LA-UR- 11-02581

Approved for public release; distribution is unlimited.

Title:	Development of Pattern Recognition Options for Combining Safeguards Subsystems
Author(s):	T. Burr M. Hamada
Intended for:	2011 MPACT (Material Protection and Control) Meeting May 4-5, 2011 ABQ, NM NE

Los Alamos
NATIONAL LABORATORY
EST. 1943

Form 836 (7/06)

Development of Pattern Recognition Options for Combining Safeguards Subsystems

Initial focus on 2 subsystems:

nuclear materials accounting (NMA), process monitoring(PM)

Figure of merit: Alarm probability

$$AP = P(\text{alarm} \mid \text{diversion scenario})$$

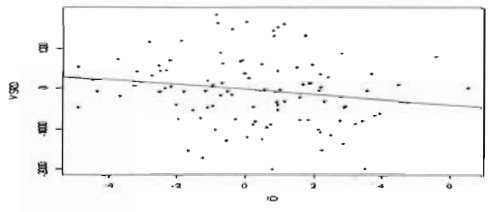
Diversion scenario:

- For NMA: how much, over what balance periods
- For PM: how much, over what time frame, and HOW

Tom Burr, Mike Hamada, George Tompkins

MPACT review meeting May 4-5, 2011

Pattern Recognition for Safeguards Data Integration



Volume shipper-receiver difference (SRD) from solution monitoring (SM) vs inventory difference from nuclear material accounting (NMA)

POCs: Tom Burr (LANL), Mike Hamada (LANL)

Statement of the MPACT Problem:

- Evaluate custom pattern recognition methods for combining process monitoring (PM) data with nuclear material accounting data (NMA) to recognize facility misuse with high detection probability

Key Outcomes:

- An approach to combine multiple safeguards systems to enhance pattern recognition

Technical Challenges

- Some subsystems such as SM and NMA are not Independent
- "Scores" from SM do not have a normal distribution
- Subsystems report data/scores/decisions at different rates. Novel pattern recognition methods must recognize a wide range of patterns and not be tuned to particular diversion scenarios.

Planned Accomplishments

- For a particular aqueous reprocessing facility, (a) describe PM and NMA subsystems to be evaluated, and (b) describe two misuse scenarios involving the misdirection of special nuclear material occurred, in terms of "how," "how much," and "over what time frame."
- For the chosen PM and NMA subsystems, develop, implement, and test candidate pattern recognition options for specified misuse scenarios.

What is PM?

- What is PM? Depends on who you ask.
- Near-real-time-accounting (NRTA) and PM:
  often both rely on frequent but short-cut less than full accountability measurements.

  PM: no attempt to "flush out" material and not necessarily mass-balance based.

  Ex1: Curium-based neutron monitor for leached hulls in aqueous reprocessing head end. The detected neutrons originate mostly from Cm, but it is assumed that Cm and Pu cannot be separated in the head end.

What is PM?

Ex 2: Bulk weight (cheap/fast measurements) and historical purity/correction factor to infer Pu mass

Ex 3: SM: in-tank volume and density, so

{Volume, Mass} every few minutes

and possibly also

{flow rates, flow acid concentration}

Maybe someday: in-line SNM concentration, gamma spec (Multi-isotope PM from PNNL) to detect off-normal conditions

PM

PM often includes data being collected for what is traditionally regarded as non-safeguards reasons such as process control/safety.

NRC/DOE/IAEA all encourage PM but DOE/IAEA have no formal basis to evaluate the safeguards benefits of PM.

Data-rich example: SM

The IAEA has tools for inspectors to manually view SM data (level, density, temperature every few minutes from many tanks), but no basis to quantify its benefit. Subjectively, most agree the SM data has great safeguards benefit.

Where is PM going?

PM is increasingly used:

- Aqueous reprocessing: neutron-based Cm monitoring, SM, pulsed column holdup assay, in-line NDA of Pu conc
- MOX fuel fab: glove box assay, distributed source term analysis, net weight checks
- Gas centrifuge enrichment: load cell monitoring, on-line gamma-based enrichment monitoring

- How to decide what type of PM has good cost/benefit?

PM benefits

Possibilities:

1) NMA remains objective/quantitative basis for AP.
   PM used to resolve alarms, support error models.
2) PM in driver's seat to trigger physical inventory taking.
3) **PM and NMA on "equal footing." FY11 effort here.**

Role/need for **Models**:

- ITV-type measurement error models (ran. and sys.)
- "Process variation" effects for SM. Example: pump carryover.
- Separation areas: predict effluent/waste stream isotopics and Pu as function of input streams (SRS/SEPHIS) and holdup models.
- Dissolver model (ANL): rather have a dissolver model and monitor dissolver temperature and acid concentration, or monitor the hulls? Diversion path is to the hulls.

