Continuous Low-Power Ammonia Monitoring Using Long Short-Term Memory Neural Networks

Zhenhua Jia
zhenhua@winlab.rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

Xinmeng Lyu
xinmeng.lyu@rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

Wuyang Zhang
wuyang@winlab.rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

Richard P. Martin
rmartin@scarletmail.rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

Richard E. Howard
reh@winlab.rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

Yanyong Zhang
yyzhang@winlab.rutgers.edu
Wireless Information Network
Laboratory
Rutgers University, USA

ABSTRACT

Accurate and continuous ammonia monitoring is important for laboratory animal studies and many other applications. Existing solutions are often expensive, inaccurate, or unsuitable for long-term monitoring. In this work, we propose a new ammonia monitoring approach that is low-power, automatic, accurate, and wireless.

Our system uses metal oxide sensors which change their electrical resistance due to an induced reduction reaction with ammonia at high temperatures. Traditional methods infer the ammonia level by measuring the sensor's electrical resistance after it reaches equilibrium. Such a system consumes a significant amount of energy because reaching equilibrium requires heating the sensor for minutes. Our proposed approach does not wait for equilibrium, but tries to predict the resistance at equilibrium using the sensor's initial resistance response curve in a very short heating pulse (as short as $200ms$). The prediction model is built on long short-term memory (LSTM) neural networks.

We built 38 prototype sensors and a home-grown gas flow system. In a 3-month in-lab testing period, we conducted extensive experiments and collected 13,770 measurements. Our model accurately predicts the equilibrium state resistance value, with an average error rate of 0.12%. The final average estimation error for the ammonia concentration level is $9.38ppm$. Given the ultra low power consumption and accurate measurements, we have partnered with cage vendors and deployed our system at two animal research facilities (NIH and Cornell University) for month-long medical trials.

CCS CONCEPTS

• **Computer systems organization** → **Embedded systems**; • **Hardware** → **Power estimation and optimization**; *Robustness*;

KEYWORDS

Ammonia Sensor, Low-power, Transient Response, Long Short-term Memory, Neural Networks

1 INTRODUCTION

Rodents play a significant role in animal testing. For example, the estimated number of rodents in U.S. laboratories alone is around 80 million in 2001 [6]. Unfortunately, researchers have recently raised the concern of irreproducible rodent-based experiments due to varying environmental factors [14, 22, 33, 40].

Among these factors is ammonia (NH_3) that has important impacts on both rodent health and research outcomes. As a reactive and sticky molecule, it easily accumulates to toxic levels. Studies show that prolonged exposure to ammonia of 50 parts per million (ppm) concentration for a 2-week period can cause epithelial degeneration in mice [41]. The standard of the care and use of laboratory animals defined by the U.S national research council [9] lists ammonia concentration as one of the most important factors while deciding the cycle of cage changes. While infrequent cage changes may lead to high ammonia levels, too frequent cage changes are also undesirable for two main reasons. First, it can adversely impact animals' physical health (such as cardiovascular activity [13]) and can induce stress as their nesting sites are disturbed [4, 7, 29, 31]. Second, cage changing is a laborious and expensive procedure: direct labor (such as changing cages, food, water and bedding), materials, and administrative costs account for up to 50% of the total animal care costs [30].

Considering a regular laboratory animal center may have hundreds of cages, we believe an automatic and continuous ammonia
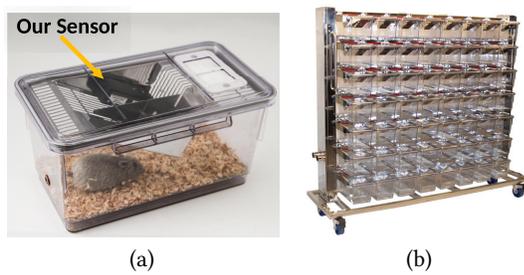
Figure 1: (a) A cage with our ammonia monitoring system and (b) a rack holding 56 cages. Our system can support several such racks simultaneously, which is suitable for animal research facilities.

Figure 2: The metal oxide sensor has three layers from bottom to top: the diaphragm layer, the heating layer and the sensing layer.

monitoring system is essential to ensure the health of the laboratory animals, lower the maintenance cost, and further improve the reproducibility of the experiments.

Unfortunately, ammonia sensing techniques haven't changed much in the last few decades. Current solutions include common ammonia test strips, test liquid tubes [11], metal oxide sensors [37, 45, 46], electrical-chemical sensors [21, 38, 42], and fiber-coupled optical sensors [27]. None of these techniques offers automatic and continuous monitoring.

In this study, we seek to address this challenge by building such an ammonia monitoring system using metal oxide sensors. A metal oxide sensor is made through a quantum tunneling technique [32], and thus reusable. A typical metal oxide sensor has a reduction reaction with ammonia at high temperatures (usually at a few hundred ℃) and an oxidization reaction with oxygen even at room temperature. The reduction reaction coverts metal oxide to metal, while the oxidization reaction returns the metal back into metal oxide. People then measure the sensor's electrical resistance at the chemical equilibrium state and map the resistance value to the corresponding ammonia concentration level. There are several major hurdles we need to overcome to make metal oxide sensors suitable for continuous ammonia sensing. First, metal oxide sensors are power hungry – it has to maintain a high temperature to keep the reduction reaction going for minutes until a chemical equilibrium is reached. Second, metal oxide sensors have dramatically different sensitivity in terms of their responses to ammonia gas.

To solve these challenges, we devise an accurate prediction model based on Long short-term memory (LSTM) neural networks [16, 19, 20]. The LSTM networks focus on the sensor's transient resistance measurements in a short time window (e.g., a window of 0.2s) and can accurately predict the final resistance value at the chemical equilibrium state which may take minutes to achieve. Our system precisely controls the heating power and duration, and accurately measures the ammonia sensor's transient resistance values in analog-to-digital converter (ADC) samples. Meanwhile, the system sends the ADC samples wirelessly to a remote Raspberry Pi for processing.

LSTM neural networks are one type of recurrent neural networks (RNNs), which are developed to deal with the exploding/vanishing gradient problem when training traditional RNNs. By adding the concept of a memory state, LSTM networks can learn the dependencies from time-series data, hence well suited for our problem. Our
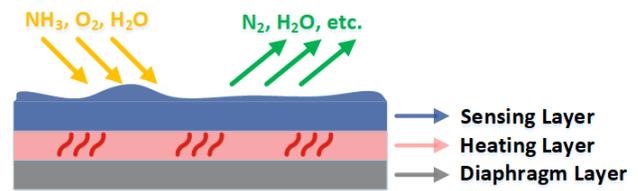
LSTM based model can perform accurate prediction from only a few samples that are collected from a couple hundreds of milliseconds. Significant power saving can be achieved in this way. In fact, our model only needs to heat the sensor for 0.2s (during this period, our sensor samples at 40Hz), which cuts down about 99.6% of the total energy from the usual measurement method that has to heat the sensor for 100s. With this improvement, our system can make a measurement once every 3 hours for at least 20 years using a single 3.6V Tadiran AA battery [5]. Also, our system is compact enough to fit into any cage and can support continuous ammonia monitoring over racks of cages, as shown in Figure 1. Finally, our wireless design demands little additional effort to routine cage changes.

We have built a home-grown ammonia gas flow system and conducted 13,770 measurements using 38 ammonia sensors within a 3-month period. While conducting measurements, we have varied the ammonia concentration from 0ppm to 240ppm. Our model proves to be very precise: across different ammonia concentrations, the average prediction error rate for the equilibrium state resistance ADC value is 0.12% and the average absolute estimation error for the ammonia concentration is 9.38ppm. Also, we can train our model using one sensor's data, test the model using data from other sensors, and still achieve equally accurate prediction results.

We deployed our ammonia monitoring system into two animal research labs with a total of 38 rodent cages and completed a 4-month trial at the National Institutes of Health (NIH) and a second 6-month breeding-related trial at Cornell University. The trials prove that our system indeed offers a viable solution for accurate and continuous ammonia monitoring for large-scale laboratory animal facilities. The observed ammonia level changes accurately reflect the cage change events logged by the facility staff. Finally, we note that our ammonia sensing system can be readily used in other application areas that require ammonia sensing, including air quality in other animal care locations, such as stables, and ammonia leak detection in industrial and factory settings. A thorough survey of ammonia sensing applications can be found in [39].

Our work has the following main contributions:

(1) We built a low-power, automatic, accurate and wireless ammonia monitoring system. The system easily fits into any rodent cage and can last for 20 years with a single Tadiran AA battery.

