Autonomous Online Predictive Monitoring System for hazardous air pollution prediction: A Case Study

AYED A. SALMAN & FAWAZ S AL-ANZI

ayed.salman@ku.edu.kw & alenzi.f@ku.edu.kw
COLLEGE OF COMPUTING SCIENCES AND ENGINEERING
KUWAIT UNIVERSITY

GET 2016 - Kuwait - 26th - 28th Jan, 2016.

Environment – Human Relationship

Human activities

economic development

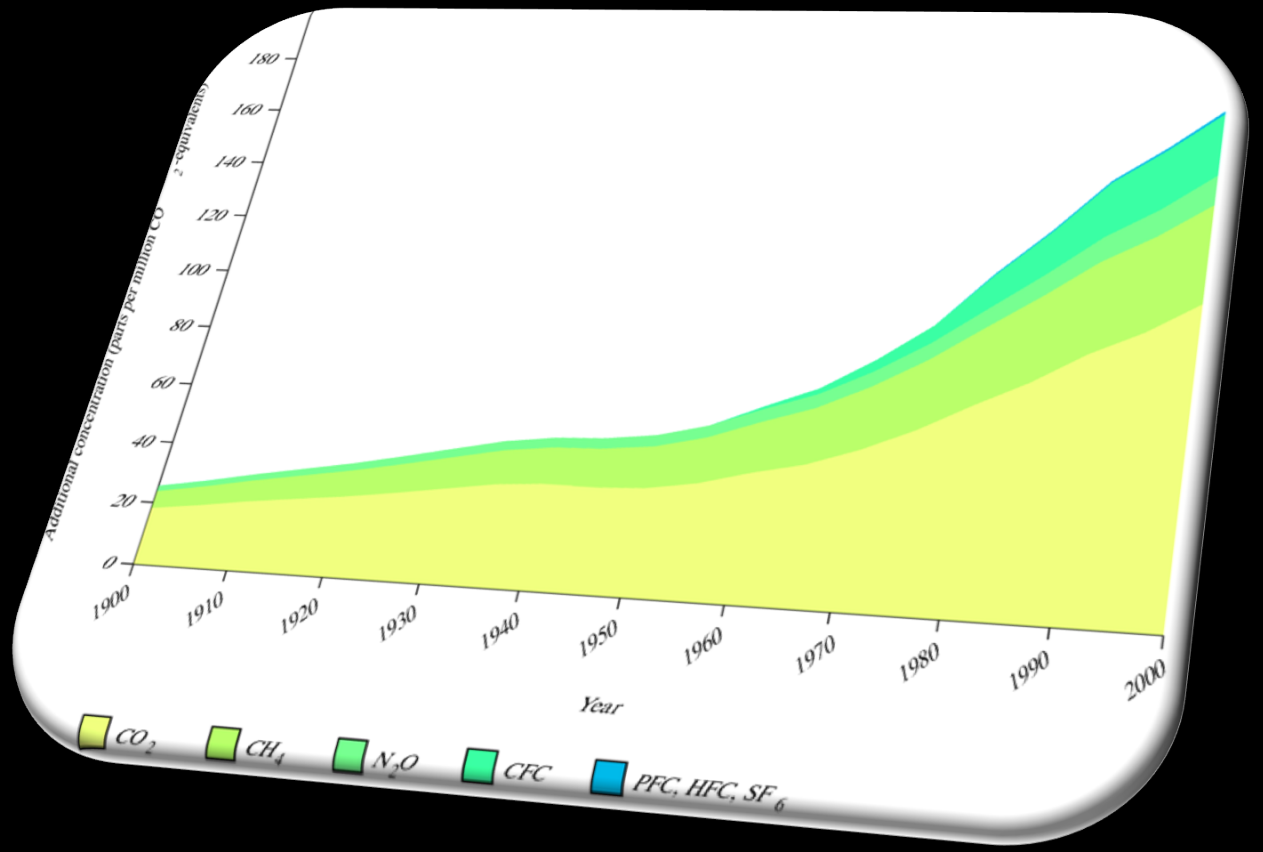
Environmental hazards and pollutions

Human Quality of Life

illness

Air pollution

- An average of **2 million people** are killed worldwide every year due to air pollution
- Air pollution in China alone is related to around **656,000 deaths every year**
- In a survey of more than 100 cities, the World Bank has found that the air in many urban areas remains unhealthy
- Air pollution increasingly affects climate changes directly by preventing sun heat from escaping from atmospheric back to the space



