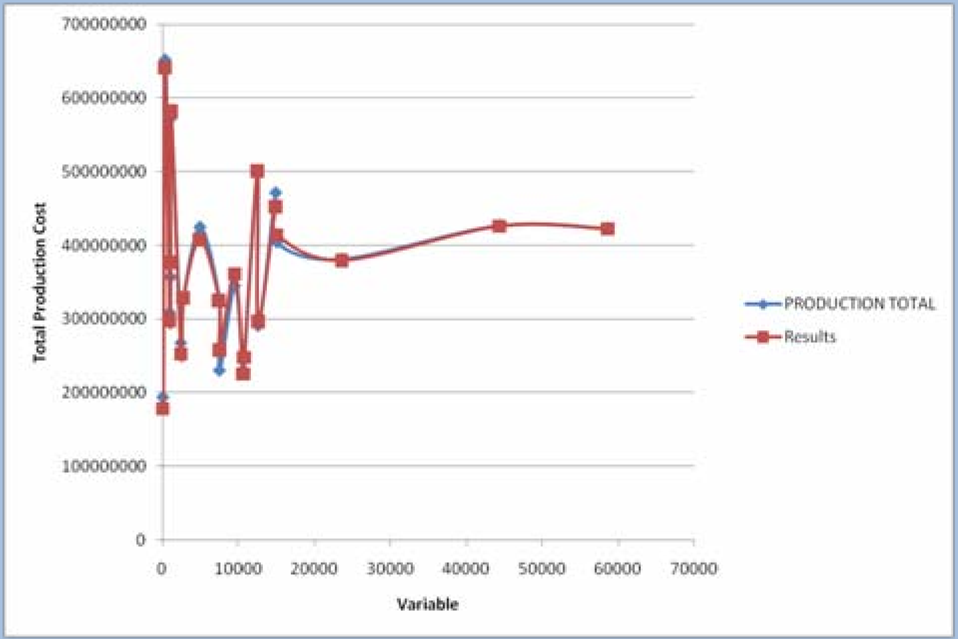
FAP – Correlation

- The Regression tool outputs y-intercept and a coefficient for each of the 16 terms
- Combining these yields a Fourier series of the form: y= Int + coefficient1*cos(constant1*var) + … coefficient16*sin(constant16*var)
- This Fourier series is applied to each variable in the data set
- The Correlation tool in the Data Analysis add-on is used to find the R^2 value between the Fourier series and the data set

	A	B	C
1		Column 1	Column 2
2	Column 1	1	
3	Column 2	0.800627	1

FAP – Iteration

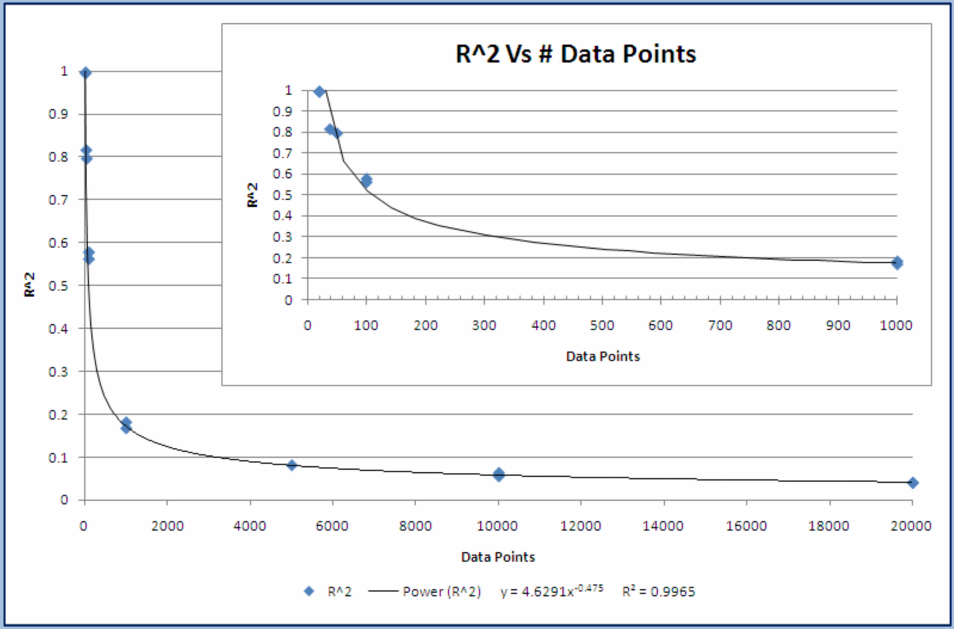
- The previous steps are run for a number of iterations.



