LA-UR-12-24305
Approved for public release; distribution is unlimited.

Title: Development of Pattern Recognition Options for Combining Safeguards Subsystems

Author(s): Burr, Thomas L.
Hamada, Michael S.

Intended for: MPACT technical review, 2012-08-28/2012-08-31 (Idaho Falls, Idaho, United States)

Development of Pattern Recognition Options for Combining Safeguards Subsystems

Tom Burr, Michael S. Hamada, Misha Skurikhin

MPACT review meeting

Idaho National Engineering Laboratory

August 28-30, 2012

Work sponsored by MPACT

Abstract: This talk reviews project progress in combining process monitoring data and nuclear material accounting data to improve the over nuclear safeguards system.

LAUR 2012-xxxx

Development of Pattern Recognition Options for Combining Safeguards Subsystems

Focus on 2 subsystems:

- nuclear materials accounting (NMA)

- process monitoring(PM)

Figure of merit: Detection probability DP

DP = P(alarm | diversion scenario)

Diversion scenario:

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

NRC requirements for reprocessing facility

1) NMA: Detect 2 kg Pu within 3 days with 0.95 DP (assume 0.05 FAP but per what period? 10 CFR 74 for PM does not mention FAP)

- Cannot stop/clean every 3 days, so engineering estimates via models/data for in-process inventory →

relatively advanced PM needed to "close balance"

(PM in "support to NMA" role)

2) Diversion path analysis: HOW diverted

Result: high DP for some specified scenarios, small reduction in DP for wide-spread in time/space diversions.

Gaps related to MPACT:

gap 17 re diversion path analysis

gap 18 re requirements for MB magnitude, σ_{MB}, and MB frequency

Relations among: PM ,NRTA, NMA (PM currently enables NRTA)

Talk Topics

- Project overview
- PM residuals:

-many options for generating PM residuals

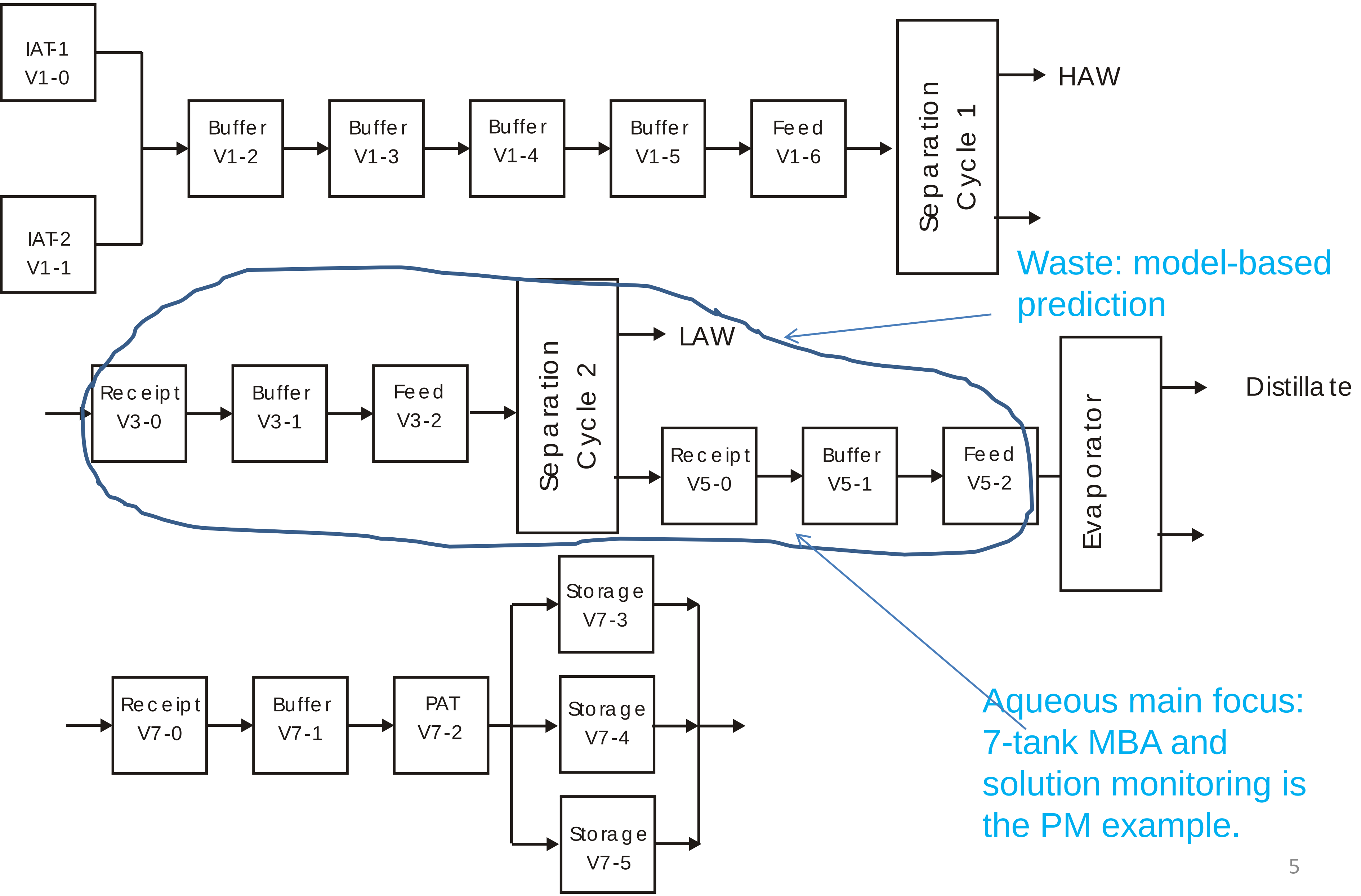
-some non-Gaussian behavior

- Combine PM and NMA residuals

- suggest hybrid of period-driven and data-driven

- FY12 summary
- FY13 plans

Generic aqueous reprocessing, “MBA-2” (SNL, Glasgow, LANL)



Bird's eye view of pyro and aqueous

MBA1: head-end

Aqueous: IAT

Pyro:
SF oxide powder

Relatively large σ_{MB} if rely on reactor calcs for input Pu. Some PM concepts apply. Example: hull monitor, heavy metal balancing

MBA2: main process

Aqueous: Increasing Pu conc in downstream tanks; separations cycles; massive pipework

Pyro: few pipes, mostly batch operations and crane/other visual transfers

→

U ingot
U/TRU ingot
Waste

MBA3: storage

Aqueous could be PUREX, UREX (no declared separation of Pu), other
Pyro: batch, but equipment holdup and salt recycle → "noise source" for NMA and obvious role for PM to support NMA via "ARBI"

Pyro

MBs in NMA:

- by batch
- by campaign with no, partial, or “total” cleanout.

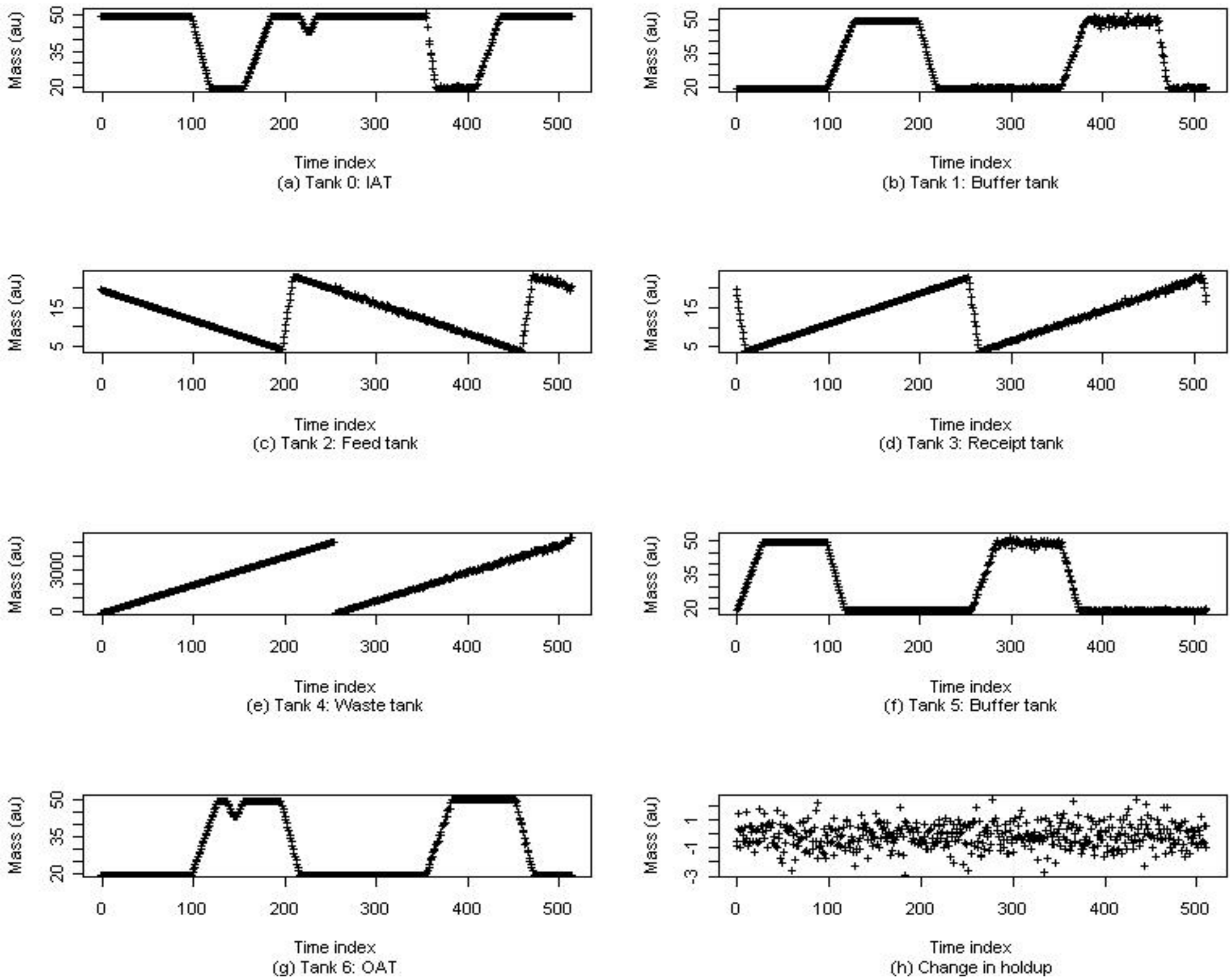
“Total cleanout ” might leave ~1kg Pu in some equipment.

Residuals around unit operations in PM:

- Example: gross weight checks in/out of electrorefiner (ER) where U, Pu and minor actinides are separated from other fission products.
- Example: ER model leads to consistency checks.

