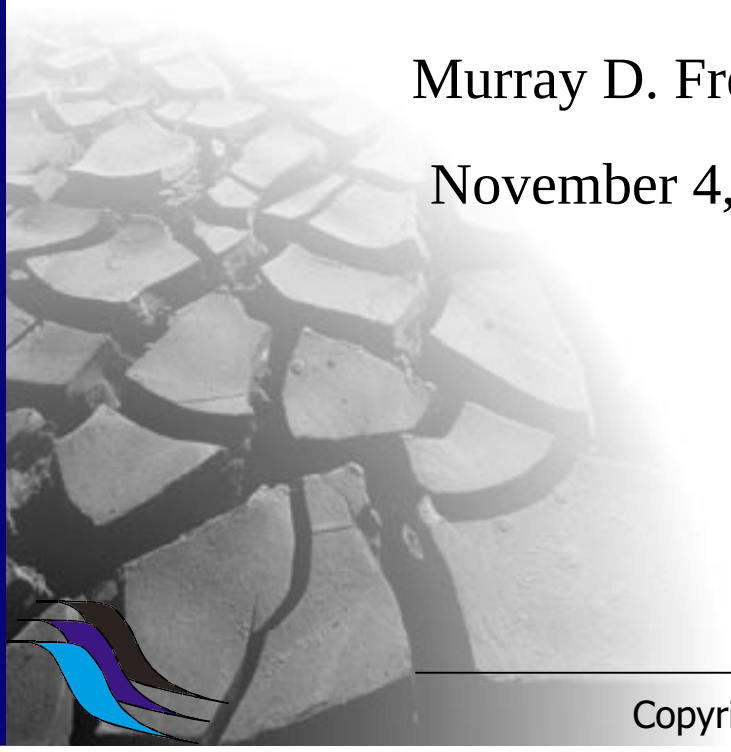
The SoilVision® Data

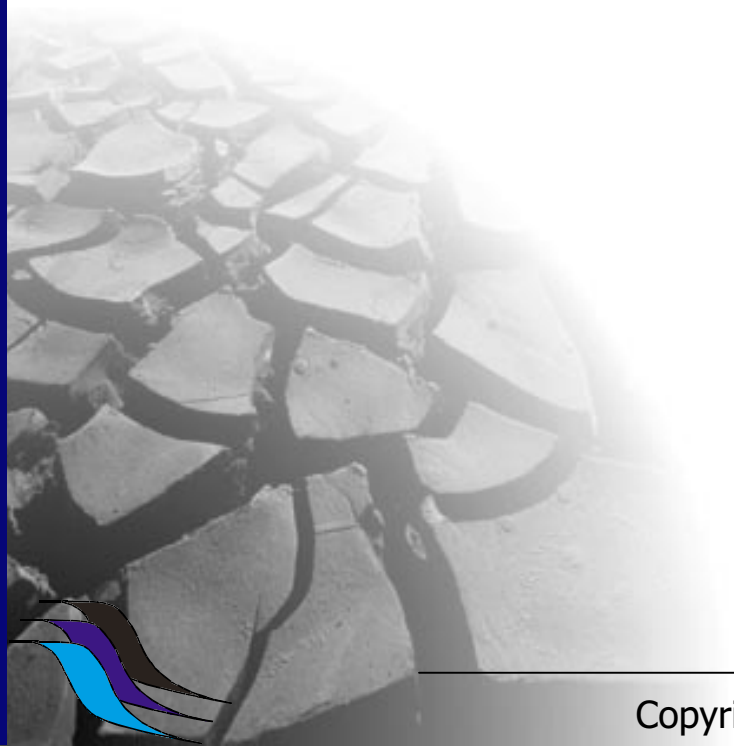
Murray D. Fredlund

November 4, 1999



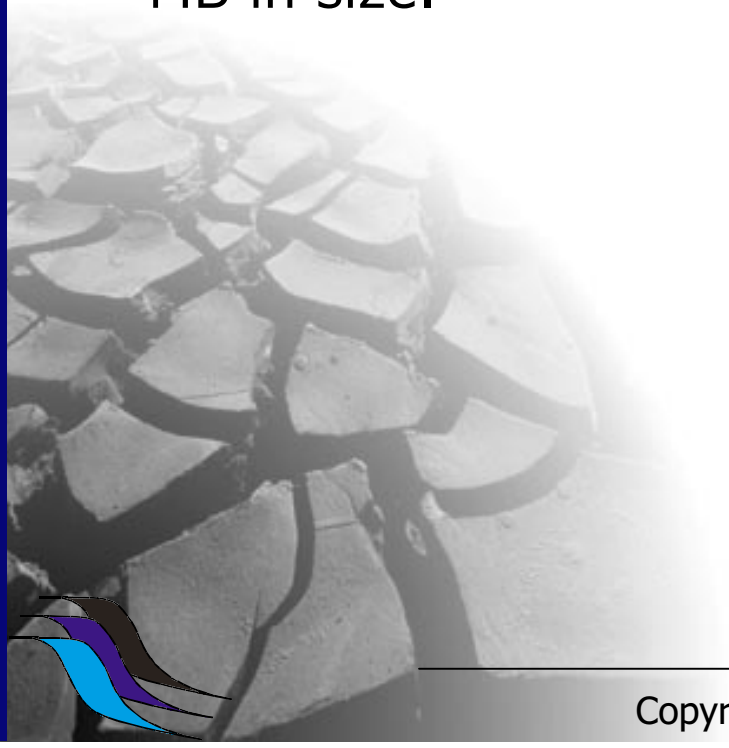
Why is the data important?

- What is the purpose of a large experimental database?
- The data provides us with knowledge.
- Once we have knowledge of the potential limits of soil behavior, this allows us to make educated decisions.
- A database is like a perfect memory.
- Even a perfect memory is limited by experience.
- It is important to realize where the data has come from since it forms the basis of our judgement.