(Source: National Geographic News)

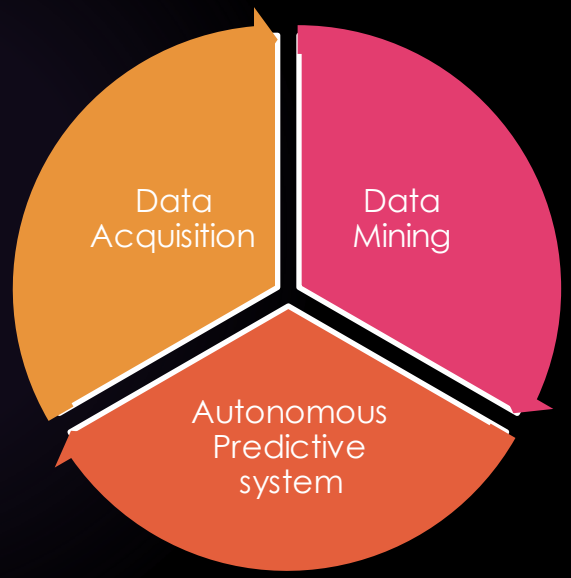
Source: European environmental Agency: http://www.eea.europa.eu/data-and-maps/figures/rise-of-greenhouse-gases-concentration-compared-with-the-year-1750

Kuwait Air Quality

- Current population of Over 4 million
- Industrial areas are of closed proximity to residential areas.
- A recent study showed that concentrations of the following air pollutants CH4, CO, O3, SO2, NOx and total sulfur (TS), had increased over the period (1998-2004)
- The same study shows that NOx and SO2 exceed the permitted standard levels.
- Alwadhi (2014) stated that factories like the one in oil industry are a major contributor to pollution in Kuwait.
- **Sources:**
  - **Al-Mutairi, N. and Koushki, P, Potential Contribution of Traffic to Air Pollution in the State of Kuwait. American, Journal of Environmental Sciences, 5, 218-222, . (2009).**
  - **Jasem M. Al-Awadhi, Measurement of Air Pollution in Kuwait City Using Passive Samplers, Atmospheric and Climate Sciences, 4, 253-271, , 2014.**

Data mining

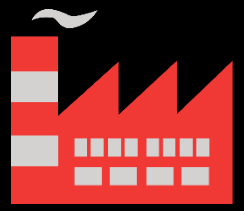
- **Data mining is a computer science field in which a computational process is used for discovering patterns in large data sets using methods from artificial intelligence.**



- Classification
- Prediction
- Clustering
- Regression
- Association

building an autonomous system for Hazardous rates prediction

System Details (1):