- With each iteration
  - The constants of each sine and cosine term are randomly varied
  - Regression and correlation are evaluated
  - When the R-squared value is better than the previous best, the data table and regression results are saved.
- The best Fourier series produced can be used to predict the desired cost using the user designated variable.

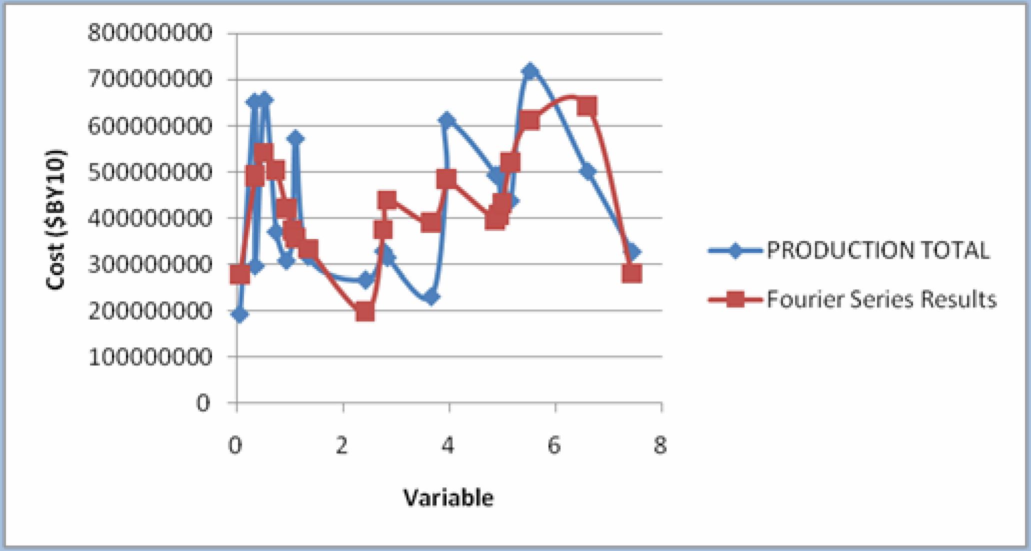
PM UPAS Production Cost Run

- The data collection tool was used to produce two seeds of 20,000 data points.
- The FAP was run on various numbers of data points to test the quality of the Fourier series that can be produced for different size data sets.
- Results show that smaller data sets, which are typically available in estimating the cost of aircraft, produce better R-squared values.



PM UPAS Run - Statistics

- A Fourier analysis was performed on a set of 50 data points from the first seed of the PM UPAS data.
- The correlation between the Fourier series and the data set generated an R-squared value of 0.801.
- The second seed of 20,000 data points was tested using the Fourier series .
- The percent difference between the Fourier series results and the actual cost was recorded.

