Information Sciences 378 (2017) 26–47

Contents lists available at ScienceDirect

Information Sciences

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

An intelligent system for livestock disease surveillance

Omolbanin Yazdanbakhsh, Yu Zhou, Scott Dick*

Dept. of Electrical & Computer Engineering, University of Alberta, 9107-116 Street, Edmonton, AB, Canada T6G 2V4

A R T I C L E I N F O

Article history:
Received 4 September 2014
Revised 2 September 2016
Accepted 12 October 2016
Available online 13 October 2016

Keywords::
Livestock disease surveillance
Machine learning
Wavelet analysis
Time-series analysis
Wireless sensors
Intelligent systems

A B S T R A C T

Currently, cattle in feedlots are monitored for disease through a manual, labour-intensive process. Physical checkups are administered approximately once per week, and pen riders are hired to watch over the herd, looking for behaviors that indicate an animal is sick. Contagious diseases thus have considerable freedom to spread before they are first detected, leading to increased morbidity and mortality in the herd. We propose the use of animal-mounted sensors, coupled with an intelligent surveillance system, to automatically and continuously monitor the health of each animal; in essence, this is a case study of designing an intelligent condition-monitoring system, in the form of an inferential sensor. In an empirical trial of the system on an Alberta feedlot, sensor data was used to forecast illness up to seven days in advance. Using an ensemble classifier in the wavelet domain, we obtain a sensitivity of 80.8% and specificity of 80%.

1. Introduction

Epidemic disease – or even just the threat of it – can have a devastating impact on a nation's livestock industry. For instance, in May 2003 a single cow in Canada was discovered to be suffering from Bovine Spongiform Encephalopathy (BSE). BSE is not usually transmitted between live animals; known BSE infections arise from contaminated cattle feed. Nonetheless, all major trade partners immediately banned all trade in Canadian cattle and beef products, leading to over $5 billion CDN in losses (Canada's cattle industry is highly export-oriented) [108]. In another recent example, the United Kingdom suffered an outbreak of Foot and Mouth Disease (FMD) in 2001. FMD is one of the most contagious animal diseases. It can be spread through direct, indirect, and airborne transmission [21]; there are historical records showing that large virus plumes from a farm can remain infectious as far as 300 km from their source. In the 2001 U.K. epidemic, an outbreak of FMD spread to over 120 separate locations before it was detected. Containing and ending the epidemic, along with the associated economic losses due to import bans, cost the U.K. roughly $16 billion USD [12,117]. Endemic disease, meanwhile, is much more significant in a producer's day-to-day operations. In North America, the Bovine Respiratory Disease Complex (BRDC) was estimated to cost $800–900 million USD per year in the United States alone in 2010 [107].

The most efficient way of managing livestock disease is to detect and treat it early on, before it can spread to other animals in a herd. Modern intensive farming practices (in which livestock are confined to high-density pens to maximize weight gain) provide an ideal environment for the spread of diseases like BRDC, meaning that early detection and prompt treatment become even more vital. Sick animals must be removed from the regular pen, so that they themselves can be treated, and to break the chain of contagion. At the present time, the state-of-practice in livestock disease surveillance consists of administering regular checkups (once every 1 or 2 weeks), and hiring pen riders (literally, men on horseback

* Corresponding author. Fax: +780 492 1811.
E-mail addresses: yazdanba@ualberta.ca (O. Yazdanbakhsh), yzhou4@ualberta.ca (Y. Zhou), dick@ece.ualberta.ca (S. Dick).

http://dx.doi.org/10.1016/j.ins.2016.10.026
0020-0255/

within the animal pens) to look for any disease symptoms between each checkup. Plainly, this manual system has potential weaknesses; FMD for instance rapidly spreads among animals with an incubation period of 2–12 days [1]. If pen riders miss the behavioral signs of the disease (which may not appear immediately), it can spread through the pen and even through other pens located far away before the next checkup. That single mistake can lead to a disaster for the country. More prosaically, BRDC is commonly associated with cattle being sold and shipped to feedlots (the stress apparently weakens their immune systems). Thus, there is a need for automated approaches to detect the onset of disease in between checkups. One possibility is using sensor technology. The Alberta Research Council (ARC), in collaboration with two industrial partners and ourselves, has developed an intelligent sensor system for livestock disease surveillance in cattle pens. This system consists of (1) an animal-mounted sensor platform that collects animal and ambient temperature, and detects watering and feeding events; and (2) a machine-learning system that predicts the onset of disease in these cattle up to 7 days in advance. This system is believed to represent the state-of-the-art in the livestock disease surveillance field. ARC has completed a trial deployment of 31 sensor devices for four months on an Alberta feedlot, and supplied this data to the authors. To our knowledge, no comparable dataset currently exists, and so this is a crucial resource for exploring the limits of sensor-based disease surveillance. Furthermore, while this dataset focuses only on cattle in feedlots, there is no intrinsic reason why an animal-mounted sensor package could not be used on free-range farms.

Our goal in this article is to select and parameterize a classifier to serve as the machine-learning component above, and determine how well it will be able to forecast livestock disease events. We will apply a variety of machine-learning algorithms to this problem, including neural networks, support vector machines, and decision trees. We will limit our analysis to the temperature sensors, as these are believed to be the most sensitive indicators of illness. We will explore multiple feature sets in both the time and wavelet domains, and evaluate our classifiers on the resulting datasets. In an earlier work [131], the instantaneous animal and ambient temperature data were used to predict whether or not a cow requires medical intervention within the next 7 days. However, the results showed that simply fusing the instantaneous measurements did not work well. In this paper, we derive the instantaneous difference between animal and ambient temperatures; and we detrend and de-season the data to deal with an apparent concept drift. We further transform this data into the wavelet domain, using wavelet coefficients as the new features. We then train classifiers on all of these different feature sets, and compare the sensitivity and specificity of each approach. We finally determined that a hierarchical classifier trained on wavelet-domain features worked best, yielding a sensitivity of 80.8% and specificity of 80%.

Our work offers two major contributions to the intelligent systems literature. Firstly, we develop an intelligent system to solve the specific, economically important problem of livestock disease surveillance. As with other application-oriented papers (e.g. [2,5,52,130,134]), this means that the resulting system is heavily tailored to the specific characteristics of the problem. Secondly, our design process and final results can reasonably be viewed both as a case study of the larger problem of *condition monitoring* via intelligent systems; and as a case study of the development of an *inferential sensor*. Both of these more general problems are high-value, well-known intelligent-systems tasks.

The remainder of this paper is organized as follows. Section 2 reviews related work and essential background in livestock disease surveillance, wavelet decompositions, and time-series analysis, as well as condition monitoring and inferential sensing. We describe our methodology in Section 3, and present our experimental results in Section 4. We discuss threats to validity in Section 5, and offer a summary and discussion of future work in Section 6.

2. Related work

2.1. Livestock disease management

Any livestock operation is vulnerable to disease, and one can reasonably expect that there will be occasional outbreaks. Within the North American system of intensive farming, BRDC represents a particular risk. Beef cattle in North America will be born on a traditional, extensive farm, and raised there until weaning. Then they will usually be sold to feedlot operations that specialize in maximizing their weight gain, until they are ready to be sold for slaughter. It is fairly common for a cow to pass through multiple feedlot operations, with each transfer exposing them to a different cattle population, having different microbial agents in circulation. Between feedlots, they may also be housed for a time in a sale barn at an auction site, alongside cattle from a wide variety of origins. Separation from their dams, transportation in a vehicle with limited fresh air, mobility, and water, introduction to a new herd, castration and dehorning procedures, etc. are all stressful events for young cattle, and appear to play a role in the development of the complex. BRDC is associated with reduced weight gain, failure to thrive, and even death in beef cattle; at slaughter the lungs are often noticeably damaged, reducing the price paid for the carcass [114].

BRDC is not a single disease, but a multi-factorial syndrome that has been associated with a group of bacterial pathogens, including *Mycoplasma bovis, Mannheimia haemolytica, Histophilus somni*, and *Pasteurella multocida.* However, these bacteria are also commonly found in the respiratory tract of healthy cattle, and experimental inoculation with these organisms does not lead to the classic clinical presentation of BRDC. Thus, mere exposure to these organisms does not seem to cause BRDC. It appears that the stressors associated with transporting and managing cattle in a feedlot predispose cattle to BRDC; the widespread view in the industry is that stress weakens the cows' immune systems, making them more susceptible. The fact that other stressors (co-morbidity with other pathogens, extreme or rapidly-changing weather, dust, dehydration, etc.) have also been identified as correlates of BRDC tends to support this notion. The pattern of cattle becoming sick immediately after transport is so common that cattle arriving from auction will normally be isolated from the rest of a producer's herd until the danger of infection subsides [114].

Containing disease outbreaks is an important loss-mitigation strategy for the stockman; every sick animal represents an economic loss, both from the reduced value of the sick animals, the cost of treating them, and the spread of sickness to other, currently healthy animals. Thus, every stockman must utilize some kind of livestock disease management system; whether this system is made up of ad-hoc techniques or scientifically designed, it must address prevention and treatment. Prevention includes animal replacement strategies (raising animals internally, purchasing from other producers, or buying at auction), livestock disease surveillance (checkups, behavioral observation, monitoring water intake and feeding), vaccinations and quarantining sick animals. For example, changes in milk characteristics and activity can be a sign of mastitis and lameness [80]; temperature variation can signal BRDC (or even worse, FMD) in cattle [46]. Treatment involves blood tests, temperature checkups, antibiotics, veterinary care, etc. [17,28].

There are a number of cattle diseases that are particularly important, due to their economic cost and/or the possibility of trade sanctions under criteria set by the World Organization for Animal Health (known as OIE, for its French name). FMD is one of these kinds of disease in animals such as cattle, sheep, pigs and goats. Caused by a virus from the Picornaviridae family, FMD symptoms take 2–3 days to appear, and can persist for 7–10 days. Fever, lameness, and vesicular lesions on the tongue, feet, snout, and teats are signs of FMD. Its transmission can be direct, indirect (aerosol droplets), or airborne; a large airborne virus plume can infect animals as far as 300 km away [48]. The 2001 outbreak in Europe led to culling millions of animals and trade sanctions; losses in the U.K. alone were estimated at $16 billion USD [117]. One important point concerning this outbreak is the shortage of veterinarians that the U.K.'s Ministry of Agriculture, Fisheries and Food was experiencing at the time of the outbreak; this is believed to have delayed control of the epidemic [13]. It seems logical that sensor-based systems might have helped catch the outbreak earlier (by focusing attention on the sick animals), when it might have been easier (and less costly) to manage. Another disease, mastitis, is perhaps the most common disease in dairy cattle. Bacteria migrate from the udder to the mammary gland causing inflammation, decreased milk yield and quality, and systemic signs (fever, depression, trembling, loss of appetite and weight loss). It can be spread during the milking routine by a contaminated milking machine or contaminated hands. The disease costs the U.S. dairy industry $1.7 billion annually [103]. Bovine Spongiform Encephalopathy (BSE), is a kind of fatal transmissible spongiform encephalopathy disease (TSE), caused by a misfolded protein called a prion, which attacks the central nervous of cattle and damages brain tissues. Its symptoms are nervous or aggressive behavior, unusual posture, lack of coordination or difficulty in rising from a lying position, decrease in milk production and weight loss in spite of increase in appetite. The disease can be transmitted to other animals and humans by consuming the flesh (and particularly nervous system tissue) of an infected animal. Variant Creutzfeldt-Jakob Disease (vCJD) is an incurable, fatal disease in humans that results from eating BSE-contaminated beef [18,44].

At the present time, most livestock disease management systems are based on visual evaluation and clinical tests. Besides weekly checkups, pen riders watch for any symptoms of disease; plainly, this is a labour-intensive and error-prone process, as the riders will only be able to spot changes in animal behavior and obvious clinical symptoms. In particular, pen riders are substantially slower to recognize a disease state compared with sensed data; researchers in [99] found that infrared thermography could detect a disease state 4–6 days before trained pen riders, while electronic monitoring of feeding behavior also detected disease an average of 4.1 days before the riders could [92]. There is furthermore often considerable variation in assessments of the same cow by different riders (Kappa statistic < 0.25 across multiple trials) [8], and the cows themselves, as prey animals, instinctively attempt to mask any weakness [129]. Finally, a large number of BRDC infections appear to be entirely missed, based on lung lesions observed by necropsy [102,128]. Those missed infections directly impact the final price at slaughter for the cow, both due to inferior perceived quality [45] and reduced weight [102]. One estimate put the sensitivity and specificity of the system of pen rider observation at just 61.8% and 62.8%, respectively [125].

Sensor technologies have recently been gaining favor as a means to automate disease-management systems, providing a stockman with earlier warnings of potential illness in an animal. We can differentiate between sensor-based examination of samples taken from an animal (e.g. milk, blood draws), and remote-sensing approaches. Sample-based monitoring approaches attempt to speed up lab throughput, in order to rapidly detect disease. Espada et al. [38] used color sensors in order to detect abnormalities in milk. Automatic somatic cell counters are used for mastitis detection [82,126]. Electrical conductivity, ultrasonic and electromagnetic sensors, biosensors and chemical sensors are also used for milk abnormalities and mastitis detection [15,53,72]. Pastell and Kujala [85] applied a probabilistic neural network model for lameness detection using measurements of the weight of a cow's legs. Cavero et al. [23] designed a neural network for early detection of mastitis by measuring maximal electrical conductivity of milk, milk yield, milk flow and days in milk. Kamphuis et al. [61] applied cost-sensitive C4.5 decision trees to detect mastitis from sensor readings of the color and electrical conductivity of milk.

