Quasi Real time Data Analysis for Air Quality Monitoring with an Electronic Nose

H. Zhou, A. V. Shevade, C. C. Pelletier, M. L. Homer, M. A. Ryan

Jet Propulsion Laboratory
California Institute of Technology
Pasadena CA 91109
Outline

• JPL Enose program overview
• Data analysis task Overview
• Data analysis approach
• Data Preprocessing
  – Baseline drift estimation
  – Humidity subtraction
• Sensor calibration and selection
• Functional group identification
• Model building
WHAT IS AN ELECTRONIC NOSE?
An array of non-specific, conductometric chemical sensors which mimics the mammalian nose by recognizing patterns of response.

1. Measure baseline resistance in each polymer-carbon composite sensor to establish $R_0$.

2. Compound from a leak or a spill comes into contact with sensors on the sensing head.

3. The sensors change resistance $\Delta R$ in response to change in composition of the air.

4. Resistance is recorded, $\Delta R/R_0$ is computed, and the distributed response pattern generated.

5. Responses of the sensor array are analyzed and quantified using software developed for the task.

2-propanol - 300 ppm
WHY USE AN ENose for Cabin Air Quality Monitoring?

- Fills gap between an alarm (no id or quantification) and complex analytical instruments
  - ID and quantification of trained-for chemical species
  - Wide dynamic range: fractional ppm to 10,000 ppm
  - Array based sensing; can be trained to detect new species and training data uplinked
  - Runs continuously (30 to 360 data points/hr) and autonomously

- Simple design, robust, rugged, microgravity-insensitive

- Minimal crew interaction required

- Requires no consumables

- Low mass, volume, power

- Readily integrated with larger devices and with monitoring/control systems

- Can analyze volatile aerosols as well as vapors
THREE GENERATIONS OF ENoses:

**Generation 1**
Experiment on STS-95, TRL 6-7
funded by AEMC

- Volume: 2000 cm³ inc. computer
- Mass: 1.4 kg including computer
- Power: 1.5 W ave., 3 W peak
- Detect/ID/Quant 10 compounds at 1 hour SMAC. No real-time data analysis; data acquisition and device control with HP 200LX computer.
- 6 day flight experiment successful.

**Generation 2**
Ground Testing, TRL 5-6
funded by AEMC

- Volume: 750 cm³ w/o computer
- Mass: 800 g w/o computer
- Power: 1.5 W ave., 3 W peak
- Detect/ID/Quant 21 compounds at 24 hour SMAC. Data acquisition and device control possible with PDA computer; real time data analysis with ultra micro computer.
- Extensive ground testing in environmental chamber.

**Generation 3**
Tech. Demo. on ISS, TRL 4
funded by AEMC

- Volume: 3-4 L inc. computer
- Mass: 3-4 kg inc. computer
- Power: 5-8 W ave., ~ 14 W peak
- Detect/ID/Quant 10 compounds at defined concentrations, including Hg, SO₂. Deconvolute mixtures, id unknowns by functional group.
- Data acquisition, device control, real time data analysis, display included.
- Extensive ground testing in environmental chamber; then six month test on-orbit.
CABIN AIR QUALITY MONITORING

FUNCTIONS

- Incident monitor for targeted contaminants exceeding targeted concentrations. Identify and quantify.
- Monitor for presence of compounds associated with fires or overheating electronics
- Monitor clean-up process

CHARACTERISTICS

- Low mass, low power device
- Requires little crew time for maintenance and calibration
- Detects, identifies and quantifies selected chemical species at or below 24 hour SMAC
Enose Data Analysis - Overview

• **Challenges:**
  – Quasi real time, autonomous detection and analysis (Id &Qn)
  – Deconvolution of mixtures of up to three
  – Functional group identification
  
  – Poor consistency in sensor responses
  – Mixed linear and nonlinear sensor responses
  – Similar (correlated) response patterns among target analytes
  – Large dynamic response range (magnitude difference in the order of 4)
  – Baseline drift
  
  – Only partially controlled environment: humidity, temperature, pressure

Multi-channel time-series resistance data

Extracted response pattern at a given time

Software analysis result

>>> 35 ppm toluene+
50 ppm benzene
Development History - 1st Gen

- Investigated / compared various approaches
  - PCA, DFA, NNBP, LA, GA, LM NLS
  - Off-line data analysis (ID and Qn only, no event detection)

- Selected LM – NLS (Levenberg Marquart Nonlinear Least Squares) method
  - Adaptive to both linear and nonlinear sensor responses
  - Recognition on untrained mixtures (of trained single analytes)
    - Significantly less training time, less training data needed
  - Robust to noisy data

- Verification w/ 6-day STS-95 flight experiment data (1998)
Development History - 2nd Gen

- LM-NLS:
  - Expanded analyte list (from 10 to 22)
  - Developed quasi real time data analysis
  - Investigated humidity subtraction technique to uncover small concentrations of analyte in presence of humidity change
  - Investigated temperature effect
  - Investigated func’l group identification w/ sub sensor array

- Extensive Ground Testing
**Development Goals/Requirements - 3rd Gen**

- **Third Generation:**
  - Quasi real time event detection, identification and quantification
  - Ability to uncover small analyte signals under large humidity
  - Ability to classify unknowns by functional group (alcohols or aromatics)
  - Ability to identify compounds not trained for as “unknown”
  - Improvement on data analysis speed, memory efficiency and performance for embedded implementation

- Will undergo extended (6-month) testing on ISS (2008)
Data Analysis - Approach

**Flow Chart (current)**

- **Enose Sensing unit output**
  - Read in new raw data
  - Extract & Preprocess raw data
  - Detect any gas event?
    - Yes: ID. & Qn.
      - Known targets?
        - Yes: Report and Record result
        - No: Classify by funct’l grp?
          - Yes: Model building
          - No: Report and Record result
    - No: Model data

- Mostly developed; need improvement & modification for new sensors /analytes
- Mostly New development
Approaches Investigated

• Discriminant Function Analysis (DFA):
  – Attempt to find hyperplanes dividing the groups.
  – Not suitable when groups are not separable (by a hyperplane); not well suited for mixtures.

• Back Propagation Neural Networks (BPNN):
  – Try to find a best-fit function (linear or nonlinear; no models needed) for the training data.
  – Good for generalization of functions to cases outside the training set.
  – Not suitable when data groups overlap;
  – Normally requires “clean” training data; not well suited for mixtures

• Linear Algebra (LA):
  – Try to find the best-fit (in a least square sense ) parameters vector x from an observation vector y by pseudo-inverse y=Ax. A: system characteristics matrix from training data.
  – Suitable only for linear model.

