

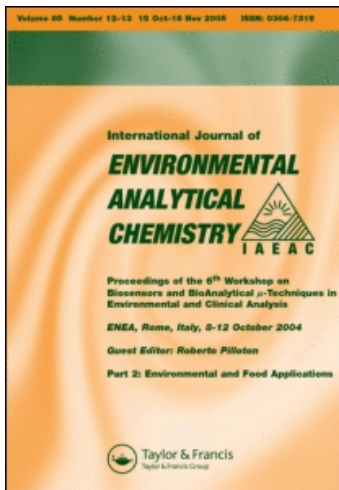
This article was downloaded by:

On: 30 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Environmental Analytical Chemistry

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713640455>

An Application of Pattern Recognition and Infrared Spectroscopy to Water Analysis

Thomas Witelski^a; Phillip Ng^a; Jackie Ying Joseph Lundy^a; John Bove^a

^a The Cooper Union for the Advancement of Science and Art, New York, New York, USA

To cite this Article Witelski, Thomas , Ng, Phillip , Lundy, Jackie Ying Joseph and Bove, John(1991) 'An Application of Pattern Recognition and Infrared Spectroscopy to Water Analysis', *International Journal of Environmental Analytical Chemistry*, 44: 2, 127 – 136

To link to this Article: DOI: 10.1080/03067319108027542

URL: <http://dx.doi.org/10.1080/03067319108027542>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

AN APPLICATION OF PATTERN RECOGNITION AND INFRARED SPECTROSCOPY TO WATER ANALYSIS

THOMAS WITELSKI, PHILLIP NG, JACKIE YING JOSEPH LUNDY
and JOHN BOVE*

*The Cooper Union for the Advancement of Science and Art, 41 Cooper Square, New
York, New York 1003, USA*

(Received 22 August, 1990; in final form 5 December 1990)

A mathematical methodology is presented that shows potential for the interpretation of infrared spectra through a technique of pattern recognition. A number of water samples and simple alkanes were studied to examine the sensitivity and discriminating qualities of the methodology. With 2500 comparisons each of tap water versus tap water or distilled water versus tap water, the technique performed well in selecting the targeted reagent. The same is true for the comparisons of *n*-pentane, as the window compound, versus *n*-heptane, *n*-octane, *n*-nonane and *n*-decane. Comparisons of these *n*-alkanes to *n*-pentane gave fitting tolerances of 15.3, 21.2, 20.9 and 24.7%, respectively. When *n*-pentane was compared to itself, the tolerance fit was 2.7%, showing the ease of discrimination.

These results suggest that this computer-aided phase space transformation method is sensitive, offers good analytical precision, and is capable of detecting small differences in the infrared spectra of compounds and mixture studied. Preliminary data also suggest that the method has potential for monitoring the quality of water.

KEY WORDS: Infrared spectra, pattern recognition, water analysis.

INTRODUCTION

Chemists traditionally make use of the infrared (IR) portion of the electromagnetic spectrum between 4000 and 400 cm^{-1} (2.5–25 microns) to assist in the identification of organic compounds. Even simple compounds generate complex spectra in this energy region. Such spectra are graphically represented by plots of per cent transmittance versus wavelength. Generally, the compounds analyzed must be reasonably pure, and the spectra produced sufficiently resolved and of reasonable intensity. It is also assumed that the spectrophotometer is calibrated and generates stable, replicable measurements. The spectra are often unique, allowing the experimenter to make peak-to-peak comparisons with copies of authentic spectra.

With the introduction of analytical instruments interfaced to computers, new

*To whom correspondence should be addressed.

and more sophisticated techniques have been explored to collect and analyze data. A number of methodologies using pattern recognition have been suggested.¹⁻⁵ It has been pointed out that there is a paucity of examples in the literature dealing with the application of pattern recognition to the field of chemistry.⁶ One such reported example is a deterministic-type approach to spectroscopic data.⁷ In another example, a probabilistic approach of pattern recognition is applied to soil data.⁸ Other examples should also be noted.⁹⁻¹²

While no preset procedure has been established for this process, work is now in progress in these laboratories to explore the feasibility and general usefulness of IR interpretation through techniques of pattern-recognition transformations as an analytic methodology for environmental and organic reagents. This mathematical approach was invented by one of us and later used successfully in the detection of some preclinical coronary artery diseases, and also for the detection of some chronic diseases suggested from EKG tracings.¹³⁻¹⁵ In an earlier report,¹⁶ we tested our working model with data collected from some 100 organic compounds, including nine functional groups (alkanes, cycloalkanes, ketones, aldehydes, alcohols, ethers, carboxylic acids, esters and amines). These transformations were applied to a standard library of FT-IR spectra digitized with a Summa Graphics Model MM1103.¹⁷ Each of the tested functional groups resulted in successfully unique transformation morphologies.

