Economic Value of Precision in the Monitoring of Linear Systems

Miguel Bagajewicz
School of Chemical Engineering and Materials Science, University of Oklahoma, Norman, OK 73019

Mariusz Markowski and Andrzej Budek
Warsaw University of Technology, Plock, Poland

DOI 10.1002/aic.10372
Published online March 1, 2005 in Wiley InterScience (www.interscience.wiley.com).

Keywords: instrumentation design, process monitoring

Introduction

In a recent conference paper[1] it was suggested that in the case of linear systems, the value of unbiased instrumentation is directly connected to the precision of the estimators of the flows measured. These estimators could be obtained directly from measurements or from data reconciliation. In fact, it is well known that data reconciliation provides estimators with smaller standard deviations when there is enabling software redundancy. These estimators have better precision and therefore correspond to higher values.

The conference paper suggested that the downside expected financial loss (*DEFL*), associated with a product stream i, is given by

$$DEFL(\hat{\sigma}_i) = \gamma_p K_S T \hat{\sigma}_i \tag{1}$$

where $\gamma_p \approx 0.2$, when the downside deviation is calculated with probability $p = 0.5$, that is, when all negative deviations from the target are considered. The value of instrumentation is therefore found by subtracting the downside financial loss before and after the new instrumentation is added. The paper stated the expression but offered no proof. In addition, its validity was limited to the case of no process variations.

This short article

(1) Provides the proof of Eq. 1.

(2) Extends Eq. 1 to the case where process variability takes place.

(3) Provides means to calculate the associated probability of loss when process variability is considered.

(4) Generalizes the probability and the associated expected downside financial loss to consider finite deviations from the targeted production values.

Probabilities

The analysis concentrates on one product rate and the downside financial losses obtained can be added up for all products. The expected value of total production of product i, it was argued, is related to the true value of the true flow rate m_i throughout time. The probability of not meeting the targeted production is equal to the probability of the true value of m_i being smaller than the targeted production m_i^*, that is, $P\{m_i(t) \le m_i^*\}$. We then considered that production is adjusted to meet the targeted value, based on the estimator's value $\hat{m}_i$. That is, if $\hat{m}_i < m_i^*$, production is increased and vice versa, if $\hat{m}_i > m_i^*$, production is decreased. Consider the case $\hat{m}_i > m_i^* > m_i^*$, that is, the estimator indicates that the target has been met. In such a situation, we assume for simplicity that the operator would not do any correction to the set points. Such an assumption can be relaxed. The probability of being wrong is given by the following conditional probability: $P\{m_i \le m_i^* | \hat{m}_i \ge m_i^*\}$, that is, the probability of having missed the target, given that the estimator is larger than the target. Because these are independent, the above probability is equal to $P\{m_i \le m_i^*\} P\{\hat{m}_i \ge m_i^*\}$, which was argued to be equal to 0.25. The reason given was that the probability of the true value being lower than the target $P\{m_i \le m_i^*\}$ and the probability of the measurement to be larger than the true value $P\{\hat{m}_i \ge m_i^*\}$ are independent and equal to 0.5. We now derive a more general expression. First, note that

