Master in Mathematics and Computing

A MODEL TO PREDICT WATER CONSUMPTION IN GROWING PIGS

Maria Marroco Ríos

Supervised by: Camilo Palazuelos

September 2024

Abstract

Water consumption plays a vital role in pig farming operations, and accurate prediction of water use is essential for efficient resource management. This study, part of the European DECIDE project—a five-year initiative aimed at developing data-driven decision support tools that provide robust and early signals of disease onset and options for diagnostic confirmation—aims to develop a comprehensive model to predict water consumption in pig herds, considering various external variables and incorporating historical water consumption data from previous batches.

First, the key parameters influencing water consumption in growing pigs, such as feed intake, humidity, indoor farm temperature, and days of growth, are identified, a task not done before. These factors are integrated into the predictive model, which was developed using both R and Python programming language.

To capture the daily seasonal pattern in water consumption, four harmonics were incorporated into the model. Through experimentation, it was discovered that adding a fourth harmonic improved prediction accuracy.

Additionally, the model was enhanced by incorporating a parameter that included historical water consumption data from other batches produced on the same day and at the same growth stage. This additional feature proved effective in improving the model's performance.

Beyond creating this model and capturing the upward trend in water consumption of a batch and daily seasonality, we also determined the daily water consumption patterns on a warm summer day and a cold winter day, noting that the pattern varies depending on the season.

The developed model was validated using historical data from several batches of pigs on the same farm. The predicted water consumption values were compared with actual values, and it was observed that all actual values were within the confidence intervals calculated using the Delta method, validating the model's reliability and accuracy.

Thus, this study successfully achieved the objective of modelling water consumption in pig herds. The findings of this research contribute to the efficient management of water resources in pig farming and lay the groundwork for future advancements in this field.

Contents

1 Introduction

Pork meat has high demand worldwide and is expected to increase in the coming years [1]. This projection presents a set of challenges for the pig sector related to developing sustainable breeding systems while ensuring animal welfare and health. However, over the years, livestock production has been subjected to increasing industrialisation, leading to much larger and more intensive production units, with less time available to attend to animals individually.

Sensor-based monitoring and early warning systems can help farmers daily by identifying animals or groups of animals that need priority attention. Ideally, the system can generate timely alerts so that farm staff can decide on appropriate interventions and prevent any condition that reduces welfare or at least mitigate its consequences. Early warning systems for livestock production, also known as detection models, have been developed over the last twenty years, often aiming to detect various conditions in individual animals, such as clinical mastitis or lameness in cows. Additionally, modelling changes in animal behaviour monitored by sensors has been an increasing focus in precision livestock farming over the past decades. For example, animal behaviour modelling has been used as an early indicator of diseases or as a decision-making tool for managing groups of growing pigs.

Specifically, some studies conducted by *The National Committee for Pig Production*, *Danish Bacon* and *Meat Council* [2], [3], and [4] have highlighted the potential of monitoring the water consumption pattern of growing pigs. Under normal conditions, pigs exhibit a stable diurnal water consumption pattern, whereas disease outbreaks, changes in feed quality, or ventilation problems often cause deviations from this pattern. Therefore, real-time monitoring of water consumption in finishing pigs seems to be a potential way to improve the management and handling of these animals [3]. Thus, to detect changes in water consumption behaviour, a well-founded model to predict expected behaviour is essential.

Although this field has been growing in recent years, very few studies on the water consumption pattern of pigs have been found in the literature [10] - [12], while their feeding behaviour seems better described. It is worth mentioning that in most studies, the intake pattern of finishing pigs housed in groups (batches) is calculated as the average over the entire test period. For example, in Hyun et al. (1997) [7], the diurnal pattern is estimated as an average of 10 weeks, or in Slader et al. (1998) [8], an average of 53 days is used. In others, such as Nienaber et al. (2011-2018) [9], the analysis was made more flexible by dividing the test period into five intervals. For each interval, the intake pattern is estimated as averages of 7-day periods. Thus, a common pattern is observed in all the aforementioned studies: the lack of dynamics. The daily pattern is estimated as an average over a number of days. The disadvantage of this method is that it does not reflect a change in the feeding pattern over the period.

Water consumption is, of course, a continuous process, but for monitoring purposes, consumption must be measured in discrete intervals, i.e., the amount of water consumed in a given period. In previous studies, such as Madsen et al. (2001; 2005) [5] [6], different time intervals are considered, and it is concluded that hourly sums are the preferable option when modelling observed water consumption. For water consumption predictions for the previous hour, a model with a dynamic nature is required, allowing for development in both the diurnal pattern and the overall water consumption period as the pigs grow, i.e., the time the batch is on the farm. A good example is the auto-regressive moving average (ARMA) models, where more attention is paid to recent information than past information. Thus, as time passes, the information loses its value. Monitoring the hourly sums of water consumption of growing pigs slightly complicates the model's structure compared to the studies described above, i.e., the model must include not only linear or quadratic growth as the pigs grow but also a cyclic effect to describe the diurnal pattern. Cyclic or periodic models are widely used in commercial business, for example, to model annual deviations in oil and gas demand.

The objective of this work can be separated in two. On the one hand, describing the water consumption pattern (for example, the daily water consumption patterns on a warm summer day and a cold winter day), as well as other variables that could have an impact on such pattern. And, on the other hand, making the most of the knowledge obtained in the descriptive part, a learning ARMA mathematical model was developed for predicting the hourly water consumption pattern in growing pigs pigs. Additionally, the goal is for the model to eventually become part of a computer-based monitoring system for the comprehensive management of the farm that provided the data, located in the province of Lleida.

1.1 Objectives

As it has been mentioned, the general objective of this study is to describe the water consumption pattern in growing pigs, considering variables that may influence this pattern, and to develop an ARMA mathematical model to predict hourly water consumption, with the aim of integrating it into a computer-based monitoring system for the comprehensive management of the farm located in the province of Lleida.

Moreover, more specific objectives are pretended to be achieved:

- To describe daily water consumption patterns in growing pigs under different weather conditions, such as warm summer days and cold winter days.
- To identify and analyse additional variables that may influence the water consumption pattern in order to better understand its behaviour.
- To develop an ARMA mathematical model based on the knowledge obtained from the water consumption pattern description, allowing for the prediction of hourly water consumption in growing pigs.

2 Material and methods

2.1 Farm, herd and data

2.1.1 Fattening Farm

The herd from which the data was obtained comes from a farm owned by a renowned agricultural and livestock company based in Spain. The company has distinguished itself in the livestock sector through the implementation of advanced technologies and innovative practices. In particular, the company has invested in digitisation and automation in production processes. This includes the use of real-time monitoring and control systems, optimisation of feed and animal welfare, and the implementation of efficient management systems.

All of this relates to the farm from which the data was obtained. This particular farm combines the latest technologies with sustainable practices to transform the way pigs are raised and to reduce the negative impact on the environment. The farm is characterised by the implementation of technologies such as artificial intelligence, robotics, and data analysis. These tools allow for more precise animal monitoring, automated control of feeding systems, and decision-making based on real-time information. They also promote animal welfare and responsible production, considering ethical and quality aspects in animal breeding and management.

The farm in question is located in the province of Lleida. Structurally, it consisted of three separated buildings with two units each, as it can be seen in Figure 1. The two units in each building are separated but connected by a central corridor. Each barn has 56 pens of $9m^2$ that house approximately 12-14 pigs (around 700 pigs per barn). The farm allowed and all-in/all-out management system, with the pigs entering at a weight of between 21-25kg and leave the facilities at a final weight of about 110kg.

2.1.2 Herd

Water consumption data was obtained from several groups of finishing pigs (referred to as batches). When a group or batch of pigs arrives at the farm, the animals are distributed among the pens in the six units, as shown in Figure 1. However, in this study, data is used per barn (this aspect is explained further in section 2.1.3). It is important to note that all pigs in a batch enter the farm at the same time and are of the same age relative to the weaning date. When a barn is emptied by sending the pigs to the slaughterhouse, it is cleaned and dried for biosecurity reasons before introducing a new batch of pigs. For our study group, the growth period (20 − 110kg) is approximately 20 weeks, including one week for cleaning.

2.1.3 Data

In relation to the development of a new monitoring system, the farm has installed water flow meters and microcomputers in each of the six barns. Data on water consumption from various batches has been collected, with the following characteristics:

- For each of the six barns, data from five consecutive different batches has been collected. In total, data from 30 batches has been used, resulting in 30 time series. Taking into account that there were between 660-730 pigs per batch, the water consumption of more than 2000 pigs was analysed.

Figure 1: Representation of the six buildings of the farm under study.

- The study period spans from December 2020 to January 2024.
- Hourly water consumption (in litters) has been recorded, providing 24 data points per day.
- The cleaning weeks between the departure of one batch and the arrival of the next have not been used in the model construction.
- The number of pigs per batch is similar but not constant. For this reason, the total water consumption of the barn has been divided by the number of pigs in the respective barn, so that each time series indicates water consumption per animal per hour of the day.
- The data contains some random noise due to biological variation and reading errors.

Figure 2 shows one of the time series, corresponding to the five batches in barn 1.01. This data has been used to find a suitable model for hourly water consumption per pig.

2.1.4 Indicator Parameters

In addition to having data on the animals' water consumption every 24 hours, we also had external parameter variables that could be associated with water consumption and, therefore, be useful in predicting this consumption in finishing pigs. In particular, these variables were:

- **Feed consumption**: There is a direct relationship between feed consumption, which has also been divided by the number of animals in the pen, and water consumption in pigs. When pigs consume more feed, they likely need more water to facilitate digestion and stay hydrated. Insufficient water supply leads to a decrease in feed intake and, consequently, a decrease in performance and feed efficiency [14].

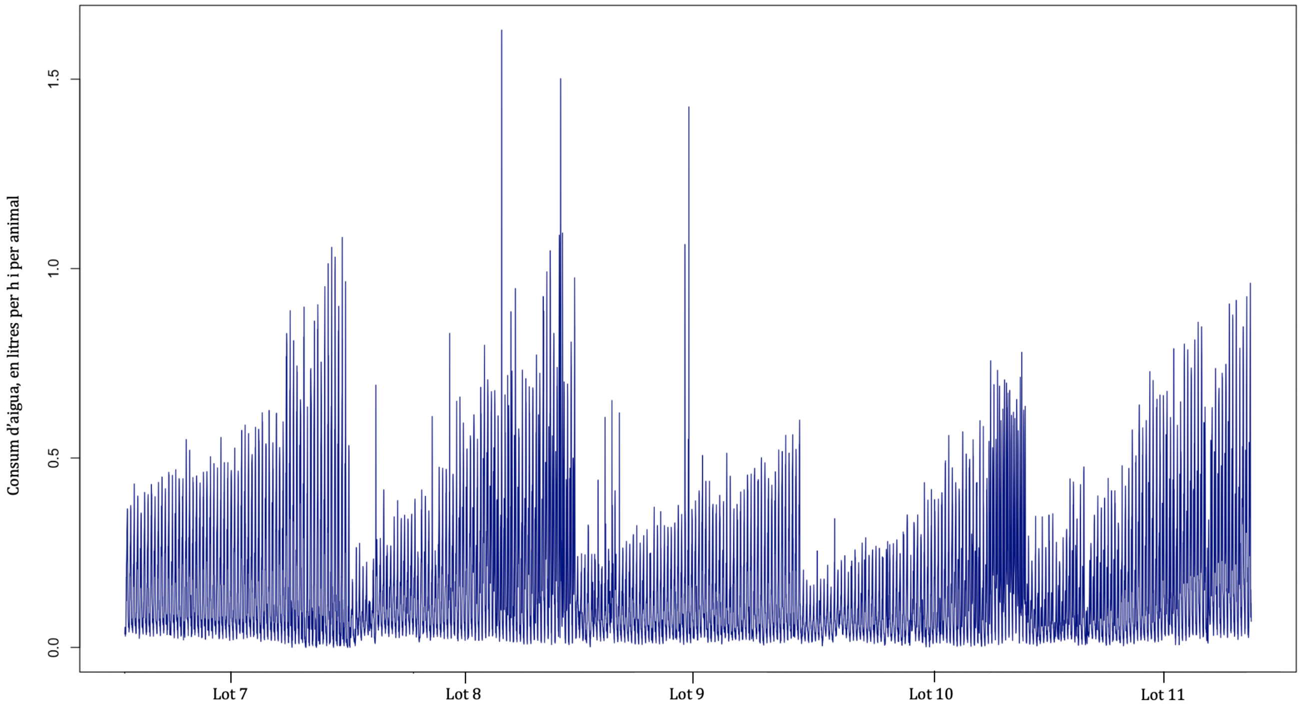


Figure 2: Time series of water consumption per animal per hour for the five batches in barn 1.01.

