Use of Internet of Things (IoT) technologies to monitor quality in fresh produce supply chains

Xuezhen Guo, Joost Snels and Seth Tromp

Abstract

The Internet of Things (IoT) is a revolutionary technology that will change the way how supply chain organizes in the future. Traditionally fresh produce supply chains are regarded as different to manage due to the perishable nature of products with quickly decayed quality. This product nature requires a dynamic decision-making process in the supply chain based on real-time product quality into account to realize the so-call quality-controlled logistics. Traditional IT solutions in the fresh produce supply chain lack the capacities of instantaneous data collection, transmission and translation, which hampers the realization of quality-controlled logistics. However, the emergence of IoT technologies with remoting sensing provides a new opportunity for companies to monitor product quality in the supply chain and make quality-informed logistic decisions. In this research, we have investigated the relevant aspects of IoT applications in fresh produce supply chains toward quality-controlled logistics. Key components of this innovation such as sensor selections, quality-decay models, integrated information platform, digital twin realization as well as soft enablers are addressed. We conclude that IoT-powered quality-controlled logistics has huge potential for efficiency improvement in fresh produce supply chains and concrete user cases at the business scale should be developed to demonstrate the validity of this innovation.

Keywords: Fresh produce supply chain, Quality-controlled logistics, Internet of Things, Digital twin

Corresponding author: Xuezhen Guo (Xuezhen.guo@wur.nl)

1. Introduction

Internet of Things (IoT) is an emerging technology that combines radio frequency identification (RFID), infrared sensors, global positioning systems, laser scanners, and other information sensing devices to intelligently identify, locate, track, monitor and manage a network with smart objects (Sun, 2012). IoT can impose significant impacts on logistics and supply chain management. With IoT applications, manufacturers/producers can send real-time information to their business partners about the status of their goods without human intervention (Minerva et al., 2015). The preliminary IoT application which aims to help data processing can already bring in significant benefits (Hribernik et al., 2010). It can facilitate strategic supply chain decisions by improving the visibility and traceability of the supply chains in a real-time manner. It can also help the optimization of the operational activities in the supply chains. For example, with IoT, port managers can optimally plan the usage of the port capacity and human resources to gain system efficiency (Lacey et al., 2015)

Although IoT has brought significant changes to some sectors, for the fresh food industry, there is not yet much research devoted to exploring the potential of IoT in enabling quality-controlled logistics. One of the understudied areas is related to IoT applications for the fresh fruit and vegetable sector of great social and economic significance. According to the world bank's statistics, the F&V sector accounted for 1.5% of the global GDP (2018) with a value of 1,249.8 billion US dollars (Fresh Plaza). Due to the highly perishable nature, about half of the global food losses and wastes were from F&V and mainly happened in the postharvest chains (Guo et al., 2020). As a result, quality-controlled logistics powered by IoT with a combination of sensors and quality decay algorithms (de Keizer et al., 2017; Jedermann et al., 2014; Van Der Vorst et al., 2009) provide ample opportunities to improve fresh food chain efficiency, reduce logistic costs and mitigate food loss & waste issues. Since food loss & waste is a major source of the void GHG emissions, reducing food loss & waste in fresh food chains via IoT applications also has significant impacts on climate change.

This paper addresses the IoT application in the fresh produce postharvest chains to enable quality-controlled logistics by combining sensor technology with quality decay models to predict product quality. This can help the supply chain practitioners to make informed logistic decisions based on real-time information, which can therefore improve the general performance of the fresh food business.

2. Sensing and quality decay modeling in fresh produce supply chain

Instead of reviewing general IoT technologies, in this section, we will address the literature on the basic components of IoT related to quality-controlled logistics. The interface area of sensor technologies to measure the environmental conditions of the products in storage and transportation and the quality-decay models that are used to monitor the product quality is the focus. A knowledge gap is identified based on this review.

The fundamental driver for developing quality-controlled logistics is the consumer's expectation for products with higher quality and longer shelflife (Smith and Sparks, 2004; Van Der Vorst et al., 2009; van der Vorst et al., 2005; van der Vorst et al., 2007). The underlying assumption for quality-controlled logistics is that even though the quality of the fresh produce is decaying along the supply chain, it is still predictable. As a result, companies can proactively control the good flows and design the logistics to improve product availability and quality to reduce losses and costs (van der Vorst et al., 2011). Quality-controlled logistics pertains to dynamic planning, implementing and controlling of the efficient goods and information flows between the primary food production to storage and transportation and finally reaching the point of consumption to meet consumers' requirements (van der Vorst et al., 2007). With the application of IoT powered by the sensor, communication, and computer modeling technologies, advanced logistic decision-making based on real-time product become feasible which lays the function for quality-controlled logistics (van der Vorst et al., 2011; Verdouw et al., 2013).

Sensor and quality decay models are two critical components in establishing quality-controlled logistics. We will therefore dive a bit in-depth into these two topics in the following two subsections.

2.1 Sensors for data collection in postharvest chain

We start with a recent paper (Ben-Daya et al., 2019) which has conducted a comprehensive study on IoT applications in supply chain management and uses it as the basis for snowballing.

Time Temperature Integrators or indicators (TTI) are the mostly applied sensors in the postharvest supply chains and are considered cost-effective in monitoring the environmental temperature of fresh products (Taoukis and Labuza, 2003). This is on the one hand because the temperature is the major factor that affects the shelf life and quality of the fresh produce and on the other hand, TTI is also a relatively cheap sensor compared to other more sophisticated sensors such as gas sensors. A study that evaluates the applicability of TTI to control frozen vegetable chains was conducted by Giannakourou and Taoukis (2003) with the aim to optimize consumer-end product quality. It demonstrates that the information logged by TTI is a more reliable indicator than quality measurement conducted in the traditional way. The authors also studied the least shelf-life first out (LSFO) inventory issuing policy based on the information collected by TTI with a Monte Carlo simulation model. The results show that significant benefits can be delivered by the TTI-enabled shelf-life-driven system with respect to improved product uniformity and quality, low product rejection and higher consumer satisfaction.

Another simulation study was conducted by Dada and Thiesse (2008) to investigate the effects of a TTI-based inventory control system to issue the products according to the shelf life information. They claim that with the automatically collected expiry date of the products, the system can improve the average product quality at the consumer end and improve quality uniformity. Shih and Wang (2016) have developed a conceptual framework to guide the design of a TTI-based cold-chain system consisting of wireless sensors applied against Critical Control Point (CCP) criteria in the whole supply chain. The proposed approach has been proven to be able to increase sales, reduce energy use and extend distribution channels.

Although TTI provides essential information on fresh produce quality development, it can not explain the whole decay process because temperature and time is not the only factor that determines product quality. Therefore, in addition to the literature on time-temperature sensors, other types of sensors have been also studied. For example, Abad et al. (2009) integrated multiple types of sensors to measure the parameters including temperature, relative humidity and light in both laboratory and field to see if it can deliver better performance on quality tracing in the cold chain. The experiment shows that the developed smart-tag system outperforms the conventional traceability tools because it has more memory, better reusability, high automation and high robustness to extreme environmental conditions. The advantages of including wireless humidity sensing during fruit transportation and storage were investigated by Ruiz-Garcia et al. (2008). They especially focus on battery life analyses and tested the reliability of communications and measurements under cooling conditions. Draganić et al. (2017) investigated a compressive sensor that targets to detect the CO2 levels inside the modified atmosphere packages for table grapes along the cold supply chain. The sensor is claimed robust enough to maintain good monitoring during the situation of communication system overloading. Sklorz et al. (2012) developed a new sensor to detect ethylene in fruit logistics and shows that it ensures 16 times higher detection sensitivity for ethylene changes.

More general studies without focusing on certain types of sensors but the impacts of adopting sensing technologies in postharvest supply chains have been also observed in the literature. For example, Bogataj et al. (2017) designed a coupled cyber-physical system from the extension of EMRP to enable the FEFO stock issuing policy. The system can enable real-time dynamic decision-making for quality-controlled logistics such as automatic rerouting based on the expected product shelf life. Jedermann and Lang (2008) find that costs are not the only barrier to supervising quality changes with ubiquitous sensors but also energy consumption. This also explains why temperature and humidity sensors are preferred in practice because they require the least energy and the resolution of data collection is several minutes. Haass et al. (2015) conducted research on the intelligent container in the banana supply chain to see if automatic decisions can be made by the containers themselves according to the production quality situation. The results are encouraging because the intelligent containers were shown to be able to assure high quality, low losses and low GHG emissions. It can help supply chain practitioners, especially the trading companies to improve business efficiency and sustainability. In addition to intelligent containers, intelligent packaging can also bring significant values to the fresh produce supply chains. Intelligent packaging can automatically sense and adjust the micro-environment of the product within the package to preserve product quality. Through conceptual analysis and computer simulation, Heising et al. (2017) demonstrate the usefulness of intelligent packaging in reducing food waste, especially when dynamic pricing based on the predicted expiry dates is implemented.

2.2 Quality-decay modeling

Only gathering the environmental data about the fresh produce in the postharvest chain is not enough. The raw sensor data should be translated into quality indicators that make sense to supply chain practitioners with the help of the quality decay/prediction models. The quality indicators could be some concrete fruit-related parameters such as firmness, color, starch & sugar contents. They can also be some fabricated indexes such as visual deterioration. These quality indicators are either difficult to measure (or can not be directly measured at all) and therefore need a mathematical algorithm to transform the raw data. In such a sense, we actually adopt the so-called "soft sensing" technology to do the quality decay modeling.

Fresh produce quality decay is essentially from a series of biochemical reactions within the fresh products governed by the kinetic process under the Arrhenius law which specifies the temperature dependence of reaction rates. Therefore, temperature and the associated time duration are the major drivers for the quality decay of fresh produce. As a result, the most applied models in quality decay specification are the mechanistic (kinetic) model which specifies the relationship between the quality indicator and time-temperature with a logistic curve (e.g., Tijskens and Polderdijk, 1996; Tromp et al., 2016; Tromp et al., 2017; Tromp et al., 2012). The strongest point of the classical kinetic model is its biological foundation (Tijskens and Polderdijk, 1996) because, in the natural world, logistic kinetics behaviors are very often observed. The major drawback of the classical kinetic model is its fixed model structure, which makes it difficult to incorporate the extra explanatory variables which may help to explain more data variance. It is also not good at capturing the with- and between-bath variety of fresh produce. Moreover, the classical kinetic model with ordinary differential equations usually does not have a closed-form solution and therefore has to be solved numerically (e.g., Abad et al., 2009; Hernández et al., 2021) unless some extra rigid assumptions such as constant temperature and gas condition are made (Xiao and Li, 2022). As well known to all, the numeric solution is more time-consuming to obtain compared to the analytical solution.

Given the aforementioned drawbacks, other data-driven approaches have been also introduced to model the fresh produce quality decay. For example, Lammertyn et al. (2000) have employed multi-nominal logistic regression to flexibly incorporate more variables to capture more within and between batch product variability for conference pear quality prediction. Other examples include Zhu et al. (2017) which predict the kiwifruit quality attributes using effective wavelengths with multiple linear regression, partial least squares regression and least squares support vector machine. A similar study

(Li et al., 2017) used hyperspectral imaging to predict the quality of the kiwifruits with partial least square and support vector machine regressions. A big drawback of those data-driven models is that there is no biochemical foundation to back up the model formulation, which makes the modeling process a bit arbitrary. Moreover, most of the data-driven models are essentially generalized linear models. They are not good at capturing the complex non-linear relationships between dependent and independent variables. This makes the data-driven models easily omit the structural relationship specified by the classical kinetic models. To overcome this issue, the hybrid model which combines the data-driven model with the mechanistic model has also been developed. For example, Xiao and Li (2022) integrate the closed-form solution of the kinetic model with the smoothed near-infrared spectral curve by functional regression to predict the firmness of the products. The hybrid model is proved to be better than the generalized linear data-driven models.

The classical kinetic model and the generalized linear models are essentially parametric, which means the model structure and parameters do not change as the data change. There also other non-parametric methods such as Bayesian Networks (e.g., Zhang et al., 2020) and Neural Networks (Chen et al., 2001; Singh et al., 2009; Soltani et al., 2015; Wang et al., 2017) were developed for quality decay modeling. Such machine-learning models have the advantages in terms of flexibly modeling the complex non-linear relationship between parameters, which may enable a better fit of the model to the data. The disadvantages of the non-parametric machine-learning models are also obvious. First, such models are data-demanding and usually require a long computational time to estimate the model parameters. Second, those models have bad explainability and are essentially black boxes.

3. IoT application in the quality-controlled logistics

3.1 Implementation at the operational level

The IoT implementation with sensor-based quality control is most relevant to be applied at the operation level to improve the efficiency of daily logistics. Figure 1 demonstrates a generic procedure to be followed for daily logistic operations in a postharvest supply chain.

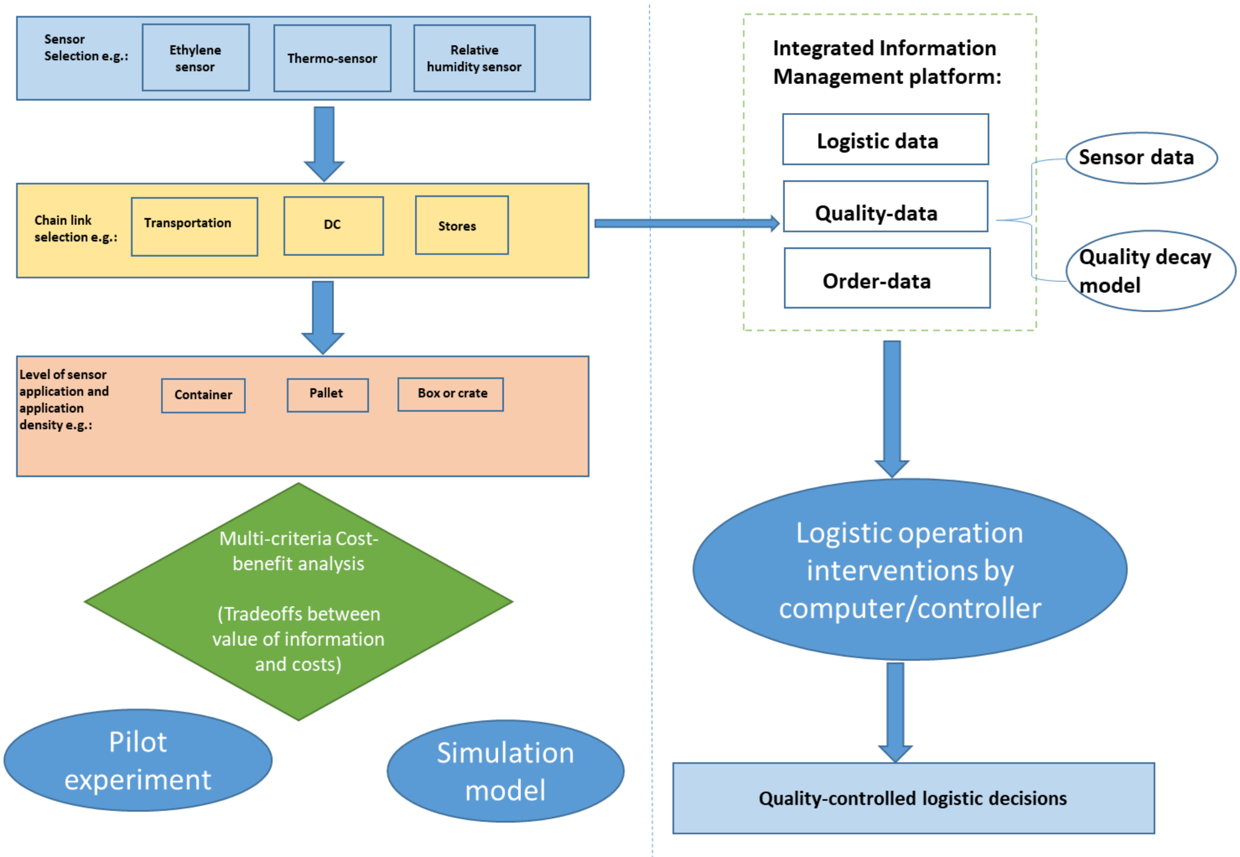

Figure 1. The generic procedure for the application of sensor-based quality-controlled logistics at the operational level

3.1.1. Sensor selection and application

Before implementing quality-controlled logistics in a real business setting, careful evaluation of different sensors and ways of applications is essential.

First of all, the type of sensors needs to be selected. This selection should depend on the most relevant information that we want to measure and the costs of installing the sensors. A good balance must be maintained between the marginal value of extra information and the extra costs associated with it. In general thermos and humidity sensors are cheap and most relevant for quality decay, and therefore should be applied in most cases. The gas sensors are more expensive but can also provide more information associated with product physiological change. For example, ethylene concentration is a strong indicator for climacteric product ripening and its added value should therefore be carefully analyzed. For most of the climacteric products (such as mongo, and bananas), ethylene information is valuable and therefore is more justified to include the ethylene sensors. A more detailed illustration of sensor selection and calibration from a statistical point of view will be addressed in the digital twin section (section 4).

The location of sensor application is relevant because there are many stages of the postharvest supply chain where sensors can be applied. Traditionally sensors are normally applied to the storage and transportation stages. For storage, the sensors stay in one place and data transmission is therefore not a problem. For transportation, since the sensors are moved together with the products (Verdouw et al., 2013), its connectivity is heavily relying on if there is a base station in the neighborhood. Therefore, it is very common to have a disconnection problem during the shipment.

The sensors can also be distinguished if it is fixed in a location of the storage room or the container. A common drawback of the fixed-location sensors is it measures the parameters related to the macro environment within the room or container but not the micro-environmental parameters close to the products. This can bring in structural measurement errors for the follow-up modeling. The mobile sensors are more flexible in terms of application location and therefore can be put closer to the products, which provides more accurate product-related parameters.

With respect to data transmission, sensors (loggers) that can not perform real-time communication are mostly used to conduct post-ante analysis (Scheer et al., 2011). Low costs and easy-for-application are the advantages of those sensors but since there is no real-time data transmission, they can not help supply chain practitioners make on-time logistical interventions. On the other hand, the sensors with real-time communication capacity require more energy to power the data transmission (so battery life is important), which also increases the costs. Tradeoffs must be made between the gains from using the real-time information and the costs and energy consumption associated with the real-time sensing.

After selecting the chain link or location to install the sensors, the next step is to determine the level and density of the sensor application. Sensors for quality control can usually be applied at multiple levels including the room/container level, pallet level, box (or crate) level and product level (theoretically possible but rarely implemented in practice). The density of the sensor application at the selected level will determine the accuracy and precision of the product quality prediction because of the intrinsic biological variation between products and batches. High-density sensor application on the one hand can capture more within- and between-batch variances of the products and on the other hand, it can also filter out the noise through sensor/data fusion to get closer to the true value of measured parameters. For example, simply averaging the temperature values of all the sensors within a batch will smooth the temperature curve along the time dimension and make it a more accurate estimator for the mean quality change of that batch. The disadvantage of the high-density application is again related to the extra costs and energy consumption compared to the low-density application. Another justification for high-density applications is the malfunction of the sensors, especially during project transportation. In practice, it is very often to have a situation where a part of the sensors stops recording and transmitting the data during the shipment due to reasons like low battery, deep position in the pallet, etc. As a result, having more sensors will give us some buffers to keep the system functional.

3.1.2. Multi-dimensional cost-benefit analysis

As is shown in the previous section, sensor selection needs to be addressed from multiple angles. Separate assessment against an individual criterion is relatively simple but can not capture the tradeoff and interaction between different criteria. To do a simultaneous evaluation against all criteria, a cost-benefit analysis with computer simulation supported by a real-life pilot should be conducted. Computer simulation has the advantage of flexibly evaluating different scenarios without implementing costly real-life experiments. It can also help to scale up the results by simulating a simplified real-life business environment. The real-business scale implementation is too risky for an unproved technology and companies are usually not willing to take the risk before an in-depth evaluation has been conducted. Computer simulation can partly fill the gap. However, since the simulation model is a simplification of the real-life situation, it must be combined with a physical pilot study to correct the errors of the model so as to amplify its effectiveness. Such kind of pilots can be of small size (e.g., several shipments) but can already bring in significant value to calibrate the simulation model.

The traditional cost-benefit analysis mainly focuses on evaluating the monetary values of different supply chain options. However, more and more attention has been drawn to the sustainability aspects including food security, climate change, biodiversity loss, etc. in the tradeoff-balancing equation. One common method to quantify such impacts is to calculate the footprint of the relevant parameters (e.g., Greenhouse emission footprint, biodiversity loss footprints) per unit product along the whole supply chain. Using this method, we can compare the costs and benefits of different sensor application strategies in a comprehensive way.

3.1.3. Integrated information platform

Through the comprehensive cost-benefit analysis, the most suitable sensor application strategy can be chosen. The next step is to implement the selected strategy in a real business environment. To successfully achieve this objective, there must be a centralized information management platform that integrates real-time information on order data, logistic data and quality data. Such integration will help companies to explore the opportunities in their supply chain system and design effective interventions to enable quality-controlled logistics. For example, in the context of long-distance international transportation, knowing the expected fruit quality at arrival in the early stage of the chain will be of great help to determine potential interventions such as shifting modality, switching distribution channels, planning sales programs, etc. With smart remote-control technologies, people can even change the temperature and humidity of the reefer container during transportation. In such a case, real-time sensing will enable a dynamic adjustment of the environmental parameters in the container to improve quality preservation and reduce costs as well as energy consumption. In the wholesale and retail environment, real-time sensing provides opportunities for companies to perform quality-based inventory control. For example, the first-in-first-out (FIFO) stock issuing policy can be replaced by the lowest-quality-first-out policy (LQFO) to minimize food loss and waste.

As mentioned in section 2.2, the quality decay model is the tool that translates the environmental parameters or product phenotyping information into quality indicators (e.g., firmness, visual deterioration). The technique that converts the easy-to-measure parameters to hard-to-measure parameters is named "soft sensing". Quality decay modeling is therefore a type of soft sensing technique. Developing suitable algorithms catering to the nature of the data collected by the sensors is crucial for the success of quality-controlled logistics. For example, functional data such as infrared spectral data are high-dimensional ordered data. The ordering of the observation itself retains also important information. Simply applying the classical linear regression model will lose the ordering information, which is therefore not an optimal modeling technique for such data. A barrier to applying different modeling technologies in quality-decay modeling is the path dependency of modelers from different backgrounds. For classical mechanistic modelers who have strong model-driven thinking, the value of the data-driven models is often questionable because they consider the data-driven model as a lack of underlying process. Modelers from a statistical background, usually regard the non-parametric machine-learning models as with low generalizability and even often biased in a strict statistical sense. Such methodological discrepancy leads to confusion among the end users, which hampers the adaptation of the quality-decay model in the industry. Therefore, selecting a suitable modeling approach that is accepted by different business partners and at the same time can make full use of different types of data is extremely challenging but must be addressed. Only then, a useful and reliable quality decay model can be constructed to convert the sensing parameters to the business-sensible product quality parameters so that it can be employed in daily business practice. Moreover, in a real-time sensing and decision-making environment, modeling approaches with high computational efficiency are required. In this sense, parametric methods could be preferred over non-parametric approaches.

After obtaining product quality parameters, algorithms that can integrate the information from the ordering, logistic and quality datasets in a computationally efficient manner must be developed to ensure instant parameter optimization. This is because the optimization opportunities can only be identified with a quick and accurate combination of information. Interventions based on the identified opportunities should be implemented by the joint efforts of automated computer protocols and human supervision to improve the efficiency and resilience of the real-time decision-making process.

3.2 Strategic network design with accumulated quality data

Quality-controlled logistics with IoT can also bring long-term benefits with respect to strategic network design. Without major mergers and acquisitions, large companies usually re-evaluate their supply chain network with certain frequencies (e.g., every 5 years) based on the cumulative data from the last time window. Traditional network design usually does not take into account the product quality distribution within the network because of the lack of such data. Quality-controlled logistics can fill this gap. After a while of the quality-sensing application, large amounts of product quality data can be accumulated. With such data, we can derive rich insights into the quality decay processes (via the mean and standard division) in each node and lane of the supply chain network. The network reconfiguration model can thereby consider the quality constraint as an extra limiting factor for network structure optimization. For example, based on the average decay processes of the fresh produce from the origins to the destinations, the company can (re-) design an optimal supply chain network structure with the best depots' locations that minimize the overall volume-quality-corrected shipment distances. It can also help to (re-) direct good flows within a fixed multi-level network according to the quality on top of volume consideration (e.g., planning the flow consolidation taking product quality into account).

4. Digital twin

A higher level of thinking for IoT applications in quality-controlled logistics is to create a digital twin that enables the interactions between the physical twin (quality distribution in the container or warehouse) and a digital twin (computer visualization of the quality distribution for a simulated container or warehouse). A digital twin can create a more interactive interface between the human and computer which can significantly increase the efficiency of quality-controlled logistics.

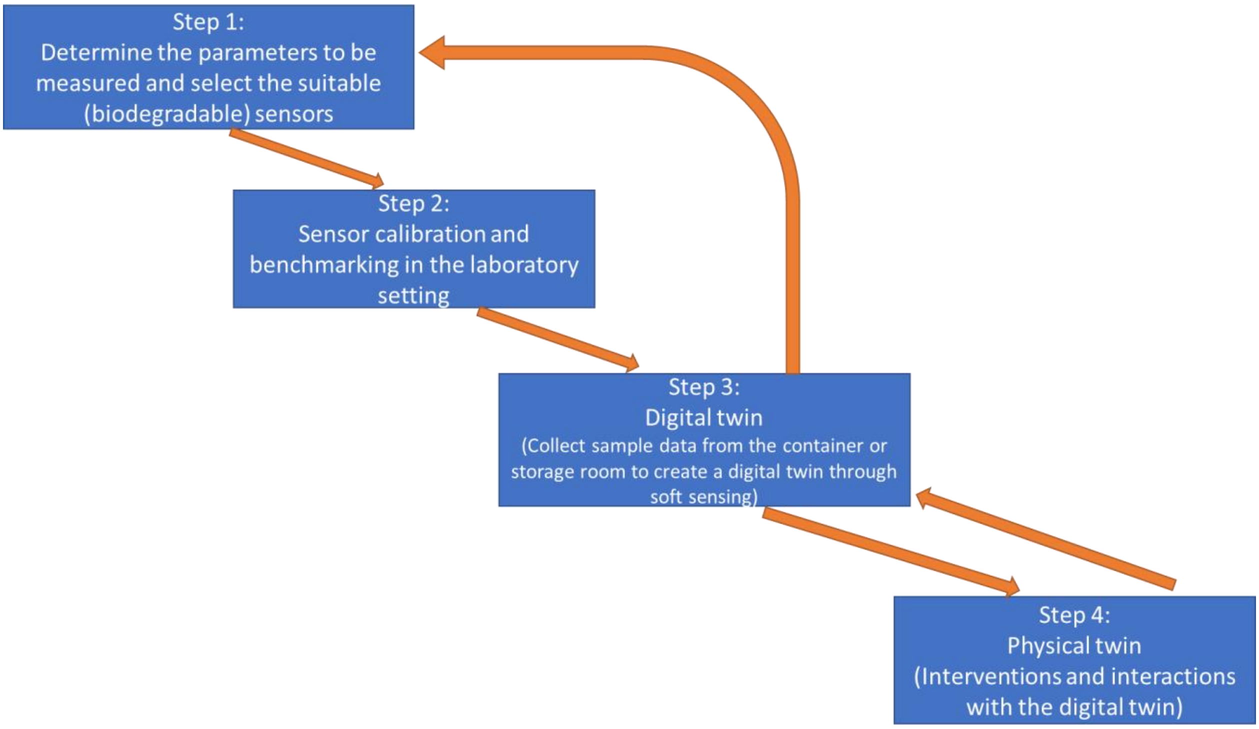

Figure 2. Steps for developing a digital twin in the context of quality-controlled logistics

Figure 2 presents a graphical overview of the project including four iterative steps. The details of each step are presented as follows:

Step 1: Determine the parameters to be measured and select the suitable sensors

Step 1 is related to subsection 3.1.1 but with more considerations from a statistical point of view. Before starting to analyze any sensors and their potential applications, we should always first ask ourselves what kinds of data we want to collect and why we need to collect them. Only after that, we can then select suitable sensors according to our ultimate goal. For example, when many parameters are correlated to the response (dependent) variable, we may only need to measure the major ones (which can already explain a large portion of the variance) and use them as explanatory (independent) variables to make the prediction. If temperature and time are highly correlated to the quality indicator for some product, say they can already explain 90% of the total variance of the data, then it may not be necessary to add other types of sensors. We may also develop models to structurally analyze trade-offs between the marginal gain (added value) of the newly measured variables and the efforts (costs) in collecting the extra data.

Step 2: Sensor calibration and benchmarking in the laboratory setting

After selecting the suitable sensors (based on the parameters to be measured), we need to do sensor calibration and compare them with the well-calibrated sensors for benchmarking purposes against a list of criteria in the laboratory setting. The criteria could include “accuracy”, “precision”, “resolution”, “linearity”, etc. Accuracy can be analogous to the concept of “unbiasedness” in statistics which reflects if the sensor can on average measure the true parameter value of the underlying population from the samples. Precision can be analogous to the concept of “efficiency” (related to the sampling variance). When two sensors are both able to measure the true value of a population parameter, the one with a smaller sampling variance is more efficient. The efficient sensor enables more accurate parameter estimation with the same sample size compared to the inefficient one. The resolution of the sensor is another important criterion because it measures the frequency of the data collection, which affects the levels of detail that can be derived from the data for the follow-up modeling. Linearity is also relevant because we want to have a linear relationship between the sensor-measured values and the true values for more convenient extrapolation. If the relationship is essentially non-linear, we may need to linearize the relationship.

To statistically calibrate a selected sensor with the benchmarking sensor, we can apply the paired sample t-test for a pair of sensors with repeated measurements. Based on the paired sample distributions, the t-statistics, and corresponding p-value can be calculated to see if there is a statistically significant difference between the measured values from the two types of sensors. Moreover, to capture individual sensor quality variability, we can employ the generalized linear mixed model (GLMM) where the between-sensor variation is modeled as a random effect. This will increase our ability to describe how fixed effects (i.e., the types of sensors) relate to the outcomes and allow simultaneous comparison of multiple selected sensors to the benchmarking sensor.

Step 3: Digital twin (Collect sample data from the field to create a digital twin through soft sensing)

Steps 3 and 4 adopt a “digital twin” approach. The digital twin and physical twin are interlinked and form a feedback loop to continuously improve the computer models and the quality-controlled logistics practices in the field.

Step 3 focuses on the construction of the digital twin whereas step 4 addresses the interventions on the physical twin as well as its interaction with the digital twin.

To create a digital twin that represents the quality distribution within a space (e.g., a container or storage room), we can homogeneously deploy the sensors (validated from the previous steps) in the space with a certain density to collect the sample data. Since the real-world conditions will be different from the well-controlled laboratory conditions, we must take the prominent real-world disturbing factors into account. One key factor is the spatial autocorrelation in different locations of the space due to the spatial dependency within the space of the storage room (or container). For example, suppose we want to use an easy-to-measure variable (X) (e.g., temperature) as the explanatory variable to predict a difficult-to-measure variable (Y) (say product firmness) at different locations of the storage room (or container). Simply regressing Y on X will give biased estimations on the intercept and slope coefficients of the linear model. To obtain the real effects of X on Y, we need to filter out the spatial dependence between y_i at different locations using spatial autoregressive models (e.g., spatial lag (or error) regression, and spatial random effect models). This lays the foundation to create a good digital twin taking into account spatial autocorrelation in the storage room (or container).

Another practical issue related to mapping/visualizing the product quality distribution in a certain space is the density limitation for sensor application. The deployed sensors can only cover a limited number of locations in the space due to practical constraints such as costs. To create a digital twin that represents the overall quality distribution of the entire storage room (or container), we can use the developed spatial autoregressive models to predict the *y* values at the locations without sensor deployments (or alternatively do kriging interpolation based on the spatial autocorrelation structure). In this way, we can create a product quality distribution map of a higher resolution, using the data from the same number of sample (sensing) locations.

Since the parametric spatial autoregressive models are essentially generalized linear models (GLM), they have difficulties in capturing the non-linear relationships between parameters. Moreover, the model specification becomes highly complex if we add the time dimension by simultaneously considering the spatial autocorrelation and serial correlation. In such a case, we actually conduct the dynamic spatial panel regression which is a hard case of the spatiotemporal regression. However, due to the linearity and complexity of the classical spatiotemporal regression, the model may not be able to accurately map the real spatial quality distribution situation through time. To create a more accurate digital twin, we can consider employing the "transfer learning" technique to integrate the parametric spatial autoregressive model with the non-parametric machine learning model, which is good at dealing with non-linear relationships. The process is as follows:

1) We first use spatial autoregression to map the spatial quality distribution of a storage room (or a container) and treat it as a "synthetic dataset" which is widely used for machine-learning model training.

2) Then, we can use the "synthetic dataset" to train e.g., a "Convolutional Long Short-Term Memory" (ConvLSTM) model which is a type of recurrent neural network designed for spatiotemporal prediction to further update the digital twin.

The benefit of the "transfer learning" approach is it can combine the strengths of both types of models. Since the neural network is good at capturing the complex non-linear relationship between parameters, it may help to extract more useful information from the data to improve the accuracy of the digital twin on top of the results from the classical spatial autoregressive model.

After obtaining a reliable digital twin, we can use it to simulate real-life scenarios without doing a lot of costly field experiments. For example, we can simulate the sensor failure (e.g., no data collected or bad data quality) in certain locations of the storage room (or container) and to see its impacts on the output variable.

Step 4: Physical twin (Interventions and interactions with the digital twin)

Developing a good digital twin is not the end of the game because the ultimate goal of using a digital twin is to guide the interventions of the physical twin. Therefore, we should create a feedback mechanism that can utilize the insights from the digital twin to optimize quality-controlled logistic activities. For example, logistic controllers can remotely change the environmental parameters to adapt to the new situation suggested by the digital twin. After the interventions, the updated physical twin information will be available through the sensor network and used to upgrade the digital twin again. If the physical twin shows that the current digital twin misses some important parameters, we can also get back to step 1 to see if more information needs (and can) be collected by other types of sensors and restart the whole process. In this way, an iterative feedback loop will be created, which will continuously refine the digital and physical twins to improve the performance of the whole system.

5. Soft aspects of the enabling environment

Previous sections focus on addressing the infrastructural and technical aspects of IoT-enabled quality-controlled logistics. However, as is for many other innovations, technology itself will not ensure a successful implementation of quality-controlled logistics. The soft aspects such as environmental, social and governance aspects (ESG) should be also taken into account.

Different from floricultural products, fruits and vegetables do not have a universally accepted sectorial standard (related to supply chain governance) to evaluate their quality. Every company or supply chain has its own way of product quality evaluation (Verdouw et al., 2013). Even for the more matured floricultural sector, data interoperability is a big problem that hampers the integration of different systems, for example, the integration of logistic systems with product quality control systems (Verdouw et al., 2011). Therefore, establishing sector-wise standards for fruit and vegetable products is the prerequisite for the successful implementation of quality-controlled logistics because supply chain partners should speak the same language with a common understanding of product quality.

To establish the standards, data sharing is key because a good understanding of the quality parameters of different supply chain players is the basis for quality indicator development. However, enabling data sharing is not an easy task. It requires a high level of trust (related to the social aspects) between supply chain partners because product quality data are usually very sensitive to fresh produce business, which makes the companies reluctant to share the data. One way to improve trust in data sharing is to apply the so-called federated learning approach (e.g., Li et al., 2020; Niknam et al., 2020). With federated learning, the model will learn the data from different companies separately to estimate the model parameters and make sure no raw data will be copied from the IT system of the companies. This technology can significantly improve the trust between different supply chain partners and provide the basis for data sharing.

Environmental aspects such as greenhouse gas (GHG) emissions from the postharvest supply chain also need to be considered because cold chain application in general pertains to high energy consumption which has negative impacts on climate change. Quality-controlled logistics with IoT require extra energy to power the sensors and real-time data transmissions and can be regarded as negative for sustainability under certain conditions. However, since the avoided food loss & waste (FLW) due to IoT application could also mean reduced carbon footprints associated with the lost food, a net result when taking both into account may be positive. Moreover, even though quality-controlled logistics adds extra energy for real-time sensing, with the possibility of real-time dynamically adjustable temperature, it may consume less energy compared to traditional cold chain logistics with a constant temperature.

Finally, the quality-controlled logistic system must be resilient, which means a certain level of human intervention still is still necessary. The advanced IoT system should be highly intelligent and can actively make decisions by itself. However, considering the development level of the fresh food supply chains with many human practitioners and the non-standard nature of fruit and vegetable products, a fully automated system can hardly be implemented in practice. Certain levels of human interventions are still necessary to improve the resilience of the postharvest supply chains to avoid the collapse of the system in some unexpected situations. Involving human decisions can retain job opportunities in the postharvest chains, which is good from a social perspective.

6. Discussion

In this study, we have investigated IoT applications in postharvest fresh produce chains to monitor product quality and enable quality-controlled logistics. Fresh produce such as fruits and vegetables are by nature perishable. Their quality after harvest will go down and can not be recovered once lost. Therefore, the primary goal of postharvest management is to preserve product quality as much as possible so as to maintain the retail price of the products, reduce food losses and improve consumer satisfaction. However, it is easy to say but hard to do. One of the big barriers to product quality preservation is it is so much dependent on the environmental conditions and the product's intrinsic nature with high individual variations. This makes the prediction of product quality based on the initial product parameters at harvest very unreliable. To solve this problem, real-time monitoring of the environmental conditions of the products in the postharvest supply chain is necessary. This brings IoT technology into the picture.

The two most important components for IoT-enabled quality-controlled logistics are 1) the sensors to collect the data and 2) quality-decay models to convert the collected data into useful quality parameters. There are many different types of sensors are available for data collection in the postharvest supply chains. The types of sensors to be selected are not standard. A tailor-made solution should be developed based on specific features of the products and requirements of the business. For the business targets in the high-end market, more sophisticated sensors with higher costs may be justified while for the low-value products, standardized cheap sensors may be more appropriate to be applied. Similar logic can also be applied to selecting the location, level and density of the sensor application. High-value products desire a high-density sensor application at the lower unit level (i.e., box level instead of pallet level) in more locations of the supply chains. Moreover, the way of sensor application should also take into account the variation of the products. Namely, for the batch of products with higher variation, more sensors should be applied at a lower level to obtain more reliable information with the same level of accuracy.

Quality decay models are the algorithms that translate the easy-to-measure parameters (e.g., temperature, humidity) into the difficult-to-measure (but practically relevant) quality indicators (e.g., firmness, visual deterioration). Such algorithms can have different forms ranging from mechanistic models to statistical/machine learning models with the aim to capture as much as possible the explained variances of the data to ensure an accurate quality prediction. On the other hand, since the collected data are essentially a sample of the fruit population, a too-good fit will lead to the classical overfitting problem at least in the short term when the sample size is small. To overcome this short-term problem, a regularization term that punishes the number of parameters used in the model can be added especially for the high-flexibility machine learning models. In the end, we still need to use cross-validation to test the model’s predictability on a different validation dataset from the training dataset.

The cost-benefits of sensor application strategy need to be evaluated in a structural way before implementing it in the practice. Computer simulation combined with small-size pilots needs to be carried out to compare different scenarios. The added values of extra information collected by the sensing system need to be quantified and compared with the extra costs from different perspectives. This will reduce the scaling-up risks in the real business environment.

The implementation of IoT-based quality-controlled logistics requires seamless integration of order data, logistic data and quality data in one centralized information management platform. Such a platform should be smart enough to spot the fleeting opportunities from the integrated data and propose on-time intervention strategies to improve the performance of the quality-controlled logistic system. Long-term data accumulation from the quality-controlled logistic system can also bring strategic benefits to fresh produce supply chains. Companies can have a better understanding of product qualities at different locations and shipment lanes within their distribution network, which allows them to perform quality-informed supply chain network design.

The advanced level of IoT application complies with digital-twin thinking. A digital twin is a computer-simulated representation/visualization of the physical twin. The “twins” are constantly interacting with each other. In the context of quality-controlled logistics, a digital twin maps the fruit-quality distributions within the spaces of e.g., cold-storage rooms and reefer containers throughout the fresh produce supply chains. The insights derived from the digital twin are fed back to the physical twin through real-time logistical interventions to change the real quality distributions in the spaces. In changed quality distribution will be captured again by the digital twin through remote sensing. In such a way, a constantly ongoing refinement mechanism on the twin systems will be established to continuously improve the quality-controlled logistics in the fresh produce supply chain.

It has been widely believed that technologies and infrastructure themselves will solve the problems in fresh produce supply chains. The enabling environment with essential soft aspects needs also to be in place. Sectorial standards, trust & data-sharing schemes, social & environmental impacts must be carefully analyzed to ensure the successful implementation of quality-controlled logistics.

To conclude, quality-controlled logistics with IoT has great potential of improving fresh produce chain efficiency at both operational and strategic levels. Remote sensing technologies, quality-decay models, an integrated information platform, real-time logistic interventions (with a digital twin) are the critical technical components, which need to be combined with the soft enablers to make the system functional. Future research should focus on developing concrete user cases at a business scale to demonstrate the usefulness of the concept described in this research.

Acknowledgment

This paper is a research output of the dinalog project "Quality Controlled Logistics in IoT-enabled Perishable Supply Chains" sponsored by NWO (Netherlands Organisation for Scientific Research).

Reference

Abad, E., Palacio, F., Nuin, M., Zárate, A.G.d., Juarros, A., Gómez, J.M., Marco, S., 2009. RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain. Journal of Food Engineering 93(4), 394-399.
Ben-Daya, M., Hassini, E., Bahroun, Z., 2019. Internet of things and supply chain management: a literature review. International Journal of Production Research 57(15-16), 4719-4742.
Bogataj, D., Bogataj, M., Hudoklin, D., 2017. Mitigating risks of perishable products in the cyber-physical systems based on the extended MRP model. International Journal of Production Economics 193, 51-62.
Chen, C., Ramaswamy, H., Alli, I., 2001. Prediction of quality changes during osmo-convective drying of blueberries using neural network models for process optimization. Drying Technology 19(3-4), 507-523.
Dada, A., Thiesse, F., 2008. Sensor applications in the supply chain: the example of quality-based issuing of perishables, The Internet of Things. Springer, pp. 140-154.
Draganić, A., Orović, I., Stanković, S., Zhang, X., Wang, X., 2017. Compressive sensing approach in the table grape cold chain logistics, 2017 6th Mediterranean Conference on Embedded Computing (MECO). IEEE, pp. 1-4.
Guo, X., Broeze, J., Groot, J.J., Axmann, H., Vollebregt, M., 2020. A worldwide hotspot analysis on food loss and waste, associated greenhouse gas emissions, and protein losses. Sustainability 12(18), 7488.
Haass, R., Dittmer, P., Veigt, M., Lütjen, M., 2015. Reducing food losses and carbon emission by using autonomous control–A simulation study of the intelligent container. International Journal of Production Economics 164, 400-408.
Heising, J.K., Claassen, G., Dekker, M., 2017. Options for reducing food waste by quality-controlled logistics using intelligent packaging along the supply chain. Food Additives & Contaminants: Part A 34(10), 1672-1680.
Hernández, I., Uarrota, V., Paredes, D., Fuentealba, C., Defilippi, B.G., Campos-Vargas, R., Meneses, C., Hertog, M., Pedreschi, R., 2021. Can metabolites at harvest be used as physiological markers for modelling the softening behaviour of Chilean "Hass" avocados destined to local and distant markets? Postharvest Biology and Technology 174, 111457.
Hribernik, K.A., Warden, T., Thoben, K.-D., Herzog, O., 2010. An internet of things for transport logistics–an approach to connecting the information and material flows in autonomous cooperating logistics processes, Proceedings of the 12th international MITIP conference on information technology & innovation processes of the enterprises. pp. 54-67.
Jedermann, R., Lang, W., 2008. The benefits of embedded intelligence–tasks and applications for ubiquitous computing in logistics, The internet of things. Springer, pp. 105-122.
Lacey, M., Lisachuk, H., Giannopoulos, A., Ogura, A., 2015. Shipping smarter: IoT opportunities in transport and logistics. Deloitte University Press.
Lammertyn, J., Aerts, M., Verlinden, B.E., Schotsmans, W., Nicolaı, B.M., 2000. Logistic regression analysis of factors influencing core breakdown in 'Conference'pears. Postharvest Biology and Technology 20(1), 25-37.
Li, M., Pullanagari, R.R., Pranamornkith, T., Yule, I.J., East, A.R., 2017. Quantitative prediction of post storage 'Hayward'kiwifruit attributes using at harvest Vis-NIR spectroscopy. Journal of Food Engineering 202, 46-55.
Li, T., Sahu, A.K., Talwalkar, A., Smith, V., 2020. Federated learning: Challenges, methods, and future directions. IEEE signal processing magazine 37(3), 50-60.
Minerva, R., Biru, A., Rotondi, D., 2015. Towards a definition of the Internet of Things (IoT). IEEE Internet Initiative, Torino, Italy.
Niknam, S., Dhillon, H.S., Reed, J.H., 2020. Federated learning for wireless communications: Motivation, opportunities, and challenges. IEEE Communications Magazine 58(6), 46-51.
Ruiz-Garcia, L., Barreiro, P., Robla, J.I., 2008. Performance of ZigBee-Based wireless sensor nodes for real-time monitoring of fruit logistics. Journal of Food Engineering 87(3), 405-415.
Scheer, F., Snels, J., Beulens, A., 2011. Business Opportunities Tracking & Tracing Tuinbouwproducten. Tuinbouw Digitaal/Wageningen UR Food & Biobased Research, 60p.
Shih, C.-W., Wang, C.-H., 2016. Integrating wireless sensor networks with statistical quality control to develop a cold chain system in food industries. Computer Standards & Interfaces 45, 62-78.
Singh, R.R.B., Ruhil, A.P., Jain, D.K., Patel, A.A., Patil, G.R., 2009. Prediction of sensory quality of UHT milk – A comparison of kinetic and neural network approaches. Journal of Food Engineering 92(2), 146-151.
Sklorz, A., Janßen, S., Lang, W., 2012. Gas Chromatograph based on Packed μGC-Columns and μ-Preconcentrator Devices for Ethylene Detection in Fruit Logistic Applications. Procedia Engineering 47, 486-489.

Smith, D., Sparks, L., 2004. Temperature controlled supply chains. Bourlakis, MA and PWH Weightman, Food Supply Chain Management, 179-198.
Soltani, M., Omid, M., Alimardani, R., 2015. Egg quality prediction using dielectric and visual properties based on artificial neural network. Food analytical methods 8(3), 710-717.
Sun, C., 2012. Application of RFID technology for logistics on internet of things. AASRI Procedia 1, 106-111.
Taoukis, P., Labuza, T., 2003. Time-temperature indicators (TTIs). Novel food packaging techniques, 103-126.
Tijskens, L.M.M., Polderdijk, J.J., 1996. A generic model for keeping quality of vegetable produce during storage and distribution. Agricultural Systems 51(4), 431-452.
Tromp, S.-O., Haijema, R., Rijgersberg, H., van der Vorst, J.G., 2016. A systematic approach to preventing chilled-food waste at the retail outlet. International Journal of Production Economics 182, 508-518.
Tromp, S.-O., Harkema, H., Hogeveen, E., Rijgersberg, H., Woltering, E.J., 2017. On the validation of improved quality-decay models of potted plants. Postharvest Biology and Technology 123, 119-127.
Tromp, S.-O., Rijgersberg, H., da Silva, F.P., Bartels, P., 2012. Retail benefits of dynamic expiry dates—Simulating opportunity losses due to product loss, discount policy and out of stock. International Journal of Production Economics 139(1), 14-21.
Van Der Vorst, J.G., Tromp, S.-O., Zee, D.-J.v.d., 2009. Simulation modelling for food supply chain redesign; integrated decision making on product quality, sustainability and logistics. International Journal of Production Research 47(23), 6611-6631.
van der Vorst, J.G., Tromp, S., van der Zee, D.-J., 2005. A simulation environment for the redesign of food supply chain networks: modeling quality controlled logistics, Proceedings of the Winter Simulation Conference, 2005. IEEE, p. 10 pp.
van der Vorst, J.G., van Kooten, O., Luning, P.A., 2011. Towards a diagnostic instrument to identify improvement opportunities for quality controlled logistics in agrifood supply chain networks. International journal on food system dynamics 2(1), 94-105.
van der Vorst, J.G., van Kooten, O., Marcelis, W., Luning, P., Beulens, A.J., 2007. Quality controlled logistics in food supply chain networks: integrated decision-making on quality and logistics to meet advanced customer demands.
Verdouw, C., Beulens, A., Trienekens, J., Van der Vorst, J., 2011. A framework for modelling business processes in demand-driven supply chains. Production Planning & Control 22(4), 365-388.
Verdouw, C.N., Beulens, A., Van Der Vorst, J., 2013. Virtualisation of floricultural supply chains: A review from an Internet of Things perspective. Computers and electronics in agriculture 99, 160-175.
Wang, J., Yue, H., Zhou, Z., 2017. An improved traceability system for food quality assurance and evaluation based on fuzzy classification and neural network. Food control 79, 363-370.
Xiao, X., Li, M., 2022. Fusion of data-driven model and mechanistic model for kiwifruit flesh firmness prediction. Computers and Electronics in Agriculture 193, 106651.
Zhang, H., Feng, H., Cui, Y., Wang, Y., 2020. A Fuzzy Bayesian Network Model for Quality Control in O2O e-Commerce. International Journal of Computers, Communications & Control 15(1).
Zhu, H., Chu, B., Fan, Y., Tao, X., Yin, W., He, Y., 2017. Hyperspectral imaging for predicting the internal quality of kiwifruits based on variable selection algorithms and chemometric models. Scientific reports 7(1), 1-13.