Pattern Recognition

- Develop data-driven (operating as declared or misuse A or misuse B) and period-driven (at the end of each day or balance period, make a judgment) pattern recognition.
- Attention to within-data-type temporal correlations (due for example to systematic errors) and to between-data-type correlations (due for example to shared dip tubes by SM and NMA).
- Consider data fusion at feature, score, or decision level
- Include issues involving multiple testing and associated false alarm rates. Allowing a high "false discovery" rate might be acceptable, provided there is efficient, timely, anomaly resolution, such as J. Howell's model-based diagnostic methods.
- Regardless whether anomaly resolution is used, need to assess false negative rates for a given low overall false positive rate* and misuse data requires simulation, so quantifying and dealing with computer model uncertainty will be a large issue.

*(low false positive after anomaly resolution)

7-tank Example

- SM every few minutes, filtered to "SM scores" based on treating each tank as sub-MBA, so monitor wait and transfer modes.
- {M,V} checks for wait and transfer modes;

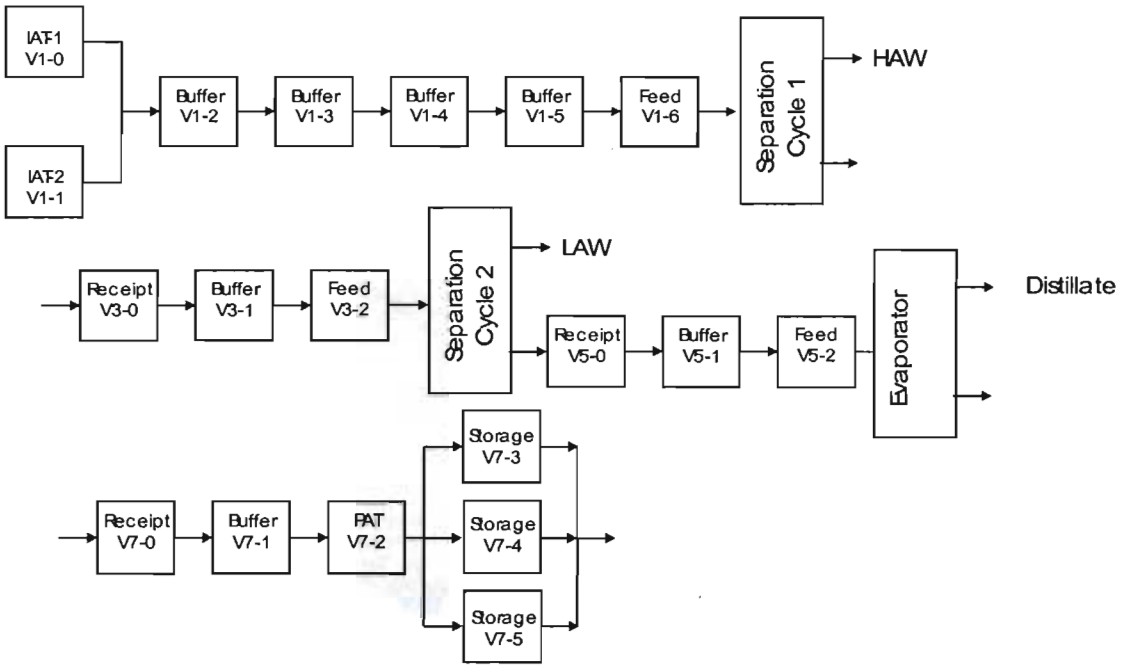
  MB every 10 days, and

within MB setting,

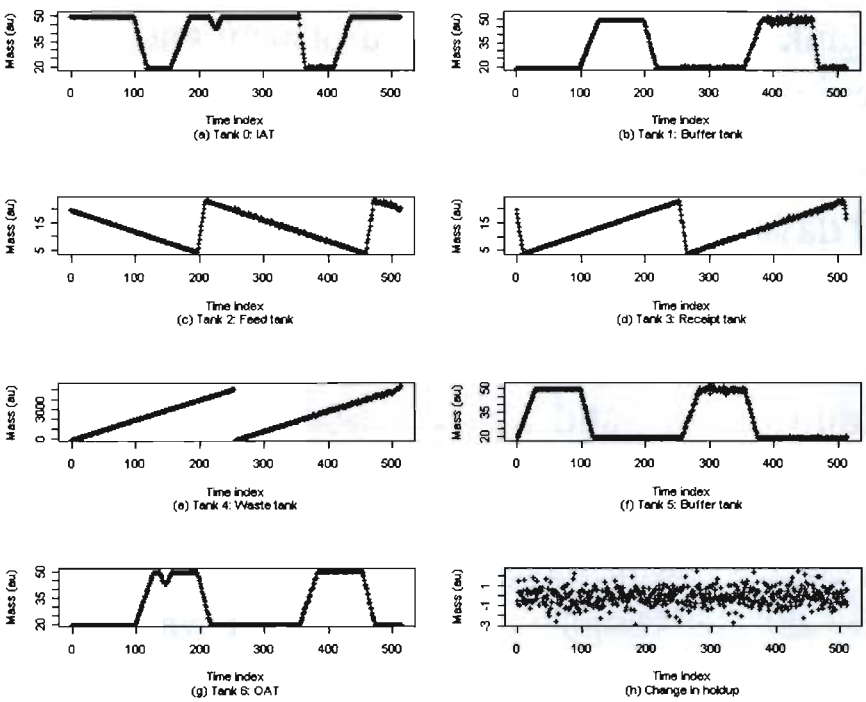
models provide book values for waste and holdup, so check measured holdup vs book and waste vs book.

