POLITECNICO DI TORINO
Repository ISTITUZIONALE

Experiment data: Human-in-the-loop decision support in process control rooms

Terms of use:

Publisher copyright

(Article begins on next page)

16 August 2024

ELSEVIER

Data Article

Experiment data: Human-in-the-loop decision support in process control rooms

Chidera Winifred Amazu [a,b], Joseph Mietkiewicz [b,c],
Ammar N. Abbas [b,d,*], Houda Briwa [b,c], Andres Alonso Perez [b],
Gabriele Baldissone [a], Micaela Demichela [a], Davide Fissore [a],
Anders L. Madsen [c], Maria Chiara Leva [b]

[a] Politecnico di Torino, Corso, Duca degli Abruzzi, 24, Turin 10129, Italy
[b] Technological University Dublin, Dublin D07 EWV4, Ireland
[c] Hugin Expert, Aalborg, Denmark
[d] Software Competence Center Hagenberg, Hagenberg 4232, Austria

ARTICLE INFO

ABSTRACT

These datasets contain measures from multi-modal data
sources. They include objective and subjective measures
commonly used to determine cognitive states of workload,
situational awareness, stress, and fatigue using data collec-
tion tools such as NASA-TLX, SART, eye tracking, EEG, Health
Monitoring Watch, a survey to assess training, and a think-
aloud situational awareness assessment following the SPAM
methodology. Also, data from a simulation formaldehyde pro-
duction plant based on the interaction of the participants in
a controlled control room experimental setting is included.
The interaction with the plant is based on a human-in-the-
loop alarm handling and process control task flow, which in-
cludes Monitoring, Alarm Handling, Recovery planning, and
intervention (Troubleshooting, Control and Evaluation). Data
was collected from 92 participants, split into four groups
while they underwent the described task flow. Each partic-
ipant tested three scenarios lasting 15–18 min with a –10-
min survey completion and break period in between using
different combinations of decision support tools. The decision
support tools tested and varied for each group include alarm
prioritisation vs. none, paper-based vs. Digitised screen-based

---

* Corresponding author at: Technological University Dublin, Dublin D07 EWV4, Ireland.
  E-mail address: D21127032@mytudublin.ie (A.N. Abbas).

procedures, and an AI recommendation system. This is relevant to compare current practices in the industry and the impact on operators' performance and safety. It is also applicable to validate proposed solutions for the industry. A statistical analysis was performed on the dataset to compare the outcomes of the different groups. Decision-makers can use these datasets for control room design and optimisation, process safety engineers, system engineers, human factors engineers, all in process industries, and researchers in similar or close domains.

Specifications Table

Subject	Chemical Engineering
Specific subject area	Control and Safety Engineering, Human Factors and Ergonomics, Human-Computer Interaction, and Artificial Intelligence
Data format	Raw, Analysed, Filtered
Type of data	CSV File (.csv), Matlab File (.mat), Excel(.xlsx), Table
Data collection	The dataset contains behavioural, cognitive, and performance data from 92 participants, including system data under each participant from three scenarios, each simulating a typical control room monitoring, alarm handling, planning and intervention tasks and subtasks. The participants consented to participate on the test day, after which the researchers trained them. They performed functions under three scenarios, each lasting 15–18 min. During these tests, the participant wore a watch for health monitoring, including an eye tracker. They were asked situational awareness questions based on the SPAM methodology at specific periods within 15 min, especially at the 6th, 8th and 12th minutes. These questions assessed the three levels of situational awareness: perception, comprehension and projection. This feedback collection process on situational awareness differed for one of the groups that used an AI-based decision support system. The question for this group was asked right after specific actions. Therefore, for the overall study, the following performance-shaping factors are considered: type of decision support system, alarm design, procedure format, AI support, communication, situational awareness, cognitive workload, interface design, experience/training, task complexity, and stress. In both cases, communication was excluded as a factor considered in the first and second scenarios based on this absence. The data collected was normalised using the Min-Max normalisation.
Data source location	Politecnico di Torino, Turin, Italy Latitude: 45° 03′ 28.28″ N, Longitude: 7° 39′ 23.91″ E
Data accessibility	Repository name: Zenodo Direct URL to data: https://zenodo.org/records/10569181 DOI:10.5281/zenodo.10569181 [1]
Related research article	[2] Amazu C. W., Briwa H., Demichela M., Fissore D., Baldissone G., and Leva M. C., 2023. "Analysing" Human-in-the-Loop" for Advances in Process Safety: A Design of Experiment in a Simulated Process Control Room," 33rd European Safety and Reliability Conference (ESREL 2023), no. November, pp.2780–2787

1. Value of the Data

- Data needed for appropriate human reliability assessment is scarce. The data collected would be an addition to existing available data in assessing human reliability for human-in-the-loop process control rooms.

- These datasets provide psycho-physiological data to holistically assess the cognitive state of operators, performance, and safety and to understand how this differs between groups that use different decision support tools in process control rooms.
- The datasets can be used to develop human performance models and process safety models and potentially develop a digital twin simulating human–machine interaction in process control rooms.
- Decision-makers can use them to optimize the process control room, systems and safety engineers, human factor engineers, and teams developing guidelines and standards. Researchers involved in similar topics across related domains like nuclear, aviation, rail, energy industries, and more can use the data to develop models or test their ideas.
- Optimizing Human-AI Interaction: The dataset provides an opportunity to study the integration of human-in-the-loop configurations with AI systems in safety-critical industries. By examining the data, researchers can identify the factors necessary for successful collaboration between humans and AI. This knowledge can lead to the development of optimized interaction mechanisms, ensuring that the strengths of both humans and AI are leveraged effectively to enhance decision-making in critical scenarios.
- The dataset allows for qualifying and quantifying the performance and effectiveness of the AI-enhanced decision support system incorporating Deep Reinforcement Learning (DRL) using a Specialized Reinforcement Learning Agent (SRLA) framework [3–6]. By analysing the data, researchers can assess how well the system performs in safety-critical process industries with human-in-the-loop configurations, which is rarely observed. This evaluation can provide insights into the potential benefits, scope, and limitations of utilizing DRL in such contexts.

2. Background

These collected datasets aim to comprehensively assess cognitive states, workload, situational awareness, stress, and fatigue in human-in-the-loop process control rooms. The datasets include objective and subjective measures from various data collection tools such as NASA-TLX, SART, eye tracking, EEG, Health Monitoring Watch, surveys, and think-aloud situational awareness assessments. They also incorporate data from a simulation of a formaldehyde production plant based on participants' interactions in a controlled control room experimental setting.

