ABSTRACT:

The generation of Non Conformances (NCs) is an unfortunate fact of life at any pharmaceutical manufacturing facility. Significant resources are applied to document, quantify and (hopefully) resolve the cause(s) of each and every NC to assure the quality of the finished product. The elusive goal of continuous improvement strives to reduce the rate of NC generation. The early identification of any trends to the contrary will hopefully "nip in the bud" any systematic deficiency in procedures that could lead to a significant increase in NC generation. On a more positive note, early identification of NC generation reduction will provide incentive to sustain the continuous improvement effort. This paper outlines methodology to measure both positive and negative trends in NC generation.

Process

NCs occur at what appears to be random intervals throughout the year. Investigations of each NC results in classification in several categories (e.g.):

- Building
- Area
- Functional Group
- Root Cause
- ….

Any subset of NCs generated over a time period will provide performance measurement on that subset, allowing for the uncovering negative trend which then can be further investigated with the goal of identifying and eliminating common factors to reduce future NC generation. Valid subsets of NCs are comprised of NCs sharing one or more like categories as illustrated above.

The trend is identified as follows:

Let "N" be the number of NCs sharing like categories. Each NC has occurred at a particular calendar day "D". Therefore we have an ordered set of dates:

$$D_1 , D_2 , D_3 , \ldots , D_{n-1} , D_n , D_{n+1} , \ldots , D_{N-1} , D_N \qquad (1)$$

The potential exists for the subset to be defined with such granularity that the value of N will be too small to allow the application of this (or any) methodology. Also, the time span over which the subset is assembled may be defined to be too long, such that equipment and process changes may be so significant during the time span that comparisons of NCs at the beginning and the end of the time span is not valid. It is proposed that the time span for the standard analysis be set at 1 year and the minimum sample size is 12 NCs. This can be stated as follows:

NC generation trends can only be reliably measured if the subset of NCs under analysis has an average generation rate of one a month.

The time (t) between each NC is calculated. Note that if two or more NCs occur on the same day, the time would be zero. Note the special handling of the case n = 1:

$$t_1 = D_2 - D_1 \quad (2)$$

$$t_n = D_n - D_{n-1} \qquad 1 < n \le N \quad (3)$$

Decaying Memory Averaging

We now define an averaging scheme that weights more recent data heavier than older data, and the weighting factor is proportional to the sample size.

$$\overline{T}_n = \frac{t_n + (a-1)\overline{T}_{n-1}}{a} \quad (4)$$

Here
a = weighting factor
$\overline{T}_n$ = decaying memory average time at sample n

The proper choice of the weighting factor is critical to a successful analysis. Too small a number and the noise of the resultant graph would be too great. Too large a number, and time dependent effects are not visible. For this case, the weighting factor is defined to be 25% of the total sample size. In the case of steady-state generation of NCs, this would represent the number of NCs generated in a quarter of a year.

$$a = \text{Greatest integer in } \frac{N}{4} \quad (5)$$

Note that application of our lower bound of

$$N \ge 12 \quad (6)$$

yields

$$a \geq 3 \tag{6}$$

A closer inspection of equation (4) reveals that it is impossible to calculate $\overline{T}_n$ for the case of n = 1. Experience dictates that to achieve a reasonable trend, one defines the following startup condition:

$$\overline{T}_n = \frac{\sum_{n=1}^{a} t_n}{a} \tag{7}$$

Or simply put, the first 25% of the samples use the simple arithmetic average of the subset to determine the baseline average against which to compare recent performance.

Frequency

The calculations above calculate the average time interval between NC generation. Note that as the manufacturing operation improves, one would expect to see a reduction in the number of NCs being generated, thus the value for $\overline{T}_n$ would be expected to increase. The accepted convention when performing metrics on NC performance is "the lower the number, the better the performance". In order to meet this convention, the inverse of $\overline{T}_n$, the average frequency of generation , is considered. Also note that the unit of measure of t_n and $\overline{T}_n$ is days, or more specifically days between NC generation. For frequency, it is preferred to use units of number of NCs per month. A 30 day month is assumed.

$$\overline{\upsilon}_n = \frac{30}{\overline{T}_n} \tag{8}$$

Normalization

Each subset of the data will have a different sample size N. In order analyze all subsets in a uniform fashion, each data set is normalized using the initial NC generation frequency:

$$\overline{Y}_n = \frac{\overline{\upsilon}_n}{\overline{\upsilon}_1} \tag{9}$$

Intervention Rules

In an ideal environment, the rate of generation of NCs should be decreasing. This is true when:

$$\overline{Y}_n < 1 \tag{10}$$

Steady state NC generation exists when

$$\overline{Y}_n = 1 \tag{11}$$

Further investigation and possible intervention is warranted when

$$\overline{Y}_n > 1 \tag{12}$$

The magnitude of $\overline{Y}_n$ and other circumstances surrounding the subset will assist the responsible person to determine the course of action.

