ARTICLE IN PRESS

Computers and Electronics in Agriculture xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Multiple linear regression modelling of on-farm direct water and electricity consumption on pasture based dairy farms

P. Shine[a], T. Scully[b], J. Upton[c], M.D. Murphy[a,*]

[a] Department of Process, Energy and Transport Engineering, Cork Institute of Technology, Cork, Ireland
[b] Department of Computing, Cork Institute of Technology, Cork, Ireland
[c] Animal and Grassland Research and Innovation Centre, Teagasc Moorepark Fermoy, Co. Cork, Ireland

ABSTRACT

An analysis into the impact of milk production, stock numbers, infrastructural equipment, managerial procedures and environmental conditions on dairy farm electricity and water consumption using multiple linear regression (MLR) modelling was carried out. Electricity and water consumption data were attained through the utilisation of a remote monitoring system installed on a study sample of 58 pasture-based, Irish commercial dairy farms between 2014 and 2016. In total, 15 and 20 dairy farm variables were analysed on their ability to predict monthly electricity and water consumption, respectively. The subsets of variables that had the greatest prediction accuracy on unseen electricity and water consumption data were selected by applying a univariate variable selection technique, all subsets regression and 10-fold cross validation. Overall, electricity consumption was more accurately predicted than water consumption with relative prediction error values of 26% and 49% for electricity and water, respectively. Milk production and the total number of dairy cows had the largest impact on electricity consumption while milk production, automatic parlour washing and whether winter building troughs were reported to be leaking had the largest impact on water consumption. A standardised regression analysis found that utilising ground water for pre-cooling milk increased electricity consumption by 0.11 standard deviations, while increasing water consumption by 0.06 standard deviations when recycled in an open loop system. Milk production had a large influence on model overprediction with large negative correlations of −0.90 and −0.82 between milk production and mean percentage error for electricity and water prediction, respectively. This suggested that overprediction was inflated when milk production was low and vice versa. Governing bodies, farmers and/or policy makers may use the developed MLR models to calculate the impact of Irish dairy farming on natural resources or as decision support tools to calculate potential impacts of on-farm mitigation practises.

1. Introduction

GDP (Gross Domestic Product) growth within developing countries is fuelling a forecasted 20% increase in global consumption of milk and dairy products by 2050 (Bruinsma and Alexandratos, 2012). In preparation for the abolishment of the European Union milk quota system in April 2015, the Irish government identified the potential for a 50% increase in milk production by 2020 over 2007–09 levels (DAFM, 2010). With dairy products and ingredients valued at €3bn to the Irish economy in 2014, the increased production should be sensitive to the use of natural resources to ensure the sustainable growth of Ireland's dairy industry (DAFM, 2016). The related impact is twofold: (1) much of Ireland's dairy farm water is supplied by groundwater boreholes to safeguard a consistent, reliable supply of adequate pressure (O'Connor and Kean, 2014). Since 7.42 L of water per litre of milk are consumed on average (Shine et al., 2018), the water demand will rise dramatically in line with milk production which may cause local water shortages during periods of little rainfall, thus placing additional pressure on the public water supply. (2) in Ireland, a strong positive correlation exists between milk production and electricity consumption with 38.84 watt-hours (Wh) per litre of milk consumed on average (Shine et al., 2018). Similarly, the electricity consumption of three Finnish dairy farms varied between 37 and 62 Wh kg^{-1} milk, with milk cooling and milk harvesting being the two largest energy consuming processes (Rajaniemi et al., 2017). Without an effective mitigation strategy, dairy farm electricity costs per litre of milk may increase, as dairy farm infrastructure may not be optimally configured for the increased milk production levels. Similarly, increased electricity consumption during daytime or peak hours may have negative effects on national grid loads as well as on dairy farm electricity costs in a dynamic pricing

* Corresponding author.
E-mail address: michaeld.murphy@cit.ie (M.D. Murphy).

https://doi.org/10.1016/j.compag.2018.02.020
Received 7 December 2017; Received in revised form 10 February 2018; Accepted 17 February 2018

environment (Upton et al., 2013).

Milk price volatility and environmental constraints are forcing farmers to produce milk at lower costs with a lower overall environmental footprint. Empirically predicting overall electricity and total dairy farm actual water consumption (leak inclusive) may allow for an efficient and comprehensive means of calculating dairy farm electricity and water (E&W) consumptions as metering equipment and/or large input mathematical models are replaced by a small number of high prediction yielding, empirically derived coefficients, allowing consumption values to be easily calculated. The ability to efficiently predict dairy farm E&W consumption supports the sustainable growth of Ireland's dairy farms, which is of benefit to governing bodies, policy makers, dairy farmers and Irish dairy industry stakeholders.