The objective is to compare the performance and safety outcomes of different groups of participants exposed to varying decision support tools. These tools include alarm prioritisation, paper-based vs. digitised screen-based procedures, and an AI recommendation system. Statistical analysis was performed to compare the outcomes among the groups.

3. Data Description

This section provides a comprehensive overview of each raw data file in the dataset, presenting them individually with detailed explanations of the various types of data they contain.

3.1. General Introduction

The data repository is organized hierarchically, as depicted in the diagram shown in Fig. 1. The participants are divided into four distinct control groups. Within each group, there are three different scenarios that each participant experiences. As a result, the participant data in each group is segmented according to the specific scenario they underwent. This segmentation is essential for facilitating a more insightful and comprehensive data analysis.
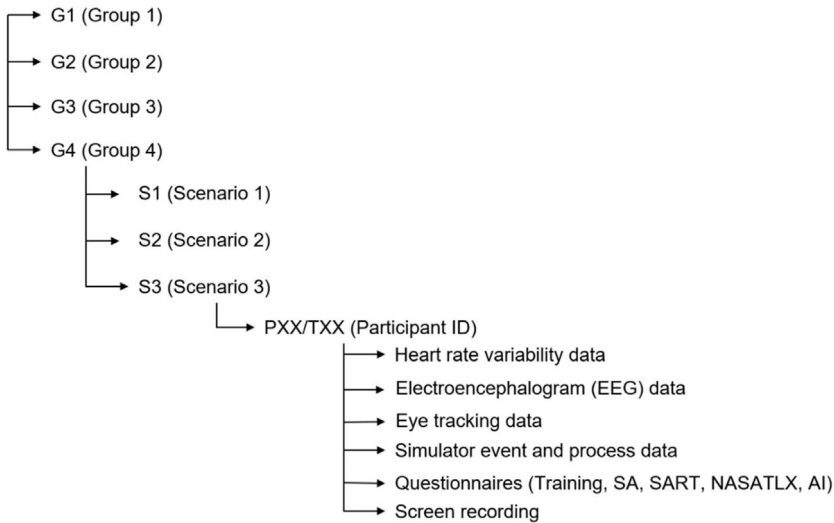
```
├──→ G1 (Group 1)
│
├──→ G2 (Group 2)
│
├──→ G3 (Group 3)
│
└──→ G4 (Group 4)
         │
         ├──→ S1 (Scenario 1)
         │
         ├──→ S2 (Scenario 2)
         │
         └──→ S3 (Scenario 3)
                  │
                  └──→ PXX/TXX (Participant ID)
                           ├──→ Heart rate variability data
                           ├──→ Electroencephalogram (EEG) data
                           ├──→ Eye tracking data
                           ├──→ Simulator event and process data
                           ├──→ Questionnaires (Training, SA, SART, NASATLX, AI)
                           ├──→ Screen recording
```

Fig. 1. Contents of the dataset (Hierarchy Diagram).

3.2. Groups

In our experiment, we divided participants into four distinct groups, each designed to evaluate different levels of support in a simulated control room environment. This division allowed us to assess the impact of various support features systematically. Here's a description of each group:

1. Group 1 - Baseline without Alarm Rationalization: This group operated under standard conditions without any alarm rationalization system. They served as the baseline for the experiment, providing a control group against which the other groups' performances could be compared.
2. Group 2 - Introduction of Alarm Rationalization: Unlike Group 1, this group was equipped with an alarm rationalization system. The primary objective was to assess how the rationalization of alarms could affect the operator's workload and decision-making process. This system was expected to filter out non-critical alarms, thereby reducing the cognitive load on operators and enabling them to focus on the most pertinent issues.
3. Group 3 - Transition to On-Screen Procedures: This group marked a shift from the paper-based methods used by Group 2, as they had access to procedures via an on-screen interface. The aim was to evaluate whether digital access to procedures could enhance operational efficiency and quick response times compared to traditional paper methods.
4. Group 4 - Integration of AI Decision Support: Representing the most advanced level of support, this group incorporated an AI decision support system as designed by Mietkiewicz et al. [7]. The focus was to measure AI's incremental benefits in improving operator performance, reducing errors, and enhancing overall system safety and efficiency, especially compared to Group 3′s digital procedures.

Each group's unique configuration allowed us to explore the effectiveness of different support levels in a control room setting, providing valuable insights into the potential benefits and challenges of integrating various technological tools into control room operations.

3.3. Scenarios

In our experimental study, we designed three distinct scenarios to evaluate the effectiveness of the decision support system in a simulated control room environment. Each scenario was crafted to challenge the participants in different aspects of control room operations. Here's a fresh description of each scenario:

5. Automatic Pressure Management Failure Scenario: This scenario simulated a tank's automatic pressure management system failure. Participants were required to adjust the nitrogen inflow to maintain stable pressure levels manually. The challenge intensified when the nitrogen flow unexpectedly stopped, leading to a critical drop in pressure. This scenario tested the participants' ability to quickly adapt to manual controls and effectively manage a sudden change in operational conditions.
6. Nitrogen Supply Disruption Scenario: In this scenario, participants faced a malfunction in the primary nitrogen supply system. The task was to swiftly switch to an auxiliary nitrogen supply while managing the gradual activation of the backup system. Participants needed to carefully adjust the pump power to slow the pressure decrease in the tank, balancing the need to maintain pressure without disrupting the overall system stability.
7. Heat Recovery System Malfunction Scenario: This scenario presented a temperature control failure in the Heat Recovery section. Participants initially attempted to rectify the issue by manually adjusting the cooling water flow. However, when these efforts proved ineffective, they consulted with a supervisor and subsequently focused on managing the reactor's temperature to prevent potential overheating. This scenario assessed the participants' problem-solving skills and ability to prioritise tasks under pressure.

Each scenario was meticulously planned to mimic real-world challenges faced in control rooms, providing a comprehensive test of the decision support system's utility in aiding operators during complex and high-pressure situations.

3.4. Participant IDs

Each participant is assigned a unique identifier, known as the Participant ID. The ID is structured using the letters "P" and "T," representing the two distinct institutes where the experiments were conducted and data collection occurred. Following the institute identifier, a numerical value is appended, representing the chronological order of the participants.