(2) We have developed a prediction model based on LSTM neural networks that accurately estimated the equilibrium state resistance value given a few transient resistance samples collected within the first 0.2s. To our best knowledge, this is
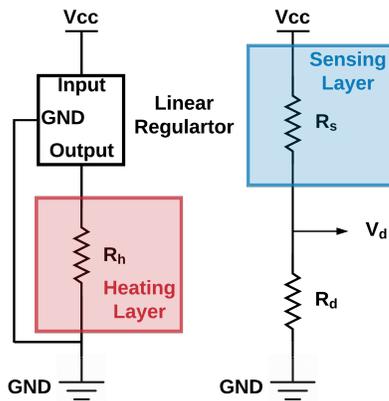
Figure 3: Our sensor circuit design. The entire circuit is powered by 2 AA lithium battery with $V_{cc} = 3V$. R_h is a heating resistor powered by a $2.2V$ linear regulator. R_s is the metal oxide resistor. R_d is a voltage divider in-series with R_s. We infer the resistance of R_s through measuring the voltage drop V_d across R_d.

Figure 4: Our ammonia sensor prototype. The sensor is compact and can be powered by either 2 typical AA lithium batteries or a $3.6V$ Tadiran AA lithium battery.

the first research work showing two orders of magnitude reduction in energy consumption of metal oxide sensor-based ammonia measurement.

(3) We built a home-grown ammonia gas flow system and conducted many measurements over long time periods. We partnered with cage manufacturers and deployed our system into two animal research laboratories[1] and completed two field trials.

2 BACKGROUND AND MOTIVATION

2.1 Metal Oxide Based Ammonia Monitoring

A metal oxide sensor is a reusable sensor for measuring the ammonia level. At high temperature, it leads to a reduction reaction with combustible gases such as ammonia. This type of reaction converts metal oxide to metal that has much smaller electrical resistance. Meanwhile, metal may have oxidation reaction by absorbing oxygen on the surface and forms metal oxide, even at room temperature [10]. As a result, by measuring how the electrical resistance changes, we can infer the ammonia level.

Shown in Figure 2, the sensor consists of three layers from bottom to top: a machined diaphragm base layer, an embedded heating layer and a sensing layer. When making a measurement, we first heat the sensor to a few hundred ℃. Next, we keep the temperature constant, let metal oxide on the sensing layer have reduction-oxidation reaction (redox) with ammonia and oxygen simultaneously, and wait till the redox process reaches its chemical equilibrium. When the chemical equilibrium is reached, the rate of the metal oxide being reduced by ammonia is equal to the rate of the metal being oxidized by oxygen. Finally, we measure the resistance of the sensing layer through a 10-bit ADC on the microprocessor and convert the ADC sample to the corresponding

ammonia concentration, as shown in Figure 3. The entire measurement is rather time-consuming as it might take a few minutes to reach the chemical equilibrium. Keeping the sensor's temperature steady at a few hundred ℃ for this period of time can be quite power consuming.

2.2 Motivation for Transient-Predict

In this study, we propose an automatic ammonia monitoring system based on a metal oxide sensor. Thanks to our low-power design, the ammonia monitoring system is compact and can be easily put into a regular rodent cage. The system automatically measures the ammonia concentration inside the cage without any additional human effort. Below we highlight the salient features of our proposed system.

Low Power Ammonia Measurement: The metal oxide sensor demands high temperature over a few minutes in order to trigger and keep the reduction reaction. This may consume a significant amount of power. In fact, heating alone costs more than 99% of the total amount of energy for a 100-second heating period on average. This is the exact reason why the current metal oxide ammonia measurement tools operate on large batteries and are usually hand-held devices.

In this work, we address this challenge by significantly reducing the amount of energy required for each measurement. One of our main contributions is to design a prediction model which can greatly shorten the time required for measurement. Our approach takes the transient ADC samples collected in the first $0.2s$ and can accurately predict the ADC measurement in a few minutes. We thus refer to the proposed approach as Transient-Predict. Transient-Predict consumes much less power and requires a much smaller battery, so it can be made compact enough to fit into a standard cage and provide continuous wireless monitoring for years, as shown in Figure 1(a).

Accurate Prediction of the Equilibrium Resistance: Our approach needs to predict the ADC value in the equilibrium state (which usually takes a few minutes to arrive at) from the first few transient ADC samples collected in less than a second. This challenge is made even harder by the fact that each metal oxide sensor has drastically different characteristics. In fact, metal oxide sensors are made by the quantum tunneling technique [32] and the growth
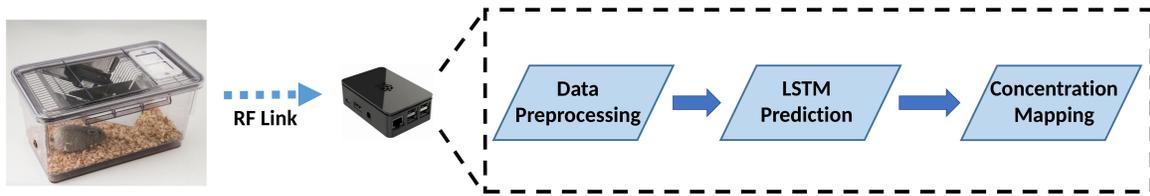
Figure 5: The overview of Transient-Predict.

of metal oxide on the sensing layer is hard to control. As such, the sensitivity of the sensors varies, sometimes by a factor of 10 [34].

Further, the process of reaching the chemical equilibrium is impacted by several factors: the initial state of the sensing layer (such as the percentage of metal in the form of metal oxide and the amount of ammonia stuck to the surface), ammonia concentration in the air, oxygen concentration, humidity level, heating temperature, ambient temperature, etc. Considering these factors one by one in a prediction model can be an onerous task as each factor is non-linear with respect to the ammonia concentration level.

In this work, we solve this challenge by developing suitable LSTM neural networks to learn the relationship between transient ADC values and the final equilibrium state value. We will present the design detail in Section 3.

2.3 Prototyping and Power Consumption Profiling of Transient-Predict

We have built 38 prototype ammonia sensors. We integrate an SGX Sensortech MiCS-5914 metal oxide sensor [34] with a general-purpose embedded platform (called PIP tags [17]) which has a Texas Instrument (TI) MSP430 microcontroller with a 10-bit ADC and a TI CC1100 radio transceiver module. PIP tags have been used in multiple applications, such as wireless passive indoor localization [43, 44], motion detection [2, 3], vital sign detection [23, 24], wireless energy transfer [15], etc. Shown in Figure 4, the dimension of our sensor is $9cm \times 3.5cm \times 2cm$. The sensor is protected with a filter paper and powered by two AA lithium batteries[2]. Our sensor measures the electrical resistance value through ADC, and the ADC samples are wirelessly transmitted to a receiver that is connected to a Raspberry Pi or a computer for subsequent processing. Our signal processing mainly consists of three steps: (1) preprocessing, (2) predicting the equilibrium resistance ADC value using a few transit resistance ADC values, and (3) converting the predicted equilibrium resistance value to the corresponding ammonia concentration level. The overview of our Transient-Predict system is shown in Figure 5.

Dominance of ammonia in a rodent cage: According to the sensor datasheet [34], the metal oxide sensors are reactive to some gases (i.e., ammonia, ethanol and hydrogen) at low concentration (starting around $1ppm$) and other gases (i.e., propane and iso-butane) at high concentration (above $1000ppm$). Considering the environment inside a rodent cage, the reaction between ammonia and the sensor dominates overall sensor reading, as the concentrations of other reactive gases are much lower than ammonia.

[2]We use two AA lithium batteries for the purpose of collecting ground truth. For our own measurement, a single Tadiran AA battery is sufficient.

Energy Profiling: Before presenting our detailed design, we first profile the power consumption of our sensor in both the sleep stage and the measurement stage. We mount a fixed resistor in series with our ammonia sensor and a direct current (DC) power supply, and use the voltage drop across the fixed resistor to infer the power consumption of our ammonia sensor. We further report the profiling results of our system and the traditional minutes heating approach in Table 1. After sleeping for 3 hours, our sensor wakes up for $0.2s$. During the measurement period, the sensor heats up, takes ADC samples at $40Hz$, and transmits the samples to the receiver. After $0.2s$, the sensor goes back to sleep. Our system consumes $0.64\mu A$ during the sleep stage and about $30.7mA$ for each measurement. As a result, our system saves 99.6% of energy compared to the traditional approach and can last for at least 20 years with a Tadiran AA battery, assuming it makes one measurement in every 3 hours.

