Multi-Task Artificial Neural Network for Post-Storage Grapes Quality Prediction

Noam Koenigstein[a,*], Yotam Givati[a,b], Abiola Owoyemi[c,d], Amnon Lichter[c], Ron Porat[c], Yael Salzer[b]

[a] *Department of Industrial Engineering, Tel Aviv University, Tel Aviv 6997801, Israel*
[b] *Institute of Agricultural Engineering, Agricultural Research Organization - Volcani Institute, Rishon LeZion, 7505101 Israel*
[c] *Department of Postharvest Science of Fresh Produce, Agricultural Research Organization - Volcani Institute, Rishon LeZion, 7505101 Israel*
[d] *Robert H. Smith Faculty of Agricultural, Food and Environmental Sciences, Hebrew University of Jerusalem, Rehovot 76100, Israel State Two Israel*

Abstract

Food loss has a major negative impact on food security, food quality, and the environment. As such, reducing food loss is an utmost important challenge for our society. To this end, a large-scale post-harvest experiment was performed and a novel neural network architecture was designed for post-harvest quality prediction. In the experiment, 924 grape clusters were stored in four cold rooms with varying temperatures and relative humidity. Grapes quality was measured prior to storage, and again after 3, 6, 9, and 12 weeks. This process yielded a post-storage fruit quality dataset, the largest of its kind. Based on this outstanding data collection effort, a novel Multi-Task Post-Storage Artificial Neural Network (MTPS-ANN) model was developed to predict multiple post-storage quality measures simultaneously. However, while the amount of data collected in this research is outstanding within the literature on post-storage quality research, it is still relatively small in the context of deep-learning models. Hence, the model's architecture was tailored specifically to address the problem of data sparsity. The model's performance was evaluated and compared to classical machine learning algorithms, where MTPS-ANN showed the best coefficient of determination score and lowest Root Mean Square Error (RMSE). Importantly, while the

*Corresponding author. Email: noamk@tauex.tau.ac.il, Telephone: +972-52-2653599.

MTPS-ANN model was developed and evaluated using a dataset of post-storage table-grape measurements, its architecture is described in general terms and can be easily adapted for many types of agricultural produce.

Keywords: Grapes, FEFO, artificial neural networks, intelligent logistics, modelling, postharvest

Funding: This research was funded by The Israel Innovation Authority, grant number 70076.

Abbreviations:
ANN - Artificial Neural Network
MTPS-ANN - Multi-Task Post-Storage Artificial Neural Network
VPD - Vapour-Pressure Deficit
RH - Relative Humidity
TSS - Total Soluble Solids
PCA - Principal Component Analysis
SVM - Support Vector Machines

1. Introduction

Food loss is a major challenge for our modern society. Research shows that roughly one-third to one-half of all food produced in the world is estimated to be lost and wasted along the different stages of the food supply chain—production, post-harvest handling and storage, processing, distribution, and consumption (Meybeck et al., 2011; Parfitt et al., 2010; Stenmarck et al., 2016; Gunders & Bloom, 2017; Porat et al., 2018). Fruits and vegetables, biodegradable organic products with a relatively short shelf life, account for about 38% of this food loss (Statista, 2018). Of all fruits and vegetables produced, 45% - 55% is lost or wasted within the food supply chain Meybeck et al. (2011).

On 2019, grapes, the fresh shape of the fruit intended for immediate consumption, raisins, and winemaking, were the fifth largest fruit in world production volume (Statista, 2020). Grapes that are intended for local distribution are stored for up to three months. Exported grapes may spend as much as 40 days from harvest day to market (Zoffoli & Latorre, 2011). Quantitative evaluations of post-harvest grapes loss i.e., the storage and distribution phases, vary according to the cold chain handling and marketing goals (Blanckenberg et al., 2021; Hertog et al., 2014). Surveys from different countries report large variations in the amount of produce loss: 21% of Pakistan's and 53% of Iran's annual grapes harvest are wasted annually due to gaps in the cold chain (Aujla et al., 2011). In high-medium income countries, where retailers aim to meet consumers' high-quality standards, grapes are more likely to be disposed of because of unsatisfactory appearance rather than being inedible (Centre for the Promotion of Imports from developing countries (CBI) Ministry of Foreign Affairs, 2021). A detailed survey of the Israeli grapes supply chain reported 7% of the grapes yield lost during production, 7% at the warehouse, 11% during marketing, and an additional 17% at consumers' homes, summing up to a total of 42% of all production (Porat et al., 2016).

One of the main causes of food loss is post-harvest logistical management and marketing decisions that do not take into account fruit shelf life (Parfitt et al., 2010). In fact, most fresh agricultural produce storage methods follow the First In, First Out (FIFO) principle: the produce release schedule from storage to markets is based on the product's warehouse entry date, irrespective of the remaining shelf life or time and distance to final destination (Jedermann et al., 2014; Hertog et al., 2014). In contrast, the First

Expired, First Out (FEFO) principal dictates produce distribution according to its shelf life expectancy. Hence, FEFO has the potential to reduce loss and improve fruit quality at consumers' homes. However, employing such intelligent logistic management policies requires estimating the produce storage and distribution potential i.e., the fruit's expected quality or remaining shelf life. This calls for the development of predictive models that estimate future produce quality based on pre-harvest quality and storage conditions (Hertog et al., 2014). Such models need to estimate the accumulated effect of storage conditions (temperature, humidity), and initial product quality on a complex of subjective and objective quality attributes, such as taste, color, appearance, and composition (Hertog et al., 2014).

A variety of machine learning models have been developed to improve different post-harvest processes and decision making such as monitoring, grading, fruits and vegetables classification, modulating microbial growth rate, modeling and predicting physical and chemical properties, and quality metrics during processing and storage (Salehi, 2020). Maftoonazad *et al.* presented an Artificial Neural Network (ANN) to predict quality indicators in avocados, manipulating duration and storage conditions (Maftoonazad et al., 2011). Sayyari *et al.* developed an Adaptive Neuro-Fuzzy Inference System integrated with an ANN to predict pomegranate fruit quality indices in varying storage duration lengths with and without methyl jasmonate—a plant hormone that protects against drought and cold stress (Sayyari et al., 2017). Li *et al.* optimised radial basis function neural network to predict table grapes' remaining shelf life (Li et al., 2019).

One key challenge in utilizing ANNs stems from their high number of learned parameters which dictates the need for more training data than classical machine learning algorithms (Goodfellow et al., 2016). Unfortunately, this requirement poses a substantial challenge to many agricultural studies. Unlike industries such as e-commerce, online entertainment, and banking, where data is collected automatically in abundance (Dua & Graff, 2017), in the agriculture industry data is seasonal and often requires a substantial manual collection effort. Even more so, research data such as in this work is collected at designed laboratory experiments which are limited in their capacity. For example, Maftoonazad *et al.* presented a model based on 509 laboratory-tested fruits (Maftoonazad et al., 2011) which is a relatively large number in the field. In comparison, the models of Sayyari *et al.* (Sayyari et al., 2017) and Li *et al.* (Li et al., 2019) were based on datasets of 225 and 271 fruits, respectively.

The primary aim of the present study was to develop a post-storage grapes quality predictive model based on an ANN that takes into account the initial product quality parameters, storage conditions, and duration. Data was sampled from 924 grape clusters and stored for 3, 6, 9, or 12 weeks, under different storage conditions. The grapes were tested for a variety of objective and subjective quality measures before and after the storage experiment. Although the data-set of 924 samples doubles similar experimental studies (Maftoonazad et al., 2011; Sayyari et al., 2017; Li et al., 2019), it is still considered relatively small in the context of deep-learning models (Hu et al., 2021). To this end, a simple yet-highly effective layered *Multi-Task Post-Storage Artificial Neural Network (MTPS-ANN)* was proposed. First, the raw data was fed into an encoder that performs supervised dimension reduction which cancels correlations and sheds away irrelevant information with respect to the prediction task. Then, the lower dimensional representation was fed into a multi-objective prediction head capable of predicting several post-storage quality measures. The multiple labels employed by the multi-task prediction head serve as additional supervision to assist the learning process. MTPS-ANN model performance was then compared with classical machine learning algorithms. Finally, an ablation study was presented in which the architectural choices of the model are analyzed and the contribution of each component is demonstrated.

2. Materials and Methods

2.1. Grapes Preparation and Data Collection

Grapes are best stored at -0.5 °C to 0 °C, keeping relative humidity of 95% (Crisosto & Crisosto, 2020; Yahia, 2011). Typically, grapes are packed in liner bags and boxed with slow release SO_2 sheets to slow berries decay and delay rachis browning (Lichter et al., 2016; Nelson, 1979; Zutahy et al., 2008). Grapes quality decline is manifested in excessive weight loss, berry shatter, cracked berries, berry abscission, flesh browning or splitting, rachis browning—caused by dehydration, and berry bleaching—caused by SO_2 excess (Yahia, 2011; Luvisi et al., 1992; Zutahy et al., 2008; Perera et al., 2018; Lichter et al., 2011; Lichter, 2016; Crisosto et al., 2001). Other post-harvest deterioration agents are fungal disease and pathogens, such as the Gray Mold, Blue Mold and Alternaria (Lichter & Romanazzi, 2017; Yahia, 2011; Zoffoli et al., 2009).

The current research was carried out on the *Vitis vinifera* cv. Scarlotta Seedless® (SunWorld, CA, USA) table grapes. 'Scarlotta' is a seedless variety characterised by late ripening with large red elliptical berries Admane et al. (2015). The vineyards were located in the Lachish area (lat. 31°33', long. 34°51'), about 250 m above sea level. The climate in the area is considered Mediterranean ranging from 9°C in February to 39°C in August, and annual rainfall of about 400 mm. In October 2020, the fruit was harvested from ten different vineyards. Vine spacing was 3.5 m between rows and 2.0 m within rows. Vines were cane-pruned and trained to the Y trellis system, drip-irrigated, and the top of each row was covered with plastic film or shade net to protect the canopy and clusters from the negative effects of excess radiation, heat, rain, hail, and wind.

On harvest day, a total of 924 clusters were sampled from the vineyard for the experiment. Unhealthy berries were removed from the vineyard prior to placing the produce in cardboard boxes and transporting it, in an air-conditioned van, to the laboratory at a distance of 45 km. The grapes were prepared, examined and stored in the Department of Post-harvest Science, Agriculture Research Organization - Volcani Institute. In the laboratory, clusters were further inspected, trimmed to mean weight of 668 $\pm$ 81 g, and inserted into a cluster bag. Seven cluster bags were placed in a liner bag which was then put into a cardboard box. The liner had sixty 4-mm diameter holes and the cardboard box (40 cm x 30 cm x 13 cm) was open at the top with ventilation holes to allow rapid cooling. An SO_2 pad (Uvasys Dual Release Green, Tessara, South Africa), covered with a paper sheet (37.5 cm x 25 cm, 1.7 g), was placed on top of the cluster bags to absorb excess humidity. The liner bag was closed by folding it into one side of the cardboard box.

After the above packaging process, the grape clusters were placed in different cold rooms for varying storage times. Temperature and humidity were continuously logged during storage time. Once a cluster was pulled out of storage, it was associated with the average and standard deviation of humidity, temperature, and vapor-pressure deficit (VPD) accumulated over storage time. The grape clusters were randomly assigned to one of four cold storage rooms with different storage conditions as follows: (1) 112 clusters were stored in a room with a mean temperature of 3.57 $\pm$ 0.8 °C, relative humidity (RH) of 88.1% $\pm$ 5%, and VPD of 0.09 $\pm$ 0.04. (2) 112 clusters were stored in a room with a mean temperature of 2.8 $\pm$ 0.4 °C, RH of 69.5% $\pm$ 9.5%, and VPD of 0.23 $\pm$ 0.07. (3) 112 clusters were stored in a room with a mean temperature of 5.7 $\pm$ 0.3 °C, RH of 93.5% $\pm$ 1.3%, and VPD of 0.06 $\pm$ 0.01.

(4) Finally, 588 clusters were stored in a room with a mean temperature of 0.7 ± 0.8 °C, RH of 86.1% ± 3.5%, and VPD 0.09 ± 0.032. Clusters that were assigned to rooms (1),(2), and (3), were evenly removed after 3, 6, 9 and 12 weeks in order to test their post-storage quality at varying storage lengths. Each time 28 clusters were taken out of each of the 3 rooms and placed for 3 more days at room temperature of 21.5 °C and 65% RH in order to check their post-storage shelf-life. The 588 clusters that were assigned to a room (4) were also part of the experiment described in (Owoyemi et al., 2022). Their removal from storage was performed according to the following schedule: 84 clusters were removed after 3 weeks, another 84 clusters were removed after 6 weeks, 336 clusters were removed after 9 weeks, and the final 84 clusters were removed after 12 weeks. After removal, the clusters from room (4) were also placed at room temperature of 21.5 °C and 65% RH for 3 more days in order to test their post-storage shelf-life. Altogether, each of the 924 clusters above was tested for quality pre- and post-storage in order to create the largest dataset of its kind.

2.2. Pre-Storage Quality Analysis

Prior to storage, the following analysis was carried out to evaluate the initial fruit quality: First, each cluster was weighted. Then, two berries were sampled from each cluster for berry weight and total soluble solids (TSS) content, measured with a digital refractometer (Atago, Japan). The number of berries was estimated by dividing the cluster weight by an average weight of a single berry. The firmness of one sampled berry was measured using a texture analyzer (Stable Micro Systems, UK) with a 3 mm probe, 5% strain and speed of 2 mm $^{-1}$. Four firmness features were extracted: force— a berry's resistance to probe penetration, area— the sum of force over time until penetration, distance— the distance of a 5% penetration probe, and strain height— the berry diameter. The cluster's emission of fluorescent light was collected by a Multiplex 3 proximal sensor (ForceA, France) Ghozlen et al. (2010). Some of the signals of the system listed in Table 1 were correlated to anthocyanin, flavonoid, and chlorophyll composition of grapes and other plants (Bahar & Lichter, 2018; Bahar et al., 2017; Lichter et al., 2014). Finally, two experts evaluated rachis browning level, i.e., *rachis score*, and assessed the cluster's suitability for marketing, i.e., *acceptance score*. Rachis browning score was rated on a scale of 0 to 1, with 0 denoting rachis and pedicels completely green, 1 denoting completely brown, and intermediate values representing the percent coverage of brown coloration (Bahar et al., 2017). Acceptance score was rated on a scale of 5 to 1, with 5 denoting the produce being very suitable for marketing and 1 denoting a product unlikely to be purchased. The acceptance scores represent the opinions of domain experts from academia who aggregated the cluster's quality and quantified it into a single representative index. For example, a mild rot of up to 2 decayed berries in a cluster, will be scored 4. If the mild rot is accompanied by bleaching or cracks, the cluster's score will be lower. Quality evaluation parameters and storage condition features are summarised in Table 1.

Feature/Label	**Units**	**Description**	**Model input (features)**	**Model output (labels)**
Weight	grams	Cluster weight	✓	
Firmness distance	millimeter (mm)	Distance probe penetrates 5% of berry's diameter	✓	✓
Firmness force	N	Berry's resistance to probe penetration	✓	✓
Firmness area	N*second	Sum of force over time until probe penetration	✓	✓
Firmness strain height	millimeter (mm)	Berry diameter	✓	✓
Total soluble solids (TSS)	Brix	Sugar level	✓	
Temperature (mean)	°C	Mean storage temperature	✓	
Temperature (STD)	°C	Standard deviation of storage temperature	✓	
Relative humidity (mean)	%	Mean storage relative humidity	✓	
Relative humidity (STD)	%	Standard deviation of storage relative humidity	✓	
Vapour-pressure deficit (VPD, mean)	kilo pascal (kPa)	Mean difference between the amount of moisture in the storage room air and how much moisture the air can hold when it is saturated	✓	
Vapour-pressure deficit (STD)	kilo pascal (kPa)	Standard deviation VPD	✓	
Storage time	weeks	Storage duration	✓	
Acceptance score	Value range 1 to 5	The marketing suitability level: 5 very suitable for marketing, 1 unlikely to be purchased	✓	✓
Rachis score	Value range 0 to 1	Rachis browning level: 0 rachis and pedicels completely green, 1 rachis and pedicels completely brown	✓	✓
Bleaching	value range 1 to 5	1 - no bleaching; 2 - 2 to 5 berries; 3 - 6 to 10 berries; 4 - 11 to 20 berries; and 5 - over 20 berries with bleaching symptoms		✓
Cracks	%	The ratio of berries showing cracks to total berries count		✓
Shatter	%	The ratio of detached berries to total berries count		✓

Weight loss	%	The ratio of cluster weight change pre to post storage over pre storage weight		✓
Decay	%	The ratio of berries with Botrytis symptoms to total berries count		✓
RF_R	mV	Red fluorescence excited by red led	✓	
FRF_G	mV	Far red fluorescence excited by green led	✓	
FRF_UV	mV	Far red fluorescence excited by UV led	✓	
SFR_G		Simple fluorescence red: the ratio of FRF_G to RF_G. Index of chlorophyll	✓	
RF_UV	mV	Red fluorescence excited by UV led	✓	
RF_G	mV	Red fluorescence excited by green led	✓	
YF_G	mV	Yellow fluorescence excited by green led	✓	
SFR_R		Simple fluoresence red: The Chlorophyll index: the ratio of FRF_R to RF_R	✓	
RF_B	mV	Red fluorescence excited by blue led	✓	
ANTH_RB		Log Anthocyanins index: the ratio of FRF_R to FRF_B	✓	
FRF_R	mV	Far Red fluorescence excited by red light	✓	
ANTH_RG		Log Anthocyanin index: the ratio of FRF_R to FRF_G	✓	
FLAV		Log (FRF_R / FRF_UV): Index of compounds which absorbs at 375 nm (flavonoids)	✓	
FRF_B	mV	far Red fluorescence excited by blue led	✓	
YF_B	mV	Yellow fluorescence excited by blue led	✓	
YF_UV	mV	Yellow fluorescence excited by UV led	✓	
FER_RG	Ratio	Anthocyanin index: ratio of FRF_R to FRF_G	✓	
FERARI	Log(1/mV)	Anthocyanin index: Log (5000 / FRF_R)	✓	
YF_R	mV	Yellow fluoresence excited by red led	✓	

Table 1: Pre- and post-storage quality measurements, storage conditions, and storage duration, which constitute the novel dataset collected in this research. Measurements that are part of the model's inputs (features) were measured prior to storage, and model's outputs (labels) were measured post-storage and after an additional 3 days of "shelf-life". Note that some parameters were measured twice i.e., Rachis Score, Acceptance Score, and the four Firmness metrics were measured once when the produce was put into storage (i.e., to be used as input) and once again when the produce was taken out of storage (i.e., to be used as output). Overall 924 clusters participated in the storage experiment making this novel dataset the largest of its kind.

2.3. Post-storage quality analysis

After removal from the cold storage rooms, the produce was kept for 3 more days of "shelf life" after which quality was measured and assessed as follows: Each cluster was examined for SO_2 damage (i.e., bleaching), cracks, shatter, firmness, weight loss, and decay. SO_2 damage was scored from 1 to 5 where 1 denote no bleaching, 2 denote 2 to 5 berries with bleaching symptoms, 3 denote 6 to 10 berries with bleaching symptoms, 4 denote 11 to 20 berries with bleaching symptoms, and 5 denote over 20 berries with bleaching symptoms. Cracks visible to the naked eye were scored according to the same scale as the bleaching scale above. Weight loss was calculated as the ratio of the cluster weight change pre- to post-storage over the original pre-storage weight. Shatter was calculated as the ratio of detached berries to total berries count. Decay was evaluated as the ratio of berries with Botrytis symptoms to total berries count. The 4 firmness features were measured in a similar fashion to the pre-storage measurements described in Section 2.2. Finally, the domain experts evaluated the clusters' post-storage rachis and acceptance scores in a similar fashion to the pre-storage measurements described in Section 2.2.

2.4. Data-set Preparation

The pre- and post-storage analyses produced 924 labeled training data points. Each datum consists of 34 input features— these are the pre-storage quality measures described in Section 2.2 plus the storage cold-room conditions and duration, all detailed in Table 1. The 11 output labels are the post-storage quality parameters described in Section 2.3 and detailed in Table 1 as well. All the features and labels underwent a min-max scaling process (Larose & Larose, 2014). Finally, a $80\% - 20\%$ train-set test-set split was performed in order to enable model evaluation and comparison.

3. Model and Methodology

The primary aim of the present study is to develop a post-storage grapes quality predictive model. To this end, the Multi-Task Post-Storage Artificial Neural Network (MTPS-ANN) was developed as described in this section. At a glance, the MTPS-ANN model takes a feature vector representing the pre-storage produce quality, storage conditions, and storage duration, in order to output a prediction vector representing multiple post-storage quality measures. The exact inputs and outputs for the model are those described in Section 2 and summarised in Table 1.

The MTPS-ANN model employs state-of-the-art deep-learning in order to maximize prediction accuracy. However, while the dataset size collected in this research is outstanding within post-storage experimental studies, its size is still limited in the context of deep-learning techniques which are highly prone to overfitting (Ying, 2019). To this end, the model's architecture was specifically designed to tackle the problem posed by data sparsity in experimental research in agriculture. Importantly, the design and architecture of the MTPS-ANN model are general, highly flexible, and can be applied in the future to different inputs or outputs and even to other types of agricultural produce.

In what follows we describe the MTPS-ANN model in detail. We begin with a general discussion of two design principals of MTPS-ANN which help mitigate data sparsity: (1) its *multi-task* property is discussed in Section 3.1, and (2) its *encoder* that performs supervised dimension reduction is discussed in Section 3.2. Finally, the formal model description is detailed in Section 3.3.

3.1. Multi-task Prediction

Most classical machine learning models perform supervised prediction on a single prediction task. While it is possible to extend classical models to perform multiple tasks, such extensions are difficult and less common (Zhang & Yang, 2017). Usually, when multiple predictions are required, separate models are trained, one for each prediction task. In contrast, deep learning models are naturally capable of performing multiple predictions simultaneously (Crawshaw, 2020). As such, the MTPS-ANN model in this study incorporates an optimized multi-task prediction head in order to predict multiple post-storage quality measures simultaneously as detailed in Table 1.

With multi-task models, it is often the case that not all prediction tasks are equally important. Within the scope of the present study, *acceptance score*, *rachis score*, and *decay* were the main objectives for the prediction model, while the other labels—bleaching, cracks, shatter, weight loss, and firmness features (area, force, strain height, and distance)—were auxiliary outputs. To this end, the MTPS-ANN model incorporates a weight vector that specifies the relative importance of each prediction task. Tuning the weight vector is achieved by a 5-fold cross-validation (Arlot & Celisse, 2010) in order to find an optimal setting in which all the main objectives of the model, *acceptance score*, *rachis score*, and *decay*, are predicted with high accuracy.

By predicting multiple quality measures simultaneously, MTPS-ANN gives a more detailed projection of post-storage produce quality. However, the multi-task property of the MTPS-ANN model serves an additional important goal: As explained in Section 1, the main challenge in employing deep-learning techniques for agricultural forecasting is coping with limited-size datasets. The inclusion of multiple prediction labels provides additional mitigation to over-fitting. Multiple objectives give additional supervision and guidance to the model and are known to serve a similar role as regularization in order to improve the model's ability to generalize and avoid overfitting (Goodfellow et al., 2016). In fact, it was been shown that for a multi-task model with k different tasks, the amount of data required to learn the shared parameters is approximately k times less than in the case of a single task Baxter (1997). Therefore, the multi-task prediction head serves two goals simultaneously; by using a single model it enables forecasting several post-storage quality measures simultaneously, and it helps mitigate the data sparsity problem, typical of experimental research in agriculture.

3.2. Supervised Dimensionality Reduction

Correlations in the input features indicate the need to perform dimensionality reduction, especially when the dataset size is limited with respect to the number of parameters, as in this work (Fodor, 2002). Arguably, the most well-known dimensionality reduction technique is Principal Component Analysis (PCA) (Jackson, 2005). In PCA, the eigenvectors of the correlation matrix are used to find an uncorrelated linear projection for dimensionality reduction. It can be shown that PCA is the optimal linear dimension reduction technique in terms of squared error minimization (Jackson, 2005). However, PCA suffers from several shortcomings. First, PCA's optimality holds only for the linear case, so non-linear techniques often provide superior results. Moreover, PCA's optimality holds only for the squared loss reconstruction error but not for any other criterion. Finally, and most importantly, PCA is an unsupervised technique with the objective of preserving the original signal. As such, PCA considers all input features equally important regardless of their relevance to the ultimate goal of the model. However, in most cases, not all signals are relevant to the prediction task, and ideally should be removed rather than preserved.

To overcome the first and second shortcomings, one may perform unsupervised dimensionality reduction using modern non-linear techniques such as auto-encoders (Bank et al., 2020). Auto-encoders are artificial neural networks that learn to perform non-linear dimensionality reduction with respect to any user-defined differentiable reconstruction loss. As such, auto-encoders can outperform PCA with respect to a large array of criteria. However, similar to PCA, auto-encoders are unsupervised and cannot discern between signals that are relevant (informative) to the prediction task, and those that are irrelevant (non-informative) with respect to the prediction task.

To overcome the aforementioned shortcomings, the MTPS-ANN model employs an encoder that performs non-linear *supervised* dimension reduction with respect to the forecasting problem. To this end, MTPS-ANN's encoder is optimized together with the model's prediction head via an end-to-end optimization procedure. Hence, the encoder learns to discern the informative components of the input signal with respect to the forecasting problem at hand.

3.3. A Formal Model Description

The model is described in general mathematical terms in order to make the description flexible and to generalize the set-up for future implementations using other features and different types of agricultural produce. Later, in Section 3.3.5, a description of the specific implementation details of the instance of MTPS-ANN used in this research is provided.

3.3.1. Mathematical Notation

In what follows, matrices and vectors are distinguished from scalars by using **bold** letters for the former. Additionally, matrices are distinguished from vectors by using capitalized letters for the former and minuscule letters for the latter. For example, $\boldsymbol{X}$ is a matrix, $\boldsymbol{x}$ is a vector, x and X are scalars.

3.3.2. The Encoder

In general, our *training* dataset consists of N tuples of feature (input) vectors and label (output) vectors. and We denote the i'th feature and label vectors by $\boldsymbol{x_i}$ and $\boldsymbol{y_i}$, respectively, where $\boldsymbol{x_i} \in \mathbb{R}^{d_{in}}$ and $\boldsymbol{y_i} \in \mathbb{R}^{d_{out}}$, and d_{in} and d_{out} are the number of features and labels, respectively. Therefore, we collectively denote the dataset of N tuples of feature vectors $\boldsymbol{x_i}$ and label vectors $\boldsymbol{y_i}$ as $\{\boldsymbol{x_i}, \boldsymbol{y_i}\}_{i=1}^{N} \in \mathcal{D}$.

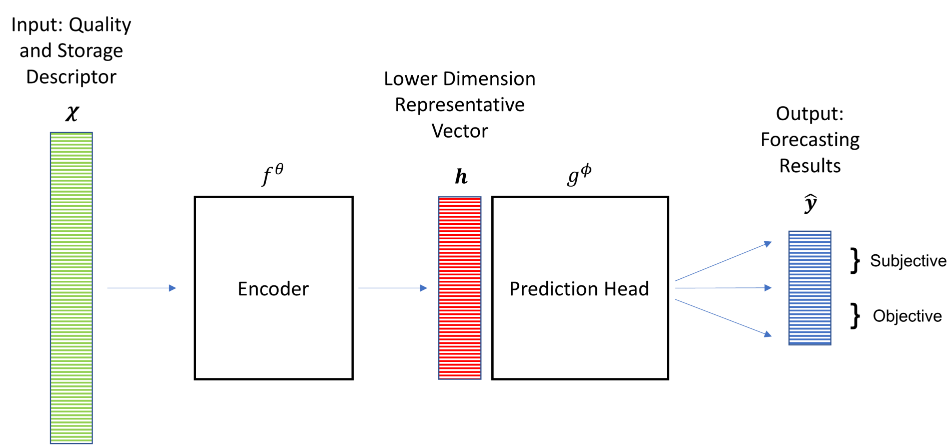

Figure 1: The Multi-Task Post-Storage Artificial Neural Network (MTPS-ANN) architecture. The produce quality and storage descriptors—input features $\boldsymbol{x}$—are fed into the encoder $f^{\boldsymbol{\theta}}$ to produce the low-dimensional representation $\boldsymbol{h}$. Then, the representation $\boldsymbol{h}$ is fed into the multi-task prediction head $g^{\boldsymbol{\phi}}$ to output the forecasting results vector $\hat{\boldsymbol{y}}$.

Let $f^{\boldsymbol{\theta}} : \mathbb{R}^{d_{in}} \to \mathbb{R}^{d_h}$ be the encoder function parameterized by $\boldsymbol{\theta}$. The encoder's role is to perform dimensionality reduction by mapping the input feature vector $\boldsymbol{x_i}$ into a lower dimension representation $\boldsymbol{h_i} \in \mathbb{R}^{d_h}$, where d_h is the lower dimension *i.e.* $d_h < d_{in}$. The input feature vector $\boldsymbol{x_i}$ is fed into the dimension reduction function $f^{\boldsymbol{\theta}}$ to produce the latent representation $\boldsymbol{h_i}$, formally: $\boldsymbol{h_i} = f^{\boldsymbol{\theta}}(\boldsymbol{x_i})$. The encoder's parameters $\boldsymbol{\theta}$ are optimized in the training phase with respect to the prediction task. Therefore, by optimizing $\boldsymbol{\theta}$, the encoder "learns" a dimensionality reduction transformation that serves the final prediction task. This can only be achieved by preserving the relevant information with respect to the prediction task. Hence, the encoder must remove redundant information in order to achieve a low dimensional representation while preserving and improving the prediction accuracy. In Section 4.3 we showcase the contributions of the encoder and further investigate its properties.

3.3.3. The Prediction Head

As explained above, the MTPS-ANN model is a multi-task model. Hence, the model's outputs consist of d_{out} different predictions collectively known as the prediction vector. For a single datum $(\boldsymbol{x_i}, \boldsymbol{y_i})$ the model predicts a single prediction vector $\boldsymbol{\hat{y}_i}$ consisting of d_{out} different predictions. We denote by $\hat{y}_{ij}$, the j'th prediction in $\boldsymbol{\hat{y}_i}$, namely $\hat{y}_{ij} \triangleq \boldsymbol{\hat{y}_i}(j)$. Similarly, we denote by y_{ij} the j'th value in label vector $\boldsymbol{y_i}$, namely $y_{ij} \triangleq \boldsymbol{y_i}(j)$.

Let $g^{\boldsymbol{\phi}} : \mathbb{R}^{d_h} \to \mathbb{R}^{d_{out}}$ be the multi-task prediction head parameterized by $\boldsymbol{\phi}$. The prediction head $g^{\boldsymbol{\phi}}$ takes the representation $\boldsymbol{h_i}$ and predicts the vector (output) $\boldsymbol{\hat{y}_i}$, formally: $\boldsymbol{\hat{y}_i} = g^{\boldsymbol{\phi}}(\boldsymbol{h_i})$. Hence, the full model can be described as the composition: $\boldsymbol{\hat{y}_i} = g^{\boldsymbol{\phi}}(f^{\boldsymbol{\theta}}(\boldsymbol{x_i}))$. Figure 1 presents the model's architecture as described above.

3.3.4. MTPS-ANN Optimization

Our goal is to minimize the difference between the prediction $\boldsymbol{\hat{y}_i}$ and the actual label vector $\boldsymbol{y_i}$, with respect to some criteria on the magnitude of the error known as the *loss* function. Hence, we define a loss term for each of the d_{out} different predictions as follows: Let, $\boldsymbol{\mathcal{L}}^j : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be the loss function for the j'th prediction (output). We denote the loss for the j'th prediction of the i'th datum by $l_i^j \triangleq \boldsymbol{\mathcal{L}}^j(y_{ij}, \hat{y}_{ij})$. Let $\boldsymbol{w} \in \mathbb{R}_{>0}^{d_{out}}$ be the non-negative importance weight vector for the different prediction tasks. We denote by w_j the j'th importance weight *i.e.* $w_j \triangleq \boldsymbol{w}(j)$. Let $\boldsymbol{\mathcal{L}} : \mathbb{R}^{d_{out}} \times \mathbb{R}^{d_{out}} \to \mathbb{R}$ be the model's overall loss function. We denote by l_i The model's overall loss term for the i'th datum which is given by $l_i \triangleq \mathcal{L}(\boldsymbol{y_i}, \hat{\boldsymbol{y}}_{\boldsymbol{i}}) = \sum_{j=1}^{d_{out}} w_j \cdot l_i^j$. Namely, the overall loss term is a weighted sum of the individual loss terms of each of the prediction tasks.

As explained above, the dimension reduction function $f^{\boldsymbol{\theta}}$ and the multi-task prediction function $g^{\boldsymbol{\phi}}$ are parameterized by $\boldsymbol{\theta}$ and $\boldsymbol{\phi}$, respectively. These parameters are jointly optimized (learned) in the training process. Given the training data-set $\mathcal{D}$, the following optimization objective is minimized:

$$(\boldsymbol{\theta}^*, \boldsymbol{\phi}^*) = \min_{\boldsymbol{\theta},\boldsymbol{\phi}} \sum_{(\boldsymbol{x_i},\boldsymbol{y_i})\in\mathcal{D}} \mathcal{L}\Big(\boldsymbol{y_i}, g^{\boldsymbol{\phi}}(f^{\boldsymbol{\theta}}(\boldsymbol{x_i}))\Big). \tag{1}$$

Optimization is achieved via stochastic gradient descent in a process known as *training* (Le et al., 2011). Once the training is concluded, the model's parameters $\boldsymbol{\theta}^*$ and $\boldsymbol{\phi}^*$ are fixed and the model is ready for evaluation and later real-world predictions.

3.3.5. Specific Implementation Details

In the general case, $f^{\boldsymbol{\theta}}$ and $g^{\boldsymbol{\phi}}$ can take many different forms depending on the input data, the prediction goals, and the size of the training data-set. In this research, we performed a 5-fold cross-validation process (Arlot & Celisse, 2010) in order to set the optimal architecture and hyper-parameters. Given the data-set described in Section 2 and the problem at hand, it was found that the optimal architecture for $f^{\boldsymbol{\theta}}$ consists of a single fully connected dropout layer (Srivastava et al., 2014) with 102 neurons followed by a ReLU (Schmidt-Hieber, 2020) activation function. The dropout probability was kept to its default value of $p = 0.5$. We note that we have experimented with different values of p, but did not notice a significant improvement. The optimal dimensionality of the representation $\boldsymbol{h_i}$ was found to be $d_h = 18$. The optimal architecture found for $g^{\boldsymbol{\phi}}$ consists of 18 neurons with a linear activation function. Note that this set-up is relatively "frugal" in terms of parameters, which seems to feet the size of the data-set. Finally, we performed a grid search (Liashchynskyi & Liashchynskyi, 2019) in order to optimize the values of the importance weights vector $\boldsymbol{w}$ (from Section 3.3.4).

4. Results and Discussion

The results section begins in Section 4.1 which provides an analysis of the experimental data collected in the study which shed some light on the post-storage prediction task. Specific emphasis is placed on exposing correlations between the input features which motivates the supervised dimension reduction, and between the input features and the prediction outputs. In Section 4.2 the model's prediction accuracy is analyzed and compared against alternative baselines. Finally, Section 4.3 provides an in-depth discussion of the contribution of MTPS-ANN's supervised dimension reduction component in comparison with alternative techniques.

4.1. Pre- and Post-Storage Quality Measures

First, we describe some statistics on the novel dataset collected in this research. The the pre-storage and post-storage quality measures statistics are summarised in Table 2 and Table 3, respectively. These tables provide statistics on the range of the values which were measured, their, mean, median, and standard deviation (STD).

Figure 2 presents the correlations between the pre-storage features which make up the input to the model. The figure reveals many groups of correlated features: As expected, the multiplex signals, which are indicative of the level of specific chemicals in the berries, were found to be highly correlated among each other. For example, the group of 12 raw signals (e.g., RF_R, FRF_G, FRF_UV, etc.) are all positively correlated. However, this group of features is negatively correlated to the group of seven computed log or ratio signals, such as FERARI, FER_RG, etc. Two signals associated with chlorophyll, SFR_G and SFR_R, showed high correlations to each other, but low correlation with all other signals. As expected, the two firmness measures, area, and strain height are highly correlated, and relative humidity is negatively correlated with VPD. The above correlations indicate a high degree of information redundancy in the input features, which motivates the supervised dimension reduction of the MTPS-ANN encoder described earlier.

The correlations between the input features, pre-storage quality measures, storage conditions, and duration, and the post-storage quality measures are presented in Figure 3. Pre-storage firmness features were found to moderately correlate with the post-storage firmness measurements. As could be expected, weight loss has a negative correlation with relative humidity and a positive correlation with VPD– the higher the humidity the lower the weight loss. Unsurprisingly, rachis score and acceptance score are both highly correlated with storage time– longer storage time is associated with lower produce quality i.e., lower acceptance score. Longer storage time is also associated with browning of the rachis, i.e., higher rachis browning score. Bleaching was

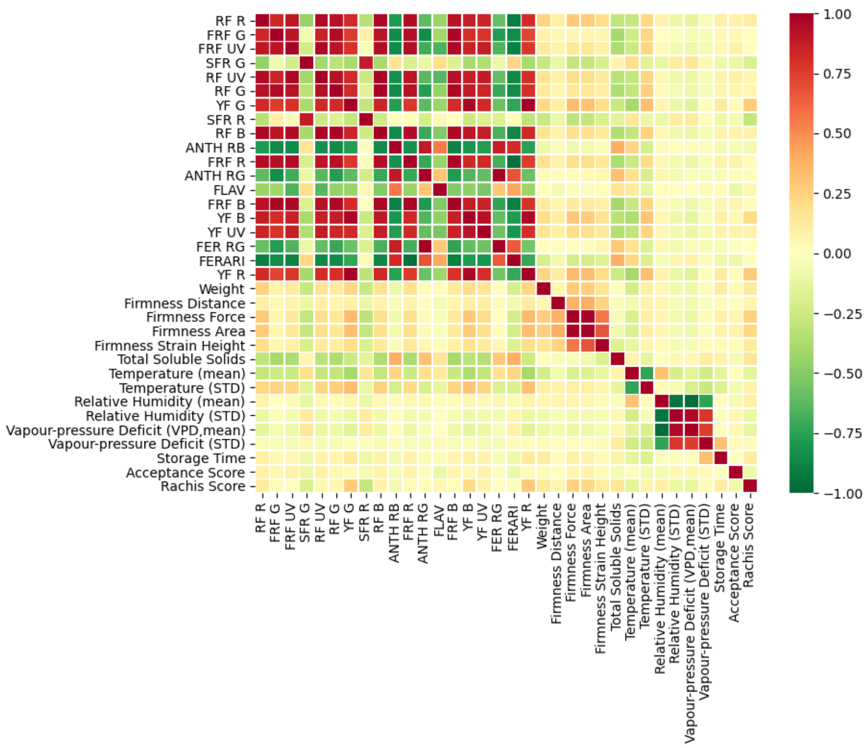

Figure 2: Pearson correlation matrix of the grapes pre-storage quality measures and storage conditions.

found to be mildly correlated with color signals. High bleaching damages the structure of the red anthocyanin and green chlorophyll molecules. In general, it may be noted, that the pre-storage measurements and storage conditions are weakly correlated with the post-storage forecasting results. Hence, it is plausible to assume that a non-linear predictive model, such as an ANN, is more suitable in this case, and will outperform linear-based models, e.g., linear regression with multiple variables.

4.2. Prediction Accuracy

The MTPS-ANN model was evaluated with respect to its prediction accuracy on of the following outputs: (1) *acceptance score*, (2) *rachis score*, and (3) the percentage of *decay*. In Table 4 the model's accuracy is compared to multiple baselines: (1) a simple linear regression model based on the storage duration only (*Storage duration*), (2) a linear regression model

	Mean	Median	STD	Minimum	Maximum
Weight (cluster weight)	668.08	686.89	81.14	305.77	840.54
Firmness distance	1.02	1.03	0.24	-3.52	1.32
Firmness force	247.66	235.49	81.18	3.98	636.14
Firmness area	70.45	65.83	27.44	-0.20	196.64
Firmness strain height	20.87	20.75	1.73	16.09	26.53
Total soluble solids (TSS)	18.23	18.30	1.59	13.10	22.60
Acceptance score	4.98	5.00	0.10	4.50	5.00
Rachis score	0.02	0.00	0.06	0.00	0.30
RF_R	371.27	341.10	161.47	50.19	1107.00
FRF_G	78.02	61.27	55.71	8.98	388.40
FRF_UV	567.48	487.25	299.99	72.30	1815.00
SFR_G	0.61	0.59	0.13	0.31	1.31
RF_UV	438.42	365.80	248.40	37.62	1602.00
RF_G	48.64	36.50	36.39	3.29	252.00
YF_G	142.94	121.55	61.86	52.77	349.60
SFR_R	0.57	0.57	0.07	0.33	0.91
RF_B	346.48	277.50	220.98	27.21	1460.00
ANTH_RB	0.08	0.08	0.05	-0.05	0.22
FRF_R	574.09	528.60	235.76	106.90	1583.00
ANTH_RG	0.49	0.49	0.13	0.13	0.82
FLAV	0.25	0.24	0.08	0.05	0.58
FRF_B	454.46	372.10	277.10	54.03	1707.00
YF_B	5.57	4.83	2.36	2.07	14.53
YF_UV	11.45	10.57	4.15	3.80	27.59
FER_RG	3.20	3.12	0.91	1.36	6.67
FERARI	0.45	0.45	0.18	-0.03	1.14
YF_R	186.66	157.80	85.60	54.55	477.10

Table 2: Statistics of the pre-storage quality features which make up the MTPS-ANN input vector $\boldsymbol{x}$. A detailed description of the above features is provided in Table 1.

	mean	50%	std	min	max
Acceptance score	2.68	3.00	1.06	1.00	4.50
Rachis score	0.58	0.60	0.27	0.10	1.00
Decay	3.82	1.15	7.50	0.00	80.60
Bleaching	1.27	1.00	0.51	1.00	4.00
Cracks	1.09	0.00	2.82	0.00	24.00
Shatter	0.75	0.00	1.40	0.00	14.00
Firmness distance	1.03	1.03	0.09	0.07	1.31
Firmness force	215.03	208.84	77.66	19.64	614.01
Firmness area	60.04	56.80	25.05	0.45	215.68
Firmness strain height	20.88	20.83	1.65	15.50	26.42
Weight loss	2.78	2.13	2.09	0.17	13.68

Table 3: Statistics of the post-storage quality labels for MTPS-ANN's output vector $\boldsymbol{y}$. We denote the labels which constitute the main objectives of the model in bold. A detailed description of all the above labels is provided in Table 1.

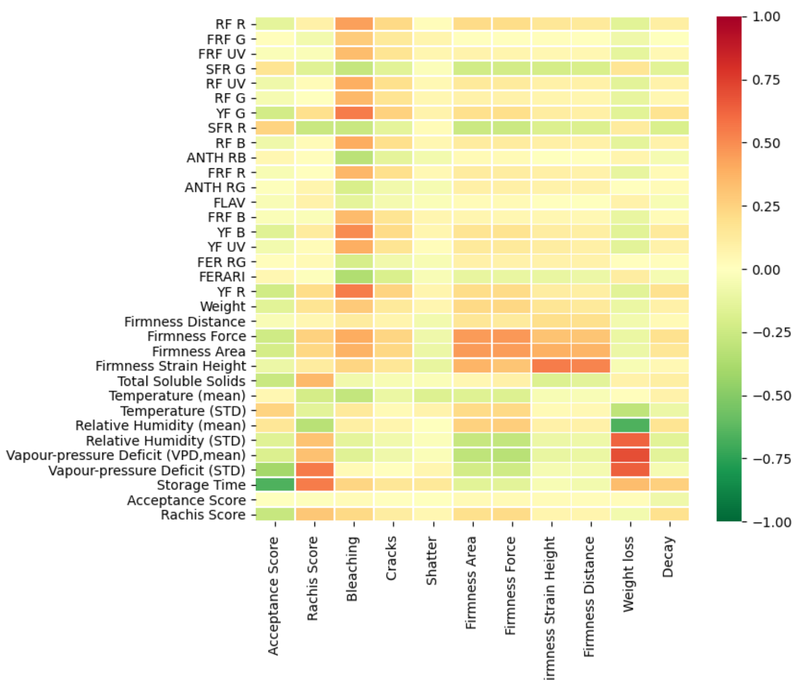

Figure 3: Pearson correlation matrix between the grapes pre-storage quality measures and storage conditions (y-axis) and the post-storage quality labels (x-axis).

Algorithm	**Acceptance Score**	**Rachis Score**	**Decay**
Storage duration	0.796	0.223	7.248
Linear Regression	0.182	0.174	0.086
SVM	0.181	0.167	0.093
Gradient Boosting	0.178	0.163	0.088
Random Forests	0.175	0.165	0.088
ANN Single-objective	0.181	0.161	0.083
MTPS-ANN	**0.161**	**0.148**	**0.070**

Table 4: The Root Mean Squared Error (RMSE) of the proposed MTPS-ANN model and multiple alternative baselines: (1) a simple linear regression model based on the storage duration only (*Storage duration*), (2) a linear regression model based on all the input features (*Linear Regression*), (3) Support Vector Machines (*SVM*), (4) *Gradient Boosting*, (5) *Random Forests*, (6) An ablation model of an artificial neural trained on a single objective (*ANN Single-objective*), and finally (7) the proposed model (*MTPS-ANN*).

based on all the input features (*Linear Regression*), (3) Support Vector Machines (*SVM*) (Smola & Schölkopf, 2004), (4) *Gradient Boosting* (Friedman, 2002), and (5) *Random Forests* (Biau & Scornet, 2016). As can be seen, the MTPS-ANN model outperforms all baselines on each of the prediction tasks. Specifically, noted the low accuracy of the model based on the storage duration only. This emphasises the importance of the pre-storage quality features and storage conditions to the overall prediction accuracy.

A distinct property of the MTPS-ANN model is the fact that it is a multi-task model that learns to predict all three outputs simultaneously; namely, a single trained instance of the MTPS-ANN model was used to predict all 3 outputs. This is in contrast to the above baselines which require different instances to be trained for each of the prediction tasks; namely, for each baseline, three different models were trained to predict either acceptance score, rachis score, or decay, separately. Therefore, the multi-task property of the model has a significant advantage in terms of scalability and ease of use. However, as explained in Section 3.1, the multi-task property of MTPS-ANN also serves to mitigate data sparsity and improves overall accuracy. In order to showcase the contribution of the model's multi-task property to the overall accuracy, an ablated version of MTPS-ANN is included in which 3 different models were trained in order to predict each of the 3 objectives separately. This ablated version is presented Table 4 and dubbed *ANN Single-objective*. As can be seen, without the additional regularization provided by the multi-

Algorithm	Acceptance Score	Rachis Score	Decay
Storage duration	0.433	0.284	0.06
Linear Regression	0.891	0.897	0.935
SVM	0.892	0.906	0.923
Gradient Boosting	0.896	0.911	0.931
Random Forest	0.900	0.908	0.931
ANN Single-objective	0.892	0.912	0.939
MTPS-ANN	**0.914**	**0.926**	**0.957**

Table 5: The R^2 (coefficient of determination) of the proposed MTPS-ANN model and multiple alternative baselines: (1) a simple linear regression model based on the storage duration only (*Storage duration*), (2) a linear regression model based on all the input features (*Linear Regression*), (3) Support Vector Machines (*SVM*), (4) *Gradient Boosting*, (5) *Random Forests*, (6) An ablation model of an artificial neural trained on a single objective (*ANN Single-objective*), and finally (7) the proposed model (*MTPS-ANN*).

task property of MTPS-ANN, its results deteriorate significantly. This serves to highlight the importance of the multi-task property to unlock the full potential of the neural networks in the face of limited training data.

The RMSE values allow for a comparison between competing alternatives. However, this metric does not provide insight into the amount of explained variance. To this end, Table 5 compares the R^2 (coefficient of determination) for the MTPS-ANN model and the respective baselines. As can be seen, the MTPS-ANN model is able to explain the majority of the uncertainty (variance); namely, $\approx$ 90%-95% depending on the prediction task. In addition, all the aforementioned trends from Table 4 seem to repeat themselves: the superiority of MTPS-ANN over the baselines, the importance of the quality and storage features, and the significant contribution of multi-task training.

4.3. Comparing Dimension Reduction Techniques

In PCA, eigenvalue analysis can be performed in order to expose the amount of variance explained by each dimension (Burges et al., 2010). Figure 4 depicts PCA eigenvalue analysis of the pre-storage quality measures and storage conditions. As can be seen, the variance in the input features can be almost entirely explained using just 20 dimensions. This showcases the high degree of correlations that exists within the model's input features. Interestingly, the MTPS-ANN model's encoder component transforms the raw input signal into an 18-dimensional vector which is very close to the 20 dimensions found by PCA. In the MTPS-model, the dimensionality of the encoder is a hyper-parameter that was set using cross-validation optmization (see Section 3.3.5). Hence, a close relation is exposed between classical PCA's 20 dimensions and the MTPS-model's encoder that requires 18 dimensions.

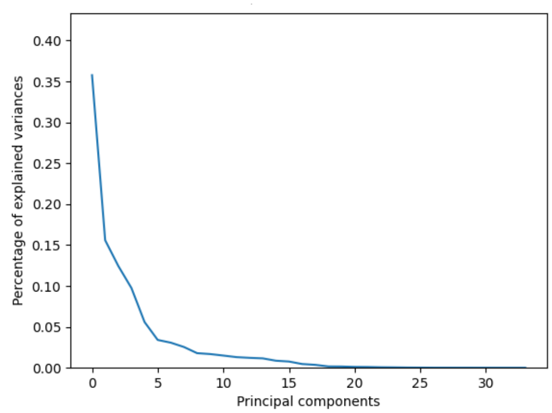

Figure 4: PCA eigenvalue analysis of the pre-storage quality measures and storage conditions. The graph demonstrates a high amount of correlation in the input features which can be accurately reduced to ≈ 20 dimensions.

Importantly, the MTPS-ANN encoder extracts the relevant information and optimally arranges it for the prediction head. A visual demonstration of MTPS-ANN encoder's ability to extract and arrange the relevant information is presented in Figure 5. The top image depicts a t-SNE (Van der Maaten & Hinton, 2008) visualization of the original pre-storage features for some portion of the dataset. The bottom image visualizes MTPS-ANN encoder's representations of the exact same data points. In both images, each point corresponds to a pre-storage datum colored by its corresponding post-storage acceptance score. The color scale ranges from 1 (low acceptance score in blue) to 5 (high acceptance score in red). The top image represents the data arrangement prior to the encoder where the relation between the input values (location) and the acceptance score (color) is complex and difficult to comprehend. In contrast, the bottom image, depicting the data after being processed by the MTPS-ANN encoder, exposes a clear trend— the MTPS-ANN encoder successfully mapped the input data that correlates with a high post-storage acceptance score to the upper-right area of the image, and data that correlates with low scores are mapped to the lower-left side of the

Dimension Reduction	**Algorithm**	**Acceptance Score**	**Rachis Score**	**% Decay**
PCA	Linear Regression	0.185	0.176	0.086
	SVM	0.190	0.166	0.095
	Gradient Boosting	0.179	0.174	0.092
	Random Forests	0.179	0.170	0.088
Auto-Encoder	Linear Regression	0.187	0.181	0.086
	SVM	0.178	0.169	0.095
	Gradient Boosting	0.185	0.176	0.093
	Random Forests	0.184	0.171	0.089
MTPS-ANN's Encoder	Linear Regression	0.165	0.17	0.084
	SVM	0.168	0.164	0.098
	Gradient Boosting	0.173	0.163	0.084
	Random Forests	0.174	0.165	0.084
MTPS-ANN		**0.161**	**0.148**	**0.070**

Table 6: Comparing the RMSE of different dimension reduction techniques—Principal Component Analysis, Auto-encoder, and employing the MTPS-ANN's encoder for dimension reduction.

image. This visualization indicates that the MTPS-ANN encoder has learned to arrange the data in space in order to assist the acceptance score prediction task.

Classical machine learning algorithms do not perform supervised dimension reduction as part of the learning process. As a final experiment to showcase the added-value of MTPS-ANN's encoder, the classical machine learning baselines models (Multiple Linear Regression, SVM, Gradient Boosting, Random Forests) were trained again, now with the data first going through three comparable dimensionality reduction techniques: (1) PCA (Jackson, 2005), (2) Auto Encoder (Bank et al., 2020), and (3) the pre-trained MTPS-ANN's encoder. Clearly, option (3) is presented for experimental purposes only, as it requires first training a MTPS-ANN model, and then use its encoder in order to train the other models.

Table 6 and Table 7 present the RMSE and R^2 achieved by employing the above dimension reduction options (PCA, Auto-Encoder, or MTPS-ANN's Encoder) and different classical machine learning models (Linear Regression, SVM, Gradient Boosting, and Random Forests). The results in Table 6 and

Dimension Reduction	**Algorithm**	**Acceptance Score**	**Rachis Score**	**Decay**
PCA	Multiple Linear Regression	0.887	0.896	0.935
	SVM	0.882	0.907	0.919
	Gradient Boosting	0.895	0.897	0.925
	Random Forests	0.895	0.902	0.931
Auto-Encoder	Multiple Linear Regression	0.884	0.890	0.935
	SVM	0.896	0.904	0.920
	Gradient Boosting	0.887	0.895	0.924
	Random Forests	0.889	0.902	0.929
MTPS-ANN's Encoder	Multiple Linear Regression	0.911	0.902	0.938
	SVM	0.907	0.910	0.914
	Gradient Boosting	0.902	0.910	0.937
	Random Forests	0.901	0.909	0.937
MTPS-ANN		**0.914**	**0.926**	**0.957**

Table 7: Comparing the R^2 (coefficient of determination) of different dimension reduction techniques—Principal Component Analysis, Auto-encoder, and employing the MTPS-ANN's encoder for dimension reduction.

Table 7 indicate that the classical algorithm successfully utilises the MTPS-ANN encoder's representations in order to achieve better results than non-supervised dimension reduction methods (PCA and Auto-Encoder). Furthermore, when compared to the auto-encoder, the results indicate that the non-linearity by itself does not explain the encoder's advantage, and the main advantage of the encoder stems from its supervision with respect to the prediction task. This experimental set-up further showcases the superiority of performing supervised dimension reduction such as in MTPS-ANN encoder over classical dimension reduction techniques.

4.4. MTPS-ANN as a decision support system component

Conceptually, the MTPS-ANN can be a component in a decision support system that recommends the order in which the agricultural produce is taken out of storage. To this end, the degree of monotony, an estimate of the order of the model, was examined: The clusters were arranged according to experts' post-storage acceptance score (AS), and the Spearman correlation test (Myers & Sirois, 2004) was performed in order to examine how well

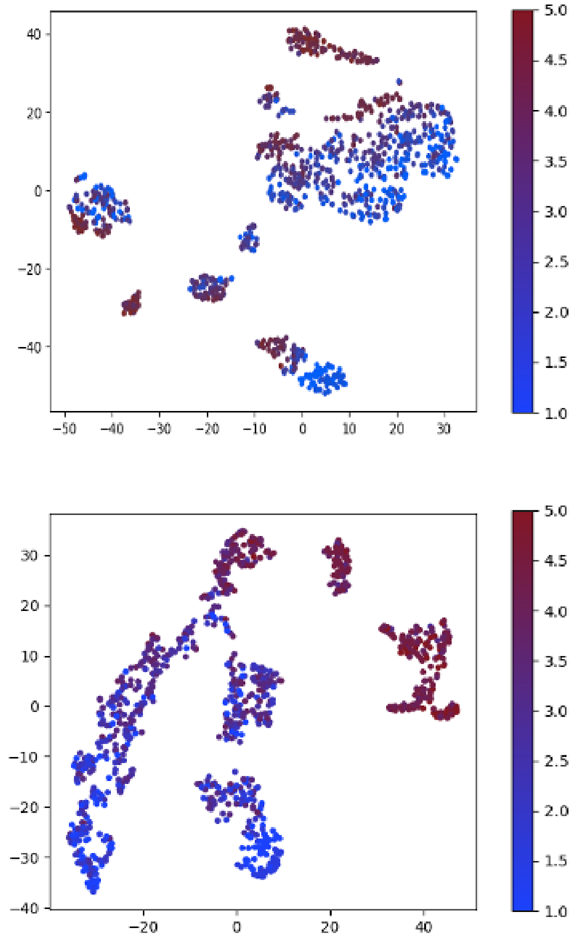

Figure 5: t-SNE images of the original input features (top) and their representations extracted by MTPS-ANN's encoder (bottom). Each point represents a pre-storage datum colored by its corresponding post-storage acceptance score. The color scale ranges from 1 (blue) to 5 (red). The upper image visualizes the original data (prior to the encoder) and does not expose a straightforward relation between the data points arrangement and the acceptance score. In contrast, the bottom image visualizes the same data points after going through MTPS-ANN's encoder. The bottom image reveals the ability of the encoder to arrange the data according to the acceptance score where high acceptance score values are mostly concentrated in the upper right and low acceptance score values are mostly concentrated at the lower left area of the image.

the order was maintained by the prediction model. A Spearman correlation coefficient of $R = 0.86$, $p < 0.0001$ was found which suggests high utility for the MTPS-ANN model in supporting storage management decisions. To explore how well the order of acceptance score was maintained at different score ranges, subsets of data were tested as well: a high-quality subset ($4 < AS \leq 5$), a medium-high quality subset ($3 < AS \leq 4$), a medium quality subset ($2 < AS \leq 3$), and finally a low-quality subset ($1 \leq AS \leq 2$). Spearman correlation tests were conducted for each subset and results were as follows: $R = 0.53$, $p < 0.0001$ for the high quality, $R = 0.62$, $p < 0.0001$ for the medium-high quality, $R = 0.3$, $p < 0.0001$ for the medium quality subset, and $R = 0.11$, $p = 0.07$ for the low quality subset.

5. Conclusions

Food loss has a major negative impact on food security, quality, safety, and on the environment Meybeck et al. (2011). For agricultural produce, a substantial amount of loss occurs at the post-harvest stages, i.e., storage and distribution. When it comes to meeting future food demand, the investment in streamlining processes that will reduce food loss is lower than the expected cost of additional resources needed to increase production (Meybeck et al., 2011). Therefore, reducing food loss is a major challenge that deserves more attention from the academic community. Currently, postharvest storage management of fruits and vegetables is principally governed by the First In, First Out (FIFO) logistics strategy, meaning that marketing decisions are based solely on storage time irrespective of the initial quality of the produce and its remaining potential shelf life. Implementation of the MTSP-ANN into an intelligent First Expired, First Out (FEFO) logistics-management system, that provides reliable information regarding the remaining shelf-life of each batch of produce, will enable efficient inventory management based on shelf-life predictions for each particular batch of produce. Applying FEFO strategy will ensure that only high-quality produce will reach the distinct marketing destinations, and is expected to reduce food loss during the post-harvest, storage, and distribution phases.

The goal of the current work is to employ state-of-the-art deep learning techniques to predict post-storage produce quality based on different measurable signals and under different storage conditions. A key challenge in this line of work stems from the difficulty of obtaining large data-sets of pre-storage and post-storage quality measurements under different storage conditions since the data collection effort is time-consuming and extremely laborious. The current research is based on a large-scale storage experiment using the Scarlotta Seedless® table grape. Notably, the data-set size in this research exceeds that of any previous post-storage quality prediction research. Nevertheless, the dataset size is still relatively small in the context of deep-learning models.

Based on the data collected, a Multi-Task Post-Storage Artificial Neural Network (MTPS-ANN) model was developed which uses post-harvest quality measurements of harvested produce in order to predict objective and subjective post-storage quality measures. MTPS-ANN mitigates data sparsity by employing two complementary techniques: First, an encoder performs supervised dimensionality reduction in order to remove redundancies and distill the informative signal. Second, a multi-task prediction head is employed to enhance supervision by utilizing several labels for each datum. Moreover, as a multi-task model, MTPS-ANN simplifies the post-storage prediction task by removing the need to train different models per each prediction task. Evaluations against common alternative approaches from classical machine learning showcase the advantage of the MTPS-ANN model in terms of prediction accuracy. Additionally, an in-depth analysis of MTPS-ANN's supervised dimension reduction is presented in order to analyze its contribution to the overall model accuracy.

Importantly, while the MTPS-ANN model was developed for the task of post-storage grapes quality prediction, its architecture is general and can be useful for other agricultural produce. MTPS-ANN is a general neural network for multi-task regression problems, with an architecture designed specifically to tackle common issues in many agricultural learning tasks: (1) A limited-size dataset compared to classical deep-learning tasks. (2) Correlations in the input features. (3) Unknown, potentially non-linear, relations between the inputs and outputs. (4) The ability to support multiple types of predictions using a single model. As such, it is a general model, which does not take dependencies on the specific inputs or outputs of the post-storage grapes quality problem. Therefore, the MTPS-ANN model can be easily adapted to tackle other agricultural learning tasks which exhibit similar properties to those described above. The general design guidelines that were explained and motivated in Section 3 are likely to be effective at treating other agricultural machine learning tasks where similar challenges arise.

References

Admane, N., Altieri, G., Genovese, F., Di Renzo, G., Verrastro, V., Tarricone, L., & Ippolito, A. (2015). Application of high carbon dioxide or ozone combined with map on organic late-season 'scarlotta seedless®'table grapes. *Acta Horticulturae*, *1079*, 193–199. doi:10.17660/ActaHortic.2015.1079.21.

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys*, *4*, 40–79. doi:10.1214/09-SS054.

Aujla, K. M., Shah, N. A., Ishaq, M., & Farooq, A. (2011). Post-harvest losses and marketing of grapes in Pakistan. *Sarhad Journal Agriculture*, *27*, 487–490.

Bahar, A., Kaplunov, T., Alchanatis, V., & Lichter, A. (2017). Evaluation of methods for determining rachis browning in table grapes. *Postharvest Biology and Technology*, *134*, 106–113. doi:10.1016/j.postharvbio.2017.08.016.

Bahar, A., & Lichter, A. (2018). Effect of controlled atmosphere on the storage potential of ottomanit fig fruit. *Scientia Horticulturae*, *227*, 196–201. doi:10.1016/j.scienta.2017.09.036.

Bank, D., Koenigstein, N., & Giryes, R. (2020). Autoencoders. *arXiv*, *abs/2003.05991*.

Baxter, J. (1997). A bayesian information theoretic model of learning to learn via multiple task sampling. *Machine learning*, *28*, 7–39. doi:10.1023/A:1007327622663.

Biau, G., & Scornet, E. (2016). A random forest guided tour. *Test*, *25*, 197–227. doi:10.1007/s11749-016-0481-7.

Blanckenberg, A., Opara, U. L., & Fawole, O. A. (2021). Postharvest losses in quantity and quality of table grape (Cv. crimson seedless) along the supply chain and associated economic, environmental and resource impacts. *Sustainability*, *13*. doi:10.3390/su13084450.

Burges, C. J. et al. (2010). Dimension reduction: A guided tour. *Foundations and Trends in Machine Learning*, *2*, 275–365. doi:10.1561/2200000002.

Centre for the Promotion of Imports from developing countries (CBI) Ministry of Foreign Affairs (2021). Entering the european market for table grapes. `www.cbi.eu/market-information/fresh-fruit-vegetables/table-grapes/market-entry`. Accessed: 2021-12-01.

Crawshaw, M. (2020). Multi-task learning with deep neural networks: A survey. *ArXiv*, *abs/2009.09796*.

Crisosto, C. H., & Crisosto, G. M. (2020). Manual on postharvest handling of mediterranean tree fruits and nuts. chapter 7. CABI.

Crisosto, C. H., Smilanick, J. L., & Dokoozlian, N. (2001). Table grapes suffer water loss, stem browning during cooling delays. *California Agriculture*, *55*, 39–42. doi:10.3733/ca.v055n01p39.

Dua, D., & Graff, C. (2017). Uci machine learning repository. `http://archive.ics.uci.edu/ml`.

Fodor, I. K. (2002). *A survey of dimension reduction techniques*. Technical Report Lawrence Livermore National Lab., CA (US). doi:10.48550/arXiv.1403.2877.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational statistics & data analysis*, *38*, 367–378. doi:10.1016/S0167-9473(01)00065-2.

Ghozlen, N. B., Cerovic, Z. G., Germain, C., Toutain, S., & Latouche, G. (2010). Non-destructive optical monitoring of grape maturation by proximal sensing. *Sensors*, *10*, 10040–10068. doi:10.3390/s101110040.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. `http://www.deeplearningbook.org`.

Gunders, D., & Bloom, J. (2017). *Wasted: How America is losing up to 40% of its food from farm to fork to landfill*. Technical Report Natural Resources Defense Council.

Hertog, M. L., Uysal, I., McCarthy, U., Verlinden, B. M., & Nicolaï, B. M. (2014). Shelf life modelling for first-expired-first-out warehouse management. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*. doi:10.1098/rsta.2013.0306.

Hu, C., Thomasson, J. A., & Bagavathiannan, M. V. (2021). A powerful image synthesis and semi-supervised learning pipeline for site-specific weed detection. *Computers and Electronics in Agriculture*, *190*, 106423. doi:/10.1016/j.compag.2021.106423.

Jackson, J. E. (2005). *A user's guide to principal components* volume 587. John Wiley & Sons. doi:10.1002/0471725331.

Jedermann, R., Nicometo, M., Uysal, I., & Lang, W. (2014). Reducing food losses by intelligent food logistics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*. doi:10.1098/rsta.2013.0302.

Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* volume 4. John Wiley & Sons. doi:10.1002/9781118874059.

Le, Q. V., Ngiam, J., Coates, A., Lahiri, A., Prochnow, B., & Ng, A. Y. (2011). On optimization methods for deep learning. In *The International Conference on Machine Learning (ICML)*. doi:10.1063/5.0066319.

Li, Y., Chu, X., Fu, Z., Feng, J., & Mu, W. (2019). Shelf life prediction model of postharvest table grape using optimized radial basis function (RBF) neural network. *British Food Journal*, *121*, 2919–2936. doi:10.1108/BFJ-03-2019-0183.

Liashchynskyi, P., & Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: a big comparison for nas. *arXiv preprint arXiv:1912.06059*, .

Lichter, A. (2016). Rachis browning in table grapes. *Australian journal of grape and wine research*, *22*, 161–168. doi:10.1111/ajgw.12205.

Lichter, A., Kaplunov, T., Zutahy, Y., Daus, A., Alchanatis, V., Ostrovsky, V., & Lurie, S. (2011). Physical and visual properties of grape rachis as affected by water vapor pressure deficit. *Postharvest Biology and Technology*, *59*, 25–33. doi:10.1016/j.postharvbio.2010.07.009.

Lichter, A., Kaplunov, T., Zutahy, Y., Daus, A., Maoz, I., Beno-Mualem, D., & Lurie, S. (2014). Effects of gibberellin on cracking and postharvest quality of the seeded table grape 'zainy'. *Acta Horticulturae*, (pp. 265–271). doi:10.17660/ActaHortic.2015.1079.31.

Lichter, A., Kaplunov, T., Zutahy, Y., & Lurie, S. (2016). Unique techniques developed in israel for short-and long-term storage of table grapes. *Israel Journal of Plant Sciences*, *63*, 2–6. doi:10.1080/07929978.2016.1151289.

Lichter, A., & Romanazzi, G. (2017). Postharvest diseases of table grapes. In D. Prusky, & J. Adaskaveg (Eds.), *Postharvest Pathology of Fruits and Vegetables*. APS Press.

Luvisi, D. A., Shoery, H. H., Smilanick, J. L., Thomposon, J. f., Gump, B. H., & Knutson, J. (1992). *Sulfur dioxide fumigation of table grapes*. 1932. Division of Agriculture and Natural Resources, University of California.

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-sne. *Journal of machine learning research*, *9*.

Maftoonazad, N., Karimi, Y., Ramaswamy, H. S., & Prasher, S. O. (2011). Artificial neural network modeling of hyperspectral radiometric data for quality changes associated with avocados during storage. *Journal of Food Processing and Preservation*, *35*, 432–446. doi:10.1111/j.1745-4549.2010.00485.x.

Meybeck, A., Cederberg, C., Gustavsson, J., van Otterdijk, R., & Sonesson, U. (2011). *Global Food Loss and Food Waste: extent, causes and prevention*. Technical Report Swedish Institute for Food and Biotechnology (SIK) Gothenburg, Sweden and FAO Rome, Italy and Food and Agriculture Organization of the United Nations.

Myers, L., & Sirois, M. J. (2004). Spearman correlation coefficients, differences between. *Encyclopedia of statistical sciences*, *12*. doi:10.1002/0471667196.ess5050.pub2.

Nelson, K. (1979). *Harvesting and handling California table grapes for market*. UCANR Publications.

Owoyemi, A., Lapidot, O., Kochanek, B., Zahavi, T., Salzer, Y., Porat, R., & Lichter, A. (2022). Sour rot in the vineyard is an indicator of botrytis rot in grapes after storage. *Postharvest Biology and Technology*, *191*, 111980. doi:10.1016/j.postharvbio.2022.111980.

Parfitt, J., Barthel, M., & MacNaughton, S. (2010). Food waste within food supply chains: Quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B*, *365*, 3065–3081. doi:10.1098/rstb.2010.0126.

Perera, E., Silva, R. G. B. e., Spagnol, W. A., & Silveira Junior, V. (2018). Water loss in table grapes: model development and validation under dynamic storage conditions. *Food Science and Technology International*, *38*, 473–479. doi:10.1590/1678-457x.08817.

Porat, R., Greenhot, Z., & Fridkeen, Z. (2016). *Summary of a survey of the loss of fresh fruits and vegetables*. Technical Report Ministry of Agriculture and Rural Development (Israel).

Porat, R., Lichter, A., Terry, L. A., Harker, R., & Buzby, J. (2018). Postharvest losses of fruit and vegetables during retail and in consumers' homes: Quantifications, causes, and means of prevention. *Postharvest Biology and Technology*, *139*, 135–149. doi:10.1016/j.postharvbio.2017.11.019.

Salehi, F. (2020). Recent Advances in the Modeling and Predicting Quality Parameters of Fruits and Vegetables during Postharvest Storage: A Review. *International Journal of Fruit Science*, *20*, 506–520. doi:10.1080/15538362.2019.1653810.

Sayyari, M., Salehi, F., & Valero, D. (2017). New Approaches to Modeling Methyl Jasmonate Effects on Pomegranate Quality during Postharvest Storage. *International Journal of Fruit Science*, *17*, 374–390. doi:10.1080/15538362.2017.1329051.

Schmidt-Hieber, J. (2020). Nonparametric regression using deep neural networks with relu activation function. *The Annals of Statistics*, *48*, 1875–1897. doi:10.1214/19-AOS1875.

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and computing*, *14*, 199–222. doi:10.1023/B%3ASTCO.0000035301.49549.88.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

Statista (2018). Share of food wasted globally as of 2017, by food category. www.statista.com/statistics/519611/percentage-of-wasted-food-by-category-global/ Accessed: 2022-03-17.

Statista (2020). Global fruit production in 2019, by selected variety. www.statista.com/statistics/264001/worldwide-production-of-fruit-by-variety. Accessed: 2021-12-01.

Stenmarck, A., Jensen, C., Quested, T., & Moates, G. (2016). *Estimates of European food waste levels*. Technical Report EU Fusions.

Yahia, E. M. (2011). Postharvest biology and technology of tropical and subtropical fruits: fundamental issues. chapter 9. Elsevier.

Ying, X. (2019). An overview of overfitting and its solutions. In *Journal of Physics: Conference Series* (p. 022022). IOP Publishing volume 1168. doi:10.1088/1742-6596/1168/2/022022.

Zhang, Y., & Yang, Q. (2017). A survey on multi-task learning. *arXiv preprint arXiv:1707.08114*, .

Zoffoli, J., & Latorre, B. (2011). 9 - table grape (vitis vinifera l.). In *Postharvest biology and technology of tropical and subtropical fruits* (pp. 179–214e). Elsevier.

Zoffoli, J. P., Latorre, B. A., Rodriguez, J., & Aguilera, J. M. (2009). Biological indicators to estimate the prevalence of gray mold and hairline cracks on table grapes cv. thompson seedless after cold storage. *Postharvest Biology and Technology*, *52*, 126–133. doi:10.1016/j.postharvbio.2008.11.010.

Zutahy, Y., Lichter, A., Kaplunov, T., & Lurie, S. (2008). Extended storage of 'red globe'grapes in modified so2 generating pads. *Postharvest Biology and Technology*, *50*, 12–17. doi:10.1016/j.postharvbio.2008.03.006.