DOI 10.5162/iCCC2026/P14

Analyse industrieller Prozessdaten mit KI Agenten
AI Agentic Analysis of Industrial Process Data

Kai Gehrke, Sarah Spiesecke, Til Bartel, Ferdinand-Braun-Institut (FBH), Berlin, Germany, til.bartel@fbh-berlin.de

Kurzfassung

Wir zeigen, wie ein Large Language Model (LLM) allein durch Texteingaben komplexe Sensordaten einer Einzelwafer-Ätzanlage analysiert. Das Modell identifiziert Prozessläufe, normalisiert Zeitverläufe und erkennt Abweichungen mittels ±3σ-Hüllkurven. Die Analyse basiert auf MES- und Trace-Daten und erfolgt vollständig ohne Programmierung durch den Nutzer. Das LLM liefert innerhalb von Minuten visuell aufbereitete Ergebnisse und markiert Ausreißer. Der Ansatz demonstriert das Potenzial agentischer LLMs für schnelle, skalierbare Analysen in industriellen Anwendungen.

Abstract

We demonstrate how a commercially available Large Language Model (LLM) can execute complex data manipulation via simple text prompts. We prompt the LLM to analyze trace sensor data from a production equipment with the aim to establish a standard trace and assess possible process deviations. The LLM executes these tasks based on simple natural language input, demonstrating the increasingly agentic capabilities of LLMs. It is expected that this type of analysis will rapidly become a standard tool in R&D and industrial environments.

1 Motivation

Modern R&D and industrial environments are increasingly saturated with data. Thanks to automated systems and networked sensors, virtually every aspect of production and experimentation can be recorded and - in theory - analyzed. However, valorizing this data requires significant effort to clean and digest it. This typically demands domain knowledge and significant manual effort, often making it a bottleneck in data-driven optimization.
To address this challenge, we explore how a Large Language Model (LLM, the model), specifically ChatGPT, can act as an agentic analytical partner: a system capable of autonomously decomposing tasks, reasoning over data, and making purposeful decisions without explicit programming. In contrast, prior work using LLMs for data analysis [1,2] focuses on architectural adaptations and custom training to develop new machine learning (ML) algorithms, rather than leveraging the model's agentic capabilities.

2 Case Study: Analysis of trace data to identify process deviations

In this case study, we chose trace data from a reactive ion etching tool extensively used for processing wafers in our clean-room. Process lots comprise multiple wafers processed by the tool in individual runs. Metadata, including wafer names, recipe, and tracking times (for both lot and wafer-level events), is recorded in our manufacturing execution system (MES). The raw sensor time series, are stored via SECS-GEM protocol in an Influx data base.
Many different recipes are run on the tool, some of them of fixed length (etching by time), some of them of variable length (etching by endpoint detection). Contextual information from the MES (process start and end time stamps, recipe and wafer identification) have yet to be matched with the time series.

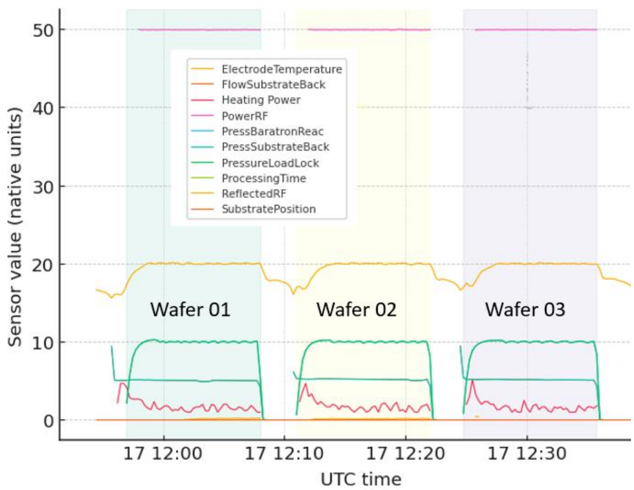

Fig. 1 ChatGPT o3-pro analyzes years of trace data and accurately identifies different sensor traces corresponding to single-wafer processing. In this example, a batch of four wafers is segmented correctly despite imprecise tracking timestamps. A visually compelling graph is generated.

Browsing this heterogenous data to identify comparable runs and aligning them temporally is expected to hold important information regarding tool stability and health. An important use case is also helping to identify possible deviations in past and future runs.

2.1 Task description & prompting

The task can be divided in the following actions:
1) Ingest and understand both sources of data (Times series and MES context data)
2) Parse the meta data for a chosen run or recipe and resolve their time stamps
3) Select the trace data corresponding to these time stamps and identify typical traces, if needed, segment by wafer runs.