• Differential Genetic Algorithm (DGA) :
  – Imitates principles of genetics & natural evolution: e.g., recombination, inversion, mutation and selection. Various paths to the optimum are checked; info about them exchanged.
  – Simple concept, easy to use, fast convergence
  – Limited control ability
Data analysis Approach

**Approaches Investigated**

- **Partial least squares (PLS)** †
  - Multivariate signal averaging; inverse linear regression
  - Non-iterative, rapid calculation
  - Can be robust to environmental variations with properly designed calibration

- **Support Vector Machine (SVM)** *
  - Decomposing N-classes classification into multiple 2-classes tasks
  - Need training data of every possible combination of gas compounds
  - Good result on trained data, but poor on untrained ones (poor generalization)

- **Maximum likelihood** *
  - Use covariance of error matrix to construct likelihood function
  - Two step (classification and quantification) procedure

- **Probabilistic Neural Network** *
  - Soft decision (PDF) based classification
  - Much less successful than SVM

---

* Independent work by I-A-I, Inc., on preprocessed JPL 1st Generation ENose data
† Still under investigation
Data Analysis - Approach

**Approach Selected**

- LM-NLS
  - Global & local; classification & quantification - simple code and tree structure
  - Recognition on untrained mixtures (of trained single analytes)
    - Significantly less training data needed (than many other approaches)
  - Robust to noisy data
  - Adaptive to both linear and nonlinear sensor responses:
    - Increase damping (reduce step) for a highly nonlinear problem
    - Decrease damping (increase step) for a linear problem
  - Good generalization (on future data)
Data Analysis - Approach

**Approach - LM NLS Method**

- **Objective**
  - Find best-fit concentration vector $c$ from an observed resistance change vector $dr/r$, which is related to $c$ through a known function, $dr/r=f(A, c)$. i.e., $dr/r=A_1c+A_2c^2$.
  - $A_1$ and $A_2$: sensor coefficient matrix pre-established from training data.

- **Procedure - Iterative optimization:**
  - Establish $A_1$ and $A_2$ from training data (curve fitting)
  - From a given starting point of $c$, calculate discrepancy of the fit:
    $$residual=wt \ast (computed \ f(A, c)-observed \ dr/r)$$
  - update $c$ with a better-fitted (smaller residual) concentration $c$ at each step till max iteration number reached.

  - Repeat for multiple starting point of $c$, and update better-fitted $c$ that favors less number of significant elements
Data Analysis - Approach

Flow chart

- ID and Qn based on characteristic response if any
- Global (all analytes) ID and Qn
  - For most gases, one step global ID & Qn w/ LM-NLS
- Second step local (selective) ID and/or Qn
  - For easily confused gases, after the first step
  - For analytes w/ only few reliable responsive sensors
- Functional group check
  - When residual is too large to be reliable for first step result
- ID of unknowns based on model building

Detected event

Charc. ID & Qn → Global ID & Qn

Local ID and/or Qn

Calibration & modeling data

funct’l grp ID & Qn

Report and Record result

Model building
Data Preprocessing

Enose Sensing unit output

Read in new raw data

Extract & Preprocess raw data

Detect any gas event?

ID. & Qn.

Calibration data

Known targets?

Classify by funct’l grp?

Report and Record result

Mostly developed; need improvement & modification for new sensors /analytes

Mostly New development

No
Data Preprocessing

Converting time-series $R$ to resistance response pattern $dR/R_0$

- General Noise removal
  - Appropriate filtering
- Baseline drift estimation
  - Multiple causes, e.g., changes in temperature, humidity, flow pressure, etc.
  - Need fast and accurate estimation for real time signal detection
- Calculation of resistance change
  - Dynamic relative resistance change: $dR/R_0$
- Event detection

Multi-channel time-series resistance data

Relative resistance calculation

Extracted response pattern
• Baseline drift estimation
  – Investigate temperature and/or humidity caused baseline drift estimation
  – Use appropriate digital filtering (w/ known events if any) to estimate slowly varying baseline drift
  – Additionally,
    • Use clean air reference cycle info to establish piecewise baseline
    • Correct abnormal baseline change due to sudden operating temperature change
Investigation of baseline estimation using temperature & humidity info:

- Correlation (between humidity or temperature or their combination and the actual baseline drift) exists but...
  - The degree of correlation varies greatly among different sensors
  - The degree of correlation also varies greatly among different files (for same or different sensors/analytes)

- Temperature & humidity info (or their combination) not reliable enough to establish actual baseline drift
Correlation exists but... varies from time to time
- Same sensor (sensor #12), different analytes (ammonia, toluene)
Also varies among different sensors
  
  Different sensors (# 23, # 19), same analyte, same data files

Baseline by filtering

raw resistance baseline (green)

humidity (blue)

temperature
• Digital filtering w/ automatic peak rejection
  – Use known (calibration mode) or detected (real time mode) event info to identify peak regions
    • Blank out peak/event regions
    • Retain non-peak regions as baseline points (A)
    • Fill in peak regions with piece-wise Linear fit (B)
  – Estimate baseline by filtering points made of (A+B)
Preprocessing – Baseline Estimation

Baseline estimation -- calibration mode

Green – raw resistance data $R$
Red – smoothed $R$

Red -- -- smoothed $R$
Blue + – points retained from $R$
Green – points filled in for peak areas
Black – baseline estimation by filtering
Preprocessing – Baseline Estimation

**Baseline estimation -- Real time mode**

- Use detected events info to block peak areas (recent past events)
- Piece-wise Linear fit for all peak region
- Data of 2 ~4hrs retained

---

Red - \( \text{-- smoothed R} \)

Blue + \( \text{-- points retained from R} \)

Green \( \text{-- points filled in for peak areas} \)

Black \( \text{-- baseline estimation by filtering} \)

Dec0903Xb, Benzene, sensor 25 minutes
Preprocessing – Event Detection

- Event detection & Pattern Extraction
  - Threshold test of sensors response strength
    - Use selected sensors based on sensor’s limit-of-detection (LOD)
      \[ \text{LOD} = 3\sigma / \text{sensor sensitivity} \]
      \[ \sigma \text{ – standard deviation of a sensor’s response to an analyte} \]

- Humidity subtraction
  - Humidity can easily overwhelm weak signals
    - \( 10^4 \) vs single digit in response magnitude
  - Improve pattern extraction by subtracting large humidity response using measured humidity info
Preprocessing – Humidity Subtraction

- Subtracting humidity response using measured humidity info
- Improved detection/identification of weak analyte signal accompanied with large humidity changes by about 50% in a test done with 30 propanol events

![Graphs a, b, c, d showing propanol, H2O, propanol + H2O, and humidity subtraction](image)
Sensor Calibration & Selection

Enose Sensing unit output

Read in new raw data

Extract & Preprocess raw data

Detect any gas event?

- Yes
  - ID. & Qn.
  - Known targets?
    - Yes
      - Report and Record result
    - No
      - Classify by funct'l grp?
        - Yes
          - Report and Record result
        - No
          - Model building

- No
  - Calibration & Sensor selection
  - Model data
Sensor Calibration - Singles

- Single gases
  - Outliers removal
  - Model fitting and coefficients calculation for each analyte, each sensor
    - e.g., Least square fit of $2^{nd}$ order polynomial, passing (0,0)
  - Statistics:
    - Reliability (inconsistency): variation between actual response vs. fitted response (scattering within an analyte)
    - Selectivity: diversity to different analytes
    - Sensitivity: responsiveness to different analytes
    - Correlation: between concentration and response strength
      - (Euclidean) Fingerprint distances between analyte pair
  - Calibration for different environment conditions
    - Base humidity, temperature, pressure
    - Experimental design?
Sensor Calibration - Singles

Resistance change v.s. Concentration, **Indole**; 2nd order polynomial fit: \( \frac{\Delta r}{r} = A_1c + A_2c^2 \)

- Response vs Concentration, **Indole**
- **fingerprint** @ median conc. = 236ppm, **Indole**

- Experimental data
- Polynomial fit
- If linear fit
Fingerprints of 2nd Gen Analytes