Current state-of-art electricity prediction on Irish dairy farms is mechanistic in nature (Upton et al., 2014), capable of predicting total CO_2 emissions, electricity costs and consumption, predicting the latter to < 10%. Large scale mechanistic modelling would require an exhaustive data collection process to collect specific data related to plate heat exchanger milk: water ratios, water temperatures, milking times, installed lighting capacity etc. (Upton et al., 2014). On the other hand, an empirical model for dairy farm electricity consumption would replace this large number of input variables with a small number of highly predictive numerical coefficients. This approach was undertaken in Southern Italy, where a MLR model was developed for predicting annual dairy farm related electricity consumption to within 11.4% utilising 285 dairy farms (Todde et al., 2017). Similarly, MLR models were developed to predict the individual and combined electricity consumption of the main electrical components (vacuum pumps, refrigeration compressors, water heaters, and air compressor) of a single farm using 14 years of metered data (Edens et al., 2003). Edens et al. (2003) were capable of explaining 62% of the variability of the combined electricity consumption of the main electrical components. Concurrently, Edens et al. (2003) concluded that quantity as opposed to the quality of milk produced had a far greater impact on electricity consumption with the number of cows milked and ambient temperatures having lesser statistical impacts.

Research regarding dairy farm water prediction has primarily focused on developing linear regression models for daily cow drinking water requirements within confined dairy systems (coefficient of determinations ranging from 0.39 to 0.75) (Cardot et al., 2008; Meyer et al., 2004; Murphy et al., 1983). However, linear regression functions have been developed for predicting annual Irish dairy farm green and blue water volumes to 11.3% and 3.4%, respectively (Murphy et al., 2017). These studies required meticulous data collection including concentrates fed, direct water demand (for blue water prediction), dry matter (DM) feed intake, live weight and sodium intake. In contrast to a meticulous data collection approach, Higham et al. (2016) utilised variables related to milk production, stock, environmental conditions and infrastructural equipment from 23 New Zealand pasture based dairy farms. Using this approach, they developed a partial least square regression model (r = 0.90) for predicting total daily water consumption (drinking plus parlour).

MLR modelling has also been applied within the agricultural domain for forecasting milk production with accuracies ranging from 9% to 26% (Murphy et al., 2014; Sharma et al., 2006; Zhang et al., 2016). MLR models utilising data related to milk production, stock, farm infrastructure, managerial processes and environmental conditions for dairy farm E&W prediction can provide; (1) a decision support tool for dairy farmers and/or policy makers to calculate the impact of potential on-farm mitigation practises. (2) governing bodies with the ability to conduct macro level E&W analysis or water risk assessments as variables may be gathered on a large scale without the requirement of specialised equipment. (3) governing bodies and state agencies such as the Department of Agriculture, Food and the Marine (DAFM) and Bord Bia with the means of calculating the impact of Irish dairy farming on natural resources for sustainability reporting and marketing Irish dairy products abroad (DAFM, 2016).

This work utilised E&W consumption data collected from 58 Irish commercial dairy farms and corresponding data related to milk production, stock, farm infrastructure, managerial processes and environmental conditions. Utilising this data, the objectives of this work were to: (1) develop MLR models for both electricity and water (leak inclusive) consumption from an initial selection of variables, by employing a range of data mining techniques to extract variable subsets which offer the greatest prediction accuracy on unseen consumption data. (2) analyse the impact of the optimum variables on E&W consumption through a standardised regression analysis. (3) analyse the monthly prediction bias of each MLR model to determine factors, which may influence model performance.

2. Materials & methods

In this study, data were acquired via both automated and manual recording of E&W consumption through Teagasc, Moorepark (Cork, Ireland). In total, 58 Irish pasture based dairy farms were monitored throughout the period 1st Jan 2014 – 31st May 2016. Electricity consumption was monitored on 56 farms (autonomous = 55, manual = 1) while total water consumption was monitored on 51 farms (autonomous = 20, manual = 31). Farms that had part of total consumption acquired through manual reporting were classified as manually recorded dairy farms. The dairy farms used in this study had a mean herd size of 116 cows and annual milk production of 621,702 litres in 2015 as described by Shine et al. (2018). Data utilised for MLR model development are summarised in Table 1 below. This data includes milk production, cow numbers, E&W consumption data and related key performance indicators, at a monthly resolution. On average, 51,421 L of milk was produced per month, which resulted in a consumption of 2094 kWh of electricity and 361 m^3 of water on average. Regarding dairy farm infrastructure, 47 (84%) of the dairy farms utilised for developing the electricity MLR model employed a direct expansion bulk

Table 1
Population descriptions for monthly electricity and water consumption, milk production, dairy cows and related key performance indicators.

Variable	Unit	Min	Mean	Median	Max	IQR	SD	SEM
Milk yield	Litre	213	51,421	48,016	204,756	42,407	32,452	863
Dairy cows	n	28	114	102	300	50	41	1
Electricity	kWh	199	2,094	1,818	7,786	1,350	1,094	31
	$Wh\ L_m^{-1}$	8.10	73.19	39.82	3,314.64	31.68	164.97	4.70
	$kWh\ Cow^{-1}$	1.98	18.35	18.14	45.36	7.78	6.02	0.17
Water	m^3	51	361	308	1575	217	218	7
	$L_w\ L_m^{-1}$	1.67	13.42	6.63	542.22	6.82	34.18	1.14
	$m^3\ Cow^{-1}$	0.51	3.29	2.95	23.50	1.83	1.90	0.06

IQR = Inter-quartile range, SD = Standard deviation, SEM = Standard error of the mean.
$Wh\ L_m^{-1}$ = Watt-hours per litre of milk. $kWh\ Cow^{-1}$ = Kilowatt-hours per dairy cow.
$L_w\ L_m^{-1}$ = Litres of water per litre of milk. $m^3\ Cow^{-1}$ = Cubic meter of water per dairy cow.