- **Exterior and interior temperature of the farm**: Temperature significantly impacts the behaviour and metabolism of pigs. Since pigs sweat very little, they achieve thermoregulation by seeking opportunities to cool down (e.g., soaking in water or mud) or through panting [15]. In hot conditions, pigs can experience heat stress, which typically leads to a greater need for water to regulate their body temperature and stay cool. Conversely, in cold conditions, pigs need more feed to warm up and consequently, more water [16].

- **Concentration of NH_3 and CO_2 in the air**: The environment in which pigs are housed is crucial for their welfare and health. Both the concentration of NH_3 (ammonia) and CO_2 are related to air quality in the farm and the degree of ventilation. High levels of these gases can negatively affect the pigs' welfare and, consequently, their water consumption. For example, high concentrations of ammonia could cause tail biting, significantly reducing animal welfare [17].

- **Humidity**: Humidity affects water consumption in pigs in several ways. Firstly, high humidity conditions increase the pigs' hydration needs due to evaporation and body water loss. Additionally, heat stress caused by a combination of high humidity and high temperatures reduces pigs' water consumption. Water quality can also deteriorate in humid environments, affecting the appeal and safety of water for pigs. Finally, pigs' behaviour may change in high humidity conditions, potentially decreasing their access and motivation to drink water. In summary, it is important to ensure an adequate supply of clean and fresh water during periods of high humidity to maintain optimal pig welfare and performance.

Therefore, the aforementioned variables can be associated with pigs' water consumption and their welfare, as well as being good predictors of this consumption for the early detection of abnormal water consumption behaviours that may be related to, for example, respiratory diseases. This association will be studied in this work, along with the capacity of these variables to predict the water consumption of finishing pigs. By modelling water consumption based on these variables, the water needs of pigs can be understood more realistically, enabling informed decisions to optimise their welfare and productive performance.

The following section presents a model designed to adapt to the characteristics of the water consumption data.

2.2 Methods

As mentioned in previous sections, water consumption is a crucial aspect of pig production since an adequate water supply is necessary to maintain their health and well-being. Given the nature of the data described above (i.e., data collected hourly over a period of time), time series modelling techniques will be used to describe and predict pigs' water consumption. In particular, autoregressive moving average models (ARMA models) will be employed, which are described in more detail below.

However, some basic concepts in classical time series analysis are introduced first, necessary for an accurate description of ARMA models.

2.2.1 Basic Concepts

A univariate time series is a set of observations $\{x_t\}$, each recorded at a time $t \in \mathcal{T}_0$. This work focuses on the case where $\mathcal{T}_0$ is a discrete and finite set, e.g., $\mathcal{T}_0 = \{1, \cdots, T\}$. Setting aside more theoretical details, it is understood that, for a fixed time t, the observation x_t is a realisation of a random variable X_t and that the set of observations $\{x_t\}_{t\in\mathcal{T}_0}$ is part of a realisation of a stochastic process $\{X_t\}_{t\in\mathcal{T}}$, $\mathcal{T}_0 \subseteq \mathcal{T}$. In practice, however, the term time series [18] is used to refer to both the stochastic process and the observations described in Section 2.1.3.

Weakly Stationary Processes and the Autocorrelation Function

Let $\{X_t\}_{t\in\mathcal{T}}$ be a time series with $\mathbb{E}\left[X_t^2\right] < \infty$. Then, this time series is weakly stationary if: (i.) the mean function of the series, $\mu_X(t) = \mathbb{E}\left[X_t\right]$, is independent of t, $\forall\, t \in \mathbb{Z}$, and (ii.) the autocovariance function (ACFV) of the series, $\gamma_X(h) = \mathrm{Cov}\left[X_t, X_{t+h}\right] = \mathbb{E}\left[(X_t - \mu_X(t))(X_{t+h} - \mu_X(t))\right]$, is independent of t, $\forall\, t \in \mathbb{Z}$ and of each lag $h \in \mathbb{Z}$. Henceforth, when stationarity is mentioned, it will always be in the weak sense unless otherwise specified.

Let $\{X_t\}_{t\in\mathcal{T}}$ be a stationary time series. Its ACFV with lag $h \in \mathbb{Z}$ satisfies: (i.) $\gamma_X(0) \geq 0$, that is, $\gamma_X(0) = \mathbb{V}\left[X_t\right]$, (ii.) $|\gamma_X(h)| \leq \gamma_X(0)$ for all $h \in \mathbb{Z}$ [19] and, (iii.) $\gamma_X(-h) = \gamma_X(h)$ for all $h \in \mathbb{Z}$.

The autocorrelation function (ACF) of a stationary time series $\{X_t\}_{t\in\mathcal{T}}$ is defined as:

$$\rho_X(h) = \mathrm{Cor}\left[X_t, X_{t+h}\right] = \frac{\mathrm{Cov}\left[X_t, X_{t+h}\right]}{\mathbb{V}\left[X_t\right]} = \frac{\gamma_X(h)}{\gamma_X(0)}, \forall\, t, h \in \mathbb{Z},$$

with $\rho_X(h) \in (-1, 1)$, $\forall h \in \mathbb{Z}$.

In practice, the functions $\mu_X(t)$, $\gamma_X(h)$, and $\rho_X(h)$, $\forall t, h \in \mathbb{Z}$ are estimated using the sample mean, sample autocovariance function, and sample autocorrelation function of $\{x_t\}_{t\in\mathcal{T}_0}$.
That is, if we observe a time series $\{x_t\}_{t\in\mathcal{T}_0}$, $\mathcal{T}_0 = \{1, \cdots, T\}$ as a realization of a part of a stationary process $\{X_t\}_{t\in\mathcal{T}}$, $\mathcal{T} \in \mathbb{Z}$, the mean, autocovariance, and autocorrelation functions of the stationary

process are estimated such that:

$$\widehat{\mu}_X(t) = \widehat{\mu}_X = \frac{1}{T}\sum_{i=1}^{T} x_i, \quad \widehat{\gamma}_X(h) = \frac{1}{T}\sum_{i=1}^{T-h}(x_{t+h} - \widehat{\mu}_X)(x_t - \widehat{\mu}_X), \quad \text{and} \quad \widehat{\rho}_X(h) = \frac{\widehat{\gamma}_X(h)}{\widehat{\gamma}_X(0)}.$$

Note that $\widehat{\gamma}_X(0) = \widehat{\mathbb{V}}(X_t)$. For more details, see [20].

White Noise Processes

A stationary process $\{\varepsilon_t\}_{t\in\mathcal{T}}$ is a white noise process, $\{\varepsilon_t\}_{t\in\mathcal{T}} \sim WN(0, \sigma^2)$, if: (i.) $\mathbb{E}[\varepsilon_t] = 0$, $\forall t$ and (ii.) $\mathbb{E}[\varepsilon_t\varepsilon_s] = \sigma^2$ if $s = t$ or $\mathbb{E}[\varepsilon_t\varepsilon_s] = 0$ if $s \neq t$ for $s, t \in \mathcal{T}$. Given the previous points (i.)-(ii.), the process $\{\varepsilon_t\}_{t\in\mathcal{T}}$ is clearly stationary since $\mu_\varepsilon(t) = \mathbb{E}[\varepsilon_t]$ is independent of $t, \forall t$, $\gamma_\varepsilon(h)$ only depends on the lag $h \in \mathbb{Z}$, and $\mathbb{E}\left[\varepsilon_t^2\right] = \mathbb{V}[\varepsilon_t] = \sigma^2 < \infty$.
Moreover, if ε_t and ε_s are independent for all $t \neq s$, the process $\{\varepsilon_t\}_{t\in\mathcal{T}}$ is an independent white noise process (i.e., $\{\varepsilon_t\}_{t\in\mathcal{T}} \sim IWN(0, \sigma^2)$). Finally, if $\{\varepsilon_t\}_{t\in\mathcal{T}}$ are independent and identically distributed random variables following a Normal$(0, \sigma^2)$ distribution (i.e., $\{\varepsilon_t\}_{t\in\mathcal{T}} \sim GWN(0, \sigma^2)$), the process is clearly white noise and independent white noise. For more details on these processes, see [20].

The Partial Autocorrelation Function

The partial autocorrelation function (PACF), $\alpha_X(h)$, $h \in \mathbb{Z}$, of a stationary process $\{X_t\}_{t\in\mathcal{T}}$ is defined as the correlation between X_t and X_{t+h} adjusting for the intermediate variables $X_{t+1}, \cdots, X_{t+h-1}$ for $h \geq 2$. Formally:

$$\begin{aligned}\alpha_X(h) = \text{Cor}(&X_{t+h} - \widehat{\mathbb{E}}(X_{t+h}|X_{t+1}, \cdots, X_{t+h-1}),\\ &X_t - \widehat{\mathbb{E}}(X_t|X_{t+1}, \cdots, X_{t+h-1})),\end{aligned}$$

where $\widehat{\mathbb{E}}(X_{t+h}|X_{t+1}, \cdots, X_{t+h-1})$ and $\widehat{\mathbb{E}}(X_t|X_{t+1}, \cdots, X_{t+h-1})$ are the predictions of X_{t+h} and X_t respectively. More details about these predictions are provided below. See also [20]. Note that $\alpha_X(0) = 1$ and $\alpha_X(1) = \rho_X(1)$.

2.2.2 ARMA Models

ARMA(p,q) models (where p is the order of the autoregressive part and q of the moving average part) are widely used in the analysis of univariate and stationary time series. These models allow, among other things, to capture the dynamic relationships and dependency structure of a time series based on past observations and past error terms. See, for example, [21].
Formally, an ARMA(p,q) process satisfies the following equation:

$$X_t = c + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + \epsilon_t + \vartheta_1 \epsilon_{t-1} + \cdots + \vartheta_q \epsilon_{t-q}, \tag{1}$$

where $\{\varepsilon_t\}_{t\in\mathcal{T}} \sim WN(0, \sigma^2)$, $t \in \mathcal{T}$, $\varphi_p \neq 0$, $\vartheta_q \neq 0$ and $c, \varphi_1, \cdots, \vartheta_q$ are parameters.

In ARMA processes, stationarity and causality depend only on the autoregressive part of the process. In fact, the process will be stationary and causal if it can be written as an infinite moving average process (i.e., an MA(∞)). An MA(∞) process is defined by the expression:

$$X_t = \sum_{j=0}^{\infty} \psi_t \varepsilon_{t-j} = \psi(B)\varepsilon_t,$$

where $\psi_0 = 1$ and $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, $BX_t = X_{t-1}$ (backshift operator). This process is well-defined (almost surely) if $\sum_{j=0}^{\infty} |\psi_j| < \infty$. Additionally, it is a causal process since X_t depends only on the

past and not the future.
On the other hand, we can also discuss the invertibility property of the process, which in this case, depends only on the moving average part. An ARMA(p,q) process is invertible if it can be written as an AR(∞) process. That is, if the process $\{\varepsilon_t\}_{t\in\mathcal{T}}$ can be written as a linear combination of the time series $\{X_t\}_{t\in\mathcal{T}}$. For more information on MA(∞) and AR(∞) processes, see [22].

For a stationary and causal ARMA(p,q) process, the mean and covariance functions are defined by the following expressions:

- $\mathbb{E}\left(X_t\right) = \mu_X(t) = \mu_X = \dfrac{c}{\left(1 - \displaystyle\sum_{j=1}^{p} \varphi_j\right)}$,
- $\gamma_X(h) = \begin{cases} \sum_{j=1}^{p} \varphi_j \gamma_Y(h-j) + \sigma^2 \sum_{k=h}^{q} \vartheta_k \psi_{k-h}, & 0 \leq h \leq q, \\ \sum_{j=1}^{p} \varphi_j \gamma_Y(h-j), & h \geq q+1. \end{cases}$,

where ψ are the coefficients of the MA(∞) representation of the ARMA(p,q) process.