- 2-Butanone
- Chloro benz
- DichloroMethane
- Furan
- Hexane
- DichloroEthane
- tetrahydrofuram
- 111-tce
- Xylenes
- %H20
- Ethyl benz
- Acetonitrile
Sensor Calibration - Singles

Variation, Diversity and Responsiveness of Sensors in Array
Averaged over all analytes at target concentration

Reliability:
Measure of scatter
Variation < 1 preferred

Selectivity (Diversity)
(max-min of norm. fingerprint)
Higher number preferred

Responsiveness
(Sensitivity)
Relative response strength
Higher number preferred
## Sensor Calibration - Singles

### Analyte fingerprint distance

<table>
<thead>
<tr>
<th>analytes #:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>21</th>
<th>22</th>
<th>23</th>
<th>24</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia (1)</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Benzene (2)</td>
<td>34</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethanol (3)</td>
<td>4</td>
<td>28</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Freon113 (5)</td>
<td>34</td>
<td>31</td>
<td>7</td>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indole (6)</td>
<td>19</td>
<td>29</td>
<td>23</td>
<td>31</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methane (7)</td>
<td>18</td>
<td>27</td>
<td>11</td>
<td>16</td>
<td>11</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methanol (8)</td>
<td>25</td>
<td>28</td>
<td>13</td>
<td>10</td>
<td>13</td>
<td>29</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Propanol (9)</td>
<td>7</td>
<td>29</td>
<td>15</td>
<td>9</td>
<td>16</td>
<td>29</td>
<td>15</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toluene (10)</td>
<td>20</td>
<td>29</td>
<td>16</td>
<td>22</td>
<td>14</td>
<td>27</td>
<td>19</td>
<td>22</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetaldehyde (11)</td>
<td>8</td>
<td>35</td>
<td>17</td>
<td>22</td>
<td>17</td>
<td>26</td>
<td>15</td>
<td>18</td>
<td>24</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetone (12)</td>
<td>14</td>
<td>35</td>
<td>28</td>
<td>25</td>
<td>28</td>
<td>35</td>
<td>27</td>
<td>29</td>
<td>23</td>
<td>28</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetonitrile (13)</td>
<td>21</td>
<td>31</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td>27</td>
<td>10</td>
<td>9</td>
<td>16</td>
<td>23</td>
<td>14</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-Butanone (14)</td>
<td>16</td>
<td>30</td>
<td>21</td>
<td>19</td>
<td>22</td>
<td>32</td>
<td>22</td>
<td>23</td>
<td>11</td>
<td>13</td>
<td>23</td>
<td>32</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorobenz (15)</td>
<td>10</td>
<td>30</td>
<td>17</td>
<td>22</td>
<td>15</td>
<td>26</td>
<td>21</td>
<td>22</td>
<td>14</td>
<td>8</td>
<td>29</td>
<td>25</td>
<td>24</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DichloroMethane (16)</td>
<td>15</td>
<td>33</td>
<td>17</td>
<td>16</td>
<td>17</td>
<td>30</td>
<td>19</td>
<td>19</td>
<td>15</td>
<td>18</td>
<td>26</td>
<td>12</td>
<td>19</td>
<td>15</td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furan (17)</td>
<td>42</td>
<td>29</td>
<td>9</td>
<td>15</td>
<td>11</td>
<td>26</td>
<td>10</td>
<td>11</td>
<td>16</td>
<td>20</td>
<td>13</td>
<td>31</td>
<td>7</td>
<td>24</td>
<td>20</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hexane (18)</td>
<td>27</td>
<td>30</td>
<td>7</td>
<td>16</td>
<td>7</td>
<td>24</td>
<td>10</td>
<td>10</td>
<td>16</td>
<td>16</td>
<td>17</td>
<td>28</td>
<td>10</td>
<td>20</td>
<td>16</td>
<td>16</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DichloroEthane (19)</td>
<td>28</td>
<td>28</td>
<td>13</td>
<td>15</td>
<td>15</td>
<td>29</td>
<td>19</td>
<td>15</td>
<td>16</td>
<td>14</td>
<td>27</td>
<td>28</td>
<td>20</td>
<td>17</td>
<td>16</td>
<td>16</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tetrahydrofuram (21)</td>
<td>20</td>
<td>27</td>
<td>18</td>
<td>19</td>
<td>19</td>
<td>32</td>
<td>22</td>
<td>23</td>
<td>13</td>
<td>11</td>
<td>32</td>
<td>28</td>
<td>25</td>
<td>9</td>
<td>14</td>
<td>19</td>
<td>22</td>
<td>18</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111-tce (22)</td>
<td>31</td>
<td>30</td>
<td>7</td>
<td>15</td>
<td>7</td>
<td>24</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>13</td>
<td>20</td>
<td>30</td>
<td>12</td>
<td>18</td>
<td>13</td>
<td>17</td>
<td>9</td>
<td>6</td>
<td>12</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xylenes (23)</td>
<td>6</td>
<td>38</td>
<td>26</td>
<td>34</td>
<td>24</td>
<td>27</td>
<td>32</td>
<td>35</td>
<td>27</td>
<td>18</td>
<td>37</td>
<td>31</td>
<td>35</td>
<td>29</td>
<td>17</td>
<td>31</td>
<td>32</td>
<td>28</td>
<td>29</td>
<td>27</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>%H2O (24)</td>
<td>15</td>
<td>39</td>
<td>23</td>
<td>30</td>
<td>23</td>
<td>30</td>
<td>21</td>
<td>26</td>
<td>31</td>
<td>33</td>
<td>12</td>
<td>42</td>
<td>21</td>
<td>38</td>
<td>35</td>
<td>33</td>
<td>20</td>
<td>24</td>
<td>32</td>
<td>37</td>
<td>27</td>
<td>40</td>
<td></td>
</tr>
<tr>
<td>Ethylbenz (25)</td>
<td>4</td>
<td>47</td>
<td>37</td>
<td>43</td>
<td>33</td>
<td>36</td>
<td>42</td>
<td>44</td>
<td>41</td>
<td>35</td>
<td>44</td>
<td>40</td>
<td>44</td>
<td>42</td>
<td>35</td>
<td>40</td>
<td>44</td>
<td>38</td>
<td>41</td>
<td>41</td>
<td>36</td>
<td>21</td>
<td>49</td>
</tr>
</tbody>
</table>

### Avg distance:

<table>
<thead>
<tr>
<th></th>
<th>19</th>
<th>23</th>
</tr>
</thead>
</table>

### min btwn-class distance:

<table>
<thead>
<tr>
<th></th>
<th>6</th>
</tr>
</thead>
</table>
Sensor Calibration - Mixtures

- Mixture gases
  - Additive linearity (superposition) validation
    polynomial with no cross interaction effect:
    \[ \frac{\Delta r}{r} = A_{1a}c_a + A_{2a}c_a^2 + A_{1b}c_b + A_{2b}c_b^2 \]
  - Alternative mixture model if necessary:
    e.g., polynomial with 2-way cross interaction effect:
    \[ \frac{\Delta r}{r} = A_{1a}c_a + A_{2a}c_a^2 + A_{1b}c_b + A_{2b}c_b^2 + A_{ab}c_ac_b \]
  - Dominant effect of one component?
    - Subtraction technique

Additive linearity holds reasonable well in a previous study
Sensor Calibration - Other Conditions

- Calibration for different environment conditions:
  Nominal ISS Atmosphere conditions:
  - Relative Humidity: 25% - 75%
  - Temperature: 20°C - 34°C
  - Partial Pressure: ~15mmHg

- Investigation of sensor calibration needs at different RH conditions
  - Difference of fitted normalized fingerprints at 5k ppm, 18k ppm, or 23k ppm, against 10k ppm with 2nd gen data
## Sensor Calibration - Other Conditions

### Changes for diff humidity: 10k vs 5k

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Fngr chg</th>
<th>Mag chg</th>
<th>Calib?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>33</td>
<td>32</td>
<td>x</td>
</tr>
<tr>
<td>Benzene</td>
<td>61</td>
<td>11</td>
<td>yes</td>
</tr>
<tr>
<td>Ethanol</td>
<td>17</td>
<td>17</td>
<td>no</td>
</tr>
<tr>
<td>Freon113</td>
<td>25</td>
<td>1</td>
<td>x</td>
</tr>
<tr>
<td>Indole</td>
<td>18</td>
<td>16</td>
<td>no</td>
</tr>
<tr>
<td>Methane</td>
<td>28</td>
<td>9</td>
<td>x</td>
</tr>
<tr>
<td>Methanol</td>
<td>22</td>
<td>20</td>
<td>no</td>
</tr>
<tr>
<td>2-Propanol</td>
<td>30</td>
<td>8</td>
<td>x</td>
</tr>
<tr>
<td>Toluene</td>
<td>17</td>
<td>2</td>
<td>no</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>15</td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>Acetone</td>
<td>21</td>
<td>1</td>
<td>x</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>29</td>
<td>24</td>
<td>x</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>22</td>
<td>11</td>
<td>no</td>
</tr>
<tr>
<td>Chloro bnz</td>
<td>14</td>
<td>9</td>
<td>no</td>
</tr>
<tr>
<td>DichloroMethane</td>
<td>29</td>
<td>23</td>
<td>x</td>
</tr>
<tr>
<td>Furan</td>
<td>23</td>
<td>20</td>
<td>no</td>
</tr>
<tr>
<td>Hexane</td>
<td>33</td>
<td>5</td>
<td>x</td>
</tr>
<tr>
<td>DichloroEthane</td>
<td>30</td>
<td>27</td>
<td>x</td>
</tr>
<tr>
<td>tetrahydrofuram</td>
<td>46</td>
<td>41</td>
<td>x</td>
</tr>
<tr>
<td>111-tce</td>
<td>38</td>
<td>29</td>
<td>x</td>
</tr>
<tr>
<td>Xylenes</td>
<td>15</td>
<td>12</td>
<td>no</td>
</tr>
<tr>
<td>%H2O</td>
<td>25</td>
<td>25</td>
<td>x</td>
</tr>
<tr>
<td>Ethyl bnz</td>
<td>4</td>
<td>1</td>
<td>no</td>
</tr>
</tbody>
</table>

*yes – need recalib
x – borderline
no – no need recalib

*Change in average success rate if use 10k coefficient for 5k data: - 4%

Most...Least changed Sensors: 8, 16, 4, 2, 12, 6, 27, 5, 14, 25, 17, ..., 18, 23, 30, 22, 31, 24, 28, 21

(Mean % change in fngrprnt): 60, 55, 48, 44, 39, 39, 36, 35, 34, 33, ..., 16, 15, 15, 13, 13, 12, 10, 7
### Sensor Calibration - Other Conditions

#### Changes for diff humidity: 10k vs 18k

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Fngr chg</th>
<th>Mag chg</th>
<th>Calib?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>34</td>
<td>1</td>
<td>x</td>
</tr>
<tr>
<td>Benzene</td>
<td>20</td>
<td>1</td>
<td>no</td>
</tr>
<tr>
<td>Ethanol</td>
<td>25</td>
<td>18</td>
<td>no</td>
</tr>
<tr>
<td>Freon113</td>
<td>29</td>
<td>3</td>
<td>x</td>
</tr>
<tr>
<td>Indole</td>
<td>40</td>
<td>22</td>
<td>x</td>
</tr>
<tr>
<td>Methane</td>
<td>99</td>
<td>99</td>
<td>yes</td>
</tr>
<tr>
<td>Methanol</td>
<td>30</td>
<td>3</td>
<td>x</td>
</tr>
<tr>
<td>2-Propanol</td>
<td>43</td>
<td>36</td>
<td>x</td>
</tr>
<tr>
<td>Toluene</td>
<td>48</td>
<td>26</td>
<td>x</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>21</td>
<td>18</td>
<td>no</td>
</tr>
<tr>
<td>Acetone</td>
<td>27</td>
<td>26</td>
<td>x</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>26</td>
<td>5</td>
<td>x</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>31</td>
<td>28</td>
<td>x</td>
</tr>
<tr>
<td>Chloro benz</td>
<td>38</td>
<td>31</td>
<td>x</td>
</tr>
<tr>
<td>DichloroMethane</td>
<td>49</td>
<td>47</td>
<td>x</td>
</tr>
<tr>
<td>Furan</td>
<td>34</td>
<td>24</td>
<td>x</td>
</tr>
<tr>
<td>Hexane</td>
<td>36</td>
<td>34</td>
<td>x</td>
</tr>
<tr>
<td>DichloroEthane</td>
<td>23</td>
<td>17</td>
<td>no</td>
</tr>
<tr>
<td>tetrahydrofuram</td>
<td>78</td>
<td>40</td>
<td>yes</td>
</tr>
<tr>
<td>111-tce</td>
<td>31</td>
<td>27</td>
<td>x</td>
</tr>
<tr>
<td>Xylenes</td>
<td>28</td>
<td>23</td>
<td>x</td>
</tr>
<tr>
<td>%H2O</td>
<td>19</td>
<td>5</td>
<td>no</td>
</tr>
<tr>
<td>Ethyl benz</td>
<td>6</td>
<td>5</td>
<td>no</td>
</tr>
</tbody>
</table>

- yes – need recalib
- x – borderline
- n – no need recalib

---

*Change in Average success rate if use 10k coefficient for 18k data: -15%*

Most...Least changed Sensors: 8   16   5   6   12   4   17   14   2   25   27 ....   31   28   1   29   24   22   30   21

(Mean % change in fngrprnt): 92 74 70 65 59 54 51 49 46 45 40 .... 19 18 17 16 15 12 10
Changes for diff humidity: 10k vs 23k

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Fngr</th>
<th>Mag</th>
<th>Calib?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>46</td>
<td>10</td>
<td>x</td>
</tr>
<tr>
<td>Benzene</td>
<td>37</td>
<td>26</td>
<td>x</td>
</tr>
<tr>
<td>Ethanol</td>
<td>40</td>
<td>31</td>
<td>x</td>
</tr>
<tr>
<td>Freon113</td>
<td>45</td>
<td>41</td>
<td>x</td>
</tr>
<tr>
<td>Indole</td>
<td>46</td>
<td>22</td>
<td>x</td>
</tr>
<tr>
<td>Methane</td>
<td>100</td>
<td>100</td>
<td>yes</td>
</tr>
<tr>
<td>Methanol</td>
<td>34</td>
<td>26</td>
<td>x</td>
</tr>
<tr>
<td>2-Propanol</td>
<td>46</td>
<td>34</td>
<td>x</td>
</tr>
<tr>
<td>Toluene</td>
<td>58</td>
<td>37</td>
<td>yes</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>31</td>
<td>9</td>
<td>x</td>
</tr>
<tr>
<td>Acetone</td>
<td>61</td>
<td>7</td>
<td>yes</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>40</td>
<td>26</td>
<td>x</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>56</td>
<td>53</td>
<td>yes</td>
</tr>
<tr>
<td>Chloro benz</td>
<td>54</td>
<td>44</td>
<td>yes</td>
</tr>
<tr>
<td>DichloroMethane</td>
<td>58</td>
<td>58</td>
<td>yes</td>
</tr>
<tr>
<td>Hexane</td>
<td>39</td>
<td>34</td>
<td>x</td>
</tr>
<tr>
<td>DichloroEthane</td>
<td>36</td>
<td>27</td>
<td>x</td>
</tr>
<tr>
<td>tetrahydrofuram</td>
<td>51</td>
<td>43</td>
<td>yes</td>
</tr>
<tr>
<td>111-tce</td>
<td>40</td>
<td>40</td>
<td>x</td>
</tr>
<tr>
<td>Xylenes</td>
<td>50</td>
<td>9</td>
<td>x</td>
</tr>
<tr>
<td>Ethyl benz</td>
<td>71</td>
<td>26</td>
<td>yes</td>
</tr>
</tbody>
</table>

Avg=50%

Avg=33%

* Change in average success rate if use 10k coefficient for 23k data: - 20%

Most...Least changed Sensors: 5 17 25 6 8 12 14 4 2 3 16 .... 31 30 22 28 24 9 1 21

(Mean % change in fngprnt): 94 92 87 85 80 73 67 60 58 .... 31 30 25 24 24 24 23 19
### Changes for diff humidity: 18k vs 23k

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Fngr</th>
<th>Mag</th>
<th>Calib?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>31</td>
<td>15</td>
<td>x</td>
</tr>
<tr>
<td>Benzene</td>
<td>28</td>
<td>23</td>
<td>x</td>
</tr>
<tr>
<td>Ethanol</td>
<td>19</td>
<td>15</td>
<td>no</td>
</tr>
<tr>
<td>Freon113</td>
<td>41</td>
<td>21</td>
<td>x</td>
</tr>
<tr>
<td>Indole</td>
<td>14</td>
<td>0</td>
<td>no</td>
</tr>
<tr>
<td>Methane</td>
<td>61</td>
<td>44</td>
<td>yes</td>
</tr>
<tr>
<td>Methanol</td>
<td>33</td>
<td>25</td>
<td>x</td>
</tr>
<tr>
<td>2-Propanol</td>
<td>60</td>
<td>5</td>
<td>yes</td>
</tr>
<tr>
<td>Toluene</td>
<td>15</td>
<td>15</td>
<td>no</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>14</td>
<td>7</td>
<td>no</td>
</tr>
<tr>
<td>Acetone</td>
<td>17</td>
<td>8</td>
<td>no</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>20</td>
<td>14</td>
<td>no</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>47</td>
<td>35</td>
<td>x</td>
</tr>
<tr>
<td>Chloro benz</td>
<td>29</td>
<td>18</td>
<td>x</td>
</tr>
<tr>
<td>DichloroMethane</td>
<td>39</td>
<td>20</td>
<td>x</td>
</tr>
<tr>
<td>Hexane</td>
<td>20</td>
<td>1</td>
<td>no</td>
</tr>
<tr>
<td>DichloroEthane</td>
<td>20</td>
<td>15</td>
<td>no</td>
</tr>
<tr>
<td>tetrahydrofuram</td>
<td>123</td>
<td>118</td>
<td>yes</td>
</tr>
<tr>
<td>111-tce</td>
<td>26</td>
<td>19</td>
<td>x</td>
</tr>
<tr>
<td>Xylenes</td>
<td>45</td>
<td>24</td>
<td>x</td>
</tr>
<tr>
<td>Ethyl benz</td>
<td>75</td>
<td>32</td>
<td>yes</td>
</tr>
</tbody>
</table>

- **Yes** – need calib
- **X** – borderline
- **no** – no need calib

**Most...Least changed Sensors:**

5  8  16  17  2  12  25  6  14  4  13 ....  31  20  19  24  1  9  21  28

(Mean % change in fngrprnt):

104  95  76  67  66  64  63  59  48  48  41 ....  19  18  17  15  14  12  10

**Avg=37%**

**Avg=23%**
Sensor Calibration - Other Conditions

Sensor Coeff vs Base humidity – Linear approximation

Ethanol
Avg correlation:
$r^2=0.95$
Sensor Calibration - Other Conditions

Summary:

• Base humidity change affects data analysis results
  – Within 5k are almost tolerable; need calibration for bigger change

• Sensor coefficient vs base humidity can be roughly modeled as linear
  – Correlation >0.80 mostly; could be higher if we have better calibration data set

• Calibration needs / strategy:
  – Calibration at two base humidity levels, e.g., 10k & 20k (or 18k)
  – Interpolate/Extrapolate to get the rest
  – Similarly, calibration at two temperature levels
Sensor Selection

- Poorly behaved sensors can be more detrimental to data analysis than being helpful.

- Optimal sensor array selection and weighting
  - For global classification and quantification
  - For local classification and/or quantification
  - For functional group classification and quantification
  - For event detecting
  - Assistance in sensor material selection

- ENose sensor selection challenge:
  - Selectivity is important while reliability and sensitivity can be major limiting factors in detecting target compounds reliably
  - Meaningful selection that will actually improve data analysis result
Sensor Selection

• Sensor selection approaches
  – PCA, Cluster Analysis
  – ANOVA-based approach
  – Genetic algorithm optimization w/ appropriate statistic measures

• Sensor selection by ANOVA-based technique:
  \[ f = (\text{selectivity} / \text{reliability}) \times \text{sensitivity} \times \text{correlation} \]
  – Selectivity (cross-class variation):
    • Difference b/w max, min relative response strengths among all analytes
    • Mean of differences of relative response strengths among analyte pairs
  – Reliability (within-class variation): mean variation b/w relative actual vs. fitted responses
  – Sensitivity: mean of a sensor’s relative response strengths to all analytes
  – Correlation: b/w concentration and response strength
Sensor Selection

• Sensor array optimization:
  – Class separation vs threshold of $f$ (worst sensors removed)
    
    Class separation = cross-class fingerprint distance / within-class fingerprint variation

![Class separation vs threshold](image)

2nd generation sensors:

The Best sensors:

30 31 22 25 23 10 17 27...

Best class separation w/ these worst sensors removed:

4 12 8 5 16 20 13 14 1 6
Functional Group Classification

- Enose Sensing unit output
- Read in new raw data
- Extract & Preprocess raw data
  - Detect any gas event?
    - Yes
      - ID. & Qn.
      - Known targets?
        - Yes
          - Classify by funct’l grp?
            - Yes
              - Report and Record result
            - No
              - Calibrate & Sensor Selection
              - Model data
        - No
          - Model building
    - No
      - Model data

Mostly developed; need improvement & modification for new sensors/analytes
Mostly New development
Analytes can be grouped into “families” or functional groups

- Based on similarities in physical structure /externally observed properties, e.g.,
  - Alcohol: ethanol, 1-, 2- Butanol, 1-, 2- propanol, etc.,
  - Aromatic: benzene, toluene, chlorobenzene, furan, etc.,
  - One nitrogen: ammonia, acetonitrile, indole, etc.
  - Ketone: acetone, 2-butanone, etc.