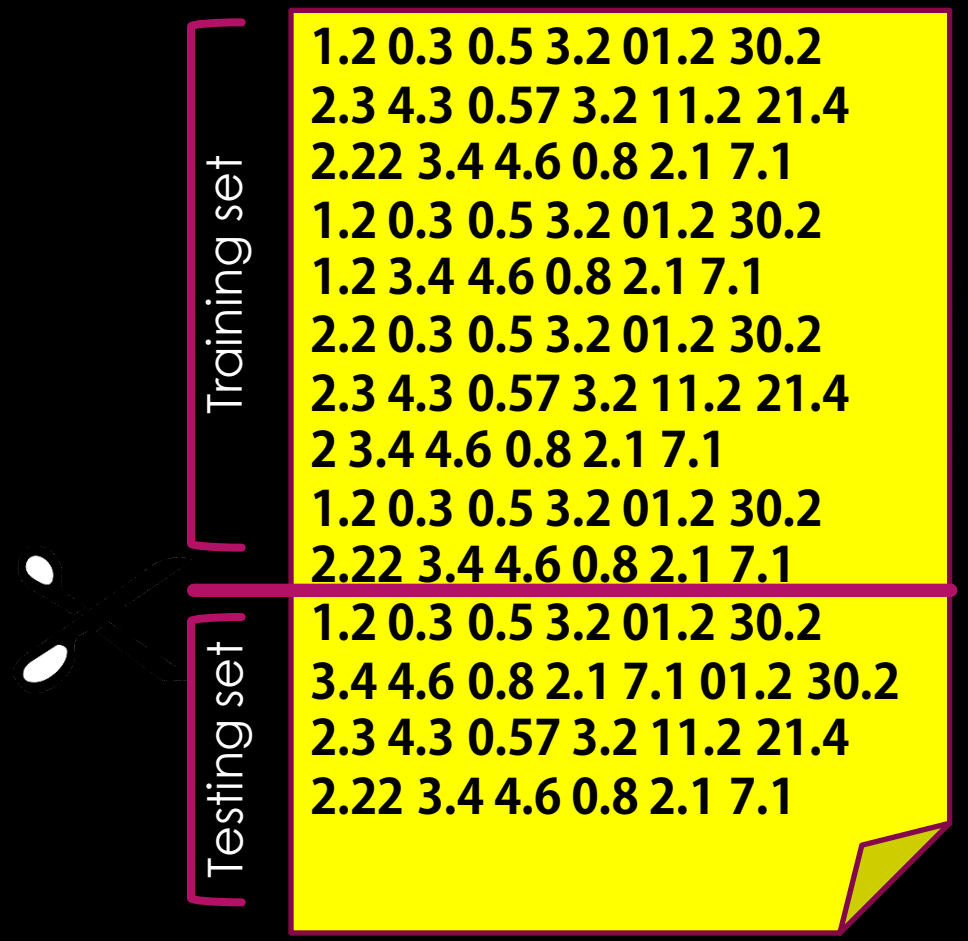


- Phase 1: Data acquisition & preparation:
  - **every 15 minutes**: Raw data readings of gases such as CH4, CO, CO2, NO, NO2, SO2, and H2S is acquired using mobile sensors around the industrial area for a period of time (**~ 2 months time**).

  - Data is transferred to a central server