Scope

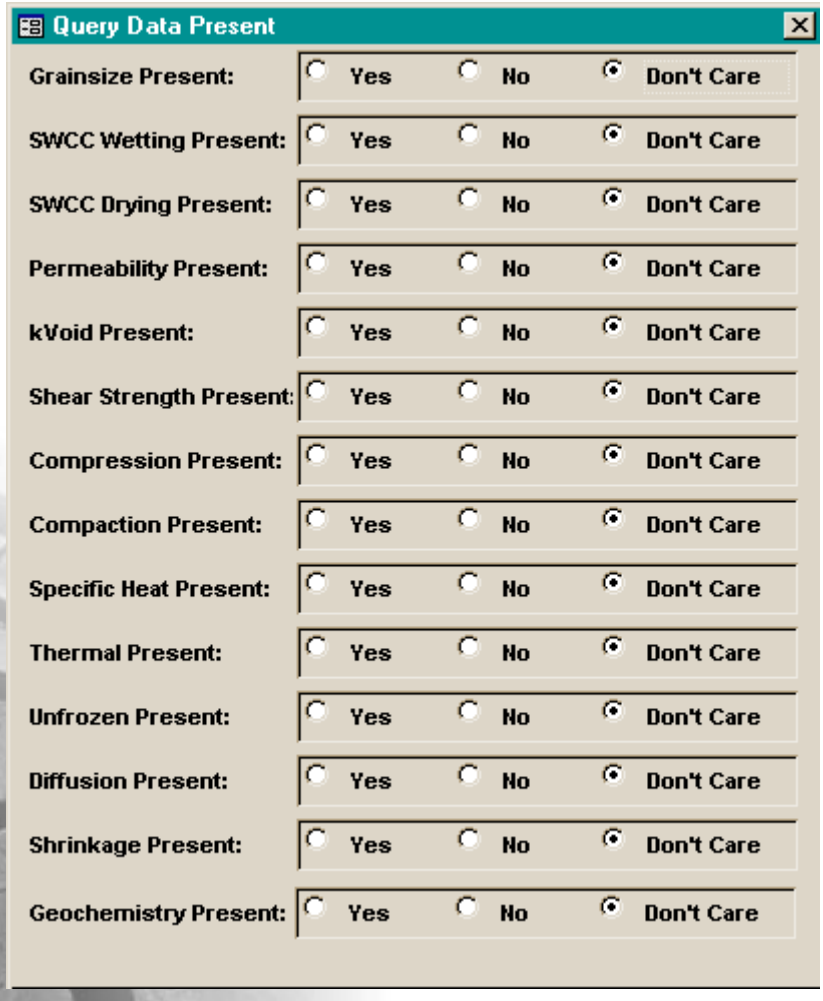
- SoilVision® is the only commercial database application which comes with experimental data.
- The database currently represents the largest collection of unsaturated soils information in the world.
- The data has been collected over a period of five years.
- The data origins are primarily soil science with a smaller percentage of geotechnical engineering data.
- The SoilVision application is roughly 35 MB in size and the data is roughly 17 MB in size.