...But not strictly member of one and only one functional group

Desire to identify an “unknown” (not-trained-for) analyte according to its functional group
- e.g., identify methanol as an alcohol without being trained for
NLS with sensor subset optimized for functional groups

- Select optimum subset sensors that minimize the distance within a functional grp and maximize the distance between the rest of analytes.

- Result in the 1st (regular Id&Qn) step be given preference in the 2nd (functional group Id) step:
  - e.g., If the 1st step results in “methanol + freon” with an unacceptably large residue, then in the 2nd step, check first if “alcohol+freon” would result in an acceptable residue.

- Other possibilities will also be checked and the result will be accepted only if the associated residue is less than that in the above.

- If all residues are greater than acceptable levels → unknowns.
Funct’l Group Classification – Case Study

- Case studied:
  - Use simulated sensor response (from molecular modeling)
    - 10 trained target analytes: methanol, ethanol, 1-proponal, methane, ammonia, benzene, formaldehyde, freon 113, indole, toluene
    - 4 untrained alcohols: 2-propanol, butanol, pentanol, phenol
    - 5 untrained aromatic: anthracence, ethyl benzene, p-xylene, and m-xylene
  - 170 simulated binary mixture test events:
    - One trained analyte + one untrained analyte, or two untrained analytes
    - Different concentration combinations
  - Two possible functional groups
    - alcohol and aromatic
  - Average success rate ~ 84%
    - 2-propanol (86%), butanol (67%), pentanol (76%), phenol (67%),
    - Anthracene (86%), Ethyl benzene (95%), p-xylene (95%), o-xylene (86%), and m-xylene (100%).
Model Building

Enose Sensing unit output

- Read in new raw data
  - Extract & Preprocess raw data
    - Detect any gas event?
      - Yes
        - ID. & Qn.
          - Known targets?
            - Yes
              - Calibration & Sensor selection
                - Model data
                  - Model building
              - Report and Record result
        - No
          - Model building
            - Classify by funct’l grp?
              - Yes
                - Model building
              - No
                - Report and Record result
  - No
    - Mostly developed; need improvement & modification for new sensors /analytes
    - Mostly New development
• Can we determine the identity of an unknown chemical which causes the sensor array to respond by models built from simulation?
  – Sensor modeling support
Sensor Modeling Support

• Objectives
  – Develop the ability to determine the identity of an unknown chemical which causes the sensor array to respond.
  – Provide characteristic sensor response of untrained analytes for functional group identification based on known responses to trained ones
  – Predict sensor’s response under different base humidity given response under nominal humidity

• Approach
  – Quantitative Structure Activity Relationships (QSAR)
MODEL OF SENSOR RESPONSE

- **Goal**
  - Develop representative response equation (using known experimental data) for each polymer-carbon sensing film in the array (16 equations for the array) to generate virtual training sets for any given sensor array.

  ![Graph](image)

  **Experimental sensor response fit:**
  
  Resistance change ($y$) = $A_1 x + A_2 x^2$

  $A_1$ and $A_2$ : constant; $x$: concentration

  Correlate the constant $A_1$ from experimental response fit with molecular descriptors using Quantitative Structure-Activity Relationships (QSAR).
MODEL OF SENSOR RESPONSE

QSAR METHODOLOGY

Generate molecular models for polymer-carbon sensing film
Calculate analyte interactions energies w/ polymer-carbon sensing film

Generate QSAR table
Calculate molecular descriptors (analyte and sensor response)

Correlate sensor response to molecular descriptors
- Generate QSAR equations using Genetic Algorithms
- Select the statistically most significant equation containing the polymer-analyte interaction term

Experimental sensor response (training set)
# MODEL OF SENSOR RESPONSE

## QSAR DESCRIPTORS

<table>
<thead>
<tr>
<th>Intrinsic analyte descriptors</th>
<th>Sensor interaction energy descriptors</th>
</tr>
</thead>
<tbody>
<tr>
<td>molar Refractivity</td>
<td>polymer-analyte</td>
</tr>
<tr>
<td>size of the molecule</td>
<td>polymer-water</td>
</tr>
<tr>
<td>H-bond donor/acceptor</td>
<td>analyze-water</td>
</tr>
<tr>
<td>dipole</td>
<td>analyze-analyte</td>
</tr>
<tr>
<td>principal moment of inertia</td>
<td>polymer-carbon black</td>
</tr>
<tr>
<td></td>
<td>analyte-carbon black</td>
</tr>
<tr>
<td></td>
<td>(Electrostatic, vdW, H-bond contributions)</td>
</tr>
</tbody>
</table>

### Descriptor calculations

**Sensor interaction energy**

- **Monte Carlo technique (sorption calculation)**

**Intrinsic analyte descriptors**

- **Quantitative Structure-Property Relationships (QSPR)**
MODEL OF SENSOR RESPONSE

Array Response to Ammonia

Polymer list (sensors #)

1. Poly (4-vinylphenol-co-methyl methacrylate)
2. Poly (ethylene-co-acrylic acid)
3. Poly (styrene-co-maleic acid)
4. Poly (2-vinyl pyridine)
5. Poly (4-vinylpyridine)
6. Vinyl alcohol/vinyl butyral copolymer
7. Cyanoethyl hydroxyethyl cellulose
8. Soluble polyimide
9. Polyepichlorohydrin
10. Poly (epichlorohydrin-co-ethylene oxide)
11. Polyethylene oxide
**Predictive Model**

Correlate polymer and analyte interactions to $A_1$ from analysis algorithm.

**PEO-carbon composite film**

Calculated Coefficient $A_1 =$

$$+ 0.116 \quad HB_{D^2}$$

$$+ 2.4E-03 \quad MR^2$$

$$+ 0.152 \quad E_{pa}$$

$pKa$ (site1, 25°C) = 15.24 ± 0.10

**pKa calculations:**

ACD/Lab online software
Backup Material
Data Analysis – PCA analysis

- PCA plot of 2\textsuperscript{nd} generation Enose data on 22 analytes
Data Analysis – Cluster Analysis

- Cluster Analysis of 2nd generation Enose data, Single linkage
Singles:

- Overall success rate ~ 85%.
  - "Success" is correct identification and quantification (within ± 50%).
- Most failed cases (~10%) are no-event identification -- concentration below identifiable threshold

<table>
<thead>
<tr>
<th>Gas Compound</th>
<th>Conc Range tested (ppm)</th>
<th>Success Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>10 - 50</td>
<td>100</td>
</tr>
<tr>
<td>Benzene</td>
<td>20 - 150</td>
<td>88</td>
</tr>
<tr>
<td>Ethanol</td>
<td>10 - 130</td>
<td>87</td>
</tr>
<tr>
<td>Freon 113</td>
<td>50 - 525</td>
<td>80</td>
</tr>
<tr>
<td>Formaldehyde</td>
<td>50 - 510</td>
<td>100</td>
</tr>
<tr>
<td>Indole</td>
<td>.006 - .06</td>
<td>80</td>
</tr>
<tr>
<td>Methane</td>
<td>3000 - 7000</td>
<td>75</td>
</tr>
<tr>
<td>Methanol</td>
<td>10 - 300</td>
<td>63</td>
</tr>
<tr>
<td>Propanol</td>
<td>75 - 180</td>
<td>80</td>
</tr>
<tr>
<td>Toluene</td>
<td>30 - 60</td>
<td>50</td>
</tr>
<tr>
<td>%Relative Humidity</td>
<td>5 - 65</td>
<td>100</td>
</tr>
<tr>
<td>Wipe</td>
<td>500 - 4000</td>
<td>100</td>
</tr>
</tbody>
</table>
Data Analysis Results - 1st generation

