Bayesian Case-Exclusion and Personalized Explanations for Sustainable Dairy Farming (Extended Abstract)*

Eoin M. Kenny1,2,4†, Elodie Ruelle3,4, Anne Geoghegan3, Laurence Shalloo3,4, Micheál O’Leary3, Michael O’Donovan3, Mohammed Temraz1,2,4 and Mark T. Keane1,2,4

1University College Dublin, Dublin, Ireland
2Insight Centre for Data Analytics, UCD, Dublin, Ireland
3Teagasc, Moorepark, Co. Cork, Ireland
4VistaMilk SFI Research Centre
{eoin.kenny, mohammed.temraz, mark.keane}@insight-centre.org, {elodie.ruelle, anne.geoghegan, laurence.shalloo, michael.oleary, michael.odonovan}@teagasc.ie

Abstract

Smart agriculture (SmartAg) has emerged as a rich domain for AI-driven decision support systems (DSS); however, it is often challenged by user-adoption issues. This paper reports a case-based reasoning (CBR) system, PBI-CBR, that predicts grass growth for dairy farmers, combining predictive accuracy and explanations to improve user adoption. PBI-CBR’s novelty lies in the use of Bayesian methods for case-base maintenance in a regression domain. Experiments report the tradeoff between predictive accuracy and explanatory capability for variants of PBI-CBR, and how updating Bayesian priors each year improves performance.

1 Introduction

Although the promise of artificial intelligence (AI) in SmartAg is usually advertised as increasing productivity, in the future it may be more about improving sustainability [Gafsi et al., 2006; Lindblom et al., 2017]. As climate change accelerates, AI’s main future contribution may be more about helping farmers to measure, profile, and predict the outcomes of farm-management decisions in ways that mitigate their environmental impacts. However, this future depends on the development of AI-enabled, decision-support systems (DSS) that are both predictively accurate and explainable to the end user. Here, an existing DSS, called PastureBase Ireland (PBI), is enhanced using case-based reasoning (CBR) techniques, in the so-called PBI-CBR system. This new DSS predicts grass growth for dairy farmers and offers explanations designed to improve user adoption. PBI-CBR’s key technical novelty is its use of Bayesian Case-Exclusion, to “clean up” user-entered data; a technique that by excludes outlier cases from the prediction process using prior beliefs about data distribution(s), thus reducing error and improving explanations. In the remainder of this introduction, the sustainability context for this work is briefly described, before outlining the structure of the paper.

1.1 Context: Agriculture, Sustainability and AI

Concerns about the impact of agriculture on climate change and the development of sustainable models are growing [Lindblom et al., 2017]. The agricultural sector and consumers are faced with opposing views from climate change denial, to proposals that animal agriculture is responsible for 18-51% of greenhouse grass emissions [Steinfeld et al., 2006; Goodland et al., 2009]. However, a recent middle ground has emerged pushing for a quick move to sustainable farming systems [Poux and Aubert, 2018]; the so-called agroecology perspective. For example, in the dairy sector, agroecology proposes a move to pasture-based systems, where animals are predominantly fed on grass outdoors rather than on meal and supplements indoors. This pasture-based system has the potential to be sustainable if there is better grass management (e.g., using grass as a carbon sink). However, these innovations depend on precision technologies, using AI, for monitoring variables such as the climate and grass growth.

1.2 PastureBase Ireland (PBI) & Dairy Farming

SmartAg often depends upon providing new DSSs for farmers to aid them in making complex decisions about how to manage their farms, balancing productivity and sustainability [Lindblom et al., 2017]. These systems need to be predictively accurate, easy to use (and interpretable), and they need to be able to support actionable decision-making in the face of increasing climate disruption. The present work enhances the existing PBI DSS used for grass-fed, pasture-based dairy farming systems in Ireland.

Since 2013, Ireland’s national agricultural research organization, Teagasc, have provided PBI to support Irish dairy farmers (6,000+ users). Among other features, the PBI database has weekly records of grass covers for individual farms (2013-present). A farm’s “grass cover” is the amount of grass available on that farm for cows to eat. PBI-CBR uses historical grass-cover data to predict grass growth rates on a farm from one week to the next. These calculations allow farmers to budget feed for their herds, to determine if there

†Contact Author.
is sufficient (inexpensive) grass available or whether (expensive) meal needs to be bought, meal that incurs additional carbon costs as it tends to be imported (e.g., from the USA).

