A Data Description Model for Reusing, Sharing and Integrating Geo-Analysis Models

OpenGMS Team

Key Laboratory of VGE (MOE)

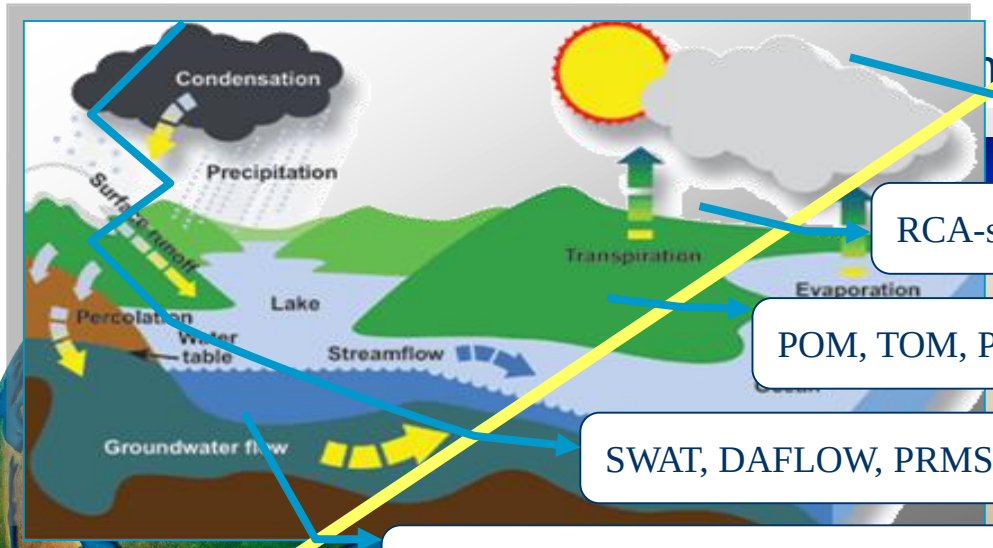
Nanjing Normal University

Outline

- Aims
- Data Description Model
- Architecture Design and Practice
- Conclusions

1. Aims

- Geo-analysis Model



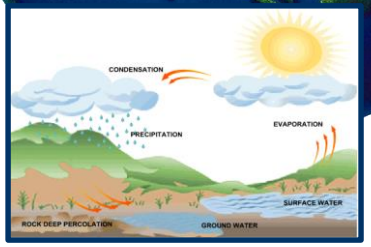
MM5, WRF, OPA, RCO, UMO…

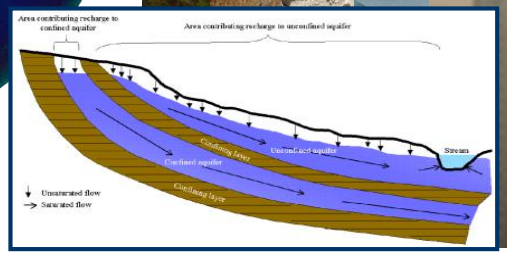
RCA-soil, MOSES, ISBA, LSS, ORCHIDEE…

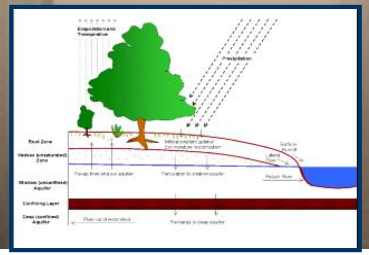
POM, TOM, PPTAR, MOM, FVCOM, ECOM…

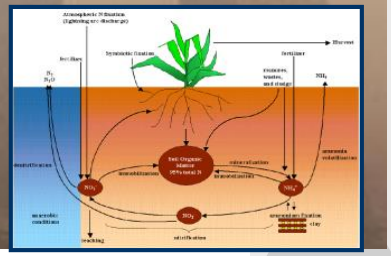
SWAT, DAFLOW, PRMSModFlow, BASINS…

SARABAND, VOLAN, CORIBAND…









Geo-analysis models are widely used to describe geographic phenomena, simulate geo-processes, and depict regular geographic patterns in the environment.

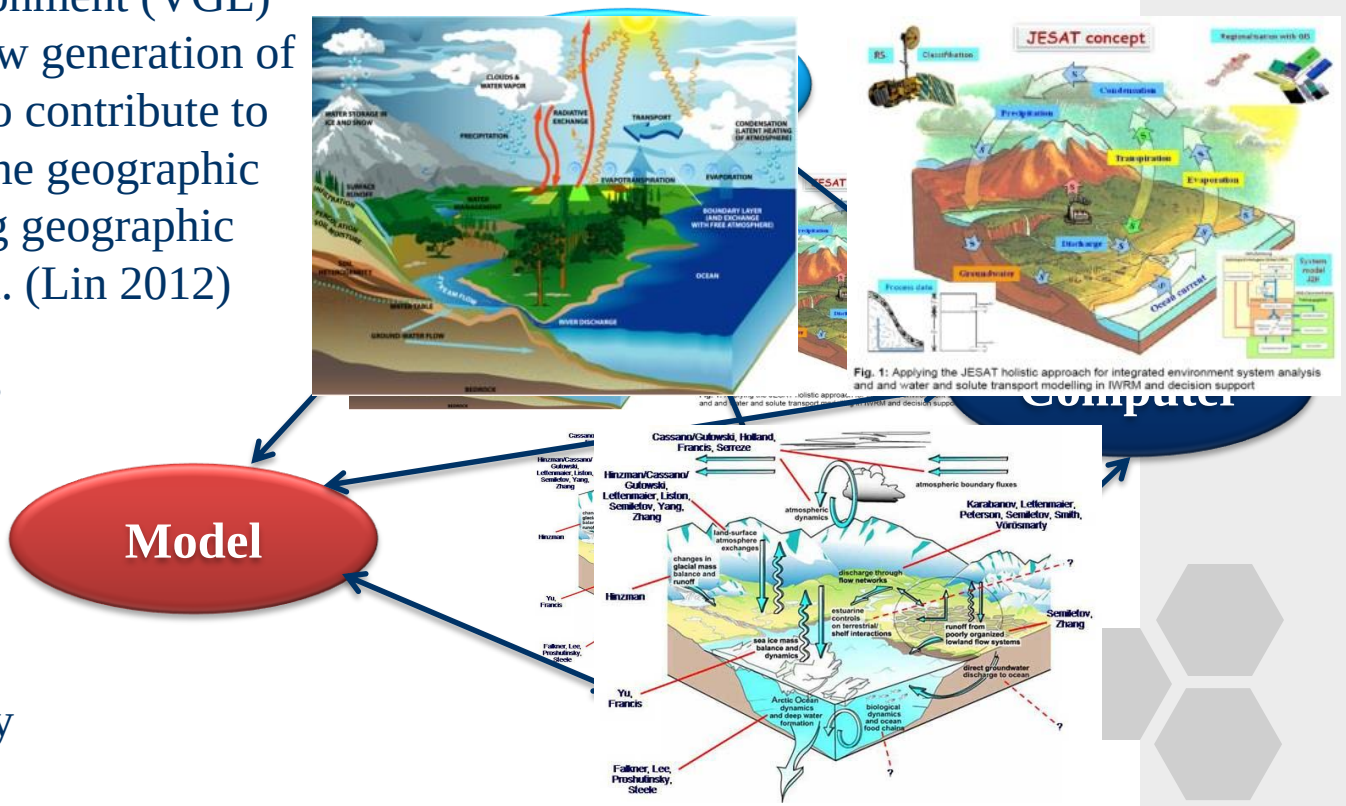
1. Aims

❖ Geo-analysis Model Sharing and Integrating

However, to **Simulate synthetic geo-processes** and **Solve complicated environmental problems**, any single study field is inadequate; the **collaboration** of different fields of study is required.