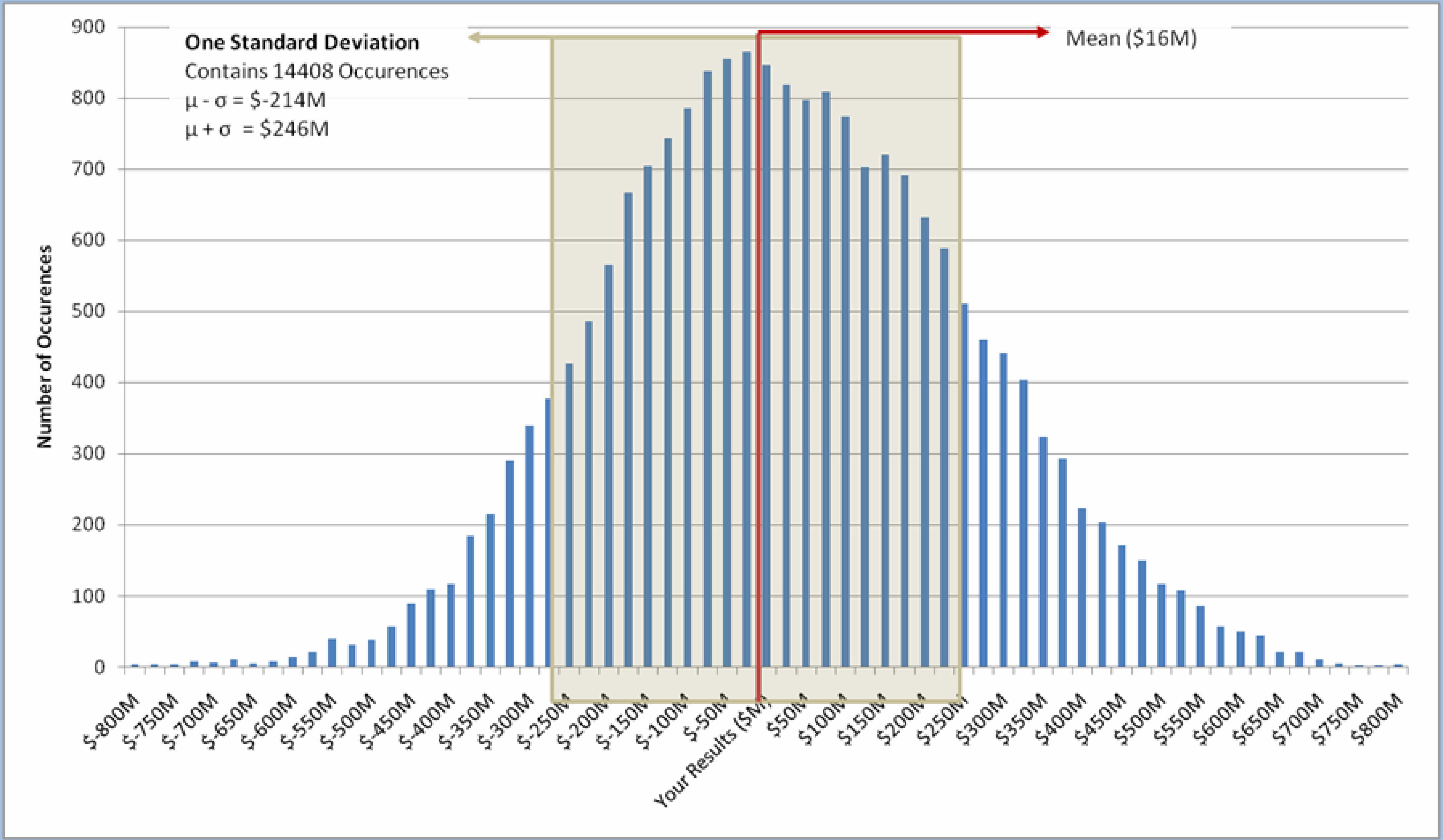


Fourier Analysis – Statistics

- A statistical analysis was performed on the 20,000 data points to determine the quality of the costs predicted by the Fourier analysis.
- The distribution of the differences between the actual and predicted costs is normal.