SoilVision® data

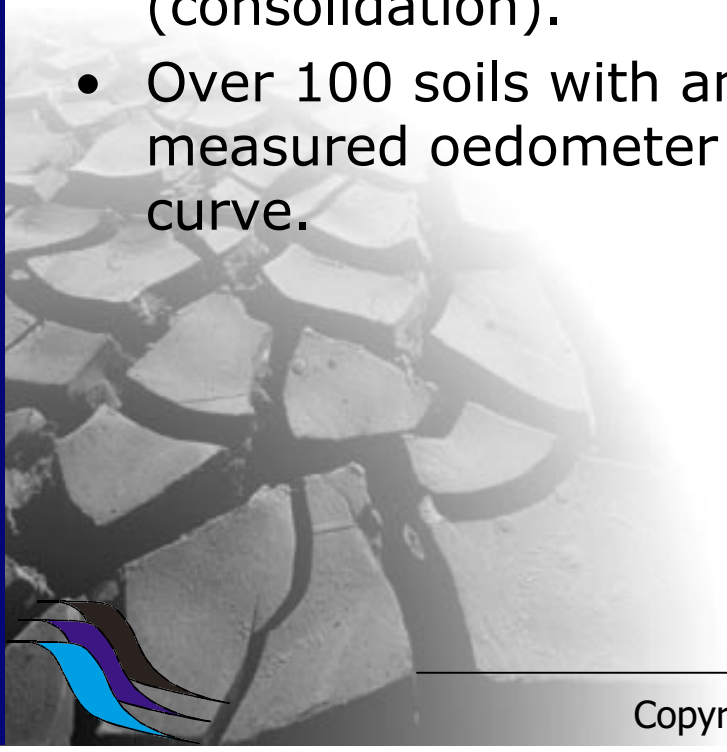
Amount of data

- It is unusual to find a soil for which there exists experimental data on more than several areas.
- SoilVision allows for easy querying for soils information based on measured properties.
- Any combination of soil properties may be selected.