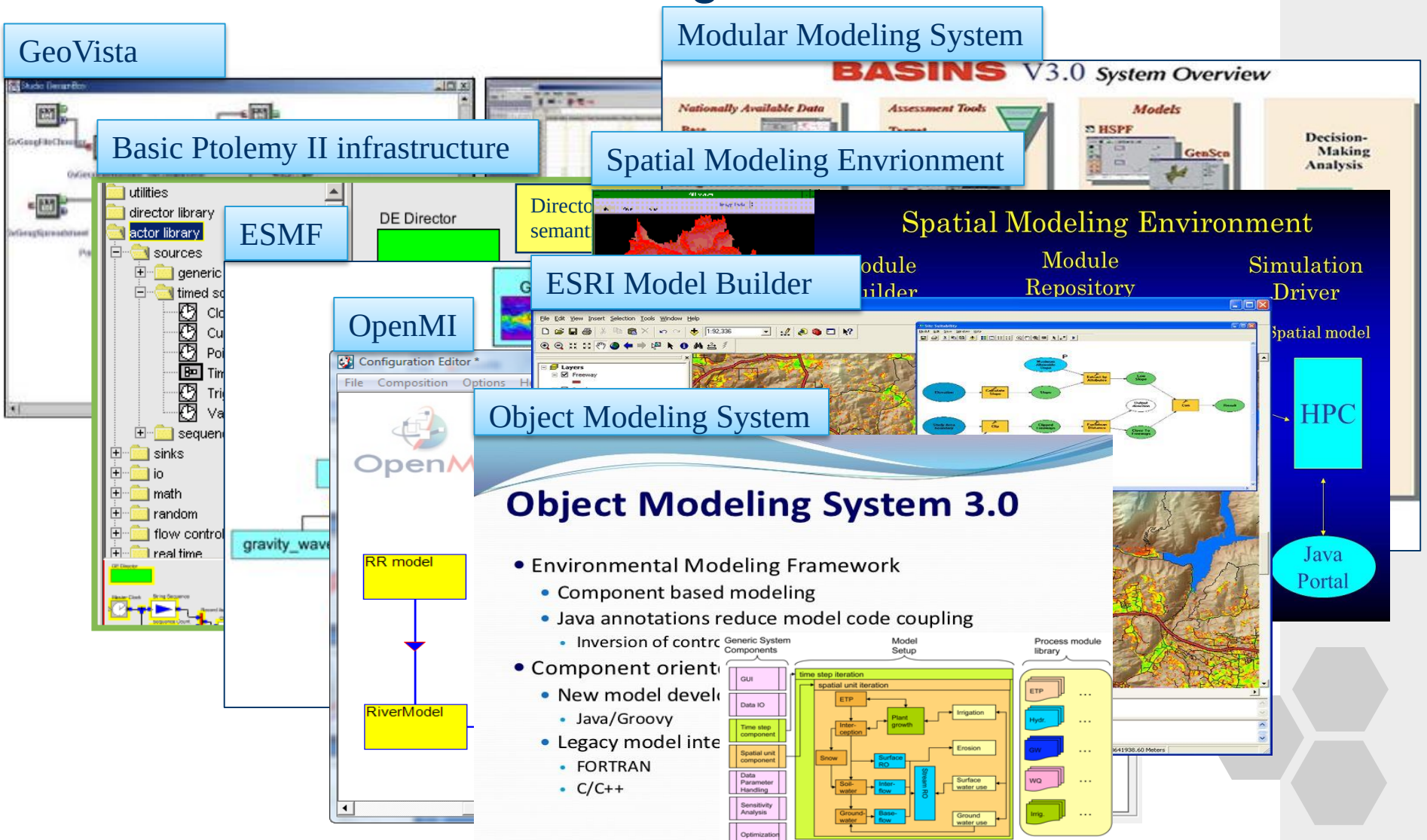
Virtual Geographic Environment (VGE) has been proposed as a new generation of geographic analysis tool to contribute to human understanding of the geographic world and assist in solving geographic problems at a deeper level. (Lin 2012)

Studying, reusing, sharing and integrating geo-analysis models can provide an interactive interface for researchers performing geographical experiment collaboratively in VGE.



1. Aims

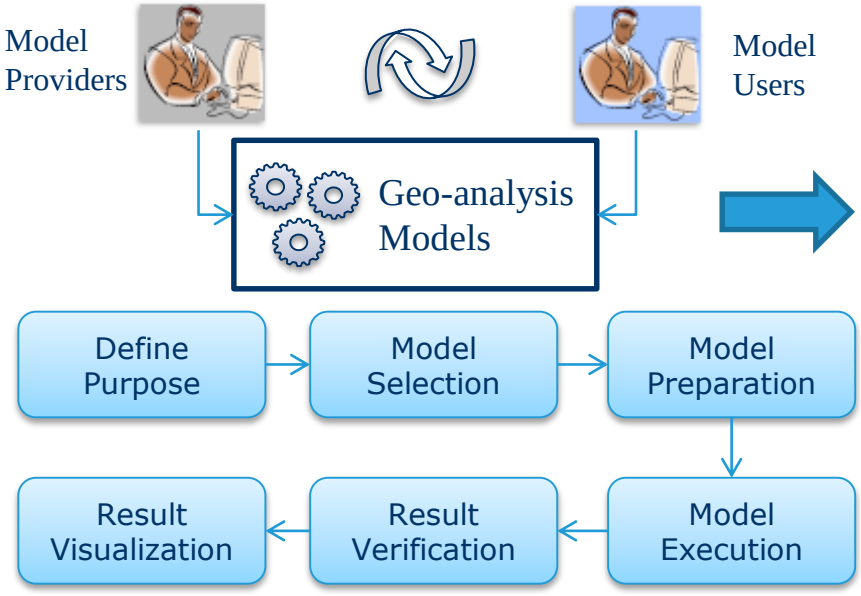
- Research of Model Integration Frameworks

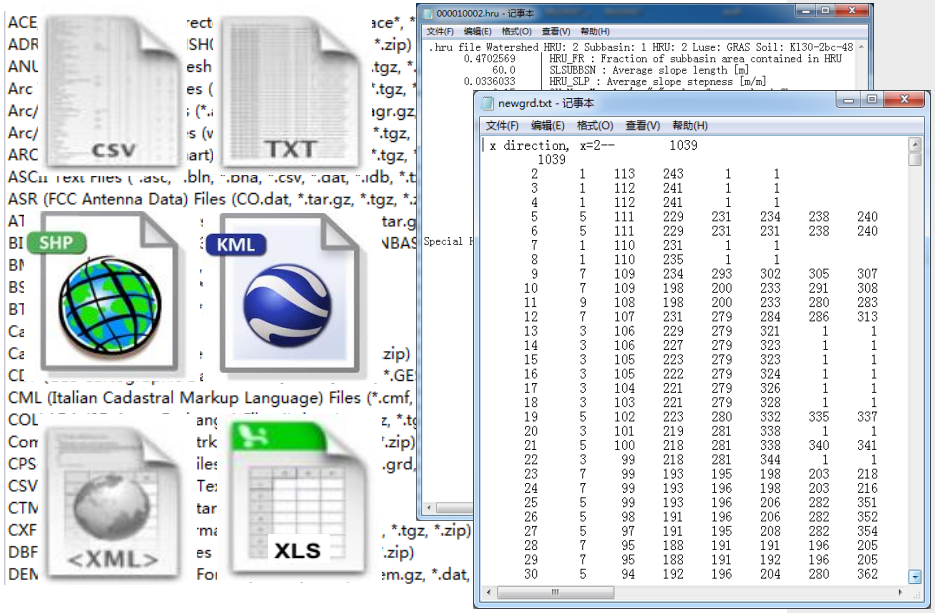


1. Aims

Difficulties in Sharing and Integrating Models

- Data preparation and pre-processing work

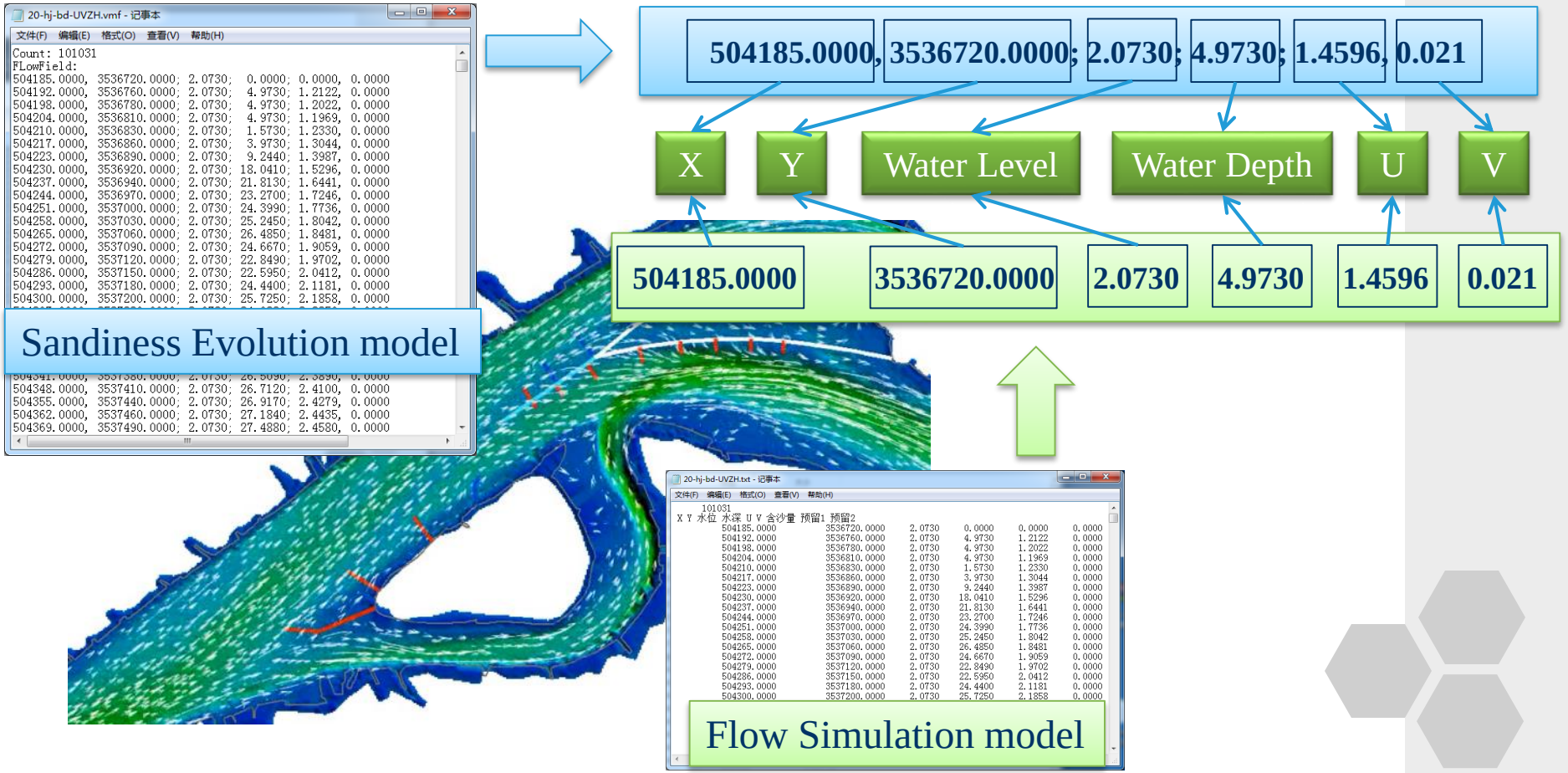




While the integration frameworks of geo-analysis models could permit their practical use, it is essential for model users to **prepare data according to the specific requirements** of the different geo-analysis models.