M. Bagajewicz's e-mail address is bagajewicz@ou.edu.

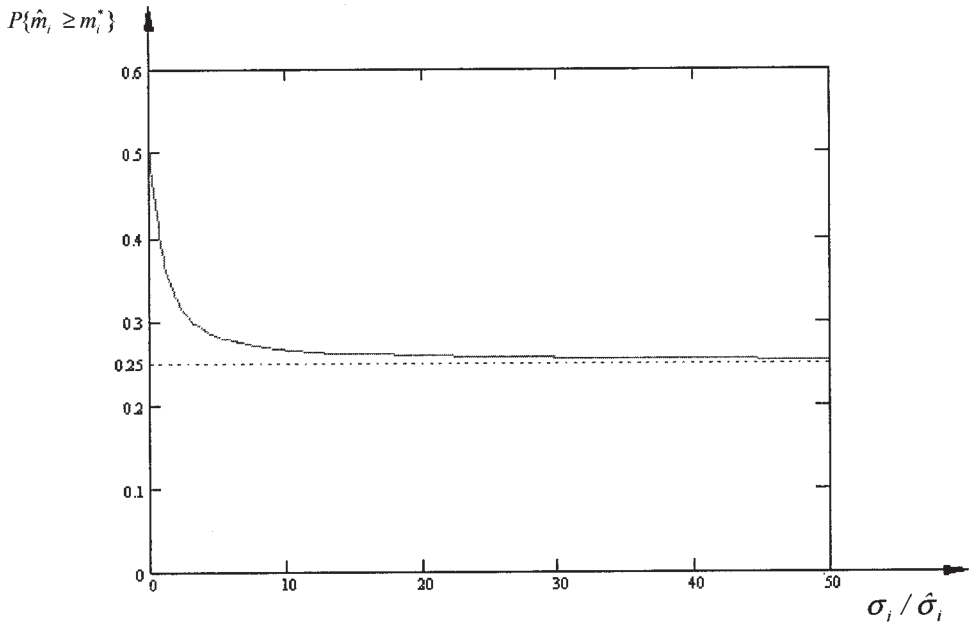

Figure 1. Probability of measurement being below target.

$$P\{m_i \leq m_i^*\} = \int_{-\infty}^{m_2^*} g_P(\xi;\, m_i^*,\, \sigma_i)d\xi \tag{2}$$

where $g_P(\xi;\, m_i^*,\, \sigma_i)$ is the probability distribution of the true process values ξ around the targeted mean m_i^* with standard deviation σ_i. Note that we assume here that the control system is such that the mean is the target value. Clearly, for any symmetric distribution, including the usual common choice of normal distributions, we have $P\{m_i \leq m_i^*\} = 0.5$. However, because this distribution is tied to the type of control system used, some distribution other than the normal, possibly non-symmetric, is possible. In turn, the probability of the measurement being larger than the target $P\{\hat{m}_i \geq m_i^*\}$ depends on both the value of m_i and the quality of the measurement, which in turn depends of the precision of the estimator ($\hat{\sigma}_i$) and is given by its probability distribution $g_M(\xi;\, m_i,\, \hat{\sigma}_i)$ around the real value m_i. Thus, for a fixed value of m_i we have

$$P\{\hat{m}_i \geq m_i^* | m_i\, fixed\} = \int_{m_i^*}^{\infty} g_M(\xi;\, m_i,\, \hat{\sigma}_i)d\xi \tag{3}$$

Now, because m_i follows a distribution we have

$$P\{\hat{m}_i \geq m_i^* | m_i\} = P\{\hat{m}_i \geq m_i^* | m_i\, fixed\}\, g_P(m_i;\, m_i^*,\, \sigma_i) \tag{4}$$

or, integrating over all possible values of the true rate,

$$P\{\hat{m}_i \geq m_i^*\} = \int_{-\infty}^{m_i^*} \left\{ \int_{m_i^*}^{\infty} g_M(\xi;\, m_i,\, \hat{\sigma}_i)d\xi \right\} g_P(m_i;\, m_i^*,\, \sigma_i)dm_i \tag{5}$$

Notice that the integral is taken over all possible values of m_i below the target because of the underlying assumption that m_i is lower than the target. When both distributions are normal, we obtain

$$P\{\hat{m}_i \geq m_i^*\} = \frac{1}{4} + \frac{1}{2\sqrt{\pi}} \int_0^{\infty} \mathrm{erfc}(z\sigma_i/\hat{\sigma}_i)e^{-z^2}dz \tag{6}$$

This value depends on the standard deviations of the true value and the measurement (Figure 1) when $\sigma_i/\hat{\sigma}_i \rightarrow 0$ the above probability has a limit of 0.5, which corresponds to the assumption made by Bagajewicz and Markowski.[1] When $\sigma_i/\hat{\sigma}_i \rightarrow \infty$, $P\{\hat{m}_i \geq m_i^*\} \rightarrow 0.25$. However, as soon as $\sigma_i/\hat{\sigma}_l$ increases to a value around 2, there is a decrease of this probability to values between 0.4 and 0.5. For example, for $\sigma_i/\hat{\sigma}_i = 2$, we have $P\{\hat{m}_i \geq m_i^*\} = 0.387$.

Financial Loss

We now determine what is the expected financial loss associated with this probability, that is a generalization of Eq. 1. The distributions of the true value and the estimator are depicted in Figure 2 for one instance of the true value m_i.