We now wish to report an improved technique, which makes use of phase-space pattern-recognition transformation of spectra generated from an analytical system that consists of a Fourier Transform Infrared spectrophotometer with an internal reflectance cell and interfaced to a personal computer.

The efforts of this new work will concentrate on the ability of this transformation technique to discriminate between some simple *n*-alkanes that are normally difficult, if not impossible, to identify from their IR spectra by eye inspection and the detection of samples of distilled water and tap water. Each of these tests was selected because of its inherently difficult nature. Statistical survival would indicate significant potential for this phase-space transformation methodology.

Normally, the quantitative analysis of water solutions using IR spectroscopy has been avoided since water is such a strong absorber in the middle IR region. In addition, such analysis calls for IR cells of short pathlengths, and these cells are difficult to maintain. This, therefore, introduces a multitude of problems in attempting to provide the necessary experimental precision.

In recent years a new technique for providing reproducible IR spectra has become commercially available for use by Spectra-Tech. The Q-CIRCLE cell avoids the problems of window wear, temperature, and pressure inherently present in the conventional fixed-pathlength cells. It also makes possible the routine analysis of water and its mixtures. The main principle of the internal reflectance technique involves the passage of IR radiation through an IR transmittance crystal of high refractive index; the radiation reflects off the crystal walls once or several times. In this way, an evanescent wave is introduced into the surrounding medium, producing a spectrum of the liquid in contact with the crystal. A more detailed description of the Q-CIRCLE cell appears in the literature.¹⁸⁻²⁰

EXPERIMENTAL

Instrumentation

Spectra were collected using a Perkin-Elmer FT-IR 1600 (2 cm^{-1} resolution) with a Spectra-Tech Q-CIRCLE internal reflectance cell. Samples were introduced into the cell via a funnel, which is part of the cell. An automated pump, which is also part of the Q-CIRCLE, draws the sample into the cell, and it is later used to empty it. In addition, a control is provided to maintain the cell's temperature. Analyses for this report were performed at 26°C . Between analyses the cell was emptied, flushed several times with acetone, and dried. An RS232 cable interfaced the FT-IR to the AT&T 6300 (40 MB hard disk). The entire analytical operation is controlled from the computer keyboard.

Chemical Reagents

All chemical reagents were purchased from Aldrich Chemical Co. and were of HPLC solvent grade (99.5% purity).

Analytical Methodology

Fifty analytical runs were conducted for *n*-pentane, *n*-heptane, *n*-octane, *n*-nonane, and *n*-decane. Each run was scanned once (4 secs) through an IR energy range of 4400 cm^{-1} – 700 cm^{-1} . Fifty runs were additionally conducted for distilled water (Aldrich, HPLC quality), and fifty runs were conducted for tap water (obtained from the fifth floor drinking fountain in The Cooper Union Albert Nerken School of Engineering). These water samples were also scanned once from 4400 cm^{-1} to 700 cm^{-1} .

Domain Transformation

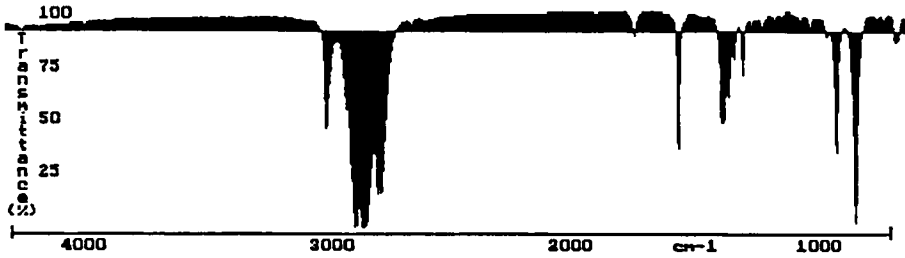
An IR spectrum generated by a spectrophotometer (represented by a plot of transmittance vs. wavelength) is first normalized to an absorbance range of 0 to 1.5 (cf. Eqs. 1 and 2 below). In addition, the range of wavelengths is normalized from 0 to 2π (cf. Eq. 3). These steps are summarized in the following equations:

$$T(\mu) \rightarrow A(\mu), \text{ where } A(\mu) = -\log_{10} T(\mu) \quad (1)$$

$$(1.5/A_{\max}) \cdot A(\mu) \rightarrow A'(\mu) \quad (2)$$

$$2\pi \cdot \left(\frac{\mu - 2.5}{15.0 - 2.5} \right) \rightarrow \mu' \quad (3)$$

with $T(\mu)$, per cent transmittance as a function of wave number; $A(\mu)$, absorbance



$$A'_{\text{avg}} = \frac{1}{2\pi} \int_0^{2\pi} A'(\mu') d\mu'$$

$$I(\mu') = \int_0^{\mu'} |A'(\lambda) - A'_{\text{avg}}| d\lambda$$

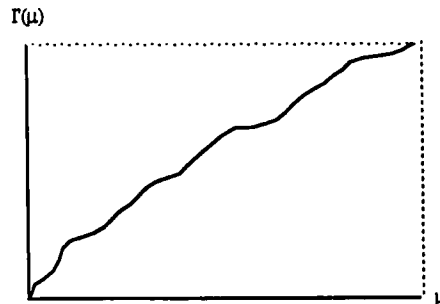


Figure 1 Infrared spectrum with integration function.

as a function of wave number; $A'(\mu)$, normalized absorbance; μ , wavelength in microns; λ , a wavelength variable in microns.

From the resulting normalized spectrum (Eq. 2), a new function, $I(\mu')$, is generated by integrating the absolute difference between the IR spectra and the average absorbance of the spectra:

$$A'_{\text{avg}} = \frac{1}{2\pi} \int_0^{2\pi} A'(\mu') d\mu', \text{ average absorbance} \quad (4)$$

$$I(\mu') = \int_0^{\mu'} |A'(\lambda) - A'_{\text{avg}}| d\lambda, \text{ integral function of absorbance.} \quad (5)$$

The average was empirically established as the one giving the proposed methodology the most sensitivity. This integral of the spectrum is the area bounded by the spectrum and the line of average absorbance (Figure 1). This resulting non-decreasing function, $I(\mu')$, is normalized between 0 and 2π :

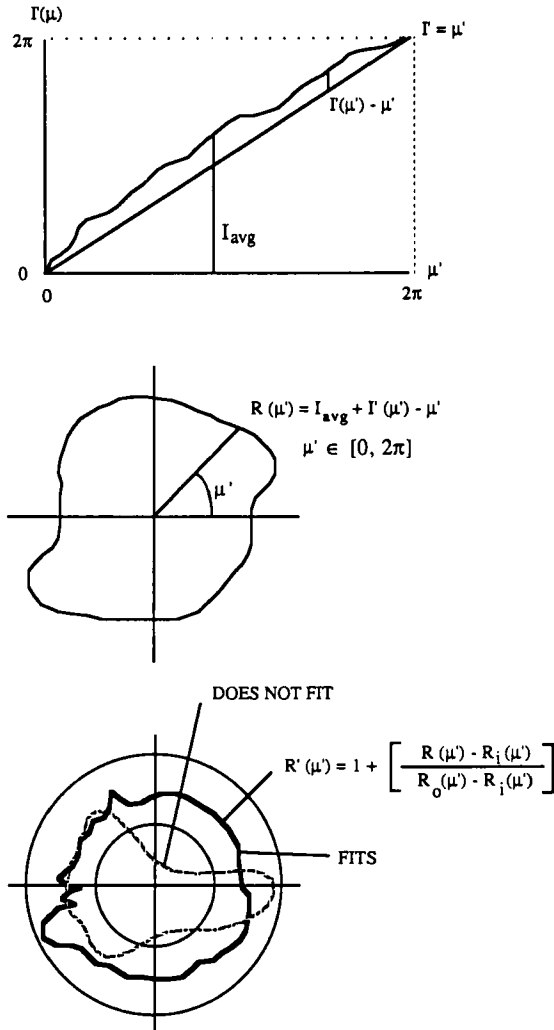


Figure 2 Steps leading to the polar circular transformations.

$$2\pi \cdot \frac{I(\mu')}{I(2\pi)} \rightarrow I'(\mu'). \tag{6}$$

The difference between the straight line, $I' = \mu'$, and the function $I'(\mu')$ is then plotted in polar coordinates to yield a closed continuous curve, $R(\mu')$:

$$I_{avg} = \frac{1}{2\pi} \int_0^{2\pi} I'(\mu') d\mu', \text{ average value of integral function} \tag{7}$$

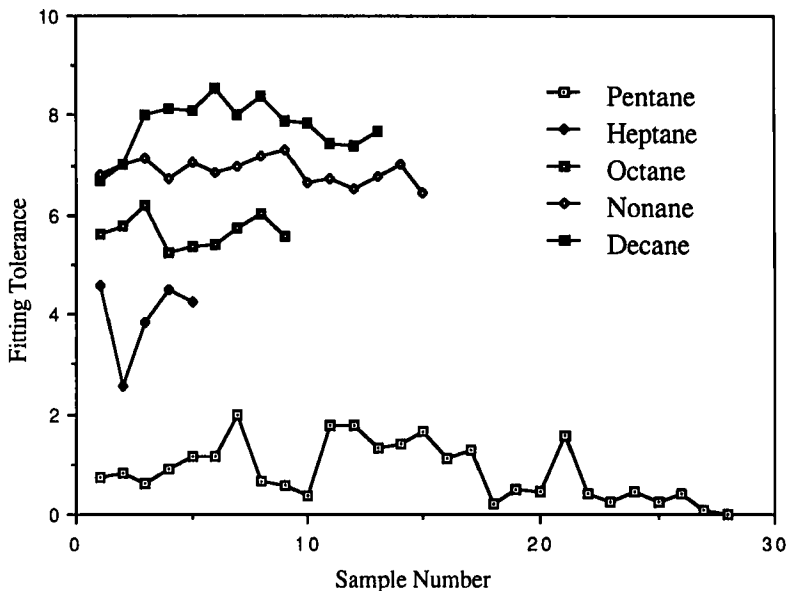


Figure 3 Fitting tolerances of several *n*-Alkanes with respect to a standard *n*-pentane.

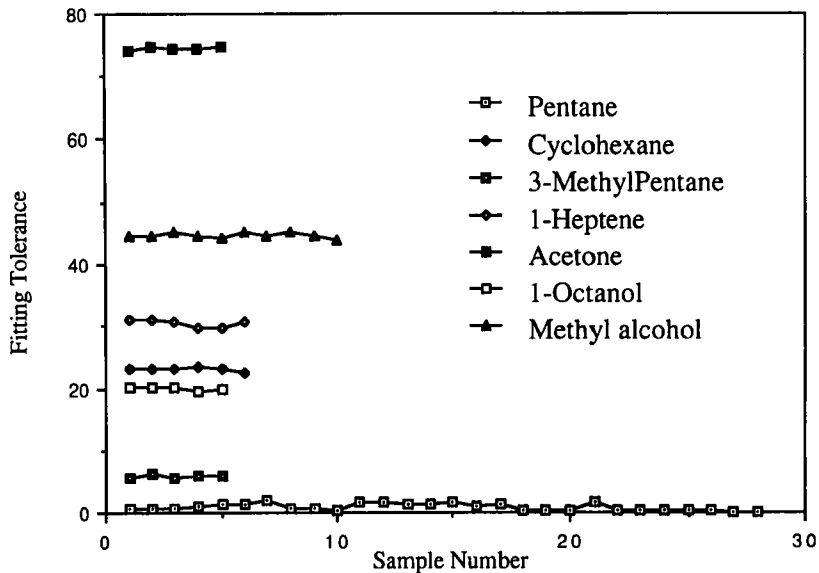


Figure 4 Fitting tolerances of several organic reagents with respect to a standard *n*-pentane.

$$R(\mu') = (I'(\mu') - \mu') + I_{\text{avg}}, \text{ radius function of } I'(\mu'). \quad (8)$$

Finally, the resulting curve is topologically fitted into a circular window that can be used to accept related compounds and reject the others (Figure 2).

where $R_i(\mu')$ is the inner boundary of the window, and $R_o(\mu')$ is the outer boundary of the window.

The condition for pattern recognition at wavelength μ' with tolerance ρ is given by:

$$R_i(\mu') = 1 + \left(\frac{R(\mu') - R_i(\mu')}{R_o(\mu') - R_i(\mu')} \right). \quad (9)$$

The overall tolerance for which one spectrum matches another is the minimum ρ , called the fitting tolerance, such that Eq. (10) is true for all wavelengths. For each new compound or groups of compounds a phase-space transformation window $R_i(\mu')$ and $R_o(\mu')$ can be generated:

$$(1 - \rho) \cdot R_i(\mu') \leq R(\mu') \leq (1 + \rho) \cdot R_o(\mu') \quad (10)$$

RESULTS AND DISCUSSION

The reproducibility of the analytical procedure was tested using one sample of pentane as a reference compound. When 28 individual samples were analyzed over several weeks, the *n*-pentane fitted the transformation reference compound with an average fitting tolerance to 0.88 (with a variance of $\pm 0.30\%$). These results are summarized in Figure 3.

A comparison of the reference window of *n*-pentane to four other *n*-alkanes (*n*-heptane, *n*-octane, *n*-nonane, and *n*-decane) using an integration factor of 1.0 resulted in fitting tolerances of 4.3, 5.7, 7.3, and 7.8%, respectively. These differences in minimum tolerance are large enough so that there is no difficulty identifying *n*-pentane from the other four *n*-alkanes. These data are summarized in Figure 3.

Further comparisons of three other alkanes (cyclohexane, 2,2,4-trimethylpentane and 3-methylpentane) gave fit differences of 23.5, 12.2, and 5.9%, respectively—again indicating easy discrimination for the alkane family of compounds. Additional comparisons of the reference *n*-pentane to 1-heptane, acetone, 1-octanol, and methyl alcohol gave fits of 30.5, 74.4, 20.0, and 44.7%, respectively, indicating easy discrimination. These values are summarized in Figure 4.

The reproducibility of data and the discrimination quality of the methodology was further examined when another selected group of *n*-alkanes (pentane, heptane, octane, nonane, and decane) was compared. Additional comparisons (50 runs)

Table 1 Circular transformation comparisons of a selected group of *n*-alkanes (*n*-pentane and the target reagent)

<i>n</i> -Alkane	Per cent fit	Standard deviation
Pentane	2.7	1.7
Heptane	15.3	2.2
Octane	21.2	1.4
Nonane	20.9	0.84
Decane	24.7	0.80

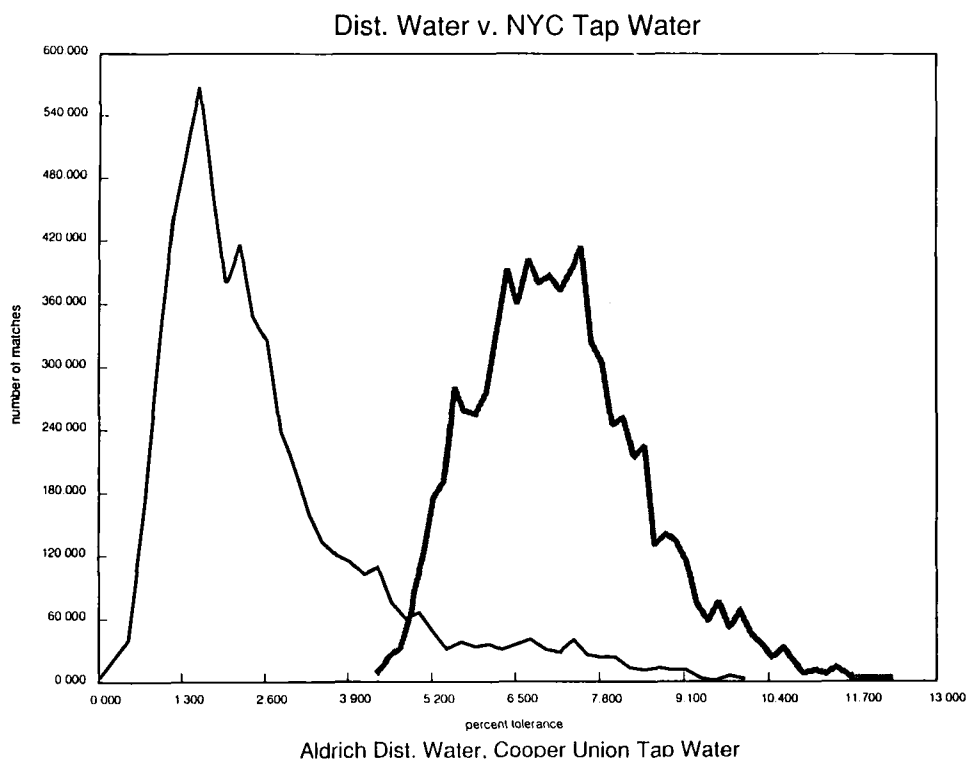


Figure 5 Comparison of distilled water to NYC tap water.

were also conducted using distilled water and tap water, using an integration factor of 0.333.

The resulting data were subjected to statistical hypotheses testing.^{21,22} When comparisons were conducted of the *n*-alkanes, there was no difficulty distinguishing between these alkanes. It should be noted that a total of 2500 comparisons were conducted. These data are summarized in Table 1. 2,500 statistical comparisons of distilled water with New York City tap water and New York City tap water with New York City tap water gave unambiguous identifications.^{23,24} The

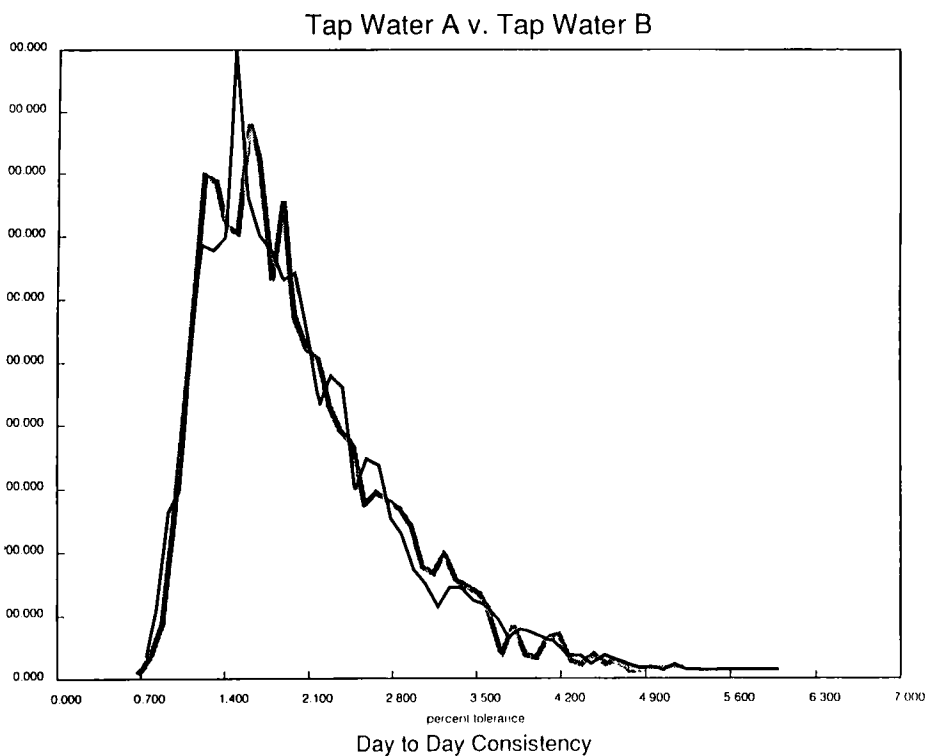


Figure 6 Comparison of New York City tap water to itself.

phase-space transformation technique resulted in only small error I/error II confusion. The vast majority of the samples were clearly segregated into their proper populations (distilled water on tap water). These data are summarized in Figs 5 and 6, respectively.