Aqueous example: 7-tank SM data
Input, Buffer, Feed, Receipt, Waste, Buffer, Output

One PM option: parse in-tank data into "wait" and "transfer" modes



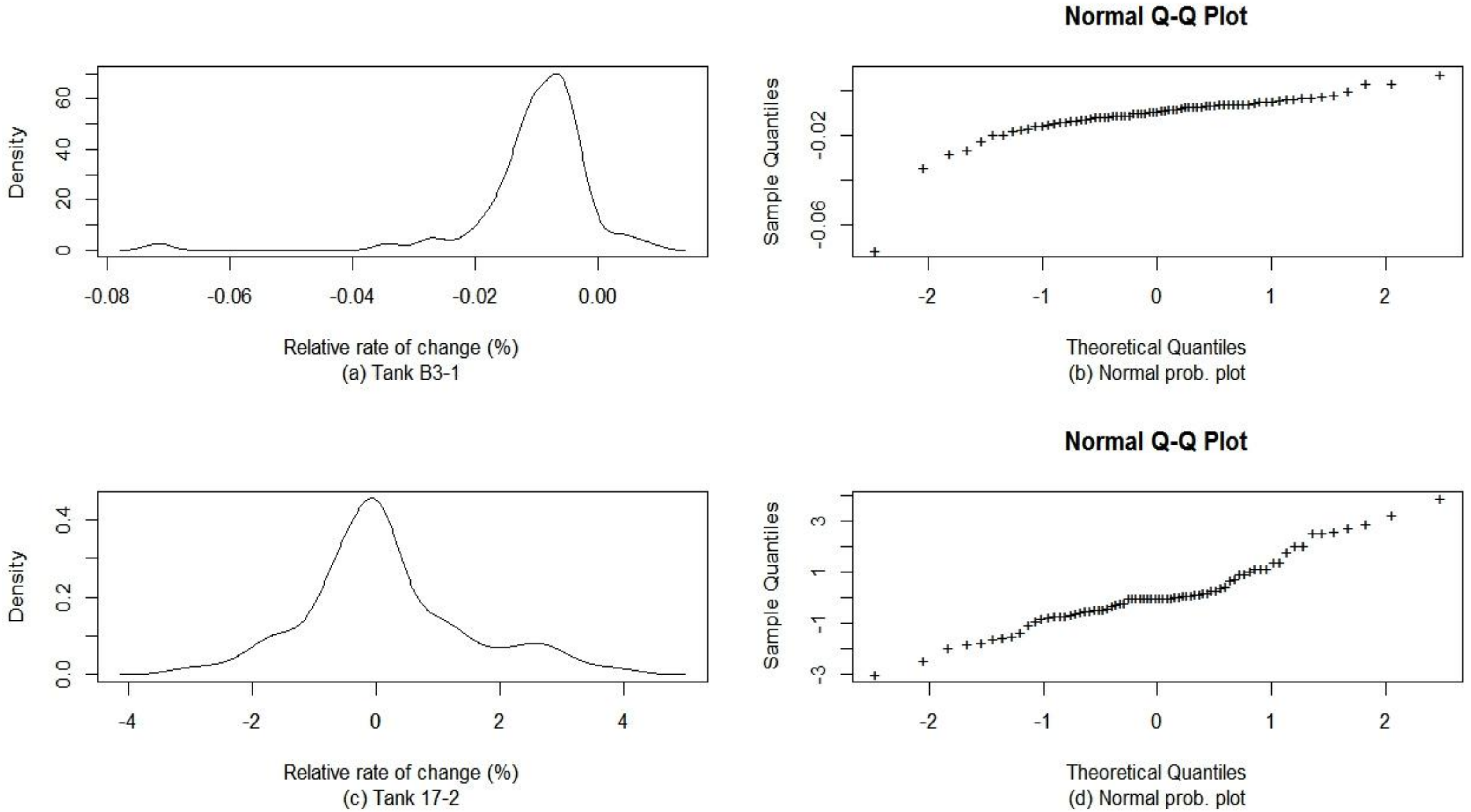
7-tank MBA:
B/B input
B/B buffer
B/C feed
C/B receipt
C/B waste
B/B buffer
B/B output

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

and

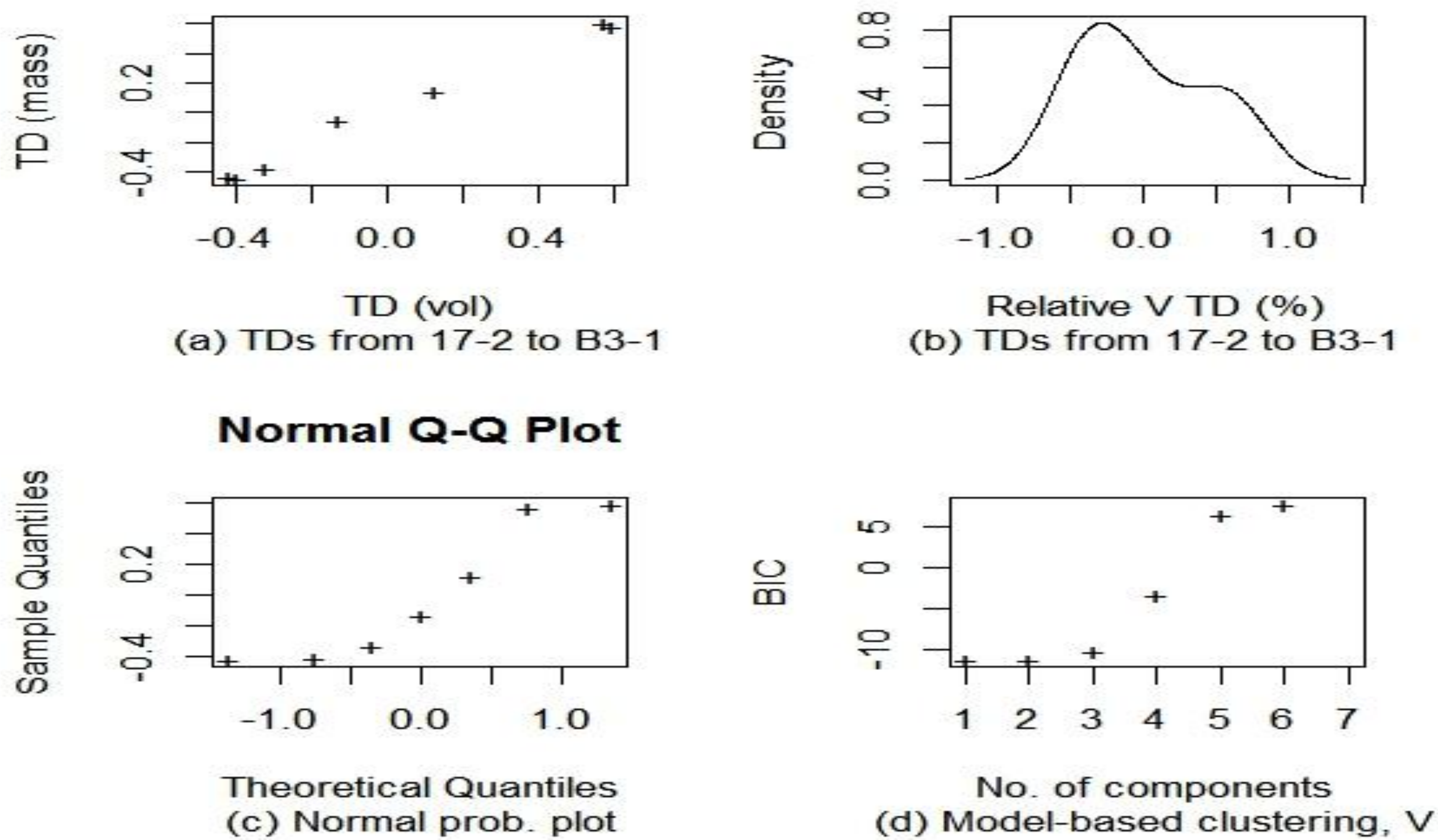
model for Pu flow to waste that provides a book value for waste

Real SM data example
wait mode: mixture distribution, not single Gaussian

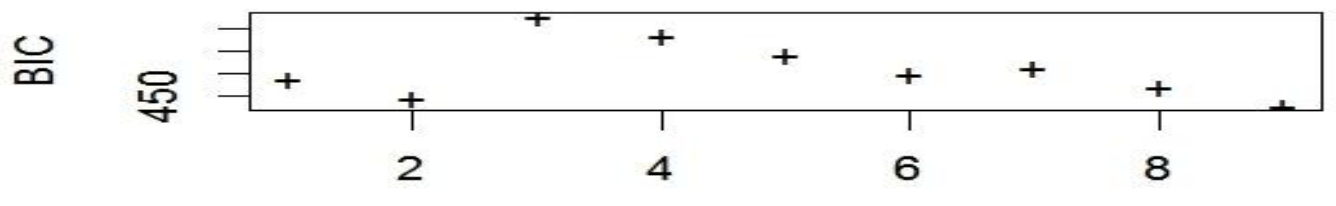


Marked wait modes in SM data from 2 tanks at SRS U storage

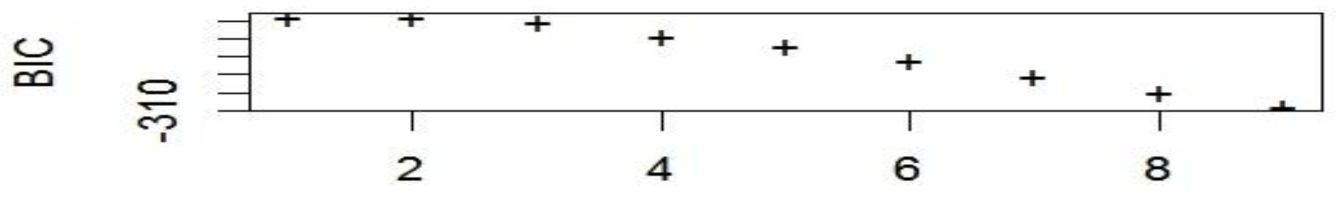
Example 2: Transfer differences in U tanks



Bayesian Information Criterion to **estimate the number of mixture components**

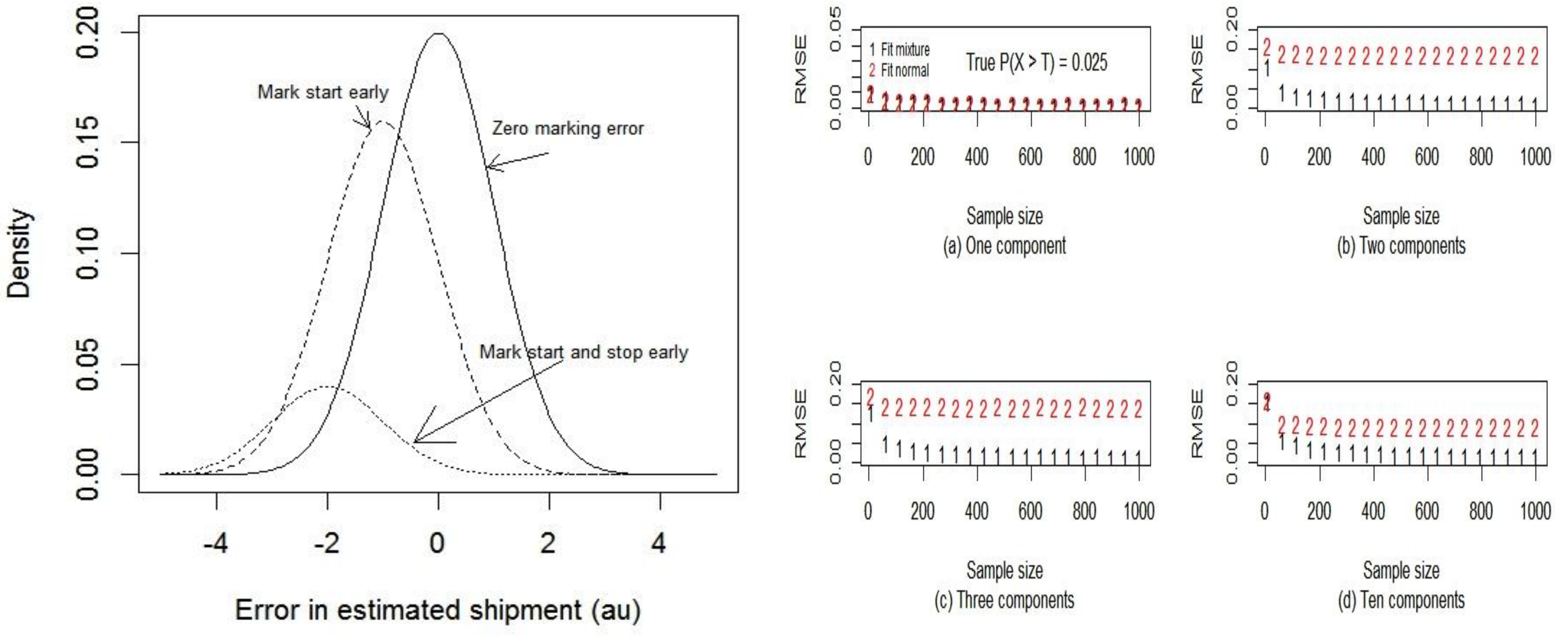


No. of components
(a) Model-based clustering, Tank B3-1



No. of components
(b) Model-based clustering, Tank 17-2

Quantile estimation via mixture fitting



Root mean square error in estimate of tail probability
True tail probability = 0.025 in this example

Batch cross talk in PM residuals: aqueous case

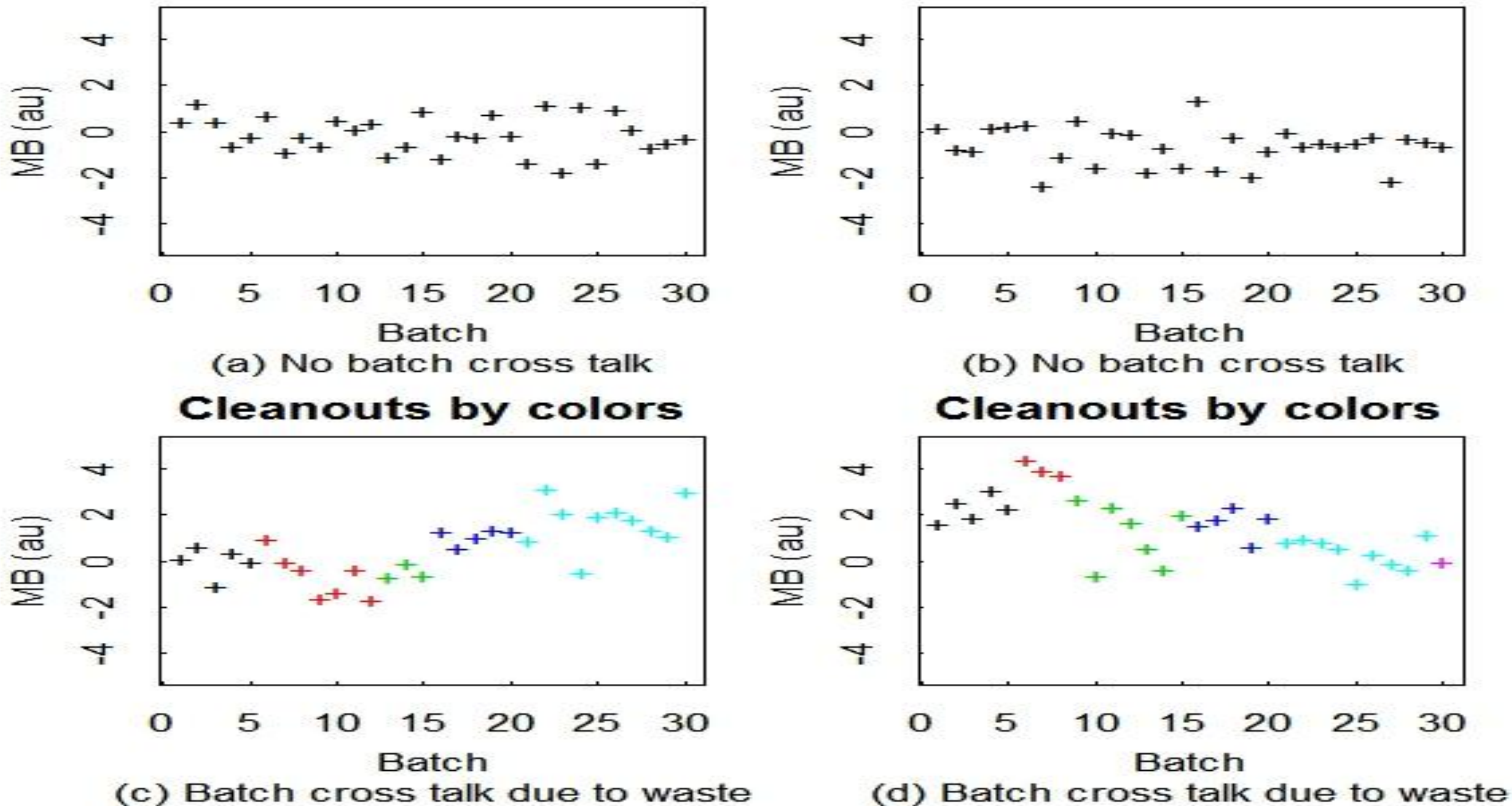
- Pump/pipe carryover → some batches donate to holdup and some batches withdraw from holdup

→ batch cross talk occurs in transfer differences.

→measurement-based batch cross talk also always occurs due to shared systematic measurement error (smaller effect)

→Bottom line: batch cross talk is a "noise source" for PM and the "noise" can be non-Gaussian.

Batch cross talk example
Clean unit every 2-7 batches
Assign same measured waste value to each batch.



Example: ER unit in pyro; HAW in aqueous, TA55 examples in Pu238 (power) and Pu239 (dissolution) processes, FY2013 task

Aqueous: PM on "equal statistical footing" with NMA as another way to generate residuals

New concept 1: PM and NMA generate residuals that can be monitored. The NMA scores are the familiar MBs.

PM example: hull monitoring in head end.

monitor neutrons emitted primarily from the Cm in leached hulls (waste stream) in head end

assume no ability to separate Pu from Cm

New concept 2: use process model to predict Pu in some streams. Examples: dissolver model for amount of Pu to hulls as function of cycle time, temperature, and [HNO_3]; ER model for Pu flow.

RESULT: "nominal" value for Pu in stream from PM based on model-based prediction→ if only diversion scenario is Pu to monitored stream, then a "win" for PM.

Note: cannot meet protracted diversion detection goals using PM as poor man's NRTA which is poor man's NMA.

New concepts 1&2 feed to an old concept (Jones, 1980s) in response to Avenhaus & Jaech

- Avenhaus & Jaech (1981): "not so fast" in response to "more frequent MBs is better."
- B. Jones: Use 2 Page's tests (1 for abrupt, 1 for protracted)

$P_t = \max(0, P_{t-1} + x_t - k)$ Large k, small h smaller k, larger h

- Pattern recognition for specified scenarios using

residuals = measurements – predictions with PM and NMA residuals, some of which are correlated

- Don't reduce DP much for worst case loss spread over time/space ("nod to A&J" but monitor for easy wins such as abrupt loss)
- 0.05 FAP /year or use 20 year ARL for "data-driven"

NMA + PM for aqueous

- Frequent PM residuals (resid = meas – pred in each tank)

- New concept: predicted value for waste streams via model for process units (dissolver, separations areas, evaporator,..)

- PM examples:

{time,temperature, nitric acid conc of each dissolver batch} → predict amount of Pu to hulls

{medium or high resolution gamma detection as in multi-isotope process monitor} → various off-normal operations

{event marking or flow monitoring in tanks} → bulk mass or volume residuals, and possibly Pu residuals if in-line Pu

NMA + PM for Pyro

- PM residuals every "batch" and some more frequently (such as Pu resid = meas – pred in ER)
- Batch "cross-talk" largely due to U and U/TRU behavior in ER and other process equipment.
- "Close batch" across each unit with help from PM.
- Period-driven vs data-driven as with aqueous

Note: Williamson and Willit, ANL 2011 for pyro process description.

FY12 accomplishments

- Pattern recognition project to combine NMA and PM accomplished:

(1) 5 journal articles and 3 conference papers describing candidate options to combined NMA and PM residuals;

(2) extended diversion and misuse models;

(3) introduced a novel concept involving a predicted value for key streams such as high-value waste streams, resulting in high detection probability for specified diversions, and

(4) initial steps toward a combined PM and NMA approach for a pyro facility, although the project has been focused on aqueous facilities.

Pyro report emphasized NMA + PM residuals, just as those on aqueous.

- October 2012 milestone: Pattern recognition options into SNLs SPM.

FY13 activities

1) Document and implement pattern recognition options. Attention to system false alarm probabilty.

2) Participate in LANL's new goal to quantify the benefit(s) of newly-developed PM at TA-55. Use some of the pattern recognition options under development in a case study. COMMON THEME: "cross-talk" between batches makes batch closure "more noisy."

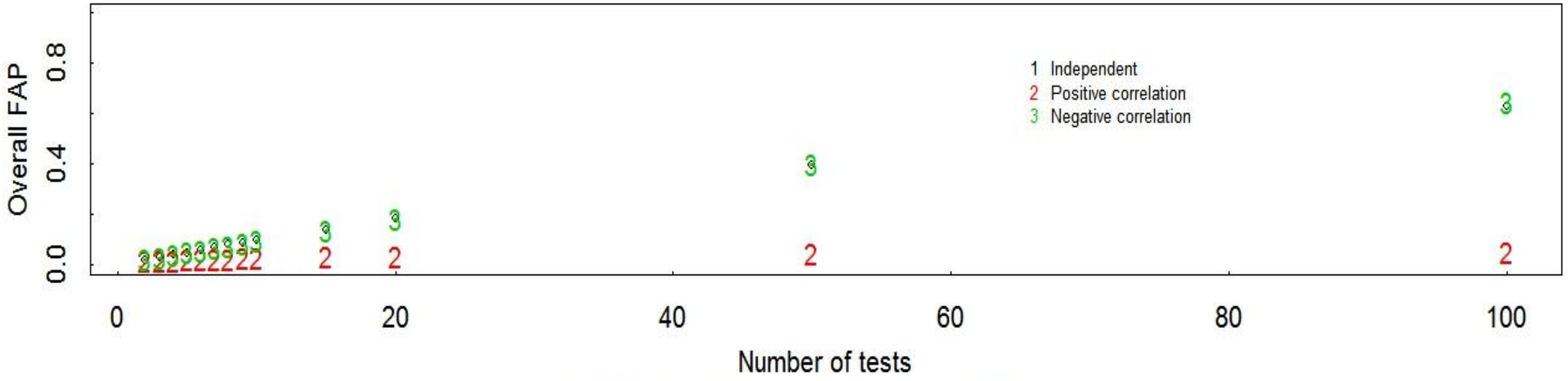
CARROT to facility for PM:

extend from 2 to 6 month physical inventory for NMA balance closure.

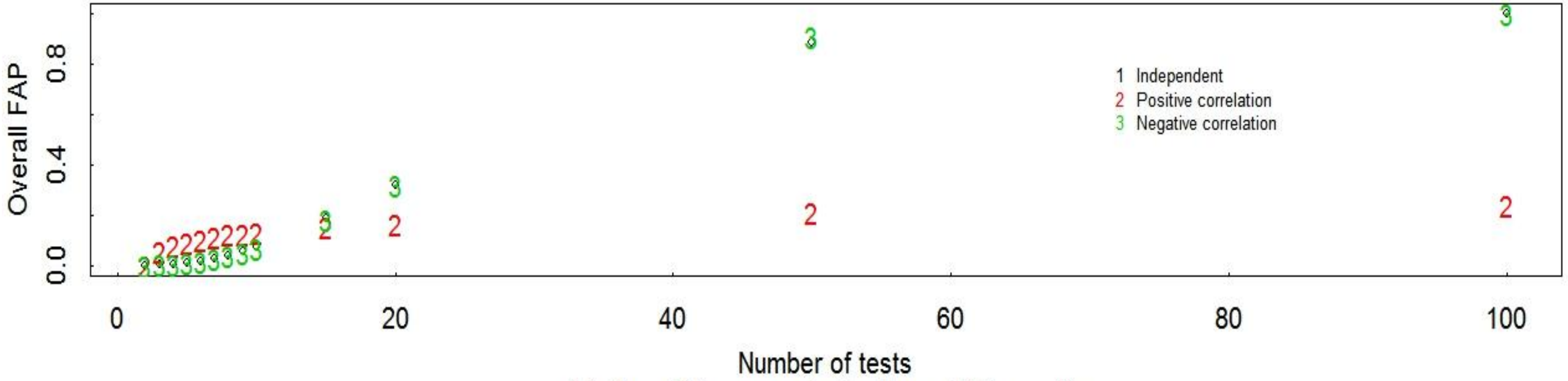
3) Training data requirements to "learn" alarm rules for various PM data types.

4) Uncertainty quantification (UQ) for new NMA and PM data types for NMA and/or PM. Examples: MIPM, microcal

False Alarm Probability versus no. of tests



(a) Alarm if 1 or more tests alarm, 0.99 quantile



(b) Alarm if 3 or more tests alarm, 0.90 quantile

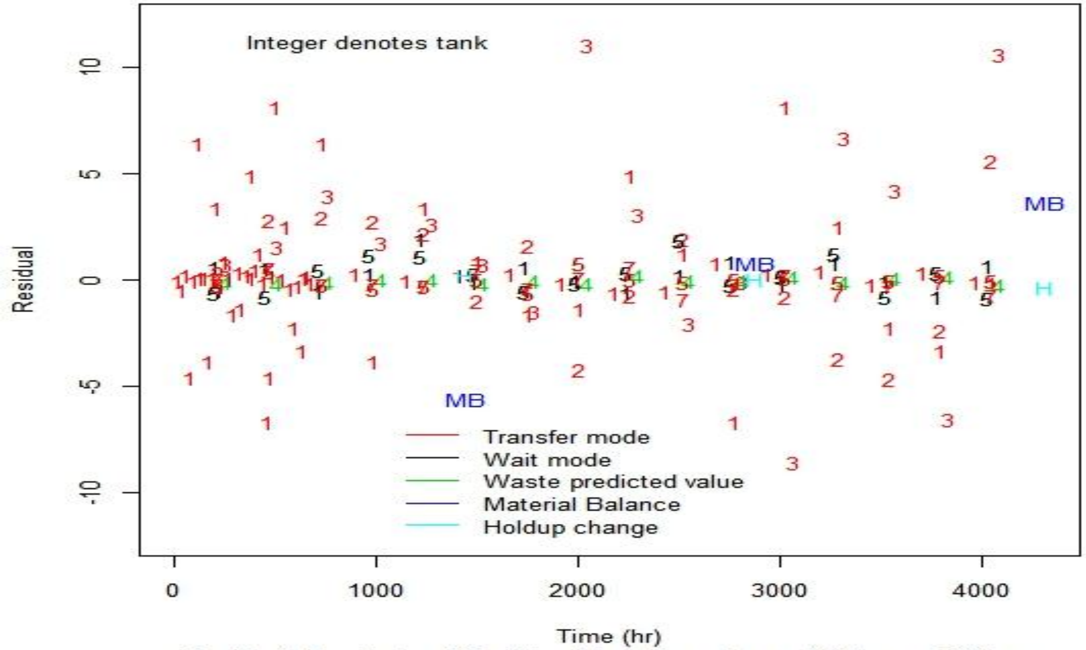
Snapshot of project emphasis

Emphasis: aqueous, but 2012 look at pyro.

Pattern recognition to assess residuals from PM and NMA (MB).

Residuals from simulated data, with NA22 project to add fidelity to simulations

What is probability distribution of PM and NMA residuals?



Residuals from tanks 1-7 wait and transfer modes and MBs over 30 days

Residual = prediction – measurement

NMA:
meas: end physical Inventory
prediction: Book Inventory
MB = Book Inv - End physical Inv

Book Inv = $T_{in} – T_{out} + Inv_{begin}$

PM: depends on facility type and choices made, but examples are:

Aqueous:
tank-to-tank transfer difference
mass change during "wait" mode
Dissolver model → hulls prediction

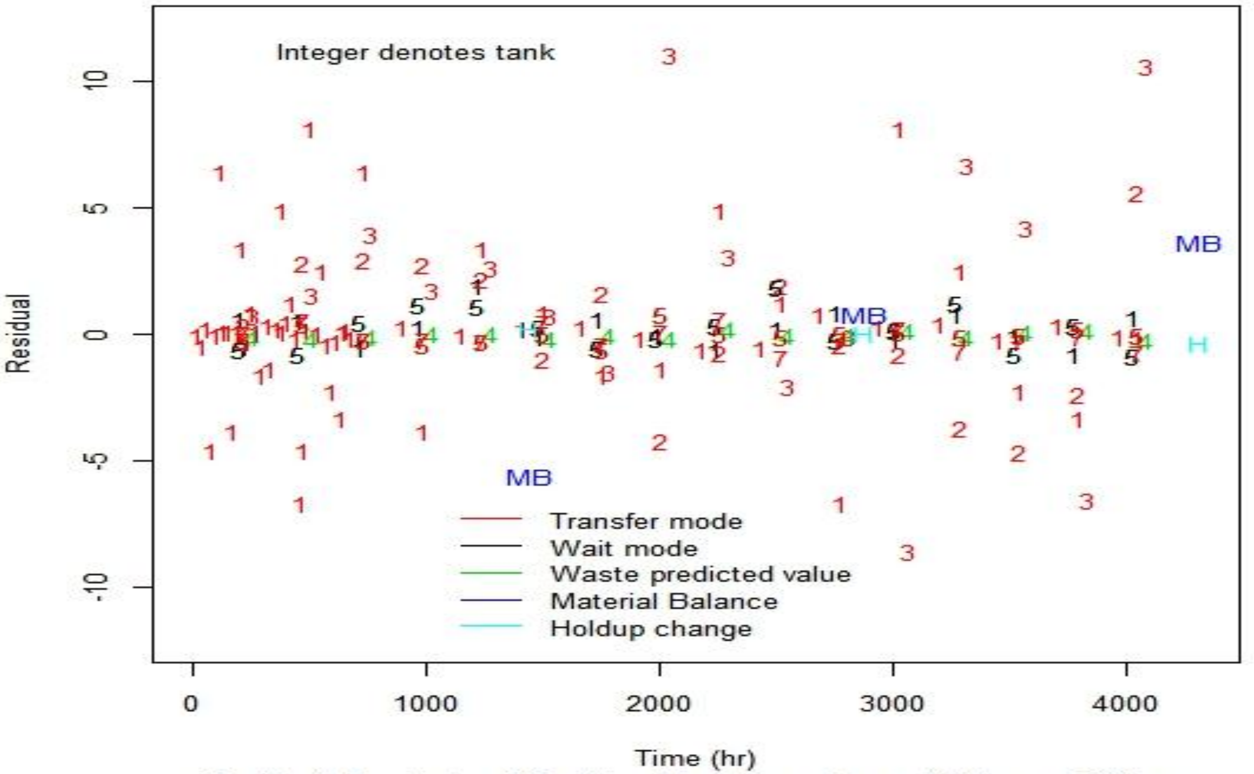
Pyro:
predict – meas of Pu mass in electrorefiner (ER)

PM residuals: (1) detect off-normal operation, or (2) related to SNM flow

NMA residuals: the MBs

Comment regarding holdup:
NMA: impacted by how holdup measured.
Direct or "by-difference".
PM: could have separate residual stream

PM residuals for off-normal operation:
- nitric acid concentration
- flow rates
- multi-isotope PM analyses



Residuals from tanks 1-7 wait and transfer modes and MBs over 30 days

PM residuals for SNM flow:

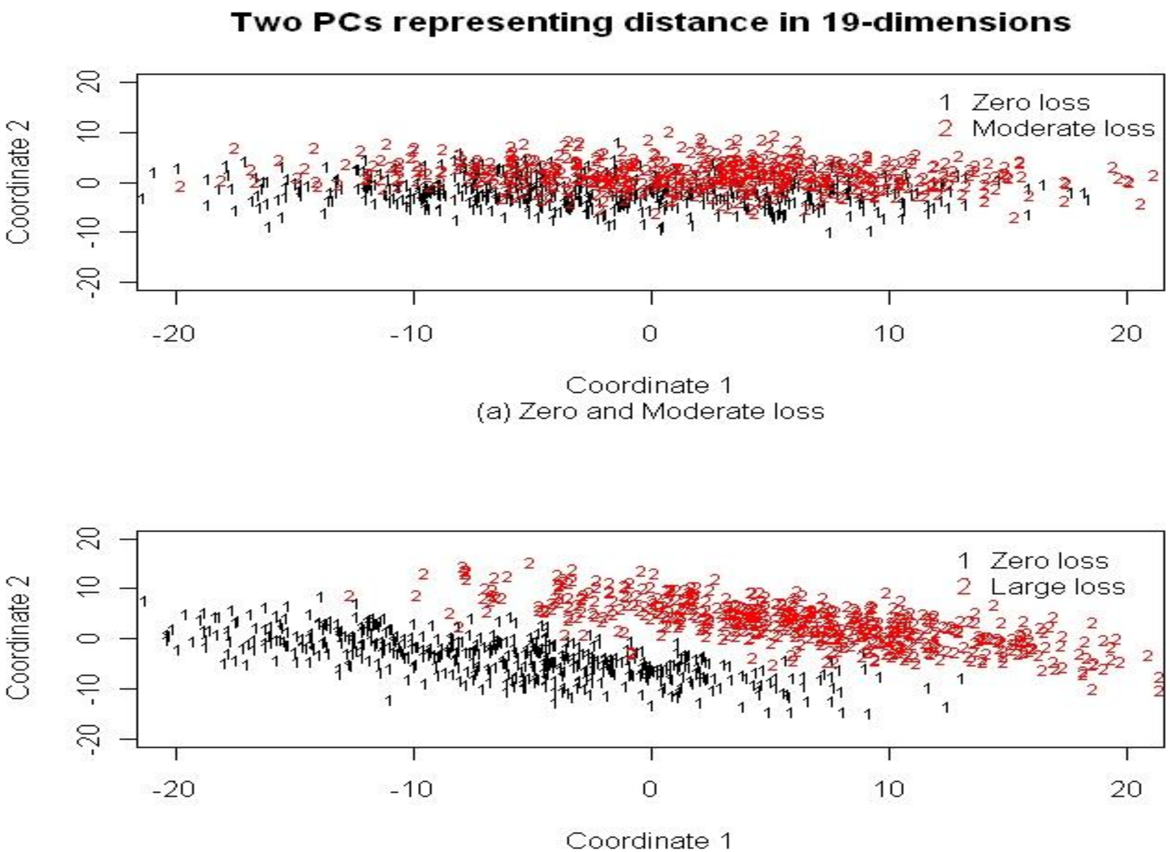
Option 1: material balance every 6 min., and

Option 2: event marking

Shown here:
Related to SNM flow using option 2 PM and Model-based prediction of SNM to waste and in holdup

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

Simulated realizations plotted are from the corresponding 19 Page tests, each applied over the 30-day period → "period-driven" pattern recognition



(a) Zero and Moderate loss

(b) Zero and Large loss

Moderate loss is approx 1% of 30 day thruput

Large loss is approx 3% of 30-day thruput

Alternate, data-driven:
Apply Page's cusum to each of 19 components:
10 from: 3 wait regions for tank1, 2 wait for tank 5, rec1, rec2, rec3, rec5, rec6
3 from mbseq, choldup, wastebook. Alternate: 33 = (4 wait,+ 4 transfer) *x* 3 + 9

Principle coordinates:
if use Euclidean distance, principle coordinates same as principle components.

DPs via MD for "vanilla" simulated data
No process variation, only ran + sys measurement errors

Loss Scenario	False Alarm Probability		
	0.01	0.05	0.10
Loss from Tank 7			
Small	0.02	0.07	0.12
Moderate	0.05	0.21	0.32
Large	0.50	0.85	0.92
Loss from Tank 7*, Use: 3 MBs only			
Small	0.12	0.06	0.10
Moderate	0.10	0.17	0.25
Large	0.29	0.50	0.64
Loss from Tank7*, Use: wait mode in tank7			
Small	0.08	0.08	0.12
Moderate	0.27	0.50	0.60
Large	**0.93**	**0.99**	**1.0**
Loss from all tanks			
Small	0.01	0.05	0.10
Moderate	0.01	0.07	0.12
Large	0.01	0.07	0.12
Very large	0.04	0.32	0.55
Loss from all tanks* Use: 3 MBs only			
Small	0.01	0.05	0.10
Moderate	0.01	0.05	0.10
Large	0.01	0.05	0.12
Very large	**0.88**	**0.95**	**0.98**
Loss from all tanks*, Use: wait mode in tank 7			
Small	0.01	0.05	0.10
Moderate	0.01	0.05	0.10
Large	0.01	0.05	0.11
Very large	0.01	0.05	0.11

Note: substantial effort to generate PM scores, NMA (MB) scores, and do various pattern recognition options. Updates ongoing.

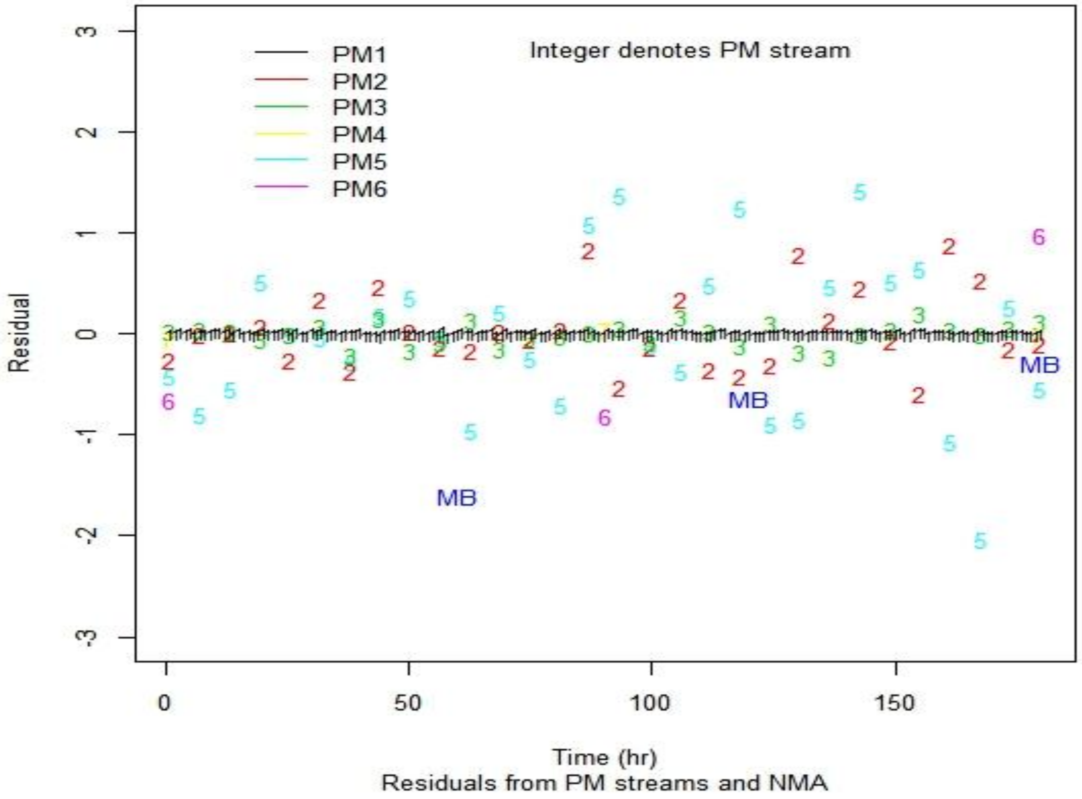
A&J: NRTA will not help detect small widespread loss

Response 1 to A&J: unless the loss is from stream having low nominal Pu and a predicted value

Response 3 to A&J: more frequent MEASUREMENT reduces random error variance. Issue is how frequently to test

Response 3 to Avenhaus and Jaech's "wait a minute:"
PM will dramatically help detect abrupt loss, and can be designed to very minimally hurt DP for widespread loss.
Here, FAP fixed at 0.01, 0.05, or 0.10 per 30 days, so "period-driven."

Pyro cartoon
PM and NMA residuals for illustration, not realistic.



PM1: frequent Pu conc residual from ER model

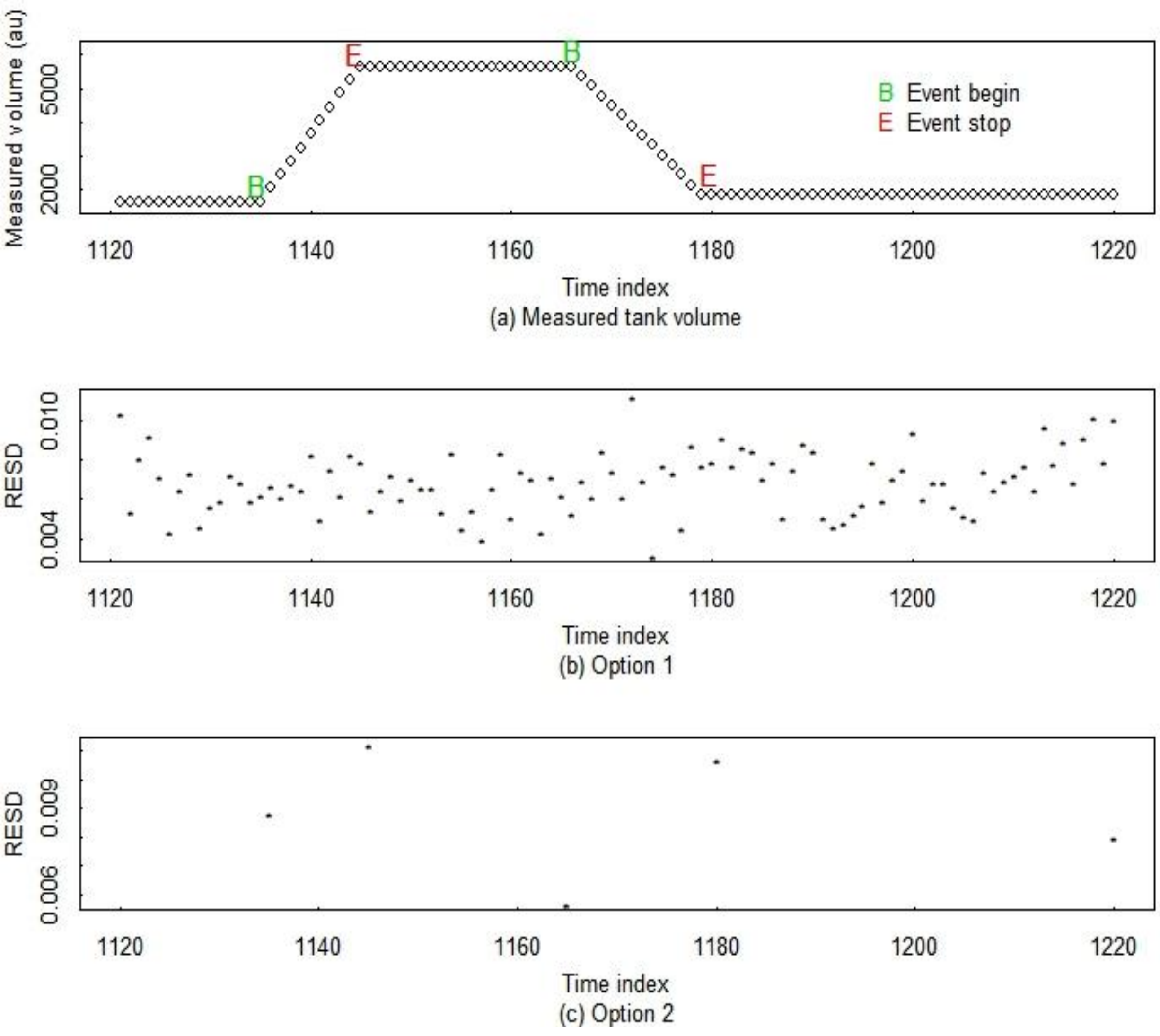
PM2-PM6: other PM streams by batch or sub-batch

MB = material balance every few days (or by batch)

Invoke: models and predictions → residuals

Note: need to know/measure spent fuel properties for ER model and dissolver model → ambitious modeling

Bulk volume or mass residuals: 2 options

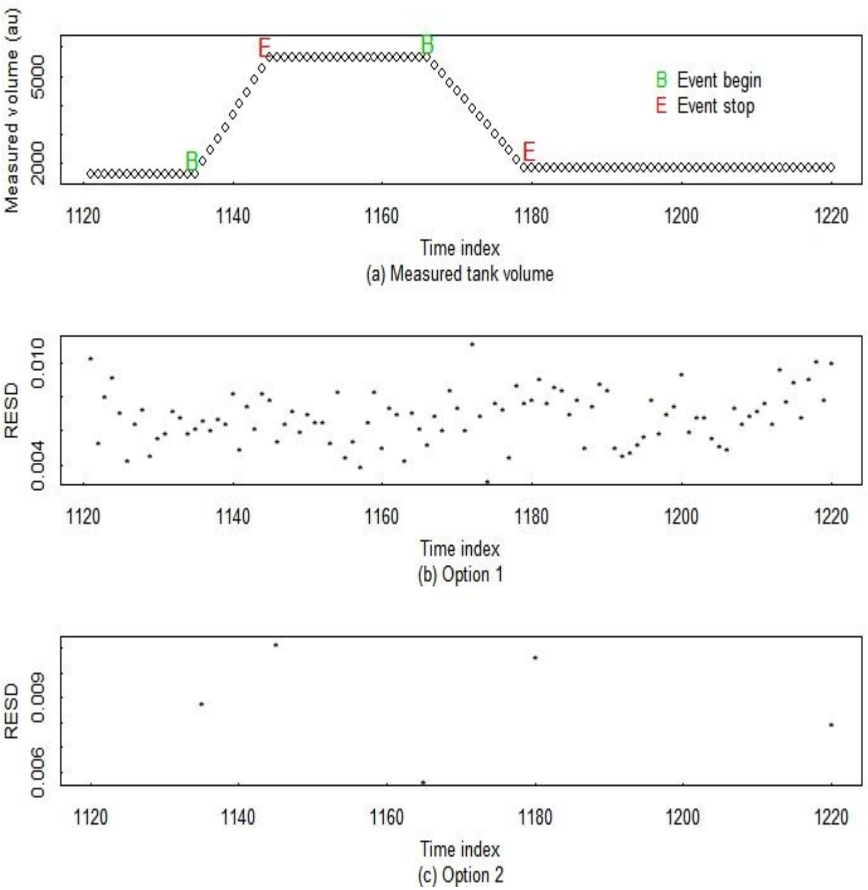


Option 2: event marking → parse volume time series into tank “wait” and “transfer” modes

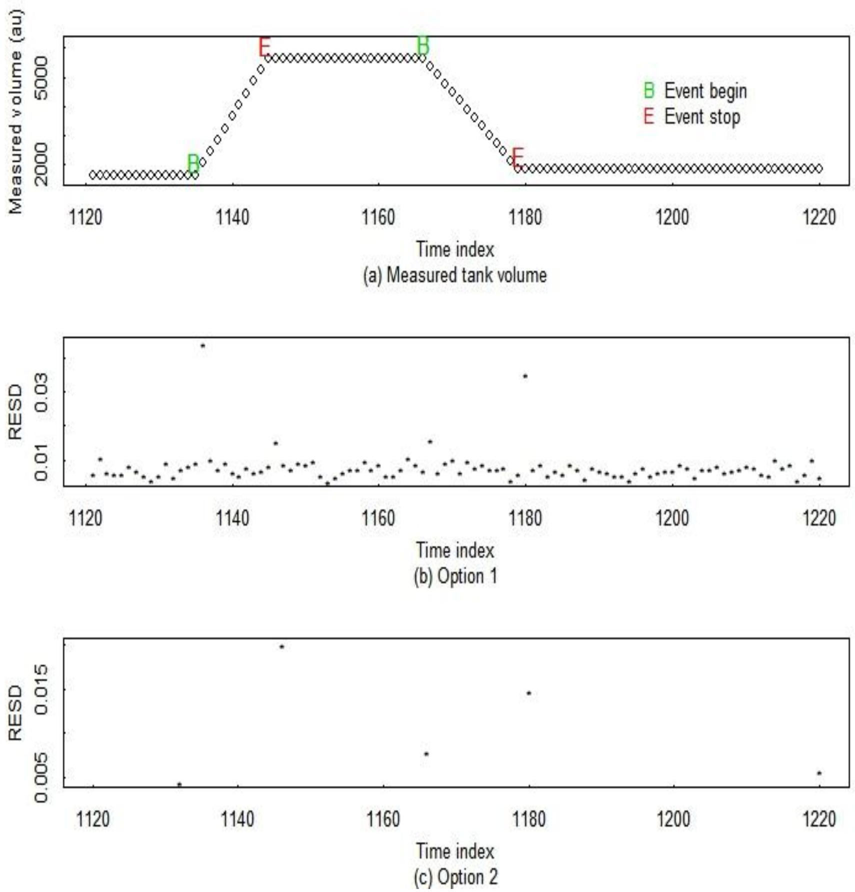
Option 1: measured flows and in-tank volumes every 6 minutes, used by SPM from SNL → residual every 6 minutes from each tank

Option 2: event marking as in several solution monitoring and evaluation systems → Residual after every wait and transfer mode.

Bulk volume or mass residuals: Option 1

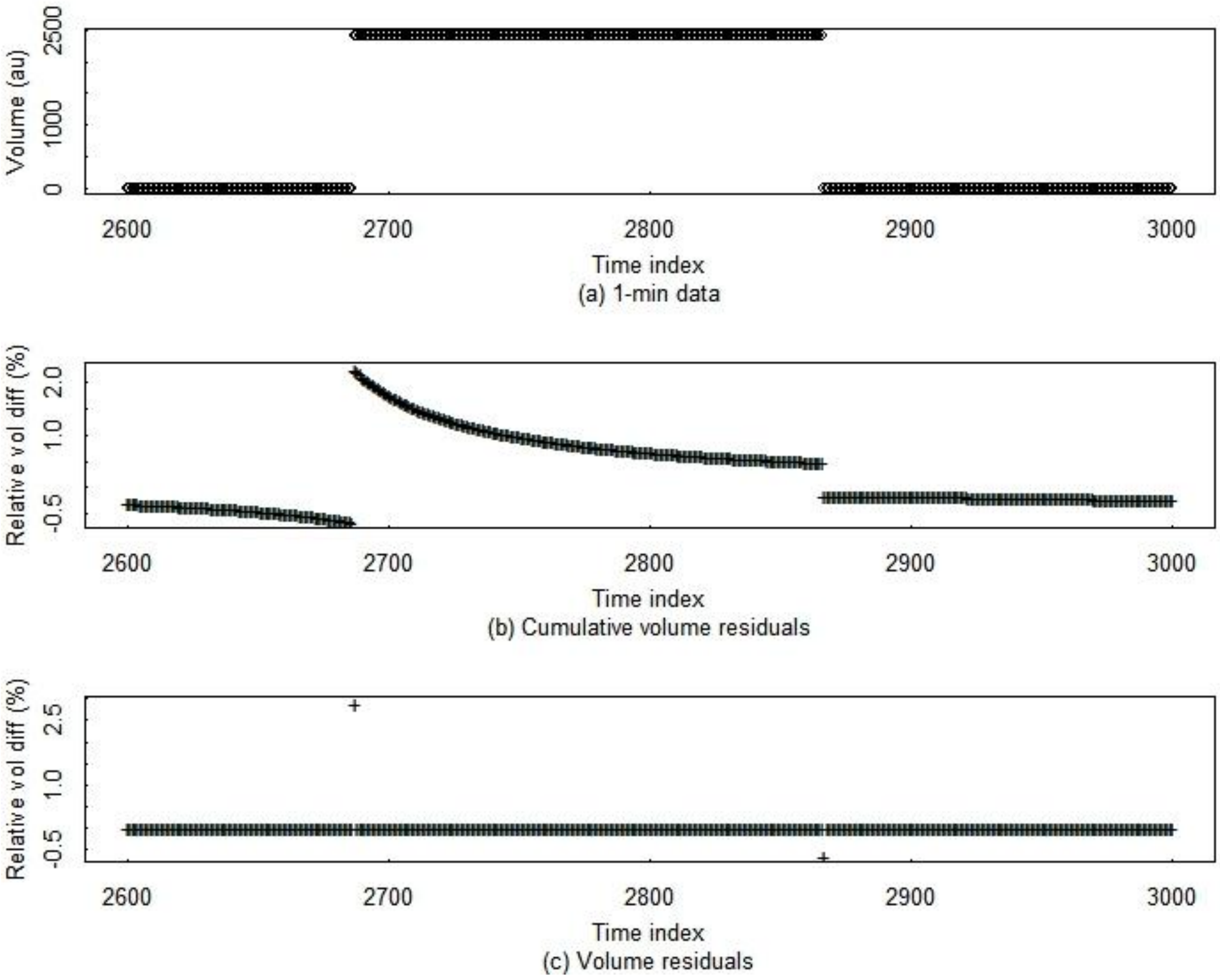


Zero offset



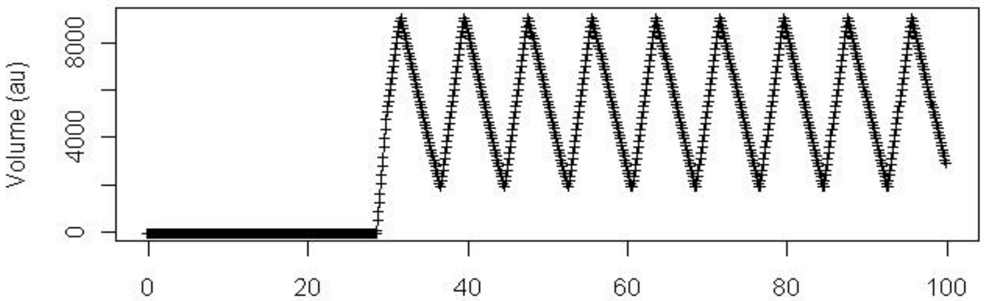
Nonzero offset: 20% of reporting interval

Synchronization error: offset between data recording times and flow rate change times

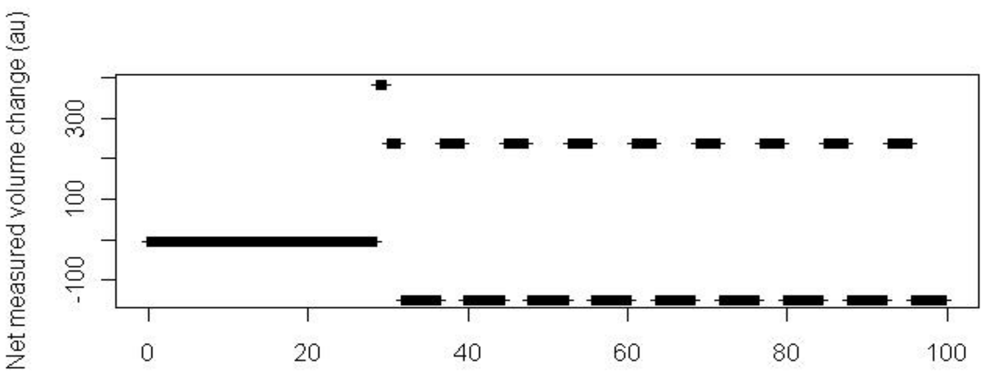


Use measured flow rates as recorded every few minutes to predict in-tank volume.
(a) in-tank volume
(b) raw residuals
(c) cumulative residuals

Example data from SPM: in-tank and flow measurements leading to residuals every 6 min. from each tank



(a) Zero measurement error



(b) V change measured by net flow rates with snychronization effects

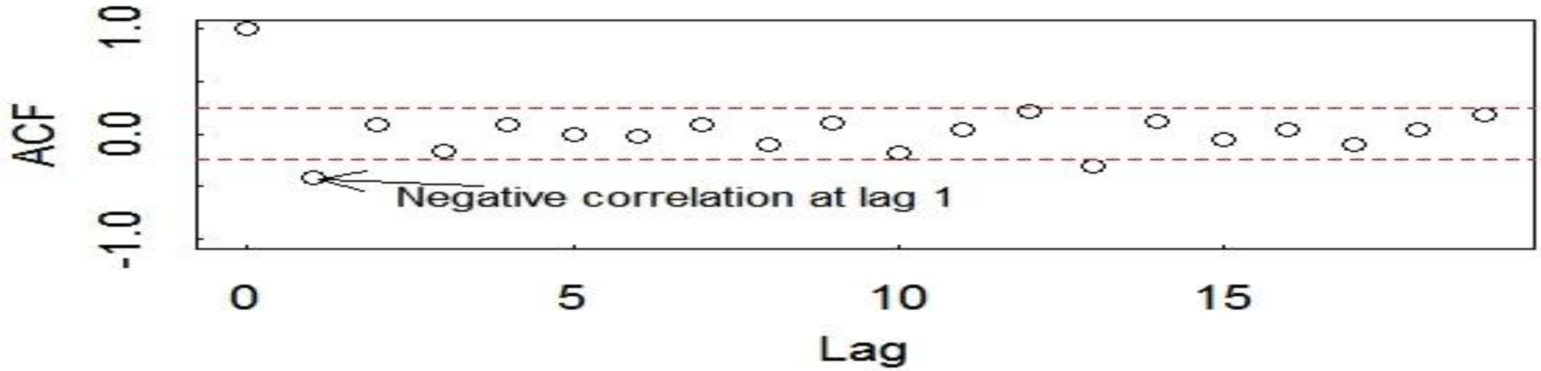
Ex. of synchronization effects: flow rate changes at unknown time between minute 101 and 106, and simulated {V,M} in tank knows exactly when flow changed, but only observe the instant flow rates every 6 minutes.

