Database And Analysis Tools

S. Kraus[1], S. Beirle[2], C. Frankenberg[2], C. von Friedeburg[2],
M. Grzegorski[2], J. Hollwedel[2], M.F. Khokhar[2], S. Kühl[2],
U. Platt[2], S. Sunghavi[2], W. Wilms-Grabe[2], and T. Wagner[2]

[1]Interdisciplinary Center of Scientific Computing
[2]Institute of Environmental Physics
University of Heidelberg

Overview

- Evaluation Result Storage
  - Current solution of GOME evaluation
  - Database storage: Pros and Cons
- Evaluation Tools
  - DOAS Fit: Old and new implementation

Evaluation Result Storage

- Current solution used in GOME evaluation:
  - Proprietary binary file format
  - One file per orbit and species
  - Pros:
    - Quick data storage
    - Low system requirements
  - Cons:
    - Data access only possible via special tools (e.g. conversion to ASCII data)
    - File system limits (number of files per folder)
    - Modification of the evaluation algorithm requires a new binary file format and new processing tools

Evaluation Result Storage

- New Database solution:
  - SQL (Structured Query Language) database
  - Access via ODBC (Open DataBase Connectivity)
  - Relational database (should be sufficient)
  - Various systems available:
    - Non-commercial: e.g. MySQL, Postgress
    - Commercial: e.g. Oracle, Microsoft SQL-Server

Evaluation Result Storage

- Why a database solution?
  - Flexible extension/modification of result structure
  - Flexible access from various clients via ODBC
    - direct data access from Excel, Origin, etc.
  - Flexible correlation of results
    - Data extraction defined by boolean and relational conditions
      - ‚select SO2 where NO2 > 1e15‘
  - No redundant data storage
    - E.g. pixel coordinates stored once and linked to species result
  - No file system limits
    - Depends on the database system used

Evaluation Result Storage

- What problems do we expect by using databases?
  - Huge amount of data slows down system and data retrieval
  - Internal database limits
    - Maximum table size
    - Maximum database size
    - Mostly non-commercial systems
- We need a fast database server with a lot of memory and disk space and a performant database system.

DOAS Fit

- DOAS fit used in GOME evaluation:
  - Implementation of C. Leue
  - B-Splines used to determine the model function's derivation
- Current limitations of the old implementation:
  - Unable to fix arbitrary model parameter
  - Unable to limit arbitrary model parameter
  - Modification of the model function requires new implementation of the function itself and its derivation
  - Only works with equidistant reference spectra that are based on channel numbers.