Prediction of ARMA(p,q) Processes

There are several different ways to express the predictions of these models. Thus, assuming that x_t is a causal and invertible ARMA(p,q) process, $\varphi(B)x_t = \vartheta(B)\varepsilon_t$, where $\varepsilon_t \sim$ i.i.d $N(0, \sigma_\omega^2)$. In the case of a nonzero mean, $\mathbb{E}(x_t) = \mu_x$, we simply substitute x_t by $x_t - \mu_x$ in the model. First, we consider two types of predictions. We write x_{T+m}^{T} to indicate the minimum mean squared error predictor of x_{T+m} based on the data $\{x_1, \ldots, x_n\}$, that is:

$$x_{n+m}^{n} = \mathbb{E}(x_{n+m} | x_n, \ldots, x_1),$$

For ARMA models, it is easier to compute the prediction of x_{n+m} assuming we have the complete history of the process $\{x_n, x_{n-1}, \ldots, x_1, x_0, x_{-1}, \ldots\}$. We will denote the estimation of x_{n+m} based on the *infinite past* as:

$$\widehat{x}_{n+m} = \mathbb{E}(x_{n+m} | x_n, x_{n-1}, \ldots, x_1, x_0, x_{-1}, \ldots).$$

In general, x_{n+m}^{n} and $\widehat{x}_{n+m}$ are not the same, but the idea here is that for large samples, $\widehat{x}_{n+m}$ will provide a good approximation to x_{n+m}^{n}. Now, if we write x_{n+m}^{n} in its causal and invertible forms:

$$x_{n+m} = \sum_{j=0}^{\infty} \psi_j \omega_{n+m-j}, \qquad \psi_0 = 1 \tag{2}$$

$$\omega_{n+m} = \sum_{j=0}^{\infty} \pi_j \omega_{n+m-j}, \qquad \pi_0 = 1 \tag{3}$$

Then, taking the conditional expectations of (2), we have

$$\widehat{x}_{n+m} = \sum_{j=0}^{\infty} \psi_j \widehat{\omega}_{n+m-j} = \sum_{j=m}^{\infty} \psi_j \omega_{n+m-j} \tag{4}$$

because, due to causality and invertibility,

$$\widehat{\omega}_t = \mathbb{E}(\omega_t | x_n, x_{n-1}, \dots, x_1, x_0, x_{-1}, \dots) = \begin{cases} 0, & t > n \\ \omega_t, & t \le n \end{cases} \tag{5}$$

Similarly, taking the conditional expectations of (3), we have

$$\widehat{x}_{n+m} = -\sum_{j=1}^{m-1} \pi_j \widehat{x}_{n+m-j} - \sum_{j=m}^{\infty} \pi_j x_{n+m-j} \tag{6}$$

where we have used that $\mathbb{E}(x_t | x_n, x_{n-1}, \dots, x_1, x_0, x_{-1}, \dots) = x_t$ for $t \le n$.

The prediction is then achieved recursively using (6), starting with the *one step ahead* prediction, $m = 1$, and then continuing for $m = 2, 3, \dots$. With (4), we can write

$$x_{n+m} - \widehat{x}_{n+m} = \sum_{j=0}^{m-1} \psi_j \omega_{n+m-j} \tag{7}$$

so that the mean squared prediction error can be written as

$$P_{n+m}^n = \mathbb{E}(x_{n+m} - \widehat{x}_{n+m})^2 = \sigma_\omega^2 \sum_{j=0}^{m-1} \psi_j^2 \tag{8}$$

We also note that for a fixed sample size n, the prediction errors are correlated. That is, for $k \ge 1$,

$$\mathbb{E}\{(x_{n+m} - \widehat{x}_{n+m})(x_{n+m+k} - \widehat{x}_{n+m+k})\} = \sigma_\omega^2 \sum_{j=0}^{m-1} \psi_j \psi_{j+k} \tag{9}$$

Maximum Likelihood and Least Squares Estimation

For general ARMA models, it is difficult to write the likelihood as an explicit function of the parameters. Instead, it is advantageous to express the likelihood in terms of innovations or one-step-ahead prediction errors, $x_t - x_t^{t-1}$.

For a normal ARMA(p,q) model, let $\beta = (\mu, \varphi_1, \dots, \varphi_p, \vartheta_1, \dots, \vartheta_q)'$ be the $(p + q + 1)$-dimensional vector of model parameters. The likelihood can be written as

$$L(\beta, \sigma_\omega^2) = \prod_{t=1}^{n} f(x_t | x_{t-1}, \dots, x_1)$$

The conditional distribution of x_t given $x_{t-1}, \dots, x_1$ is Gaussian with mean x_{t-1} and variance P_{t-1}. Recall that $P_{t-1} = \gamma(0) \prod_{j=1}^{t-1} (1 - \phi_{jj}^2)$.

For ARMA models, $\gamma(0) = \sigma_\omega^2 \sum_{j=0}^{\infty} \varphi_j^2$, and in this case we can write

$$P_t^{t-1} = \sigma_\omega^2 \left\{ \sum_{j=0}^{\infty} \varphi_j^2 \prod_{j=1}^{t-1} (1 - \phi_{jj}^2) \right\} = \sigma_\omega^2 r_t$$

where r_t is the term inside the braces. Note that the terms r_t are functions of the regression parameters only and can be computed recursively as $r_{t+1} = (1 - \phi_{tt}^2) r_t$ with the initial condition $r_1 = \sum_{j=0}^{\infty} \varphi_j^2$. The likelihood of the data can now be written as

$$L(\beta, \sigma_\omega^2) = (2\pi\sigma_\omega^2)^{-n/2} [r_1(\beta) r_2(\beta) \cdots r_n(\beta)]^{-1/2} \exp\left[-\frac{S(\beta)}{2\sigma_\omega^2}\right] \tag{10}$$

where

$$S(\beta) = \sum_{t=1}^{n} \frac{(x_t - x_t^{t-1}(\beta))^2}{r_t(\beta)} \tag{11}$$

Both x_{t-1} and r_t are functions that depend only on β, as seen in expressions (10) and (11). Given values of β and σ_ω^2, maximum likelihood estimation would proceed by maximizing (10) with respect to β and σ_ω^2.

Delta Method

If $\eta = g(\theta)$ is a one-to-one transformation of θ, then the asymptotic distribution of its MLE $\widehat{\eta} = \widehat{g}(\theta)$ is given by

$$\widehat{\eta} \sim N(\eta, g'(\theta)^2 (F^*(\theta))^{-1})$$

if $g(\theta)$ is continuously differentiable with $g'(\theta) \neq 0$.

For a $\boldsymbol{\theta}$ d-dimensional and the function $\boldsymbol{g} = (g_1, \cdots, g_r)$, where $r \leq d$, the formulation is

$$\widehat{\boldsymbol{\eta}} \sim N(\boldsymbol{\eta}, \boldsymbol{G}'\boldsymbol{G}(\boldsymbol{F}^*(\theta))^{-1})$$

where the ij-th entry of $\boldsymbol{G}$ is equal to $\frac{\partial g_i(\boldsymbol{\theta})}{\partial \theta_j}$.

2.2.3 Model Description

An ARMA(p,q) model will be used to model the water consumption of growing pigs, as it provides several advantages. Here are some reasons why these models are useful in this context:

- **Capturing patterns and trends**: ARMA(p,q) models can identify patterns and trends in water consumption data over time. This can help understand how water consumption varies at different times and whether there are seasonal patterns or long-term trends.
- **Predicting future values**: These models allow for forecasting future water consumption. Based on past observations and past error terms, these models can estimate how water consumption will evolve in the coming periods. This is particularly useful for resource planning and decision-making in pig production.
- **Identifying influential factors**: By incorporating regressors or other variables into ARMA(p,q) models, it is possible to account for other factors that may influence pig water consumption. For example, variables such as outside temperature, feed consumption, or the interior temperature of the barn can be included to assess their impact on water consumption. This provides an opportunity to better understand the relationships and factors affecting water consumption and optimize production practices.

- **Evaluating policies and strategies**: Using ARMA(p,q) models to model the water consumption of growing pigs can also serve as a tool to evaluate different management policies and strategies. Hypothetical scenarios can be simulated to analyze how they would affect water consumption and pig welfare. This provides valuable information for decision-making in production and the implementation of more efficient and sustainable practices.

ARMA(p,q) Models with Parameters

ARMA(p,q) models that include regressor parameters are known as ARMAX(p,q) models. Adding regressors is useful for several reasons. Firstly, it improves the model's fit to the data by considering the influence of other variables that might be related to the time series of interest. This allows for better capturing of the relationships between variables and obtaining a more accurate model.

Additionally, it is possible to control or account for important factors that may affect the time series. These external factors can introduce noise or bias into the analysis, and by including them in the model, their impact is reduced, leading to more reliable results.

Another advantage of adding regressors is the ability to explore causal effects and cause-and-effect relationships between variables. This allows for analyzing how changes in indicator variables translate into changes in the target time series, which helps to better understand the underlying mechanisms and interactions between variables.

Moreover, considering these external variables in an ARMA(p,q) model improves the accuracy of predictions. These additional variables allow for better capturing of fluctuations and trends present in the data, resulting in more accurate and reliable forecasts.

Finally, including regressors facilitates optimization and informed decision-making. If the regressor variables are relevant to the study context, they can be used to predict trends and make decisions based on the model's results.

Thus, including them will improve data fit, control for confounding factors, explore causal effects, enhance prediction accuracy, and facilitate informed decision-making. These advantages make ARMA(p,q) models with regressor variables a valuable tool in analyzing time series with multiple external variables available.

An ARMAX(p,q) model with parameters is thus defined as follows:

$$\begin{aligned}\log(X_t) =& c + \phi_1 \log(X_{t-1}) + \phi_2 \log(X_{t-2}) + \ldots + \phi_p \log(X_{t-p}) \\ &+ \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \\ &+ \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_k Y_{t-k} + \varepsilon_t\end{aligned}$$

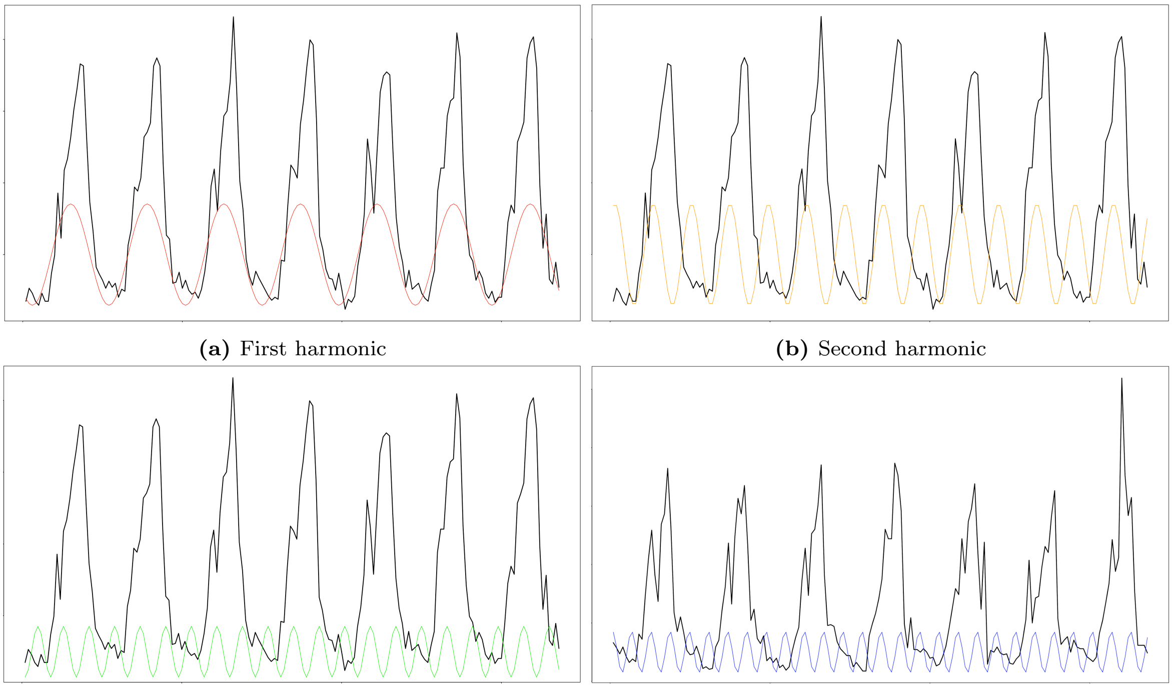
where X_t is the value of the time series at time t, c is a constant, $\phi_1, \phi_2, \ldots, \phi_p$ are the autoregressive coefficients, ε_t is the error term at time t, $\theta_1, \theta_2, \ldots, \theta_q$ are the coefficients of the past error terms, $Y_{t-1}, Y_{t-2}, \ldots, Y_{t-k}$ are the values of the regressor variables at previous times, and $\beta_1, \beta_2, \ldots, \beta_k$ are the coefficients of the regressor variables.

Cyclic Components

As observed in Figure 3, the daily consumption pattern (Fig. 6) can be modeled by four cyclic components [5], each describing a harmonic wave. One way to express these waves in an ARMA(p,q) model is through trigonometric functions for representing seasonality. Thus, the harmonic waves can be described as:

$$\begin{aligned} H_1(t) &= \sin(\omega t) + \cos(\omega t) \\ H_2(t) &= \sin(2\omega t) + \cos(2\omega t) \\ H_3(t) &= \sin(3\omega t) + \cos(3\omega t) \\ H_4(t) &= \sin(4\omega t) + \cos(4\omega t) \end{aligned}$$

where $\omega = 2\pi/24$.



(a) First harmonic **(b)** Second harmonic **(c)** Third harmonic **(d)** Fourth harmonic

Figure 3: The daily consumption pattern over a week (black line) is shown along with the first four harmonic waves: 24h (H_1), 12h (H_2), 8h (H_3), and 6h (H_4).

The daily water consumption pattern in pigs can be modeled using four harmonics with periods of 24, 12, 8, and 6 hours due to the cyclic and repetitive nature of this behavior.

The 24-hour harmonic captures the general daily variation in water consumption. The 24-hour cycle is the basic period representing a full day, and it is natural to assume that the water consumption of pigs will follow a regular pattern over this period. This harmonic allows modeling the general trend of water consumption during the day, with peaks at specific hours and periods of lower consumption during the night.

However, water consumption in pigs may be influenced by additional factors that follow shorter patterns. For example, pigs often have regular feeding routines with specific times for receiving feed. This can generate variations in water consumption at shorter intervals, such as every 12 hours (period 12), every 8 hours (period 8), or every 6 hours. These additional harmonics allow capturing these periodic variations and improving the model's accuracy.

Considering the three harmonics together (Figure 4) provides a more complete and accurate representation of the daily water consumption pattern in pigs. Each harmonic contributes with its own frequency and amplitude, allowing the model to capture both the general daily variations and the faster, more specific variations. This facilitates the analysis of water consumption at different times of the day and helps to better understand the water needs of pigs under various conditions and circumstances.

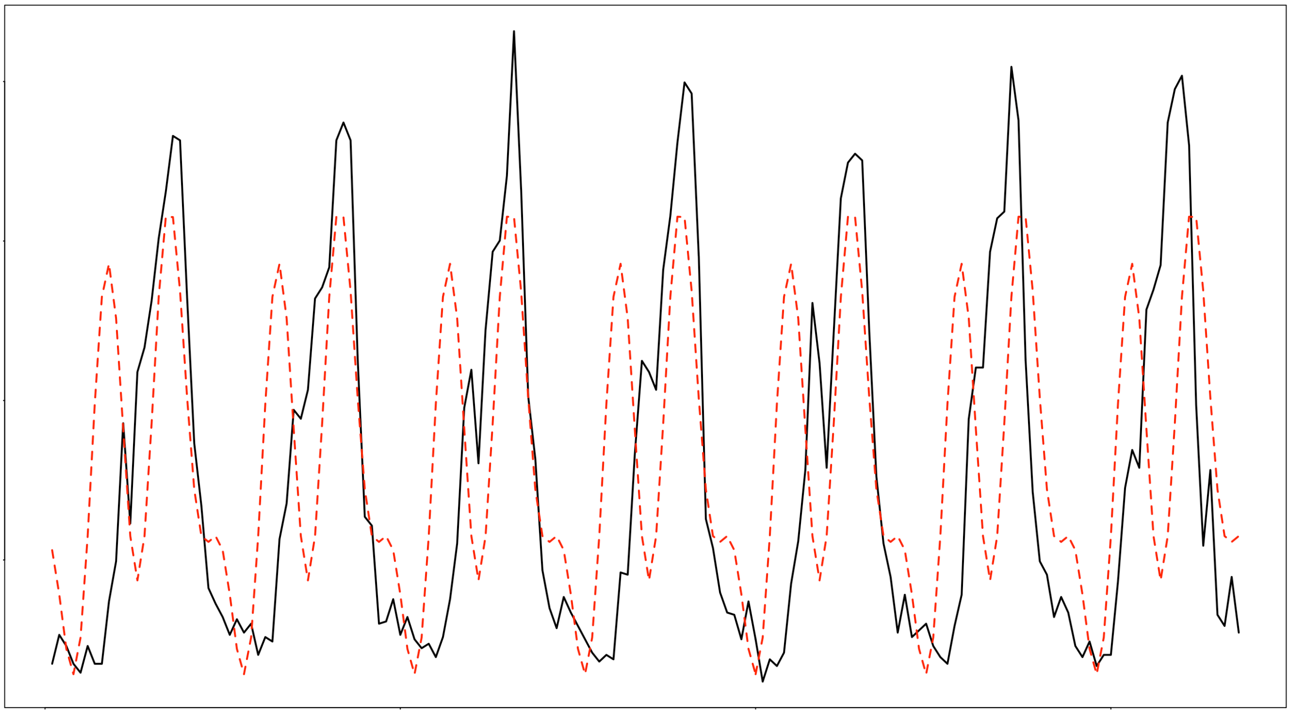


Figure 4: The daily consumption pattern over a week (black line) is shown along with the sum of the four harmonic waves.

3 Results and Discussion

We have thirty time series, the characteristics of which have been described in section 2.1.3.

First, let us recall Figure 2, which showed five consecutive batches from barn 1.01.

It is clearly observed that, within the same batch, water consumption per animal shows an increasing trend and then decreases sharply until a new batch of pigs enters. Thus, we can affirm that water intake increases as the pigs grow.

On the other hand, if we look at the water consumption of any of the batches (Figure 5), in addition to the mentioned increasing trend, a daily oscillation is noticeable, following a pattern that repeats day after day; during the night, consumption is minimal and gradually increases throughout the day until it peaks at around 17-18 hours (6b). Moreover, the daily drinking pattern is relatively stable, as shown in Figures 6a and 6b.

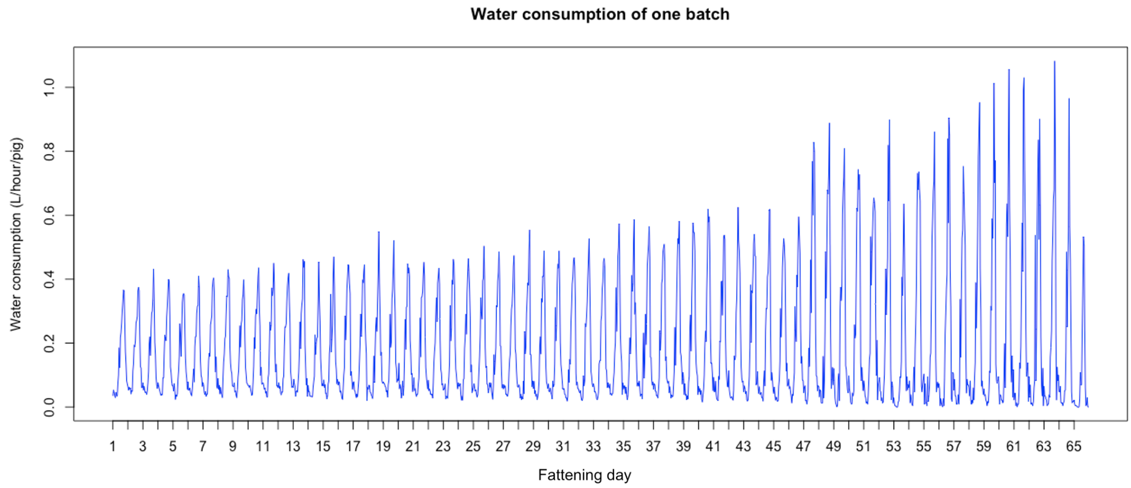
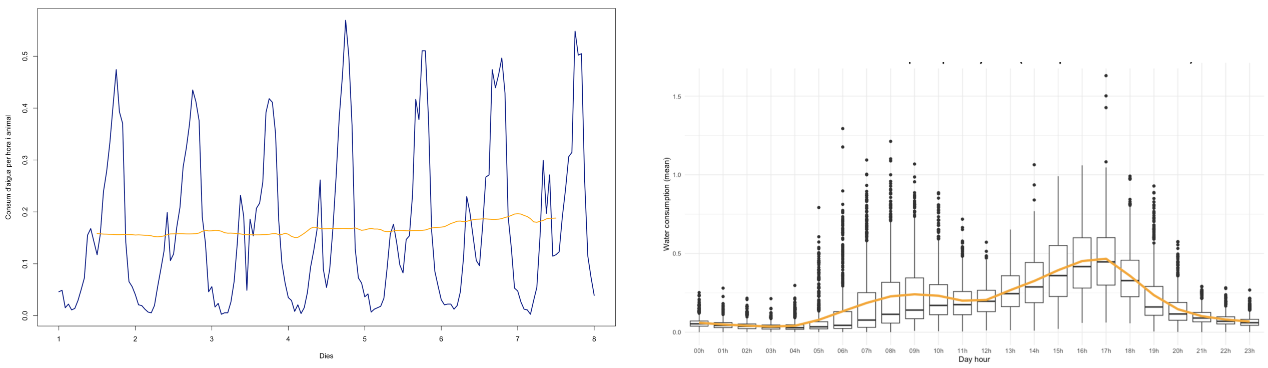


Figure 5: Time series of water consumption per animal and hour for batch 10 of barn 1.01



(a) Time series of average daily water consumption per day of the week considering the 30 analyzed batches.

(b) Boxplot of average daily water consumption considering the 30 analyzed batches.

Figure 6: Water consumption per animal and hour for batch 10 of barn 1.01

The time series for all batches can be found in Section 5.1 of the Appendix.

Thus, after analyzing the characteristics of our data, we proceed to explain, on one hand, which

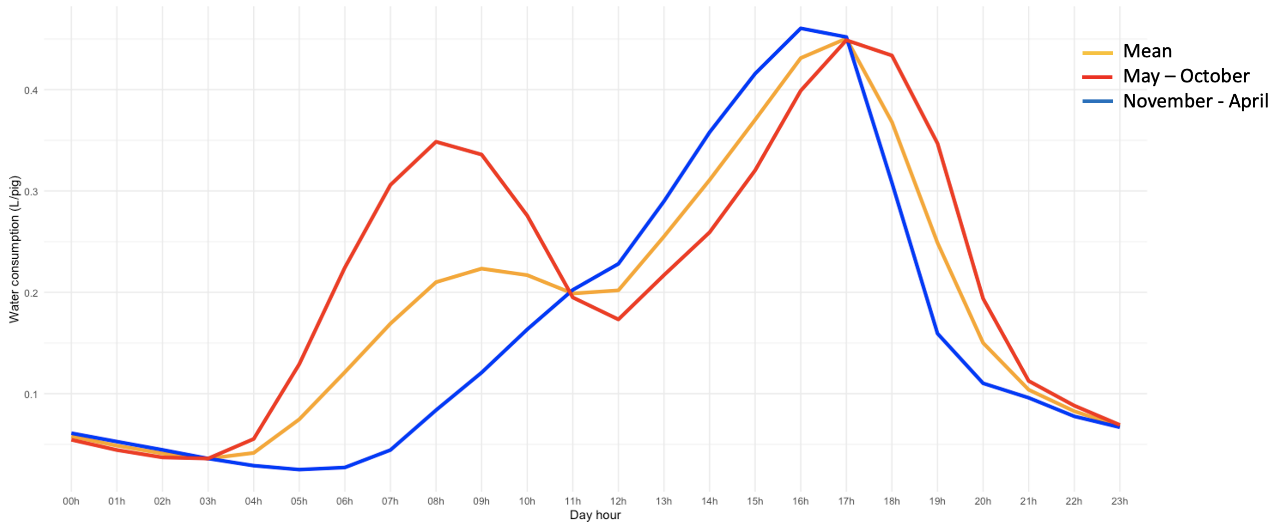


Figure 7: Water consumption per day hour and season (mean per hour of all batches)

indicator parameters have been chosen to include in the model and, subsequently, describe the model applied to the data to achieve the objective of predicting the hourly water consumption of a batch.