The downside financial loss incurred for a fixed value of m_i is given by $K_S T(m_i^* - m_i)$, where K_S is the value of the products sold and T is the period of time under consideration. One could argue that, in reality, the remedy for the shortcomings of production is solved by keeping inventory. This is indeed the case in many industries such as those engaged in refining. In such a case, the value of K_S needs to be reassessed to reflect the cost of keeping inventory. We keep the former interpretation in the understanding that the theory does not change, only the value of K_S does. Now, one could easily integrate this expression to obtain an expected value using the process distribution as follows: $K_S T \int_{-\infty}^{m_i^*} (\hat{m}_i^* - m_i) g_P(m_i;\, m_i^*,\, \sigma_i)dm_i$. However, this

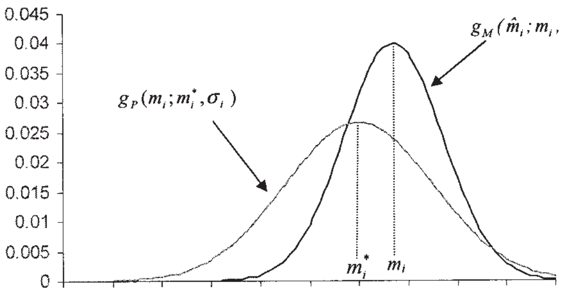

Figure 2. Example of process and measurement distributions.

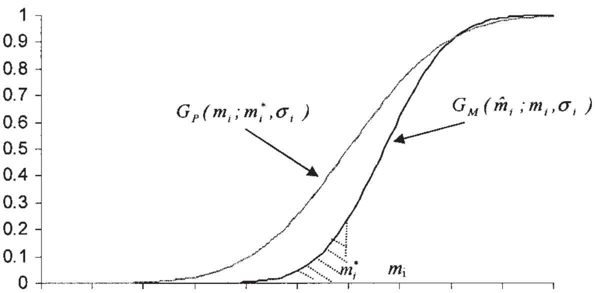

Figure 3. Process and measurement cumulative distributions.

would be wrong because it would calculate just an expected value for negative deviations from the target regardless of the information available, that is, not using the estimator $\hat{m}_i$ suggested by the measurement directly or, in general, by data reconciliation. Finally, when the operator performs adjustments when the production target is exceeded ($\hat{m}_i > m_i^*$), then an additional term for financial loss should be calculated. We leave this for future work.

To incorporate the notion that an estimator is being obtained (a measurement or a reconciled value) and therefore some information is added that allows the operator to make corrections during the time of operations, the proper integral over all possible values of m_i below the target that needs to be used to obtain the downside expected financial loss is the following:

$$DEFL(\hat{\sigma}_{i,m}|\hat{m}_i) = K_S T \int_{-\infty}^{m_i^*} (m_i^* - m_i)\, \hat{g}_L(m_i, \hat{\sigma}_{i,m}|\hat{m}_i) dm_i \tag{7}$$

where $\hat{g}_L(m_i, \hat{\sigma}_{i,m}\,|\,\hat{m}_i)$ is the likelihood function, that is, the probability distribution of the state m_i, given the estimator $\hat{m}_i$. However, the likelihood function is equal to the distribution of the measurement, that is

$$\hat{g}_L(m_i, \hat{\sigma}_{i,m}|\hat{m}_i) = g_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m}) \tag{8}$$

Now, it can be easily shown that Eq. 7 is equivalent to

$$DEFL(\hat{\sigma}_{i,m}|m_i) = K_S T \int_{-\infty}^{m_i^*} G_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m}) d\hat{m}_i \tag{9}$$

where $G_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m})$ is the cumulative distribution corresponding to $g_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m})$. This relationship was indicated by Barbaro and Bagajewicz.[2] The area shaded in Figure 3 shows the area corresponding to the integral in Eq. 9.

Now, because m_i is not fixed, we need to integrate over all possible values below the target m_i^* multiplied by the probability density of each state $g_p(m_i^*, \sigma_i, m_i)$. We integrate below the target because those values are the ones that will lead to financial loss. Thus

$$DEFL(\hat{\sigma}_{i,m}, \sigma_i) = \int_{-\infty}^{m_i^*} g_p(m_i^*, \sigma_i, m_i) \times \left\{ K_S T \int_{-\infty}^{m_i^*} (m_i^* - \hat{m}_i)\, g_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m}) d\hat{m}_i \right\} dm_i \tag{10}$$

