Modeling Sensor Data For Knowledge Discovery And Explainable Decision-Making In Fruit Storage

Viacheslav Barkov

December 2023

1 Introduction and Previous Works

The cold storage of apples is a critical phase in post-harvest management, aiming to maintain quality and extend shelf life. Current practices often fail to dynamically adjust to the changing physiological responses of apples under storage and face challenges such as respiration rate fluctuations and moisture condensation [1]. The advent of sensor technologies offers novel ways to face these challenges and enhance the monitoring and control of storage conditions [2]. Prior studies have laid the groundwork by demonstrating the importance of precise monitoring using sensors [3]. To effectively manage the factors such as fruit respiration, atmospheric conditions, and surface wetness, the sensor technologies have emerged as essential tools for capturing critical data within the storage environment. These sensor-based systems are not only applicable in the experimental and research settings but are also scalable to real-life commercial storage facilities.

This proposal prioritizes the deployment and integration of two specific sensors developed by the Leibniz Institute for Agricultural Engineering and Bioeconomy (ATB): the Respiration Measuring Sphere (RMS88) and Wetness Sensor. These sensors have demonstrated feasibility in the real-time monitoring of critical parameters affecting the storage quality of apples.

1.1 Respiration Measuring Sphere (RMS88)

The RMS88 Respiration Measuring Sphere [4] [5] [6] [7] provides an in-depth view into the respiratory processes of apples. The respiration rate of apples is an indicator of their metabolic activity, which determines their post-harvest life and quality. Fluctuations in the respiration rate can lead to inconsistent quality of the stored apples, complicating inventory management and leading to potential waste.

In the previous works, the application of the respirometer involved measuring the respiration rate of Pinova apples in commercial controlled atmosphere storage for 32 days [4]. This demonstrated the system's ability to detect real-time changes in respiration rate and respiration quotient.

The RMS88 respirometer can be used for real-time assessment of respiration rates concentrations with the aim to determine the impact of extrinsic factors (such as temperature, gas composition, and storage time) on respiration rates.

1.2 Wetness Sensor

Water vapor condensation on the surfaces of stored apples presents a prominent challenge within storage scenarios, particularly because of its impact on product quality due to potential microbial growth [8].

To monitor these condensation events, a Wetness Sensor attached to the apple surface was developed. The study successfully identified water vapor condensation on the surfaces of stored apples and documented how the refrigeration cycle, comprising of cooling and re-warming phases, can be interrupted by defrosting processes and how this affects condensation. Different types of cycles were observed, with varying durations of condensate presence on the apple surface [8].

Measuring the condensation processes with the Wetness Sensor can be used to evaluate the impact of refrigeration system operation cycles on the condensation process and adjust the control strategies for refrigeration components to mitigate moisture condensation.

2 Proposed Project

How can real-time sensor data from advanced sensor systems like the RMS88 respirometer and Wetness Sensor be efficiently processed? To what extent can predictive AI and machine learning models improve the decision-making? How can we develop and validate a Digital Twin to refine storage conditions for apples?

The principal objective of this research project is to develop and validate the scalable, data-driven Digital Twin, utilizing a high throughput of streaming sensor-generated data and various AI and machine learning models to create a predictive, adaptable system for improving the long-term cold storage. By leveraging large amounts of streaming data from advanced sensor systems like RMS88 respirometer and Wetness Sensor, the system could be used for smart decision-making to continually refine storage conditions, ensuring apple freshness and extending shelf life, while also addressing energy efficiency, sustainability and explainability concerns.

The research will be grounded in multiple components: data acquisition, predictive modelling, Digital Twin, model explainability.

2.1 Data Acquisition

Data streams will be sourced from a network of RMS88 and Wetness Sensors deployed throughout the storage facility, collecting diverse metrics such as gas concentrations, temperature, humidity, and surface wetness in real time. This streaming approach will result in a dense, uninterrupted flow of information fundamental for further analysis.

2.1.1 Respiration Measuring Sphere (RMS88) Data Collection

Table 1: Example of data from RMS88 respiration measuring sphere

Date, Time	O_2, %	T, °C	P, Bar	CO_2, %	VCC, V	H, %
04.12.2019 16:10	7,17	5,4	1016	3,3509	4.075	86,2
04.12.2019 16:15	7,17	5,4	1016	3,3569	4.065	86,2
04.12.2019 16:20	7,16	5,5	1017	3,3579	4.072	86,3

Data from the RMS88 sensor (Table 1) provides insights into the respiratory function of stored produce. The following data is collected:

- Concentrations of Oxygen, which are measured using a fluorescence-based optical sensor within the RMS88, with a range extending from non-existent (0%) to a saturation threshold (25%).
- Ambient Temperature.
- Pressure, indicating of the gas exchange dynamics.
- Levels of Carbon Dioxide, featuring dual nondispersive infrared CO_2 sensors for precise readings in a lower range (up to 0.5%) as well as a higher range (up to 20%).
- Supply voltage metadata, indicating the operational status of the sensor.
- Relative Humidity.