3.5. Screen recording

The experiment utilised three screens, each serving a distinct purpose. To observe participant responses and behaviors for each event, it was imperative to simultaneously record all three screens, as depicted in Fig. 2. OBS Studio was the chosen tool for screen recording, and sub-
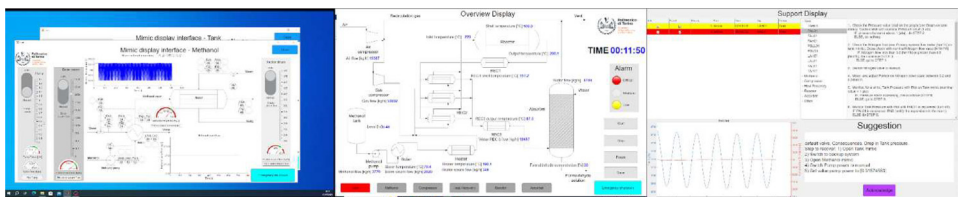

Fig. 2. Screenshot of screen recording.

timestamp_unix	timestamp_iso	participant_full_id	pulse_rate_bpm	missing_value_reason
1.69142E+12	07.08.2023 16:00	234-1-1-46	75	
1.69142E+12	07.08.2023 16:01	234-1-1-46	63	
1.69142E+12	07.08.2023 16:02	234-1-1-46	59	
1.69142E+12	07.08.2023 16:03	234-1-1-46	68	
1.69142E+12	07.08.2023 16:04	234-1-1-46	66	
1.69142E+12	07.08.2023 16:05	234-1-1-46	61	
1.69142E+12	07.08.2023 16:06	234-1-1-46	62	
1.69142E+12	07.08.2023 16:07	234-1-1-46	60	
1.69142E+12	07.08.2023 16:08	234-1-1-46	60	
1.69142E+12	07.08.2023 16:09	234-1-1-46	63	
1.69142E+12	07.08.2023 16:10	234-1-1-46	68	
1.69142E+12	07.08.2023 16:11	234-1-1-46	64	
1.69142E+12	07.08.2023 16:12	234-1-1-46	69	
1.69142E+12	07.08.2023 16:13	234-1-1-46	66	
1.69142E+12	07.08.2023 16:14	234-1-1-46	63	
1.69142E+12	07.08.2023 16:15	234-1-1-46	59	
1.69142E+12	07.08.2023 16:16	234-1-1-46	63	

Fig. 3. Example Screenshot of Watch Recording.

sequently, videos were saved and divided into segments based on individual scenarios for each participant.

3.6. Psycho-physiological data

3.6.1. Heart rate data
The health measurements were taken with a watch. The watch measures the pulse rate, electrodermal activity, movement intensity, respiratory rate, and temperature using EmpaticaPlus CareLab. A sample data snapshot is shown in Fig. 3.

3.6.2. Eye tracking data
A head-mounted device in the form of eye-tracking glasses containing tiny cameras that can measure the wearer's pupil size and eye movements is used to collect eye-tracking data using Tobii Glasses 3. It records various eye-related parameters such as eye movements, gaze patterns, and pupil dilation during usage. The collected raw recordings are organised separately for each participant and scenario, resembling other data files. These raw recordings can be conveniently uploaded to Tobii Pro Lab software, enabling researchers to perform in-depth analysis and extract valuable metrics.

3.7. Operational data

After each experimental scenario, the simulator data is recorded and saved, comprising variables from the process, alarms, and the decision support tool. The data files are categorised based on different process subsystems, and they include corresponding variables with the simulation time represented on the x-axis. Table 1 presents a comprehensive overview of the saved files, their associated variables (with acronyms and complete forms), and the corresponding y-axis representations, including the range of values or units for each variable.

Table 1

Operational log variables and meaning.

Section/File	Acronym	Full form	Y-axis	Value
Alarm	All2	Alarm activation	Alarm label	0 = not run, 1 = run
	Ackall	Alarm action	Status	0 = not run, 1 = run no action, 2 = run and action
	Sall	Alarm sound (no master silence)	Status	0 = not run, 1 = run sound, 2 = run single sound off
	Proclis	Procedure visualised	Procedure number	0 = No Procedure 1001-1999 Tank 2001-2999 Methanol 3001-3999 Heat Recovery 4001-4999 Reactor 5001-5999 Absorber 6001-6999 Other
	T2	Duration		[s]
	Intop	Open interface	1 = tank 2 = Methanol 3 = Compressor 4 = Heat recovery 5 = Reactor 6 = Absorber	0 = Close, 1 = Open
	Emmero	Emergency	Emergency status	0 = Close, 1 = On
	Choice	Case study		
	sugOld	Recommendation system		List of the recommendation sentence
	adsugOld	Operators accept the recommendation.		0 = not adopt 1 = Adopt
Tank	Lserb	Level	Level	[m]
	Pserb	Pressure	Pressure	[ata]
	Fn2serb1o	Primary nitrogen flow	Flow	[Nm³/h]
	Fn2serb2o	Back-up nitrogen flow	Flow	[Nm³/h]
	Flo2	Airflow	Flow	[Nm³/h]
	Nitsel	Nitrogen system selector	Status	0 = off, 1 = Primary, 2 = Back up
	Mnitsel	Manual primary nitrogen selector	Status	0 = auto, 1 = manual
	X	Manual primary flow value	Flow	0 = off, >0 = [Nm³/h]
Methanol	Fmeps	Methanol liquid pumped	Flow	[kg/h]
	Wpomp	Pump power	Power	[Kw]
	Mpumpold	Manual pump selector	Status	0 = auto, 1 = manual
	Mwpopold	Manual pump power value	Power	0 = off, >0 = [Kw]
	Mliqbolo	Methanol mass in the boiler	Mass	[kg]
	Fbol	Methanol boiled flow	Flow	[kg/h]
	Fbv	Boiler steam flow	Flow	[kg/h]
	Tbol	Methanol boil temperature		[°C]
	Mboilvapo	Manual boiler steam selector	Status	0 = auto, 1 = manual
	Fvapbolo	Manual boiler steam value	Flow	0 = off, >0 = [kg/h]
	Tris	Heater methanol output temperature	Temperature	[°C]
	Frisv	Heater steam	Flow	[kg/h]

(*continued on next page*)

Table 1 (*continued*)