	A	B
1	*Virtual POE Statistics*	
2		
3	Mean	16,052,339
4	Standard Error	1,625,793
5	Median	8,366,544
6	Mode	#N/A
7	Standard Deviation	229,921,886
8	Sample Variance	52,864,073,472,204,200
9	Kurtosis	-0.087
10	Skewness	0.082
11	Range	1,692,729,663
12	Minimum	-893,370,769
13	Maximum	799,358,894
14	Sum	321,046,787,008
15	Count	20,000

Fourier Analysis – Statistics

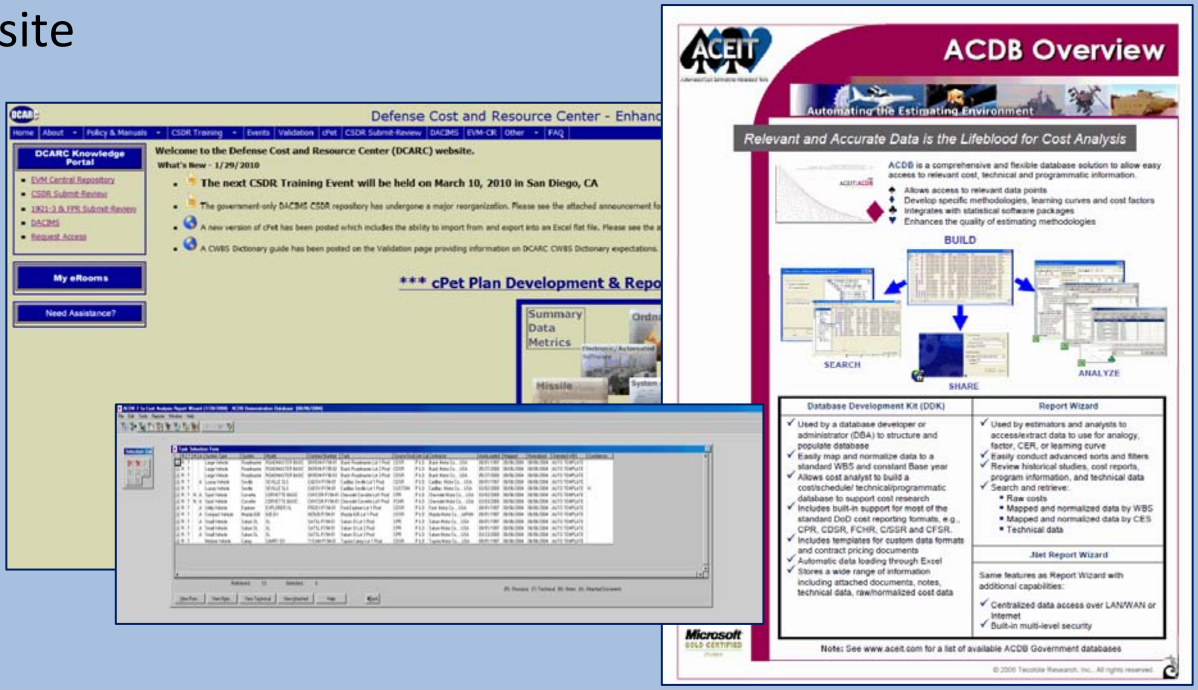


Testing the Method - Actual Data

- Fourier Analysis produced good results from the UPAS Virtual POE data.
- Will the results be similar with actual data?
- The Fourier Analysis Method was tested on actual aircraft data.

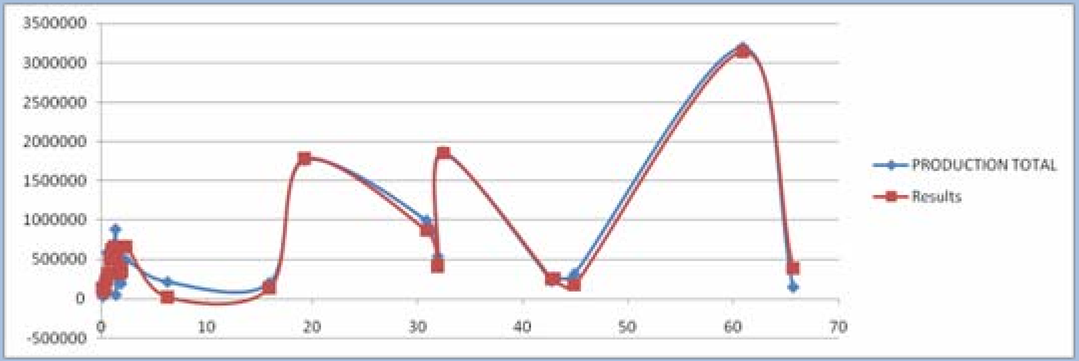
Data Sources for Actual Data

- Data for the actual cost of aircraft systems was gathered.
  - ACEIT Automated Cost Database (ACDB) Report Wizard
  - Defense Automated Cost Information Management System (DACIMS) Library located on the Defense Cost and Resource Center (DCARC) website