For example, in Figure 7 we plot again the daily consumption pattern that we just saw in Figure 6b (you can see it represented in orange), but now we also separate the data into hot and cold months, we can observe that in the hot months (the line in red) the two peaks that we mentioned before are heavier, whereas in cold months (blue line) the first peak barely exists. For this reason, we could think that temperature should be introduced into the model.

Moreover, if we observe Figure 8 we can say that the temperature not only affects the seasonality, but also the growing trend.

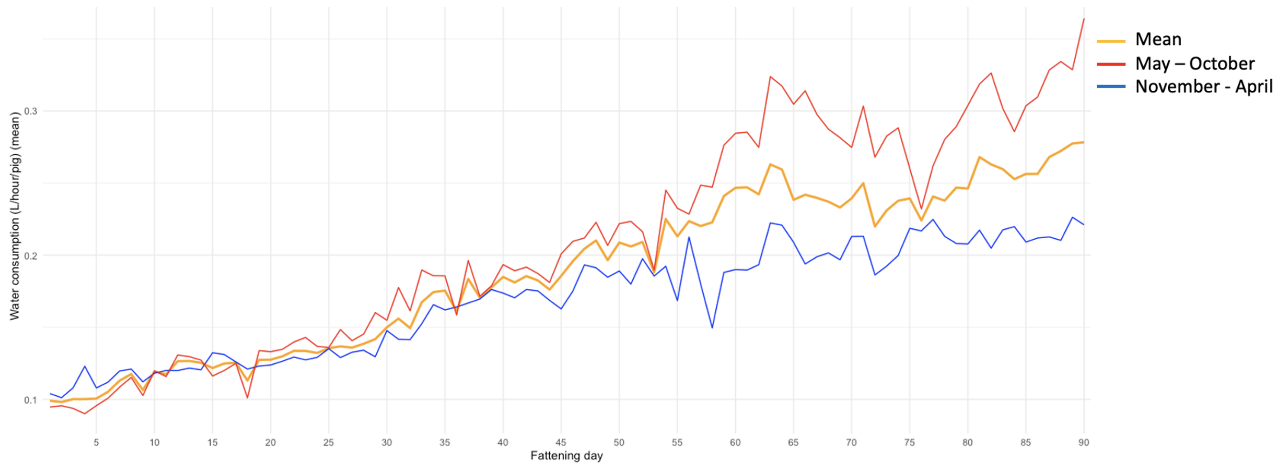


Figure 8: Water consumption per fattening and season (mean per day of all batches)

In the graph in Figure 8, we can see the water consumption trend throughout the fattening period. The yellow line represents the average of water consumption per day and pig of all batches. Whereas, if we split the data again like we did before, the red line represents the trend in hot months and the line in blue in cold months.

It can be seen that also the trend varies according to the temperature, being the trend in hot months more pronounced. So this was another evidence to add temperature to the model.

We proceed to study the possible indicator variables in more detail below.

3.1 Influence of Indicator Parameters

The following explains which parameters were introduced into the model while deciding the orders p, q of the model. This choice was made by selecting a model with the lowest possible AIC while also making sensible predictions. Thus, the decision to include or exclude a regressor variable was made mathematically by analyzing the significance level of the coefficient attributed to each parameter, i.e., if zero appears in its 95% confidence interval, it is considered null and, consequently, will not be included in the model. However, even if a parameter was not entirely significant mathematically, if it made sense from a biological perspective and/or helped improve prediction accuracy, it was still included in the model.

Feed Consumption

Feed consumption can affect the water consumption of growing pigs due to the relationship between feeding and hydration. As we know and have discussed, feed consumption is directly related to the nutritional needs of pigs. When pigs consume feed, they obtain the necessary nutrients for growth and development. The digestion and metabolism of these nutrients generate heat in the pig's body. As a result, pigs may experience an increase in body temperature, which can lead to increased water consumption to regulate their temperature and stay hydrated.

Additionally, feed consumption is associated with saliva production in pigs. Saliva is important for digestion and also helps keep the mouth and throat hydrated. As pigs eat more feed, there is increased saliva production, which may increase the need for water to maintain proper hydration.

Another factor to consider is the composition of the feed. Some feeds may contain ingredients that are drier and require more water for digestion. Therefore, if pigs consume a feed with a higher dry matter content, they may need to increase their water intake to compensate for the hydration loss associated with digesting these ingredients.

Thus, as shown in Figure 9, feed consumption can affect the water consumption of growing pigs due to physiological processes related to digestion, saliva production, and body temperature regulation. This is why Figure 9 resembles Figure 5, which showed the water consumption of the same batch. It is observed that the animals eat in a similar pattern to their drinking, with a growing trend as the pigs gain weight and a daily cyclical component, similar to what happens with water.

Indoor Temperature of the Barn

Figure 10 shows that summer batches of pigs consume slightly more water than winter batches. This difference in water consumption can be explained by various factors related to the exterior and interior temperatures of the barn, which increase during the summer months.

Firstly, the external temperature has a direct impact on the behavior of pigs. During the summer months, when temperatures are higher, pigs may experience thermal stress. Thermal stress leads to a greater need for pigs to regulate their body temperature and stay cool. As a result, pigs may actively seek water to drink and cool down, which is reflected in an increase in water consumption.

It is observed that indoor temperatures are less extreme and more stable than outdoor temperatures. This is because, during the summer months, cooling systems are used in the barn to maintain an appropriate temperature for the pigs. These systems help reduce thermal stress and provide a more comfortable environment for the animals. Similarly, in winter, underfloor heating is applied to mitigate the negative impact of low temperatures.

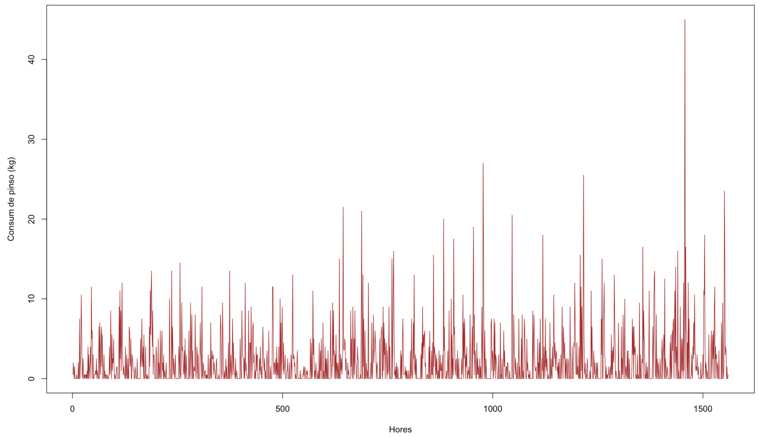


Figure 9: Observations of feed consumption per animal and hour for batch 10 of barn 1.01.

Thus, the increase in water consumption of summer pig batches compared to winter ones is due to the warmer climatic conditions during the summer months. Especially, the indoor temperature of the barn, which is where pigs are mostly exposed, seems to play an important role in this increase, which is why it has been decided to include it in the model.

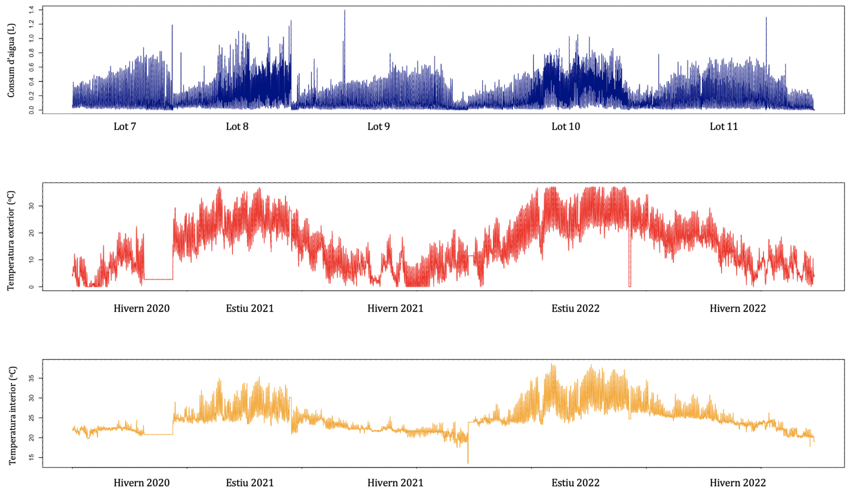


Figure 10: Observations of water consumption per animal and per hour for barn 5.01 with corresponding exterior and interior temperatures per hour, respectively, from top to bottom.

Humidity

As previously mentioned, humidity in the barn can affect the water consumption of pigs because high humidity can cause a greater feeling of heat and discomfort for the animals, as it hinders the dissipation of body heat through sweating. This can lead to an increased demand for water from pigs, as they need to drink more to stay hydrated and cool down. Thus, this parameter has also been included in the model.

Number of Days of the Batch in the Pen

Clearly, the number of days a pig has been in a pen can influence its water consumption. As the pig grows and increases in weight, its water demand also increases. This is because larger pigs require more liquid to meet their physiological needs and stay hydrated. This implies that as the pig approaches the end of the growth cycle and achieves a higher weight, its water consumption is likely to be significantly higher. For this reason, this variable has also been introduced into our model.

3.2 Orders p, q of the ARMA(p,q) Model

We now proceed to explain the model applied to the pig water consumption data. It is important to highlight that the results obtained by applying this model to three different batches from one barn are shown. Thus, we will analyze how suitable the model is and how well it predicts future batches.

The chosen model is an ARMA(24,24) and has the following characteristics:

- The AR part coefficients are all zero except for AR(1), AR(2), AR(23), and AR(24).
- The MA part coefficients are all zero except for MA(1), MA(2), MA(3), MA(6), MA(8), MA(12), MA(13), MA(14), MA(23), and MA(24).
- The included indicator parameters, as explained in the previous Section 3.1, are: feed consumption, indoor temperature of the barn, humidity, and the number of days the batch has been in the pen.
- Additionally, four harmonics with periods of 24, 12, 8, and 6 have been introduced.
- Finally, a variable has been included to account for the water consumption of other batches on the same day of growth and at the same hour.

For batch 7 of barn 1.01, one of the batches modeled, the coefficients obtained are shown in Table 1 located in Annex 5.3. It is important to note that, although some coefficients were not strictly significant at the 95% confidence level, they were considered non-zero if they improved the prediction or, on the other hand, if it was necessary from a veterinary perspective to include them.

Observing the ACF and PACF (Figures 11 and 12) of the time series for batch 7 after applying the chosen model, it is noted that for all lags, the values remain within the 95% confidence interval, showing that the chosen model is appropriate for this batch.