- Related work: Barnwell depleted U runs where real diversions occurred and SM was key component, with individual tanks regarded as sub MBAs. However, system-wide false alarm rates were not assessed although cusums and MBs were used without attention to all the within- and between- correlations.

Generic aqueous reprocessing (SNL, Glasgow, LANL)



Example 7-tank SM data
Input, Buffer, Feed, Receipt, Waste, Buffer, Output



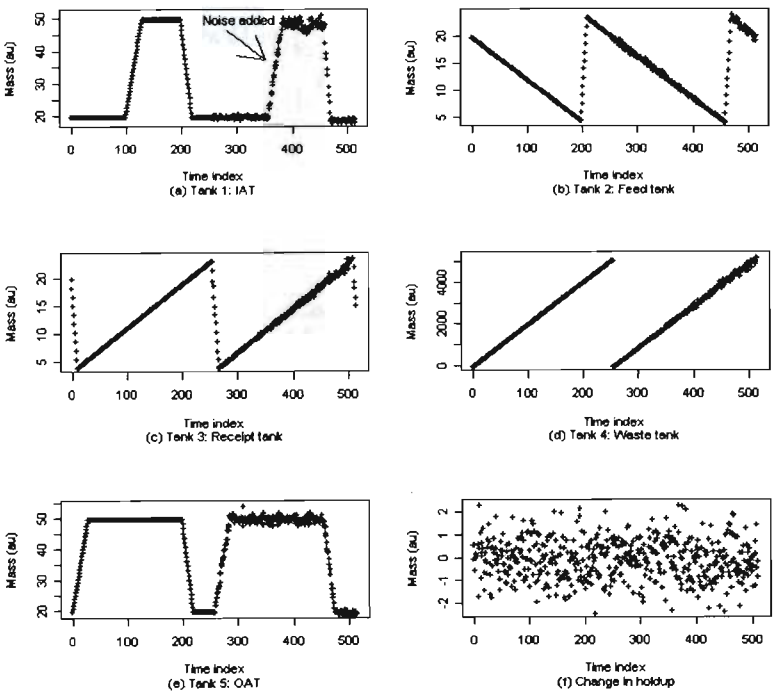
7-tank MBA:
B/B input
B/B buffer
B/C feed
C/B receipt
C/B waste
B/B buffer
B/B output

Also:
neutron-based measurement of the change in holdup,

and

model for Pu flow to waste that provides a book value for waste

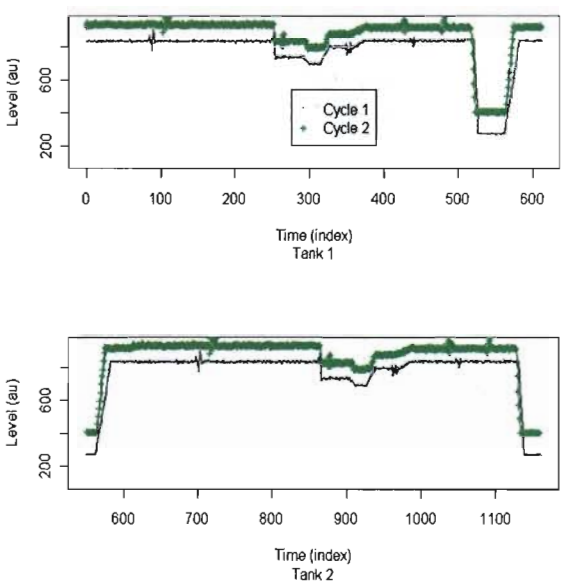
Improving quality of simulated data



Separate project: process variation and meas. error effects
Notice: second portion of simulated data is "noisier" and includes:
process variation effects such as "synchronization" effects, pump carryover, and

ITV ran/sys meas errors

2 Current Types of SM: (1) template and (2) sub-MBA

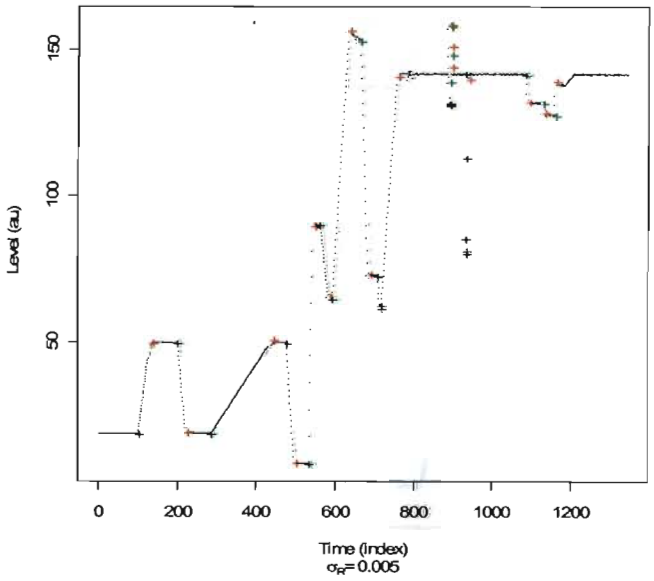


Template option and sub-MBA with event marking option both generate "residuals."

NMA also generates "residuals."

Current effort to characterize behavior of "residuals" for sub-MBA option.

Event marking for sub-MBA SM option

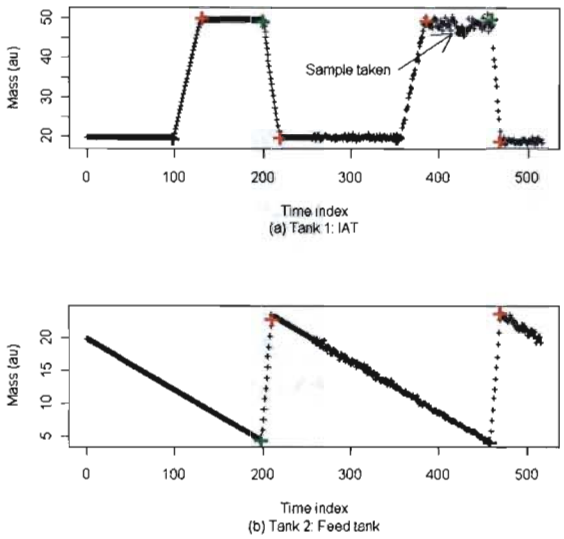


Green "+" marks start
Red "+" marks end

Residuals from "wait" modes and "transfer" modes.

Sampling, recirculation are "ignored" and occur during "wait" modes.

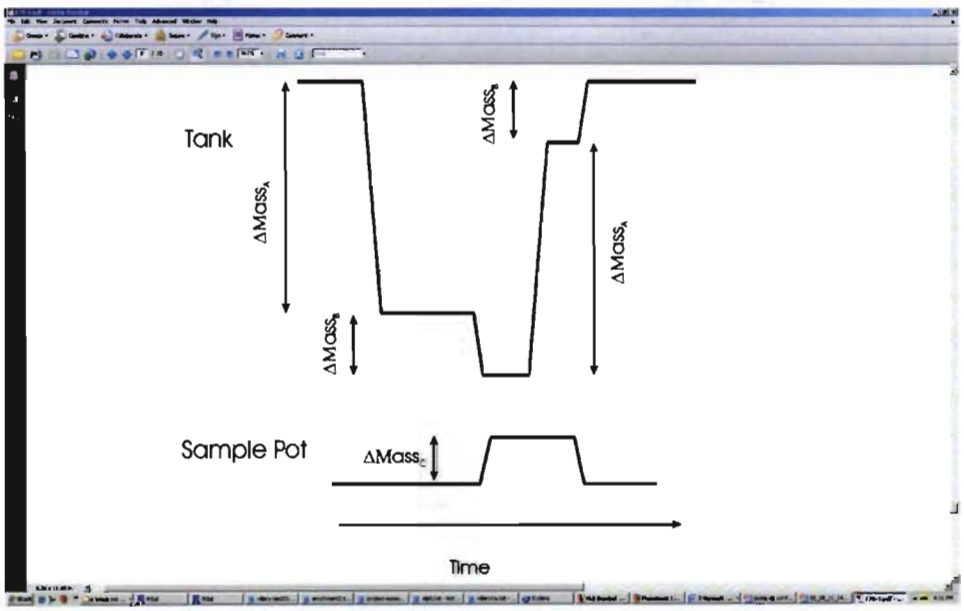
Event marking



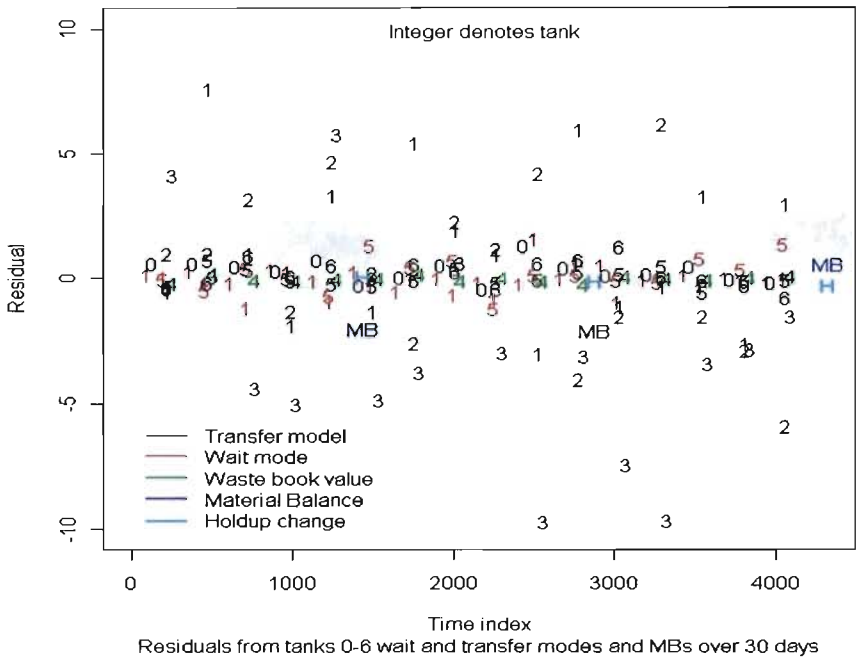
Top: "ignore" the sample event, but monitor change between start and stop of "wait" mode.

Bottom: Continuous-mode tanks have "wider" control limits on residuals

Current SM systems (J. Howell) check sample and recirculation patterns.

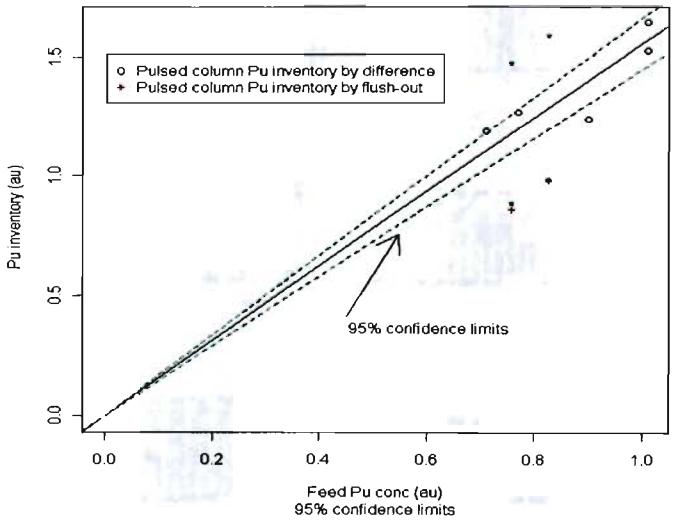


Simulated SM and NMA data for 7-tank MBA
NMA data every 10 days, SM data every 6 minutes

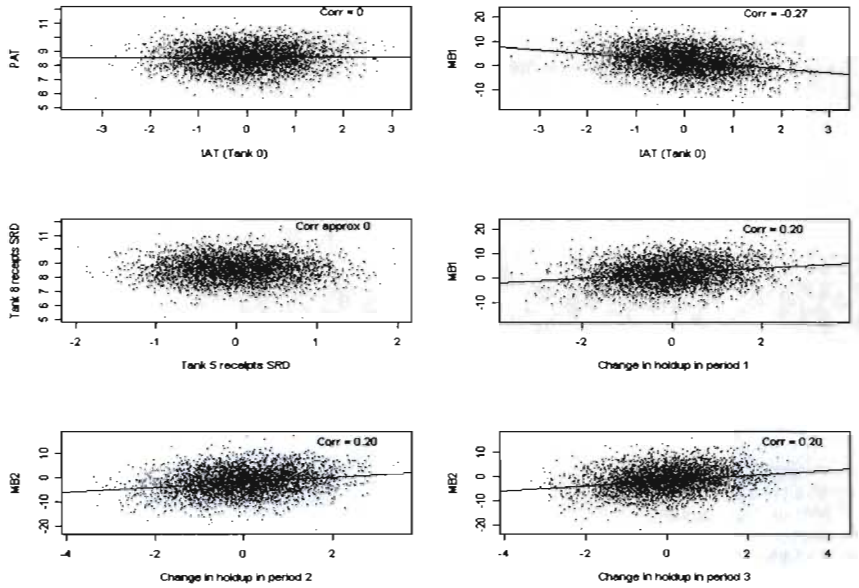


Residuals from tanks 0-6 wait and transfer modes and MBs over 30 days

Pulsed column model using input Pu to estimate holdup
Numerical example from Yamaya et al. 2009 INMM for RRP



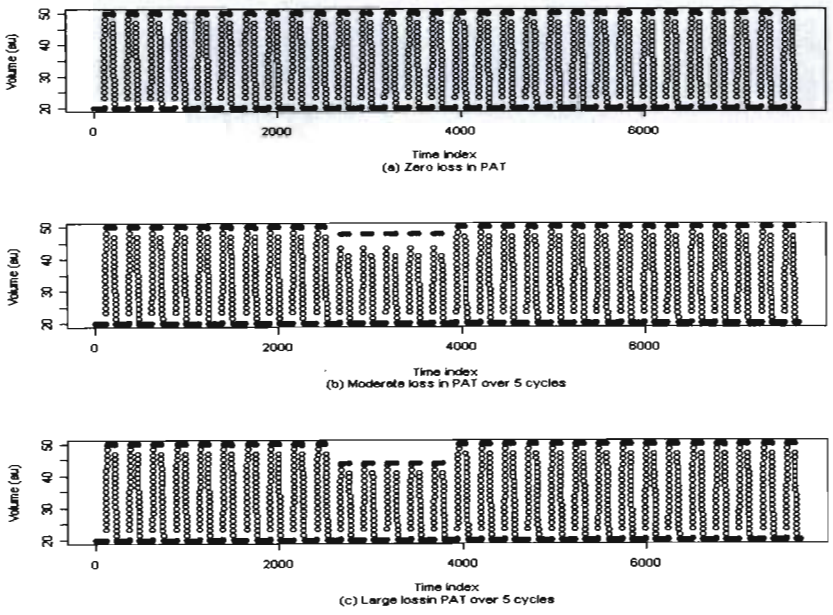
Examples of between-data correlations



Assessing between-data correlations via simulation.