1. Aims

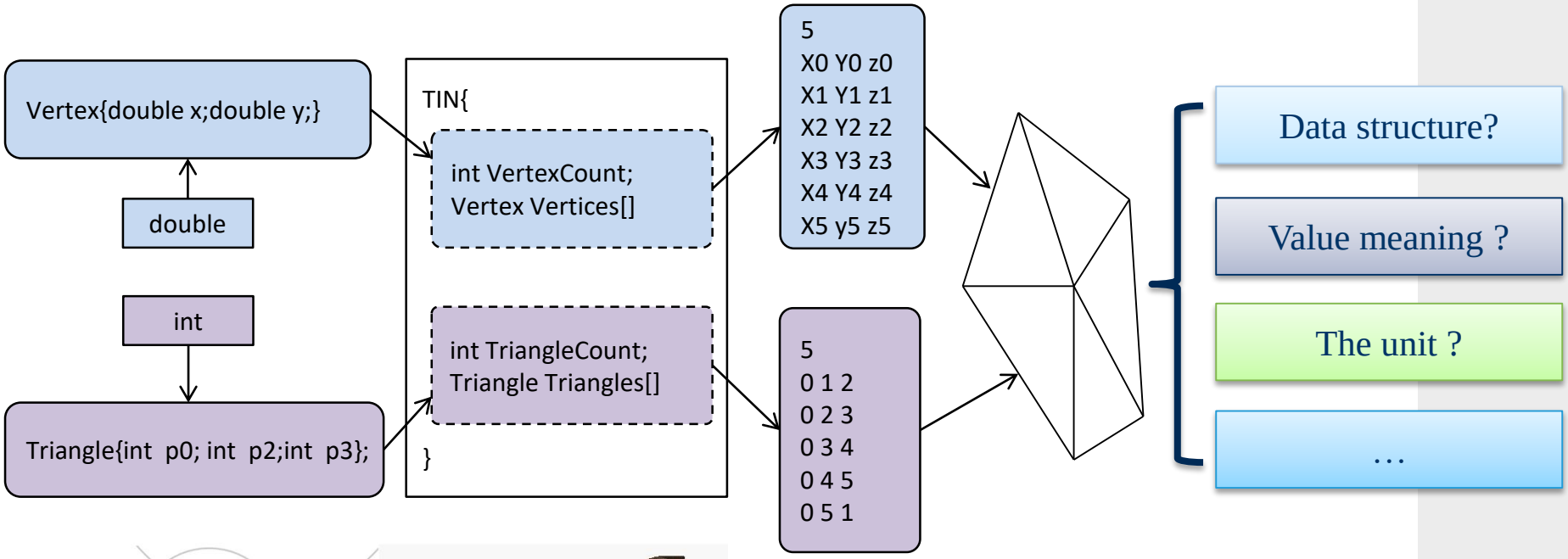
- Difficulties in Sharing and Integrating Models
  - Data preparation and pre-processing work

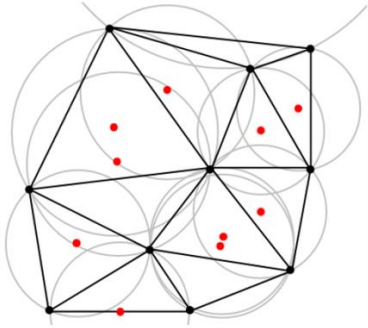


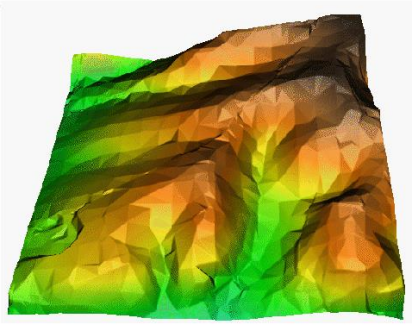
1. Aims

Difficulties in Sharing and Integrating Models

- Data preparation and pre-processing work

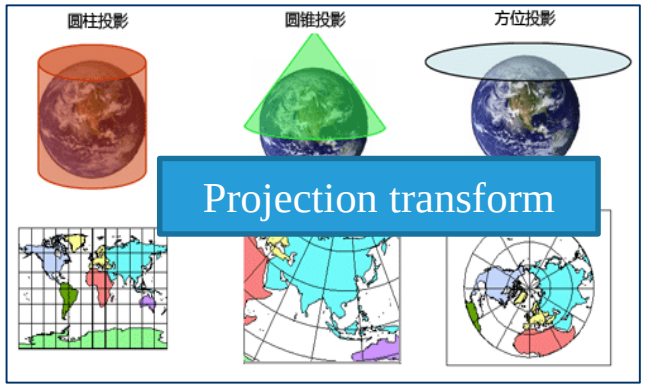


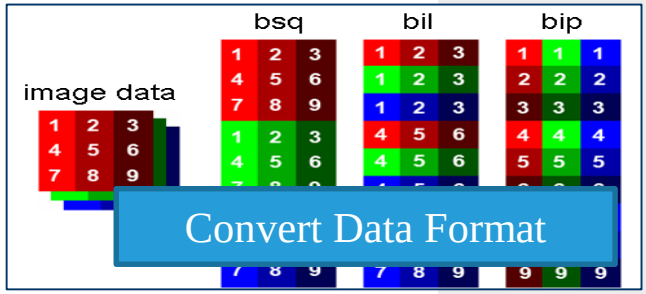


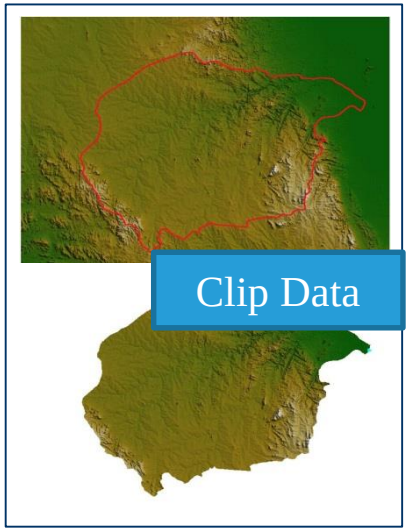


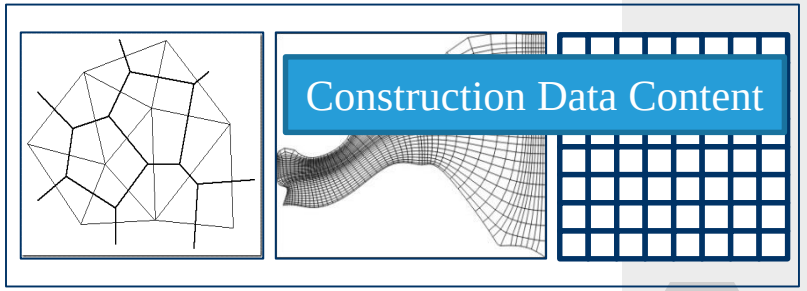
1. Aims

- Difficulties in Sharing and Integrating Models
  - Data preparation and pre-processing work









……

1. Aims

- Basic Aim of our study
  - The heterogeneities of model data lead to the difficulty in using geo-analysis models.
  - Data preparation and pre-processing work are needed to reuse, share and integrate geo-analysis models, however, the data processing methods can hardly be accumulated to be shared among modelers.

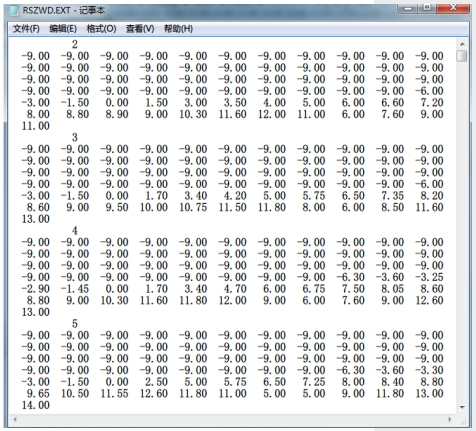
- Provide possibility and methods for modelers to **describe model data**, to **constructed model data**, and to **accumulate and share the data processing methods**.
- More concerned at modeling logic other than repetitive coding tasks.

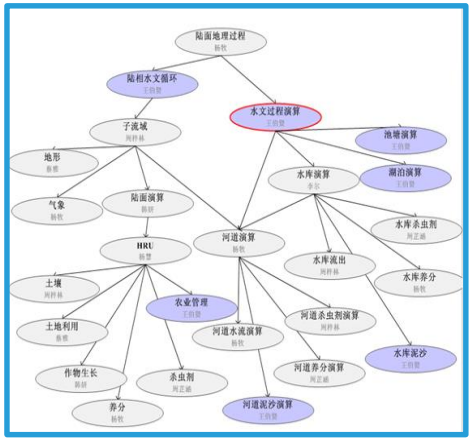
Outline

- Aims
- Data Description Model
- Architecture Design and Practice
- Conclusions

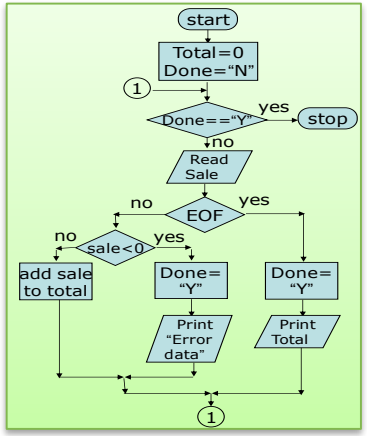
2. Data Description Model

- Geo-analysis models should be implemented by programming codes, all model data are converted into variables of the employed program languages.
- Using a series of basic variable types and the combination of these variable types, a flexible data description model (**Universal Data eXchange model, UDX**) is proposed to reduce the heterogeneity of model data.