The cattle industry has also taken note of the advantages of remote sensing; by design, these methods remove the need for direct handling of an animal, dramatically reducing the effort required for whole-herd surveillance. A review in [115] found that several novel sensing technologies have been introduced and evaluated in livestock operations; these include Infra-Red Thermography (IRT); rumen temperature boluses; tympanic (ear canal) temperature probes; intra-vaginal temperature probes; proximity sensors for feed and water locations; video capture & analysis; accelerometers; pedometers; GPS sensors; and real-time location sensors.

There have been a number of recent studies of IRT for disease surveillance and behavior monitoring of livestock and animals; it is perhaps the most widely-studied remote sensing technology. IRT involves training an infrared camera on animals in the pen, and then identifying the surface temperature of some part of the animal from the sensed wavelengths. The eye is a common choice for fever surveillance; anatomically it should correlate well with the animal's core temperature, and be less affected by ambient temperatures [83]. Studies of cattle inoculated with various disease-causing organisms (bovine viral diarrhea, BRDC, FMD, etc.) have shown that eye temperatures were a reliable indicator of fever [94]. It was used to screen for generalized fever in groups of piglets (rather than individual temperatures) in [31]. Other areas of the animal might also be selected; elevated foot temperatures were found to be a possible sign of FMD in [93] (although this has subsequently been found to be an unreliable test [47,94]). Further studies have shown that IRT is effective in detecting a wide variety of diseases in many species of animals, including bluetongue virus in sheep [88], canine bone cancer [88], FMD in deer [37], ear tag infections in lambs [63], pneumonia in pigs [79], and digital dermatitis [7], FMD [93] and mastitis in cattle [30,54,89]. Studies have found that numerous other health conditions may also be diagnosed via IRT, such as hoof lesions in cattle or sheep [6,111], lower leg inflammation in horses [105], and ligament damage in dogs [56]; note, however, that IRT was found to be poorly correlated with observed mobility in [96]. Pregnancy in horses [14], semen quality in bulls [78], and psychological stress in dogs [119] have also been detected using IRT.

However, a number of factors affect the performance of IRT. First is the physical distance between the camera and the animal; carbon dioxide in the air absorbs infrared radiation. In laboratory testing, there was a significant difference between readings taken at a distance of 0.5 m, and those taken at 2 m [83]. Another analysis in [57] confirmed a significant effect on measured temperatures beyond 1.5 m. There is a significant temperature difference between bare skin (the animal's coat being clipped) versus unclipped hair. Rapid decreases in ambient temperature will also cool the bare skin, but the animal's rectum temperature edges up. Overall, only eye temperatures were clearly correlated with rectum temperatures [83]. Wind and humidity also significantly affect IRT, as did sunlight [37,57,83,94]. Overall, these findings indicate that IRT should focus on eye temperatures, with thermograms being taken at short range (ideally 1 m or less), preferably in a controlled environment.

Rumen boluses are another means of monitoring animal temperature. A small device package (the bolus) is fed to the animal, and then lodges in the rumen. The bolus may be battery-powered as in [97], or it could by a passive RFID device [104] (the latter means cattle must pass close by an RFID reader in order for a temperature measurement to be transmitted). By definition, the temperature of its surroundings should be the animal's core temperature; these readings are reported over a wireless link. There is relatively little evidence on the efficacy of these devices; while both [104] and [97] find correlations with rectal temperatures, and the former observes fever signs when cattle are inoculated with bovine viral diarrhea and one of the BRDC pathogens, these are not clinical trials. In a more in-depth evaluation, Timsit et al. found a positive predicted value of 73% in [118]. Very recently, a pilot study on young calves found a sensitivity of 77% and specificity of 97% [123]. The device is relatively expensive, needing as it does to survive in an animal's digestive system for an extended period without causing ill effects. There are also potential environmental difficulties with picking up the bolus' radio signal [115]. Tympanic temperature probes were not very effective; the placement of the probe in the ear canal substantially affected the results, and the cows seemed to be able to partially dislodge it [77]. Intra-vaginal probes have also been proposed, but there is very little evidence on their efficacy to this point [115].

Proximity sensors for feed and water troughs have been developed and marketed by multiple vendors. Studies of cattle behavior indicate that sick animals will decrease their feeding times and increase the amount of time spent resting [51,129], and so in theory changes in feeding behaviors could signal disease. There is thus interest in searching for patterns in the proximity data that correlate with illness. Some systems are based on radio-frequency identification (RFID) tags, while others rely on transponder collars. Evidence for the utility of water trough sensors is mixed; Sowell et al. [106] found no association with illness, while Buhman et al. [19] found there was an association. In [34], C4.5 decision trees were used to detect disease from watering behavior; that paper is an earlier study of the sensor platform used in the current article. However, as also noted previously, changes in feeding behavior do seem to indicate disease; to the point where a disease state can be identified hours or days before pen riders notice distress [19,106,115,129].

Video capture and analysis generally consists of a camera recording of cow behavior – which some human must view and interpret. Protocols for doing so range from continuous monitoring (literally watching the cows at live speed) to various sampling strategies [115]. Automated behavior recognition, based on machine-vision approaches, is an alternative that attempts to relieve humans of this tedious job. Commercial systems are available to monitor animal behavior in laboratory settings, while research prototypes have been developed for detecting behaviors associated with illness in field settings [16,36,121] One important note is that continuous video monitoring is usually considered the gold standard of livestock behavior monitoring, against which other techniques are compared [115].

Accelerometers have been shown to be highly accurate (>97% accordance with video [95]) in detecting the behaviors of standing, walking, and lying. These behaviors do naturally vary by time of day and from day to day; nonetheless, these changes are also some of the signs pen riders use to identify sick cows. Pedometers have also been examined; these cheaper devices merely count the total number of steps taken and distance travelled. Nonetheless, behaviors associated with pain or stress can also be discerned in this data [33,49,115].

Technologies for monitoring the movement of cattle on a farm over time has also been investigated. Mounting GPS receivers on cattle is one obvious approach, but was found to be ineffective. The receivers in use were only accurate to roughly 10 m, and the batteries only lasted a few days. This spatial inaccuracy obscured any signs of illness [101,115,120]. An alternative approach is to deploy RF beacons and receivers around the farm, and use multilateration to determine the cow's location relative to these receivers. Signs of illness were detected in trials using these more accurate movement patterns [115,116,124].

Kramer et al. [64] developed a fuzzy logic system based on water intake, feeding behavior, milk yield and activity information of cows to diagnose mastitis and lameness. Kamphuis et al. [62] used a combination of sensor data and visual observations for mastitis detection; bagging and boosting C4.5 decision trees were studied. Martiskainen et al. [76] used multi-class support vector machines to predict how a cow would behave in a pen during the day (i.e. what is normal behavior, and what is not). De Mol et al. [32] applied time series analysis for oestrus and mastitis detection.

Transforming data from the time domain to the wavelet domain can sometimes improve a classifier's performance; partly by decreasing noise and partly by providing an alternative representation of the data. Since sensor data is inherently noisy, the wavelet transform is a plausible alternative representation, and various authors have explored it [87]. In the livestock disease management space, Kruse et al. [66] used wavelet coefficients of water intake of lactating sows to diagnose health problems. It employed differences between approximation coefficients and bootstraping. In [84], data recorded by an accelerometer goes through wavelet transformation in order to study the gait of cows. Miekley et al. [80] used wavelet analysis to detect two kinds of diseases: mastitis and lameness. Cow activity and milk electrical conductivity were transformed to the wavelet domain, and the filtered values used to predict the onset of disease.

Except for our own previous work in [34], all of the "remote sensing" solutions for temperature measurement require cattle to be in close proximity to a sensor (e.g. IRT) or wireless receiver/reader. In a feedlot, this weakness can be masked by placing those sensors/receivers near the feed and water troughs, knowing that the cattle will eventually go to those locations (although taking a thermograph of one eye in a crowd, and associating it with an ear tag also in the crowd, might be challenging). In a free-range farm, this would be far more problematic. While our own device would need to use a different wireless protocol for the free-range case (Zigbee only has a 100-m range), our transmitter's placement inside an animal collar affords easy access for swapping out the radios. Other remote sensing systems rely on less specific location or movement data, or video recordings. Close-range systems gather data from centralized sensors in an exam or feeding area, milk testing, and observations from stockmen or pen riders. In feedlot operations, the cows will only pass through exam areas once a week or less, while observations are gathered and recorded. Thus, those systems attempt to make current surveillance practices more sensitive and specific, but not more *timely.* Our system is one of the only approaches for *continuous* disease surveillance that we are aware of, and the only one that can easily be deployed on a free-range farm as well as a feedlot.

2.2. Discrete wavelet transform

Wavelets are literally defined as small waves having limited duration and an average value of zero [29,87,110]. A family of wavelets is defined as [29]:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), \quad a > 0, \quad b \in \mathbb{R} \tag{1}$$

where each wavelet, $\psi_{a,\,b}$, is obtained through translation and scaling of a basic wavelet called the mother wavelet $\psi(.)$. a and b are real values varying continuously and defining the amount of dilation and shifting of the mother wavelet, respectively The coefficients of the continuous wavelet transform, CWT, of a signal, $f(t)$, are obtained by [87]:

$$CWT(a,b) = \int_{-\infty}^{\infty} f(t).\psi^*{}_{a,b}(t)dt \tag{2}$$

where "*" denotes complex conjugation. CWT is a measurement of similarity between the signal and the set of wavelet functions.

The discrete wavelet transform (DWT) is obtained through discretizing dilation and translation parameters of the wavelet: $= {a_0}^j, \; b = kb_0{a_0}^j \; k, \; j \in \mathbb{Z}$. By choosing $a_0 = 2$ and $b_0 = 1$, the dyadic discrete wavelet is obtained [110]:

$$\psi_{j,k}(t) = 2^{-\frac{j}{2}}.\psi\left(2^{-j}t - k\right) \tag{3}$$

The DWT coefficients of a continuous time function, $f(t)$, are defined as [110]

$$w_{j,k} = \frac{1}{2^{\frac{j}{2}}}\int f(t)\psi\left(2^{-j}t-k\right).dt \tag{4}$$

Discrete wavelet coefficients can be presented through a pyramid algorithm proposed by Mallat [74]. The algorithm uses two filters: one high-pass filter, $h(n)$, and one low-pass filter, $g(n)$. These two filters are related by Percival and Walden [87]:

$$g(n) = (-1)^{1-n}h(1-n) \tag{5}$$

In signal processing, $h(n)$ and $g(n)$ are called quadrature mirror filters. The outputs of the high-pass and low-pass filters, down-sampled by a factor of 2 give the detail, D1, and approximation, A1, coefficients of the first level of decomposition, respectively. The second level of decomposition is obtained by reapplying the same procedure on the output of the low-pass filter. This process is recursively done on the approximation coefficients at each level giving sequential decompositions.

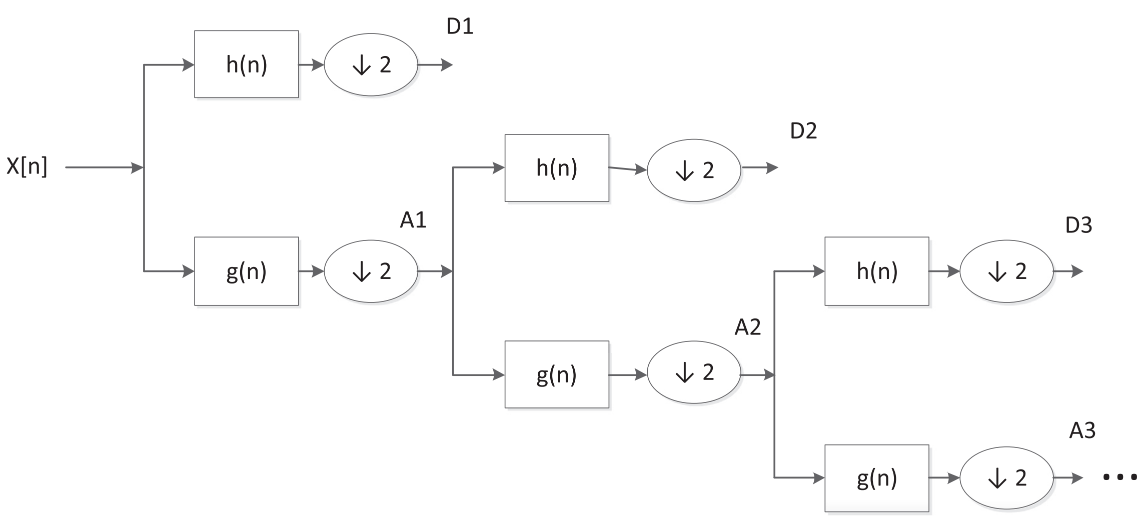

Fig. 1. Structure of a filter bank for wavelet decomposition [110].

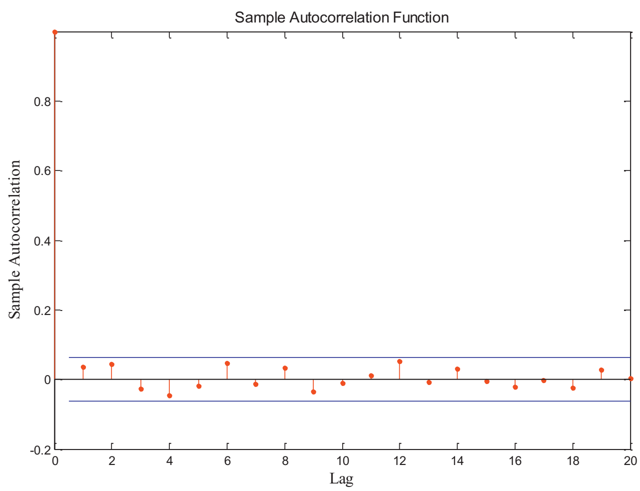

Fig. 2. ACF of a stationary signal [68].

By up-sampling the detail coefficients at each level and filtering them by $\bar{h}(n) = h(-n)$, the wavelet details of the level, are obtained and by up-sampling approximation coefficients and filtering them by $\bar{g}(n) = g(-n)$, the j-th level smoothed wavelet is obtained. The procedure is called a filter bank and is shown schematically in Fig. 1 [75]:

Fourier transforms capture the frequency contents of a signal, but time-domain information is lost. In the case of studying a non-stationary signal, whose frequency contents change over time, the Fourier transform is not a suitable tool. Short time Fourier transform (STFT) is a proposed methodology to extend Fourier transforms to non-stationary signals. One calculates the Fourier transform over a moving window in time domain, in which the signal is assumed to be stationary. However, the problem of STFT is constant time and frequency resolution. If the window length is narrow, it will give good time resolution and poor frequency resolution, and if the window is wide, poor time resolution and good frequency resolution are obtained. On the other hand, the wavelet transform studies a signal in different frequencies and time intervals, giving various time and frequency resolutions; at high frequencies, it has good time and poor frequency resolution whereas at low frequencies, it has poor time and good frequency resolution. The localization in time and frequency makes wavelet transform a popular tool in analyzing non-stationary signals [24,73,90].

Choosing a wavelet and decomposition level are essential steps in analyzing signals with DWT. The Haar wavelet is the first known and the simplest possible wavelet because it involves only differencing and averaging operations on a given signal. Its mother wavelet is defined as [113]:

$$\psi_{Haar}(u) = \begin{cases} -1 & -0.5 < u \leq 0 \\ 1 & 0 < u \leq 0.5 \\ 0 & otherwise \end{cases} \tag{6}$$

Thus, it is not a continuous wavelet, so it cannot smoothly follow a continuous signal, although this characteristic is beneficial when studying signals with sharp transitions. Moreover it is a two element wide wavelet, which reduces its resolution. In other words, the coefficients of the first level of decomposition are obtained by taking difference of two adjacent values, and average of them. On the other hand, other wavelets such as Daubechies, Mexican Hat, and Morlet are continuous with more flexible mother wavelets, which allows them to catch smoothed transitions better [3,40,65]. Nevertheless, Haar wavelets are widely used in areas such as time series analysis, stream data mining and databases because they are fast, memory-efficient, and easily computed [27,43,109,112].

Wavelet coefficients can be used as features in time series analysis, enabling dimensionality reduction and noise filtering. In dimensionality reduction, we attempt to preserve the information within a time series in a more compact representation. We can thus retain a subset of wavelet coefficients containing most of the energy of the time series and removing others. Retaining the largest coefficients and choosing the first k coefficients are other possibilities [113,132]. Likewise, noise filtering is accomplished by discarding the detail coefficients (which are more sensitive to noise), and retaining the others. The choice of decomposition level has been also discussed in some studies; Liu [70] selected the decomposition level for mass spectrometry data by examining the energy distribution of approximation coefficients. Sang et al. [98] developed a method for selection of the decomposition level for wavelet de-noising applications based on wavelet energy entropy. Liu [69] simply retained the detail coefficients from decomposition levels 2, 3 and 4. Reviews of different approaches for selection of coefficients and decomposition level can be found in [25,113]. In addition to these basic techniques, additional features can be derived from wavelet coefficients. Subasi [110] extracted statistical properties of wavelet coefficients from specific frequency sub-bands; Long et al. [71] considered statistical properties of the coefficients such as min, max, and average; Scheunders et al. [100] used wavelet energy as the features.

2.3. *Time*-domain analysis

When the statistical properties of a time series, e.g. mean and variance remain constant over time, the time series is considered stationary. In general, this is not the case for observations of real systems; they can drift, exhibit periodic behaviors, even undergo abrupt changes. Non-stationarity in a time series generally consists of three components: trends, seasonal and cyclic patterns. A trend is a long-term change in the mean of the time series; seasonal variations are changes occurring over a specific period, and cyclic variations are changes that do not have fixed periods. A time series with trend or seasonal variations is considered non-stationary whereas having only a cyclic pattern does not make a time series non-stationary (as these are treated as noise in statistical models). These two components of non-stationary signals can result in misleading interpretation of statistical results of the given time series since the repeating and normal events within the signal can obscure important information [26,86].

A non-stationary time series is transformed to a stationary one by removing trends and seasonality; differencing is one simple way to do so. In de-trending, changes between consecutive points are investigated [26,55]. Seasonality is adjusted by differencing each data point from the corresponding data point in the previous season.

To determine if a given time series is stationary or not, the autocorrelation function (ACF) correlogram graph can be used, which depicts the relation between sample autocorrelations and time lags [26]:

$$r_k = \frac{\sum_{t=1}^{N-k} (x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N} (x_t - \bar{x})^2} \tag{7}$$

where r_k is the autocorrelation coefficient at lag k, N is the number of observations, and $\bar{x}$ is the sample mean of the time series. By plotting the 95% confidence interval $\pm\frac{2}{\sqrt{N}}$ and observing the values of r_k in this region, we can determine if the time series is stationary or not; if the values converge quickly to the confidence interval, it is a stationary signal [10,26]. If the signal still shows high correlation between lags after the first differencing, a second differencing is used [26,55]. However, we rarely do differencing more than twice [10,26]. When both trends and seasonal variations appear in a time series, we can first perform seasonal differencing and then apply regular differencing. Note that de-trending and de-seasoning cannot be considered methods for deal with concept drifts, as they are determined *a posteriori*. Cyclic patterns, on the other hand, can be considered as an example of reoccurring drifts [122,133].

2.4 Inferential sensors and condition monitoring

The term *condition monitoring* traditionally means the use of non-destructive testing and sensing to detect changes in engineered systems (bridges, electric motors, etc.), on the theory that at least some of these changes may indicates damage or degradation to the system. Some advanced techniques can quantify the location, nature and severity of the damage, and forecast how it affects the remaining service lifetime of the system. Some modern systems might even be able to attempt self-repair based on these findings [22]. In this article, however, we use the term in the broadest sense: the use of non-destructive (benign) sensing to monitor the "health" of some object of interest, by collecting a stream of observations. Changes from a baseline are noted, and classified as harmless or harmful. This obviously includes classical condition monitoring and its relative, structural health monitoring; livestock disease surveillance; public health monitoring [20]; and monitoring individual *wearables*, such as the Fitbit®, to name just a few.

Inferential sensing (also known as *soft* sensing) generally refers to estimating a complex, high-value, time-varying concept by fusing together a number of simpler (and cheaper) sensor outputs [58,59,91]. Research in this topic has progressed from early (1970s) designs using Kalman filters [58] and multivariate statistical analysis [59] to modern self-evolving approaches that adapt to concept drifts automatically [9]. Since the goal of condition monitoring is to track the health of an object of interest (which is almost certainly a complex, time-varying concept), inferential sensing is an obvious methodological choice. Indeed, inferential sensors for condition monitoring have been reported in [11,60,67,91], among others. From this discussion, it is clear that the livestock disease surveillance system proposed in the current paper is also an example of an inferential sensor, applied to an instance of the general condition-monitoring problem we have described.

3. Methodology

In this section, we describe our experimental methodology. The first part describes our data collection procedure; the second and third parts discuss our methodology for time-domain analysis and wavelet analysis, respectively.

3.1. Dataset collection

The Alberta Research Council, Xanantech Inc., Ovistech Inc., and the University of Alberta have proposed an animal mounted sensor platform that collects animal and ambient temperature from sensors mounted in an ear tag the animal's ear every five minutes. The ear tag connects to a neck-mounted collar which also incorporates feeding and watering proximity sensors and a three-axis accelerometer. Data is reported back to a base station via Zigbee radios mounted in the animal collars, and then transmitted via a data cable to a computer workstation [50]. The system was mounted on 31 newly-acquired cows at an Edmonton-area feedlot between February and May 2006. The feedlot operator's experience is that the first month after an animal is acquired is the period when they are most likely to get sick; sickness is rare after that. Due to Edmonton's severe winters, temperatures measured from an animal's ear will be imprecise; however, subcutaneous or rectal sensors were rejected as creating too much distress for the animals. Instead, ambient temperature was also measured, and the difference between animal and ambient temperatures was derived, yielding for each cow a trivariate time series with a sampling period of five minutes extending over roughly five continuous weeks.

Inspection of the data revealed that there were spurious and missing data points. Spurious data points are measurements occurring in less than five minutes (likely due to re-transmissions from the sensors to the base station), and missing data points refer to a lack of measurements after a full five minutes. To deal with the problem, we deleted spurious data points and used linear interpolation to impute missing data points. Moreover, we expected 288 sequential observations for each animal per day (which we refer to as a cow-day). After cleaning the data sets, we removed cow-days with less than 90 measurements (which occurred due to adverse events, e.g. the cows eating the base station data cable), leaving an average of 35 days of measurements per animal.

Finally, a classification label was appended to each instance in the data sets, reflecting the presence or absence of disease. We applied a coding rule to determine the class labels, which is aligned with farmers' needs in predicting the onset of disease. All observations in a cow-day are labeled sick if an intervention or sickness is observed within the next k days ($k = 7$ for our experiment, following advice from the feedlot operators), otherwise the observations are labelled as healthy [131]; thus the class labels show the prediction of animal status in terms of health and sickness in the next 7 days based on features of the current day.

In this experiment, only data from cows having both sick and healthy records are considered; 9 out of 31 cows satisfy this condition. This is necessary since learning algorithms exposed to only one class will generally predict that *all* examples belong to that class. These datasets are divided equally into training and testing datasets; this is a chronologically-ordered single-split design, which ensures training data points occur earlier than testing data points. This design is a common approach in time series analysis. Figs. 3 and 4 show features from a healthy and a sick cow-day of one of the cows (C3434), respectively.

3.2. Time domain

For our time domain experiments we study each cow individually; it is a known medical fact that the "normal" temperature of different individuals can vary substantially, and this renders time-series analysis ineffective if the animals were to be pooled. Firstly, we re-examine the results from [6], in which we explored classifying each observation using only the instantaneous animal and ambient temperatures. That approach was ineffective, so we next explore the addition of a derived variable representing the difference between animal and ambient temperature (on the theory that extreme ambient temperatures may be biasing the animal temperature readings). Finally, we explore the possibility that there is a non-stationarity in our data. Intuitively, it seems reasonable that the time series could be non-stationary and contain trends and seasonality (periodic fluctuations; note that this is a generic characteristic of time series, not connected to the seasons of the year), since they record temperature measurements during consecutive months in winter and spring. In the Edmonton area, this time of the year sees a dramatic change in ambient temperatures from the beginning to the end of the study period, and also within a single day. The former may create a trend in the time series, while the latter may create a seasonality effect

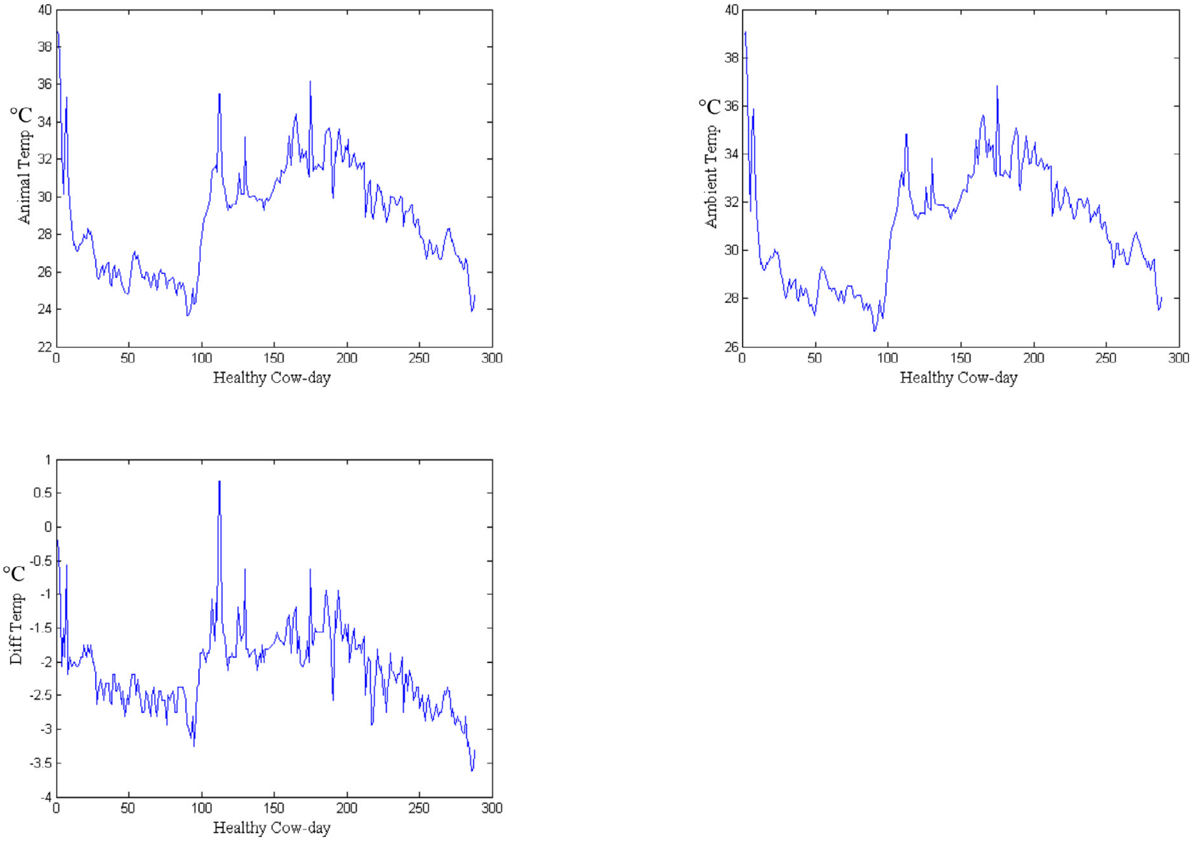

Fig. 3. Animal temperature, Ambient temperature and Difference in one healthy cow-day in C3434.

with a period of one day (288 measurements). We determine whether the time series are stationary or non-stationary using the AutoCorrelation Function (ACF).

The trend and seasonality in our data are corrected by differencing. The procedure is repeated for all the training and testing data sets. However, the labels in the new data sets must be adjusted after differencing. The problem appears in seasonal differencing where we have two consecutive cow-days labelled with two different classes. In that case, the data obtained by differencing these two days are labelled by the following rules, deduced from our coding rule for assigned label to data sets: If the first day is labelled as healthy and the second one as sick, their difference will be labelled as sick because the animal is becoming sick or needs treatment in the next 7 days; if the first day is sick and the second one is healthy, it will be labelled as healthy.

We test our four machine learning algorithms, Radial Basis Function Network (RBFN), Support Vector Machines using Sequential Minimal Optimization (SMO), C4.5, and Logistic Regression, on the obtained training data sets. The best set of parameters for the classifiers are obtained by parameter exploration on the training sets via 10-fold cross validation. This best parameter set is then used to train a fresh example of the classifier on the whole training set, which is then evaluated on the testing set. We measure performance using the well-known sensitivity and specificity metrics, as well as Cohen's Kappa statistic.

3.3. Wavelet domain

For this experiment, we design a new data set where cow-days (data for one cow, recorded for one day) and wavelet coefficients are considered as instances and features, respectively. We have three different measurements in the previous data set: animal temperature, ambient temperature and difference between them; thus, for each cow-day, we calculate wavelet coefficients of each of the measurements, separately, and then combine them. Finally, we construct a data set by combining the new data sets of all the cows (as we are examining differences between measurements instead of raw temperatures, this seems harmless). The new data set has 349 instances, with 157 instances from the original training sets of the cows, and 192 instances from the testing sets. For comparability with our time domain results, we do not re-balance our training and testing sets.

For the discrete wavelet transform, we use the Haar wavelet function, and we use the energy distribution of approximation coefficients to select the appropriate decomposition level. However, the number of wavelet coefficients is also taken

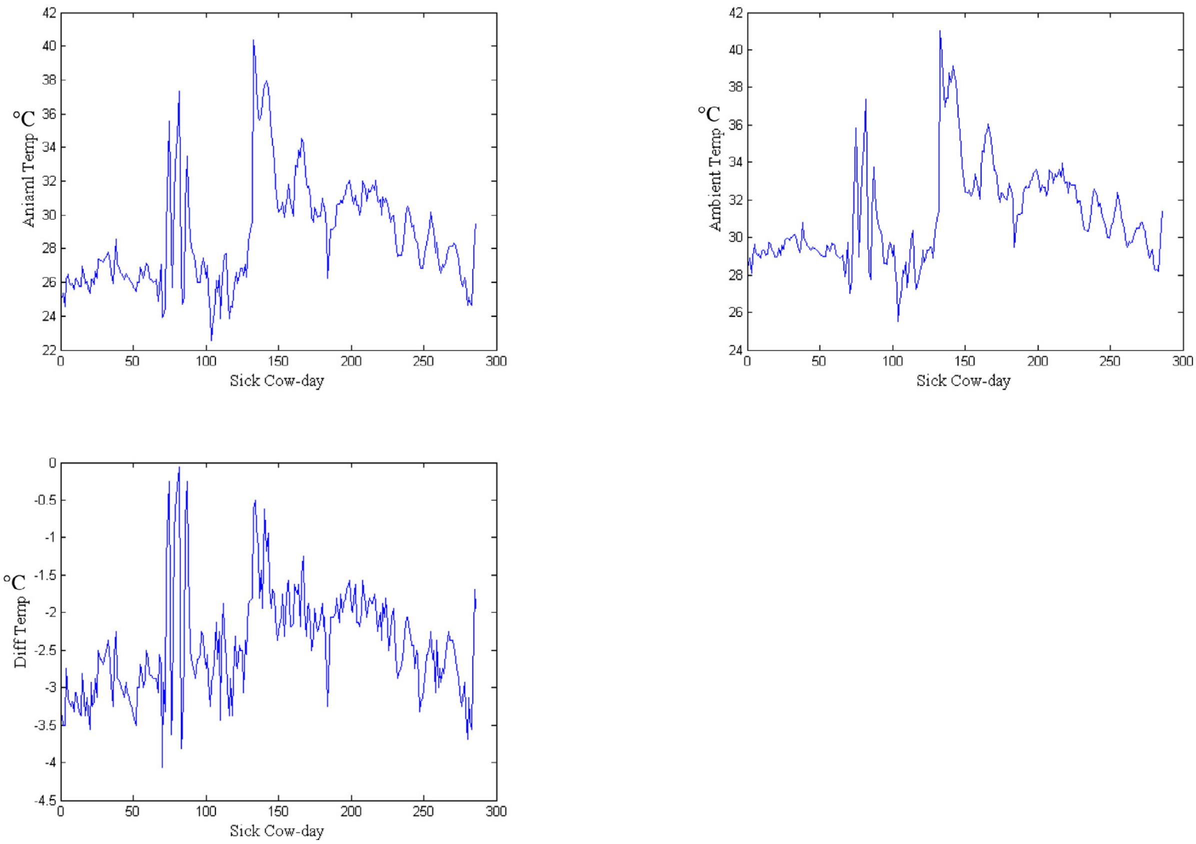

Fig. 4. Animal temperature, Ambient temperature and Difference in one sick cow-day in C3434.

Table 1
Number of approximation wavelet coefficients for a cow-day.

	Level 1	Level 2	Level 3	Level 4	Level 5	Level 6	Level 7	Level 8
# approximation wavelet coefficients	144	72	36	18	9	5	3	2

into account when selecting the decomposition level since our new data set does not have many instances (349 instances). We thus take the following steps to select the decomposition level, following [70]:

1. For each cow-day, obtain the maximum decomposition level by $M = \log_2(n)$ [98], where n is the length of a cow-day.
2. For each cow-day, calculate wavelet coefficients of all the possible levels using the MATLAB Wavelet toolbox [81]".
3. For each cow-day, calculate the percentage of energy of the cow-day preserved by approximation coefficients for each decomposition level using the L2-norm [81].
4. For each cow, calculate the mean of the energy distribution of the cow-days per decomposition level.
5. Compare the energy distribution of all the cows, and with consideration of number of wavelet coefficients, select the highest decomposition level that best preserves the energy distribution for all cows.
6. We repeat steps 1–5 for all three measurements: animal temperature, ambient temperature and temperature difference; then the common decomposition level having the greatest energy retention between them is selected.

Each cow-day has at most 288 measurements (every 5 min), resulting in a maximum of 8 decomposition levels. The number of approximation wavelet coefficients in each level for a cow-day is presented in Table 1.

The energy distribution of animal temperature and ambient temperature is around 98% for all decomposition levels; therefore, we decide the decomposition level based on the energy distribution of the temperature difference, given in Table 2. We limit this presentation to the first five decomposition levels, as the energy distribution for several cows has plainly fallen below acceptable levels by this point.

Based on Table 2, the first level of decomposition gives 144 wavelet coefficients for each of the three measurements resulting in $144 \times 3 = 432$ features, meaning we have more features than instances. Table 2 shows that except C3528, all cows retain over 90% of the signal energy at the first decomposition level. The second level of decomposition has 72 wavelet coefficients ending to $72 \times 3 = 216$ features, which is less than the number of instances, although it is still a huge number

Table 2
Mean of wavelet energy for difference temperature per decomposition level.

	Level 1	Level 2	Level 3	Level 4	Level 5
C3434	98.8506	97.8589	96.8511	95.7394	94.9155
C3933	98.1544	96.3335	94.5122	93.1605	91.5999
C3555	97.6062	95.7449	93.6795	91.7777	90.5491
C3401	97.8498	95.876	93.6362	92.1393	90.7555
C3271	90.703	80.6815	71.0712	61.8738	49.841
C3519	97.0644	94.3278	91.4267	88.2155	85.5584
C3550	95.2994	90.0873	84.1627	78.4353	73.6691
C3528	89.0877	79.7813	70.2913	60.6785	53.5275
C3570	96.1628	92.818	90.1088	88.0494	85.6926

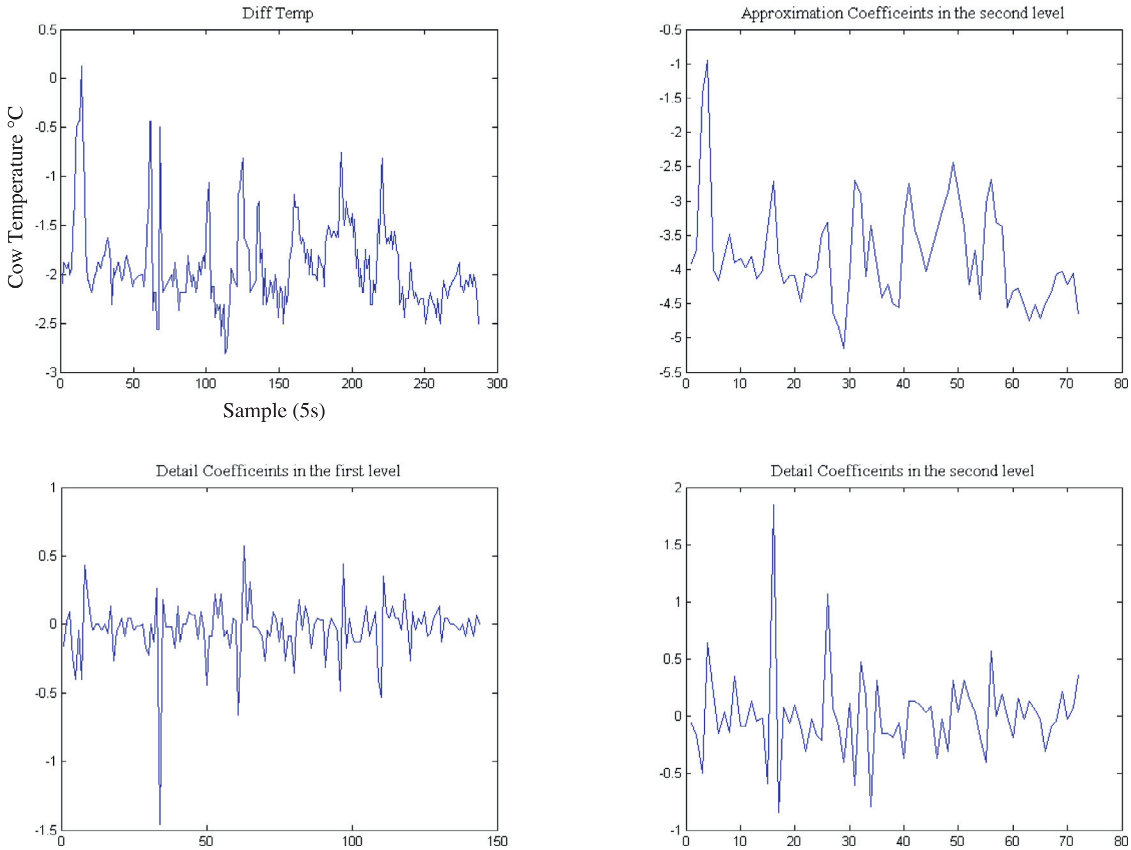

Fig. 5. The temperature difference of the first cow-day of C3434, wavelet approximation in the second level and detail coefficients in the first and second level.

for the data set. Energy retention in the second level, besides C3528, C3550 and C3271, remains over 90%. This appears to be the best balance of retained energy and dimensionality reduction. To demonstrate how the decomposition level has impact on the data sets, Fig. 5 shows the temperature difference of the first cow-day of C3434, and its decomposition in the second level.

Therefore, the new training set has 157 instances and 216 features. By applying the Information gain feature selection technique on the dataset, we reduce the number of features to 18, where the temperature difference, animal temperature and ambient temperature give 8, 6 and 4 features, respectively. The information gain method is implemented through WEKA [50], and evaluates the importance of a given attribute by measuring its information gain concerning a class:

$$InfoGain(Class, Attribute) = H(Class) - H(Class|Attribute) \tag{8}$$

where $H(.)$ indicates the information entropy of a given object. To append classification labels to the new dataset, we use the fact that in the time-domain dataset, class labels are constant within a single cow-day; thus, the new instances simply keep the labels. The training set has 109 healthy and 48 sick labels, and the testing set has 165 healthy and 26 sick labels.

Table 3
Average Kappa statistic for instantaneous measurements (training) [131].