When both distributions are normal, the expected financial loss is given by

$$DEFL(\hat{\sigma}_i, \sigma_i) = \gamma K_S T \left\{ \frac{\hat{\sigma}_i}{\sqrt{\left(\frac{\sigma_i}{\hat{\sigma}_i}\right)^2 + 1}} + \sigma_i \sqrt{\frac{1}{\left(\frac{\hat{\sigma}_i}{\sigma_i} + 1\right)}} \right\} \tag{11}$$

where $\gamma = 1/(2\sqrt{2\pi}) \int_0^\infty \xi e^{-\xi^2} d\xi = 0.19947$. When process variations are negligible ($\sigma_i \to 0$), this expression reduces to Eq. 1, which was given by Bagajewicz and Markowski.[1] Indeed, when the process variations are negligible, the probability density for the process $g_p(m_i^*, \sigma_i, m_i)$ reduces to one half of the delta function, that is, $g_p(m_i^*, \sigma_i, m_i) = 0.5\delta(m_i - m_i^*)$. The 0.5 factor comes from the fact that we are integrating in the interval $(-\infty, m_i^*)$. Conversely, when the process variations are large, that is, $\sigma_i \gg \hat{\sigma}_{i,m}$, the first term becomes irrelevant (that is, measurements have no influence) and $DEFL(\hat{\sigma}_{i,m}, \sigma_i) \to \gamma K_S T \sigma_i$. For intermediate values, one obtains some interesting results. Indeed, we rewrite the expression as follows

$$DEFL(\hat{\sigma}_{i,m}, \sigma_i) = \gamma K_S T \hat{\sigma}_{i,m} \beta(\sigma_i/\hat{\sigma}_{i,m}) \tag{12}$$

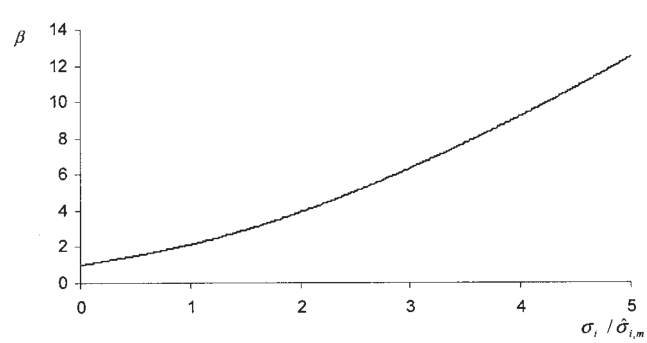

Figure 4. Correction factor for process variability.