Section/File	Acronym	Full form	Y-axis	Value
	Mhearvapo	Manual heater steam selector	Status	0 = auto, 1 = manual
	Mmvapriso	Manual heater steam value	Flow	0 = off, >0 = [kg/h]
Compressor	Maircompold	Manual air compressor selector	Status	0 = auto, 1 = manual
	Maircomppowold	Manual air compressor value	Power	0 = off, >0 = [Kw]
	Warglob	Air compressor power	Power	[Kw]
	Fariao	Airflow	Flow	[kg/h]
	Gasselectold	Gas manual selector	Status	0 = auto, 1 = Manual Vent fraction, 2 = manual Ga Compressor
	Frazventcold	Vent fraction	Fraction	[]
	Fvento	Vent flow	Flow	[kg/h]
	Wriglob	Gas compressor power	Power	[Kw]
	Fgaspomo	Gas compressed flow	Flow	[kg/h]
	Mgascompowol	Manual gas flow	Flow	0 = off, >0 = [kg/h]
	Poutcomp	Gas compressor pressure	Pressure	[Pa]
Reactor	Meold	Reactor methanol fraction in	Frac	[]
	Oold	Reactor oxygen fraction in	Frac	[]
	Nold	Reactor nitrogen fraction in	Frac	[]
	Wold	Reactor water fraction in	Frac	[]
	Foold	Reactor formaldehyde fraction in	Frac	[]
	Coold	Reactor carbon monoxide fraction in	Frac	[]
	Tinreaold	Temperature reactor in	Temperature	[K]
	Pinreaold	Pressure reactor in	Pressure	[Pa]
	Tmaxreatore	Max temperature in reactor	Temperature	[°C]
	Tmreato	Temperature along the reactor	Temperature	[°C]
	Meoutro	Reactor methanol fraction out	Frac	[]
	Ooutro	Reactor oxygen fraction out	Frac	[]
	Noutro	Reactor nitrogen fraction out	Frac	[]
	Woutro	Reactor water fraction out	Frac	[]
	Fooutro	Reactor formaldehyde fraction out	Frac	[]
	Cooutro	Reactor carbon monoxide fraction out	Frac	[]
	Toutro	Temperature reactor out	Temperature	[K]
	Poutro	Pressure reactor out	Pressure	[Pa]
	Frout	Reactor out flow	Flow	[kg/h]
	Fmw1	Steam flow cooling system	Flow	[kg/s]
	Pebre	Steam pressure cooling system	Pressure	[bar]
	Treatsteam	Steam temperature cooling system	Temperature	[°C]
	Mcoolreatold	Manual reactor cooling selector	Status	0 = auto, 1 = manual
	Mcreattempold	Manual reactor cooling value	Temperature	0 = off, >0 = [°C]
Heat recovery	Pvrec1	Pressure steam rec1	Pressure	[ata]
	Twrec1	Temperature steam rec1	Temperature	[°C]
	Mvaprec1	Flow steam rec1	Flow	[kg/h]
	Toutrec1	Temperature out rec1	Temperature	[°C]
	Tinrec1	Temperature in rec1	Temperature	[°C]

Table 1 (*continued*)

Section/File	Acronym	Full form	Y-axis	Value
	Rec1smo	Manual rec1 selector	Status	0 = auto, 1 = manual
	Rec1smpo	Manual rec1 value	Pressure	0 = off, >0 = [ata.]
	Sezrec2o	Selector number of heat exchanger in rec2	Number	1 = 1 heat exchanger used 2 = 2 heat exchangers used 3 = 3 heat exchangers used
	Toutrec2f	Reagent out temperature	Temperature	[°C]
	Toutrec2c	Temperature in rec3	Temperature	[°C]
	Toutrec3	Temperature out rec3	Temperature	[°C]
	Twrec3	Temperature steam rec3	Temperature	[°C]
	Mvaprec3	Flow steam rec3	Flow	[kg/h]
	Rec3wmo	Manual rec3 selector	Status	0 = auto, 1 = manual
	Rec3wmfo	Manual rec1 value	Flow	0 = off, >0 = [kg/s]
Absorber	Linaso	Water inflow	Flow	[kg/h]
	Masswato	Manual absorber water selector	Status	0 = auto, 1 = manual
	Masswatflowo	Manual absorber water value	Flow	0 = off, >0 = [kg/h]
	Tmeancol	Mean temperature in absorber	Temperature	[°K]
	Ttopcol	Top temperature in absorber	Temperature	[°K]
	Tliqin	Water in temperature in absorber	Temperature	[°K]
	Mfondasso	Liquid bottom mass in the absorber	Mass	[kg]
	Ymeoutasso	Absorber gas methanol fraction out	Frac	[]
	Ywoutasso	Absorber gas water fraction out	Frac	[]
	Ycooutasso	Absorber gas carbon monoxide fraction out	Frac	[]
	Yfooutasso	Absorber gas formaldehyde fraction out	Frac	[]
	Ynoutasso	Absorber gas nitrogen fraction out	Frac	[]
	Yooutasso	Absorber gas oxygen fraction out	Frac	[]
	Toutasso	Temperature absorber gas out	Temperature	[°K]
	Poutasso	Pressure absorber gas out	Pressure	[Pa]
	Gas	Absorber gas out flow	Flow	[kg/h]
	Xmasoutass	Absorber liquid formaldehyde fraction out	Frac	[%]
	Xmasoutassme	Absorber liquid methanol fraction out	Frac	[%]
	Xmasoutassw	Absorber liquid water fraction out	Frac	[%]
	Toutassol	Temperature absorber liquid out	Temperature	[°C]
	Loutasso	Absorber liquid outflow	Flow	[kg/h]
	Mliqouto	Manual absorber liquid out selector	Status	0 = auto, 1 = manual
	Mliqoutflowo	Manual absorber liquid out value	Flow	0 = off, >0 = [kg/h]

3.8. Questionnaires data

3.8.1. General Information data

This dataset provides a demographic and educational background overview of control room simulation experiment participants. It offers insights into the diversity of the participant pool, including age, gender, academic qualifications, and familiarity with industry and control room environments. This data is crucial for understanding the varied perspectives and experiences that participants brought to the experiment, influencing their interaction with the control room simulation.