3 DESIGN OF TRANSIENT-PREDICT

In this section, we present the detailed design of Transient-Predict, which is centered around predicting the equilibrium resistance value using a few transient resistance values through LSTM neural networks.

3.1 Overview of Transient-Predict

The *baseline* ammonia sensing approach involves heating the metal oxide sensor long enough (usually for a few minutes) such that the chemical reaction on the sensor reaches the equilibrium. The main drawback of this approach is the high energy requirement in heating the sensor. On the contrary, Transient-Predict only needs to heat the sensor for a very short duration – say, a few hundred milliseconds. During this duration, we sample the transient ADC values. Then we predict the final ADC value in the chemical equilibrium state

Stage	Our approach		Traditional Approach	
	Power (μA)	Time (s)	Power (μA)	Time (s)
Sleep	0.64	10,799.8	0.64	10,700
Measurement	30,700	0.2	30,670	100
Avg Power Consumption	$0.0012mA$		$0.2847mA$	

Table 1: The power consumption profiling for our Transient-Predict system. The system consumes $0.64\mu A$ during sleep and $30,700\mu A$ for a measurement. Compared to the traditional minutes heating approach, our system saves about 99.6% of energy. On average, our system can last for at least 20 years with one Tadiran AA lithium battery ($2400mAh$), assuming it makes one measurement in every 3 hours.
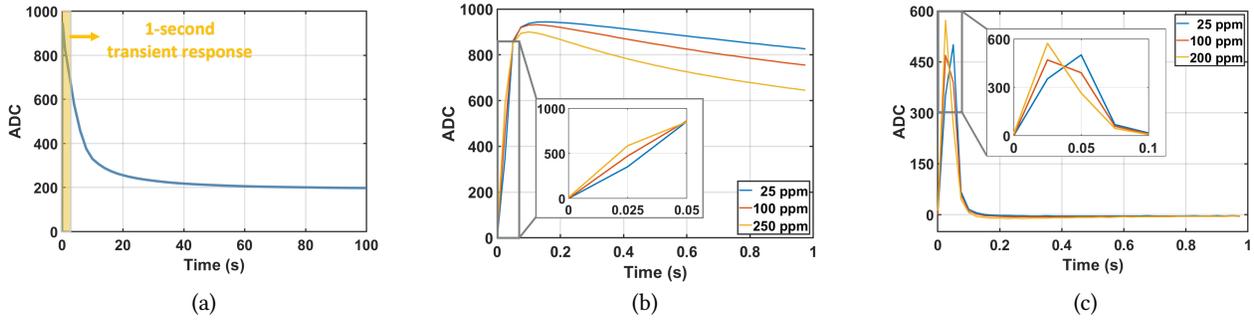
Figure 6: (a) A 100-second ADC trace until the chemical equilibrium is reached. The yellow shade marks the 1-second transient period we work with in Transient-Predict. (b) A sensor's 1-second transient ADC samples under ammonia concentration of 25, 100, 200*ppm***. (c) The first derivatives of the same transient samples as in (b). Each derivative curve has two parts: between 0 - 0.1***s***, the derivatives are positive with large variations, and after 0.1***s***, the derivatives are negative but stable.**

based upon these transient responses. Considering that we now only need to heat the sensor for a few hundred milliseconds instead of a few minutes, we can reduce more than 99% of the overall energy consumption. At the heart of the Transient-Predict approach lies in an accurate prediction algorithm that can quickly learn how the transient ADC values map to the final equilibrium ADC value.

Our data collection works as follow. First, a sensor wakes up from sleep and turns on the heater inside of the sensor. For the first second, the sensor measures its ADC samples at $40Hz$. During this 1-second period, the ADC values have significant variations due to the drastic reduction-oxidation reaction, and the samples collected in this period are thus referred to as *transient responses*. Then, the sensor turns off the heater and goes back to sleep for 3 hours till the next duty cycle. Note in order to collect the ground truth, we actually heat the sensor and measures the final ADC sample at $100s$ when the expected chemical equilibrium has occurred. In total, we collect 41 samples for each experiment. Our objective here is to find and train a model that can predict the final ADC sample based upon as few transient samples as possible.

Figure 6(a) shows an example ADC values trace in a 100-second period (the chemical equilibrium is reached by the end). On the same plot, we also mark the 1-second transient period we work with in Transient-Predict. More example transient ADC traces are shown in Figure 6(b).

3.2 Step I: Data Preprocessing

Since we transmit transient ADC values wirelessly, a small fraction of the data may become missing or polluted due to collision. To address this problem, we apply the following preprocessing steps.

Firstly, we ignore those measurements that do not have the first 5 ADC samples or the final sample. The first few samples are critically important for training our prediction model, and we use the final sample as the ground truth for evaluating the prediction model.

Secondly, we apply the spline interpolation technique on the samples such that the missing data is recovered. After the interpolation, we have 41 samples for each measurement: the first 40 samples are measured within the first second during the measurement and the last sample is measured at $100s$ when the chemical equilibrium is reached.
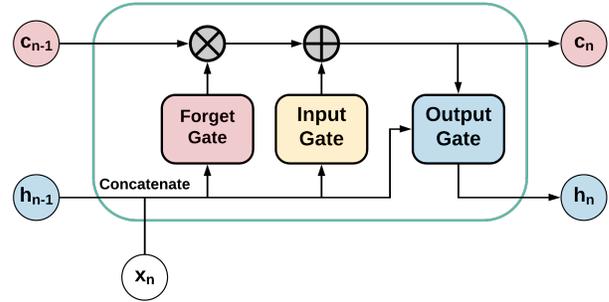
Figure 7: Illustration of the *n*-th LSTM layer.

3.3 Step II: Equilibrium State Electrical Resistance Prediction

The long short-term memory (LSTM) neural networks [19, 20] are one type of recurrent neural networks (RNNs). While common RNNs suffer from failing to learn information dependencies over a large time period, LSTM networks overcome this shortcoming by keeping the memory in a unit called the cell state and maintaining the cell state through a dynamic gating mechanism.

In fact, common LSTM networks have multiple sequential LSTM blocks where each block consists of three gates: a forget gate, an input gate and an output gate. The architecture of an LSTM block is shown in Figure 7. The forget gate in the *n*-th LSTM block uses the input x_n at time n and the previous block's output h_{n-1} to maintain the relevant memory and forget the irrelevant memory. The input gate updates the memory in the current cell state C_n based on the partial memory provided by the forget gate, the input x_n and the previous output h_{n-1}. Later, combining x_n, h_{n-1}, and the new memory, the output gate computes the new output h_n.

Our LSTM neural networks consist of two layers: an LSTM layer followed by a fully connected layer, as shown in the vertical pipeline in Figure 8. The LSTM layer at time n processes the input data x_n, together with the cell state c_{n-1} and the previous output h_{n-1}, and sends the output h_n to a fully connected layer. The fully connected layer generates the final output for the entire network. We build and train our networks with Keras, a deep learning framework
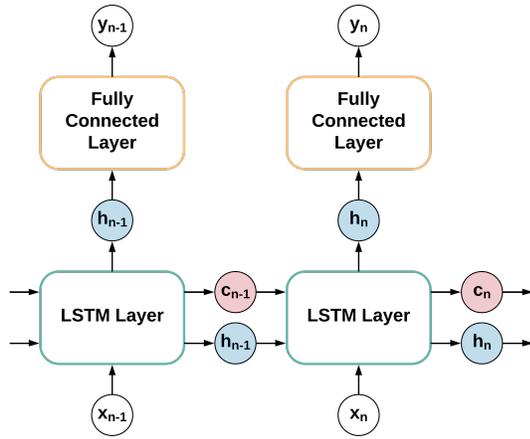
Figure 8: Our LSTM neural network structure. Each vertical column constitutes a block of the LSTM neural networks.

based on python. The model is trained with the mean squared error as the objective loss function and all the coefficients are optimized by the Adam algorithm[26]. Initially, we set the batch size to 10 and the number of hidden neurons to 500. Note, the number of hidden neurons is correlated with the input window size. We start with sufficient hidden neurons for a large window size and further tune parameters (i.e., the window size, the batch size and the number of hidden neurons) and report their impact on the performance in Section 4.1. Table 2 shows the summary representation of our well-tuned network structure.

3.3.1 Preparing the Transient ADC Samples for LSTM Networks. We carefully prepare the transient ADC samples (after the preprocessing step) to make them suitable for the LSTM networks. We find the following two techniques very helpful.

Firstly, we use the first derivatives of the ADC samples, instead of the samples themselves. For the i-th sample $s(i)$ in a measurement, we compute the first derivative as

$$d_1(i) = s(i) - s(i-1). \qquad (1)$$

The use of first derivative greatly helps the LSTM neural networks learn the desired patterns from the data. We observe that the number of epochs required for training the neural networks significantly decreases when we feed the networks with the derivatives rather than the raw data. Specifically, with the raw ADC samples, the training loss starts quite large from the first epoch and remains high for the subsequent 100 epochs, while with the derivatives, the networks quickly converge to a reasonable local minimum within 30 epochs.

Layer	Output Shape	# of parameters
LSTM	(2, 75)	23,100
Fully Connected	(2, 1)	76

Table 2: The final structure of our LSTM neural networks. Here, we set the batch size to 2 and the number of neurons to 75.

Figure 9: Our ADC to concentration mapping process involves industrial calibrated gas, a gas valve with flow rate control, a stainless steel container and necessary anti-corrosive tubes and connectors.

The reason is that the equilibrium ADC value is strongly correlated with how the previous samples change, and the derivative reflects such changes more directly. In Figures 6(b) and (c), we show three transient ADC traces and their corresponding derivatives under ammonia concentration of 25, 100, and 200ppm. Measurements in the first $0.2s$ show larger variations and thus contains more information. Also, the ADC samples decrease between $0.2s$ and $1s$, and the decreasing rate is quite consistent. By calculating the derivatives, we make these patterns more explicit to the LSTM model, which then significantly improves its performance.

Secondly, we scale the derivative $d_1(i)$ to the range of $[a, b]$ by

$$S(i) = \frac{d_1(i) - min(d_1)}{max(d_1) - min(d_1)}(b - a) + a. \qquad (2)$$

In our case, we pick $[-1, 1]$ as the specific range. This step can significantly bring down the training error from a few dozens to below 1×10^{-3}. Here, we mainly consider *sigmoid* function and *tanh* function, both of which are commonly used in neural networks. Our scaling step transforms the data into the desired ranges for both functions. For example, the output of the *tanh* function will exhibit more pronounced changes when the input data is scaled between -1 and 1. As such, the layer outputs become more distinguishable, rending it easier to reduce the objective loss.

3.4 Step III: Ammonia Concentration Calibration

In this last step, we map the estimated equilibrium ADC value to the corresponding ammonia concentration level. In our study, we place multiple sensors into a stainless steel container with fresh air. As shown in Figure 9, the container has one gas connector on each side and has high corrosive resistance to ammonia. Next, we connect one of the connectors to a valve controlled cylinder which contains calibrated ammonia, while leaving the other connectors to the open air. Then, we continuously release calibrated ammonia gas. After some time, the gas in the container reaches the same ammonia concentration inside of the calibrated ammonia cylinder due to the mechanical equilibrium effect.

We repeat this process with different ammonia concentration levels and fit the data into a curve-fitting model. Although the manufacturer suggests a polynomial model [36], we find a power law curve-fitting model fits the data better. The power law model is defined as:

$$C = a\left(\frac{1}{R_s}\right)^n, \qquad (3)$$
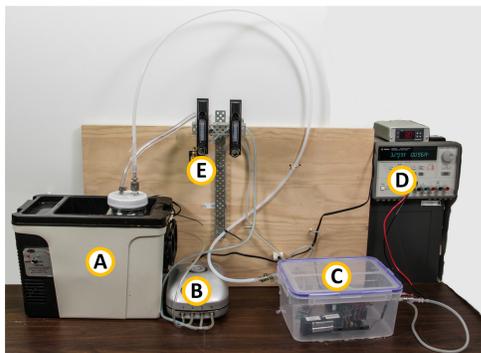
Figure 10: Our home grown gas flow system consists of Ⓐ a temperature controlled gas washing bottle, Ⓑ an anti-corrosive air pump, Ⓒ an anti-corrosive container and other parts like Ⓓ power supply, Ⓔ flow meters, etc.

where C is the concentration, a is a constant prefactor, n is a constant exponent, and R_s is calculated as

$$R_s = \frac{V_{cc} - V_d}{V_d} R_d, \qquad (4)$$

based on the circuit shown in Figure 3. Note we can solve Equation 3 by taking the log on both sides and treat it as a linear least curve fitting problem.

4 EVALUATION OF TRANSIENT-PREDICT

In this section, we evaluate the accuracy of our approach. We report the error rate in predicting the equilibrium ADC value as well as the estimation error in mapping an ADC value to the corresponding ammonia concentration level.

4.1 Evaluation of the LSTM Based Equilibrium ADC Value Prediction

We first evaluate how well our LSTM neural networks model predicts the equilibrium ADC value.

4.1.1 Experimental Setting. We build a home-grown gas flow system which consists of a gas washing bottle, an anti-corrosive air pump, an anti-corrosive container and other parts (flow meters, Teflon tubes, stainless steel connectors, etc.), as shown in Figure 10. The air pump generates the directional air flow that goes through the gas washing bottle to the container. We can switch between the self-circulation mode and no-circulation mode by connecting/disconnecting the outlet of the container to the inlet of the air pump. The gas washing bottle contains a mixture of water-based ammonia hydroxide and distilled water, serving as the source of ammonia gas. We can manipulate the ammonia gas concentration by tuning the ratio of the liquid mixture. Inside the container are a spinning fan and multiple ammonia sensors. The container works as a buffer of ammonia gas and the fan can speed up the uniform balance of ammonia gas and air. Please note that all the equipment parts are washed with the distilled water and then baked at an appropriate temperature for at least 24 hours, in order to eliminate any gas which may affect our measurements.
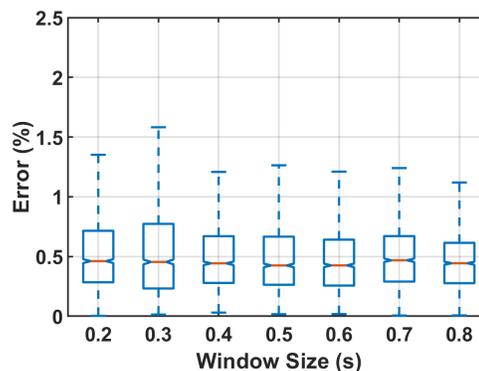
Figure 11: We vary the transient window size from 0.2s to 0.8s and report the average error rate. The transient window size of 0.6s gives the lowest average error rate at 0.89%. However, a window size of 0.2s yields an average error rate at 0.94%, which is sufficient for our application. Please note that a transient window size of 0.1s leads to much larger errors – the error rate is above 50%.

Over a period of 3 months, we collect data from a total of 38 sensors. All the sensors were placed into the anti-corrosive container for a period between 28 and 60 days. During this 3-month period, we collected a total of 13,770 measurements. To conduct the controlled experiments, we repeatedly performed different combinations of the following operations: (1) injecting distilled water in the container and collecting data; (2) injecting the mixture of water-based ammonia hydroxide and distilled water in a self-circulation mode and measuring the concentration drift over time; (3) injecting more ammonia hydroxide to increase the ammonia concentration in the container; (4) switching to a no-circulation mode and measuring the ammonia decay over time; (5) cleaning the gas washing bottle with distilled water and blowing fresh air into the container.

4.1.2 The Impact of the LSTM Neural Networks Parameters. The performance of the LSTM neural networks has a significant bearing on the overall performance of Transient-Predict. Specifically, we consider the following parameters that can impact the performance of the LSTM neural networks: the transient window size, the batch size, the number of hidden neurons, and whether to have a stateful or stateless LSTM neural networks. Empirically, we set the max number of epochs as 100 and allow an early termination of the training phase when the training loss becomes below 1×10^{-4}. Also, we use a learning rate of 0.1 in the following evaluation[3].

Transient window size: In Transient-Predict, we use the transient ADC samples collected from the transient window size to predict the equilibrium ADC value. The window size thus has a significant impact on the overall performance. Here, we program our sensor to sample at 40Hz for the first 1 second. In addition, it also collects the equilibrium ADC sample at the 100-th second. We vary the transient window size from 0.2s to 0.8s and report the resultant prediction error rate – *the ratio between the absolute prediction error and the ground truth* – in Figure 11.

---

[3]In Section 4, all the results are validated over 50 iterations.
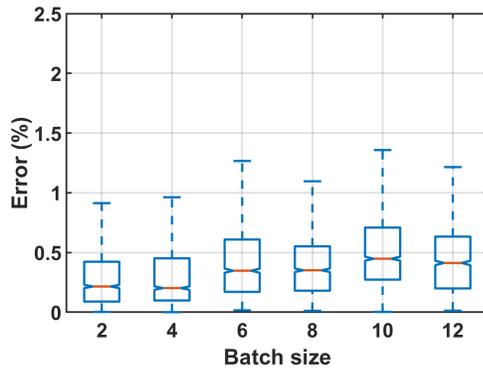
Figure 12: We vary the batch size from 2 to 12 and report the average error rate. The batch size of 2 gives the lowest average error rate at 0.42%. Also, we find a batch size of 1 leads to rather unstable training processes.

Figure 13: We vary the number of hidden neurons from 200 to 75 and report the average error rate. Having 100 hidden neurons gives the lowest average error rate of 0.08%. Also, having 75 hidden neurons yields an average error rate of 0.12%, which is sufficient for our application.

Our main observation is that any window size larger than or equal to 0.2*s* is sufficient for training and yields accurate results. When the transient window size is 0.2*s* or longer, the average prediction error rate is 0.94% or lower; when the transient window size is 0.1*s*, the average prediction error rate is 50%! The reason is that the most important information is embedded in the first few transient ADC samples – there are 8 samples for a 0.2*s* window. Capturing these few samples is the key to our LSTM model. As such, in the rest of this paper, we fix the window size to 0.2*s*.

Batch size: The batch size defines the number of samples that go to forward propagation through the LSTM networks before a backward propagation occurs. An optimal batch size is a good representation of the data set and can prevent models from overfitting. We vary the batch size from 2 to 12 and report the resultant average error rate in Figure 12. We observe that a batch size of 2 is optimal with an average error rate of 0.42%. Larger batch sizes lead to slightly higher error rates (even though the overall error rates are very low). Meanwhile, a batch size of 1 leads to an unstable training process – a small fraction of the models work well while the majority of the models have much larger errors. As such, in the rest of this paper, we fix the batch size to 2.

Number of hidden neurons: The number of hidden neurons is another important parameter for the LSTM networks. Too many hidden neurons may lead to high computation cost and over-fitting, while too few hidden neurons may cause under-fitting. Also, the suitable number of hidden neurons is usually determined by the number of input samples. That is, we may need more hidden neurons for a model trained with a larger transient window. In this set of experiments, we fix the transient window size to be 0.2*s* and vary the number of hidden neurons from 200 to 75. As shown in Figure 13, having 100 hidden neurons gives the lowest average error rate, 0.08%. Also, we notice that having 75 hidden neurons is sufficient for training a good model with an average error rate of 0.12% and having more hidden neurons doesn't significantly lower the error rate. On the other hand, we observe that cutting down the number of hidden neurons even further, such as 50 or 25, leads
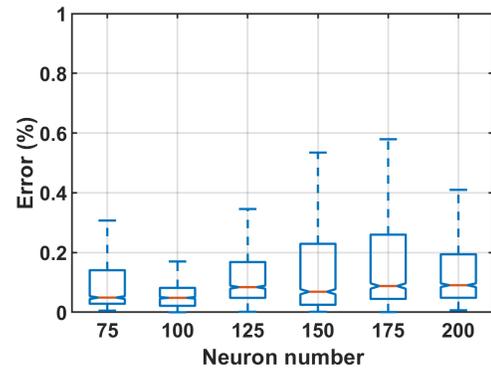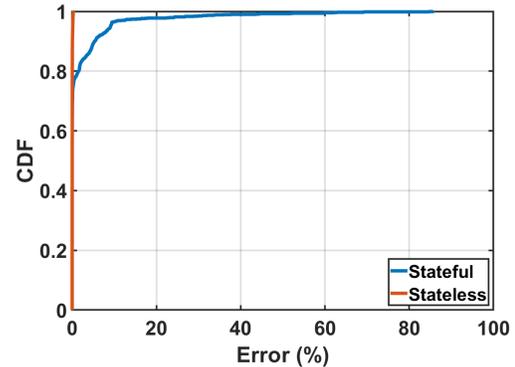
Figure 14: A comparison of stateful LSTM models and stateless LSTM models. The stateless LSTM networks yield lower error rates than the stateful LSTM networks.

to unstable training and much higher error rates. As such, in the rest of this paper, we fix the number of hidden neurons to 75.

Stateful vs stateless: Next, we compare stateful networks vs stateless networks. When processing a new batch of data, stateful networks maintain the states from the previous batch, while stateless networks reset to the initial states. Figure 14 shows the cumulative distribution function (CDF) curves for both stateful and stateless versions of our model. We find the stateless LSTM networks perform better than the stateful counterparts. The average error rates are 0.12% and 2.95%, respectively. The reason, we believe, is that the time dependencies are already well embedded inside of the transient responses, thus we can simply use stateless LSTM networks.

4.1.3 Comparing the LSTM Neural Networks with an Exponential Decay Model and a Linear Regression Model. A few studies [27, 35] mentioned that an exponential decay model might be sufficient for transient response prediction. Other regression methods, such as a simple linear regression model [12], could also be a potential solution. However, our study shows a different trend. We observe two typical types of transient responses, illustrated in Figures 15 (a)

(a)   (b)   (c)
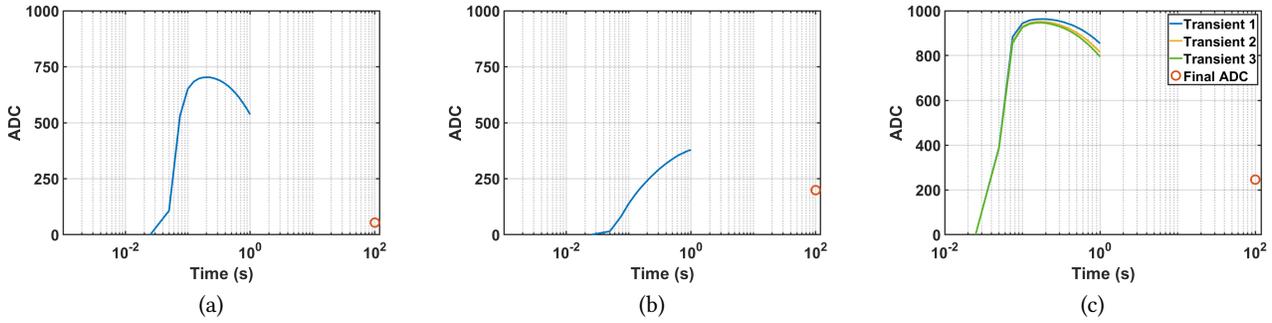
Figure 15: We observe two types of transient responses in our experiments. Within the 1-second window, (a) the transient response of a frequently used sensor has an obvious peak at around 0.2s; (b) the transient response of a sensor that has not been used for a long time may only have an increasing trend. (c) shows multiple transient responses for the same ammonia concentration. In all 3 plots, the red circles mark the equilibrium ADC sample at 100s. Obviously, we could not apply the exponential decay model or the linear regression model to cases shown in (b). Also, we expect large errors when applying the exponential decay model and the linear regression to cases shown in (c).

and (b). Figure 15(a) shows the transient response of a much-used sensor, whose pattern matches our observation in Figure 7(b). Here, the ADC samples first increase, and then decrease. Figure 15(b) shows the transient response of a sensor which has not been used for a while, whose sensing layer is mostly metal oxide. As a result, the reduction reaction is more drastic and takes a longer time to reach the equilibrium. In this case, we observe a monotonically increasing trend rather than an increase followed by a decrease as in (a). In addition, Figure 15(c) shows a sensor's transient responses in different measurements under the same ammonia concentration – the transient response varies from measurement to measurement.

Obviously, an exponential decay model will fail in cases like (b), and for the sake of fairness we only apply the model to cases like in (a) and (c) in our comparison. Specifically, we adopt the following model:

$$y(t) = ae^{-bx(t)} + c, \tag{5}$$

where $x(t)$ is the input, $y(t)$ is the output, a, b and c are constant parameters. We use the samples collected between 0.175s and 1s to estimate the parameters in the model. Once we get the parameters, we estimate the equilibrium ADC sample at 100s. We find that the exponential decay model leads to very poor results for the considered transient window size. A linear regression model faces similar difficulties and yields large errors.

Table 3 compares our approach against the exponential decay model and a linear regression model. Note that for the sake of fairness, we evaluate our model using all the data available, but evaluate the other two models only using those data where such models can be applied such as the ones shown in Figure 15 (a) and (c). The results show that our model can handle all three cases shown in Figure 15 and performs much better than the other two models – even the maximum error of our approach is smaller than the minimum error of the other two models.

Other related models, such as the heat transfer model used in temperature prediction from initial response [1], are well studied. However, it's hard to adapt the same idea in our study due to the significant differences between these two sensing approaches.