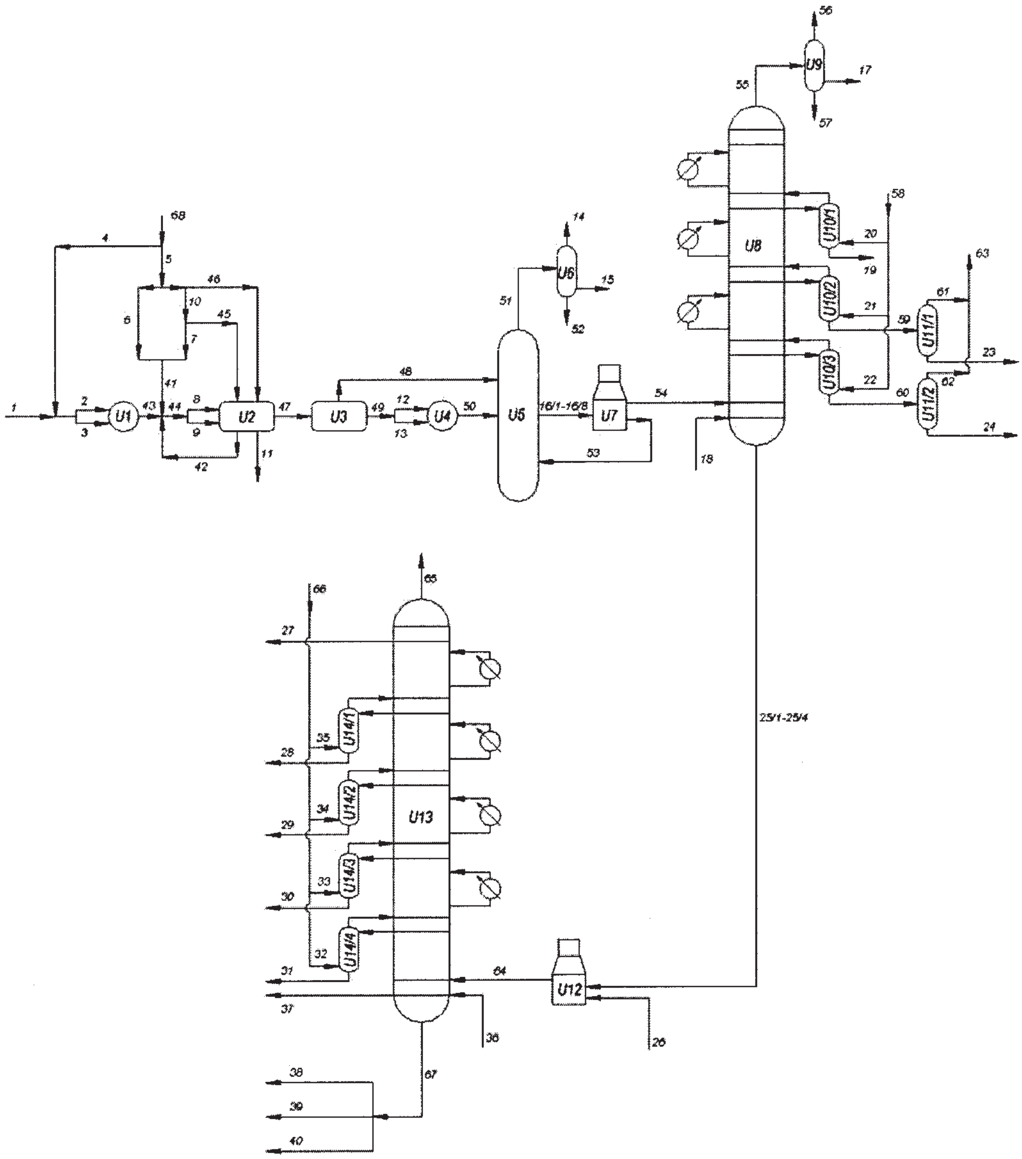

Figure 5. Flow sheet for crude distillation unit.

Units: U1, HEN; U2, desalting unit; U3, crude vessel; U4, HEN; U5, prefractionation tower; U6, condenser; U7, furnace; U8, atmospheric tower; U9, condenser; U10/1–U10/3, preflash column; U11/1, U11/2, atmospheric product dryer; U12, furnace; U13, vacuum tower; U14/1–U14/4, preflash column. *Streams*: crude oil: 1, 2, 3, 8, 9, 43, 44; desalted crude oil: 12, 13, 47, 49, 50; water 4, 5, 6, 7, 10, 41, 45, 46, 68; hydrocarbon vapors: 48, 56, 61, 62, 63, 65; sour water: 11, 42, 48; oily water: 52, 57; prefractionation products: 14, 15; atmospheric products: 17, 19, 23, 24; vacuum products: 27, 28, 29, 30, 31, 37, 67; steam: 18, 20, 21, 22, 26, 32, 33, 34, 35, 36, 58.

where

$$\beta(x) = \left\{ \frac{1}{\sqrt{x^2+1}} + x\sqrt{\frac{1}{\left(\frac{1}{x}+1\right)}} \right\} \tag{13}$$

The term $\beta(\sigma_i/\hat{\sigma}_{i,m})$ is an increasing function of $\sigma_i/\hat{\sigma}_{i,m}$ (Figure 4).

We recall that the assumption was made that the operator will not introduce corrective actions when the measurement is above the target. Therefore, the expected financial loss is associated with a given level of confidence, given by

$$P\{m_i \leq m_i^* | \hat{m}_i \geq m_i^*\} = 0.5P\{\hat{m}_i \geq m_i^*\} \tag{14}$$

which is a function of the measurements and the process variability. Thus, a given set of instruments provides both this level of confidence and its corresponding expected financial loss, which is only function of the estimator's precision. Therefore, under almost no process variations (there is always some variation), there is a 25% chance that a downside financial loss of $\gamma K_S T\hat{\sigma}_{i,m}$ is achieved. If corrections to set points are made when readings indicate that targeted productions are exceeded the result would be even worse.