Plotted data is multiple realizations of the 30-day balance period.

Example diversion: loss over 5 consecutive batches in PAT at days 10 to 15

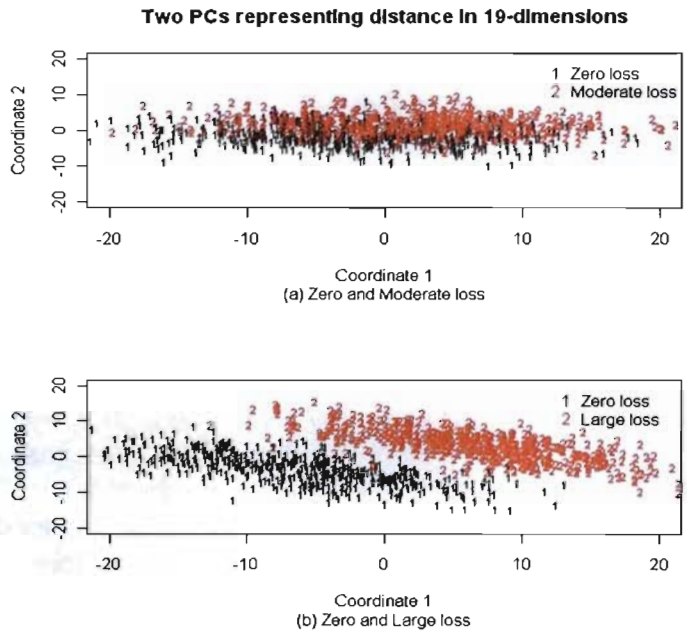


Top: zero loss

Middle: moderate loss

Bottom: large loss

Principle coordinate plots for19-components: 10 waits, transfers, 3 MBs, 3 waste book, 3 holdup book



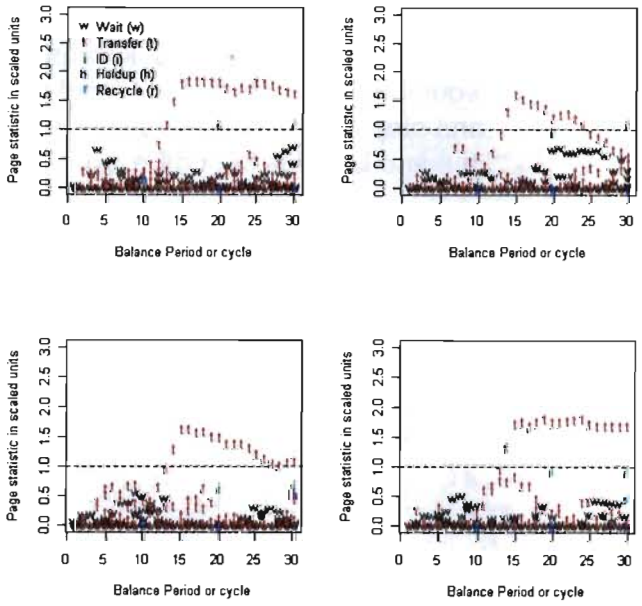
Moderate loss is approx 1% of 30 day thruput

Large loss is approx 3% of 30-day thruput

19 components:
10 from: 3 wait regions for tank1, 2 wait for tank 5, rec1, rec2, rec3, rec5, rec6
3 from mbseq, choldup, wastebook

Principle coordinates:
if use Euclidean distance, principle coordinates same as principle components.
Also using nonlinear multidim scaling such as sammon [21]

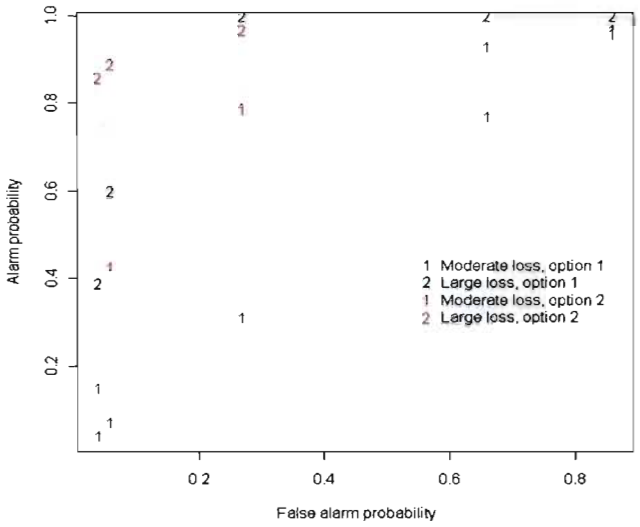
Pattern recognition



Moderate loss, 4 realizations of Page's statistic, the alarming transfer modes are for tank 7 transfers.