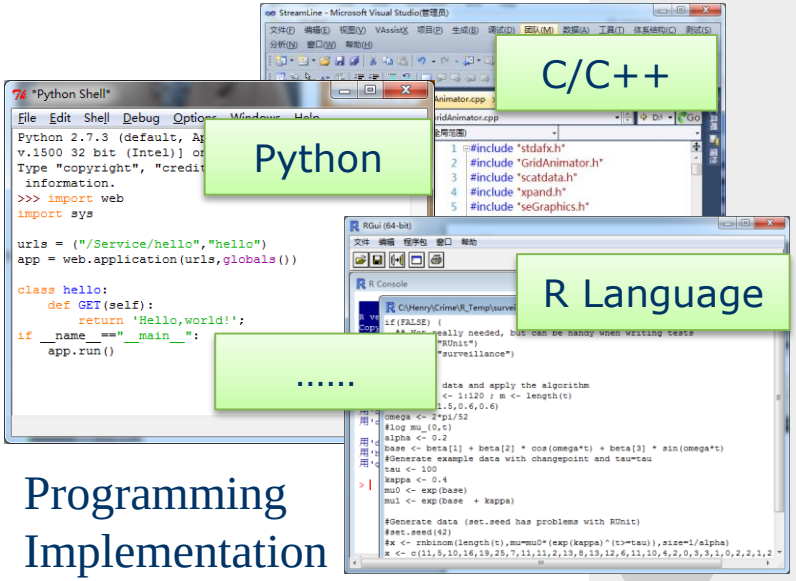




Basic concept design of model



Design of program logic

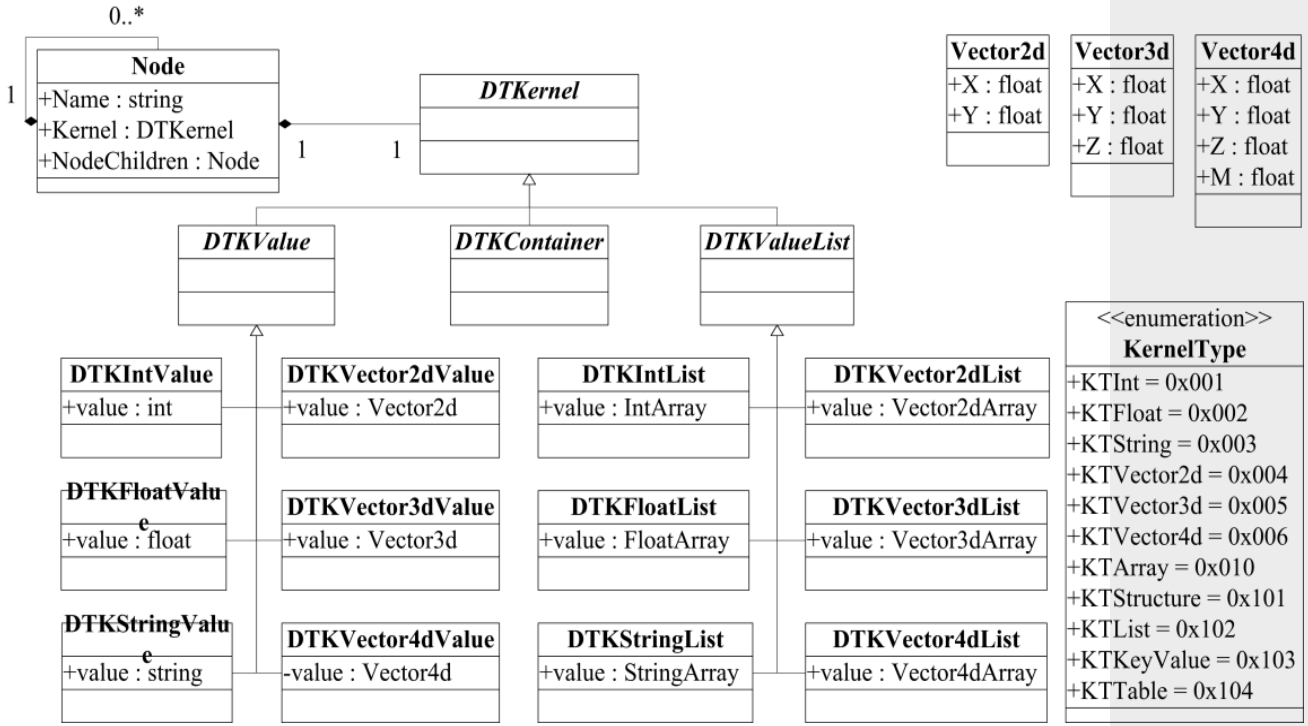


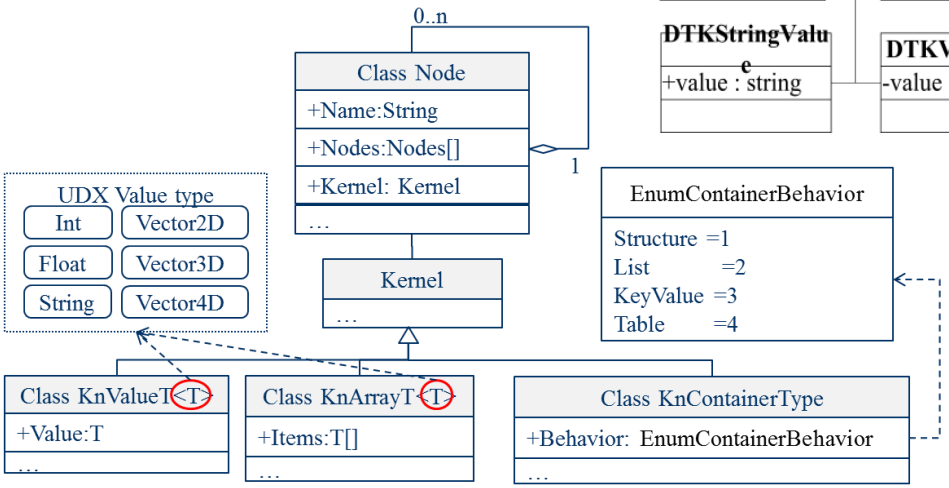
Programming Implementation

2. Data Description Model

Basic Design of UDX Model

(1) Simple variables and variable sets are the most common basic elements.
(2) A hierarchical structure and a two-dimensional table structure are the most general forms of data organisation.





(3) Naming is the simplest way to distinguish data contents.
(4) The meaning of a data structure generally relies on its explanation.

2. Data Description Model

Name=Watershed
KernelType=*KTStructure*

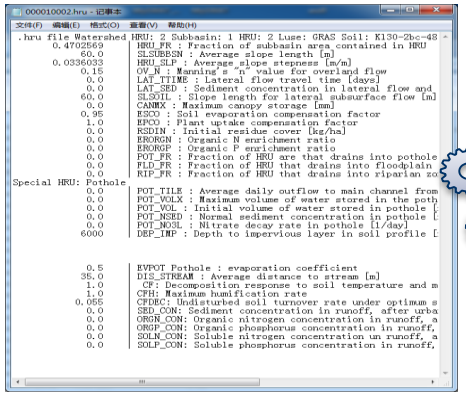
- Name=Watershed_Name
  KernelType=*KTStringValue*
- Name=Watershed_Area
  KernelType=*KTFloatValue*
- Name=Watershed_Boundary
  KernelType=*KTVector2dList*
- Name=SubbasinList
  KernelType=*KTList*
  - …
- …

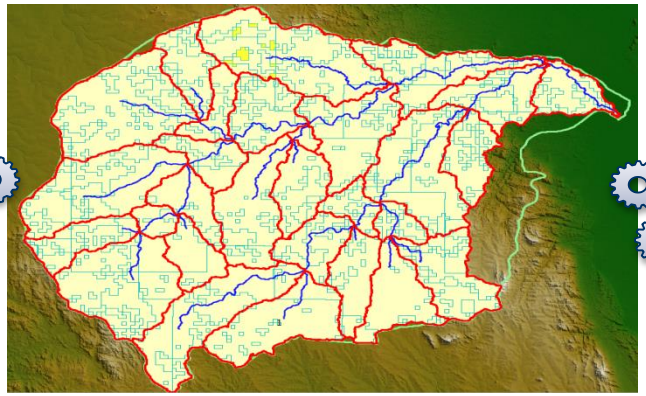
Name=Subbasin_Item
KernelType=*KTStructure*

- Name=Subbasin_ID
  KernelType=*KTIntValue*
- Name=Subbasin_Area
  KernelType=*KTFloatValue*
- Name=Subbasin_Boundary
  KernelType=*KTVector2dList*
- Name=Subbasin_Attribute
  KernelType=*KTTable*
  - …
- …

Name=Subbasin_Attribute
KernelType=*KTTable*

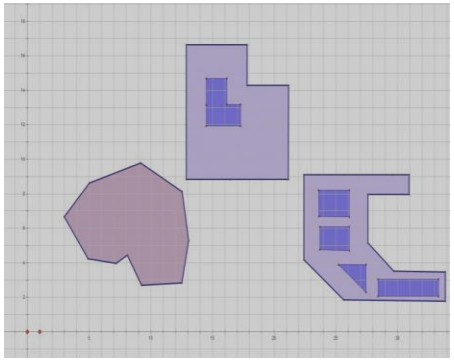
- Name=HRU_ID
  KernelType=*KTIntList*
- Name=HRU_Fraction
  KernelType=*KTFloatList*
- Name=HRU_SlopeLength
  KernelType=*KTFloatList*
- Name=HRU_AverageSlope
  KernelType=*KTFloatList*
- …





SIN	LANDUSE	SOIL	SLOPE_BAND	AREA (ha)	%SUBBASIN	HRUGIS
	CRDY	Lf91-3ab-781	0-244	2309.72037	3.81991501	NA
	CRDY	Lf88-3b-778	0-244	5979.85632	9.88974388	NA
	CRWO	Lf91-3ab-781	0-244	14919.4325	24.6743999	NA
	CRWO	Lf88-3b-778	0-244	7963.88251	13.1710118	NA
	SAVA	Lf91-3ab-781	0-244	7552.07729	12.4899506	NA
	SAVA	I-Bc-c-644	0-244	2587.49793	4.27931549	NA
	SAVA	Lf88-3b-778	0-244	19152.7624	31.6756632	000010001
	CRDY	Lf91-3ab-781	0-244	9937.49205	49.1381086	000020001
	CRWO	Lf91-3ab-781	0-244	3148.60859	15.5689856	NA
	SAVA	Lf91-3ab-781	0-244	7137.49429	35.2929057	NA
	CRDY	Lf91-3ab-781	0-244	41286.7725	70.03864	000030001
	CRDY	Fo75-2-3a-534	0-244	575.693983	0.97660391	NA
	CRWO	Lf91-3ab-781	0-244	9.72221444	0.0164927	NA

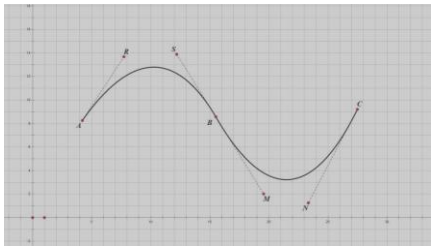
2. Data Description Model



Name =Polygon
Kernel=KTStructure

- Name =ExternalRing
  Kernel=KTVector2dList
- Name =InnerRings
  Kernel=KnList
  - Name =InnerRing
    Kernel=KTVector2dList
  - ...
  - Name =InnerRing
    Kernel=KnVector2dList

(a) Simple feature class



Name =BezierCurve
Kernel=KTStructure

- Name =BezierSegment
  Kernel=KTStructure
  - Name =StartPoint
    Kernel=KTVector2d
  - Name =EndPoint
    Kernel=KTVector2d
  - Name =ControlPoint1
    Kernel=KTVector2d
  - Name =ControlPoint2
    Kernel=KTVector2d
  - Name =ControlValue
    Kernel=KTIntValue
- ...
- Name =BezierSegment
  Kernel=KTStructure

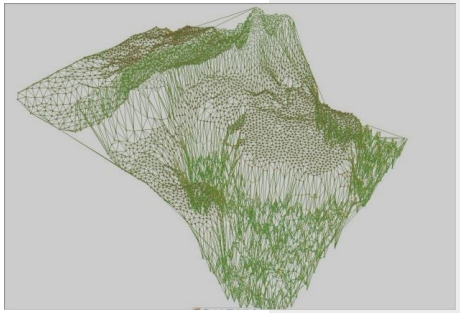
(b) Parameter curve

1	1	2	3	5	5	6	5
1	1	2	3	5	5	6	6
0	0	4	3	7	7	7	7
0	0	4	4	7	7	7	7
8	8	8	8	7	7	7	7
8	8	8	8	9	9	9	9
8	8	8	8	9	9	9	9
8	8	8	8	9	9	9	9

Name=GRID
Kernel=KTStructure

- Name=ORIGIN
  Kernel=KTVector2D
- Name=XCount
  Kernel=KTInt
- Name=YCOUNT
  Kernel=KTINT
- Name=ROWS
  Kernel=KTList
  - Name=ROW
    Kernel=KTFloatList
  - Name=ROW
    Kernel=KTFloatList

(c) Grid data

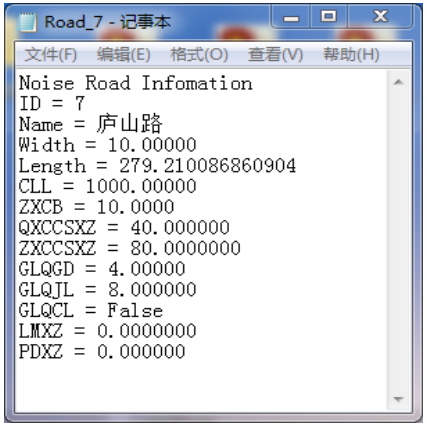


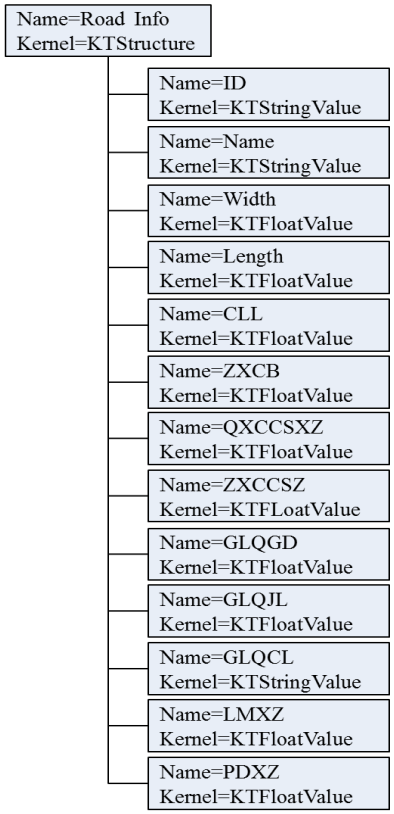
Name=TIN
Kernel=KTStructure

- Name=NODES
  Kernel=KTVector3DList
- Name=TRIANGLS
  Kernel=KTIntList

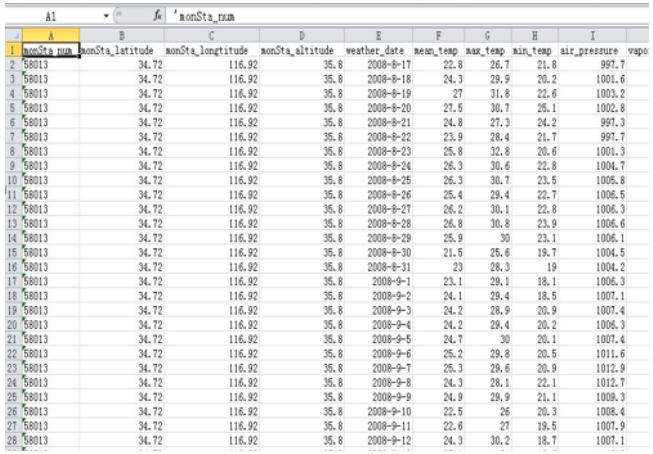
(d) TIN data

2. Data Description Model

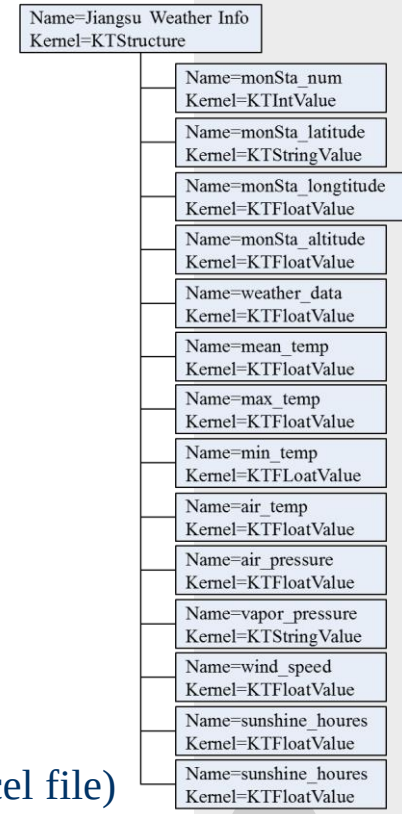




(e) Text data



monSta_num	monSta_latitude	monSta_longtitude	monSta_altitude	weather_date	mean_temp	max_temp	min_temp	air_pressure	vapo
58013	34.72	116.92	35.8	2008-8-17	22.8	26.7	21.8	997.7	
58013	34.72	116.92	35.8	2008-8-18	24.3	29.9	20.2	1001.6	
58013	34.72	116.92	35.8	2008-8-19	27	31.8	22.6	1003.2	
58013	34.72	116.92	35.8	2008-8-20	27.6	30.7	25.1	1002.8	
58013	34.72	116.92	35.8	2008-8-21	24.8	27.3	24.2	997.3	
58013	34.72	116.92	35.8	2008-8-22	23.9	28.4	21.7	997.7	
58013	34.72	116.92	35.8	2008-8-23	25.8	32.8	20.6	1001.3	
58013	34.72	116.92	35.8	2008-8-24	26.3	30.6	22.8	1004.7	
58013	34.72	116.92	35.8	2008-8-25	26.3	30.7	23.5	1005.8	
58013	34.72	116.92	35.8	2008-8-26	25.4	29.4	22.7	1006.5	
58013	34.72	116.92	35.8	2008-8-27	26.2	30.1	22.8	1006.3	
58013	34.72	116.92	35.8	2008-8-28	26.8	30.8	23.9	1006.6	
58013	34.72	116.92	35.8	2008-8-29	25.9	30	23.1	1006.1	
58013	34.72	116.92	35.8	2008-8-30	21.5	25.6	19.7	1004.5	
58013	34.72	116.92	35.8	2008-8-31	23	28.3	19	1004.2	
58013	34.72	116.92	35.8	2008-9-1	23.1	29.1	18.1	1006.3	
58013	34.72	116.92	35.8	2008-9-2	24.1	29.4	18.5	1007.1	
58013	34.72	116.92	35.8	2008-9-3	24.2	28.9	20.9	1007.4	
58013	34.72	116.92	35.8	2008-9-4	24.2	29.4	20.2	1006.3	
58013	34.72	116.92	35.8	2008-9-5	24.7	30	20.1	1007.4	
58013	34.72	116.92	35.8	2008-9-6	25.2	29.8	20.5	1011.6	
58013	34.72	116.92	35.8	2008-9-7	25.3	29.6	20.9	1012.9	
58013	34.72	116.92	35.8	2008-9-8	24.3	28.1	22.1	1012.7	
58013	34.72	116.92	35.8	2008-9-9	24.9	29.9	21.1	1009.3	
58013	34.72	116.92	35.8	2008-9-10	22.5	26	20.3	1008.4	
58013	34.72	116.92	35.8	2008-9-11	22.6	27	19.5	1007.9	
58013	34.72	116.92	35.8	2008-9-12	24.3	30.2	18.7	1007.1	

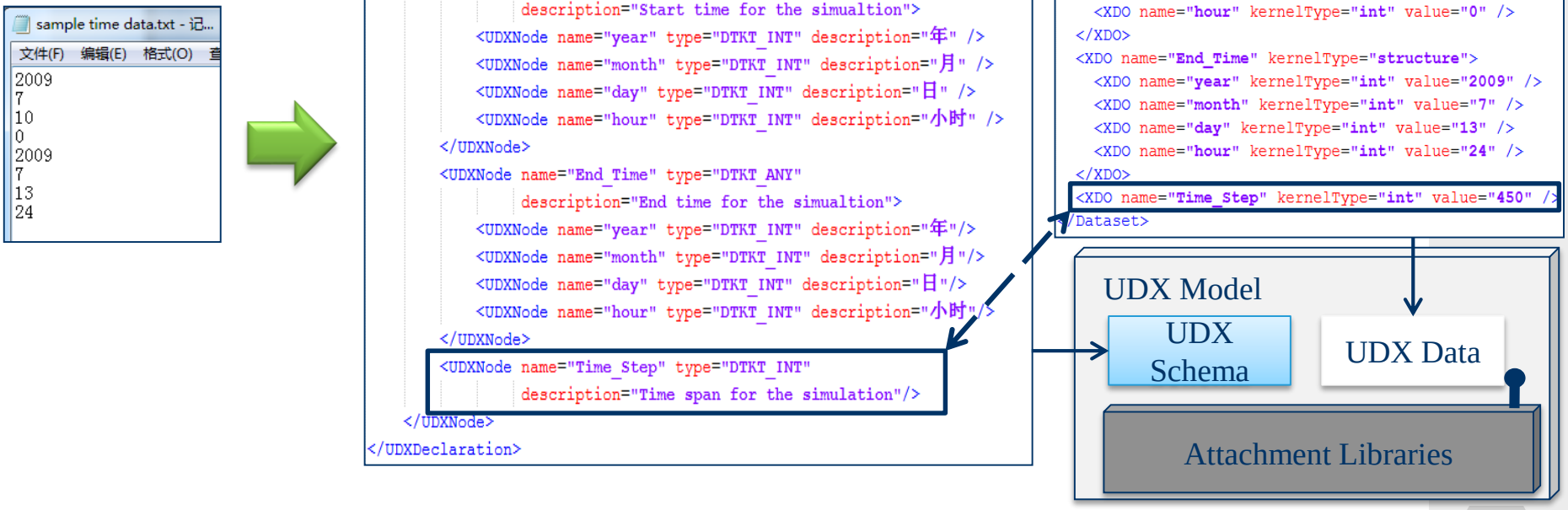


(f) Structural data (excel file)

2. Data Description Model

❖ UDX Data & UDX Schema

Derived from the UDX model: UDX Data and UDX Schema.



UDX Data is designed to interchange model data in *original format* and *UDX model*.

UDX Schema is designed to provide the skeleton frame information of model data, which should keep tightly consistent with the corresponding UDX Data.

2. Data Description Model

Group	Kernel Type	Example
DTKValue	*KTInt* ➔ (a.1) *KTFloat* ➔ (a.2) *KTString* ➔ (a.3) *KTVector2d* ➔ (a.4) *KTVector3d* ➔ (a.5) *KTVector4d* ➔ (a.6)	**(a.1)** <XDO name="Analysis_Window_Size" kernelType="int" value="3" /> **(a.2)** <XDO name="Flow_Accumulation_Threshod" kernelType="float" value="7200.0" /> **(a.3)** <XDO name="Landuse_Type" kernelType="string" value="Grass" /> **(a.4)** <XDO name="Watershed_Pour_Out" kernelType="vector2d" value="512312.4323, 412332.2989" /> **(a.5)** <XDO name="Sample_Point_Depth" kernelType="vector3d" value="542166.19, 3546342.10, 0.67" /> **(a.6)** <XDO name="Sample_Point_Velocity" kernelType="vector4d" value= "542166.19, 3546342.10, 0.845, 0.263" />
DTKValueList	*KTIntList* ➔ (b.1) *KTFloatList* ➔ (b.2) *KTStringList* ➔ (b.3) *KTVector2dList* ➔ (b.4) *KTVector3dList* ➔ (b.5) *KTVector4dList* ➔ (b.6)	**(b.1)** <XDO name="Feature_ID" kernelType="int|list" value="0,1,2,3,4" /> **(b.2)** <XDO name="Watershed_Area" kernelType="float|list" value="103.212, 453.235, 1018.14, 763.213" /> **(b.3)** <XDO name="Landuse_Code" kernelType="string|list" value="URMD,CRDY,CRIR,GRAS" /> **(b.4)** <XDO name="Boundary_Line" kernelType="vector2d|list" value="3876.676,2722.342; 4175.325,2816.684; 4473.974,2911.024; 4683.389,3098.370; 4978.841,3195.912; 3876.676,2722.342; …" /> **(b.5)** <XDO name="Section_Depth" kernelType="vector3d|list" value="166.191,6342.109,0.667; 266.191,6342.109,0.721; 366.191,6342.109,0.721; …" /> **(b.6)** <XDO name="Section_Velocity" kernelType="vector4d|list" value="166.191,6342.109, 0.845,0.263; 266.191,6342.109, 0.333,0.606; 366.191,6342.109, 0.915,-0.116; …" />

2. Data Description Model

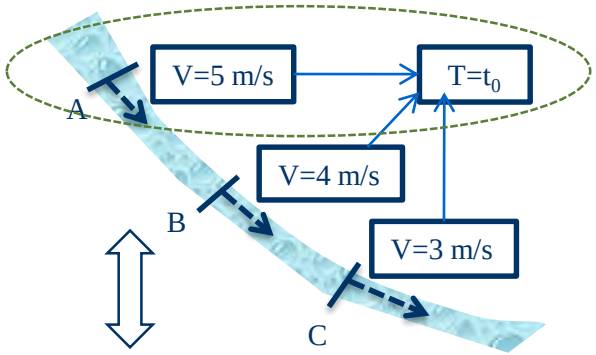
Group	Kernel Type	Example

DTKContainer	*KTStructure* ➔ (c.1)
	KTList ➔ (c.2)
	KTKeyValue ➔ (c.3)
	KTTable ➔ (c.4)
		(c.1)
		`<XDO name="Watershed_Info" kernelType="structure">`
		`<XDO name="name" kernelType="string" value="" />`
		`<XDO name="area" kernelType="float" value="" />`
		`<XDO name="perimeter" kernelType="float" value="" />`
		`<XDO name="boundary" kernelType="vector2d|list" value=" 3876.676,2722.342; 4175.325,2816.684; …" />`
		`</XDO>`
		(c.2)
		`<XDO name="Subbasin_Information" kernelType="list">`
		`<XDO name="subbasin_1" kernelType="structure">`
		`<XDO name="id" kernelType="int" value="1" />`
		`<XDO name="area" kernelType="float" value="419643.843" />`
		`<XDO name="perimeter" kernelType="float" value="353265" />`
		`</XDO>`
		`<XDO name="subbasin_2" kernelType="structure">`
		`<XDO name="id" kernelType="int" value="2" />`
		`<XDO name="area" kernelType="float" value="329113.923" />`
		`<XDO name="perimeter" kernelType="float" value="212314" />`
		`</XDO>`
		`</XDO>`
		(c.3)
		`<XDO name="Monitor_Station" kernelType="keyvalue">`
		`<XDO name="key" kernelType="int" value="1" />`
		`<XDO name="value" kernelType="structure">`
		`<XDO name="wind_speed" kernelType="float" value="321.59" />`
		`<XDO name="air_monisture" kernelType="float" value="1.242" />`
		`<XDO name="air_presure" kernelType="float" value="98.32" />`
		`</XDO>`
		`</XDO>`
		(c.4)
		`<XDO name="Basin_Feature" kernelType="table">`
		`<XDO name="id_column" kernelType="int|list" value="0,1,2,3, …" />`
		`<XDO name="area_column" kernelType="float|list" value="77.514, 51.514, 224.96, 169.4, …" />`
		`<XDO name="flow_in_column" kernelType="float|list" value="5.636, 3.915, 27.48, 14.15, … " />`
		`<XDO name="flow_out_column" kernelType="float|list" value="5.632, 3.915, 27.43, 14.13, … " />`
		`</XDO>`