Example 1

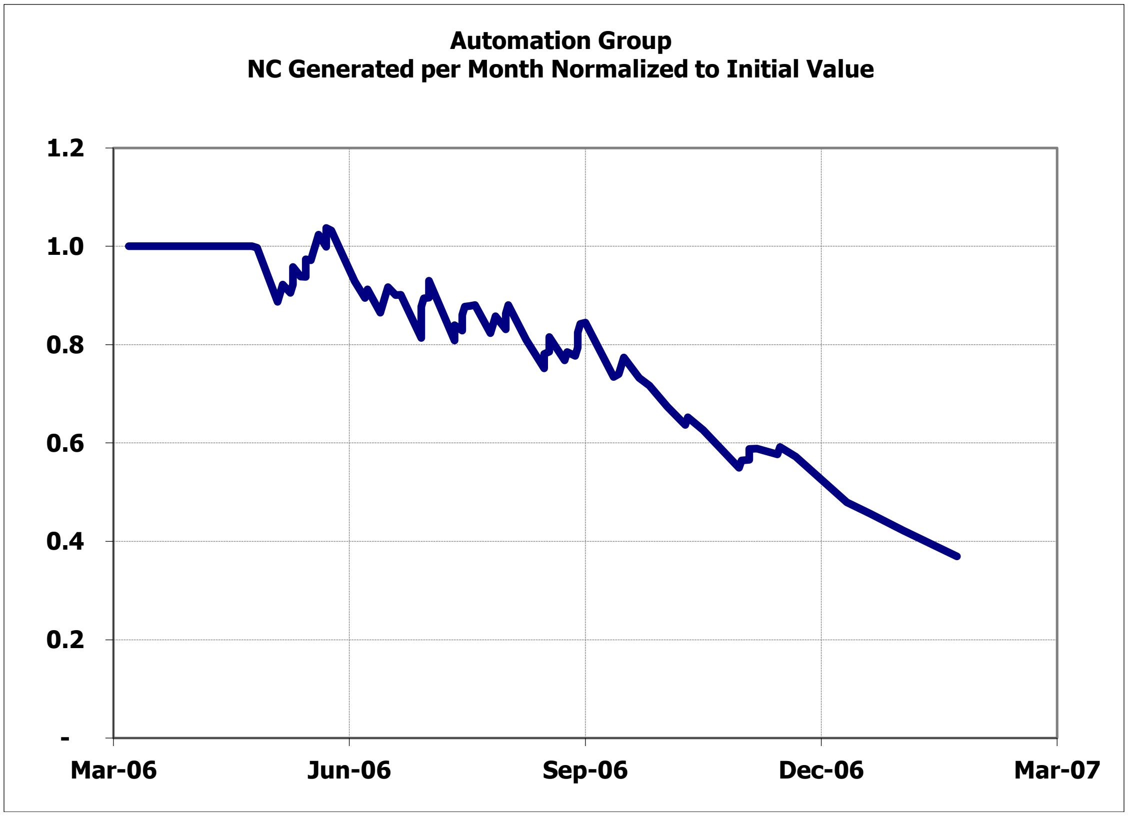
The first example is for a subset of data that is behaving in an ideal fashion. The trend is for fewer NCs generated as time passes.

The Automation group is part of the Facilities and Engineering Department at a major pharmaceutical facility on the east coast of the United States. In the time period from 3/7/2006 to 3/6/2007, 108 NCs were assigned to this group. For this data set, we know:

$$N = 108 \tag{13}$$

$$a = 27 \tag{14}$$

The results:


Automation Group
NC Generated per Month Normalized to Initial Value
1.2
1.0
0.8
0.6
0.4
0.2
-
Mar-06
Jun-06
Sep-06
Dec-06
Mar-07

The complete data with calculations:

N (Sample Number)	Date of NC	Days Between NCs	Decaying Memory Average Days Between NCs	Decaying Memory Average NCs Generated per Month	NC Generated per Month Normalized to Initial Value	
1	03/07/2006	0	1.8	16.30	1.0	
2	03/07/2006	0	1.8	16.30	1.0	
3	03/08/2006	1	1.8	16.30	1.0	
4	03/08/2006	0	1.8	16.30	1.0	
5	03/10/2006	2	1.8	16.30	1.0	
6	03/15/2006	5	1.8	16.30	1.0	
7	03/17/2006	2	1.8	16.30	1.0	
8	03/21/2006	4	1.8	16.30	1.0	
9	03/22/2006	1	1.8	16.30	1.0	
10	03/23/2006	1	1.8	16.30	1.0	
11	03/23/2006	0	1.8	16.30	1.0	
12	03/26/2006	3	1.8	16.30	1.0	
13	03/29/2006	3	1.8	16.30	1.0	
14	03/30/2006	1	1.8	16.30	1.0	Linear Average Region
15	03/31/2006	1	1.8	16.30	1.0	
16	04/03/2006	3	1.8	16.30	1.0	
17	04/06/2006	3	1.8	16.30	1.0	
18	04/11/2006	5	1.8	16.30	1.0	
19	04/11/2006	0	1.8	16.30	1.0	
20	04/12/2006	1	1.8	16.30	1.0	
21	04/14/2006	2	1.8	16.30	1.0	
22	04/15/2006	1	1.8	16.30	1.0	
23	04/20/2006	5	1.8	16.30	1.0	
24	04/22/2006	2	1.8	16.30	1.0	
25	04/22/2006	0	1.8	16.30	1.0	
26	04/22/2006	0	1.8	16.30	1.0	
27	04/24/2006	2	1.8	16.30	1.0	
28	04/26/2006	2	1.8	16.25	1.0	
29	05/04/2006	8	2.1	14.47	0.9	
30	05/05/2006	1	2.0	14.75	0.9	
31	05/06/2006	1	2.0	15.03	0.9	
32	05/09/2006	3	2.0	14.76	0.9	
33	05/10/2006	1	2.0	15.04	0.9	
34	05/10/2006	0	1.9	15.62	1.0	
35	05/13/2006	3	2.0	15.30	0.9	
36	05/15/2006	2	2.0	15.29	0.9	
37	05/15/2006	0	1.9	15.88	1.0	
38	05/17/2006	2	1.9	15.84	1.0	
39	05/18/2006	1	1.9	16.12	1.0	
40	05/19/2006	1	1.8	16.41	1.0	
41	05/20/2006	1	1.8	16.69	1.0	
42	05/23/2006	3	1.8	16.28	1.0	
43	05/23/2006	0	1.8	16.91	1.0	
44	05/25/2006	2	1.8	16.83	1.0	
45	05/29/2006	4	1.9	16.09	1.0	
46	06/03/2006	5	2.0	15.14	0.9	
47	06/07/2006	4	2.1	14.59	0.9	

Example 2

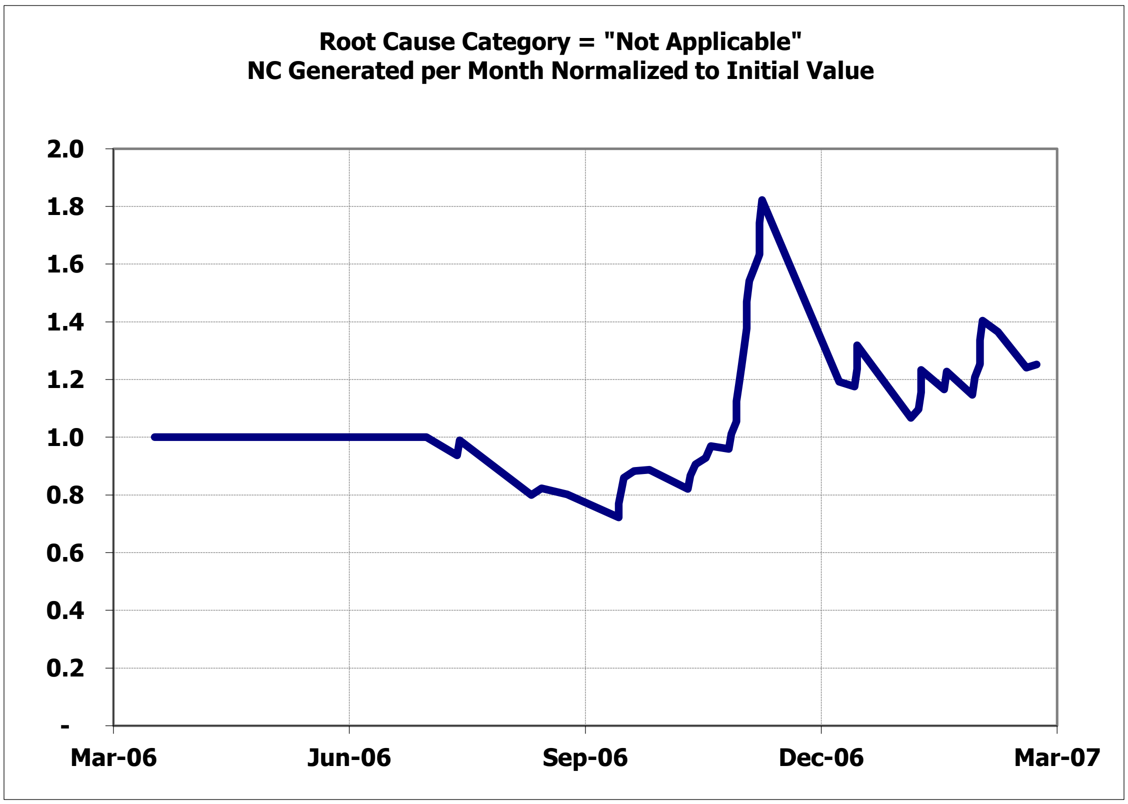
This example illustrates a subset that is displaying an unfavorable trend. The trend is for more NCs generated as time passes.