ID	Log	C4.5	RBFN	SVM
3271	−0.0352	0.225	**0.2723**	0.0295
3401	0.3119	0.4679	**0.5228**	0.345
3434	0.0987	0.222	**0.2942**	0.1175
3519	0.0092	0.3055	**0.3061**	0.0825
3528	0.2019	0.2716	**0.4026**	0.2653
3550	0.267	0.6052	**0.6396**	0.6396
3555	0.2095	0.3698	**0.4077**	0.3892
3570	0.3465	0.353	**0.4872**	0.3088
3933	0.4385	0.5043	**0.5732**	0.5027

We apply four machine learning algorithms on the training set: Radial Basis Function Network (RBFN), Support Vector Machine using Sequential Minimal Optimization (SMO), C4.5, and Logistic Regression. The classifiers are implemented via WEKA. Parameters to be explored for each algorithm are as follows: for RBFN, the number of basic functions, ridge value for regression, and minimum standard deviation for the clusters (basis functions). For SMO, we examine three different kernel functions: radial basis function (RBF), polynomial and normalized polynomial kernel; kernel parameters are gamma for RBF and the exponent for polynomials; and finally the complexity parameter C. In J48, (an implementation of C4.5 in WEKA), minimum numbers of instances per leaf and confidence factor are altered. Only the ridge parameter in logistic regression is explored. To obtain the best sets of parameters in each classifier, a parameter exploration via 5-fold cross validation is performed on the training set. The results are evaluated using Cohen's Kappa statistic which is defined in the interval of $[-1,1]$ with $\kappa = 0$ showing no improvement over random chance, $\kappa = 1$and $\kappa = -1$ denoting perfect classification and misclassification, respectively [39].

Our focus is on finding classifiers with optimal sensitivity and specificity, where sensitivity is the proportion of sick animals classified correctly by the classifiers, and specificity is the proportion of healthy animals correctly classified [42]. Moreover, in our experiment, misclassification of sick cows can lead to spreading the disease among healthy ones; while misclassifying healthy cows as sick costs money for check-ups, quarantine and treatment, and possibly the unnecessary use of antibiotics. These costs should still be less than the losses from an unchecked spread of disease; thus, sick cow misclassification should be penalized more severely than misclassifying a healthy cow. The MetaCost algorithm adds cost-sensitive classification to whatever base classifier it is applied to [35]. Thus, once the "best" set of parameters for each of the classifiers is determined, we do two experiments; first, a fresh classifier is trained using only those best parameters on the entire training set, and then the classifier is evaluated on the holdout testing set. Second, the classifier is used as the base classifier for the meta-classifier; then, we change the cost ratio for the MetaCost classifier; classifier outputs with different cost ratios are then used to construct an ROC curve. The ROC curve [41] is closely related to sensitivity and specificity, as they are directly related to the true positive rate (TPR) and false positive rate (FPR) as follows:

$$Sensitivity = TPR \times 100 \tag{9}$$

$$Specificity = (1 - FPR) \times 100 \tag{10}$$

The best sensitivity and specificity of each ROC curve is determined by the fact that the point (0,1) on the ROC curve has sensitivity and specificity of 100%; thus, the closer a point is to the perfect point (in terms of Euclidean distance), the better performance it has. We finally compare the sensitivity and specificity of those "best" points on the ROC curves. It is important to note that this usage of ROC curves is distinct from their usual task of comparing classifiers. ROC analysis is usually motivated by the desire to determine which classification *algorithms* would perform better across a range of tradeoffs between type-I and type-II errors. Instead, we are looking for the single best classifier *instance*, and so we use MetaCost to bias our learning in favor of positive-class examples, and ROC curves as a means of visualizing the sensitivity and specificity of the resulting classifiers.

4. Experimental results

4.1. Instantaneous temperatures

We begin this discussion be reviewing the detailed results from [131], which focused on the instantaneous animal and ambient temperatures for each of the nine cows. The results of the tenfold cross-validation experiments on the training set are given in Table 3.

In almost all cases, RBFN was the best classifier, albeit often only by a minor amount. As we did not observe a consistent, statistically significant superiority of any one classifier over the others (according to the *t*-test), we elected to run all classifiers on the test sets, using only the "best" parameters found during training. Those results are presented in Table 4.

Table 4
Sensitivity and specificity for instantaneous measurements (testing) [131].

ID	C4.5		Log		RBFN		SVM	
	Sn	Sp	Sn	Sp	Sn	Sp	Sn	Sp
3271	0	0.840	0	0.990	0	0.903	0	1.0
3401	0.157	0.872	0.515	0.548	0.224	0.895	0.472	0.581
3434	0.460	0.438	0.614	0.136	0.663	0.387	0.621	0.200
3519	0	0.912	0	0.999	0	0.950	0	1.0
3528	0	1.0	0	0.999	0	0.992	0	0.996
3550	0	0.956	0	1.0	0	0.948	0	0.929
3555	0.254	0.541	0.248	0.670	0.264	0.568	0.272	0.526
3570	0	0.997	0	0.978	0	0.977	0	0.997
3933	0	0.681	0	0.443	0	0.511	0	0.648

Table 5
Average Kappa statistic for instantaneous differencing (training).

ID	C4.5	Logistic	RBFN	SVM
3271	0.283	0.022	0.303	0.151
3401	0.513	0.440	0.518	0.502
3434	0.321	0.074	0.350	0.111
3519	0.321	0.136	0.338	0.008
3528	0.493	0.289	0.509	0.121
3550	0.661	0.644	0.765	0.655
3555	0.389	0.210	0.413	0.378
3570	0.462	0.311	0.525	0.296
3933	0.551	0.401	0.566	0.444

Table 6
Sensitivity and specificity for instantaneous differencing (testing).

ID	C4.5		Log		RBFN		SVM	
	Sn	Sp	Sn	Sp	Sn	Sp	Sn	Sp
3271	0	0.924	0	0.996	0	0.909	0	0.938
3401	0.151	0.882	0.080	0.874	0.403	0.770	0.054	0.922
3434	0.657	0.406	0.608	0.141	0.522	0.399	0.639	0.211
3519	0	0.921	0	0.967	0	0.921	0	1.0
3528	0	0.980	0	0.975	0	0.984	0	1.0
3550	0	0.954	0	0.977	0	0.911	0	0.935
3555	0.234	0.558	0.245	0.673	0.273	0.556	0.278	0.530
3570	0	0.991	0	0.993	0	0.987	0	0.995
3933	0	0.434	0	0.394	0	0.516	0	0.715

From Table 4, we find that the performance of all our classifiers on the holdout sample is poor; for most of the cows, the classifiers are unable to recognize sickness, and in the few exceptions, the classifiers do poorly in recognizing healthy cows. Furthermore, these results represent a precipitous decline from the tenfold cross-validation results on the training set. Plainly, this is not useful, and we must seek alternatives. The first possibility we will examine is whether the animal temperature sensor is being biased by the outside air temperature. Winters in Alberta are generally severe, with overnight temperatures below $-30\,^{\circ}$C commonplace. Since our animal temperature sensor is designed to read from the inner surface of the cow's ear, such extreme ambient temperatures might in fact depress the measured *animal* temperatures, as the ear is only a thin piece of tissue physically separated from the cow's main body mass. We attempt to correct for any such bias by adding a derived variable representing the difference between the animal and ambient temperature. In Tables 5 and 6, we examine this possibility.

Once again, we found that the performance of our classifiers was quite poor on the holdout sample; and that performance dropped sharply from our tenfold cross-validation results. Hence, this does not seem to be a simple matter of bias; we instead suspect that the data may be non-stationary. We will explore that possibility in the next sub-section using de-trending and de-seasoning techniques.

4.2. Removing trend and seasonality

In the interests of brevity, we will only describe the de-trending and de-seasoning of the data for one cow; the same procedures were followed for the other cows as well. Figs. 6 and 7 show the body temperature of one of the cows, C3401, and its ACF plot.

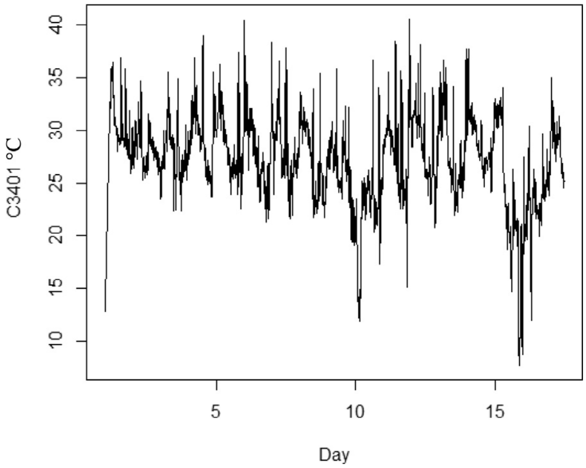

Fig. 6. The body temperature of C3401.

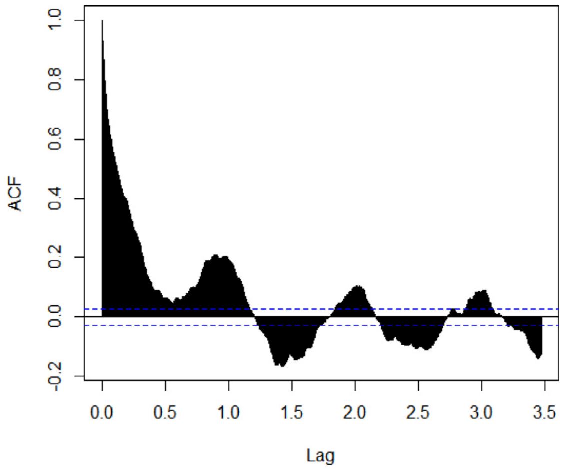

Fig. 7. ACF of the body temperature of C3401.

Table 7
Cohen's kappa statistics for the training sets.

	C4.5	RBFN	Logistic regression	SMO
C3434	0.1	0.09655	−0.0015	0.00944
C3401	0.061	0.0066	0	0
C3555	0.022	0.0242	0.00052	0

Fig. 7 indicates that there is a possibility of seasonal variations in the time series which can be removed by differencing; Figs. 8 and 9 shows the body temperature and ACF plot after de-seasoning.

Fig. 9 indicates that despite the de-seasoning, our time series is still non-stationary, and so we apply de-trending as well. Figs. 10 and 11 shows C3401's body temperature after de-seasoning and de-trending, and its ACF.

According to Fig. 11, after de-seasoning and de-trending, the trend and seasonal variations are almost zero. After assigning new labels to the obtained training and testing data sets, we observed that only three of the new data sets still had both healthy and sick labels. This is potentially due to the fact that some cows were coded as "sick" on their arrival dates, when they needed an intervention, but then were healthy at all subsequent checkups. Thus, we only applied our classifiers on the three cows still showing illnesses. Tables 7 and 8 presents Cohen's Kappa statistic on the training and testing sets. While the performance of the classifiers is now roughly the same in the training and testing sets, the quality is uniformly low; we seem to have resolved the non-stationarity, but the solution quality is unacceptable for this problem domain. We will therefore explore classification in the wavelet domain next.

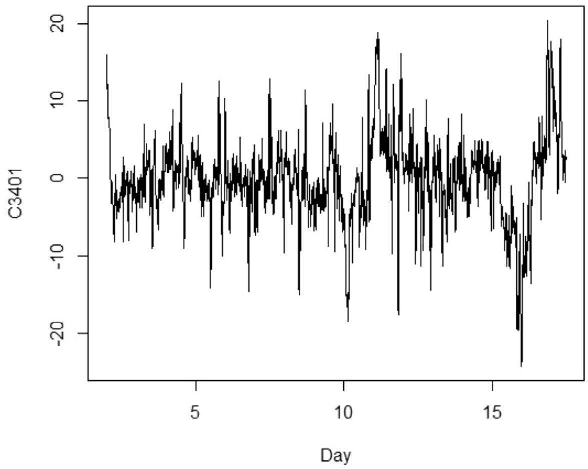

Fig. 8. The body temperature of C3401 after de-seasoning.

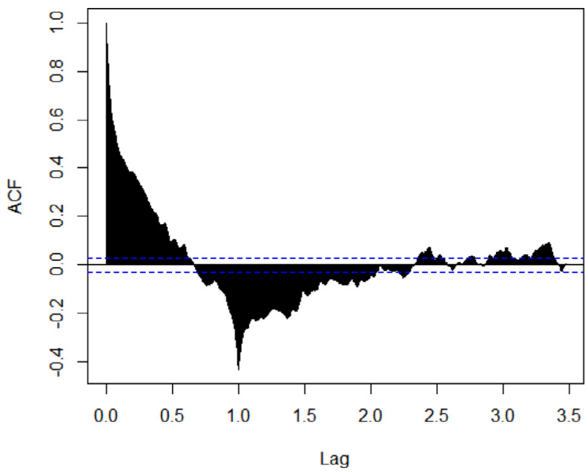

Fig. 9. ACF of the body temperature of C3401 after de-seasoning.

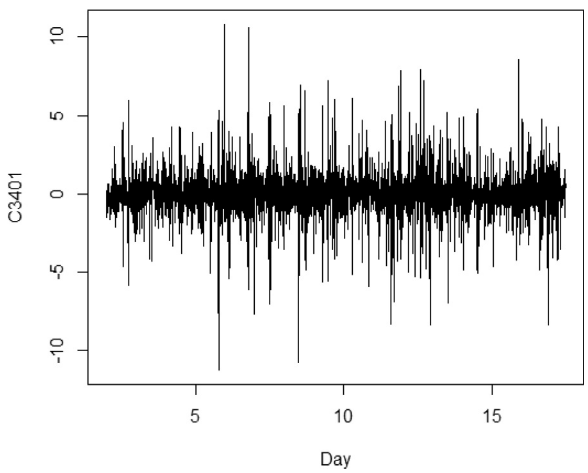

Fig. 10. The body temperature of C3401 after de-seasoning and de-trending.