We used the PBI dataset recorded from thousands of private farms in Ireland between 2013-2017. The primary feature of concern is the average grass growth rate for a farm since the last grass cover recorded, but location features (Farm ID-anonymized and County) are also important for explanation purposes. To explain its predictions PBI-CBR aims to provide explanatory cases from the same farm, or failing that, a case from a nearby farm in the same county.

1.3 Outline of Paper
Section 2 discusses noise in the PBI dataset and how a Bayesian approach is both useful and intuitive for the problem. Section 3 describes Expt. 1 which compares four systems on accuracy and explanatory success. Section 4 describes Expt. 2 which shows how updating priors using Bayesian techniques can improve prediction accuracy and may well help to deal with disruptive climate events. Section 5 reviews related work before making some final conclusions.

2 Noise: A Tale of Two Datasets

The grass-growth domain faces many of the typical problems that arise in SmartAg; notably, that the data is noisy, in part, because it has been entered by end-users (who are, often, non-technical). The PBI dataset has growth-data entered by farmer end-users, data that is known to contain errors, miss-recordings, and subjective estimates. For example, some grass-growth recordings are based on physical measurements with specialized devices (i.e., plate meters), whereas others are based on visual inspection. This noise in the dataset is dealt with by removing outlier cases, using a novel method, called Bayesian case-exclusion; this method uses a separate gold-standard dataset, gathered under controlled conditions (which is noise-free), to “clean up” the farmer-entered data, to create what we call the working-farm dataset.

2.1 The Gold Standard Dataset

The gold-standard dataset of grass-growth measurements we used, covers 28 years of carefully-controlled, weekly measurements in which samples, taken by researchers from the same pasture, were cut, dried, and weighted on a weekly basis at the Teagasc Moorepark Dairy Research Centre, Fermoy, Co. Cork. These measurements are idealized, but very accurate and can thus serve as a good benchmark for determining outlier cases in the PBI dataset, which we exclude using our Bayesian case-exclusion method.

2.2 PBI Dataset

The PBI dataset, used to construct the working-farm case-base, came from the weekly grass-covers entered by farmers in PBI; these grass-covers are farmer’s estimates of grass available on a given farm on a given day and were used to make the grass-growth predictions for one week ahead. Some of these records are known to be in-error; for example, often multiple entries are made on the same day, where the last entry of the day was the intended record. For the years 2013-2017, this dataset had 99,087 grass-cover records, that reduced to 92,635 when same-day entries were removed.

Case generation. Let a farm’s data be \( f = \{x_1, x_2, \ldots, x_n\} \), where \( x_i \) is a grass cover for a day, and \( n \) the total number of covers (in chronological order). The features of \( x_i \) used to generate a case \( (C_i) \) are the average growth rate since the previous grass cover \( (gr) \), the week \( (wk) \), month \( (mth) \), and season \( (seas) \) in which the grass cover was recorded. Weather data \( (w_i) \) at the county level was scraped from Met Éireann, and added as an average from \( x_i \) until \( x_{i+1} \). The weather information in \( w_i \) is the maximum daily temperature \( (maxt) \), the average soil temperature 10cm below the surface \( (soilt) \) on a given day, and the average global radiation \( (grad) \) on a given day. Finally, \( gr \) from \( x_{i+1} \) was added to \( C_i \) as the target feature. Thus, a case is represented as:

\[
C_i(x_i, w_i, x_{i+1}) = \langle x_i(gr, wk, mth, seas), w_i(maxt, soilt, grad), x_{i+1}(gr) \rangle
\]

Case base construction. Taking the raw-data, the cases as defined in Equation 1 were constructed; however, given that the system has to predict one week ahead, only those cases where the target \( x_{i+1}(gr) \) was recorded 5-9 days after \( x_i \) were included in the case base. Also, cases from January and December were excluded (as they tend to show zero growth). Finally, only those cases with accurate historical weather information until the next grass cover were considered (weather is crucial for predicting grass growth). These steps resulted in a working-farm case-base of 20,760 cases.

2.3 The Current Experiments

In the remainder of this section, two experiments are reported that test several variants of the Bayesian case-exclusion idea. In Expt. 1, we examine what happens in this predictive CBR-system when cases are not excluded (Control), versus experimental systems in which they are. In Expt. 2, we explore
adaptive Bayesian case-exclusion, and consider the data as a

time series where priors derived from the gold-standard dis-

tribution are updated year-on-year.

3 Experiment 1: Bayesian Case-Exclusion

In this domain an error of \( \leq 10 \) kg DM/ha is sufficient, but the

major problem is uncertainty in the working-farm case-base,

hence we use Bayesian case-exclusion to exclude likely out-

erlier cases when making predictions. PBI-CBR also explains

predictions by referencing nearest neighbouring cases from

the same farm or county used in the prediction. The four

variants of the system tested were:

- **Control.** A basic system that uses all the cases in the

  working-farm case base.

- **Exclude-2sd.** A Bayesian system that excludes cases

  two-standard deviations away from the weekly, mean

  growth-rates of the gold-standard dataset (see Fig. 1).

  The rationale being that grass growth in a given week ap-

  proximates a normal distribution. Formally, the data for

  growth rate \((GR)\) in a given week across all years in the

  gold-standard dataset approximates \( GR \sim N(\mu, \sigma^2) \),

  where \( N \) is a normal distribution with parameters \( \mu \) and

  \( \sigma \) for the mean and standard deviation, respectively. All

  cases outside \( \mu \pm 2\sigma \) are excluded. This step reduces the

  working-farm case-base by 42% (N=12,042 cases).

- **Exclude-3sd.** This is identical to the Exclude-2sd system

  but \( \mu \pm 3\sigma \) is used to exclude cases. This reduces the

  working-farm case-base by 21% (N=16,443 cases).

- **Transform-3sd.** This is a Bayesian system that trans-

  forms the growth-rates of cases using the gold-standard

  distribution. That is, the distribution of growth in a given

  week from the gold-standard dataset \([GR \sim N(\mu, \sigma^2)]\)

  is transformed to transform the growth-rate values of cases

  for the same week in the working-farm case-base, to fit to

  the parameters \( \mu \) and \( \sigma^2 \). Formally, to transform the

  growth-rate \((gr)\) in a grass cover \( x \) in any given week

  of the year:

  \[
  y_{gr} = (x_{gr} - \mu) \frac{\sigma_p}{\sigma} + \mu_p
  \]

  where \( x_{gr} \) is the growth rate in grass cover \( x \), \( y_{gr} \) is the

  transformed growth rate of \( x_{gr} \), \( \mu \) and \( \sigma \) are the mean

  and standard deviation for the overall growth rate in that

  week in the working-farm case-base, respectively, and

  \( \mu_p \) and \( \sigma_p \) are the mean and standard deviation for the

  overall growth rate in that week in the gold-standard

  dataset, respectively. The intuition being that the gold-

  standard dataset is closer to the ground-truth. Note, in

  this system cases that fall outside \( \mu_p \pm 3\sigma_p \) after the

  transformation are still excluded, so the working-farm

  case-base is reduced by 2% (N=20,282 cases).

As we shall see, exclusion methods improve prediction ac-

curacy, with varying levels of explanatory success. The trans-

form system retains as many cases as possible, aiding accu-

racy and explanatory success. Interestingly, there are indica-

tions that the transformed case-base is closer to the ground

truth as the correlation of Pearson’s \( r \) between \( grad \) and \( GR \)

across all cases increases from \( r = 0.53 \) to \( r = 0.66 \) after trans-

formation, reflecting known dependencies between such radi-

ation and grass-growth [Ruelle et al., 2018].