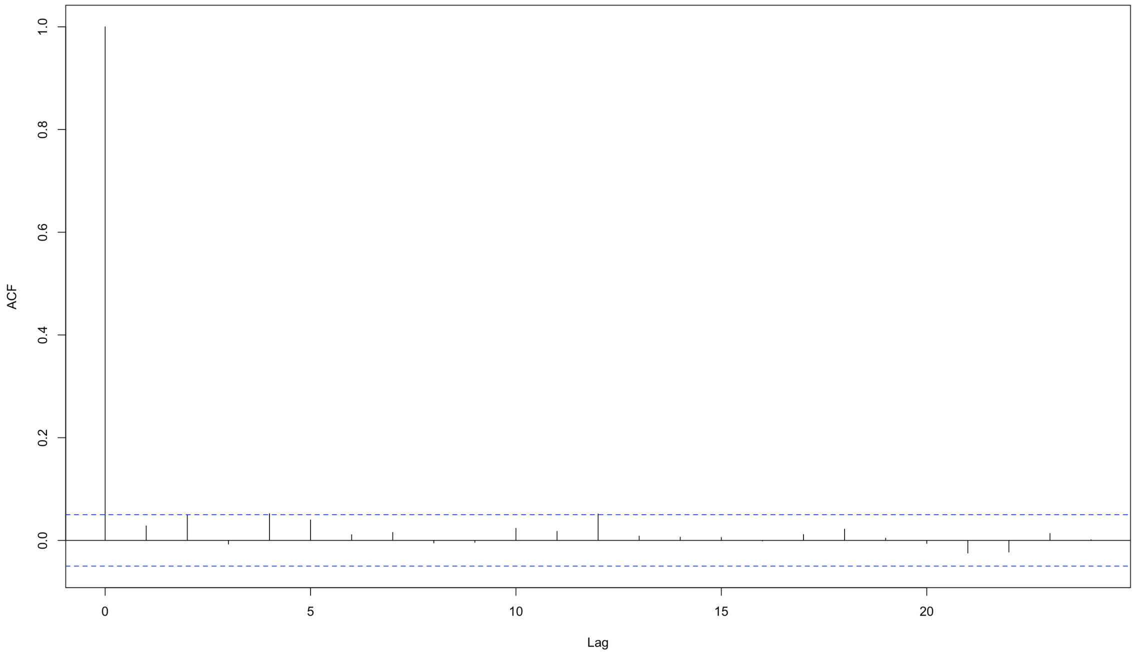


Figure 11: ACF of water consumption for batch 7 of barn 1.01 after applying the model.

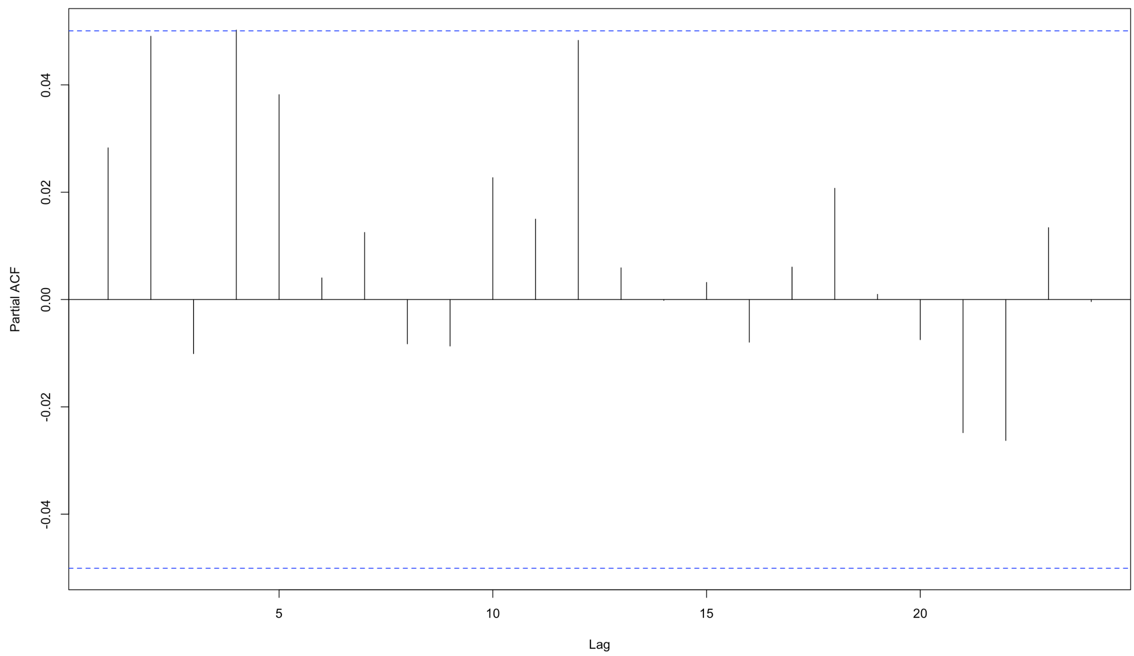


Figure 12: PACF of water consumption for batch 7 of barn 1.01 after applying the model.

To predict the water consumption of a new batch of pigs, in this case, batch 11 from the same barn, the `predict` function in R was not used, as we are not looking to continue the time series of the same batch but rather estimate the consumption per hour and day of a completely new batch. Therefore, a different approach to predicting water consumption was needed. The strategy used was to take the most recent values available from the time series each day (i.e., the last 24 values) and apply the model to obtain the prediction of water consumption for the following day.

It is important to consider that since this is a new batch of pigs with unique characteristics, the prediction of water consumption will be subject to a certain degree of uncertainty. Nevertheless, by using historical data from other batches and our model, a reasonable estimate of water consumption for the new batch 11 is obtained, which meets our goal of predicting the water consumption of an entire batch of pigs (Figure 13).

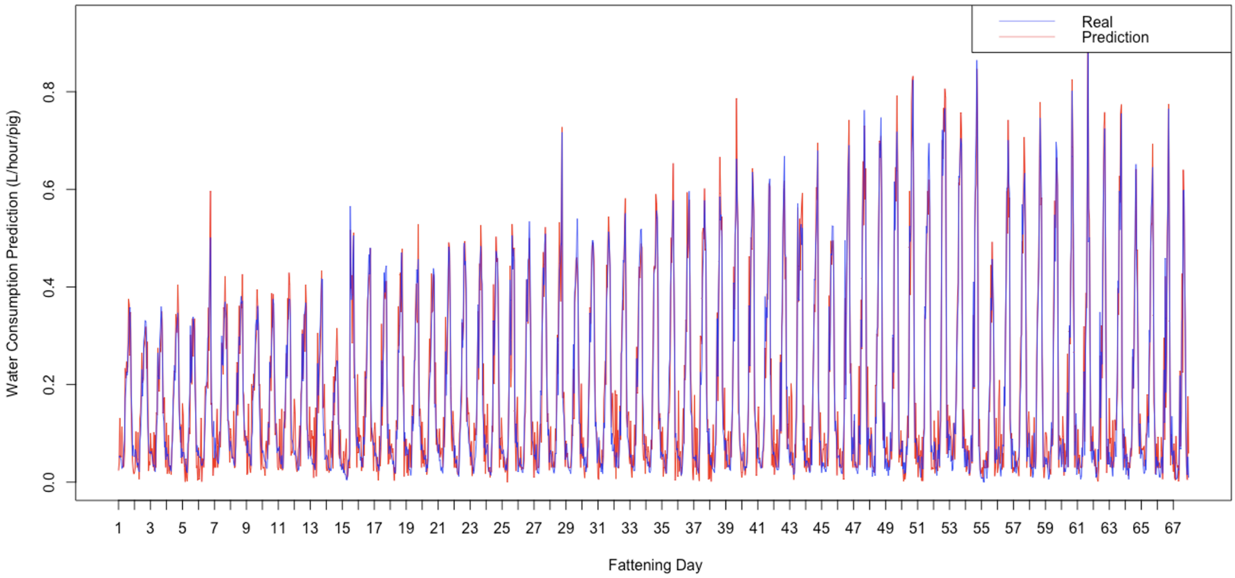


Figure 13: Prediction of hourly and per animal water consumption for batch 11 of barn 1.01 based on historical data of the same barn. Real observations are shown in blue, predictions in red.

Figure 13 shows the water consumption data for batch 11 of barn 1.01 We observe that the prediction is good and all real values fall within the 95% confidence intervals. Moreover, the model predicts very well the increasing trend in water consumption as the pigs grow and also adapts and describes very well the daily water consumption seasonality.

Similarly, the same procedure is carried out for two other barns to predict the hourly and per animal water consumption for batch 11 (see Annex 5.2).

In Figure 14 we can observe the prediction results zoomed to 1 week to appreciate the forecast better, and as you can see we can conclude that our model fits well and it is accurate. Therefore we can conclude that our model fits well and it is accurate.

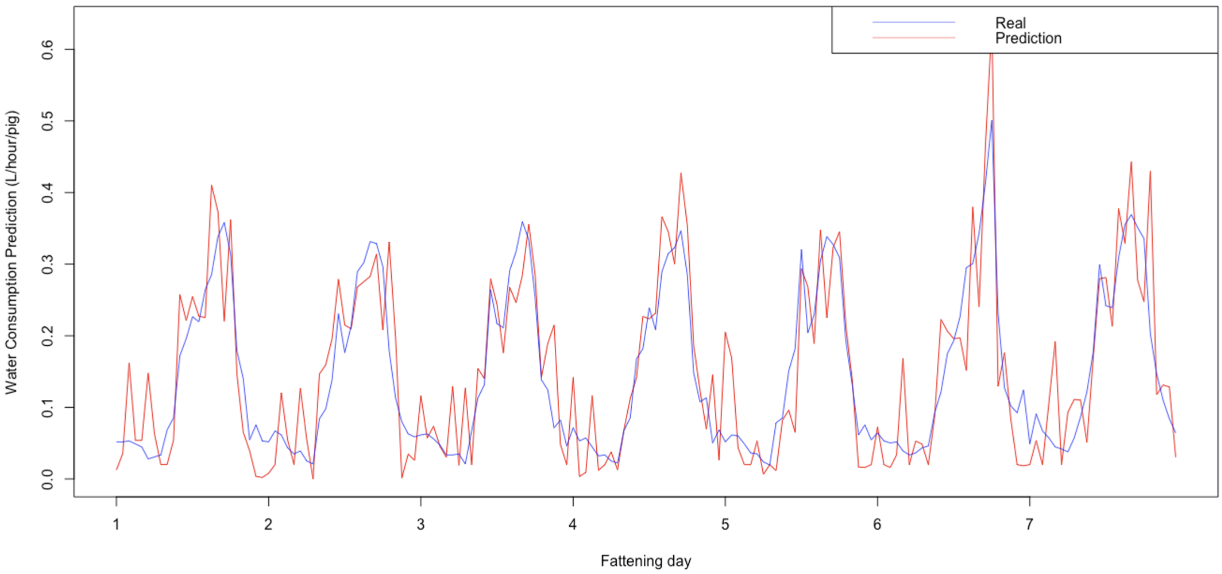


Figure 14: Prediction of hourly and per animal water consumption for batch 11 of barn 1.01 zoomed to 1 week. Real observations are shown in blue, predictions in red.

4 Conclusions

After analyzing the results, we can state that in this work, we have achieved our objective of modeling the water consumption of a complete batch of pigs. To accomplish this, we followed several steps that have proven effective in analyzing and predicting water consumption.

First, we found a model that accurately describes the drinking behavior of pigs and identified the key parameters influencing the water consumption of growing pigs. These parameters include feed consumption, humidity, indoor temperature of the barn, and the number of days of growth. By incorporating these factors into our model, we significantly improved the accuracy of our predictions. It is worth noting that no related study on modeling the drinking patterns of these animals included external parameters as we have done.

Second, we addressed the daily water consumption seasonality by using four harmonics in our model. This allowed us to capture and account for recurring and cyclic patterns in water consumption throughout the day. Moreover, by adding a fourth harmonic, we observed an improvement in the accuracy of our predictions, demonstrating the importance of considering the complexity of seasonal patterns.

Third, we discovered that including a parameter containing the historical water consumption of other batches on the same day and hour of growth also contributes to improving our model. This suggests that there are correlations between the water consumption of different batches at specific times, and leveraging this historical information can help us obtain better predictions.

Finally, we achieved the initial goal of modeling the water consumption of a complete batch using historical data from other batches in the same barn. Our approach has proven satisfactory, as all real water consumption values fell within the confidence intervals calculated using the Delta method. This supports the validity and effectiveness of our model in predicting water consumption in pig batches.

In summary, this work has contributed to advancing the understanding and prediction of water consumption in pig farming. We have identified influential factors, incorporated seasonal patterns, considered historical correlations, and obtained satisfactory results. These findings have significant practical implications for water supply management in pig production, potentially leading to better planning and optimization of water resources.

However, given that this work is part of a five-year project, there are several future proposals that could further enrich our work. For example, it would be beneficial to test the model on a larger number of batches to evaluate its performance in different contexts. Additionally, we could consider including additional variables, such as water quality or the presence of diseases, which could influence pigs' water consumption.

Another interesting proposal would be to investigate the relationship between water consumption and the health status of pigs. If significant links could be established, water consumption could be used as an early indicator of diseases, allowing for earlier detection and treatment. These data could enrich our model and improve its predictive capability.

In conclusion, the current work has successfully modeled water consumption in pig batches and opens the door to many future proposals that could expand and improve our approach. These improvements would contribute to a more efficient, sustainable production and enhance animal welfare.