The sensor's data collection can be timed from 1 to 60-minute intervals, tailored to research or monitoring needs. Data can be stored internally for durations extending up to 100 days due to its low power consumption [6]. The system also enables automated, real-time calculation of the respiration rate (RR) of stored produce. Respiration rates are quantified following Eqs. (1) and (2) below, derived from changes in gas concentration over time inside the controlled chamber, taking into account the free volume and mass of the stored produce [9]:

$$RO_2 = \frac{\Delta O_2}{100 \times \Delta t} \times \frac{V_{net}}{M_p} \tag{1}$$

$$RCO_2 = \frac{\Delta CO_2}{100 \times \Delta t} \times \frac{V_{net}}{M_p} \tag{2}$$

Here, RO_2 and RCO_2 denote the changes in Oxygen and Carbon Dioxide concentration (%), respectively; Δt represents the time interval; V_{net} is the net volume of the respirometer; and M_p is the mass of the produce in kilograms. The Respiratory Quotient (RQ) is determined as the ratio of RCO_2 over RO_2, providing additional insights into the metabolic processes occurring within the stored environment.

2.1.2 Wetness Sensor Data Collection

The function of Wetness Sensor is to monitor the moisture conditions on the surface, providing an electrical signal that correlates to the presence of water. The resistance changes when water connects the electrodes of the sensor, and as the water evaporates, the resistance increases until it indicates a dry state. This resistance-based measurement provides a timeline that illustrates the periods of wetness experienced by the fruit [8]. The concrete data provided by the wetness sensors include:

- Wetness Presence. An indication of when moisture is present on the surface of the fruit, suggesting condensation events.
- Retention Time. The length of time that moisture remains on the fruit surface, from the moment condensation forms until it completely evaporates.

Although wetness sensors do not directly measure other parameters, such as temperature, humidity, surface temperature and dew point temperature, they were also measured during the previous experiments [8].

- Relative humidity of the air was measured using digital combination sensors, specifically SHT35.
- Surface Temperature was measured with a contactless infrared thermometer, specifically the MLX 90416.
- Thermal images from ThermoCAM HD 600 camera were taken to visually examine the temperature distribution on the fruit surface.

This additional data can also be used in the research project and integrated with other sensor measurements to get a comprehensive understanding of the micro-environmental conditions affecting fruit surfaces.

2.2 Predictive Modelling

Important part of this research project is development of various AI and machine learning models with the aim for a resulting Digital Twin to be well-equipped to adapt and respond intelligently to a range of storage conditions and apple states. Some of the proposed approaches include anomaly detection, regressive modelling and time-series forecasting.

2.2.1 Anomaly Detection via Self-Supervised Learning

How can self-supervised learning techniques be effectively utilized to identify anomalies and prevent potential issues in the changing environment of apple cold storage? Given the complex variability in cold storage environments, it is important to rapidly identify conditions that deviate from the norm. Anomaly detection can be employed using self-supervised learning, a technique that learns from the data itself without explicit labelling of anomalies. The approaches could be adopted from different domains, like detecting credit card fraud [10] or network intrusions [11]. Developing these models will allow finding distinctions between normal operational patterns and outliers that could signify potential issues in storage, enabling preemptive corrective actions.

2.2.2 Regressive models

To what degree can AI and machine learning models based on historical and real-time sensor data predict key variables indicative of apple quality? A targeted approach will be developed to predict key variables indicative of apple quality, such as weight loss and ethylene production. By compiling historical and real-time sensor data, such as temperature, humidity, and respiration rates, regression models will be able to forecast these quality indicators.

2.2.3 Time-Series Forecasting

How can we understand the temporal patterns and forecast future conditions in apple storage? Many variables in apple storage consist of time-series data, recording changes over regular intervals. Time-series forecasting models will be developed to understand temporal patterns and forecast future conditions, providing valuable insight for dynamic storage adjustments.

2.3 Digital Twin

How can real-time sensor data and predictive models be combined into a comprehensive integrated system? The mentioned predictive models will be iteratively developed and integrated into Digital Twin [2] [12] [13], capable of managing the complexities of real-time apple storage. By integrating these models into a comprehensive Digital Twin, the project can not only predict but also prescribe solutions to optimize long-term storage conditions, ensuring the highest quality of apples and balancing economic and ecological sustainability.

By creating a Digital Twin, a virtual representation that mirrors the real-time state of the apple storage environment can be constructed. This virtual counterpart will not only showcase current conditions but also simulate and test various "what if" scenarios to inform strategic decisions without disrupting actual operations [2]. The predictive modelling and sensor data streams will feed into the Digital Twin, ensuring a seamless transition between virtual and physical environments.