4.1.4 Testing the LSTM Networks With Different Test Data. In this section, we demonstrate that our LSTM model can work well on different test data. From all the 38 sensors, we randomly pick a single sensor and train a model with the data from this sensor. Then, we evaluate the model using test data from the same sensor and test data from the other 37 sensors. We repeat this process for all the sensors and show the results in Figure 16. Due to the space limit, we report the performance of three models, each trained by a different sensor, as well as the performance averaged over all cases. The blue box plot represents the average error rate of the model using the test set from the same sensor and each orange box plot represents the results of the model using different test data. In general, we observe a very small performance difference between testing our LSTM networks with same-sensor test data and testing with different-sensor test data – the average error rate in the former case is 0.12% while the average error rate in the latter case is 0.20%. We believe that, although the variations in the sensors are large, there are consistent underlying hidden state representations among all our ammonia sensors that can be captured by LSTM neural networks. More detailed reasons will be studied and included as our future work.

	Time (s)	Min (%)	Max (%)	Mean (%)
Our approach	0.2	0.00	0.45	0.12
Exponential decay	1	$>1 \times 10^5$	$>1 \times 10^5$	$>1 \times 10^5$
Linear regression	1	$>2 \times 10^3$	$>7 \times 10^4$	$>9 \times 10^4$

Table 3: We compares our LSTM model against the exponential decay model and the linear regression model. Our model only requires data within a 0.2s window size and the maximum error is as low as 0.45%. Both the exponential decay model and the linear regression model could not provide any reasonable results using a 1-second window size.
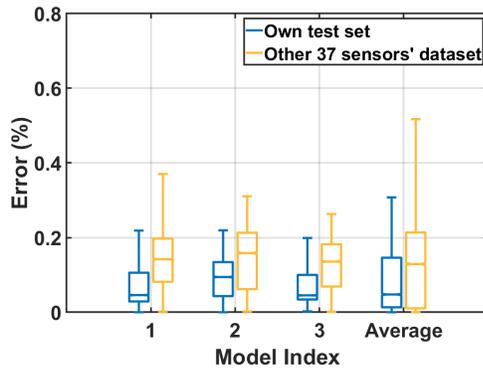
Figure 16: We evaluate how well our LSTM networks perform when using test data from different sensors. The results show that our model performs similarly when testing with same-sensor data or testing with different-sensor data.

4.2 Evaluation of the Power Law Model based Calibration

We repeat the sensor calibration process – i.e., mapping the equilibrium ADC to the corresponding ammonia level (as described in Section 3.4 – using industrial calibrated ammonia gas at different concentrations and report the calibration results. Figure 17 shows example calibration results for 3 sensors. We convert the ADC samples to the sensor's resistance according to Equation 4 and show the results in the log scale on both x and y axes. The mean absolute error across all 38 sensors is 9.38ppm. Also, the sensor has different sensitivity to ammonia. Our results and observation agree with the results reported in [28, 34, 37].

5 REAL WORLD TRIALS

A 4-month NIH Trial: We have deployed our ammonia monitoring system in a laboratory animal facility at National Institute of Neurological Disorders and Stroke (NINDS) and completed a 4-month trial[4]. This trial involves 20 ventilated cages, and we deployed one ammonia sensor in each cage. The entire trial is divided into 6 cycles, with 21 days per cycle. All cages host 5 female mice that were 12 weeks old at the beginning of the trial. We provide mice with bedding, standard food and reverse osmosis water. We measure the resultant ammonia concentration level continuously, once every 3 hours.

To illustrate the results, we plot the sensor measurements in a 21-day cycle for cage 11 in Figure 18(a). During this cycle, we never changed the cage but took measurements (e.g. temperature, humidity, ammonia, carbon dioxide, etc.) with handheld sensors on every Tuesdays and Fridays. All of the six measurement events are marked and numbered in Figure 18(a). Also, on the same days when the measurements happened, we measured adenosine triphosphate (ATP) released in duct cells, the amount of water consumed, etc. Since the cage remained the same during the entire cycle, the overall ammonia level continuously went up, except that after each measurement ammonia level decreased to a certain degree.
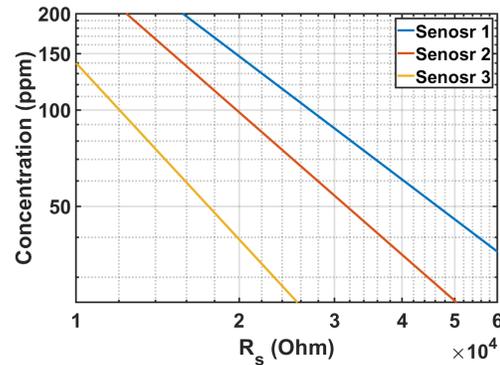
Figure 17: We show example calibration results for 3 sensors. We observe the sensor has different sensitivity for ammonia. Note that both x and y axis are in log scale. The mean absolute error across all 38 sensors is 9.38ppm.

Even though the overall trend of the ammonia level is going up, we observe a fine-grained periodic ammonia level fluctuation in Figure 18(a). Further, the interval between two adjacent local peaks is 24 hours. We hypothesize this daily fluctuation is caused by daily events in the environment, such as the building ventilation system. To confirm this observation, we remove the trend in the ammonia measurement trace and compute the FFT of the residual signal. Shown in Figure 18(b), the residual signal has an obvious spike at once per day. Similarly, we also observed the same phenomenon in our lab and in other trials we have conducted. Finally, we note that the above trend we observe for cage 11 is common across all the cages.

A 6-month Cornell Trial: The second trial we had is at Cornell University and lasts for 6 months. The objective of the trial is to study the animals' breeding related behaviors. This trial involves 16 standard cages and we deployed one ammonia sensor in each cage. In this trial, we measure each cage's ammonia level continuously, once every 3 hours.

To illustrate the results, we show cage 6's measurements in a 28-day period in Figure 18(c). On day 1, the female mouse of the couple in the cage gave birth to 8 pups. On day 2, we had our first scheduled cage change – both the two adult mice and the pups were moved to a new cage 6. Between day 2 and day 15, we observe the ammonia concentration in cage 6 increased. On day 16, we had our second scheduled cage change. After the second cage change, we observe a much faster increase of ammonia concentration in the cage as the pups grew over time. Further, the ammonia level increase from day 21 to day 25 became even faster. This is because in general day 21 is considered a pup's weaned age – after that, they can eat and latrine on their own. Thus, the production of waste in cage 6 increased and the ammonia production increased as well. After day 25, the ammonia level decreased due to some undocumented cage change events.

Discussion about the ammonia measurement frequency in real-world applications: Lab animals are used for a variety of research purposes. The ammonia concentration depends on several factors – the type of the cages (static, air ventilated), the type of the bedding, the number of mice/rats, the type of experiments (e.g.,
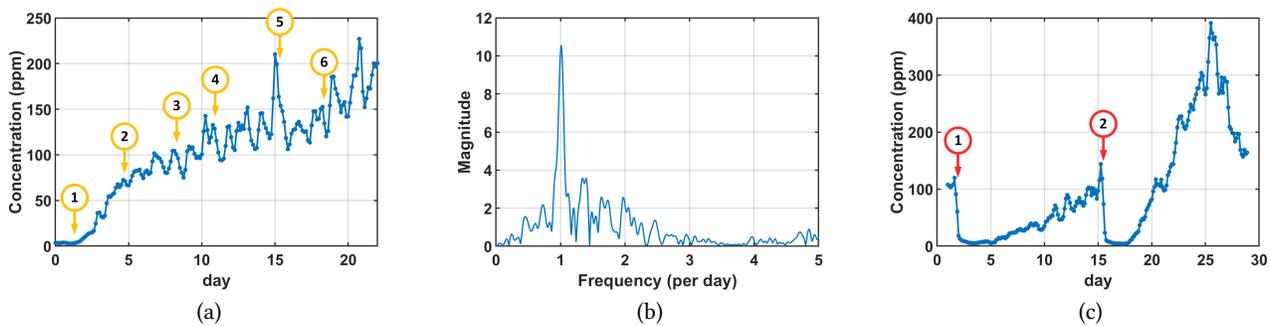
Figure 18: (a) The ammonia concentration trace in a 21-day cycle of cage 21 in the NIH trial. The yellow circles mark the scheduled operations of measuring cage environment. (b) We removed the obvious trend in (a) using a moving average filter and compute the FFT results of the residual. Clearly, there is a spike indicating a repetitive event on a daily basis. We hypothesize this is due to the daily operations of the building ventilation system. (c) The ammonia concentration trace of cage 6 in a 28 day period in the Cornell trail. The red circles show the scheduled cage change events.

drug tests may disturb digest behavior), etc. Thus, it is hard to have an exact measurement frequency that satisfies all the cases. For example, a diabetic model mouse might require sampling on the order of a few hours, while a single mouse in a cage might be 1 to 2 days. Our system can satisfy the need of frequent ammonia monitoring and still can last for at least 20 years assuming one measurement in every three hours.