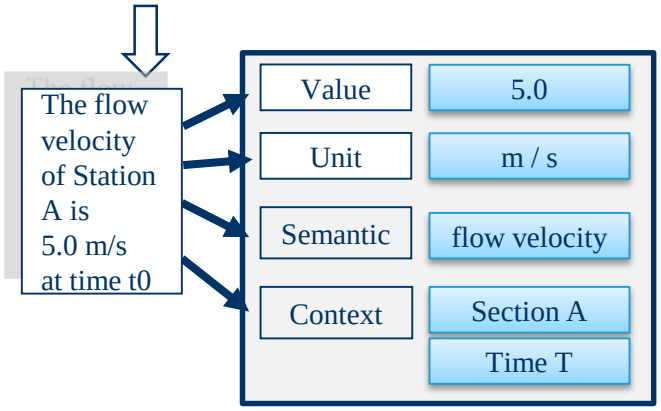
2. Data Description Model

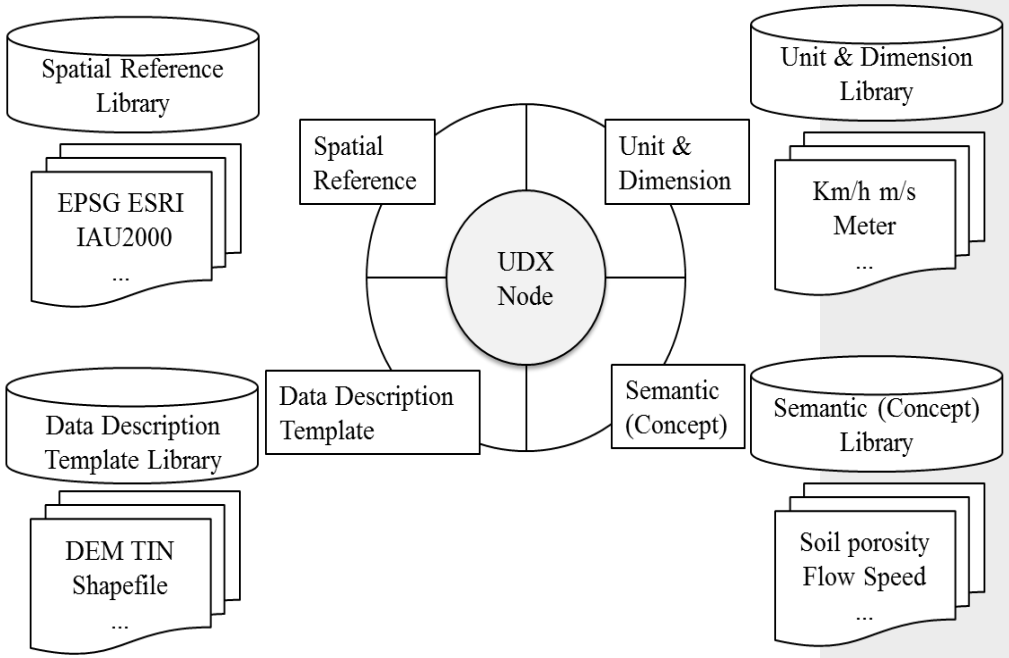
UDX Schema

Attachment Libraries



T	A	B	C
t0	5	4	3
t1	--	--	--

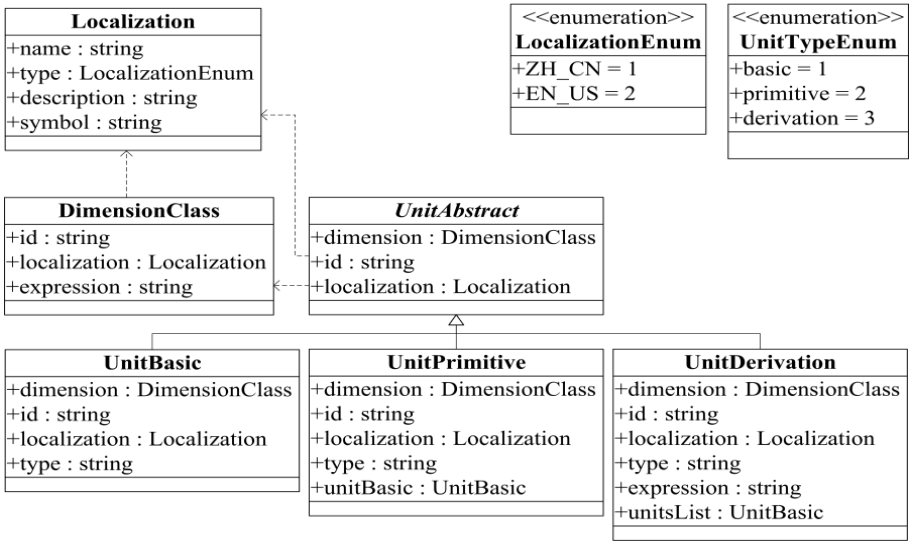




- By embedding the semantic information into the UDX model, the UDX model can be formed and employed to represent both data values and semantics.

2. Data Description Model

Unit & Dimension Library




```
<Dimension id="L" type="basic">
  <Localization local="ZH-CN" name="长度"  description="就是长度，距离，高度"/>
  <Localization local="EN-US" name="Length" description="length height"/>
</Dimension>
<Dimension id="T" type="basic">
  <Localization local="ZH-CN" name="时间" description="时
  <Localization local="EN-US" name="Time" description="Time"/>
</Dimension>
<Dimension id="SPEED" type="compound" expresion="L/T">
  <Localization local="ZH-CN" name="速度" description="距离与时间之比"/>
  <Localization local="EN-US" name="Speed" description="Length divided by Time"/>
</Dimension>
```

Dimension

```
<Unit id="M"  type="basic" dimension="L">
  <Localization local="ZH-CN" name="米" description="米,国际单位" symbol="米"/>
  <Localization local="EN-US" name="Meter" description="Meter, International Unit" symbol="m"/>
</Unit>
```

```
<Unit id="S" type="basic" dimension="T">
  <Localization local="ZH-CN" name="秒" description="秒，国际单位" symbol="秒"/>
  <Localization local="EN-US" name="second" description="Second,International Unit" symbol="s"/>
</Unit>
```

Basic

```
<Unit id="KM" type="primtive" dimension="L" expression="1000*M">
  <Localization local="ZH-CN" name="千米" description="千米，公里 距离单位" symbol="千米"/>
  <Localization local="EN-US" name="kilometre" description="kilometre" symbol="km"/>
</Unit>
```

```
<Unit id="HOUR" type="primtive" dimension="T" expression="3600
  <Localization local="ZH-CN" name="小时" description="小时" symbol="小时"/>
  <Localization local="ZH-CN" name="hour" description="hour" symbol="h"/>
</Unit>
```

Primitive

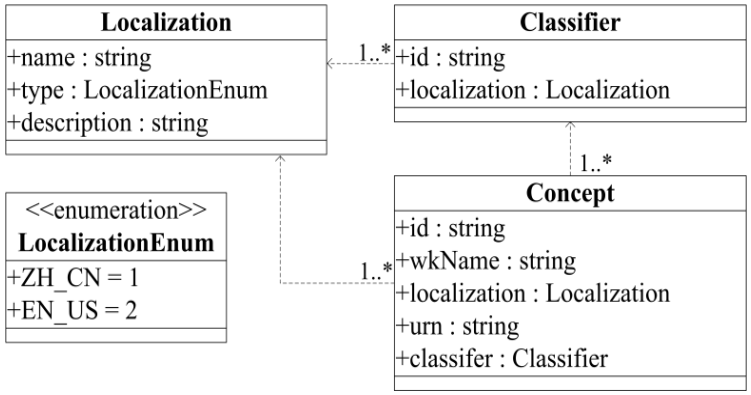
```
<Unit id="METRE_PRE_SECOND" type="derivation" dimension="SPEED">
  <Localization local="ZH-CN" name="米每秒" description="米每秒" symbol="米/秒"/>
  <Localization local="EN-US" name="m/s" description="metre per second" symbol="m/s"/>
  <Dimensions>
    <Dimension dimension="L" unit="M"/>
    <Dimension dimension="T" unit="S"/>
  </Dimensions>
</Unit>
```

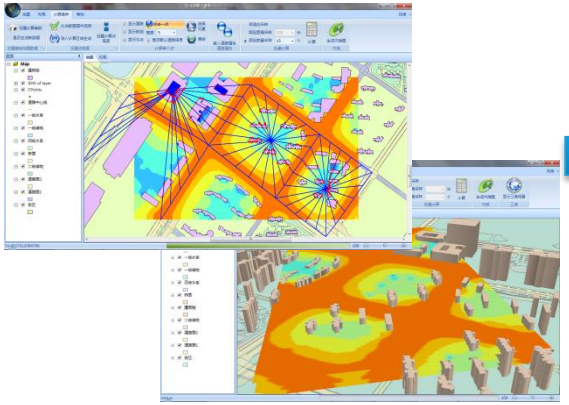
```
<Unit id="KILOMETRE_PRE_HOUR" type="derivation" dimension="SPEED">
  <Localization local="ZH-CN" name="千米每小时" description="千米每小时" symbol="千米/小时"/>
  <Localization local="EN-US" name="KM/H" description="kilometre per hour" symbol="km/h"/>
  <Dimensions>
    <Dimension dimension="L" unit="KM"/>
    <Dimension dimension="T" unit="HOUR"/>
  </Dimensions>
</Unit>
```