Mixtures

- Moderate overall success rate ~60%.
- No training data on mixtures used; total ~75 testing events.
- Sensor loosely follows additive linearity property for the tested binary mixtures.
- Those of low success rates are largely due to low concentrations of one of the mixture components (for which detecting even the single gases is difficult)

<table>
<thead>
<tr>
<th>Gas Compound 1</th>
<th>Conc. range (ppm)</th>
<th>Gas Compound 2</th>
<th>Conc. range (ppm)</th>
<th>Success Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ammonia</td>
<td>20-50</td>
<td>ethanol</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>ammonia</td>
<td>10-20</td>
<td>benzene</td>
<td>20-60</td>
<td>65</td>
</tr>
<tr>
<td>benzene</td>
<td>50-150</td>
<td>methanol</td>
<td>50</td>
<td>100</td>
</tr>
<tr>
<td>benzene</td>
<td>30-100</td>
<td>propanol</td>
<td>75 – 100</td>
<td>60</td>
</tr>
<tr>
<td>benzene</td>
<td>10-30</td>
<td>formaldehyde</td>
<td>85 – 250</td>
<td>20</td>
</tr>
<tr>
<td>ethanol</td>
<td>40-130</td>
<td>wipe</td>
<td>60-185</td>
<td>90</td>
</tr>
<tr>
<td>methanol</td>
<td>25</td>
<td>propanol</td>
<td>80-185</td>
<td>25</td>
</tr>
<tr>
<td>methanol</td>
<td>50-75</td>
<td>toluene</td>
<td>30-60</td>
<td>40</td>
</tr>
</tbody>
</table>
FLIGHT DATA ANALYSIS

• Every daily marker was correctly identified as 2-propanol. In some cases it was identified as 2-propanol plus a humidity change.

• All peaks other than daily markers were identified as humidity change.

• In every case where the independent humidity monitor in the cabin recorded a humidity change, it was detected by the ENose. Humidity change quantification by the ENose corresponded to the change recorded by the cabin monitor.

• Independent air samples analyzed by Gas Chromatography/Mass Spectrometry showed no presence of contaminants that should have been detected by the ENose.
Data Analysis Results - 2nd generation

- **Ground Testing**

<table>
<thead>
<tr>
<th>Case</th>
<th>Miss</th>
<th>False</th>
<th>Id</th>
<th>Id &amp;Qn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1: all/all:</td>
<td>4%</td>
<td>9%</td>
<td>78%</td>
<td>73%</td>
</tr>
<tr>
<td>Case 2a: half/all:</td>
<td>4%</td>
<td>10%</td>
<td>78%</td>
<td>73%</td>
</tr>
<tr>
<td>Case 3b: third/all:</td>
<td>4%</td>
<td>13%</td>
<td>78%</td>
<td>72%</td>
</tr>
<tr>
<td>Case 4c: 1st half / 2nd half</td>
<td>4%</td>
<td>11%</td>
<td>77%</td>
<td>73%</td>
</tr>
<tr>
<td>Case 5d: all / 2nd half:</td>
<td>4%</td>
<td>11%</td>
<td>77%</td>
<td>74%</td>
</tr>
<tr>
<td>Case 6e: 10k coeff/5k data:</td>
<td>3%</td>
<td>11%</td>
<td>73%</td>
<td>67%</td>
</tr>
</tbody>
</table>

- a. sort events from all data files by conc for each analytes, then use odd cased for training and all cases for testing
- b. Similar as a, but use every other 2 cases for training and all cases for testing
- c. Training on 1st (early) half data files; testing on (later) 2nd half data.
- d. Testing on 2nd half data which are part of training data
- e. Training data are of RH @10kppm, test data are of RH @5kppm
• Breakdown success rates:

<table>
<thead>
<tr>
<th>Compound</th>
<th>Concentration range</th>
<th>Miss Rate</th>
<th>False Alarm Rate</th>
<th>ID Rate</th>
<th>IQ Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonia</td>
<td>6 - 60</td>
<td>0%</td>
<td>5%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Benzene</td>
<td>3 - 75</td>
<td>4%</td>
<td>3%</td>
<td>79%</td>
<td>79%</td>
</tr>
<tr>
<td>Ethanol</td>
<td>200-6000</td>
<td>0%</td>
<td>1%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Freon113</td>
<td>15 - 500</td>
<td>10%</td>
<td>12%</td>
<td>63%</td>
<td>56%</td>
</tr>
<tr>
<td>Indole</td>
<td>25 - 415</td>
<td>13%</td>
<td>42%</td>
<td>79%</td>
<td>79%</td>
</tr>
<tr>
<td>Methane</td>
<td>1550 - 15000</td>
<td>6%</td>
<td>22%</td>
<td>64%</td>
<td>36%</td>
</tr>
<tr>
<td>Methanol</td>
<td>6 - 100</td>
<td>5%</td>
<td>13%</td>
<td>55%</td>
<td>45%</td>
</tr>
<tr>
<td>2-Propano</td>
<td>30 - 400</td>
<td>0%</td>
<td>1%</td>
<td>96%</td>
<td>96%</td>
</tr>
<tr>
<td>Toluene</td>
<td>5 - 50</td>
<td>0%</td>
<td>6%</td>
<td>86%</td>
<td>76%</td>
</tr>
<tr>
<td>Acetaldehyde</td>
<td>4 - 20</td>
<td>4%</td>
<td>5%</td>
<td>88%</td>
<td>88%</td>
</tr>
<tr>
<td>Acetone</td>
<td>65 - 600</td>
<td>3%</td>
<td>1%</td>
<td>91%</td>
<td>91%</td>
</tr>
<tr>
<td>Acetonitrile</td>
<td>4 - 25</td>
<td>0%</td>
<td>11%</td>
<td>85%</td>
<td>81%</td>
</tr>
<tr>
<td>2-Butanone</td>
<td>15 - 150</td>
<td>4%</td>
<td>8%</td>
<td>75%</td>
<td>71%</td>
</tr>
<tr>
<td>Chloro benz</td>
<td>6 - 30</td>
<td>8%</td>
<td>17%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>DichloroMethane</td>
<td>10 - 150</td>
<td>7%</td>
<td>18%</td>
<td>86%</td>
<td>79%</td>
</tr>
<tr>
<td>Furan</td>
<td>1 - 35</td>
<td>0%</td>
<td>9%</td>
<td>60%</td>
<td>53%</td>
</tr>
<tr>
<td>Hexane</td>
<td>16 - 150</td>
<td>0%</td>
<td>11%</td>
<td>71%</td>
<td>63%</td>
</tr>
<tr>
<td>DichloroEthane</td>
<td>10 - 100</td>
<td>0%</td>
<td>6%</td>
<td>44%</td>
<td>42%</td>
</tr>
<tr>
<td>Tetrahydrofuram</td>
<td>13 - 120</td>
<td>0%</td>
<td>4%</td>
<td>71%</td>
<td>71%</td>
</tr>
<tr>
<td>111-tce</td>
<td>7 - 200</td>
<td>13%</td>
<td>18%</td>
<td>49%</td>
<td>47%</td>
</tr>
<tr>
<td>Xylenes</td>
<td>33 - 300</td>
<td>0%</td>
<td>24%</td>
<td>96%</td>
<td>96%</td>
</tr>
<tr>
<td>%H20</td>
<td>1000 - 13400</td>
<td>7%</td>
<td>4%</td>
<td>89%</td>
<td>89%</td>
</tr>
<tr>
<td>Ethyl benz</td>
<td>20 - 185</td>
<td>0%</td>
<td>2%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

(conc range)  Avg= miss= 4%, false= 11%, Id= 77%, IQ=73%,
Enose Data Analysis - Overview

Sensor response recorded in an Early Human Testing (EHT) chamber test

When the EHT chamber door was opened, there is significant change in resistance in Enose sensors. Dotted vertical lines show 30 minute reference cycles with a corresponding dip in sensor response.
JPL Electronic Nose (ENose) Comparison w/State of Art

Relative to current ISS monitoring
- No real-time event monitor currently exists aboard ISS.