One way that NCs are divided into subsets is by Root Cause Category. We select the Root Cause Category "Not Applicable". The Facilities and Engineering Department for the time period 3/6/2006 – 3/6/2007 had 64 NCs with a Root Cause Category classified as "Not Applicable". For this data set, we know

$$N = 64 \quad (15)$$

$$A = 16 \quad (16)$$

The results:



The complete data set with calculations:

N (Sample Number)	Date of NC	Days Between NCs	Decaying Memory Average Days Between NCs	Decaying Memory Average NCs Generated per Month	NC Generated per Month Normalized to Initial Value	
1	03/17/2006	4	5.8	5.19	1.0	
2	03/21/2006	4	5.8	5.19	1.0	
3	03/26/2006	5	5.8	5.19	1.0	
4	03/28/2006	2	5.8	5.19	1.0	
5	03/28/2006	0	5.8	5.19	1.0	
6	04/01/2006	4	5.8	5.19	1.0	
7	04/01/2006	0	5.8	5.19	1.0	
8	04/06/2006	5	5.8	5.19	1.0	Linear Average Region
9	05/06/2006	30	5.8	5.19	1.0	
10	05/06/2006	0	5.8	5.19	1.0	
11	05/16/2006	10	5.8	5.19	1.0	
12	05/16/2006	0	5.8	5.19	1.0	
13	05/18/2006	2	5.8	5.19	1.0	
14	06/02/2006	15	5.8	5.19	1.0	
15	06/17/2006	15	5.8	5.19	1.0	
16	07/01/2006	14	5.8	5.19	1.0	
17	07/13/2006	12	6.2	4.86	0.9	
18	07/14/2006	1	5.9	5.13	1.0	
19	08/11/2006	28	7.2	4.15	0.8	
20	08/15/2006	4	7.0	4.27	0.8	
21	08/25/2006	10	7.2	4.16	0.8	
22	09/14/2006	20	8.0	3.74	0.7	
23	09/14/2006	0	7.5	3.99	0.8	
24	09/15/2006	1	7.1	4.22	0.8	
25	09/16/2006	1	6.7	4.46	0.9	
26	09/20/2006	4	6.6	4.58	0.9	
27	09/26/2006	6	6.5	4.60	0.9	
28	10/11/2006	15	7.1	4.25	0.8	
29	10/12/2006	1	6.7	4.50	0.9	
30	10/14/2006	2	6.4	4.70	0.9	
31	10/18/2006	4	6.2	4.81	0.9	
32	10/20/2006	2	6.0	5.03	1.0	
33	10/27/2006	7	6.0	4.97	1.0	
34	10/28/2006	1	5.7	5.25	1.0	
35	10/30/2006	2	5.5	5.47	1.1	
36	10/30/2006	0	5.1	5.83	1.1	
37	10/31/2006	1	4.9	6.14	1.2	
38	11/01/2006	1	4.6	6.46	1.2	
39	11/02/2006	1	4.4	6.80	1.3	
40	11/03/2006	1	4.2	7.14	1.4	Decaying Memory Average Region
41	11/03/2006	0	3.9	7.62	1.5	
42	11/04/2006	1	3.8	7.99	1.5	
43	11/06/2006	2	3.6	8.23	1.6	
44	11/08/2006	2	3.5	8.47	1.6	
45	11/08/2006	0	3.3	9.04	1.7	
46	11/09/2006	1	3.2	9.45	1.8	
47	12/09/2006	30	4.9	6.18	1.2	
48	12/15/2006	6	4.9	6.09	1.2	
49	12/16/2006	1	4.7	6.41	1.2	
50	12/16/2006	0	4.4	6.84	1.3	
51	01/06/2007	21	5.4	5.53	1.1	
52	01/09/2007	3	5.3	5.69	1.1	
53	01/10/2007	1	5.0	5.99	1.2	
54	01/10/2007	0	4.7	6.39	1.2	
55	01/19/2007	9	5.0	6.05	1.2	
56	01/20/2007	1	4.7	6.36	1.2	
57	01/30/2007	10	5.0	5.95	1.1	
58	01/31/2007	1	4.8	6.26	1.2	
59	02/02/2007	2	4.6	6.50	1.3	

Side By Side Analysis

Note how both Examples can be compared and contrasted on the same graph because of the normalization based on the initial value. Values less than 1 are favorable; values greater than 1 are unfavorable: