Anomaly Detection in Dairy Cows by Fitting Lactation Curve

Authors: Sharath Manjunath[1], Brendan Murphy[1], Donagh Berry[2]

Affiliations: [1] University College Dublin, Co. Dublin, Ireland

[2] VistaMilk SFI Research Centre, Moorepark, Co. Cork, Ireland

A World Leading SFI Research Centre

Science Foundation Ireland For what's next

Anomaly detection refers to the problem of finding patterns in data that do not match the expected behavior. Anomaly detection is widely used in various applications such as risk management, compliance, security, financial monitoring, medical and health risks, AI security, and fault detection. The importance of anomaly detection is that anomalies are often critical and carry actionable information in a wide variety of application domains.

In terms of dairy cows, the seasonal effect and animal activities play an important role in the production of milk. The change in weather, food intake, or sickness of the cow, results in some unusual patterns in the flow of data. Farmers need to detect this unusualness in cows as soon as possible in terms of both milk production efficiency and animal welfare.

INTRODUCTION

Milk production evolves during lactation following a cycle that is similar in all dairy females and usually characterized by two different phases: an **ascending phase** from parturition to **peak production** (the maximum production) and a **downward phase**, from this peak to the dry period. The slope of this phase represents the persistency of lactation.

The **lactation curve** in dairy cows provides a graphical overview of milk production over the different stages of lactation. It is **not possible** to fit the lactation curve using **a linear model** because the curve does not tend to be linear over time. Capturing the different changes in trends is **challenging**. We have implemented the most widely used lactation model and the **Wilmink** model turn out to be the best. Our aim was not just to fit a lactation curve, but also to fit a lactation curve that is **efficient**, **flexible**, and **interpretable** in detecting and analyzing the **anomalies.**

MATERIALS AND METHODS

In order to capture this unusual pattern in data, we have improved the current Wilmink model by adding a **piece-wise regression** to it. A total of **796 pasture-based cows** resulted in **1741 lactations** collected between the years 1998 and 2018, inclusive data are collected. The days in milk between day 6 and day 305, inclusive are considered in this study. Any data beyond **305 days** are excluded as the standard lactation cycle as per the study is always 305 days. The initial **5 days** of milk yield are the **colostrum** milk yield and including them will affect the modeling of the lactation curve as their milk yield will be **very high**.

Wilmink Model: $\text{Yield}_{DIM} = a + be^{-k*DIM} + c*DIM$

Breakpoint Model: $\text{Yield}_{DIM} = a + be^{-k*DIM} + c*DIM + \sum_{j=1}^{J} d_j * (\text{DIM} - \tau_j)_+$