  - average readings for **1, 12 and 24 hours** is calculated & stored as well.

System Details (2)

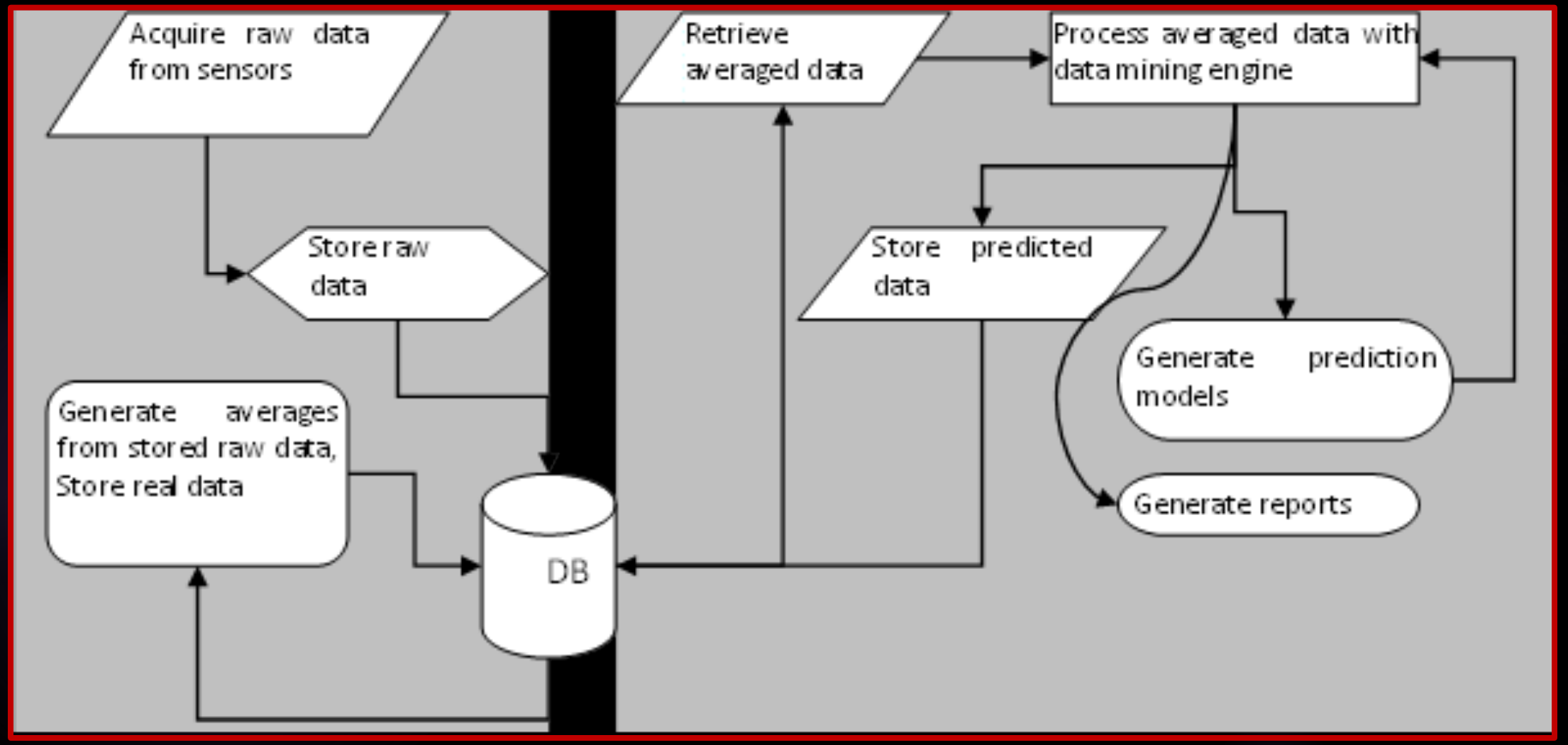
- Phase 2: Data Mining
  - Data is divided into **training** and **testing sets**
  - Data Mining algorithm is used to learn the pattern using training set
  - The same algorithm is tested using the testing set, once satisfactory results → **System goes online.**