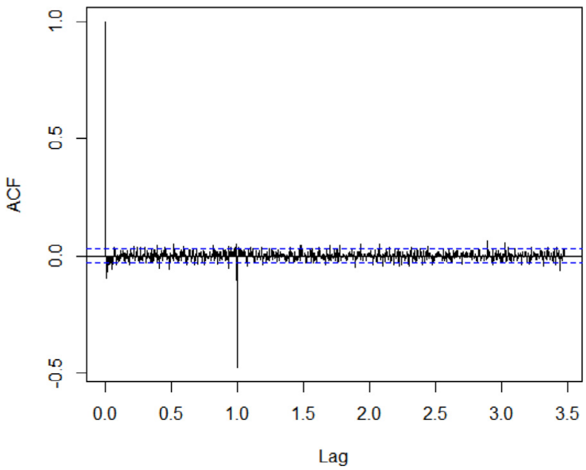

Fig. 11. ACF of the body temperature of C3401 after de-seasoning and de-trending.

Table 8
Cohen's kappa statistics for the testing sets.

	C4.5	RBFN	Logistic regression	SMO
C3434	−0.151	−0.129	−0.0009	0
C3401	−0.038	0.0015	0	0
C3555	0.038	0.0297	0	0

Table 9
The classifiers' results in terms of Kappa statistic, sensitivity and specificity.

	Training	Testing
RBFN1	$\kappa = 0.40544$	$\kappa = 0.3913$ Sensitivity = 53.8% Specificity = 89.1%
RBFN2	$\kappa = 0.41254$	$\kappa = 0.2053$ Sensitivity = 50% Specificity = 78.2%
SMO (RBF Kernel)	$\kappa = 0.30382$	$\kappa = 0.1541$ Sensitivity = 26.9% Specificity = 88.5%
SMO (Polynomial Kernel)	$\kappa = 0.29228$	$\kappa = 0.0749$ Sensitivity = 30.8% Specificity = 78.8%
SMO (Normalized Poly Kernel)	$\kappa = 0.32334$	$\kappa = 0.1054$ Sensitivity = 30.8% Specificity = 81.7%
C4.5	$\kappa = 0.30714$	$\kappa = 0.108$ Sensitivity = 38.5% Specificity = 76.4%
Logistic regression	$\kappa = 0.18288$	$\kappa = 0.3682$ Sensitivity 42.3% Specificity = 92.7%

4.3. Wavelet domain

The results for the best set of parameters for each classifier are reported in Table 9. For RBFN, we report two different sets of parameters with near-identical performance on the training set. Fig. 12 presents the corresponding ROC curve for one classifier, obtained by varying the cost ratio from 2:1 to 20:1 in favor of sick cows. The best sensitivity and specificity for each classifier, and the cost ratio corresponding to it, is presented in Table 10.

4.3.1. Sensitivity experiment

To see if the proposed approaches are robust to random noises, we conduct another experiment in which we add uniform noise to 10% of the instances in the training set; the noise is added to the last feature by preprocessing filters in WEKA [50]. We train the RBFN algorithm, using the same procedures as described in Section 4.3, on the contaminated training set, and

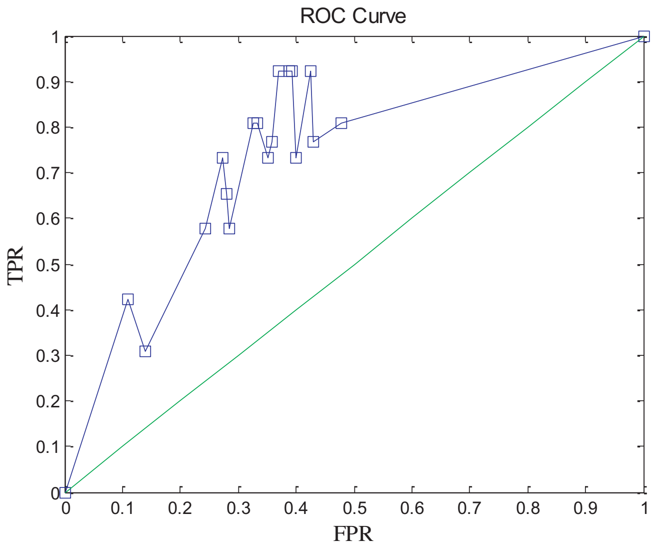

Fig. 12. ROC curve of RBFN1 when changing cost ratio from 2 to 20 in favor of sick animals.

Table 10
ROC curve results for the classifiers.

Classifier	Cost ratio	Sensitivity	Specificity
MetaRBFN1	15,16,17	92.3%	63%
MetaRBFN2	3	73.1%	78.8%
MetaSMO (RBF Kernel)	9–20	53.8%	81.8%
MetaSMO (Polynomial Kernel)	5–9	53.8%	69.1%
MetaSMO (Normalized Poly Kernel)	10–12	73.1%	65.5%
MetaC4.5	8	73.1%	70.9%
Meta Logistic Regression	4	80.8%	66.1%

Table 11
Sensitivity experiment results.

	Training	Testing	MetaCost (optimum for each classifier)	MetaCost (comparison)
RBFN1	$\kappa = 0.40544$	$\kappa = 0.3913$ Sensitivity = 53.8% Specificity = 89.1%	Cost Ratio = 15 Sensitivity = 92.3% Specificity = 63%	Cost Ratio = 4 Sensitivity = 57.7% Specificity = 75.8%
RBFN (Sensitivity Experiment)	$\kappa = 0.39458$	$\kappa = 0.1881$ Sensitivity = 34.6% Specificity = 86.1%	Cost Ratio = 4 Sensitivity = 92.3% Specificity = 67.3	Cost Ratio = 15 Sensitivity = 96.2% Specificity = 36.4%

evaluate the resulting models on the unaltered testing set. Table 11 compares results of RBFN1 from Section 4.3 and the sensitivity experiment.

From Table 11, after adding noise to the training set, the performance of the RBFN declines substantially. The picture is more complex when cost-sensitive classification is added. We found that a different cost ratio was optimal in the sensitivity experiment, and the effects could be fairly profound. We include two columns for these results in Table 11. MetaCost (optimum for each classifier) gives the sensitivity and specificity for the best cost ratio on each classifier (again, judged by Euclidean distance from the optimal (0,1) point ion ROC space), while MetaCost (comparison) gives the sensitivity and specificity if the cost ratio for the *other* classifier is used. We found that, even though the performance of the RBFN dropped when noise is added to the training set, with cost-sensitive classification there is a slight ad-hoc *improvement* in specificity, with sensitivity unchanged. While this odd result is likely dataset-specific, the evidence does indicate that our method is robust against noise in the sensor readings.

4.4. Ensemble classifiers in the wavelet domain

Tables 9 and 10 indicate that no one learning algorithm gave us both high sensitivity and high specificity, but most algorithms were strong in one or the other with an appropriate cost ratio. This is a reasonably common situation in intelligent systems, in which different classifiers minimize either false-positive or false-negative errors, but not both at once. One very

Table 12
Ensemble classifier results – voting.

Classifiers combined	Sensitivity	Specificity
MetaRBFN1 & Logistic Regression	69.2%	76.4%
MetaRBFN1 & Logistic Regression & MetaRBFN2	80.8%	80%

Table 13
Ensemble classifier results – stacking.

	Sensitivity	Specificity
MetaRBFN1 & Logistic Regression	38.5%	95.8%

Table 14
Results of sequential classification.

<table>
<tr><td>First Test</td><td></td><td>Second Test</td></tr>
<tr><td>MetaRBFN1</td><td></td><td>Logistic Regression</td></tr>
<tr><td>Sensitivity= 92.3%
Specificity= 63%
Confusion Matrix:
<table>
<tr><td colspan="2">Predicted class</td><td></td><td></td></tr>
<tr><td>healthy</td><td>sick</td><td></td><td></td></tr>
<tr><td>104</td><td>61</td><td>healthy</td><td rowspan="2">Actual class</td></tr>
<tr><td>2</td><td>24</td><td>sick</td></tr>
</table></td>
<td>Number of animals predicted as sick:
TP+FP=24+61=85</td>
<td>Sensitivity= 45.8%
Specificity= 83.6%
Confusion Matrix:
<table>
<tr><td colspan="2">Predicted class</td><td></td><td></td></tr>
<tr><td>healthy</td><td>sick</td><td></td><td></td></tr>
<tr><td>51</td><td>10</td><td>healthy</td><td rowspan="2">Actual class</td></tr>
<tr><td>13</td><td>11</td><td>sick</td></tr>
</table></td></tr>
</table>

common approach to this situation is to combine those classifiers in an ensemble, hopefully reducing both the false-positive and false-negative rates [127]. Among our results (see Tables 9 and 10), logistic regression (with a sensitivity of 42.3% and specificity of 92.7%) has the lowest false-positive rate, while MetaRBFN1 (sensitivity 92.3%, specificity 63%) has the lowest false-negative rate, and MetaRBFN2 (sensitivity 73.1, specificity 78.8) gave us the closest point in ROC space to the optimum classifier. We will now examine ensembles using these three approaches as the base classifiers, using voting and stacking as the meta-classifiers, as implemented in WEKA.

For voting, two different experiments are done; the first experiment combines logistic regression and MetaRBFN1. In the second experiment, we add MetaRBFN2 to the previous combination. Table 12 provides the results of these experiments. In the stacking method, we use MetaRBFN1 and logistic regression as base classifiers with linear logistic regression as the meta-learner. Table 13 shows the results.

We next perform a slightly different experiment to combine different learning algorithms. We use the fact that there is usually a pair of diagnostic tests used for livestock disease surveillance. The first test is a screening test, where the farmer's efforts are focused on quarantining animals and making a disease-free zone in order to prevent spread of disease. Therefore, at first, a test with high sensitivity is performed, allowing the ones classified as healthy to be removed from further consideration. The second test has high specificity and is applied to the quarantined population in order to remove false positives [4]. For the experiment, we use MetaRBFN1 and logistic regression as the first and the second test, respectively. Table 14 shows the results of this experiment (the final sensitivity is 45.8%, with specificity of 83.6%).

Among the ensemble learning results, the voting method combining MetaRBFN1, MetaRBFN2, and Logistic Regression shows the best performance, with sensitivity of 80.8% and specificity of 80%. This is the closest to the optimum point in ROC space of any classifier we have examined in this work.

5. Threats to validity

In this section, we discuss the internal and external threats to validity in our experiments. In our view, the primary threat to internal validity is the seasonal weather change that occurred during the study period; our dataset for some cows runs from late February (still deep winter in Alberta) to May (after the spring thaw). This is, however, a normal and expected part of feedlot operations in North American climates, and all livestock disease management systems must be robust against it.

We believe there are three main threats to external validity in this study. Plainly, the most significant one is the small sample size. This unfortunately is unavoidable with only 31 prototype devices having been produced and passed quality assurance checks, and a limited period of data collection (due to operational constraints). We can expect that only a fraction of those devices will be placed on a cow that becomes sick. It is thus unknown if the sensitivity and specificity of our final design will generalize to other populations of cattle.

A second threat to external validity concerns regional effects. As noted previously, Alberta's winter climate is famously extreme; to the point where only a limited part of the North American cattle herd is exposed to such temperature variations. Again, while this is unavoidable, it calls into question how generalizable the sensitivity and specificity of our classifier will be.

A third threat concerns our reported false-positive rate. We determined the FPR by treating the producer's health intervention records as a ground truth; specifically, if there was no intervention recorded for the next k days, the cow is classified as "healthy" on this day. However, the evidence in our literature review is that the traditional, manual system employed by our producer partners misses a substantial number of BRD cases [102,128], meaning that some cows labelled "healthy" may in fact be sick. This will tend to increase the false-positive rate for our classifiers, and therefore decrease the observed specificity. We are unable to determine the degree to which this impacts our results, as we have no reports on observed lung quality at slaughter.

6. Conclusions

In this paper, we have conducted a case study of condition monitoring via inferential sensing: we developed an intelligent system for livestock disease surveillance, and tested it on a cattle feedlot in Alberta. We used sensor data measuring body and ambient temperature, and attempted to differentiate cows that would be sick in the next 7 days from those who would not. We applied a variety of machine-learning algorithms to both the raw data and multiple time-domain and wavelet-domain transformations of the data. Using a wavelet-domain ensemble classifier, we finally achieved a sensitivity of 80.8% and specificity of 80%.

In future work, we plan to fuse our temperature-based classifier with data from the feeding and watering sensors integrated on the platform, in order to further improve the sensitivity and specificity of our method. Existing research tells us that feeding behaviors are also useful signals of illness, and our own research found that signs of illness can be detected in watering behavior. Beyond these, practical issues will also need to be addressed before our system can be deployed in a feedlot or on a farm. For example, we will need to match the wireless technology in the collars to the dimensions of the animals' range (Zigbee radios are likely not a good choice on a free range farm). For the free-range context, additional location-monitoring technologies would be helpful in corralling a potentially sick animal faster. A more robust base station design is needed (the cows should *not* be able to eat the power or data cables), and stockmen will need to be trained in the mounting and maintenance of the sensor packages.

Acknowledgments

This research was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant no. RGPIN 262151, and in part by the Alberta Research Council/PRECARN, Inc. under grant no. SENG9023-2005.