Fourier Analysis - Actual Data

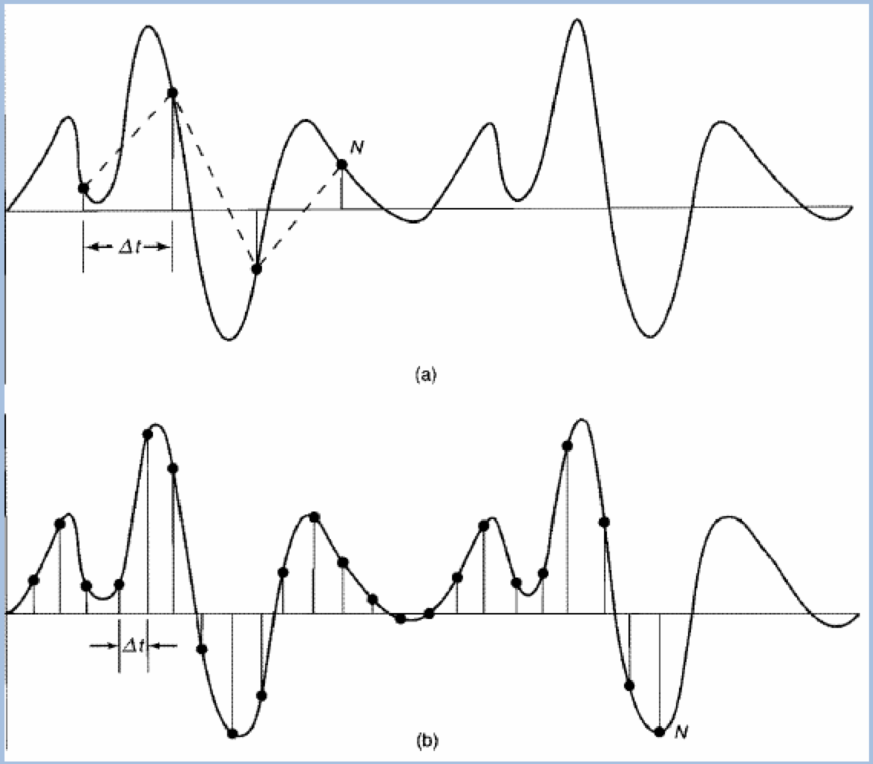
- Data was gathered for 20 rotary wing aircraft and 17 fixed wing aircraft.
  - Total Production Cost
  - Maximum Takeoff Weight
  - Maximum Cruise Speed
  - Range
  - Service Ceiling
  - Lot Quantity



- Additionally, data on 3 rotary wing aircraft and 3 fixed wing aircraft was gathered to test the Fourier Series.
- All costs were normalized to BY2010 dollars using OSD Inflation Indices and the ACEIT 7.2 Inflation Utility.
- The Fourier Series produces negative costs in regions with large spacing between the data points.