Amount of data

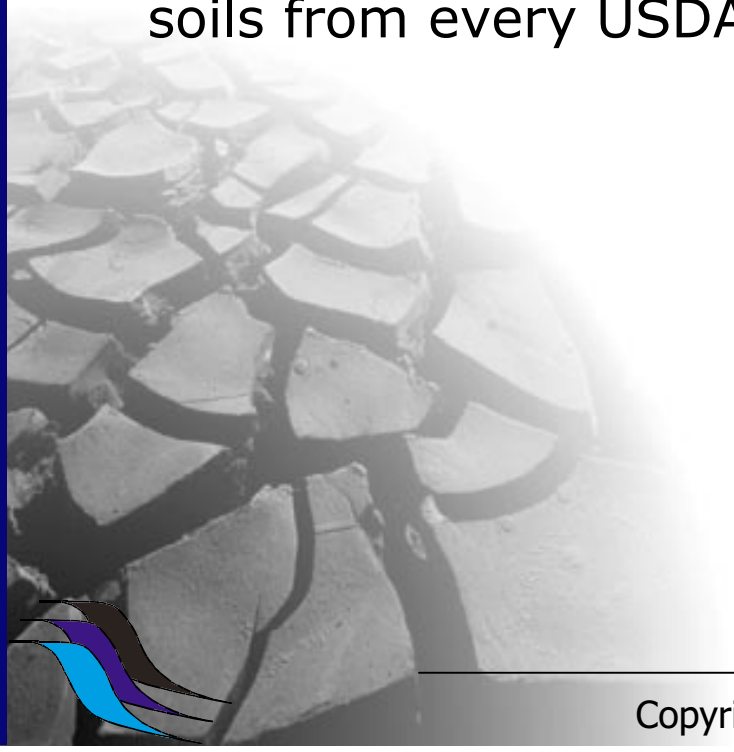
- The database currently contains over 6200 soils.
- Over 90% of the soils include an experimentally measured soil-water characteristic curve.
- Over 900 soils with a grain-size distribution.
- Over 2200 soils with a measured saturated hydraulic conductivity.
- Over 420 soils with an experimentally measured hydraulic conductivity curve.
- Over 30 soils with permeability measured as a function of void ratio (consolidation).
- Over 100 soils with an experimentally measured oedometer compression curve.



SoilVision® data

Amount of data

- Over 25 soils with an experimental defined compaction curve.
- Over 50 soils with an experimentally measured shrinkage curve.
- Over 25 soils with unsaturated-saturated shear box data.
- Over 580 soils with unfrozen water content data.
- Approximately 5 soils with thermal conductivity experimental data.
- The data provides a highly diverse set of data.
- There is significant representation of soils from every USDA soil texture.



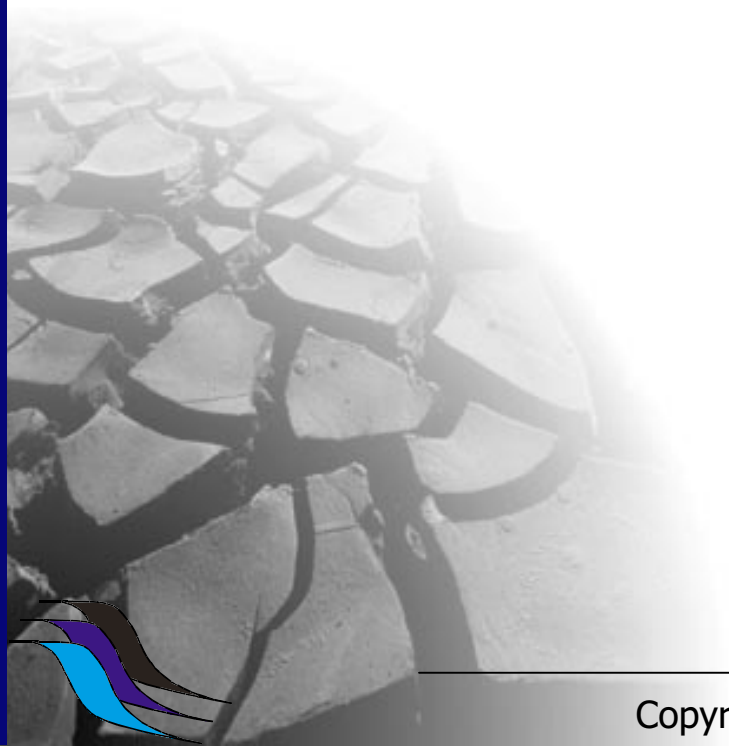
Data sets present

- SoilVision currently includes 13 data sets.
- The data sets come from a variety of university, government, and consulting sources.
- Data sets include data from the following sources:
  - University of Saskatchewan
    - Soils imported from research publications.
  - Rawls data set
    - Personal dataset of Prof. Rawls. Mostly soils from the United States.
  - CECIL data set
    - Soils tested during a governent study of the southern United States.
  - UNSODA
    - A collection of soils from around the world.
  - Mining Tailings
    - Soils collected from research of mining tailings at the University of Saskatchewan.