References

[1] *Global pork consumption expected to increase 7.2% by 2030.* Pig333, 15 March 2024.

[2] Kai, P., Madsen, T., Pedersen, B., Ruby, V., Boykel, I., Nielsen, N.. *Elektronisk overvågning af produktionen i stalde med alt-ind alt-ud produktion.* Report 420, The National Committee for Pig Production, Danish Bacon and Meat Council. (1999).

[3] Bird, N., Crabtree, H.G.. *Real-time production performance monitoring.* Pig J. **45**, 55–68. (2000).

[4] Bird, N., Crabtree, H., Schofield, C.P.. *Engineering Technologies for Pig Production.* No. 28 in British Society of Animal Science Occational Publication. BSAS, Edinburgh, pp. 105–112 (2001).

[5] Madsen, T.N., Kristensen, A.R. *A model for monitoring the condition of young pigs by their drinking behaviour.* Computers and Electronics in Agriculture **48** 138–154 (2005).

[6] Madsen, T. N. *Tools for monitoring growing pigs.* Dina research Report num 91, The Dina Research School, The Royal Veterinary and Agricultural University. (2001).

[7] Hyun, Y., Ellis, M., McKeith, F.K., Wilson, E.R. *Feed intake pattern of group-housed growing-finishing pigs monitored using a computerized feed intake recording system.* J. Anim. Sci. 75, 1443–1451 (1997).

[8] Slader, R.W., Gregory, A.M.S. *An automatic feeding and weighing system for ad libitum fed pigs.* Computers and Electronics in Agriculture **3** 157–170 (1998).

[9] Nienaber, J.A., McDonald, T.P., Hahn, G.L., Chen, Y.R. *Eating dynamics of growing-finishing swine.* Trans. ASAE **33**, (2011-2018).

[10] Thomas Nejsum Madsen, Søren Andersenb, Anders Ringgaard Kristensena *Modelling the drinking patterns of young pigs using a state space model.* Computers and Electronics in Agriculture **48** 39–62 (2005).

[11] K.N. Dominiak, L.J. Pedersenb, A.R. Kristensen *Spatial modeling of pigs' drinking patterns as an alarm reducing method I. Developing a multivariate dynamic linear model.* Computers and Electronics in Agriculture **161** 79–91 (2019).

[12] K.N. Dominiak, Jeff Hindsborg, A.R. Kristensen *Spatial modeling of pigs' drinking patterns as an alarm reducing method II. Developing a multivariate dynamic linear model.* Computers and Electronics in Agriculture **161** 92–103 (2019).

[13] Anders Ringgaard Kristensen, Erik Jørgensen, Nils Toft. *Herd Management Science. II. Advanced topics.* Academic books (2010).

[14] Schale P, Arkenau EF, Schmitt AO, Dänicke S, Kluess J, Grümpel-Schlüter A. *Estimates of Water Consumption of Growing Pigs under Practical Conditions Based on Climate and Performance Data.* Animals (Basel). 2023 May 5;13(9):1547. doi: 10.3390/ani13091547. PMID: 37174584; PMCID: PMC10177191.

[15] Ingram DL. *Evaporative cooling in the pig.* Nature. 1965 Jul 24;207(995):415-6. doi: 10.1038/207415a0. PMID: 5885859.

[16] Morrison S.R., Mount L.E. *Adaptation of growing pigs to changes in environmental temperature.* Anim. Sci. 1971;13:51–57. doi: 10.1017/S0003356100029421.

[17] Sonoda L.T., Fels M., Oczak M., Vranken E., Ismayilova G., Guarino M., Viazzi S., Bahr C., Berckmans D., Hartung J. *Tail biting in pigs-causes and management intervention strategies to reduce the behavioural disorder.* A review. Berl. und Münch. Tierarztl. Wochenschr. 2013;126:104–112. doi: 10.2376/0005-9366-126-104.

[18] Cryer, J. D., Chan, K. S. *Time Series Analysis with Applications in R.* (2nd ed.) New York Springer. (2008)

[19] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and random processes* (4th edition) pp. 669, ISBN 978-019884759, Oxford University Press (2020) Published online by Cambridge University Press: 12 October 2022.

[20] *Time Series Analysis and Its Applications: With R Examples.* Journal of the American Statistical Association, 102 (479), p. 1079.

[21] Scott H. Holan, Robert Lund, Ginger Davis. *The ARMA alphabet soup: A tour of ARMA model variants.* Statistics Surveys, Statist. Surv. 4(none), 232-274, (2010).

[22] Brockwell, P. J., Davis, R. A. *Introduction to Time Series and Forecasting.* Springer International Publishing.

5 Annex

5.1 Hourly Water Consumption Time Series per Animal from Different Farms

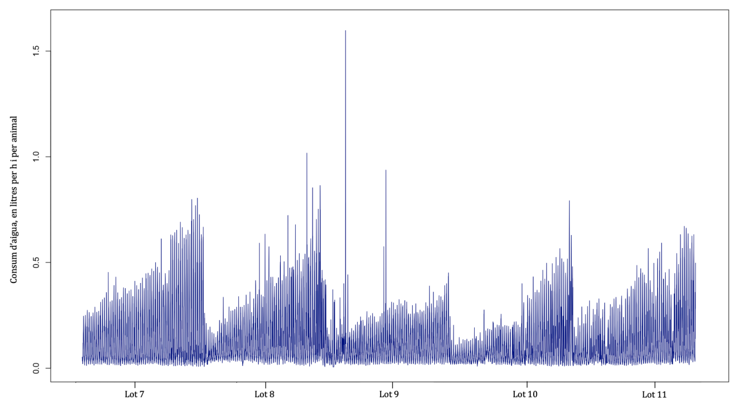


Figure 15: Hourly observations of water consumption per animal for the five batches from Farm 1.0.2.

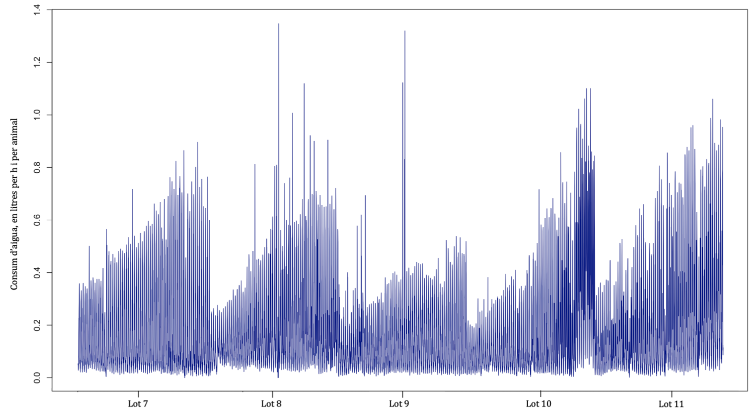


Figure 16: Hourly observations of water consumption per animal for the five batches from Farm 3.0.1.

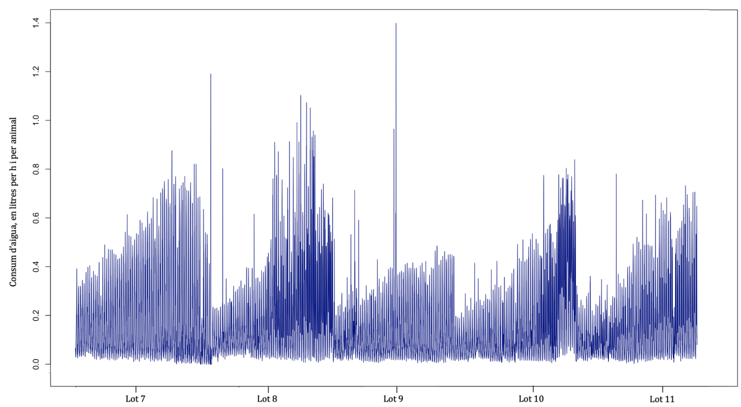


Figure 17: Hourly observations of water consumption per animal for the five batches from Farm 3.0.2.

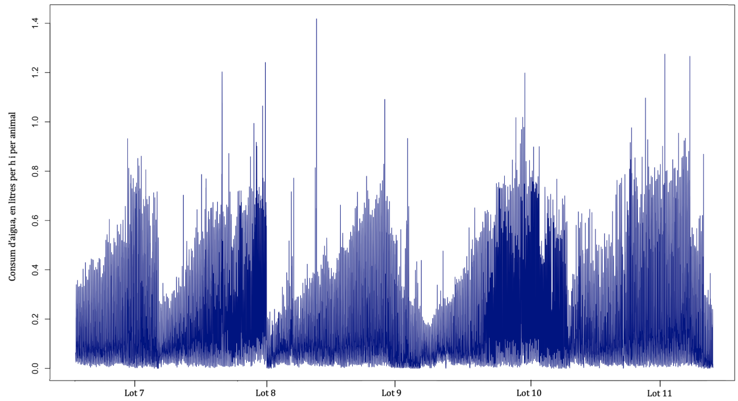


Figure 18: Hourly observations of water consumption per animal for the five batches from Farm 5.0.1.

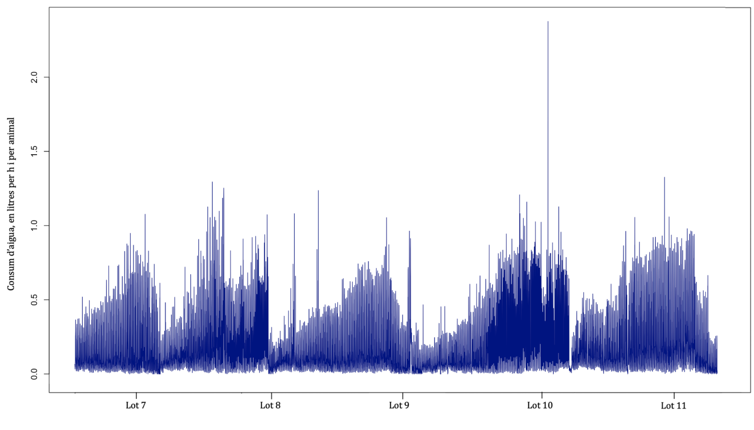


Figure 19: Hourly observations of water consumption per animal for the five batches from Farm 5.0.2.

5.2 Prediction of other batches

As shown in Figures 20 - 21, the predictions for the new batch based on historical data are similar to the originals and also fall within the selected 95% confidence intervals. The coefficients for both models can be found in Tables 2 - 3 located in Annex 5.3.

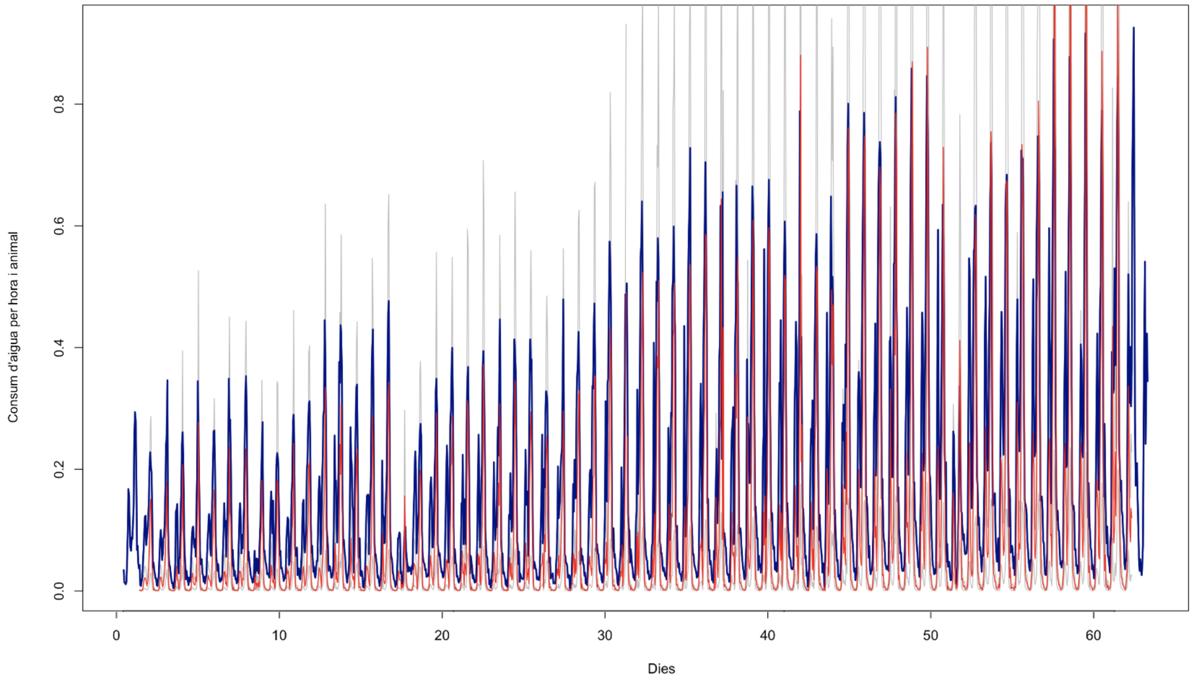


Figure 20: Prediction of hourly and per animal water consumption for batch 11 of barn 1.01 based on historical data from batch 8 of the same barn. Real observations are shown in blue, predictions in red, and 95% confidence intervals calculated using the delta method in gray.