References

[1] (2007-08-04). Foot-and-Mouth Symptom Guide. Available: http://www.fwi.co.uk/Articles/16/08/2012/105713/Foot-and-mouth-symptoms.htm.
[2] G. Acampora, D. Alghazzawi, H. Hagras, A. Vitiello, An interval type-2 fuzzy logic based framework for reputation management in peer-to-peer e-commerce, Inf. Sci. 333 (2016) 88–107.
[3] P.S. Addison (Ed.), The Illustrated Wavelet Transform Handbook: Introductry Theory, IOP, 2002.
[4] A.K. Akobeng, Understanding diagnostic tests 1: sensitivity, specificity and predictive values, Acta paediatrica 96 (2007) 338–341.
[5] M.M. Al-Rahhal, Y. Bazi, H. Al-Hichri, N. Alajlan, F. Melgani, R.R. Yager, Deep learning approach for active classification of electrocardiogram signals, Inf. Sci. 345 (2016) 340–354.
[6] M. Alsaaod, W. Büscher, Detection of hoof lesions using digital infrared thermography in dairy cows, J. Dairy Sci. 95 (2012) 735–742.
[7] M. Alsaaod, C. Syring, J. Dietrich, M.G. Doherr, T. Gujan, A. Steiner, A field trial of infrared thermography as a non-invasive diagnostic tool for early detection of digital dermatitis in dairy cows, Vet. J. 199 (2014) 281–285.
[8] D.E. Amrine, B.J. White, R. Larson, D.E. Anderson, D.A. Mosier, N. Cernicchiaro, Precision and accuracy of clinical illness scores, compared with pulmonary consolidation scores, in Holstein calves with experimentally induced Mycoplasma bovis pneumonia, Am J Vet Res 74 (2013) 310–315.
[9] P. Angelov, A. Kordon, Adaptive inferential sensors based on evolving fuzzy models, IEEE Trans. SMC, Part B 40 (2010) 529–539.
[10] Bisgaard, Time Series Analysis and Forecasting By Example, 2011 Available http://onlinelibrary.wiley.com/book/10.1002/9781118056943.
[11] C. Booth, J.R. McDonald, The use of artificial neural networks for condition monitoring of electrical power transformers, Neurocomputing 23 (1998) 97–109.

[12] J. Bouran, R. Eales, P. Thomas, D. Bostock, S. Lingard, I. Derbyshire, et al., The 2001 Outbreak of Foot and Mouth Disease: Report by the Comptroller and Auditor General, London, U.K., 2002.
[13] J. Bourn, R. Eales, P. Thomas, D. Bostock, S. Lingard, I. Derbyshire, et al., The 2001 Outbreak of Foot and Mouth Disease, National Audit Office, London, UK, 2002.
[14] S. Bowers, S. Gandy, B. Anderson, P. Ryan, S. Willard, Assessment of pregnancy in the late-gestation mare using digital infrared thermography, Theriogenology 72 (2009) 372–377.
[15] M. Brandt, A. Haeussermann, E. Hartung, Invited review: Technical solutions for analysis of milk constituents and abnormal milk, J. Dairy Sci. 93 (February) (2010) 427–436.
[16] K. Branson, S. Belongie, Tracking multiple mouse contours (without too many samples), Presented at the IEEE Conference on Computer Vision & Pattern Recognition, San Diego, CA, USA, 2005.
[17] B.J. Bricker, Diagnostic strategies used for the identification of Brucella, Vet. Microbiol. 90 (December) (2002) 433-4.
[18] P. Brown, Bovine spongiform encephalopathy and variant Creutzfeldt-Jakob disease, BMJ 322 (April) (2001) 841-4.
[19] M.J. Buhman, L.J. Perino, M.L. Galyean, T.E. Wittum, T.H. Montgomery, R.S. Swingle, Association between changes in eating and drinking behaviors and respiratory tract disease in newly arrived calves at a feedlot, Am. J. Vet. Res. 61 (2000) 1163–1168.
[20] H.S. Burkom, S. Murphy, J. Coberly, K. Hurt-Mullen, Analytic methods: public health monitoring tools for multiple data streams, Morbidity and Mortality Weekly Report 54 (2005) 55–62.
[21] CanadianFoodInspectionAgency. (2012-02-28). Fact sheet- foot and mouth disease. Available http://www.inspection.gc.ca/animals/terrestrial-animals/diseases/reportable/foot-and-mouth-disease/fact-sheet/eng/1330481689083/1330481803452.
[22] E.P. Carden, P. Fanning, Vibration based condition monitoring: a review, Struct. Health Monit. 3 (2004) 355–377.
[23] D. Cavero, K.H. Tolle, C. Henze, C. Buxade, J. Krieter, Mastitis detection in dairy cows by application of neural networks, Livestock Sci. 114 (April) (2008) 280–286.
[24] B. Cazelles, M. Chavez, G.C. de Magny, J.-F. Guégan, S. Hales, Time-dependent spectral analysis of epidemiological time-series with wavelets, J. R. Soc. Interface 4 (2007) 625–636.
[25] P. Chaovalit, A. Gangopadhyay, G. Karabatis, Z. Chen, Discrete wavelet transform-based time series analysis and mining, ACM Comput. Surv. (CSUR) 43 (2011) 6.
[26] C. Chatfield, The Analysis of Time Series: An Introduction, 6th ed., Chapman and Hall/CRC, Boca Raton, FL, 2004.
[27] C.L. Chentan Gupta, Song Wang, Abhay Mehta, Non-dyadic haar wavelets for streaming and sensor data, in: 26th International Conference on Data Engineering (ICDE), IEEE, 2010, pp. 569–580.
[28] J.W. Chi, A. Weersink, J.A. VanLeeuwen, G.P. Keefe, The economics of controlling infectious diseases on dairy farms, Can. J. Agric. Econ. Revue Canadienne D Agroeconomie 50 (November) (2002) 237–256.
[29] A. Cohen, J. Kovacevic, Wavelets: the mathematical background, Proc. IEEE 84 (April) (1996) 514–522.
[30] A. Colak, B. Polat, Z. Okumus, M. Kaya, L.E. Yanmaz, A. Hayirli, Early detection of mastitis using infrared thermography in dairy cows, J. Dairy Sci. 91 (2008) 4244–4248.
[31] N.J. Cook, B. Chabot, T. Lui, C.J. Bench, A.L. Schaefer, Infrared thermography detects febrile and behavioural responses to vaccination of weaned piglets, Animal 9 (2015) 339–346.
[32] R. De Mol, A. Keen, G. Kroeze, J. Achten, Description of a detection model for oestrus and diseases in dairy cattle based on time series analysis combined with a Kalman filter, Comput. Electron. Agric. 22 (1999) 171–185.
[33] M. Devant, S. Marti, A. Bach, Effects of castration on eating pattern and physical activity of Holstein bulls fed high-concentrate rations under commercial conditions, J. Anim. Sci 90 (2012) 4505–4513.
[34] S. Dick, C.C. Bracho, Learning to detect the onset of disease in cattle from feedlot watering behavior, in: NAFIPS 2007 - 2007 Annual Meeting of the North American Fuzzy Information Processing Society, 2007, pp. 112–116.
[35] P. Domingos, MetaCost: a general method for making classifiers cost-sensitive, in: Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1999, pp. 155–164.
[36] S. Duarte, L. Reig, J. Oca, Measurement of sole activity by digital image analysis, Aquacult. Eng. 41 (2009) 22–27.
[37] M.R. Dunbar, S.R. Johnson, J.C. Rhyan, M. McCollum, Use of infrared thermography to detect thermographic changes in mule deer (odocoileushemionus) experimentally infected with foot-and-mouth disease, J. Zoo Wildlife Med. 40 (2009) 296–301.
[38] E. Espada, H. Vijverberg, J. McLean, M. Sinclair, B. West, Milk colour analysis as a tool for the detection of abnormal milk, *First North American Conference on Robotic Milking, Toronto, Canada, 20–22 March 2002*., 2002.
[39] B.D. Eugenio, M. Glass, The kappa statistic: a second look, Comput. Ling. 30 (2004) 95–101.
[40] M.I.M. Fatma H. Elfouly, Moawad I.M. Dessouky, Salah Deyab, Comparision between Haar and Daubechies wavelet transformation on FPGA technology, Int. J. Comput. Inf. Eng. 2 (2008) 37–42.
[41] T. Fawcett, An introduction to ROC analysis, Pattern Recognit. Lett. 27 (2006) 861–874.
[42] D.F. Fegan, "Evaluation of diagnostic tests: the epidemiological approach," *DNA-based molecular diagnostic techniques. Research needs for standardization and validation of the detection of aquatic animal pathogens and diseases. FAO Fish Tech Pap*, vol. 395, pp. 30–37, 2000.
[43] A.W.-c.F. Franky Kin-Pong Chan, C. Yu, Haar wavelets for efficient similarity search of time series: with and without time wraping, IEEE Trans. Knowl. Data Eng. 15 (2003) 20.
[44] J.-D.F. Frederic Forge, Mad cow disease and Canada's cattle industry Available. http://www.parl.gc.ca/Content/LOP/ResearchPublications/prb0301-e.htm.
[45] B.A. Gardner, H.G. Dolezal, L.K. Bryant, F.N. Owens, R.A. Smith, Health of finishing steers: effects on performance, carcass traits, and meat tenderness, J Anim. Sci 77 (1999) 3168–3175.
[46] J. Gloster, K. Ebert, S. Gubbins, J. Bashiruddin, D.J. Paton, Normal variation in thermal radiated temperature in cattle: implications for foot-and-mouth disease detection, BMC Vet. Res. 7 (2011) 73.
[47] J. Gloster, K. Ebert, S. Gubbins, J. Bashiruddin, D.J. Paton, Normal variation in thermal radiated temperature in cattle: implications for foot-and-mouth disease detection, BMC Vet. Res. 7 (2011) 10.
[48] M.J. Grubman, B. Baxt, Foot-and-mouth disease, Clin. Microbiol. Rev. 17 (April) (2004) 465–493.
[49] D.B. Haley, D.W. Bailey, J.M. Stookey, The effects of weaning beef calves in two stages on their behavior and growth rate, J. Anim. Sci 83 (2005) 2205–2214.
[50] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I.H. Witten, The WEKA data mining software: an update, ACM SIGKDD Explor. Newsl. 11 (2009) 10–18.
[51] B.L. Hart, Biological basis of the behavior of sick animals, Neurosci. Biobehav. Rev. 12 (1988) 123–137.
[52] M. Hashemzadeh, N. Farajzadeh, Combining keypoint-based and segment-based features for counting people in crowded scenes, Inf. Sci. 345 (2016) 199–216.
[53] H. Hogeveen, W. Ouweltjes, Sensors and management support in high-technology milking, J. Anim. Sci. 81 (2003) 1–10.
[54] M. Hovinen, J. Siivonen, S. Taponen, L. Hänninen, M. Pastell, A.-M. Aisla, et al., Detection of clinical mastitis with the help of a thermal camera, J. Dairy Sci. 91 (2008) 4592–4598.
[55] J. Hyndman, G. Athanasopoulos, Forecasting: principles and practice Available. https://www.otexts.org/book/fpp.
[56] T. Infernuso, C.A. Loughin, D.J. Marino, S.E. Umbaugh, P.S. Solt, Thermal imaging of normal and cranial cruciate ligament-deficient stifles in dogs, Vet. Surg. 39 (2010) 410–417.