Not yet attempting to control overall false alarm rate of 13 non-independent Page statistics

Pattern recognition, controlling overall false alarm rate



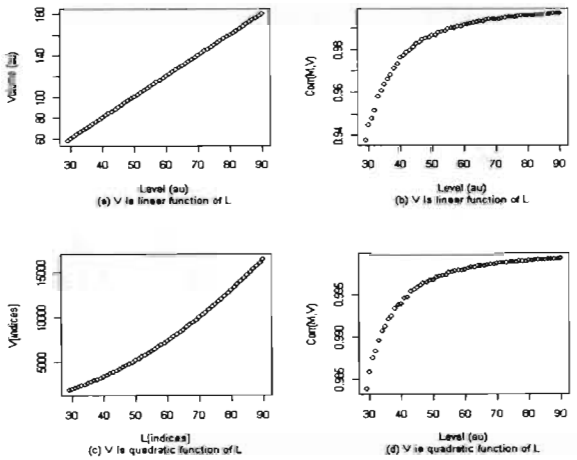
Mahalanobis distance from center of zero-loss distribution as pattern recognition option is the same as Fisher's linear discriminant analysis (very common, established method)

Option 1: 13 separate Page tests, but control for multiple tests.

Option 2: Mahalanobis distance from mean of zero-loss distribution. AP results for the small (near 0.01 to 0.05) FAPs are most relevant.

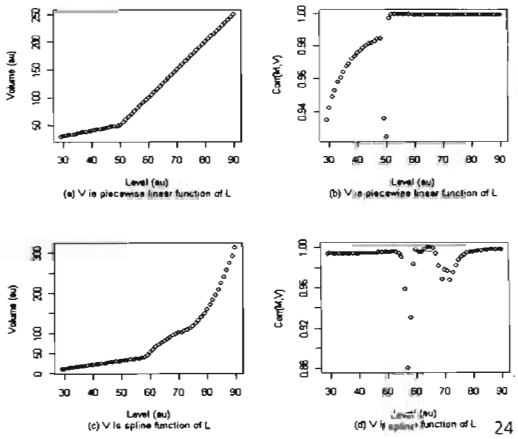
APs for the higher FAP are given for completeness and to emphasize that it is important to control for the FAP in such "multiple-testing" situations.

Work in progress: {V,M} monitoring V and M measurements are correlated in standard 3 dip-tube arrangement

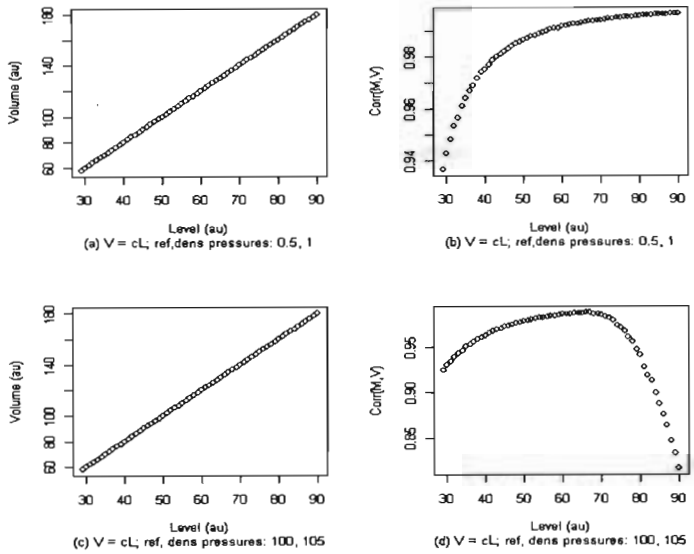


Corr(M,V) depends on dip tube separations and on relation between L and V

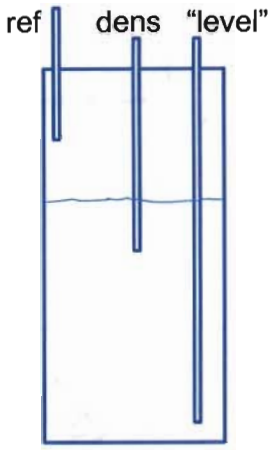
Safeguards by design: tank transfers would use consistent and friendly start and stop levels. This is actually done in some facilities based on experience.