CONCLUSIONS

Preliminary results indicate that the new computer-controlled phase-space transformation methodology is both sensitive and precise, and is capable of recognizing the small differences in FT-IR spectra. For example, it was easy to distinguish between some of the simple alkanes when the transformation method was employed, even though the alkane spectra could not be identified by visual inspection. The same was found to be true when the original IR spectra were employed to distinguish distilled water from tap water—both IR spectra appeared the same to the naked eye. The opposite proved true, however, when tap water and distilled water were compared via the transformation technique—there was a clear difference.

Future work will center on the use and evaluation of this new analytical

technique for examining environmental samples of water. This work will include the analysis of water samples for phenol and chlorinated pollutants.

Acknowledgements

We gratefully acknowledge our sponsor, The International Joint Commission for U.S. and Canada, under contract number 1761-700429. Our special thanks go to Dr Joel Fisher, Project Officer, for the smooth coordination of our efforts with the Commission.

References

1. A. J. Stuper, W. E. Brugger and P. C. Jurs, *Computer Assisted Studies of Chemical Structure and Biological Function* (Wiley-Interscience, New York, 1979).
2. K. Varmuza, in: *Computer Applications in Chemistry* (S. R. Heller and R. Potenzzone Jr., eds.) Elsevier, Amsterdam, 1983; B. R. Kowalski and S. Wold, in: *Handbook of Statistics*, Vol. 2 (P. R. Krishnaiah and L. N. Kanal, eds.) North-Holland, Amsterdam, 1982; L. Kryger, *Talanta* **28**, 871 (1981).
3. A. M. Harper, *Pyrolysis and G. C. in Polymer Analysis*. S. A. Liebman and E. J. Levy, eds. Dekker, New York, 1985.
4. S. Wold, *et al.*, *Chemometrics, Mathematics and Statistics in Chemistry*. B.R. Kowalski, ed. Reidel, New York, 1985.
5. B. R. Kowalski, *Anal. Chem.* **52**, 112R (1980); M. F. Delaney, *ibid.* **56**, 261R (1984).
6. D. Coomans and I. Broeckaert, *Potential Pattern Recognition in Chemical and Medical Decision Making* (J. Wiley, New York, 1986) **8**, 225.
7. L. N. Drozdov-Tichomirov, *Opt. Spectrosc.* **27**, 77 (1968).
8. J. D. F. Habbema and J. Hermans, *Technometrics* **19**, 487 (1977).
9. Kowalski *et al.*, *Anal. Chem.* **41**, 1945, (1969).
10. Woodruff *et al.*, *Anal. Chem.* **46**, 2150 (1974).
11. Bink *et al.*, *Anal. Chim. Acta* **150**, 53 (1983).
12. Domokos *et al.*, *Anal. Chim. Acta* **154**, 181 (1983).
13. J. R. Lundy Patent Numbers 4,570,225 and 4,812,976.
14. J. R. Lundy and Kazlauskas, "Gait Evaluation Via a Non-Linear Graphic Transform Synthesis," 37th Annual Conference of Engineering in Medicine and Biology, September 17-19, 1984, Los Angeles, California.
15. J. R. Lundy, W. M. Walsh, L. Rondinella and S. Lee, "Detection of Chronic Disease from EKG Tracings," 12th Annual Northeast Bioengineering Conference, March 1986, Yale University, New Haven, CT.
16. J. Ying, S. Lee, L. Rondinella, J. R. Lundy and J. L. Bové, "The Application of Computer-Aided Diagnostic Enhancement on Infrared Spectra," American Chemical Society National Meeting, Division of Chemical Education, New York City, April 1986.
17. *The Aldrich Library of Infrared Spectra*, Edition III, C.J. Pouchert, 1981.
18. A. Rein and P. Wilks, *P. Am. Lab* **14**, **10**, 152 (1982).
19. Barnes Analytical Division, Spectra-Tech, Inc., 652 Glenbrook Road, Stamford, CT.
20. E. G. Bartick and R. Messerschmidt, *Am. Lab.* **16**, No. 11 (1984).
21. A. D. Walen, *Detection of Signals in Noise* (Academic Press, 1971), pp. 126-131.
22. Franzen, *Chromatographia* **7**, 518 (1974).