3.1 Method: Procedure and Measures

For each system variant Monte Carlo cross-validation was

used with 30 re-sampling iterations, each time taking 80/20%

data for training and testing, respectively. A standard \( k \)-NN

algorithm was used for case retrieval and prediction. Selected

values of \( k \) ranging from 5-1000 were tested for each sys-

tem variant. For each evaluation of \( k \) in each system, three

measures were taken: (i) the mean absolute error (MAE) (ii)

the %Farm-Explanation-Success (%FES; i.e., the percentage

of times the \( k \)-nearest-neighbours contained a case from

the same farm as the query), and (iii) the %County-Explanation-

Success (%CES; i.e., the percentage of times the \( k \)-nearest-

neighbours contained a case from the same county).

3.2 Results and Discussion

Fig. 2a shows the results the system variants for all values of

\( k \) in three graphs, one for each measure: MAE, %FES, and

%CES. Overall, MAE is worst for the lowest \( k \) with some

improvement in at \( k=20-35 \). Regarding %FES all systems

are similar, though success does change for different values

of \( k \). For all systems %FES is poor for low values of \( k \), but

beyond \( k=50 \) it rises to \( 80\% \); showing that higher values of \( k \)

have enough cases from the same farm to explain the predic-

tions. For all systems, %CES starts high (\( \sim 80\% \)) and rapidly

reaches \( \sim 100\% \).

Overall, the Control system never gets lower than an

MAE of 15. Similarly, from the two exclusion-systems only

---

**Figure 2:** The tradeoff between error and explanation. (a) Expt. 1

shows that as the value for \( k \) approaches 1000, more explanatory

cases are retrieved, but the MAE for all systems also increases. 

Transform-3sd has the best MAE of \( \sim 8.6 \) kg DM/ha/day at \( k \sim 35 \),

but same-farm explanatory success is low at \( \sim 7\% \); however, at

\( k=1000 \), the tradeoff is balanced, with the MAE still acceptable and

%CES at 85%. (b) Expt. 2 shows MAE is improved for almost every

update-variant, although the improvement in the transform-system is

minimal; explanatory success and MAE are similar to Expt. 1, but

poorer, likely due to less training data. Finally, note the log scale on

the x-axis.

---
Exclude-2sd with \( k=35 \) has the somewhat acceptable MAE of \( \sim 10.01 \) kg DM/ha/day. Overall, the Transform-3sd system is the best with a MAE < 10 kg DM/ha/day for all values of \( k \); here it is capable of delivering county explanations very consistently, but more data would be needed for similar success in farm explanation-retrieval.

4 Experiment 2: Updating Priors

In Expt. 1 Bayesian exclusion/translation of cases resulted in improved performance. However, all of these systems exclude cases using static parameters from the gold-standard dataset. In recent years, climate change is impacting the distributions of grass growth, and Expt. 2 tries to rectify this by updating Bayesian priors year-on-year. So, Expt. 2 has six versions of PBI-CBR, three systems tested in Expt. 1, and three variants of these systems in which priors were updated (Update-Exclude-2sd, Update-Exclude-3sd, Update-Translate-3sd). The procedure is described next.

4.1 Updating Bayesian Priors

To perform the updating, priors are taken from the gold-standard dataset and then progressively used on each year’s data from the PBI-dataset to update them; the idea being to try to track shifts in climate over time. First, take the gold-standard dataset and, binning all its data into weeks, for any given week, let the growth rate (GR) approximate a normal distribution \( GR \sim N(\mu, \sigma^2) \), where \( \mu \) and \( \sigma^2 \) are its mean and variance, respectively. In 2013, all the data for this week was processed into cases (see Section 4.2). Then, we proceed with transformation or exclusion methods on these cases depending on the system variant (using \( \mu \) and \( \sigma^2 \) as in Expt. 1), which gives the new data \( D = \{ C_1, C_2, \ldots C_n \} \), where \( n \) is the number of cases. Take the prior to be \( \mu \sim N(\mu_0, \sigma_0^2) \), where the value \( \sigma_0^2 \) is initially chosen as 4, and \( \mu_0 \) is initially chosen as \( \mu \). Here the value for \( \sigma^2 \) is assumed to remain fixed. Bayes rule shows the posterior (for a given week) is proportional to the likelihood times the prior, in addition, because \( \sigma^2 \) and \( \sigma_0^2 \) are known we can ignore the constant of proportionality and derive that the posterior \( \mu_p \) is:

\[
\mu_p \sim N \left( \frac{\sigma^2}{\sigma^2 + \sigma_0^2 n} \mu_0 + \frac{\sigma_0^2}{\sigma^2 + \sigma_0^2 n} n \bar{x}, \frac{\sigma^2 \sigma_0^2}{\sigma^2 + \sigma_0^2 n} \right)
\]