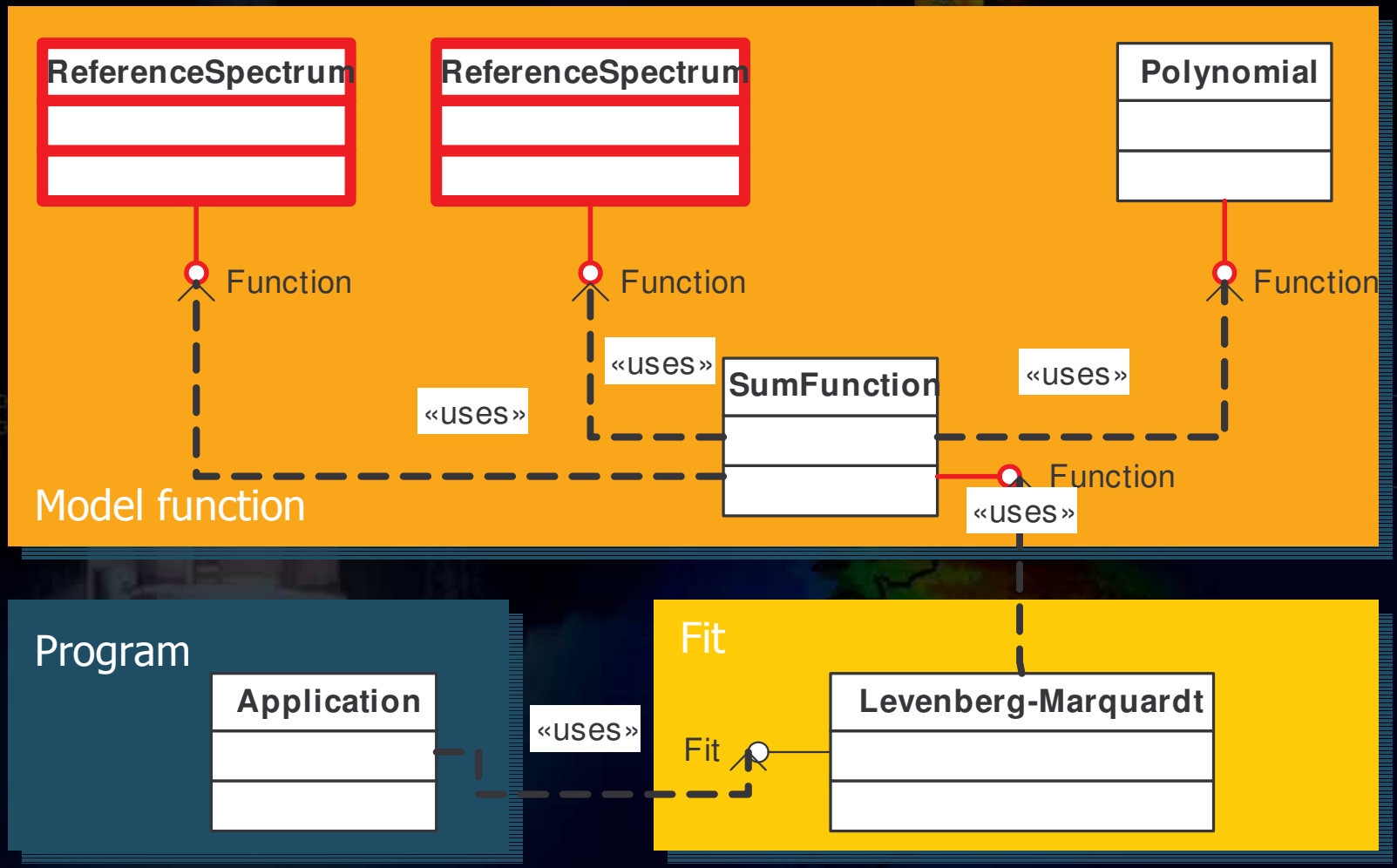
DOAS Fit – New Implementation

- Basic concept of the new implementation:
  - Separation of the fit functionality and the model function
    - Either the model function can easily be exchanged without the need to modify the fit implementation and vice versa
  - Model function is build out of a set of small basic functions
  - Basic functions can provide an analytical derivation or use a numerical method to determine its derivation
  - Linking, fixation and limitation of any model parameter should be possible
  - Non-equidistant reference spectra based on wavelength settings should be used

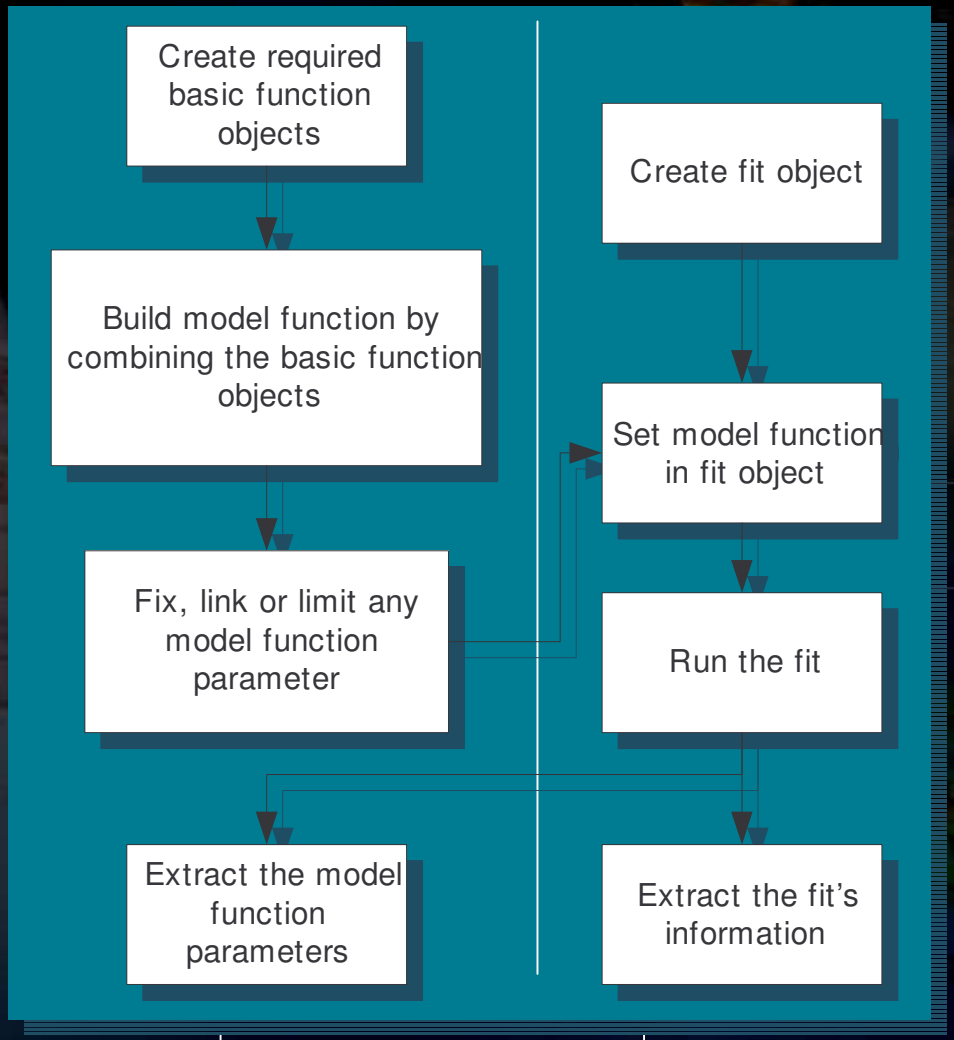
DOAS Fit – New Implementation

- Implementation of the new concept:
  - Object orientated approach used
  - Interfaces export common methods
- Function Interface:
  - Provides methods to get the function‘s data and derivation (numerical or analytical determined)
  - Allows the fixation, linking and limitation of any model parameter
- Fit Interface
  - Uses the function interface to get the model function‘s derivation and parameters

DOAS Fit – New Implementation



DOAS Fit – New Implementation



Summary

- Usage of a database solution to store the evaluation results
  - More flexible storage and retrieval of evaluation results
- Flexible evaluation algorithm to easily modify model function and fit method
  - Object orientated approach