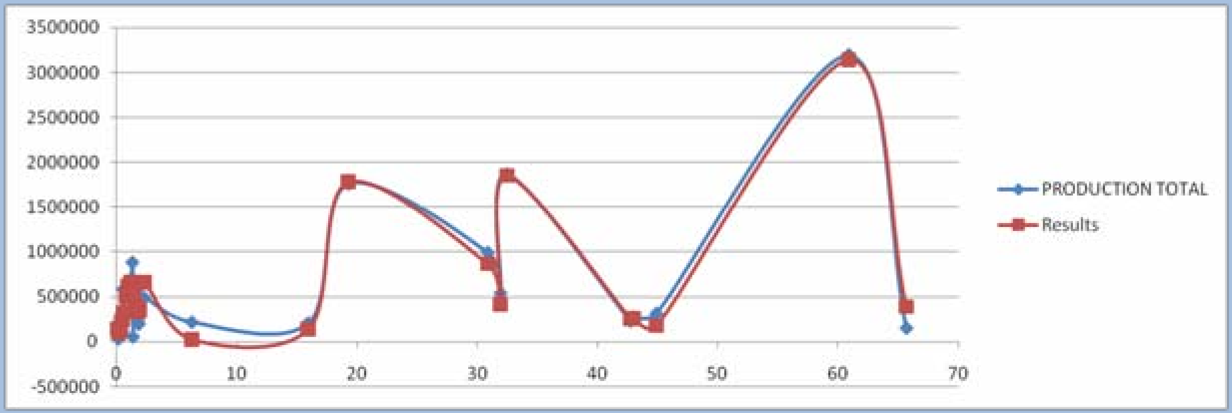
Sampling Rate

- The function behaves freely in regions with low resolution.
- Low resolution is equivalent to a low sampling rate in signal analysis.
- Sampling below the minimum sampling rate does not reveal the true behavior of the function.



Results - Actual Data

- The Fourier series produced from the actual data gave a realistic estimate of the production cost for
  - Systems similar to the systems in the data set
  - Systems with characteristics within the range of the data set (Ex. System weight must be between the lowest and highest weights in the data set).

