What’s new
SIMCA®15 Multivariate Data Analysis Solution
June, 2020
Overview

- Time in analysis of continuous process data
- Managing recurrence of batches and phases in batch processes
- SIMCA-online synergies
- SIMCA Import updates
- Minor fixes
Time in analysis of continuous process data
Time consideration in continuous processes

Changes has been made to the following functionalities:

- Lags
  - in relation to time
  - in relation to speed/distance

- Introduction of Leads (negative lags)

- Control charts grouped in time range

- Cross validation groups based on time range
Lag on time and “distance”

- **Time** = constant time lag
  - Relevant if process speed is static
- **Distance** = constant distance, varying time
  - Relevant if process is dynamic/changing speed
  - Distance the material travels through the process

- Constant speed of 1 m/s → **Time lag** of 10 s A to C and 7 s B to C
- Varying speed → **Distance lag** of 10 and 7 m respectively calculated from Time and Speed tags
Lags

- New design of Lag page in Workset
- Lags can be based on a time variable or any variable
  - Observation index = how SIMCA 14.1 and previous worked
  - Time variable → Fixed time structure
  - Variables → Fixed “distance” structure
    - Requires a time variable. A distance is calculated meaning that the lag time adjust by changes in process speed.
- Lags in relation to variable / distance
- For guidance, median step length is given
  - For Time and Distances
Leads are negative lags, i.e. lags into the future.
• Useful when creating models for SIMCA-online Control Advisor.
• Mark the Is Lead Checkbox to switch from Lag to Lead mode
• Lags and Leads can be combined in the same project
  • Leads with -time
• Due to their nature, leads are set as Y variables
Control charts in Time or variable range

- Group on observations, time or another continuously increasing or decreasing variable.
- Ex Shewhart Mean/Std.Dev.
  - Left: Group size 10
  - Right: Group range 24h
  - Group range unit relates to parsed time unit at import
  - Group size vary, statistics are adjusted. Std.Dev. cannot be calculated when group size <2
Cross validation groups on time

- Cross validation groups must be changed prior to model fitting
- Access from Model Options
  - Right click on model or Options.. button in Model window.
  - CV-groups tab
- Variable range in dataset shown for guidance
- Max range is 99 hours
- Number of CV groups cannot exceed number of variables in dataset
Customer challenges addressed

- Lagging based only on observations can be erroneous
  - As in SIMCA 14.1 and previous
  - Gaps in data was disregarded, actual lag structure in relation to time unknown

- Need support for changes in Lag structure as process dynamics vary
  - Changed speed of process $\rightarrow$ change in lag time

- Control charts should optionally be in relation to time ranges
  - Grouping by observations may hide time gaps in data

- Cross validation difficult to trust
  - Even with options for group similar etc. it was unknown if CV groups spanned over varying time gaps
Managing recurrence of batches and phases in batch processes
Managing recurrence of batches and phases in batch processes

- Two main cases:
  - The complete batch is processed in one phase more than once
    - Failure to meet Q / reworks
    - Processing failure
  - The batch is divided into sub-parts (sub-batches) which are separately processed
    - In parallel in different process units
    - In series in same process unit
Data structure requirements

Process setup

- **Data structure**
  - **Hierarchy:**
    - **Batch**
    - **Phase**
    - **Phase iteration**

![Diagram of process setup]

Batch 1

- Unit 1
- Unit 2
- Unit 3
- Unit 4

Phase 1 and phase 2  Phase 3  Phase 4

Batch 2

- Phase 1:Iter 1
- Phase 1:Iter 2
- Phase 2:Iter 1
- Phase 2:Iter 2
- Phase 3
- Phase 4

Batch 3

- Phase 1:Iter 1
- Phase 1:Iter 2
- Phase 2:Iter 1
- Phase 2:Iter 2
- Phase 3
- Phase 4
Impact of recurring batches

- A new ID has been introduced to recognize the different occurrences of the batch
  - Phase Iteration ID

- Batch evolution models with phase iterations will include several instances of the same batch
  - Number of batches varies between phases
  - Control limits are based on total number of iterations (sub-batches) in that phase

- For projects with phase iterations, separate batch level datasets will be created per phase

- There are options on how to treat phase iterations at batch level

- Max number of iterations of a phase limits prediction capabilities on Batch level
  - Prediction sets including more iterations than the workset will be cut at the number matching workset
    - First and last will not be affected.
Data import

- New observation ID for Phase iterations (slide 30)

- Hierarchy is Batch-Phase-Phase iteration
  - All data from one batch contiguous within a dataset
    - Different phases could be in different datasets
  - All data from on batch and phase must be contiguous
  - All data from on batch, phase and phase iteration must be contiguous
  - If Issues pane complains, sort on (in this order)
    1. Phase iteration ID
    2. Phase ID
    3. Batch ID

- Batch, phase and phase iteration conditions
  - First value - from a continuous dataset - in each batch, phase and phase iteration
  - Phase and phase iteration conditions from a discrete dataset requires batch, phase and phase iteration ID
Batch evolution models – BEM Workset Batch tab

- Batch tree structure
  - Batch ID
    - Phases
      - Phase iterations
  - Include/exclude Batches/ phases/ phase iterations
  - A specific phase iteration can be removed from the workset
Batch evolution models – the plots

- Example has 2 phase iterations
- For one batch there are two lines in all BC charts
  - Statistics are based on actual number of batches/iterations in that specific phase

- Use Marking mode to select batches with specific iteration ID, Unit ID etc.
  - Tools ribbon, Plot tools group, Select
The batch level datasets

- Batch level dataset with phase iterations differ from normal BL datasets
- Primary ID is a combination of BatchID, PhaseID and PhaseIterationID
- One row per Batch, phase and phase iteration
  - Several rows per batch for phases with phase iterations
The batch level workset and phase iteration options

- For BMs with phase iterations, Phase Iteration Options are available
  - Active option shown in info bar
  - Pay attention to Primary ID match

- Phase iteration options dialog – specify treatment of phase iterations
Phase iteration options

Which phase iterations to include

How to organize the data from phase iterations, if “All phase iterations” option is active

Option preview on dummy data
First, last or average phase iteration

- Gives only one iteration per phase meaning that these options equal a normal batch level model without iterations
- Examples of when these options make sense
  - First: When investigating differences between a successful first iteration with a failing one resulting in rework
  - Last: When you are only interested in modeling the final successful iteration (last and first can be equal)
  - Average: When there are varying numbers of phase iterations between batches or to scale down potential high influence of phases with many iterations in the model.
  - Block scaling can also be applied
All phase iterations

- As variables resembles normal batch level unfolding
  - One observation per batch
  - One block of variables per phase iteration instead of per phase
    - One phase is represented as many times as there are max number of iterations in workset
    - Workset limits how many phase iterations there can be in predictionset
    - Batches with fewer than max phase iterations are replaced with average phase iteration within that batch
      - Not beneficial to leave empty
- As observations resembles normal BL model of one phase only
  - One observation per batch and phase iteration
  - Only applicable when one dataset including phase iterations is selected
    - Otherwise warning
    - Low primary ID match
  - Allows investigation of drifts over time in phase iterations, differences between production units etc.
Customer challenges addressed

- **SIMCA and SIMCA-online must support modeling/monitoring of recurring phases**
  - Common in downstream processing but also occur in upstream

- **Ability to combine all data for the batch at batch level**
  - For CPV
  - Quality / output prediction
  - Compare production unit performance

- **All data from the process in One project**
  - Phase and phase iteration conditions
The customer challenge

- It should be possible for the model builder to include information about limits and alarms to be used in SIMCA-online configuration
  - Not always same person building model as configuring in SIMCA-online

- Some steps in the SIMCA-online configuration is demanding, how could info in the model building dataset help?
  - If available in dataset, UnitIDs should be possible to help with Unit configuration. The models know which models that are run in which unit and which units there are in the system

- SIMCA-online can send additional data to a SIMCA project for model update, but the way back to SIMCA-online is manual.
  - SIMCA can now share projects directly with SIMCA-online. Either for model update – then applying original configuration - or as a completely new configuration
SIMCA-online synergies

- Prepare SIMCA-online configuration file in SIMCA usp
  - Define alarm limits and rules
    - File|Info|Configure limits and alarms for SIMCA-online
    - Define UnitID at import greatly simplifies Unit configuration in SIMCA-online

- Push project from SIMCA to accessible SIMCA-online server
  - File|Share|Open in SIMCA-online
  - Review and complete configuration in SIMCA-online
Configure limits and alarms for SIMCA-online

- Vector, limits and trigger
- Stored in SIMCA usp, copied to SIMCA-online
- Window is non-modal meaning that you can still access and work with plots in the main SIMCA window.
- Icon opens plot for vector showing active prediction set and triggered alarms
  - Note: Limits and alarms must be “Applied” before alarm triggers are visible
  - Note: Only the first triggered alarm for each phase, batch and vector will show.
SARTORIUS

SIMCA® Import updates
SIMCA import updates

- Changed workflow in Home ribbon and File
- New variable and ID types
- Guidance in Issues pane
- Exchanged Excel parser for faster and more memory efficient import of Excel files
- SIMCA Import integrated into SIMCA
  - Earlier SIMCA Import was stand-alone, no visible difference in experience for user.
Home Ribbon changes including New variable and ID types

Regular project import

- Add data replace File|New spreadsheet
- Observation section

Batch project import

- Phase ID arrow includes Phase Iteration ID
- New ID type
- Batch condition arrow holds Phase and Phase iteration conditions
- Finish import Cancel is replaced by File|Discard import
- Batch project import
- Regular project import

SARTORIUS
Issues pane

- Hidden by default. Icon indicates issue
  - Click for more info

- If “Finish import” is clicked while warnings or critical issues remain, there will be a warning and the issues pane opens.

- Icons indicate severity or issue
  - ⓐ Can be handled automatically by SIMCA
  - ⓑ Issues with suggested actions
  - ⓒ Critical issues that must be managed by manual action, e.g. sort dataset
Minor fixes
Minor fixes summary

- Backstage New information
- Workset Select data page
- Batch marking mode ➔ Mark on any secondary ID
- SoV drill down to BE BCC
- Secondary Observation Identifiers are brought to Batch level datasets
- Permutation plot axis titles, change in default settings and Python support for creation
- Support project restore in case of computer crash
- Improvements in DPI awareness for high resolution screens
- Send us Feedback or report bugs directly from SIMCA
Descriptions to assist new users set up relevant type of project
Workset Select data page

- Changed layout where any potential issues are shown at the bottom of the list
- Size of datasets is included for guidance
Batch marking mode ➔ Mark on any Secondary ID

- Batch marking mode was replaced by a generic Secondary Identifier marking mode
- Available in Tools ribbon, Plot tools section, Select function
- Accessible for all project types with secondary IDs
Variable plot drill-down from Batch level

- Variable plots like loading and contribution plots drill down to BE variable BC charts
  - Also Sources of variation
- For Loading plot / SoV Loading plot drill down all workset batches are shown in BCC
- For Contribution plot / SoV contribution plot
  - Group to group contribution shows batches from both groups in BCC
  - Group / point to average contribution shows batch(es) from group / point in BCC
Permutation plot

- Axis titles added for clarity
- Default number of permutations changed from 20 to 100
- Permutation plot use same CV rules and groups as model
- Support to create permutation plots via Python interface
Give us your feedback

- Link to feedback form in SIMCA
- Upper right corner
- Report bugs, give suggestions for future improvements
Skins for SIMCA 15

- Omics skin and Spectroscopy skin are included in SIMCA 15 installation
- If not shown on the File|New page, enable them in File|Options|SIMCA options
Thank you for your interest in SIMCA® 15