SoilVision® data

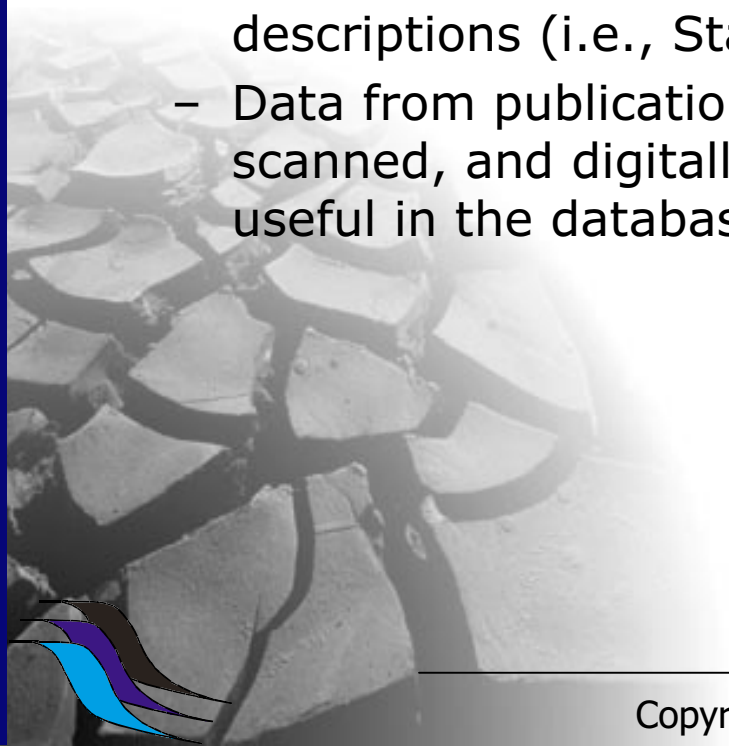
Data sets present

- Newly acquired data sets include data from the following sources:
  - Subgrade Soils
    - "Soil Water Properties of Subgrade Soils", 1980, Donald J. Janssen and Barry J. Dempsey, A report of the investigation of Soil Water Properties of Subgrade Soils, Project IHR-606, Illinois Cooperative Highway and Transportation Research Program
  - Compacted Clay Liners
    - "Estimating Hydrauic Conductivity of Compacted Clay Liners.", 1996, by C.H. Benson, Huaming Zhai, and Xiaodong Wang



Importation process

- Data was originally provided on a variety of different formats.
  - Publications
  - ASCII text files
  - older database files
  - EXCEL spreadsheets
- A significant amount of work was required to import the data.
  - Units must be consistent (i.e., kg/m^3 or pcf)
  - Fields must be consistent (i.e., gravimetric or volumetric water content)
  - Codes were often used in earlier databases and were replaced with text descriptions (i.e., State representation)
  - Data from publications needed to be scanned, and digitally rendered to be useful in the database.



Importation process

- The following standards were implemented in the database:
  - The Metric system of units was adopted as the standard for data representation. The data, however, may be displayed in any system of units.
  - Gravimetric water content was adopted as the standard for representation of the soil-water characteristic curve because it is theoretically correct.
  - Codes were replaced with full descriptions because there does not exist a consistent description of the codes.
  - A method of separating theoretical and experimental data was adopted.
  - After each dataset was imported, random checks were perfomed to ensure data integrity.
  - Datasets were stored at all stages of the conversion process to allow future checking.
  - Volume-mass properties were used to further check for data integrity (i.e., S=113%).

SoilVision® data

Growth?

- Version 1.0 of SoilVision contained over 5000 soils.
- Version 2.0 has increased the number of soils to over 6200.
- There are currently four datasets that require input into the SoilVision database.
- The number of soils has increased as people become aware of the database.
- This trend is expected to increase as the new version provides easier data exchange.