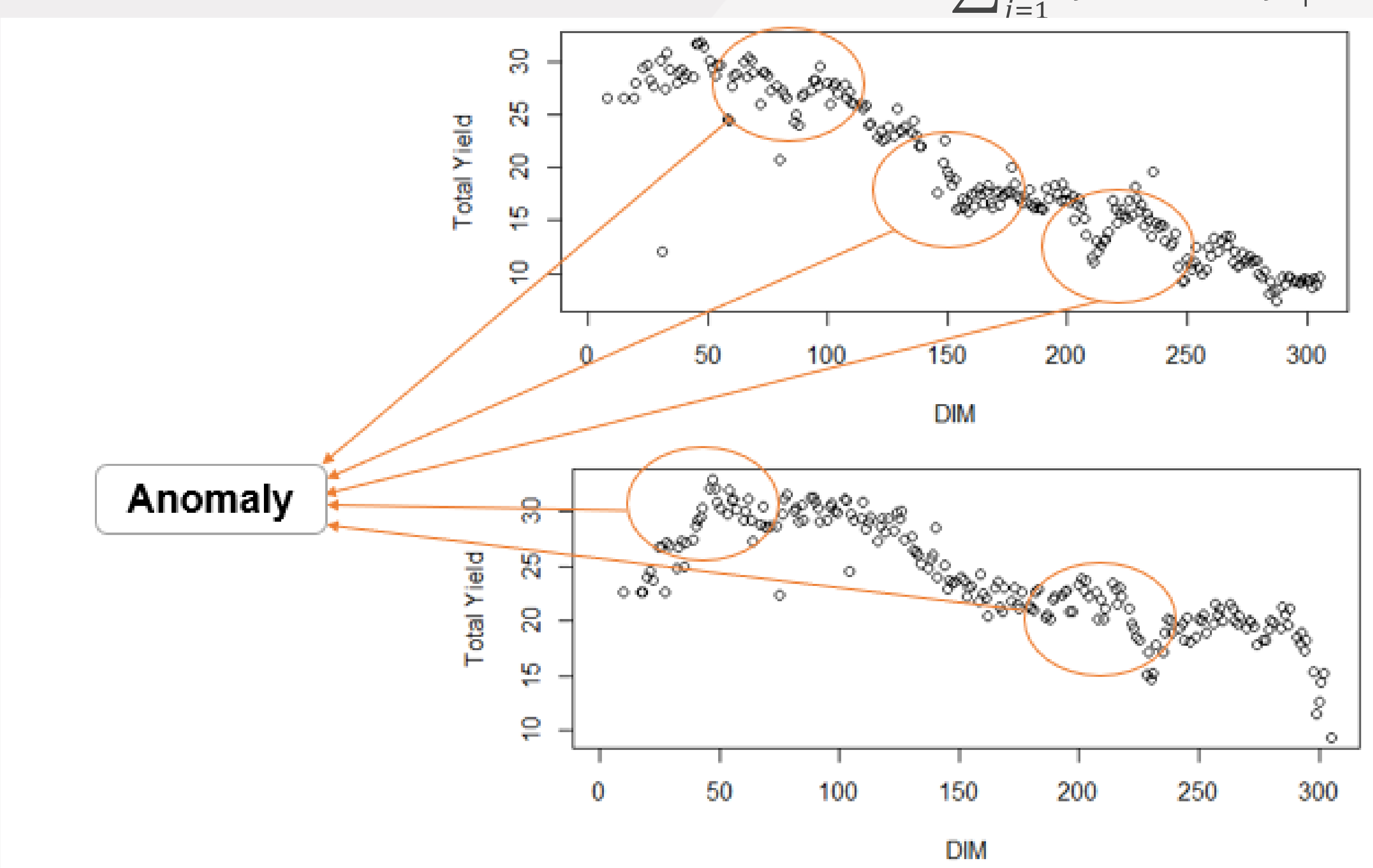
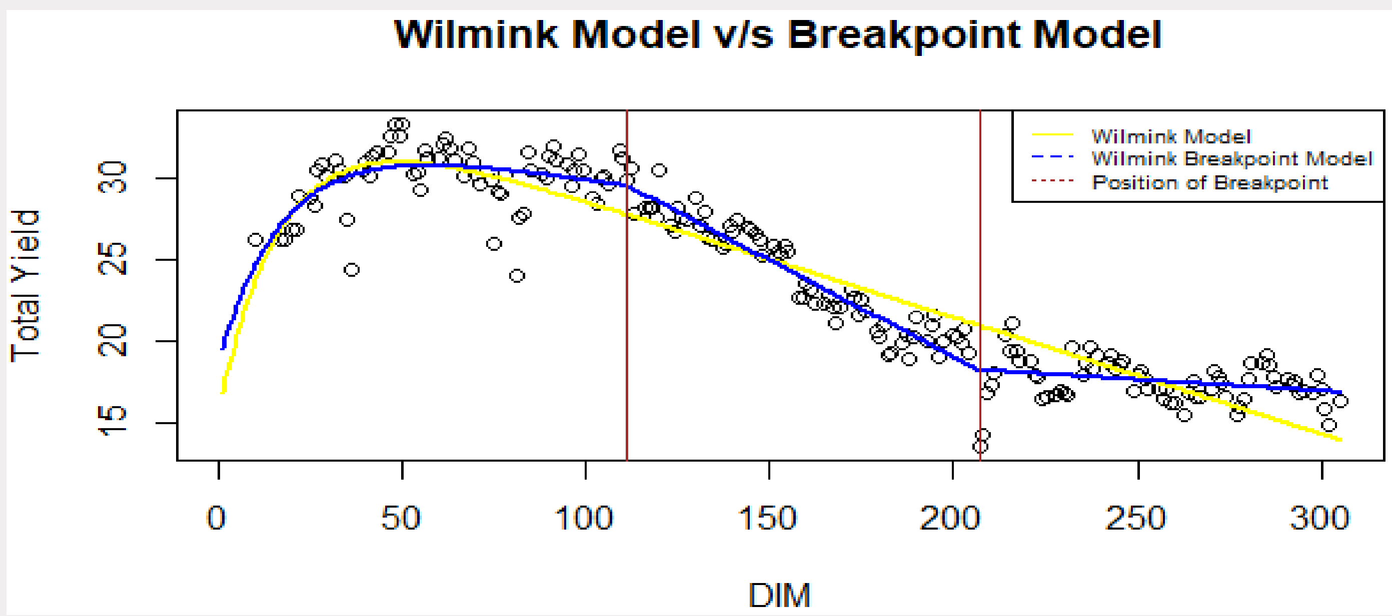