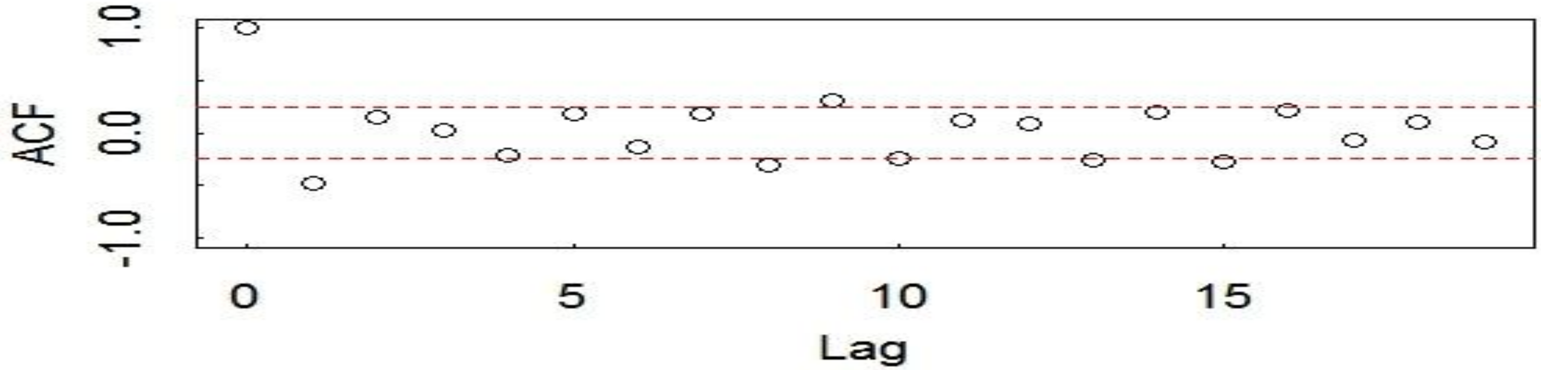
Similar synchronization occurs in other PM applications.

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)

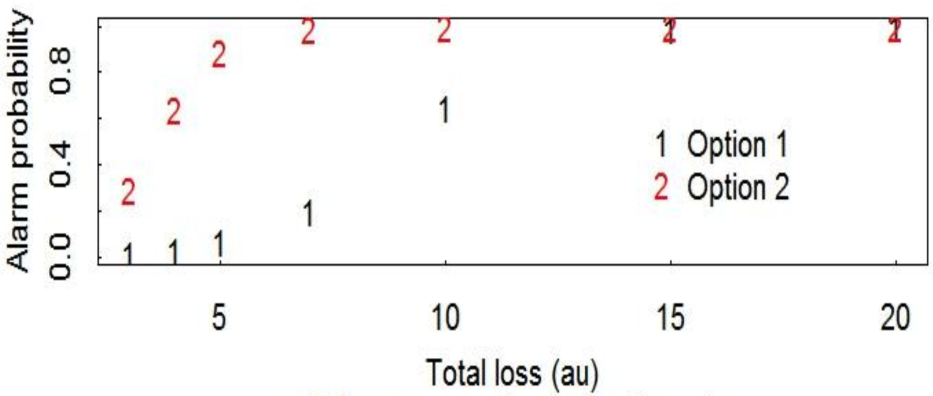
Either option: unusual patterns in residuals

Serial correlation in PM residuals
Analogous to lag-1 serial correlation in MB sequences

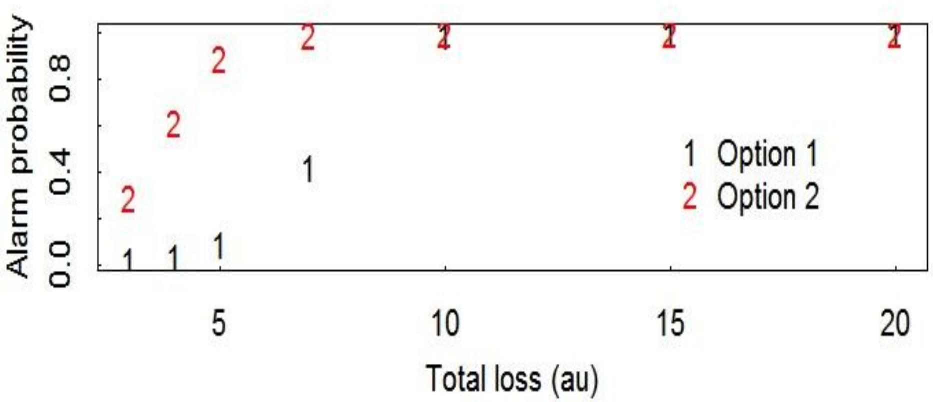




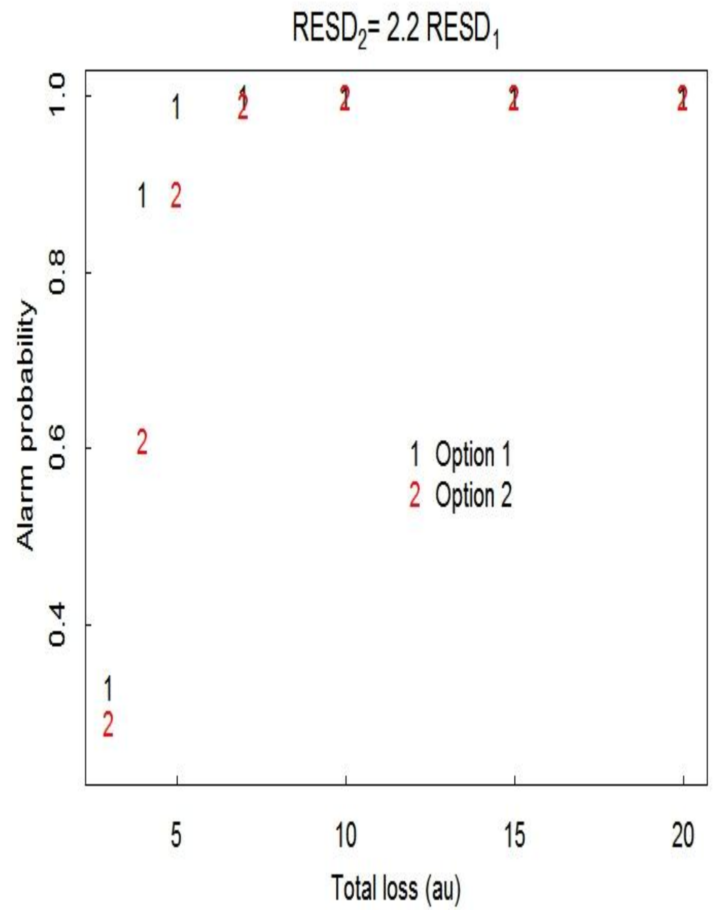
Bulk volume or mass residuals: 2 options



(a) Loss over entire first wait mode



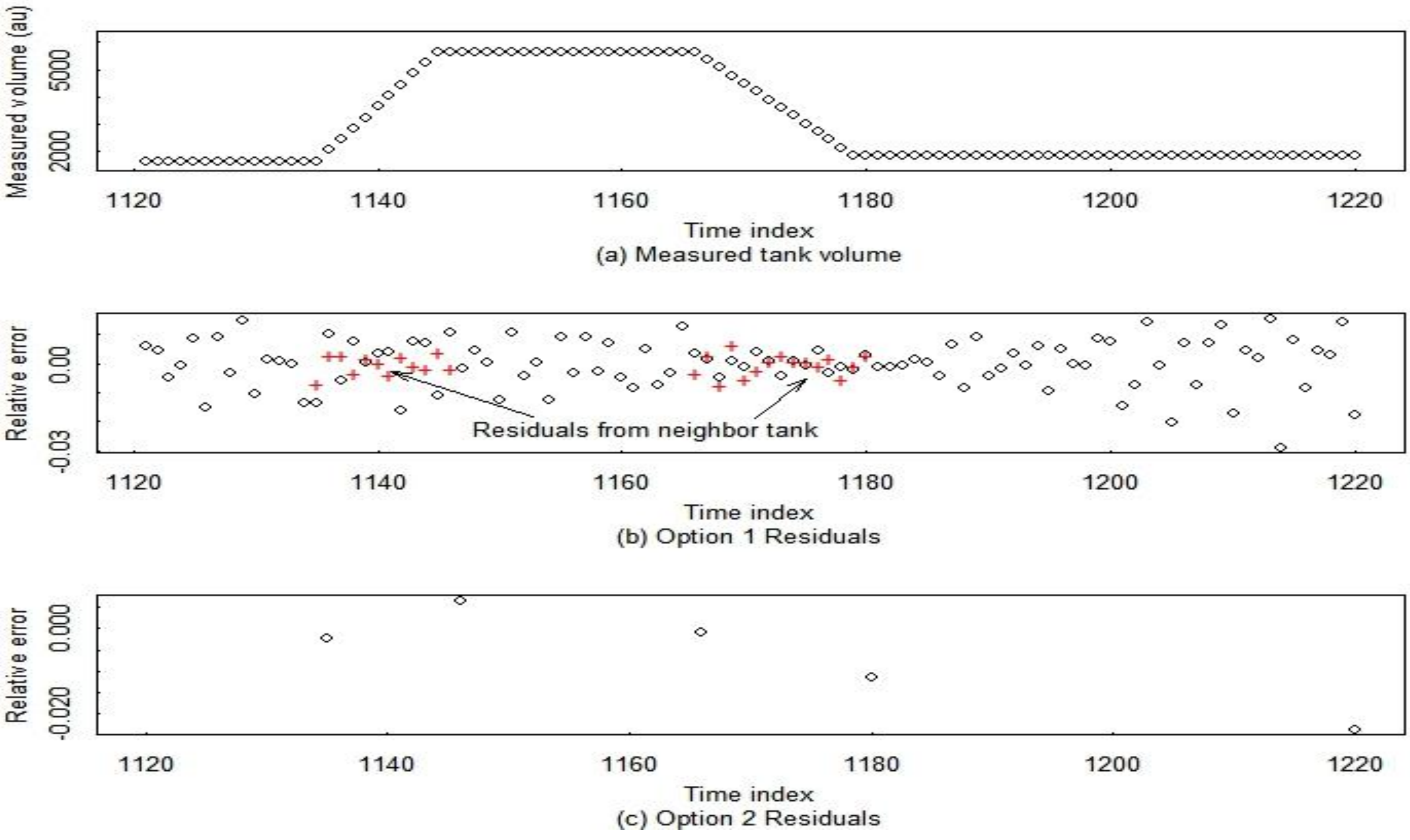
(b) Loss at one time index during wait model



(a) Loss at one time index during wait model

RESD = Relative error standard deviation in residuals

Residuals must involve neighboring tanks in options 1 and 2



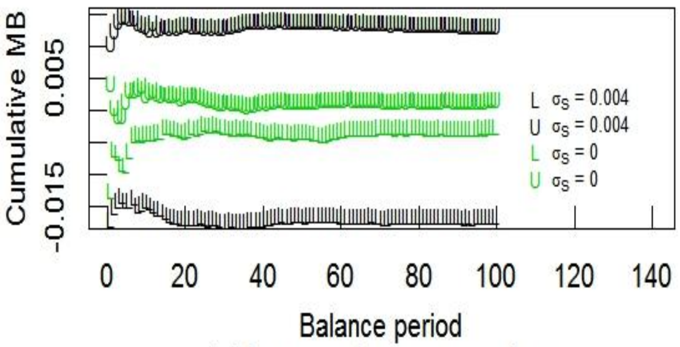
(a) Measured tank volume

(b) Option 1 Residuals

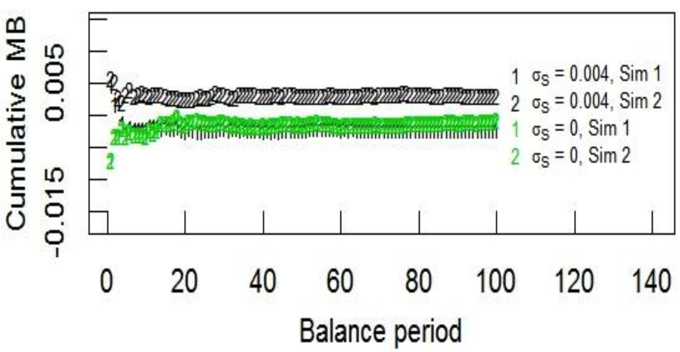
(c) Option 2 Residuals

PM is data rich →bias adjustment?