Examples of corr(M,V)

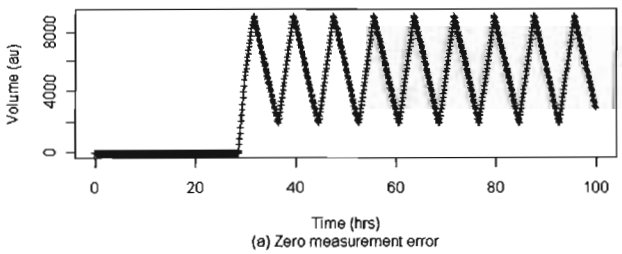


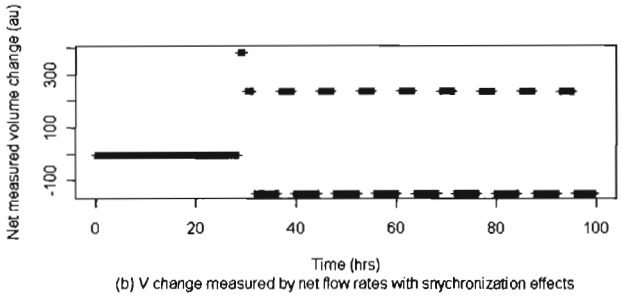
V = cL, but different dip tube separations, pressures



V to L calibration: V is rarely a linear function of L, and dip tube separations must be known, but tubes expand different amounts as function of temperature. R. Binner et al, 2009 good reference updates A. Liebetrau/others work on V to L

Other SM data: In-tank and flow measurements





Ex. of synchronization effects: flow rate changes at unknown time between minute 101 and 106, and simulated {V,M} in tank knows exactly when flow changed, but only observe the instant flow rates every 6 minutes.

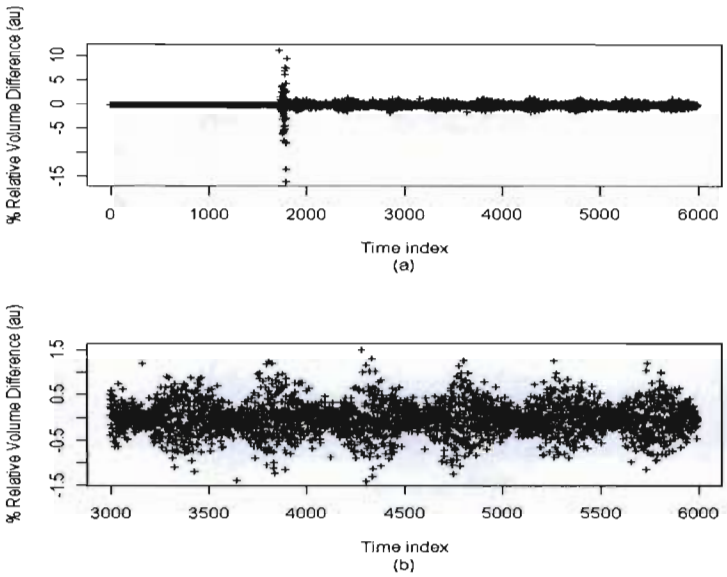
Similar synchronization occurs in other PM applications.

How to estimate tank V at given time?
1) Use previous V measurement and measured {net flow in – out}
2) Use previous V estimate and measured {net flow in – out}
(Kalman filter says use "both", but that assumes zero loss)

Either option: unusual patterns in residuals.

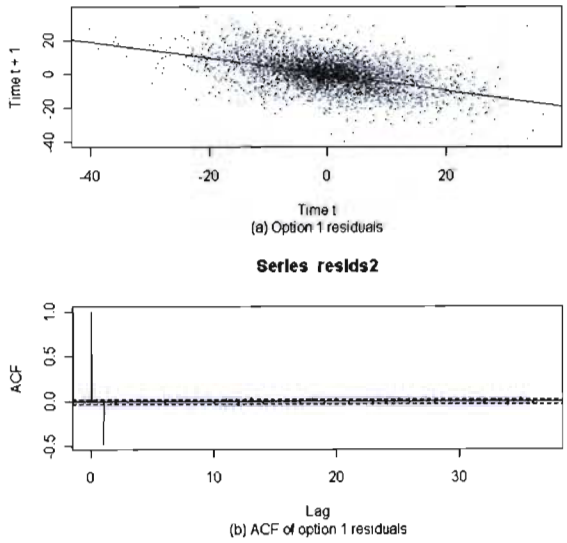
Example patterns in V residuals

Residual = $V_{meas} - V_{estimate}$



Reasons for pattern in residuals:
synchronization effect
Error magnitudes scale with true values (relative error models)

Flow rate monitoring, SM data from B. Cipiti. Two options to generate residuals. Both options result in within-data-type serial correlation.



Example tank has 2 flows in, 1 flow out, and option (1) monitoring results in classic lag-1 moving average serial correlation in residuals, which is not difficult to handle.

Pattern Recognition for Safeguards Data Integration

- Common theme among pattern recognition problems: require "up-front" data analysis, and feature creation prior to the standard n observations of p predictors setup that has inspired considerable research effort.
- Our context: real data (TRP, SRS, ICPP, Barnwell) used to add realism to simulated data, then event marking, residual generation "up front" of pattern recognition.

Current pattern recognition options:

Multiple Page cusum tests

Mahalanobis distance from 0-loss distribution.