Dataset Overview: The dataset includes selected entries from the participant pool, focusing on their demographic and educational background. Key elements of this partial dataset include:

- **Participant Identifier:** Unique codes for participants (e.g., P100, P36) to maintain anonymity.
- **Group Assignment:** Indicates the experimental group (e.g., G1) to which participants were assigned, reflecting different levels of decision support in the simulation.
- **Age and Gender:** Basic demographic information of participants, providing insights into the diversity of the participant pool.
- **Educational Background:** Details of participants' academic qualifications, including degree type (e.g., Masters, PhD), year of study, and field of study (e.g., Chemical Engineering, IT).
- **Dominant Hand:** Information on whether participants are right or left-handed, which could influence their interaction with the simulation interface.
- **Familiarity with Industry and Control Room:** Self-reported familiarity levels with the industry in general and control room environments specifically, on a scale from 1 to 5.

This dataset is valuable for assessing the influence of demographic and educational backgrounds on participants' performance and experiences in the control room simulation. It allows researchers to correlate these factors with other data collected during the experiment, such as performance metrics and questionnaire responses. The diversity in educational backgrounds, particularly in fields related to and outside of Chemical Engineering, provides a broader perspective on how different educational experiences impact interaction with control room simulations.

3.8.2. Training data

This dataset comprehensively evaluates participants' ability to perform specific tasks in a control room simulator and their self-assessed knowledge of the simulator and operating procedures. The dataset includes instructor assessments of participants' skills in changing values, identifying plant sections, acknowledging alarms, locating intervention procedures, and interacting with an AI-enhanced decision support system. Additionally, it captures participants' self-ratings of their familiarity with the simulator and operating procedures and their evaluation of the training quality.

Dataset Overview: The dataset encompasses responses and assessments from participants in a control room simulation experiment. Key components include:

- **Participant Identifier:** Unique IDs (e.g., P10, P11, P12) to maintain participant anonymity.
- **Instructor Assessments:** Ratings on a scale of 1 to 5, assessing participants' capabilities in performing specific simulator tasks, such as changing values, identifying sections, acknowledging alarms, and using the AI decision support system.
- **Self-Assessed Knowledge:** Participants' self-ratings on their understanding of the simulator and operating procedures.
- **Training Evaluation:** Participants' ratings on the quality of information provided during training and any comments or suggestions for improvement.

This dataset is crucial for evaluating the effectiveness of training programs in control room simulations and the competency of operators in handling complex tasks. The instructor assessments objectively measure participants' skills in critical control room operations, while the self-assessments offer insights into their perceived readiness and confidence. The feedback on training quality is invaluable for refining instructional methods and materials.

Participant	Group	Age	Gender	Degree	Year_of_study	Course	Dominant_hand	Familiarity_industry	Familiarity_control_room
P100	G1	22	Male	Masters	Yr2	Chemical Engineering	Right	4	1
P36	G1	22	Male	Bachelors	Yr4	IT	Right	1	1
P43	G1	31	Male	PhD	Yr1	Other engineering	Right	1	3
P44	G1	27	Male	Masters	System innovation	Mechanical Engineering	Left	2	2
P45	G1	29	Male	PhD	Yr3	Other engineering	Left	2	3
P46	G1	25	Male	PhD	Yr1	Other engineering	Right	1	1
P49	G1	35	Male	PhD	Ended	Other engineering	Right	3	5
P60	G1	24	Male	Masters	Yr2	Chemical Engineering	Right	4	2

Fig. 4. Sample of the general information data.

Participant ID	Can you show us how ... Below is the instructor (1: Low, 5: High)	Can you show us how ... Below is the instructor (1: Low, 5: High)	Can you show us how ... Below is the instructor (1: Low, 5: High)	Can you show us wher ... Below is the instructor (1: Low, 5: High)	Can you show us how ... Below is the instructor (1: Low, ...	How would you rate yo... (1: Low, 5: High)	How would you rate yo... (1: Low, 5: High)	How would you rate th... (1: Low, 5: High)	Any comment/ sugg
P10	5	5	5	5		2	1		none
P11	5	5	1	5		4	4	5	none
P12	5	5	2	5		4	3		it was good
P13	5	5	5	5		3	3	3	none

Fig. 5. Training questionnaire data sample.

3.8.3. SPAM data

This SPAM dataset presents responses from the Situation Present Assessment Method (SPAM) questionnaire used in the experiment. The dataset evaluates situation awareness among operators in the three simulated scenarios. It provides critical insights for research in human-machine interaction, control room ergonomics, and decision-making processes in safety-critical environments.

Dataset Overview: The dataset includes selected entries, focusing on specific participants and scenarios. Key elements of this partial dataset include:

- **Participant Identifier:** Unique codes for participants (e.g., P04, P06), maintaining anonymity while allowing for individual analysis.
- **Group Assignment:** Indicates the experimental group (e.g., G4, G3, G2, G1) to which participants belonged, reflecting different levels of decision support in the simulation.
- **Scenario Engagement:** Identifies the specific scenarios (e.g., S1, S2, S3) each participant encountered, representing diverse challenges within the control room simulation.
- **SPAM Metrics:** Participant ratings across three dimensions of the SPAM questionnaire, **Monitoring, Planning**, and **Intervention**, on a scale typically from 1 to 5.
- **SPAM Index:** Composite scores derived from the SPAM, indicating overall situation awareness levels experienced by participants.

This partial dataset is instrumental for focused research on specific aspects of situation awareness in control room environments. It allows for a detailed examination of how different interface designs and decision support levels impact operators' cognitive processing and decision-making abilities (Figs. 4 and 5).

3.8.4. NASA-TLX

This dataset encompasses participants' responses utilising the NASA Task Load Index (NASA TLX) questionnaire. The dataset aims to provide insights into the cognitive, physical, and emotional workload experienced by operators in simulated control room scenarios. It is a valuable resource for researchers studying human factors in high-stakes environments, particularly in process safety and control room operations.

Dataset Overview: The dataset is structured to capture a comprehensive range of metrics reflecting the participants' experiences and performance across different scenarios in a control room simulation. Key components of the dataset include:

- **Participant Identifier:** A unique alphanumeric code assigned to each participant (e.g., P04, P06), ensuring anonymity while allowing for individual response tracking.
- **Group Assignment:** Classification of participants into different experimental groups (e.g., G4, G3), each representing varying levels of decision support or experimental conditions within the control room simulation.

Participant	Group	Scenario	monitoring	planning	intervention	SPAM_index
P100	G1	S1	1	4	1	2
P100	G1	S2	2	3	1	2
P100	G1	S3	3	4	1	2.66666666667
P36	G1	S1	5	2	3	3.33333333333
P36	G1	S2	5	2	4	3.66666666667
P36	G1	S3				

Fig. 6. Sample of the SPAM dataset.