Beyond predicting future states, the Digital Twin system will leverage the predictions to suggest optimal courses of action. This will include experimenting with different control strategies, observing the effect of varying refrigeration cycles, humidity levels, and gas compositions on the predicted quality indicators, seeking the most effective preservation methods.

2.4 Explainability and Decision Support

In what ways can we increase transparency and enhance the understanding of complex data-driven decision processes for end-users? The Digital Twin amplifies the explainability of underlying AI and machine learning models by providing a virtual platform where various scenarios and interventions can be simulated and analyzed. With a Digital Twin in place, the users can visualize how different storage strategies affect key variables such as ethylene production or apple weight loss, all without risking actual stored produce. This visual representation of model outcomes helps in decoding complex AI and machine learning decisions, illustrating the impact of various factors on apple preservation and revealing the reasoning behind specific AI and machine learning driven suggestions.

Through these mechanisms, the Digital Twin not only serves as an advanced tool for storage management but also as a medium to add transparency into complex data-driven decision processes. It ensures that the application of AI and machine learning to cold storage is accessible, explainable, and aligned with the operational understanding and goals of the end user. This integration of state-of-the-art models with human-centric interfaces amplifies the potential of smart agriculture technologies and their role in transforming the future of food storage.

3 Conclusion

The proposal is based on the alignment of current cutting-edge AI and machine learning advances with the necessity to innovate in cold storage practices. Using the prior work on sensors as a foundation, the project will implement predictive modelling and build a sophisticated Digital Twin to actively manage apple storage in real-time. This comprehensive Digital Twin has the potential to bring together sensory data and AI and machine learning models to create an adaptive, responsive storage environment capable of optimizing conditions for preserving apple quality and cutting down food waste. The explainability mechanisms will make such a system accessible and understandable for the end user, potentially creating a new wave of intelligent cold storage solutions.

References

[1] Pramod V Mahajan, Namrata Pathak, Graziele G Bovi, Makgafele Lucia Ntsoane, Ali Jalali, Nandita Keshri, Guido Rux, Ulrike Praeger, and Martin Geyer. Recent advances on packaging and storage technologies for the preservation of fresh produce. 2021.

[2] Reiner Jedermann, Kunal Singh, Walter Lang, and Pramod Mahajan. Digital twin concepts for linking live sensor data with real-time models. *Journal of Sensors and Sensor Systems*, 12(1):111–121, 2023.

[3] Yogesh Bhaskar Kalnar and Pramod V. Mahajan. *Sensors for Fresh Produce Supply Chain*, pages 1238–1244. Springer International Publishing, Cham, 2023.

[4] Pramod V Mahajan. Real-time measurement of respiration rate of fresh produce. In *Encyclopedia of Smart Agriculture Technologies*, pages 1–6. Springer, 2023.

[5] Nandita Keshri, Ingo Truppel, Werner B Herppich, Martin Geyer, Cornelia Weltzien, and Pramod V Mahajan. In-situ measurement of fresh produce respiration using a modular sensor-based system. *Sensors*, 20(12):3589, 2020.

[6] N Keshri, I Truppel, WB Herppich, M Geyer, and PV Mahajan. Modular sensor-based respirometer for real-time monitoring of respiration rate. In *VI International Symposium on Applications of Modelling as an Innovative Technology in the Horticultural Supply Chain Model-IT 1311*, pages 297–302, 2019.

[7] Nandita Keshri, Ingo Truppel, Werner B Herppich, Martin Geyer, Cornelia Weltzien, and Pramod V Mahajan. Development of sensor system for real-time measurement of respiration rate of fresh produce. *Computers and Electronics in Agriculture*, 157:322–328, 2019.

[8] Manfred Linke, Ulrike Praeger, Daniel A Neuwald, and Martin Geyer. Measurement of water vapor condensation on apple surfaces during controlled atmosphere storage. *Sensors*, 23(3):1739, 2023.

[9] Susana C Fonseca, Fernanda AR Oliveira, and Jeffrey K Brecht. Modelling respiration rate of fresh fruits and vegetables for modified atmosphere packages: a review. *Journal of food engineering*, 52(2):99–119, 2002.

[10] Jian Chen, Yao Shen, and Riaz Ali. Credit card fraud detection using sparse autoencoder and generative adversarial network. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pages 1054–1059, 2018.

[11] Monowar H Bhuyan, Dhruba Kumar Bhattacharyya, and Jugal K Kalita. Network anomaly detection: methods, systems and tools. *Ieee communications surveys & tutorials*, 16(1):303–336, 2013.

[12] Abozar Nasirahmadi and Oliver Hensel. Toward the next generation of digitalization in agriculture based on digital twin paradigm. *Sensors*, 22(2):498, 2022.

[13] Warren Purcell and Thomas Neubauer. Digital twins in agriculture: A state-of-the-art review. *Smart Agricultural Technology*, 3:100094, 2023.