The above probability can also be seen as the risk one is

Table 1. Stream Data

Stream Number	Measured Flow Rate (kg/h)	Mass Flow Rate from Balance (kg/h)	Standard Deviation (kg/h)	Flow Rate after Reconciliation (kg/h)	Standard Deviation after Reconciliation (kg/h)	Stream Number	Measured Flow Rate (kg/h)	Mass Flow Rate from Balance (kg/h)	Standard Deviation (kg/h)	Flow Rate after Reconciliation (kg/h)	Standard Deviation after Reconciliation (kg/h)
1	418839	—	6283	413336	1259	31	18187	—	273	18196	191
2	212050	—	3181	209310	2339	32	312	—	5	314	4.87
3	213020	—	3195	210256	2341	33	338	—	5	340	4.87
4	6231	—	93	6231	91	34	325	—	5	327	4.87
5	20352	—	305	20327	244	35	311	—	5	313	4.87
6	7174	—	108	7174	108	36	3226	—	48	3225	48
7	7256	—	109	7256	109	37	18097	—	271	18106	190
8	230650	—	3460	230650	3460	38	15141	—	227	15154	224
9	229870	—	3448	229870	3448	39	20245	—	304	20268	297
10	10188	—	153	10188	153	40	12650	—	190	12659	188
11	26180	—	393	26243	391	41	—	14430	—	—	192
12	209170	—	3138	206932	2296	42	—	26523	—	—	5048
13	208950	—	3134	206718	2295	43	—	419566	—	—	1256
14	5122	—	77	5124	54	44	—	460520	—	—	4886
15	21434	—	322	21467	227	45	—	2932	—	—	192
16/1	62562	—	938	61562	938	46	—	2966	—	—	288
16/2	60985	—	915	60985	915	47	—	413651	—	—	1186
16/3	61253	—	919	61253	919	48	—	0	—	—	NA*
16/4	61490	—	922	61490	922	49	—	413651	—	—	1191
16/5	61009	—	915	61109	915	50	—	413651	—	—	1200
16/6	60796	—	912	60796	912	51	—	27068	—	—	192
16/7	62012	—	930	62012	930	52	478	—	7	478	7
16/8	60413	—	906	60413	906	53	—	103938	—	—	2849
17	45680	—	685	45829	478	54	—	386582	—	—	1171
18	4275	—	64	4272	64	55	—	57169	—	—	494
19	26084	—	391	26133	275	56	7130	—	107	7137	107
20	155	—	2	156	1.98	57	4200	—	63	4202	63
21	256	—	4	260	3.83	58	795	—	12	759	5.85
22	337	—	5	343	4.65	59	—	73900	—	—	752
23	73319	—	1100	73704	753	60	—	50852	—	—	538
24	50533	—	758	50716	528	61	196	—	3	196	3
25/1	45721	—	686	45902	615	62	136	—	2	136	2
25/2	45698	—	685	45878	615	63	—	332	—	—	3.6
25/3	45747	—	686	45928	615	64	—	185593	—	—	721
25/4	45671	—	685	45851	615	65	4512	—	68	4513	68
26	2035	—	31	2035	31	66	1322	—	20	1293	9
27	38515	—	578	38557	392	67	—	48081	—	—	245
28	18921	—	284	18931	198	68	26583		399	26559	249
29	19835	—	298	19846	208						
30	23864	—	358	23880	249						

*NA, not applicable.

incurring. We now generalize the above idea to an arbitrary value of probability/risk. Consider now a downside deviation Δ from the target production m_i^*. The probability of the true value not meeting the target production, given that the measured value is lower than the target minus the downside deviation, is

$$P\{m_i \leq m_i^* | \hat{m}_i \leq m_i^* - \Delta\} = P\{m_i \leq m_i^*\} P\{\hat{m}_i \leq m_i^* - \Delta\} = 0.5p \quad (15)$$

where p is the probability of the measurement being smaller than $m_i^* - \Delta$. Such probability is given by

$$P\{\hat{m}_i \geq m_i^* - \Delta\} = \int_{-\infty}^{m_i^* - \Delta} \left\{ \int_{m_i^* + \Delta}^{\infty} g_M(m_i, \hat{\sigma}_{i,m}, \xi) d\xi \right\} g_P(m_i^*, \sigma_i, m_i) dm_i \quad (16)$$

In other words, there is a direct correspondence between p and Δ. The expected downside financial loss corresponding to this probability is

$$DEFL(\hat{\sigma}_{i,m}, \sigma_i) = \int_{-\infty}^{m_i^* - \Delta} g_p(m_i^*, \sigma_i, m_i) \times \left\{ K_S T \int_{-\infty}^{m_i^* - \Delta} (m_i^* - \hat{m}_i) g_M(\hat{m}_i; m_i, \hat{\sigma}_{i,m}) d\hat{m}_i \right\} dm_i \quad (17)$$