Figure 1., shows an example of the unusual pattern of total milk yield flow over different stages of lactation. These patterns tend to be anomalies.

RESULTS



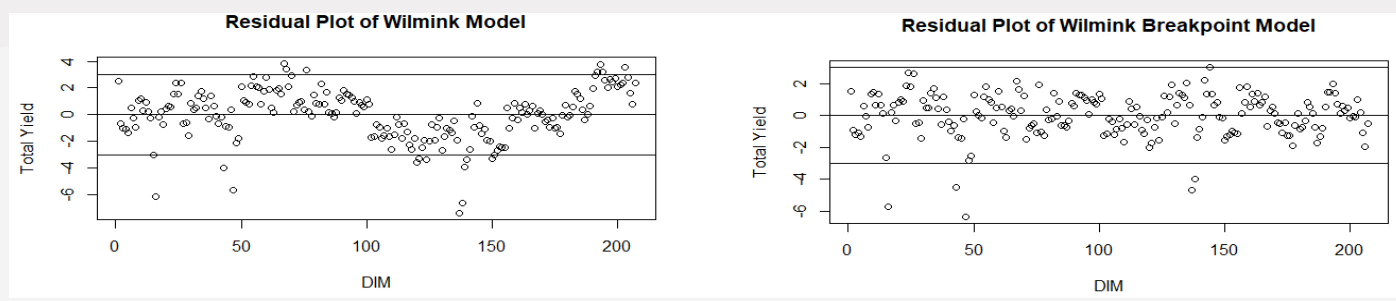


Figure 2., shows the results comparison of fit (a) and residuals (b and c) between Wilmink and the Wilmink Breakpoint Model. The breakpoints in the upgraded model indicate the change in trend. If there is a breakpoint, it indicates the position to study for anomalies. If the breakpoints are too close to each other, then it is an anomaly. If there are more breakpoints, then there are more anomalies.

CONCLUSION

The mathematical functions presented in this study to explain the milk yield during lactation can depict different shapes that may occur when specific patterns are fitted. Our study revealed that the state-of-art Wilmink model could not follow the different trends in the lactation, but the improvised version of the Wilmink model which includes the functionality of the segmented or piece-wise regression outperformed in capturing those trends, especially in the later part of the lactation. The comparison of the residual plot clearly gave evidence by highlighting the underperformance of the existing model which overestimated and underestimated the milk yield in the first and second halves of the lactations, respectively. Also, this updated model of Wilmink is giving an efficient, flexible, and interpretable insight in understanding and detecting the anomalies across the lactations.

HOST INSTITUTION

teagasc Agriculture and Food Development Authority

PARTNER INSTITUTIONS

DCU icbf Insight Tyndall UCD UCC Walton SETU

FUNDED BY:

Science Foundation Ireland For what's next

VistaMilk

Digitalising Dairy