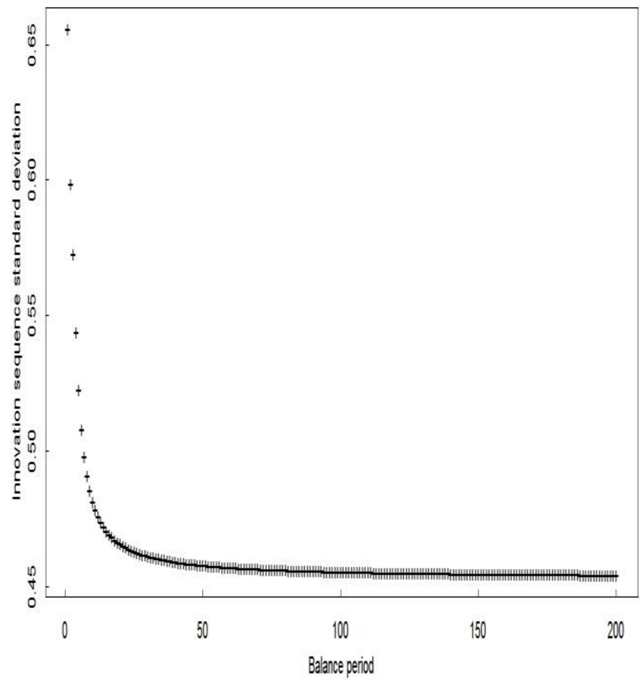
Common example:
Meas = True + Random + Sys for particular assay method →
$Meas_{biasadjust}$ = Meas – Avg(Meas-True)



(a) Lower and upper excursions



(b) Two random simulations



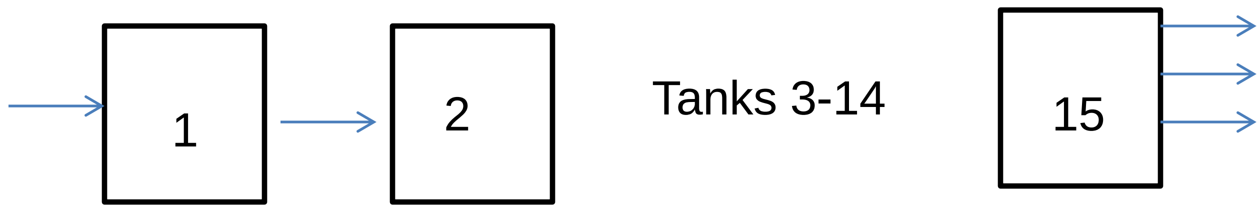
SITMUF variance is smaller than MUF variance. Caveat: assume zero loss during "training"

SITMUF transform regarded as specialized bias correction that reduces variance in NMA residuals.
$Residual_t = MB_t – E(MB_t|MB_1, MB_2,…, MB_{t-1})$

Consider using 30 days of 6-minute data to estimate overall bias in each tank's bulk mass or volume

15-tank MBA in SPM

1 flow into tank 1.   3 flows from tank 15



In-tank bulk V and M → "poor-man's NRTA" with balance closure every 6 minutes.
Extend to rich man's NRTA with in-line Pu conc or off-line Pu conc and mixing rules.
Extend to individual tank monitoring → PM and NMA example.

Error propagation to get $\Sigma_{n-\text{by}-n}$ for sequence of n balance closures that:
-ignore internal flows
-use the 1 flow into tank 1 and from tank 15, and 15 inventory changes

-FY2012 Task discovered numerical issues in converting from MUF to SITMUF sequence via Cholesky decomposition for some covariance matrices.

Support to B. Cipiti's MPACT performance model

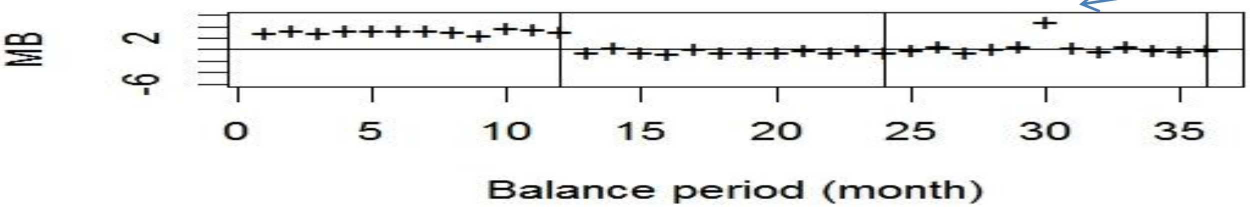
- Matlab/simulink model: NMA residual streams in each tank, assumed pipe flows in/out and in-tank volume/mass/Puconc
- Statistical issues:

- error propagation to get covariance matrix of $MB_1, \ldots, MB_n$ for each tank.

- control of "family wise" false alarm rate by brute-force simulation to set alarm thresholds
- Implement any reasonable sequential test, such as Page's test (long history in NRTA for safeguards), in period-driven fashion. Anticipate hybrid of period- and data-driven.

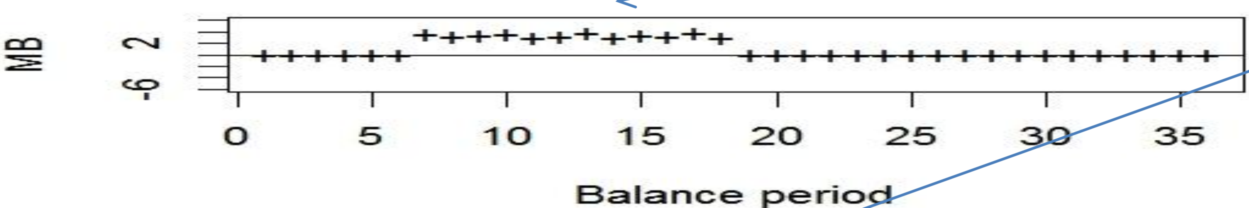
DPs via cumuf, page, scan

Close balance every 12 months → "period-driven."
Common criterion: detect 3.3 σ_{CUMUF} loss with DP 0.95



Use 2 Page's test:
1 with small k for abrupt loss
2 with larger k for protracted loss

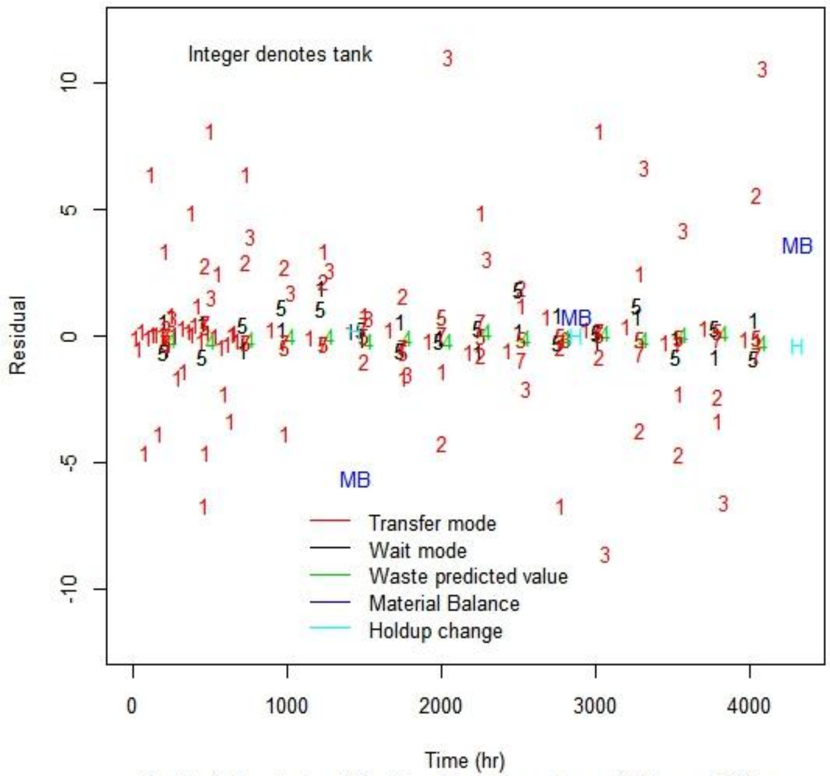
Unknown start time for 12 consecutive losses



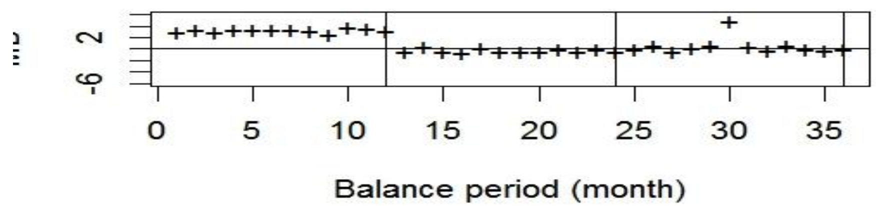
Any loss pattern is possible. A&J considered "worst case" loss pattern.
Page is often "2nd best" sequential test for wide variety of loss patterns.

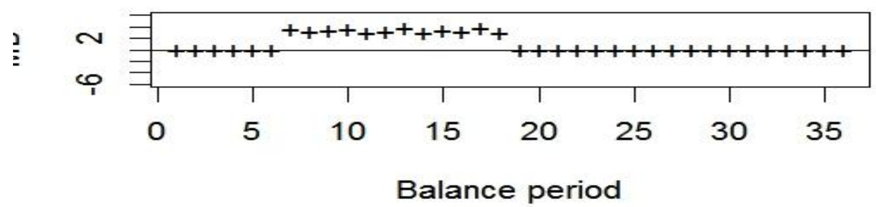
If allow unknown start time of loss sequence:
CUMUF: 0.95 → reduced to 0.75
Page's test: 0.79, but range of results (truncated at 5 yr, 0.05 FAP per yr, or could use avg run length criterion) depending on Σ_{12x12}
Scan test: 0.94 (scan statistic on "SITMUF")

Hybrid of data-driven and period-driven testing



Residuals from tanks 1-7 wait and transfer modes and MBs over 30 days





Suggest for PM and NMA residuals:

period-driven testing (make decision to alarm or not) every *year* AND carry along vector-valued residual across years for data-driven
Precedent in NMA context is using 2 Page's tests.

PM benefits and summary

Possibilities for PM roles:

1) NMA remains objective/quantitative basis for AP.

PM used to resolve alarms, support error models, support NMA (Example: adjusted running book inventory via PM to deal with unmeasured inventory in pulsed columns (aqueous) or ER (pyro)

2) PM in driver's seat to trigger physical inventory taking: TA55.

3) PM and NMA on "equal footing." This project lives here.

- Scores from NMA and SM are not independent
- Must control false alarm rate → "understand" process/data.
- OK to tune to a few loss scenarios, but include catch-all anomaly detection for scores "unlike anything seen in training."
- FOM: P(alarm|diversion scenario), easy win for small waste streams
- Hybrid of data-driven and period-driven