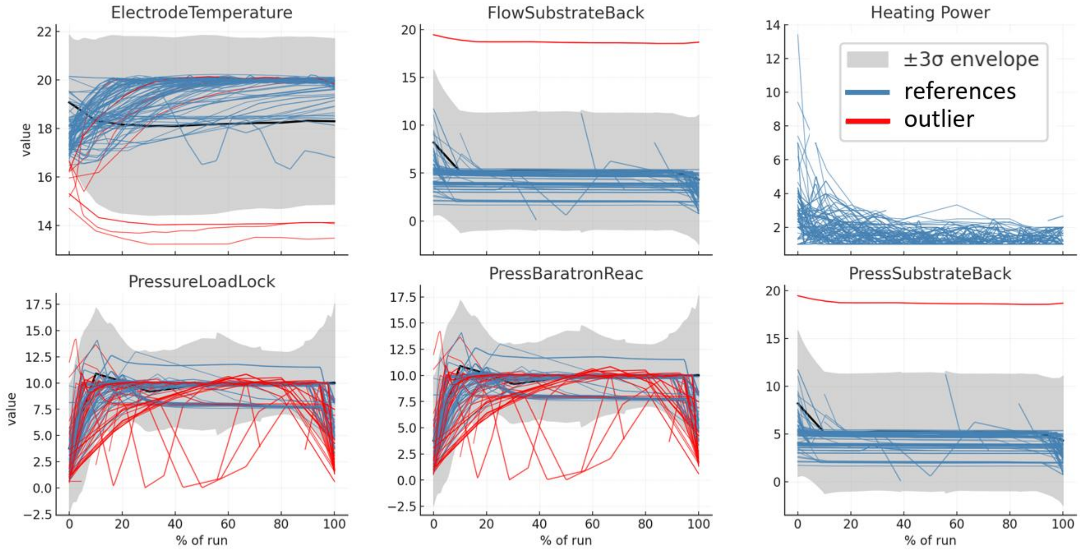

Fig. 2 The LLM has identified 157 runs processed with a given recipe on this particular machine. As the recipe has no fixed length, it was normalized and the run progress scaled from 0-100%. The 3-sigma envelope (grey shading) gives a good impression of the expected variations (healthy processes in blue). Processes that may have experienced deviations are shown in red.

4) Temporally align all selected runs, normalize to a common runtime if needed and compute a ±3σ envelope.
5) Generate plots and emphasize outliers

The entire workflow shall be driven solely through natural language prompts.

Approximately 5 MB of combined metadata and trace data were uploaded via the standard ChatGPT interface. The LLM (ChatGPT o3-pro) was prompted to perform the listed tasks.

2.2 Result and Discussion

Interpreting trace data, accurately mapping it to process runs, and establishing a reliable baseline are time-intensive tasks for human analysts. Furthermore, aligning these traces and normalizing their duration has always required manual intervention or custom algorithms. ChatGPT o3-pro, a model described as having "advanced reasoning capabilities", remarkably excels at these tasks. The model processed two years' worth of trace data, pre-filtered for a specific recipe, and successfully identified single-wafer runs in all cases, where the trace was within a 60-minute time window of the logged tracking times.

The model was able to distinguish and exclude incomplete processes, such as aborted runs or those with nonstandard profiles, and automatically filtered out data sets with unreliable tracking times, ensuring that only complete, valid traces were analyzed. The result is compellingly visualized in plots (Fig. 1 and Fig. 2). Runs that deviated beyond this envelope were automatically flagged as outliers and clearly annotated for further investigation.

Some of the tasks (run recognition, temporal alignment and skewing, plotting) were achieved in auditable Python code, that the model independently tuned both to the described tasks and the data analyzed.

The complete end-to-end analysis required less than 20 minutes, underscoring both the efficiency and power of agentic LLM-driven workflows—executed without custom coding, using only conversational prompts.

3 Outlook

We anticipate that agentic capabilities of LLMs - hosted on-premise or on trusted cloud-nodes and integrated with live access to databases and other organizational data sources (for instance using the model Context Protocol [3]) - will rapidly become standard tools in R&D and industrial environments. By combining historic and real-time data ingestion, these systems can continuously monitor process health, trigger alerts on emerging anomalies, and maybe at some point even propose corrective actions. The turnaround times for this type of analysis is reduced from days to minutes, and enables real time data-driven insights. Remaining challenges include reproducibility of results, which can often vary due to model stochasticity. These can be mitigated through advanced prompting strategies and the control of model parameters such as temperature.

As costs for running such models fall and tech-stack to integrate them mature, we expect such LLM-driven analytics platforms to transition from an experimental use towards a widespread tool, critical for advanced process control, continuous improvement and accelerating innovation cycles in manufacturing.

4 References

[1] X. Zhang, R. R. Chowdhury, R. K. Gupta, and J. Shang, "Large Language Models for Time Series: A Survey," arXiv preprint arXiv:2402.01801, 2024.

[2] Y. Jiang et al., "Empowering Time Series Analysis with Large Language Models: A Survey," arXiv preprint arXiv:2402.03182, 2024

[3] https://modelcontextprotocol.io/docs/getting-started/intro (retrieved on 31.07.2025)