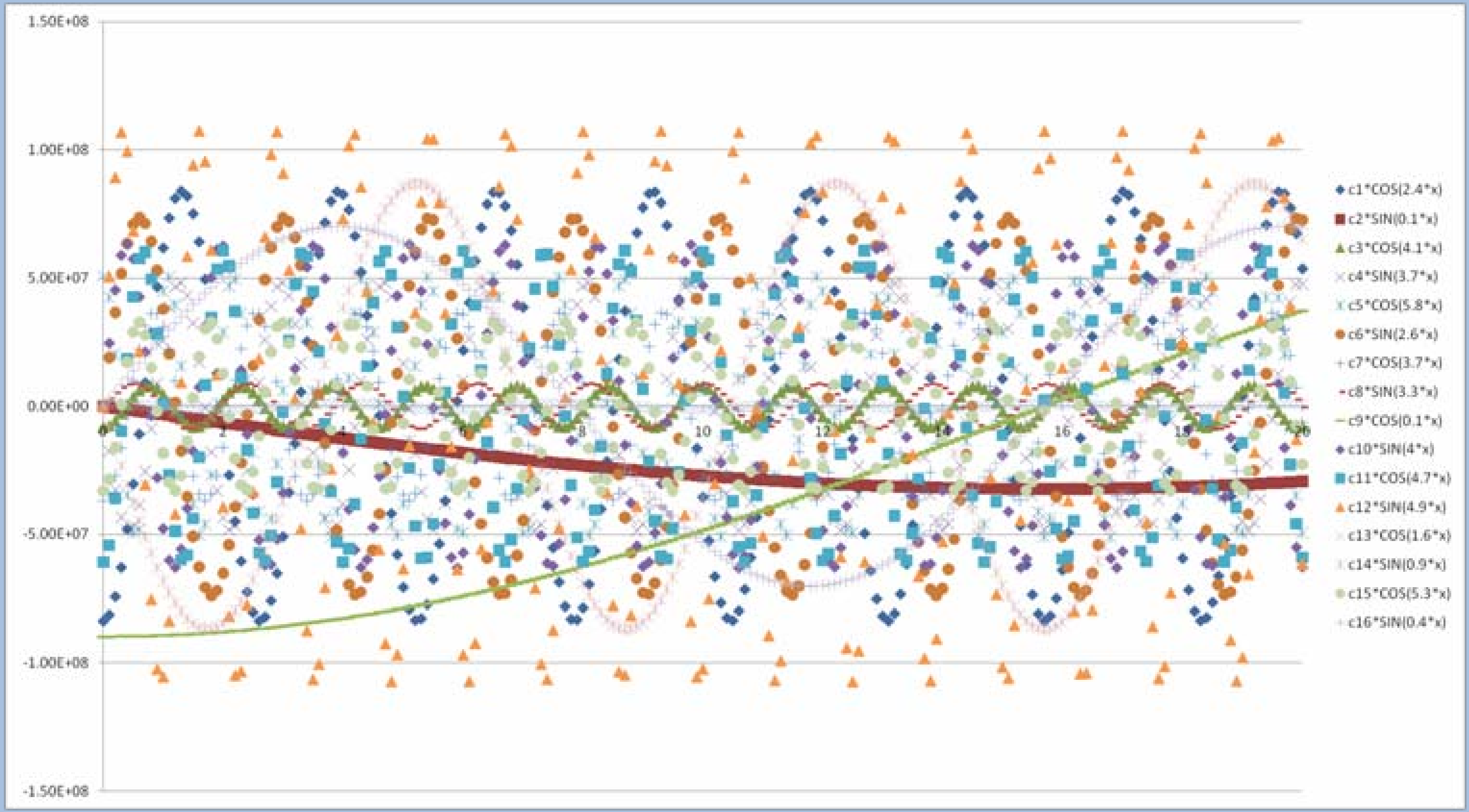


Conclusions

- Fourier analysis is a viable method for estimating a total cost when historical data is available.
- The analysis tool is useful to cost analysts because
  - It makes use of a common environment (MS Excel)
  - Produces quick and realistic results.
- This method is bound by the traditional rules for developing CERs from a data set.
- The Fourier series produced poor estimates for
  - Unlike systems
  - Systems whose characteristics were not in the range of the data set.