System Details (3)

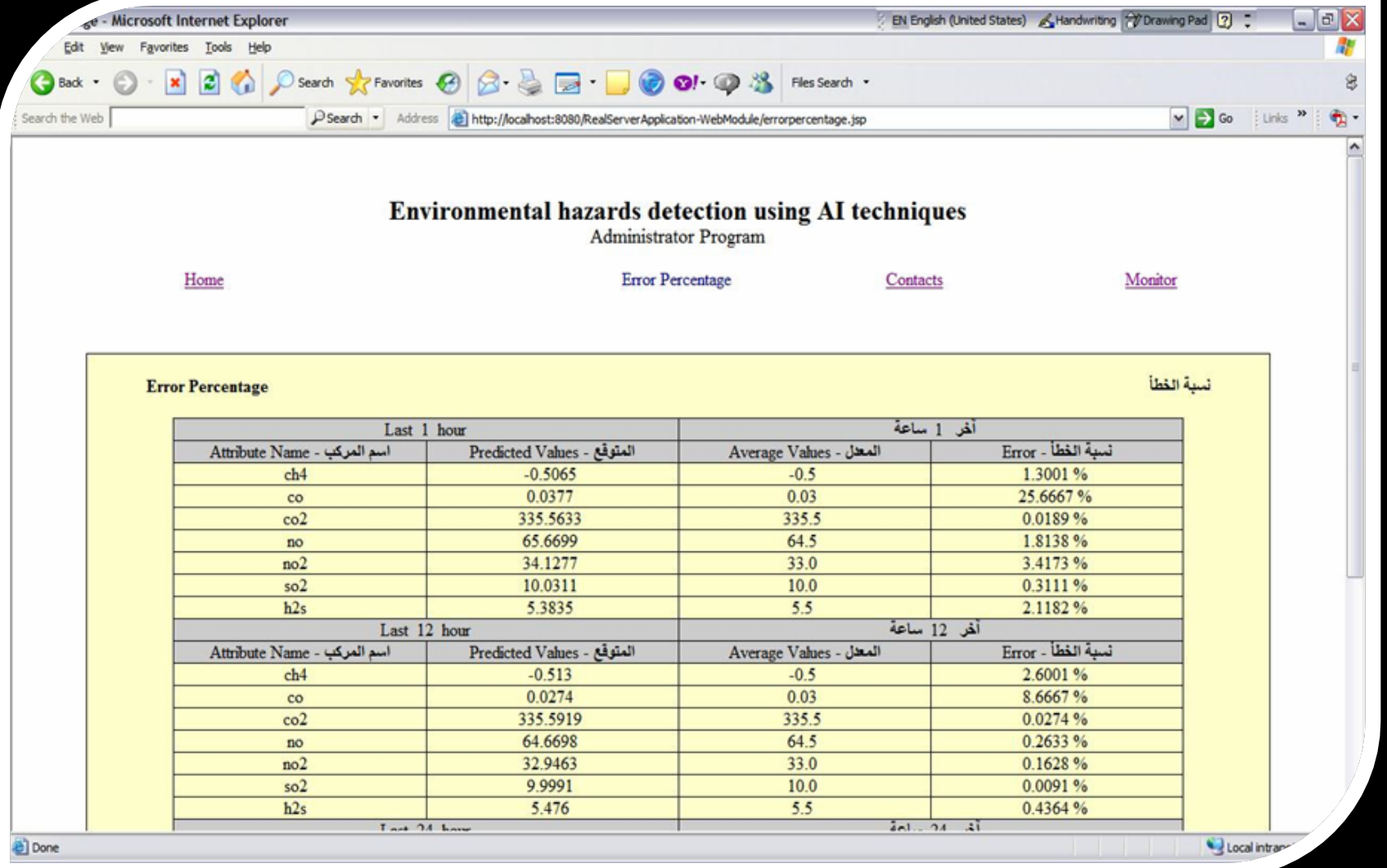
- Online Phase:
  - Continuous Data acquisition
  - Continuous prediction for 1, 12, 24 hours period.
  - Alert is sent to managers in case rates goes beyond standards
  - Reports are generated continuously