Trade-Offs between Value and Cost

We next briefly consider the trade-offs. In the case of buying a data reconciliation package, one would write

$$NPV = d_n\{DEFL(\bar{\sigma}_{i,m}, \sigma_i) - DEFL(\hat{\sigma}_{i,m}, \sigma_i)\} - \text{Cost of License} \quad (18)$$

where d_n is the sum of discount factors for n years, $\bar{\sigma}_{i,m}$ is the value of precision after data reconciliation and $\hat{\sigma}_{i,m}$ represents the actual value of precision obtained directly from measurements.

In the case of instrumentation upgrade, a similar expression holds:

$$NPV = d_n\{DEFL(\bar{\sigma}_{i,m}, \sigma_i) - DEFL(\hat{\sigma}_{i,m}, \sigma_i)\} - \text{Cost of New Instrumentation} \quad (19)$$

where $\bar{\sigma}_{i,m}$ is the value of precision after new instrumentation is added and $\hat{\sigma}_{i,m}$ is the actual value. In more complete formulations one would also discount the maintenance costs. One can, of course, continue with obtaining the net present value of reliability and other robustness network properties.

Example

A flow sheet of a crude distillation unit (CDU) is shown in Figure 5. This contains the vacuum unit that was presented by Bagajewicz and Markowski.[1] The values of mass flow for process streams and the reconciled values are given in Table 1.

Net present value of performing data reconciliation

Using the data listed in Table 1 and the costs of products of Table 2 (based on a cost of crude of 30 $/bbl), the downside financial loss for the existing instrumentation without the aid of data reconciliation is around $7.36 million, whereas after applying data reconciliation it decreases to $7.12 million. This renders a net present value (over only 5 years) of $236,817. This might justify the purchase of a data reconciliation package.

Net present value of new instrumentation

We do not explore here any strategy to optimize the addition of new instrumentation. Rather, we illustrate the effect of adding an instrument of 1.5% standard deviation to some streams. The results are shown in Table 3. The first column of this table indicates which instruments have been added. The second column indicates the net present value of the project, assuming no data reconciliation package has been installed, whereas the third column indicates the net present value for the case where the plant already has data reconciliation installed and in use.

Table 2. Stream Values

Stream	K_S ($/kg)	Stream	K_S ($/kg)
14–15	0.13	28	0.13
17	0.14	29	0.12
19	0.25	30	0.11
23	0.24	31	0.1
24	0.23	37	0.08
27	0.15	38–39–40	0.06

Table 3. Effect of New Flow Meters on Savings

	NPV	
Location of New Sensors	Without Originally Installed Data Reconciliation	With Originally Installed Data Reconciliation
14, 15, 17, 19, 23, 24, 27, 28, 29, 30, 31, 37, 38, 39, 40	$2,088,108	$1,851,290
14, 15, 17, 19, 23, 24, 27, 28, 29, 30, 31, 37, 38, 39	$2,098,762	$1,862,945
14, 15, 17, 19, 23, 24, 27, 28, 29, 30, 31, 37, 38	$2,071,202	$1,834,384
14, 15, 17, 19, 23, 24, 27, 28, 29, 30, 31, 37	$2,052,908	$1,816,091

Acknowledgments

Financial support from the National Science Foundation, Grants CTS-0075814 and CTS-0203153, is acknowledged.

Notation

K_S = cost of rejection of production
NPV = net present value
d_n = sum of discount factors for n years
$DEFL$ = downside expected financial loss
g_M = density distribution of measurements
g_p = density distribution of process values
H = holdup
m = flow rate
T = time horizon

Greek letters

σ = precision
$\hat{\sigma}$ = precision after data reconciliation
μ = mean value
Δ = threshold deviation for quality assurance

Literature Cited

1. Bagajewicz M, Markowski M. Instrumentation design and upgrade using an unconstrained method with pure economical objectives. Proc of FOCAPO 2003. Coral Springs, FL: Foundations of Computer Aided Process Operations, January; 2003.
2. Barbaro AF, Bagajewicz M. Managing financial risk in planning under uncertainty. *AIChE J.* 2004;50:963-989.

Manuscript received May 11, 2004, and revision received Aug. 2, 2004.