(3)

where \( \bar{x} \) is the empirical mean of the growth rates in the cases of \( D \), for a full derivation see [Murphy, 2007].

Using Equation 3 we update values for \( \mu_0 \) and \( \sigma_0^2 \), the new value of \( \mu_0 \) was then used to update the original \( \mu \) from the gold-standard dataset, which was used with \( \sigma^2 \) (the fixed variance from the gold-standard dataset) to repeat the whole process in 2014 for the same week. This process is repeated for all weeks of each year until the end of 2016 when all training data was collected. The latest priors in each week were again used to exclude or transform cases in 2017 for evaluation.

4.2 Method: Procedure and Measures

For each system the case base was split in a \( \sim 60/40\% \) ratio for training and testing, respectively; the former being the PBI data from 2013-2016 and the latter 2017. Crucially, these results will be different from identical systems in Expt. 1 because of the different splits (there is also less training data here). For case retrieval, a standard \( k\)-NN was again used with selected values for \( k \) ranging from 5-1000. The three measures used were again MAE, %FES, and %CES.

4.3 Results and Discussion

Fig. 2b shows the results, that generally replicate Expt. 1. Regarding MAE, as before the transformation-versions do better than the exclusion-versions, with \( k=75 \) being optimal for all systems. Expt. 2 shows that Bayesian updating improves all systems at nearly every value of \( k \). Regarding explanation the overall curve-shapes are similar to Expt. 1, with maximum values being %FES=68% and %CES=100%, in contrast to %FES=85.94% and %CES=99.98% in Expt. 1. Acceptable tradeoffs for accuracy and explanation are achieved for both of the transform systems in that at \( k=1000 \) the MAE is \( \sim 9.95 \) kg DM/ha/day with \( \sim 67.5\% \) explanatory-success rate for same-farm cases in both systems.

5 Related Work

Case base maintenance is a notable area of research in CBR [Smitti and Elouedi, 2011]. However, most methods have focused on classification [Hart, 1968; Gates, 1972; Ritter et al., 1975; Guan et al., 2009; Aha et al., 1991; Markovitch and Scott, 1988], as opposed to regression [Redmond and Highley, 2010]. Redmond and Highley [2010], did try to convert Edited Nearest Neighbors [Wilson, 1972] for regression by assigning two hyperparameters, but they acknowledge that applying the classification algorithms to regression is difficult. Our method requires no hyperparameters, though it does require the specification of a prior(s).

XAI within CBR has been shown to be important in intelligent systems [Kenny and Keane, 2019; Keane and Kenny, 2019], with some consideration of CBR recommenders in SmartAg [Cho et al., 2012]. Frameworks have been proposed for explanation in CBR XAI [Sørmo el at., 2005], but there are few CBR-applications in SmartAg (for one exception see Branting et al.’s [2001] work on grasshopper-infestation).

6 Conclusion

We have shown that a CBR system can be used for a DSS in dairy farming to predict grass growth and provide personalized case-based explanations. To deal with noise in the data, we have introduced Bayesian case-exclusion to use prior knowledge to identify and exclude noisy data. Furthermore, we have shown that Bayesian calculations for updating priors year-on-year also improves performance [Kenny et al., 2019]. These systems have the ability to improve the sustainability of dairy farming in the future. Our more recent research [Temraz et al., 2020] has shown that these techniques can continue to deliver accurate predictions in the face of climate change by using the previously-excluded “outlier” cases in later years (such as the hot summer of 2018). Hopefully, though we may experience significant climate shifts, there will always be a case somewhere in the historical record that can provide accurate predictions, and by extension explanations.
References