To have a more precise estimate of hourly and per animal water consumption for batch 11, the average of the predictions obtained from batches 7, 8, and 9 has been taken (Figure 22). These predictions fit more accurately than the previous ones to the real water consumption observations, with all of them again falling within the calculated confidence intervals. We interpret that since this forecast comes from the average of predictions based on other batches (historical data), it is expected to adjust better to what might occur in the future, i.e., the actual hourly and per animal water consumption of batch 11.

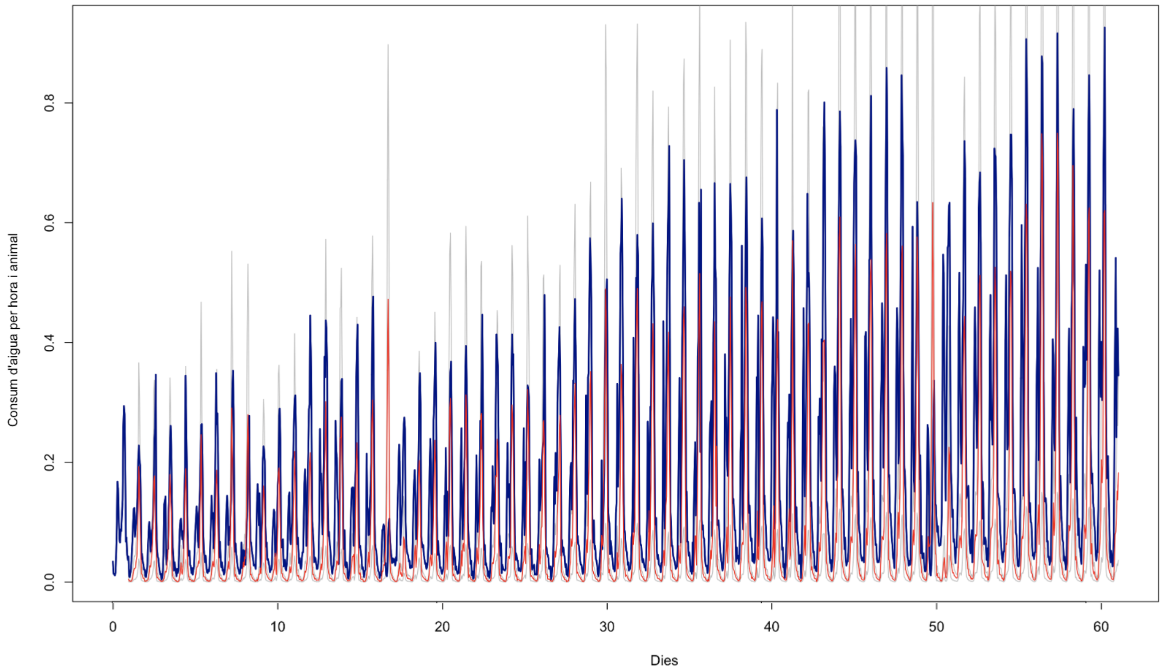


Figure 21: Prediction of hourly and per animal water consumption for batch 11 of barn 1.01 based on historical data from batch 9 of the same barn. Real observations are shown in blue, predictions in red, and 95% confidence intervals calculated using the delta method in gray.

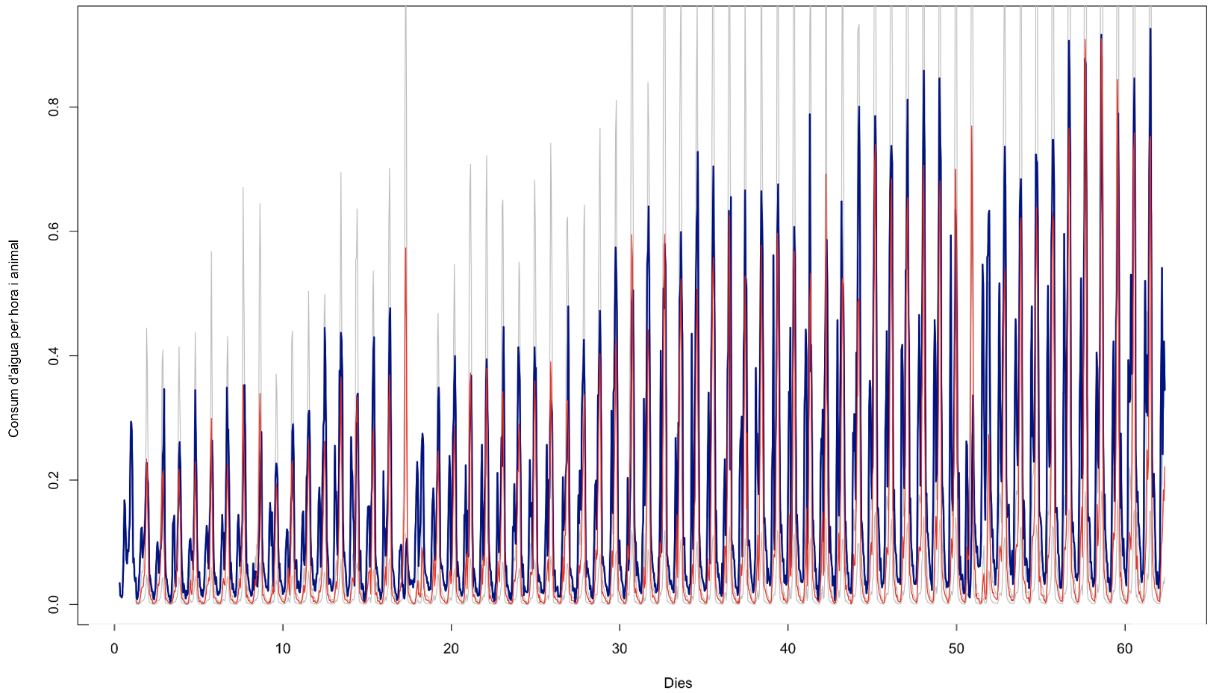


Figure 22: Prediction of hourly and per animal water consumption for batch 11 of barn 1.01 based on the average of the other three predictions. Real observations are shown in blue, the average of predictions in red, and 95% confidence intervals calculated using the delta method in gray.

5.3 Estimates of ARMA Model Coefficients for Batches 7, 8, and 9 from Farm 1.01

	Estimate	Std. Error	z value	Pr(>|z|)
ar1	0.6519	0.0871	7.4835	0.0000
ar2	-0.2524	0.0501	-5.0404	0.0000
ar23	0.2302	0.0397	5.8037	0.0000
ar24	0.3432	0.0776	4.4243	0.0000
ma1	-0.6373	0.0890	-7.1629	0.0000
ma2	0.2743	0.0576	4.7573	0.0000
ma3	0.1536	0.0258	5.9506	0.0000
ma6	0.1061	0.0197	5.3930	0.0000
ma8	0.0530	0.0181	2.9305	0.0017
ma12	0.0361	0.0249	1.4534	0.0731
ma13	-0.0449	0.0296	-1.5178	0.0645
ma14	0.0196	0.0252	0.7794	0.2179
ma23	-0.1557	0.0442	-3.5245	0.0002
ma24	-0.2421	0.0722	-3.3517	0.0004
intercept	-2.1362	0.5962	-3.5831	0.0002
s1	-1.3668	0.1681	-8.1327	0.0000
c1	-0.5833	0.1968	-2.9636	0.0015
s2	0.1574	0.0768	2.0500	0.0202
c2	0.0440	0.0865	0.5083	0.3056
s3	0.2545	0.0335	7.5957	0.0000
c3	-0.1036	0.0361	-2.8673	0.0021
v4	0.0438	0.0217	2.0149	0.0220
v5	0.0097	0.0030	3.2277	0.0006
v6	-0.0074	0.0033	-2.2595	0.0119
v8	-0.0117	0.0053	-2.2221	0.0131
V19	0.1326	0.0443	2.9903	0.0014
s4	-0.0235	0.0172	-1.3659	0.0860
c4	0.0459	0.0169	2.7086	0.0034

Table 1: Estimates of coefficients for modeling Batch 7 from Farm 1.01.

	Estimate	Std. Error	z value	Pr(>|z|)
ar1	1.3778	0.0415	33.1633	0.0000
ar2	-0.6031	0.0233	-25.9181	0.0000
ar23	0.3370	0.0055	61.0070	0.0000
ar24	-0.1902	0.0257	-7.3963	0.0000
ma1	-0.9998	0.0480	-20.8177	0.0000
ma2	0.3510	0.0380	9.2268	0.0000
ma3	0.0850	0.0350	2.4288	0.0076
ma4	0.0789	0.0246	3.2101	0.0007
ma8	0.0661	0.0172	3.8351	0.0001
ma10	-0.0500	0.0153	-3.2682	0.0005
ma18	-0.0363	0.0163	-2.2285	0.0129
ma22	0.1197	0.0264	4.5293	0.0000
ma23	-0.2980	0.0355	-8.3908	0.0000
ma24	0.0572	0.0339	1.6883	0.0457
intercept	-1.9376	0.2039	-9.5008	0.0000
s1	-0.7204	0.0794	-9.0765	0.0000
c1	-0.6163	0.0687	-8.9764	0.0000
s2	-0.3455	0.1866	-1.8517	0.0320
c2	-0.4985	0.2197	-2.2688	0.0116
s3	0.1807	0.0277	6.5352	0.0000
c3	0.1613	0.0261	6.1761	0.0000
v5	0.0025	0.0119	0.2079	0.4177
v6	-0.0057	0.0025	-2.2764	0.0114
v8	0.0133	0.0020	6.5506	0.0000
V21	0.1174	0.0433	2.7133	0.0033
s4	0.1078	0.0119	9.0743	0.0000
c4	0.0203	0.0118	1.7129	0.0434

Table 2: Estimates of coefficients for modeling Batch 8 from Farm 1.01.

	Estimate	Std. Error	z value	Pr(>|z|)
ar2	-0.1083	0.0185	-5.8396	0.0000
ar24	0.8860	0.0184	48.0576	0.0000
ma1	0.1768	0.0176	10.0617	0.0000
ma2	0.2779	0.0312	8.9176	0.0000
ma3	0.1347	0.0187	7.1970	0.0000
ma4	0.1035	0.0180	5.7466	0.0000
ma5	0.0804	0.0158	5.0956	0.0000
ma6	0.0591	0.0137	4.3126	0.0000
ma23	0.0851	0.0143	5.9429	0.0000
ma24	-0.7389	0.0247	-29.9092	0.0000
intercept	-2.0368	0.6060	-3.3609	0.0004
s1	-1.0720	0.0527	-20.3279	0.0000
c1	-0.5403	0.0517	-10.4399	0.0000
s2	0.2252	0.0395	5.7008	0.0000
c2	0.1450	0.0395	3.6718	0.0001
s3	0.2495	0.0256	9.7419	0.0000
c3	-0.0124	0.0257	-0.4839	0.3142
v4	-0.0018	0.0245	-0.0753	0.4700
v5	-0.0050	0.0121	-0.4106	0.3407
v6	-0.0095	0.0031	-3.0299	0.0012
v8	0.0060	0.0025	2.4174	0.0078
s4	-0.1137	0.0227	-4.9993	0.0000
c4	-0.0214	0.0228	-0.9391	0.1738

Table 3: Estimates of coefficients for modeling Batch 9 from Farm 1.01.