6 RELATED WORK

Many ammonia monitoring methods have been proposed. We categorize existing ammonia monitoring methods by their sensing modality. Specifically, we have ammonia test strips, test tubes [11], metal oxide sensors [37, 45, 46], electrical-chemical sensors [21, 38, 42] and fiber-coupled optical sensors [25, 27].

We note that, none of the existing systems are accurate, low-power, low-cost, compact and automatic – all at the same time. For example, ammonia test strips are inaccurate; ammonia test tubes are expensive; metal oxide sensors are power-consuming and bulky; electrical-chemical sensors and fiber-coupled optical sensors are power-consuming, expensive and bulky. Also, all the existing ammonia monitoring methods require extensive manual operations for large-scale lab cage monitoring.

Ammonia test strips: An ammonia test strip typically has two layers. When we dip a moist ammonia test strip into a cage, the first layer quickly picks up ammonia in the air onto its moist side. Next, the alkaline chemical, on the other side of the first layer, release ammonia in the gaseous state to the second layer. The second layer has a chemical reagent which reacts with ammonia gas. This reaction leads to a pH change and shows in the form of color change. Finally, we compare the color of the ammonia strip against a standard color chart and estimate the ammonia concentration empirically. The major advantage of ammonia test strips is low-cost. However, the measurement is neither accurate nor convenient. In practice, at high concentration (> 50*ppm*), the color of an ammonia strip may change quickly in one or two seconds because ammonia is a highly volatile gas. Also, the standard color chart only shows colors at a few concentrations – usually 0, 10, 25, 50, 100*ppm*. Thus,

it's rather difficult to get an accurate measurement using ammonia test strips.

Ammonia test tubes: An ammonia test tube [11], also known as gas chromatography in a handheld unit, is typically combined with a hand-held device. This device typically has an air pump, a meter which reads the color change, and a large battery. Starting from each experiment, the pump draws a fixed amount of air and mix the air with the chemical reagent in the tube. Instead of mapping the color, the device measures the duration of the color change and maps the duration to a concentration. Ammonia test tubes provide a measurement of ±10% accurate. However, the device requires manual operation and the cost per measurement is expensive (about $10).

Metal oxide sensors: A few papers [37, 45, 46] proposed to use metal oxide sensors for ammonia monitoring. At elevated temperature, usually a few hundred ℃, metal oxide may have reduction reaction with ammonia and be converted to metal. Meanwhile, metal may have oxidization reaction with oxygen and be converted back to metal oxide. Over time, the reduction and the oxidization reaction reach a chemical equilibrium. Then, we can measure the sensor resistance and map it to ammonia concentration. The sensor is reusable and accurate. However, the sensor requires minutes heating thus is power consuming. Also, like others, the existing metal oxide sensors are built into handheld devices and requires manual operation.

Electrical-chemical (EC) sensors: A EC sensor [21, 38, 42] consists of three electrodes: a reacting electrode, a counter electrode and a reference electrode. Similar as metal oxide sensors, the reacting electrode in a EC sensor have reduction-oxidation reaction with ammonia at high temperature. The reaction generates electric current between the reacting electrode and the counter electrode. We can map the voltage across the electrodes with respect to the reference electrode to an ammonia concentration. EC sensors are accurate, but expensive, power consuming and have limited lifetime.

Fiber-coupled optical sensors: Klein et al. [27] proposed a fiber-coupled optic ammonia sensor. The sensor consists of a glass chamber which contains a multimode waveguide with carefully coated chemical dye. Outside the chamber, two light sources of particular wavelength shoot lights towards the sensor. As ammonia appears in the glass chamber, the dye changes its color due to the reversible chemical reaction. Such color change greatly changes the attenuation of the light at one wavelength, but not the attenuation of the light at the other wavelength. Then, an algorithm can map the relative attenuation ratio between two light wavelength to an ammonia concentration. This sensor is reusable. However, the sensor requires expensive calibration and can be easily affected by other common gas like carbon dioxide. Karlsson et al. [25] proposed gas analysis based on photoacoustic spectroscopy. The entire system is centered around a Cobolt Odin mid-wavelength infrared unit [8] as the input laser source and Gasera's PA201 photoacoustic gas cell [18] as the sensing unit. The entire system is expensive and bulky. Also, both sensor systems require extensive manual operations as well. As such, they are not suitable for large-scale lab cage monitoring.

7 CONCLUSION AND FUTURE WORK

In this paper, we develop and evaluate a low-power, automatic, accurate and wireless ammonia monitoring system that uses metal oxide sensors. The traditional approach requires one to heat the sensor for minutes till the chemical reaction among the metal oxide, ammonia and oxygen reaches the chemical equilibrium point. Such an approach consumes a significant amount of energy, thus hindering automatic and long term ammonia sensing. We propose to greatly shorten the heating period and use the transient measurements in the short heating window to predict the final value at the equilibrium state. Our prediction model is centered around LSTM neural networks. Our approach can cut down the overall energy assumption by about 99.6%. Through extensive experiments, we show that our prediction model is accurate across 38 prototype sensors – the average prediction error rate of is 0.12% and the average absolution error against the calibrated gas is 9.38*ppm*. Also, our model can be used to accurately predict for new sensors even though the data from these sensors were not part of the training data.

As our future work, we will further investigate several important issues, including (1) integrating other sensors (such as a temperature sensor, a humidity sensor, an air pressure sensor, etc.) into our system to better understand the impact of environmental factors on ammonia sensing; (2) visualizing the hidden state representations of LSTM with our time-series input and understanding the features learned by LSTM; (3) investigating the detailed reasons for successfully applying one model (learned from a single sensor) to other sensors, despite the significant characteristic differences among all these sensors.

ACKNOWLEDGMENTS

REFERENCES

[1] NA Abukhshim, PT Mativenga, and MA Sheikh. 2006. Heat generation and temperature prediction in metal cutting: A review and implications for high speed machining. *International Journal of Machine Tools and Manufacture* 46, 7-8 (2006), 782–800.

[2] Musaab Alaziz, Zhenhua Jia, Richard Howard, Xiaodong Lin, and Yanyong Zhang. 2017. MotionTree: A Tree-Based In-Bed Body Motion Classification System Using Load-Cells. In *Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017 IEEE/ACM International Conference on.* IEEE, 127–136.

[3] Musaab Alaziz, Zhenhua Jia, Jian Liu, Richard Howard, Yingying Chen, and Yanyong Zhang. 2016. Motion Scale: A Body Motion Monitoring System Using Bed-Mounted Wireless Load Cells. In *Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2016 IEEE First International Conference on.* IEEE, 183–192.

[4] Jonathan P Balcombe, Neal D Barnard, and Chad Sandusky. 2004. Laboratory routines cause animal stress. *Journal of the American Association for Laboratory Animal Science* 43, 6 (2004), 42–51.

[5] Tadiran Batteries. 2018. Tadiran Rapid Response Lithium Batteries MODEL TL-7903. (2018). http://www.tadiranbat.com/trr.html.

[6] Larry Carbone. 2004. *What animals want: Expertise and advocacy in laboratory animal welfare policy.* Oxford University Press, USA.

[7] MJ Castelhano-Carlos and V Baumans. 2009. The impact of light, noise, cage cleaning and in-house transport on welfare and stress of laboratory rats. *Laboratory animals* 43, 4 (2009), 311–327.

[8] Cobolt. 2018. Cobolt Odin™Series. (2018). https://www.coboltlasers.com/lasers/tunable-lasers/

[9] National Research Council et al. 2010. *Guide for the care and use of laboratory animals.* National Academies Press.

[10] Bruce D Craig, David S Anderson, et al. 1994. *Handbook of corrosion data.* ASM international.

[11] Drägerwerk. 2018. Dräger-Tubes® & accuro® Pump. (2018). https://www.draeger.com/Products/Content/tubes-accuro-pump-pi-us.pdf

[12] Norman R Draper and Harry Smith. 2014. *Applied regression analysis.* Vol. 326. John Wiley & Sons.