Relative to higher-end analytical instruments (e.g. VOA, VCAM):
- Low mass, small volume, low power, rugged.
- No consumables required.
- Runs continuously (30 to 360 data points/hr).
- Excellent complement to these instruments - can serve as trigger to activate them.

Relative to other ENoses, which are based on pure conducting polymers:
- Much wider range of usable polymers - superior fingerprinting.
- Quantification included in analyte identification.
- Superior environmental stability.
- Tunable resistance properties - facilitates readout.

Relative to materials-specific sensors:
- Can be trained to recognize new materials and unknowns.
- Can be used to deconvolute and identify simple mixtures of knowns.
- One array serves to detect and identify a wide variety of analytes.
## JPL Enose Program Overview and Background

### Second Generation JPL ENose: 21 Analytes

tested concentrations

<table>
<thead>
<tr>
<th></th>
<th>24 hr SMAC (ppm)</th>
<th>1/3 SMAC (ppm)</th>
<th>3x SMAC (ppm)</th>
<th>low (ppm)</th>
<th>high (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetaldehyde (1500 ppm gas)</td>
<td>6</td>
<td>2.00</td>
<td>18</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>acetone</td>
<td>200</td>
<td>66.67</td>
<td>600</td>
<td>65</td>
<td>600</td>
</tr>
<tr>
<td>acetonitrile (1500 ppm gas)</td>
<td>4</td>
<td>1.33</td>
<td>12</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>ammonia - NH3</td>
<td>20</td>
<td>6.67</td>
<td>60</td>
<td>6</td>
<td>60</td>
</tr>
<tr>
<td>benzene (1510 ppm gas)</td>
<td>3</td>
<td>1.00</td>
<td>9</td>
<td>3</td>
<td>75</td>
</tr>
<tr>
<td>2-butanone</td>
<td>50</td>
<td>16.67</td>
<td>150</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td>chlorobenzene</td>
<td>10</td>
<td>3.33</td>
<td>30</td>
<td>3</td>
<td>30</td>
</tr>
<tr>
<td>dichloromethane</td>
<td>35</td>
<td>11.67</td>
<td>105</td>
<td>11</td>
<td>150</td>
</tr>
<tr>
<td>dichloroethane</td>
<td>0.4</td>
<td>0.13</td>
<td>1.2</td>
<td>0.12</td>
<td>98</td>
</tr>
<tr>
<td>ethanol</td>
<td>2000</td>
<td>666.67</td>
<td>6000</td>
<td>665</td>
<td>6000</td>
</tr>
<tr>
<td>ethyl benzene</td>
<td>60</td>
<td>20.00</td>
<td>180</td>
<td>16</td>
<td>184</td>
</tr>
<tr>
<td>freon 113</td>
<td>50</td>
<td>16.67</td>
<td>150</td>
<td>15</td>
<td>150</td>
</tr>
<tr>
<td>hexane</td>
<td>50</td>
<td>16.67</td>
<td>150</td>
<td>16</td>
<td>150</td>
</tr>
<tr>
<td>indole</td>
<td>0.3</td>
<td>0.10</td>
<td>0.9</td>
<td>0.10</td>
<td>0.90</td>
</tr>
<tr>
<td>methane</td>
<td>5300</td>
<td>1766.67</td>
<td>15900</td>
<td>3000</td>
<td>40000</td>
</tr>
<tr>
<td>methanol</td>
<td>10</td>
<td>3.33</td>
<td>30</td>
<td>3</td>
<td>100</td>
</tr>
<tr>
<td>isopropanol</td>
<td>100</td>
<td>33.33</td>
<td>300</td>
<td>30</td>
<td>600</td>
</tr>
<tr>
<td>tetrahydrofuran</td>
<td>40</td>
<td>13.33</td>
<td>120</td>
<td>13</td>
<td>120</td>
</tr>
<tr>
<td>toluene</td>
<td>16</td>
<td>5.33</td>
<td>48</td>
<td>5</td>
<td>50</td>
</tr>
<tr>
<td>1,1,1-trichloroethane</td>
<td>11</td>
<td>3.67</td>
<td>33</td>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>o,p-xylenes</td>
<td>100</td>
<td>33.33</td>
<td>300</td>
<td>33</td>
<td>300</td>
</tr>
</tbody>
</table>
## Analytes for Technical Demonstration Onboard ISS

<table>
<thead>
<tr>
<th>TIER</th>
<th>ANALYTE</th>
<th>Target Con (ppm)</th>
<th>JUSTIFICATION</th>
<th>QUANT. RANGE (ppm)</th>
<th>ODOR THRESHOLD* (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ammonia</td>
<td>5</td>
<td>Coolant; 24 hr SMAC 20</td>
<td>1.6 – 15</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>Mercury</td>
<td>0.01</td>
<td>Hg vapor in lighting</td>
<td>.003 - .03</td>
<td>odorless</td>
</tr>
<tr>
<td></td>
<td>Sulfur Dioxide</td>
<td>1</td>
<td>Thionyl chloride batteries</td>
<td>.3 – 3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>Acetone</td>
<td>270</td>
<td>24 Hr SMAC</td>
<td>90 – 810</td>
<td>20 - 600</td>
</tr>
<tr>
<td></td>
<td>Dichloromethane</td>
<td>10</td>
<td>24 Hr SMAC</td>
<td>3 – 30</td>
<td>200 - 300</td>
</tr>
<tr>
<td></td>
<td>Ethanol</td>
<td>500</td>
<td>Solvent; 24 hr SMAC 2000</td>
<td>166 – 1500</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Freon 218</td>
<td>20</td>
<td>Russian AC coolant</td>
<td>6 – 60</td>
<td>na</td>
</tr>
<tr>
<td></td>
<td>Methanol</td>
<td>10</td>
<td>24 Hr SMAC</td>
<td>3 – 30</td>
<td>10 - 100</td>
</tr>
<tr>
<td></td>
<td>2-Propanol</td>
<td>100</td>
<td>24 Hr SMAC</td>
<td>30 – 300</td>
<td>20 - 40</td>
</tr>
<tr>
<td></td>
<td>Toluene</td>
<td>16</td>
<td>24 Hr SMAC</td>
<td>5 – 50</td>
<td>0.2 - 2</td>
</tr>
<tr>
<td>3</td>
<td>Formaldehyde</td>
<td>TBD</td>
<td>24 Hr SMAC</td>
<td>TBD</td>
<td>0.5 - 1</td>
</tr>
</tbody>
</table>

* SOURCE:  
US Coast Guard, Chemical Hazards Response Information System; [http://www.chrismanual.com/](http://www.chrismanual.com/)