BIG Picture


Acquire raw data from sensors
Store raw data
Generate averages from stored raw data, Store real data
DB
Retrieve averaged data
Process averaged data with data mining engine
Store predicted data
Generate prediction models
Generate reports

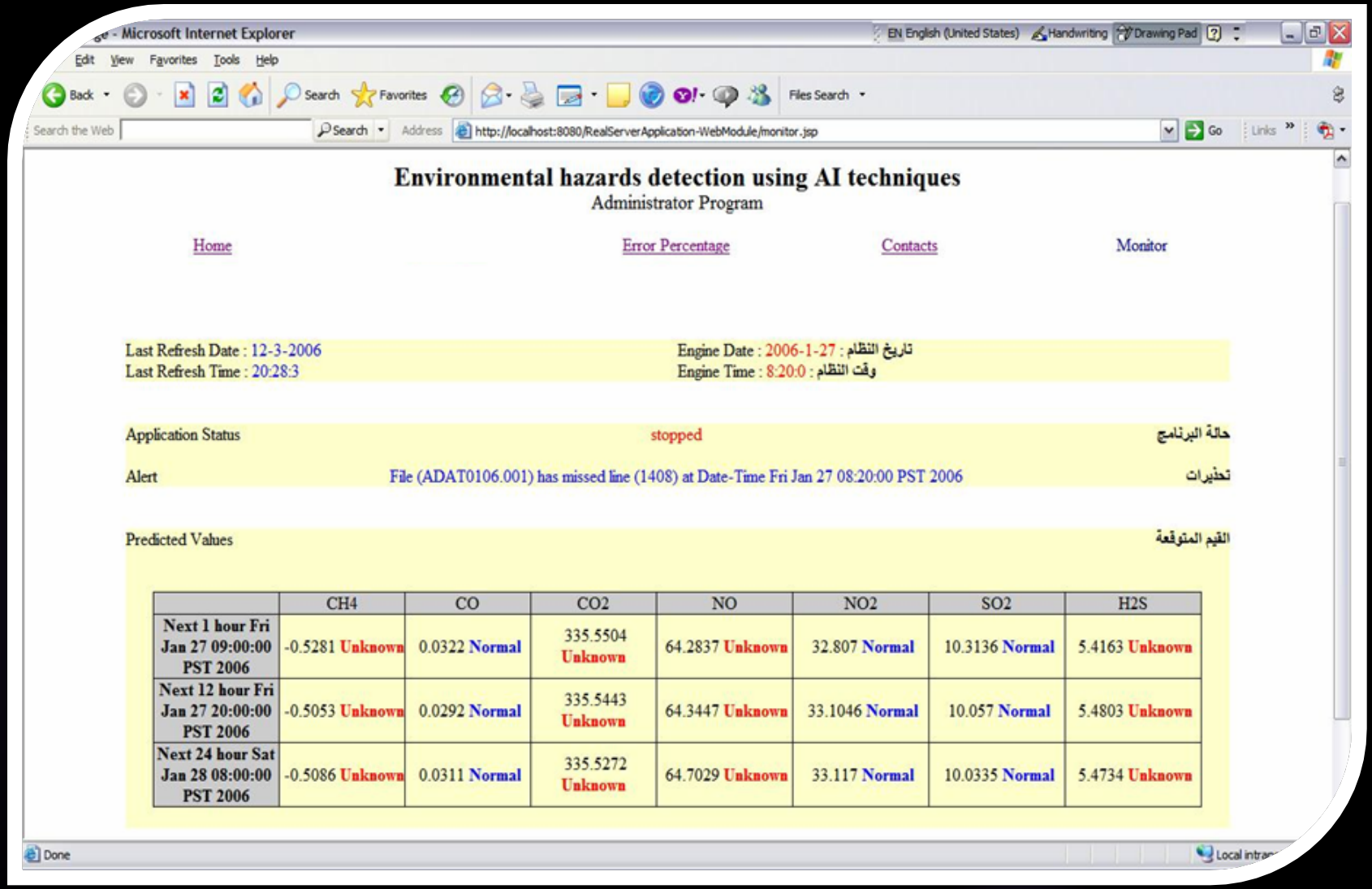
Screen Shot of Prediction Report

- Prediction Reaches as high as 99.8% for some gasses especially for larger periods (12 hours or 24 hours)



Last 1 hour		آخر 1 ساعة	
Attribute Name - اسم المركب	Predicted Values - المتوقع	Average Values - المعدل	Error - نسبة الخطأ
ch4	-0.5065	-0.5	1.3001 %
co	0.0377	0.03	25.6667 %
co2	335.5633	335.5	0.0189 %
no	65.6699	64.5	1.8138 %
no2	34.1277	33.0	3.4173 %
so2	10.0311	10.0	0.3111 %
h2s	5.3835	5.5	2.1182 %
Last 12 hour		آخر 12 ساعة	
Attribute Name - اسم المركب	Predicted Values - المتوقع	Average Values - المعدل	Error - نسبة الخطأ
ch4	-0.513	-0.5	2.6001 %
co	0.0274	0.03	8.6667 %
co2	335.5919	335.5	0.0274 %
no	64.6698	64.5	0.2633 %
no2	32.9463	33.0	0.1628 %
so2	9.9991	10.0	0.0091 %
h2s	5.476	5.5	0.4364 %