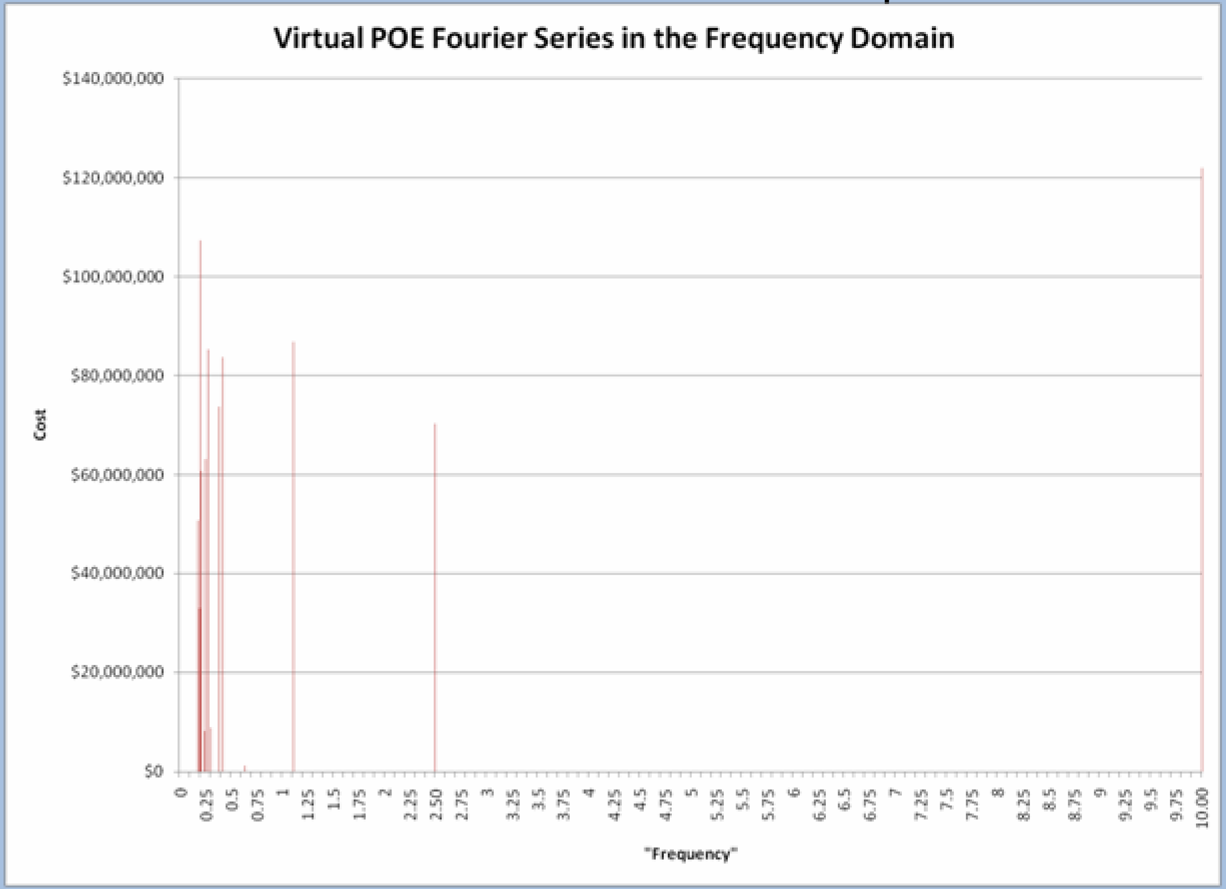
Where to next

- Extract the cost elements that make up the total cost as predicted by the Fourier series (Ex. SEPM from total production cost)

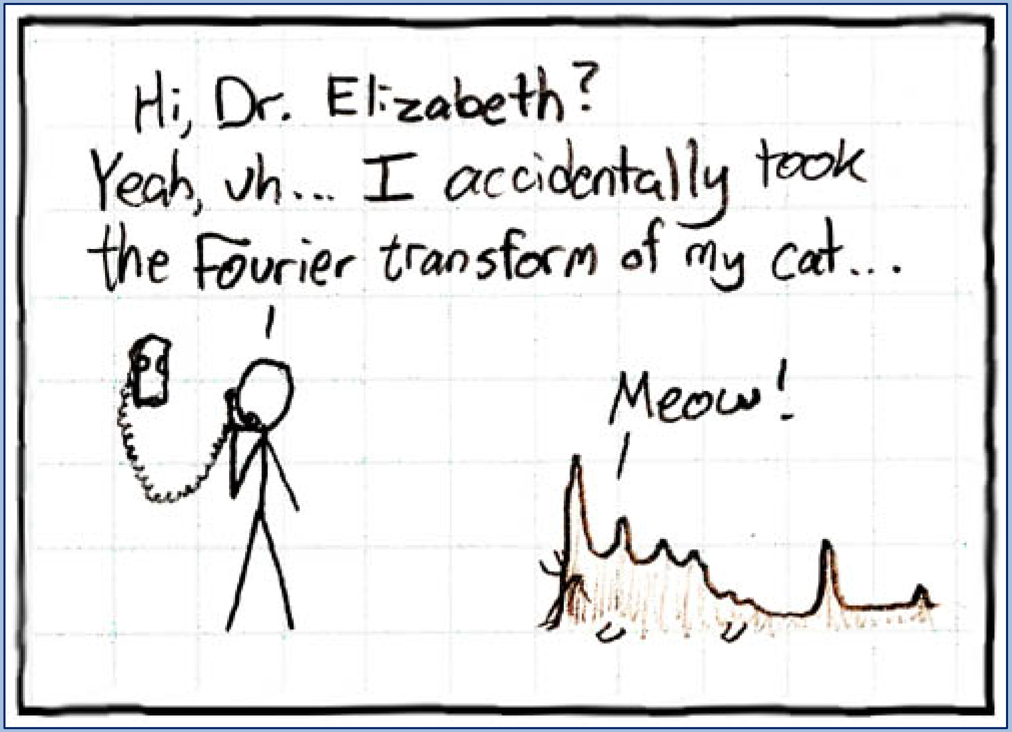


Where to next

- Transform the function from the "time" to the "frequency" domain and analyze the "frequencies"
- Do they relate to the cost elements that make up the total cost?



QUESTIONS?



xkcd comics – www.xkcd.com/26/