Participant	Group	Scenario	Mental_demand	Physical_demand	Temporal_demand	Performance	Effort	Frustration	TLX_index
P04	G4	S1	3	1	2	6	3	3	3
P04	G4	S2	3	1	5	5	6	7	4.5
P04	G4	S3							
P06	G4	S1	4	6	2	7	5	1	4.166666667
P06	G4	S2	7	7	7	1	7	7	6

Fig. 7. Sample of the NASA-TLX dataset.

- **Scenario Engagement:** Details of the specific scenario (e.g., S1, S2, S3) encountered by the participant, with each scenario representing distinct challenges or tasks in the control room environment.
- **NASA TLX Metrics:** Ratings provided by participants across six dimensions of the NASA TLX questionnaire, including Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each dimension is rated on a scale, typically 1 (low) to 7 (high).
- **TLX Index:** A composite score derived from the NASA TLX, representing the overall workload experienced by the participant. It is calculated as an average of the ratings across the six dimensions (Figs. 6 and 7).

3.8.5. SART

This dataset presents participants' responses to the Situation Awareness Rating Technique (SART) questionnaire collected during a control room simulation experiment. The dataset aims to assess operators' situation awareness in various simulated scenarios, providing insights into their perception, comprehension, and projection abilities in a high-stakes environment.

Dataset Overview: The dataset captures a range of metrics that reflect the participants' situation awareness in different simulated control room scenarios. Key components of the dataset include:

- **Participant Identifier:** A unique code for each participant (e.g., P04, P06), ensuring anonymity and enabling individual response analysis.
- **Group Assignment:** Classification of participants into experimental groups (e.g., G4, G3), each experiencing different levels of decision support or conditions in the control room simulation.
- **Scenario Engagement:** Details of the specific scenario (e.g., S1, S2, S3) each participant encountered, representing various challenges in the control room environment.
- **SART Metrics:** Ratings provided by participants across several dimensions of the SART questionnaire, including Instability, Variability, Complexity, Arousal, Spare Capacity, Concentration, Attention Division, Quantity, Quality, and Familiarity. Each dimension is rated on a scale, typically 1 (low) to 7 (high).
- **SART Index:** A composite score derived from the SART, representing the overall situation awareness experienced by the participant. It includes SART Demand, SART Supply, and SART Understanding scores.

This dataset is crucial for situation awareness, control room design, and safety management research. It provides empirical evidence on how different control room designs, decision support systems, and scenario complexities impact operators' situation awareness.

3.9. Decision support system data

This dataset analyses operators' perceived task load and interaction with various support systems in a simulated control room environment. The data were collected across different scenarios, focusing on participants' experiences with alarm lists, prioritisation, and procedural support systems. The dataset provides insights into how these support systems influence the operators' workload and efficiency in managing control room tasks. All values are on a scale of 1–5.

Dataset Overview: The dataset includes responses from participants in control room simulation scenarios. Key components of the dataset are:

- **Participant Identifier:** Unique IDs (e.g., P100, P36) for participant anonymity.
- **Group and Scenario:** Classification of participants into groups and the scenarios they encountered during the simulation.
- **Task Load:** Participants' self-assessment of their workload on a scale, indicating the perceived intensity of tasks in each scenario.
- **Support System Interactions:** Ratings on the effectiveness and utility of various support systems, including alarm lists, alarm prioritisation, and procedural support, in aiding their task management.

This dataset is instrumental in understanding the impact of support systems on operators' workloads in simulated control room environments. By analysing participants' ratings, researchers can gauge the effectiveness of alarm lists, prioritisation techniques, and procedural supports in reducing task load and enhancing operational efficiency.

3.10. AI decision support system data

This dataset encapsulates participant feedback on using an AI Decision Support System (DSS) in a control room simulation experiment. The data reflects participants' perceptions of the effectiveness, explainability, trustworthiness, and utility of the AI support. It also explores the perceived impact of the AI system on workload and the balance between its benefits and the additional workload it introduces. Additionally, the dataset includes insights into the importance of validating AI recommendations and the influence of Deep Reinforcement Learning (DRL) on trust enhancement.

Dataset Overview: The dataset presents selected responses from participants who interacted with an AI-based decision support system during a control room simulation. Key components of the dataset include:

- **Participant Identifier:** Unique codes (e.g., P97, P96) for participant anonymity.
- **Group and Scenario:** Indicates the experimental group (G4) and scenario (S1, S2, S3) where the participant interacted with the AI system.
- AI Support Ratings: Participants' ratings on various aspects of the AI system, including support, explainability, trust, and helpfulness, on a scale from 1 to 5.
- **AI Workload Impact:** Ratings on how the AI system affected the workload and the balance between its benefits and additional workload.
- A**I Validity and DRL Importance:** Participants' views on the importance of validating AI recommendations and the role of DRL in enhancing trust in the system.

This dataset is instrumental in understanding participants' subjective experiences with the AI decision support system in a simulated control room environment. The varied ratings across different aspects of the AI system, such as support, explainability, and trust, provide insights into the system's perceived strengths and areas for improvement. The data on workload impact and the balance between AI benefits and additional workload offer valuable perspectives on the practicality and efficiency of integrating AI systems in control room operations. Furthermore, the emphasis on the importance of validating AI recommendations and the role of DRL in enhancing trust sheds light on critical factors for the successful adoption of AI in high-stakes environments.

Participant	Group	Scenario	Instability	Variability	Complexity	Arousal	Spare_capacity	Concentration	Attention_division	Quantity	Quality	Familiarity	SART_demand	SART_supply	SART_understandin	SART_index
P100	G1	S1	2	2	3	6	6	7	6	7	7	6	7	24	20	37
P100	G1	S2	6	5	6	7	6	7	7	7	7	6	17	27	20	30
P100	G1	S3	7	7	7	7	7	7	7	7	6	6	21	28	19	26
P36	G1	S1	3	3	2	2	4	5	4	2	5	3	8	15	10	17
P36	G1	S2	3	3	2	2	4	5	3	2	5	4	8	14	11	17
P36	G1	S3	3	3	4	4	3	4	3	2	3	1	10	14	6	10
P43	G1	S1	2	1	1	6	4	7	1	7	7	1	4	18	15	29
P43	G1	S2	5	5	5	4	4	7	1	7	7	4	15	16	18	19
P43	G1	S3	7	7	7	6	6	6	3	5	5	3	21	21	13	13

Fig. 8. Sample of the SART dataset.

Participant	Group	Scenario	Task_load	Alarm_lists_support	Alarm_prioritization_support	Procedures_support
P100	G1	S1	2	4	4	5
P100	G1	S2	4	2	2	4
P100	G1	S3	5	4	2	1
P36	G1	S1	2	4	1	4
P36	G1	S2	2	4	1	4
P36	G1	S3	3	4		4

Fig. 9. Sample of the support system dataset.

4. Experimental Design, Materials and Methods

4.1. Introduction

This within-and between-subject experiment was conducted to test different combinations of decision support tools, alarm design, procedures, and AI decision support while investigating the impact of such setups on operators' cognitive state of situational awareness and mental workload and a possible stress condition. This research complies with the American Psychological Association (APA) code of ethics. An internal board has given ethical approval within the collaborative intelligence for safety-critical systems (CISC) network (Figs. 8 and 9).

4.2. Experiment design

The participants in this study are students from Politecnico di Torino and Technological University Dublin. A total of 92 participants, 36 females and 56 males, participated. Participants were between 21 – and 61 years old, M = 25, SD = 5.4, with varying experience levels. They were primarily junior process engineers selected voluntarily from among the students of the master courses in chemical engineering at Politecnico di Torino.

They were eligible if they were above 18, willing to participate voluntarily, and could give informed consent. The exclusion criteria for participants are any medical condition that could affect the participant's ability to perform the tasks required in the experiment, making any medication that could impact the participant's cognitive or physical abilities or any history of psychological or neurological disorders, and no history of epilepsy or seizures.

The participant is divided into four groups to test 4 different HMI. All four groups have the same interface with the addition of some other support. Group 1 has the default support interface. Group 2 has the alarm rationalisation in addition to Group 1. Group 3 has the display of the procedure on the screen. Finally, Group 4 has an AI decision support system and a screen-based procedure.

4.3. Scenarios and description

This study uses a simulated interface of a modified formaldehyde production plant based on the work done by Demichela et al. [8]. The plant generates approximately 10,000 kg/h of 30 % formaldehyde solution through the partial oxidation of methanol with air. The simulator consists of six sections: Tank, Methanol, Compressor, Heat Recovery, Reactor, and Absorber. It incorporates 80 alarms with different prioritisation levels (1: low, 2: medium, 3: high), including

Participant	Group	Scenario	AI_support	AI_explainability	AI_trust	AI_help	AI_additional_work_load	AI_benefits_vs_additional_work_load	AI_validity	DRL_importance	Trust_increase_due_to_
P97	G4	S1	2	3	4	2	3	3	3	4	3
P97	G4	S2	2	3	4	2	3	3	3	4	3
P97	G4	S3	2	3	4	2	3	3	3	4	3
P96	G4	S1	5	5	3	4	4	4	4	5	4
P96	G4	S2	5	5	3	4	4	4	4	5	4
P96	G4	S3	5	5	3	4	4	4	4	5	4

Fig. 10. Sample of the AI question.

nuisance alarms (alarms that are not relevant). The main screen of the simulator can be seen in Fig. 2. Different scenarios, with varying task complexities based on the number of alarms and the number of sub-task steps, were created to assess the efficiency of the decision support.

- Pressure indicator control failure (Tank Section).
    ○ Hazardous event: Methanol storage tank implosion.
    ○ Goal: Prevent activation of pressure switch PSL01 within 7 min of the initial alarm.
- Nitrogen valve primary source failure (Tank Section).
    ○ Hazardous event: Methanol storage tank implosion.
    ○ Goal: Prevent the implosion of the tank
- Failure of temperature indicator control (Heat Recovery section).
    ○ Hazardous event: Overheating of the reactor.
    ○ Goal: Prevent the overeating of the reactor.

4.4. Tools

4.4.1. Eye tracker

Eye tracking was conducted using Tobii eye trackers. These devices recorded the participants' gaze points and pupil diameter, allowing for the analysis of attention distribution and the duration of focus on specific areas of the screens. This information was used to generate heat maps, visual representations of where the participants directed their attention and the time spent on particular screens. The eye-tracking data analysis proved particularly valuable in studying factors such as the participants' focus on, for example, decision support and the duration of their attention to them. The participants wore eye trackers during each scenario.

4.4.2. Smartwatch

In addition to the eye trackers, each participant wore a smartwatch. The smartwatch collected various physiological measurements, including heartbeat, movement intensity, pulse rate variability (PRV), respiratory rate, oxygen saturation (spo2), temperature, and electrodermal activity (EDA). These measurements provided additional insights into the participants' physiological responses during the experiment (Fig. 10).

4.5. Task flow example

All participants completed three scenarios, each lasting 15–17 min. The tasks were based on a human-in-the-loop process control room situation where tasks, such as operating and intervention tasks, are commonly described in procedures (paper, digitised/screen-based, automated). Here, intervention procedures have been defined for all 80 alarms in the setup. An example event and task flow for scenarios 1, 2 and 3 are shown in Figs. 11–13, respectively. The participants wore the glasses, EEG, and health monitoring watch some minutes before the start of the test to create a baseline for data collection. They are worn throughout the three tests except for the eye tracker, which is removed a few minutes after a test to allow for post-test conditions.
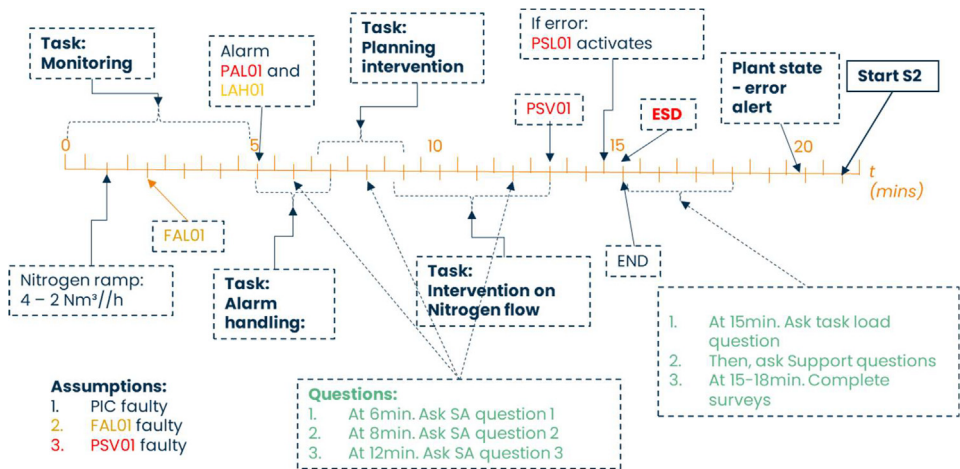
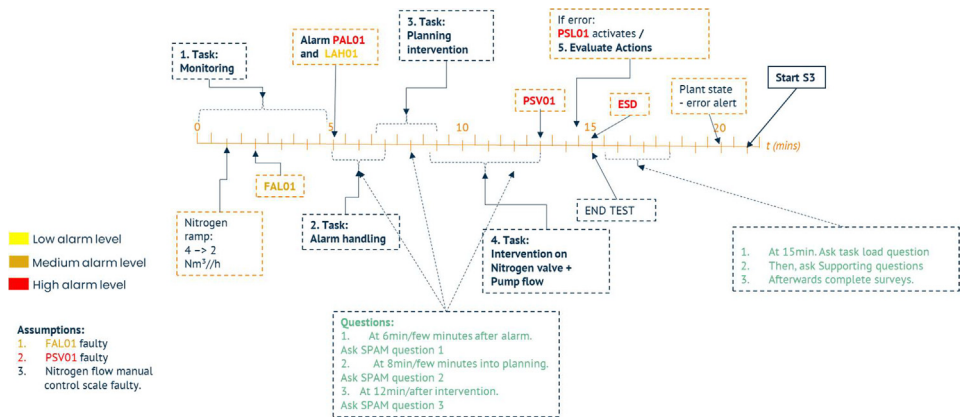**Fig. 11.** Scenario 1 task flow and events (Amazu C. W. et al., 2023).

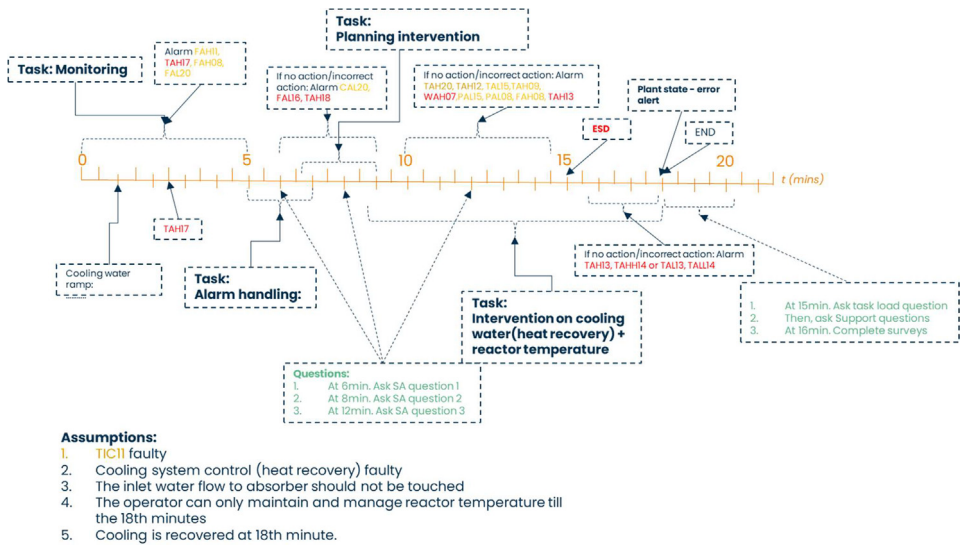**Fig. 12.** Scenario 2 Task flow and events.

Fig. 13. Scenario 3 Task flow and events.

Limitations

- Participants 1–20 do not have eye-tracking data recording, as can be observed in the repository.
- The participants incompletely filed some survey data.

Ethics Statement

The participants were briefed before the study, read a detailed description of what the study entailed via an information sheet and signed the necessary consent form before participating. These documents, together with the ethics application, were approved by the Internal Ethical Committee of the Collaborative Intelligence for Safety-Critical Systems, following a first approval by the ethical review committee of Technological University Dublin Ireland with approval number REC-20-52.

CRediT Author Statement

All authors have contributed equally to developing the work and the data collection.

Data Availability

Human-in-the-Loop Decision Support in Process Control Rooms (Original data) (Zenodo).

Acknowledgments

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] CISC LiveLab 3: data repository (2024). doi:10.5281/zenodo.10569181.
[2] C.W. Amazu, H. Briwa, M. Demichela, D. Fissore, G. Baldissone, M.C. Leva, Analysing 'Human-in-the-loop' for advances in process safety: a design of experiment in a simulated process control room, in: 33rd European Safety and Reliability Conference (ESREL 2023), 2023, pp. 2780–2787.
[3] A.N. Abbas, G.C. Chasparis, J.D. Kelleher, Hierarchical framework for interpretable and specialized deep reinforcement learning-based predictive maintenance, Data Knowl. Eng. 149 (2024) 102240.
[4] A.N. Abbas, G.C. Chasparis, J.D. Kelleher, Interpretable input–output hidden markov model-based deep reinforcement learning for the predictive maintenance of turbofan engines, in: International Conference on Big Data Analytics and Knowledge Discovery, Cham, Springer International Publishing, 2022, July2023, pp. 133–148.
[5] A.N. Abbas, G.C. Chasparis, J.D. Kelleher, Specialized deep residual policy safe reinforcement learning-based controller for complex and continuous state-action spaces, arXiv preprint arXiv:2310.14788. (2023).
[6] Mietkiewicz, J., Abbas, A.N., Amazu, C.W., Madsen, A.L., & Baldissone, G. (2023). Dynamic influence diagram-based deep reinforcement learning framework and application for decision support for operators in control rooms.
[7] J. Mietkiewicz, A.L. Madsen, Enhancing control room operator decision making: an application of dynamic influence diagrams in formaldehyde manufacturing, in: European Conference on Symbolic and Quantitative Approaches with Uncertainty, Springer Nature Switzerland, Cham, 2023, pp. 15–26, doi:10.3850/978-981-18-8071-1_P531-cd.
[8] M. Demichela, G. Baldissone, G. Camuncoli, Risk-based decision making for the management of change in process plants: benefits of integrating probabilistic and phenomenological analysis, Ind. Eng. Chem. Res. 56 (50) (2017) 14873–14887.