[57] S.R. Johnson, S. Rao, S.B. Hussey, P.S. Morley, J.L. Traub-Dargatz, Thermographic eye temperature as an index to body temperature in ponies, J. Equine Vet. Sci. 31 (2011) 63–66.
[58] B. Joseph, C. Brosilow, Inferential control of processes, AIChE J. 24 (1978) 485–509.
[59] A. Jutan, J. MacGregor, J. Wright, Multivariable computer control of a butane hydrogenlysis reactor, part II - data collection, parameter estimation, and stochastic disturbance identification, AIChE J. 23 (1977) 742–750.
[60] H. Kamohara, A. Takinami, M. Takeda, M. Kano, S. Hasebe, I. Hashimoto, Product quality estimation and operating condition monitoring for industrial ethylene fractionator, J. Chem. Eng. Jpn. 37 (2004) 422–428.
[61] C. Kamphuis, H. Mollenhorst, A. Feelders, D. Pietersma, H. Hogeveen, Decision-tree induction to detect clinical mastitis with automatic milking, Comput. Electron. Agric. 70 (January) (2010) 60–68.
[62] C. Kamphuis, H. Mollenhorst, J. Heesterbeek, H. Hogeveen, Detection of clinical mastitis with sensor data from automatic milking systems is improved by using decision-tree induction, J. Dairy Sci. 93 (2010) 3616–3627.
[63] F. Karakus, A. Duzgun, M. Karakus, L. Aslan, Can infrared thermography be used to predict ear tag infections in lambs? Sci. Pap. Ser. D. Anim. Sci. 58 (2015) 205–208.
[64] E. Kramer, D. Cavero, E. Stamer, J. Krieter, Mastitis and lameness detection in dairy cows by application of fuzzy logic, Livestock Sci. 125 (October) (2009) 92–96.
[65] S. Kruse, I. Traulsen, J. Krieter, Analysis of water, feed intake and performance of lactating sows, Livestock Sci. 135 (February) (2011) 177–183.
[66] S. Kruse, I. Traulsen, J. Salau, J. Krieter, A note on using wavelet analysis for disease detection in lactating sows, Comput. Electron. Agric. 77 (June) (2011) 105–109.
[67] R.E. Lamberson, "Apparatus and method for the remote monitoring of machine condition," USA Patent, 1998.
[68] J.P. LeSage, Econometrics toolbox for matlab. (Accessed April 28, 2005) Internet site: http://www.spatial-econometrics.com/ 2009.
[69] Y. Liu, Feature extraction and dimensionality reduction for mass spectrometry data, Comput. Biol. Med. 39 (2009) 818–823.
[70] Y. Liu, Dimensionality reduction and main component extraction of mass spectrometry cancer data, Knowl.-Based Syst. 26 (2012) 207–215.
[71] Z. Long, G. Liu, X. Dai, Extracting Emotional Features from ECG by Using Wavelet Transform, in: Biomedical Engineering and Computer Science (ICBECS), 2010 International Conference on, 2010, pp. 1–4.
[72] J.M. Lukas, J.K. Reneau, R. Wallace, D. Hawkins, C. Munoz-Zanzi, A novel method of analyzing daily milk production and electrical conductivity to predict disease onset, J. Dairy Sci. 92 (Dec) (2009) 5964–5976.
[73] O.Z. Maimon, L. Rokach, Data Mining and Knowledge Discovery Handbook, Springer, 2005.
[74] S.G. Mallat, A Theory for Multiresolution Signal Decomposition - the Wavelet Representation, IEEE Trans. Pattern Anal. Mach. Intell. 11 (Jul) (1989) 674–693.
[75] C.H. Martin Vetterli, Wavelets and filter banks: theory and design, IEEE Trans. Signal Process. 40 (1992) 26.
[76] P. Martiskainen, M. Järvinen, J.-P. Skön, J. Tiirikainen, M. Kolehmainen, J. Mononen, Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines, Appl. Anim. Behav. Sci. 119 (2009) 32–38.
[77] R. McCorkell, K. Wynne-Edwards, C. Windeyer, A. Schaefer, et al., Limited efficacy of Fever Tag® temperature sensing ear tags in calves with naturally occurring bovine respiratory disease or induced bovine viral diarrhea virus infection, Can. Vet. J. 55 (2014) 688–690.
[78] S.R.O. Menegassi, J.O.J. Barcellos, E.A. Dias, C. Koetz Jr., G.R. Pereira, V. Peripolli, et al., Scrotal infrared digital thermography as a predictor of seasonal effects on sperm traits in Braford bulls, Int. J. Biometeorol. 59 (2015) 357–364.
[79] A. Menzel, M. Beyerbach, C. Siewert, M. Gundlach, D. Hoeltig, R. Graage, et al., Actinobacillus pleuropneumoniae challenge in swine: diagnostic of lung alterations by infrared thermography, BMC Vet. Res. 10 (2014) 13.
[80] B. Miekley, I. Traulsen, J. Krieter, Detection of mastitis and lameness in dairy cows using wavelet analysis, Livestock Sci. 148 (Oct) (2012) 227–236.
[81] M. Misiti, Y. Misiti, G. Oppenheim, J. Poggi, Wavelet Toolbox, The MathWorks Inc., Natick, MA, 1996.
[82] J. Moon, H. Koo, Y. Joo, S. Jeon, D. Hur, C. Chung, et al., Application of a new portable microscopic somatic cell counter with disposable plastic chip for milk analysis, J. Dairy Sci. 90 (2007) 2253–2259.
[83] K. Okada, K. Takemura, S. Sato, Investigation of various essential factors for optimum infrared thermography, *J. Vet. Med. Sci*. 75 (2013) 1349–1353.
[84] M. Pastell, J. Tiusanen, M. Hakojarvi, L. Hanninen, A wireless accelerometer system with wavelet analysis for assessing lameness in cattle, Biosyst. Eng. 104 (Dec 2009) 545–551.
[85] M.E. Pastell, M. Kujala, A probabilistic neural network model for lameness detection, J. Dairy Sci. 90 (May 2007) 2283–2292.
[86] D. Peña, G.C. Tiao, R.S. Tsay, A Course in Time Series Analysis, 322, John Wiley & Sons, 2011.
[87] D.B. Percival, A.T. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, New York, 2000.
[88] A.C. Pérez de Diego, P.J. Sánchez-Cordón, M. Pedrera, B. Martínez-López, J.C. Gómez-Villamandos, J.M. Sánchez-Vizcaíno, The use of infrared thermography as a non-invasive method for fever detection in sheep infected with bluetongue virus, Vet. J. 198 (2013).
[89] B. Polat, A. Colak, M. Cengiz, L.E. Yanmaz, H. Oral, A. Bastan, et al., Sensitivity and specificity of infrared thermography in detection of subclinical mastitis in dairy cows, J. Dairy Sci. 93 (2010) 3525–3532.
[90] R. Polikar, The wavelet tutorial.
[91] S.J. Qin, H. Yue, R. Dunia, Self-validating inferential sensors with application to air emission monitoring, Ind. Eng. Chem. Res. 36 (1997) 1675–1685.
[92] W.F. Quimby, B.F. Sowell, J.G.P. Bowman, M.E. Branine, M.E. Hubbert, H.W. Sherwood, Application of feeding behaviour to predict morbidity of newly received calves in a commercial feedlot, Can. J. Anim. Sci. 81 (2001) 315–320.
[93] K. Rainwater-Lovett, J.M. Pacheco, C. Packer, L.L. Rodriguez, Detection of foot-and-mouth disease virus infected cattle using infrared thermography, Vet. J. 180 (2009) 317–324.
[94] S.I. Rekant, M.A. Lyons, J.M. Pacheco, J. Arzt, L.L. Rodriguez, Veterinary applications of infrared thermography, Am. J. Vet. Res. 77 (2016) 98–107.
[95] B. Robert, B.J. White, D.G. Renter, R.L. Larson, Evaluation of three-dimensional accelerometers to monitor and classify behavior patterns in cattle, Comput. Electr. Agric. 67 (2009) 80–84.
[96] A.R. Rodríguez, F.J. Olivares, P.T. Descouvieres, M.P. Werner, N.A. Tadich, H.A. Bustamante, Thermographic assessment of hoof temperature in dairy cows with different mobility scores, Livestock Sci. 184 (2016) 92–96.
[97] T.K. Rose-Dye, L.O. Burciaga-Robles, C.R. Krehbiel, D.L. Step, R.W. Fulton, A.W. Confer, et al., Rumen temperature change monitored with remote rumen temperature boluses after challenges with bovine viral diarrhea virus and Mannheimia haemolytica, J. Anim. Sci. 89 (2011) 1193–1200.
[98] Y.-F. Sang, D. Wang, J.-C. Wu, Entropy-based method of choosing the decomposition level in wavelet threshold de-noising, Entropy 12 (2010) 1499–1513.
[99] A.L. Schaefer, N.J. Cook, J.S. Church, J. Basarab, B. Perry, C. Miller, et al., The use of infrared thermography as an early indicator of bovine respiratory disease complex in calves, Res. Vet. Sci. 83 (2007) 376–384.
[100] P. Scheunders, S. Livens, G. Van de Wouwer, P. Vautrot, and D. Van Dyck, "Wavelet-based texture analysis."
[101] J.B. Schleppe, G. Lachapelle, C.W. Booker, T. Pittman, Challenges in the design of a GNSS ear tag for feedlot cattle, Comput. Electron. Agri. 70 (2010) 84–95.
[102] M.J. Schneider, R.G. Tait Jr., W.D. Busby, J.M. Reecy, An evaluation of bovine respiratory disease complex in feedlot cattle: impact on performance and carcass traits using treatment records and lung lesion scores, J. Anim. Sci. 87 (2009) 1821–1827.
[103] J.W. Schroeder, Mastitis Control Programs: Bovine Mastitis and Milking Management, North Dakota State University Fargo, North Dakota, Extension Service, 1997 Revised and Reprint July 2012.
[104] J.A. Small, A.D. Kennedy, S.H. Kahane, Core body temperature monitoring with passive transponder boluses in beef heifers, Can. J. Anim. Sci. 88 (2008) 225–235.

[105] M. Soroko, R. Henklewski, H. Filipowski, E. Jodkowska, The effectiveness of thermographic analysis in equine orthopedics, J. Equine Vet. Sci. 33 (2013) 760–762.
[106] B.F. Sowell, M.E. Branine, J.G. Bowman, M.E. Hubbert, H.E. Sherwood, W. Quimby, Feeding and watering behavior of healthy and morbid steers in a commercial feedlot, J. Anim. Sci. 77 (1999) 1105–1112.
[107] O. Staff, Bovine respiratory disease: new research to tackle major concern for cattle industry, Science Daily, Science Daily, Rockville, MD, USA, 2010.
[108] StatisticsCanada. *Canada's beef industry and BSE.* Available: http://www41.statcan.ca/2006/0920/ceb0920_001-eng.htm.
[109] A.N. Su Chen, Dynamic nonuniform data aproximation in databases with haar wavelet, J. Comput. 2 (2007) 67–76.
[110] A. Subasi, Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients, Exp. Syst. Appl. 28 (2005) 701–711.
[111] S. Talukder, G. Gabai, P. Celi, The use of digital infrared thermography and measurement of oxidative stress biomarkers as tools to diagnose foot lesions in sheep, Small Ruminant Res. 127 (2015) 80–85.
[112] X.Z. Tan Minsheng, Wang Lei, Voronoi tessellation based haar wavelet data compression for sensor network, in: International Conference on Wireless Communications, Networking and Mobile Computing, WiCOM, 2006, pp. 1–4.
[113] Q.L. Tao Li, Shenghuo Zhu, Mitsunori Ogihara, A survey on wavelet applications in data mining, SIGKDD Explorations 4 (2003) 49–68.
[114] J.D. Taylor, R.W. Fulton, T.W. Lehenbauer, D.L. Step, A.W. Confer, The epidemiology of bovine respiratory disease: what is the evidence for predisposing factors? Can. Vet. J. 51 (2010) 1095–1102.
[115] M.E. Theurer, D.E. Amrine, B.J. White, Remote noninvasive assessment of pain and health status in cattle, Vet. Clinics North Am. 29 (2013) 59–74.
[116] M.E. Theurer, B.J. White, J.F. Coetzee, L.N. Edwards, R.A. Mosher, C.A. Cull, Assessment of behavioral changes associated with oral meloxicam administration at time of dehorning in calves using a remote triangulation device and accelerometers, *BMC Vet. Res.* 30 (2012) 8.
[117] D. Thompson, P. Muriel, D. Russell, P. Osborne, A. Bromley, M. Rowland, et al., Economic costs of the foot and mouth disease outbreak in the United Kingdom in 2001, Revue Scientifique et Technique-Office International des Epizooties 21 (2002) 675–685.
[118] E. Timsit, S. Assié, R. Quiniou, H. Seegers, N. Bareille, Early detection of bovine respiratory disease in young bulls using reticulo-rumen temperature boluses, Vet. J. 190 (2011) 136–142.
[119] T. Travain, E.S. Colombo, E. Heinzl, D. Bellucci, E.P. Previde, P. Valsecchi, Hot dogs: thermography in the assessment of stress in dogs (Canis familiaris) - a pilot study, J. Vet. Behav. 10 (2015) 17–23.
[120] M.G. Trotter, D.W. Lamb, G.N. Hinch, C.N. Guppy, Global navigation satellite system livestock tracking: system development and data interpretation, *Anim.* Prod. Sci. 50 (2010) 616–623.
[121] M. Tscharke, T.M. Banhazi, A brief review of the application of machine vision in livestock behaviour analysis, J. Agri. Inf. 7 (2016) 23–42.
[122] A. Tsymbal, The Problem of Concept Drift: Definitions and Related Work, Computer Science Department, Trinity College Dublin, 2004.
[123] B. Voss, H.-J. Laue, M. Hoedemaker, S. Wiedemann, Field-trial evaluation of an automatic temperature measurement device placed in the reticulo-rumen of pre-weaned male calves, Livestock Sci. 189 (2016) 78–81.
[124] B.J. White, D.E. Anderson, D.G. Renter, R.L. Larson, D.A. Mosier, L.L. Kelly, et al., Clinical, behavioral, and pulmonary changes in calves following inoculation with Mycoplasma bovis, Am. J. Vet. Res. 73 (2012) 490–497.
[125] B.J. White, D.G. Renter, Bayesian estimation of the performance of using clinical observations and harvest lung lesions for diagnosing bovine respiratory disease in post-weaned beef calves, J. Vet. Diagn. Invest. 21 (2009) 446–453.
[126] D. Whyte, R. Orchard, P. Cross, T. Frietsch, R. Claycomb, G. Mein, An on-line somatic cell count sensor, in: A. Meijering, H. Hogeveen, C.J.A.M. de Koning (Eds.), Automatic Milking: A Better Understanding, Wageningen Academic Publishers, Wageningen, the Netherlands, 2004, pp. 235–240.
[127] I.H. Witten, E. Frank, M.A. Hall, Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques, Elsevier, 2011.
[128] T.E. Wittum, N.E. Woollen, L.J. Perino, E.T. Littledike, Relationships among treatment for respiratory tract disease, pulmonary lesions evident at slaughter, and rate of weight gain in feedlot cattle, J. Am. Vet. Med. Assoc. 209 (1996) 814–818.
[129] B. Wolfger, E. Timsit, B.J. White, K. Orsel, A systematic review of bovine respiratory disease diagnosis focused on diagnostic confirmation, early detection, and prediction of unfavorable outcomes in feedlot cattle, Vet. Clin. North Am. 31 (2015) 351–365.
[130] Y.-C. Wu, Language independent web news extraction system based on text detection framework, Inf. Sci. 342 (2016) 132–149.
[131] S.D.Y. Zhou, Toward identification of sick feedlot cattle using sensor data, Presented at the World Conference on Soft Computing, SanFrancisco state university, 2011.
[132] J. Zhao, S. Xiao, Classification of evoked potentials using wavelet coefficient features, in: Engineering in Medicine and Biology Society, Proceedings of the 19th Annual International Conference of the IEEE, 1997, pp. 1496–1499.
[133] I. Zliobaite, Learning Under Concept Drift: An Overview, Vilnius University Technical report, , 2009 techniques, related areas, applications. Subjects: Artificial Intelligence 2009.
[134] A.-M. Zou, A.H.J. de Ruiter, K.D. Kumar, Distributed attitude synchronization control for a group of flexible spacecraft using only attitude measurements, Inf. Sci. 343-344 (2016) 66–78.