[13] Jodi L Duke, Timothy G Zammit, and David M Lawson. 2001. The effects of routine cage-changing on cardiovascular and behavioral parameters in male Sprague-Dawley rats. *Journal of the American Association for Laboratory Animal Science* 40, 1 (2001), 17–20.

[14] Aaron C Ericsson, J Wade Davis, William Spollen, Nathan Bivens, Scott Givan, Catherine E Hagan, Mark McIntosh, and Craig L Franklin. 2015. Effects of vendor and genetic background on the composition of the fecal microbiota of inbred mice. *PloS one* 10, 2 (2015), e0116704.

[15] Xiaoran Fan, Han Ding, Sugang Li, Michael Sanzari, Yanyong Zhang, Wade Trappe, Zhu Han, and Richard E Howard. 2018. Energy-Ball: Wireless Power Transfer for Batteryless Internet of Things through Distributed Beamforming. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 65.

[16] Santiago Fernández, Alex Graves, and Jürgen Schmidhuber. 2007. An application of recurrent neural networks to discriminative keyword spotting. In *International Conference on Artificial Neural Networks.* Springer, 220–229.

[17] Bernhard Firner, Shweta Medhekar, Yanyong Zhang, Richard Howard, Wade Trappe, Peter Wolniansky, and Eitan Fenson. 2008. Pip tags: Hardware design and power optimization. In *Proceedings of the Fifth Workshop on Embedded Networked Sensors (HotEmNets).*

[18] Gasera. 2018. GaseraâĂŹs PA201 Research Photoacoustic Gas Cell for Laser Sources. (2018). https://www.gasera.fi/product/pa201/

[19] Felix A Gers, Jürgen Schmidhuber, and Fred Cummins. 1999. Learning to forget: Continual prediction with LSTM. (1999).

[20] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.

[21] RKI Instruments. 2018. SC-01 Operator's Manual. (2018). http://www.rkiinstruments.com/pdf/71-0136RK.pdf

[22] M Nygaard Jensen and M Ritskes-Hoitinga. 2007. How isoflavone levels in common rodent diets can interfere with the value of animal models and with experimental results. *Laboratory animals* 41, 1 (2007), 1–18.

[23] Zhenhua Jia, Musaab Alaziz, Xiang Chi, Richard E Howard, Yanyong Zhang, Pei Zhang, Wade Trappe, Anand Sivasubramaniam, and Ning An. 2016. HB-phone: a bed-mounted geophone-based heartbeat monitoring system. In *Proceedings of the 15th International Conference on Information Processing in Sensor Networks.* IEEE Press, 22.

[24] Zhenhua Jia, Amelie Bonde, Sugang Li, Chenren Xu, Jingxian Wang, Yanyong Zhang, Richard E Howard, and Pei Zhang. 2017. Monitoring a Person's Heart Rate and Respiratory Rate on a Shared Bed Using Geophones. In *Proceedings of*

the 15th ACM Conference on Embedded Network Sensor Systems. ACM, 6.

[25] Håkan Karlsson and Sauli Sinisalo. 2017. Air Quality Monitoring with Photoacoustic Spectroscopy: Tunable, narrow bandwidth mid-IR laser sources for compact trace gas analysis systems with ppb sensitivity. *Optik & Photonik* 12, 1 (2017), 36–39.

[26] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[27] Rainer Klein and E Voges. 1993. Integrated-optic ammonia sensor. *Sensors and Actuators B: Chemical* 11, 1-3 (1993), 221–225.

[28] Douglas C Meier, Steve Semancik, Bradley Button, Evgheni Strelcov, and Andrei Kolmakov. 2007. Coupling nanowire chemiresistors with MEMS microhotplate gas sensing platforms. *Applied Physics Letters* 91, 6 (2007), 063118.

[29] Angela M Mexas, Angela K Brice, Adam C Caro, Troy S Hillanbrand, and Diane J Gaertner. 2015. Nasal histopathology and intracage ammonia levels in female groups and breeding mice housed in static isolation cages. *Journal of the American Association for Laboratory Animal Science* 54, 5 (2015), 478–486.

[30] Jorge Perez-Serrano, Javier Palacin-Urquijo, Juan Soliveri de Carranza, Jose MarÃŋa Orellana-Muriana, and Jose L. Copa-Patino. 2017. Desarrollo, Caracterizacion y Optimizacion de Nuevos Materiales Biologicos Utilizados Como Lechos Para la Mejora del Bienestar Animal Durante el Alojamiento de Animales de Experimentacion. (March 2017). https://portal.uah.es/portal/page/portal/epd2_profesores/prof121192/investigacion

[31] Skye Rasmussen, Melinda M Miller, Sarah B Filipski, and Ravi J Tolwani. 2011. Cage change influences serum corticosterone and anxiety-like behaviors in the mouse. *Journal of the American Association for Laboratory Animal Science* 50, 4 (2011), 479–483.

[32] Mohsen Razavy. 2003. *Quantum theory of tunneling.* World Scientific.

[33] Sara Reardon. 2016. A mouse's house may ruin studies: environmental factors lie behind many irreproducible rodent experiments. *Nature* 530, 7590 (2016), 264–265.

[34] SGX SENSORTECH. 2018. MiCS-5914 Datasheet. (2018). https://www.sgxsensortech.com/content/uploads/2014/08/AN2-%E2%80%93-Frequently-Asked-Questions-for-MiCS-Gas-Sensors.pdf

[35] SGX SENSORTECH. 2018. MiCS Application Note FAQ. (2018). https://www.sgxsensortech.com/content/uploads/2014/07/1108_Datasheet-MiCS-5914.pdf

[36] SGX SENSORTECH. 2018. SGX Metal Oxide Gas Sensors. (2018). https://www.sgxsensortech.com/content/uploads/2014/08/AN-0172-SGX-Metal-Oxide-Gas-Sensors-V1.pdf

[37] Feng Shao, Martin WG Hoffmann, Joan Daniel Prades, Joan Ramon Morante, NuÌŊria LoÌ́Apez, and Francisco HernaÌ́Ąndez-RamiÌ́Ąrez. 2013. Interaction mechanisms of ammonia and tin oxide: A combined analysis using single nanowire devices and DFT calculations. *The Journal of Physical Chemistry C* 117, 7 (2013), 3520–3526.

[38] Bartosz Szulczyński and Jacek Gębicki. 2017. Currently commercially available chemical sensors employed for detection of volatile organic compounds in outdoor and indoor air. *Environments* 4, 1 (2017), 21.

[39] Björn Timmer, Wouter Olthuis, and Albert Van Den Berg. 2005. Ammonia sensors and their applicationsâĂŤa review. *Sensors and Actuators B: Chemical* 107, 2 (2005), 666–677.

[40] Linda A Toth. 2015. The influence of the cage environment on rodent physiology and behavior: Implications for reproducibility of pre-clinical rodent research. *Experimental neurology* 270 (2015), 72–77.

[41] Catherine M Vogelweid, Kathleen A Zapien, Matthew J Honigford, Linghui Li, Hua Li, and Heather Marshall. 2011. Effects of a 28-day cage-change interval on intracage ammonia levels, nasal histology, and perceived welfare of CD1 mice. *Journal of the American Association for Laboratory Animal Science* 50, 6 (2011), 868–878.

[42] Gustavious Williams, Mark David Wald-Hopkins, Stephen J Obrey, and Valida Dushdurova Akhadov. 2016. *Handheld Multi-Gas Meters Market Survey Report.* Technical Report. Los Alamos National Lab.(LANL), Los Alamos, NM (United States).

[43] Chenren Xu, Bernhard Firner, Robert S Moore, Yanyong Zhang, Wade Trappe, Richard Howard, Feixiong Zhang, and Ning An. 2013. Scpl: Indoor device-free multi-subject counting and localization using radio signal strength. In *Information Processing in Sensor Networks (IPSN), 2013 ACM/IEEE International Conference on.* IEEE, 79–90.

[44] Chenren Xu, Bernhard Firner, Yanyong Zhang, Richard Howard, Jun Li, and Xiaodong Lin. 2012. Improving RF-based device-free passive localization in cluttered indoor environments through probabilistic classification methods. In *Proceedings of the 11th international conference on Information Processing in Sensor Networks.* ACM, 209–220.

[45] CN Xu, N Miura, Y Ishida, K Matsuda, and N Yamazoe. 2000. Selective detection of NH3 over NO in combustion exhausts by using Au and MoO3 doubly promoted WO3 element. *Sensors and Actuators B: Chemical* 65, 1-3 (2000), 163–165.

[46] K Zakrzewska. 2001. Mixed oxides as gas sensors. *Thin solid films* 391, 2 (2001), 229–238.