Remote Sensing Information Classification
Presentation for HELIX-Israel Kick-Off Workshop April 29, 2008

Doug Rickman
NASA/Marshall Space Flight Center/Earth Science Office
The remote sensing data offers a uniform measurement over a large area.

Remote sensing provides direct measurement of various geophysical properties, such as reflectance, emission and absorption of electromagnetic energy.

These geophysical properties are *partially* controlled by things of interest to epidemiologists, such as vegetation.

The actual remote sensing data can be used directly or ....

- It can be classified.
- It can be integrated into models.
- It can be converted into products.
The sensor measures multiple wavelengths at each pixel. All energy within a pixel for a wavelength band pass is integrated to a single integer in the range 0 – 255 ($2^8$). Each band is independent. The Thematic Mapper (TM) has 7 bands. For the TM a pixel is measured in 7 dimensions with a precision of $2^8$ bits. In theory each pixel could have any one of $(2^8)^7$ values.

This is a very, very big number.

In practice the bands are correlated and all the available dynamic range is not used. Practically, each pixel has only one of ~$(2^7)^4$ possible values.

This is merely a very big number.
Classification is a way to reduce the dimensionality and precision to something a human can understand.

Classification changes SCALAR data into NOMINAL data!

The names used come from a FEATURE SPACE. The names and the feature space are abstractions!

Conversion from scalar to nominal loses information and introduces error.
The "I hope I am right" option

How strong is the relationship between the nominal designation (the class) and some "objective" standard?

Classification creates a statistical connection between scalar data and a feature space

Start with the scalar geophysical measurement

YES classified

NO

Classification is NOT required

YES supervised

NO

B

The "I hope I am right" option

YES assess accuracy

NO

B

YES relationship

NO

A
Sources of Classification Error

Most classifications have average errors in the range of 25 – 40%.

Example:
"Forcing a square peg through a round hole."
Only two classes are permitted, plain (low number) or hashed (high number). So pixel (2,3) is what?

Example:
"Variables don't and constants aren't."
The measurements always have "noise". Note the values in pixels (1,1), (1,2) and (2,1).

Note: there is a white rectangle used to clip the diagonal shading.
This can cause some fascinating and entertaining problems.

Example:

The raw satellite data are INTEGER.

Most logical algorithms assume REAL domain input.

This can cause some fascinating and entertaining problems.
Binning ALWAYS creates artifacts

What are the relationships between points C, D and E?

If the data are binned to the values $i_2$ and $i_3$, the true relationship are obscured AND spurious information has been added.

Actual function
In a review of publications applying remote sensing to epidemiology
- most were found to use classification
- none gave any information about the accuracy of the classification.

Therefore, their results
- can only reveal how their health data related to their subset,
- they don't really know what that subset is,
- nor do they know if the subset can be reproduced!

It is strongly recommended that epidemiological studies utilize the full information content of the remote sensing material.
- This means using the full dimensionality or some statistically defensible expression of the total or a derived product.
- The computational burden, which 25 years ago was huge, is now easily handled by ordinary desktop systems.
Contacts & Acknowledgement

Doug Rickman
Telephone - 256-961-7889 (United States)
Email - Douglas.LRickman@nasa.gov
Address - Earth Science Office / VP61
NSSTC/MSFC/NASA
320 Sparkman Drive
Huntsville, AL 35805 (USA)

Scientific Team Members at MSFC
Bill Crosson    Dale Quattrochi    Jeff Luvall
Maury Estes    Ashutosh Limaye    Maudood Khan

Illustrative Website
http://www.ghcc.msfc.nasa.gov/ follow Applications: Health and Environment link to
http://weather.msfc.nasa.gov/helix/helix_home1.html

Current Significant Public Health Partners
Leslie McClure, University of Alabama, Birmingham
Judith Qualters, Centers for Disease Control and Prevention
Amanda Niskar, Tel Aviv University
Bill Sprigg, University of Arizona
Stan Morain, University of New Mexico
Science and Mission Systems Office IV
Presentation Added to Database

The accompanying publication metric for the conference proceedings has been added to the database.

Instructions:
1. Print this page.
2. Submit it, along with 3 copies of the abstract, to your MSA.

The following record has been added to the database:

<table>
<thead>
<tr>
<th>Field</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Remote Sensing Information Classification</td>
</tr>
<tr>
<td>Presenters</td>
<td>D. L. Rickman</td>
</tr>
<tr>
<td>Conference Name</td>
<td>HELIX-Israel Kick-Off Workshop</td>
</tr>
<tr>
<td>Location</td>
<td>Tel Aviv University, Tel Aviv, Israel</td>
</tr>
<tr>
<td>Conference Start Date</td>
<td>5/29/2008</td>
</tr>
<tr>
<td>Conference End Date</td>
<td>5/29/2008</td>
</tr>
<tr>
<td>Date Presented</td>
<td></td>
</tr>
<tr>
<td>Conference Proceedings to Follow</td>
<td>Yes</td>
</tr>
<tr>
<td>Organization</td>
<td>VP61</td>
</tr>
</tbody>
</table>

Remote Sensing Information Classification
D. L. Rickman
No