Screen Shot of Monitoring Interface



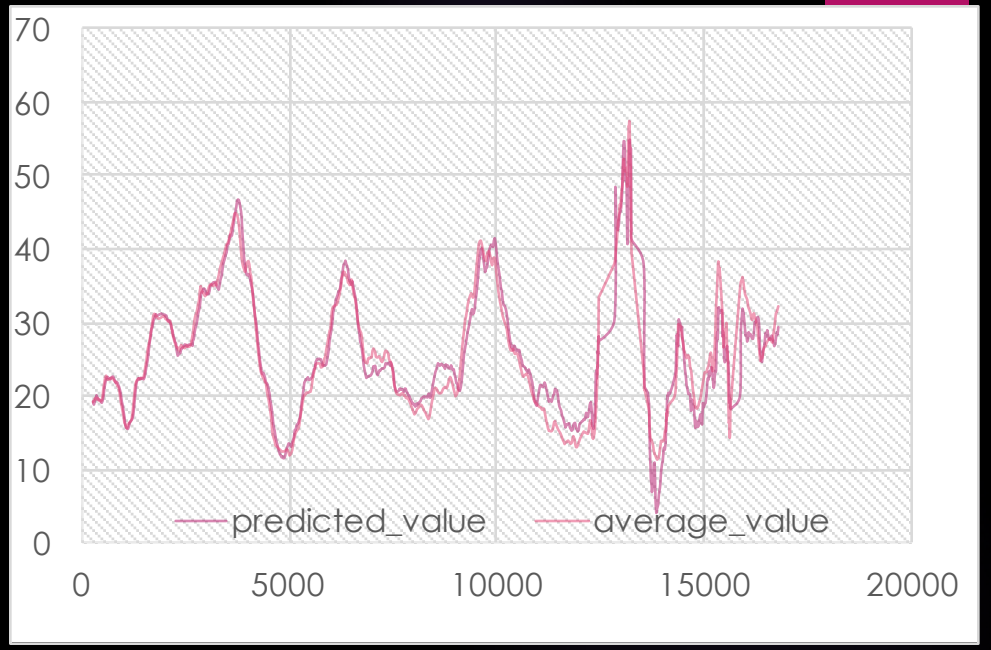
	CH4	CO	CO2	NO	NO2	SO2	H2S
Next 1 hour Fri Jan 27 09:00:00 PST 2006	-0.5281 Unknown	0.0322 Normal	335.5504 Unknown	64.2837 Unknown	32.807 Normal	10.3136 Normal	5.4163 Unknown
Next 12 hour Fri Jan 27 20:00:00 PST 2006	-0.5053 Unknown	0.0292 Normal	335.5443 Unknown	64.3447 Unknown	33.1046 Normal	10.057 Normal	5.4803 Unknown
Next 24 hour Sat Jan 28 08:00:00 PST 2006	-0.5086 Unknown	0.0311 Normal	335.5272 Unknown	64.7029 Unknown	33.117 Normal	10.0335 Normal	5.4734 Unknown

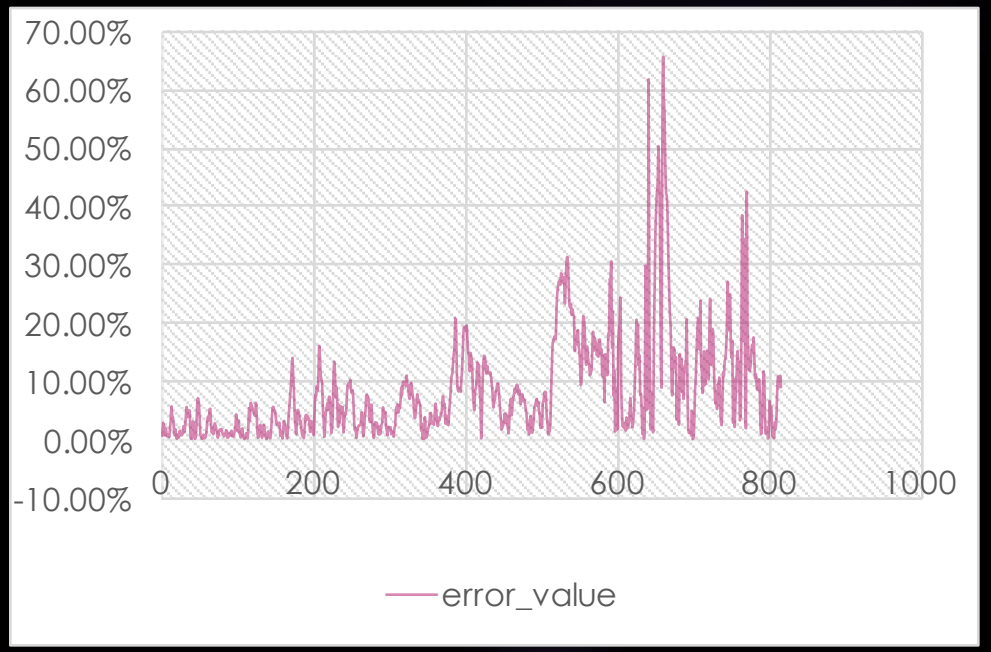
System Reliability

- Data transfer temporally disruption
- Mobile station Shutdown (intentional or unintentional)
- Concurrency problem.
- Dealing with Multi-norms

Conclusion

- System was able to achieve error percentage below 10% within a 15 days of operation only when reading does not contained abnormalities.
- System is being able to adjust itself once spikes in data reading fade away.





GET 2016 - Kuwait - 26th - 28th Jan, 2016.