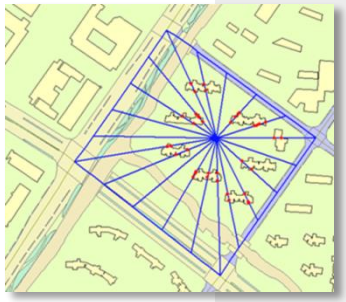
Derivation

2. Data Description Model

Semantic (Concept) Library







City Main Road

```
<Concept id="GIS.DIGITAL_ELEVATION_MODEL" wkName="DEM"
    urn="http://en.wikipedia.org/wiki/Digital_elevation_model">
  <Localization local="ZH_CN" name="数字高程模型"
      description="用网格离散的方式表达地形的高度" />
  <Localization local="EN_US" name="Digital Elevation Model"
      description="3D representation of a terrain's surface" />
</Concept>
```

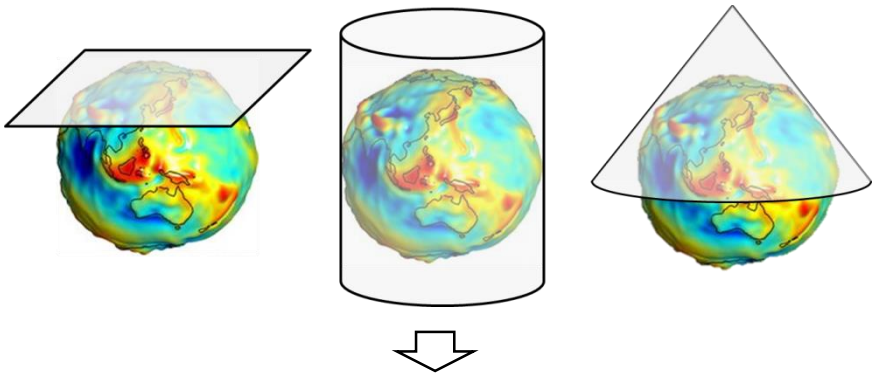
```
<Concept id="RegularRiver" wkname="常流河"
    urn="http://en.wikipedia.org/wiki/#regularriver">
  <Localizations>
    <Localization local="ZH-CN" name="常流河"
      description="一年四季一直流动的河流"></Localization>
    <Localization local="EN-US" name="RegularRiver"
      description="River flow all around a year"></Localization>
  </Localizations>
</Concept>
```

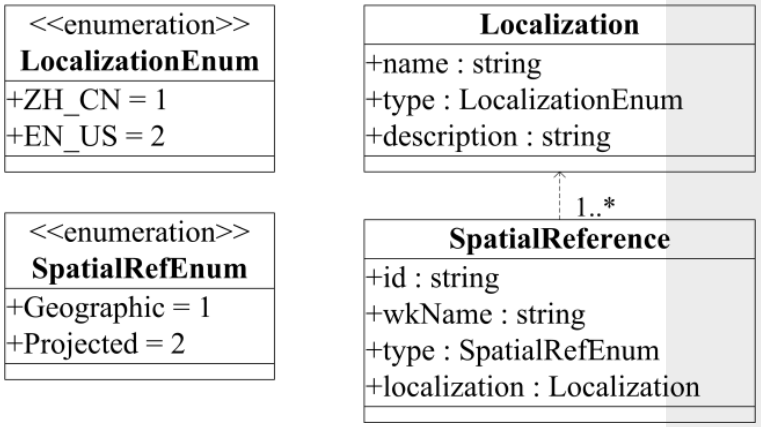
```
<Concept id="Road.CityMainRoad" wkname="MainRoad"
    urn="http://en.wikipedia.org/wiki/#road">
  <Localizations>
    <Localization local="ZH-CN" name="城市主干道"
      description="城市中机动车行驶的主干道" />
    <Localization local="EN-US" name="Main road in city"
      description="The main road in city" />
  </Localizations>

  <Classifers>
    <Classifer id="Length">
      <Localization local="ZH-CN" name="宽度"
        description="按照宽度进行分类" />
      <Localization local="EN-US" name="Length"
        description="Classify Road Based on the Width" />
    </Classifer>
    <Classifer id="Material">
      <Localization local="ZH-CN" name="路面材质"
        description="按照路面材质进行分类" />
      <Localization local="EN-US" name="Material"
        description="Classify Road Based on the Material" />
    </Classifer>
  </Classifers>
</Concept>
```

2. Data Description Model

❖ Spatial Reference Library






```
<SpatialRef id="EPSG.4326" type="Geographic" wkName="WGS_1984">
  <Localization local="ZH_CN" name="WGS_84"
        description="是为GPS全球定位系统使用而建立的坐标系统" />
  <Localization local="EN_US" name="WGS_84"
        description="World Geodetic System 1984" />
</SpatialRef>
```

```
<SpatialRef id="EPSG.2327" type="Projection" wkName="Xian1980_Zone13">
  <Localization local="ZH_CN" name="1980西安坐标系"
        description="采用1975国际椭球，
        以JYD 1968.0系统为椭球定向基准，
        大地原点设在陕西省泾阳县永乐镇，
        采用多点定位所建立的大地坐标系....." />
  <Localization local="EN_US" name="Xian1980_Zone13"
        description="Xian Geodetic Coordinate System 1980" />
</SpatialRef>
```

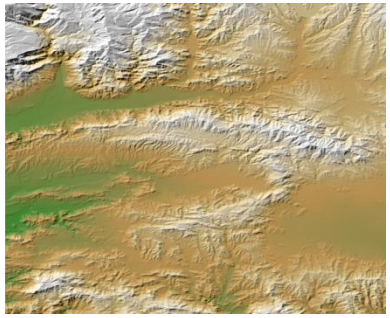
WGS 84: 93.9154, 30.5076 decimal degrees

World Mercator: 10454.6154, 3547.2713 km

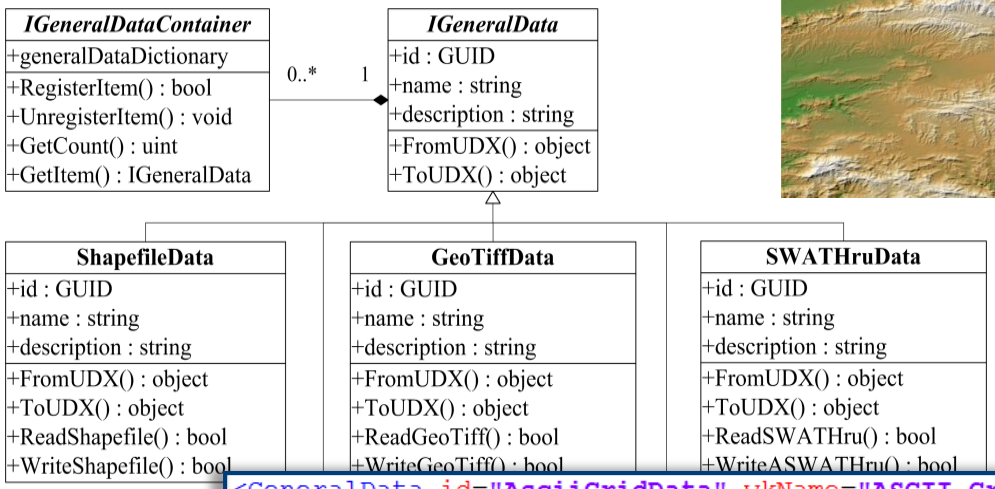
World Cylindrical Equal Area: 10454.6154, 3219.9029 km

2. Data Description Model

❖ Data Description Template Library

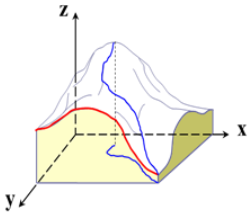


- Arc/Info Binary Grid (.adf)
- Arc/Info ASCII Grid (.asc)
- NetCDF (.nc)
- GeoTIFF (.tif)
- USGS ASCII DEM / CDED (.dem)
- ASCII Gridded XYZ (.xyz)
- Golden Software ASCII Grid (.grd)
- SRTM HGT Format (.hgt)
- .....




```
<GeneralData id="AsciiGridData" wkName="ASCII_Grid_Data"
        description="common used Grid data format (.asc)">
<GeneralData id="ShapeFileData" wkName="Shapefile_Data"
        description="common used geometry data format (.shp)">
<GeneralData id="GeoTiffData" wkName="GeoTiff_Data"
        description="common used Image or Grid data format (.tif)">
<GeneralData id="HRUData" wkName="HRU_Data"
        description="common used hydro response unit data format (.hru)">
  <Localization local="ZH_CN" name="水文响应单元数据"
        description="常用的一种SWAT水文响应单元数据格式" />
  <Localization local="EN_US" name="HRU"
        description="A common used hydro response unit format" />
</GeneralData>
```

2. Data Description Model



154	149	144	141	138
157				
161				
164				
168				

```
<?xml version="1.0" encoding="utf-8"?>
<Dataset name="SlopeAnalysis_Input" kernelType="structure">
  <XDO name="Dem_Data" kernelType="structure">
    <XDO name="head" kernelType="structure">
      <XDO name="ncols" kernelType="int" value="786" />
      <XDO name="nrows" kernelType="int" value="767" />
      <XDO name="xllcorner" kernelType="float" value="471090.08257249" />
      <XDO name="yllcorner" kernelType="float" value="208342.35339682" />
      <XDO name="cellsize" kernelType="float" value="30" />
      <XDO name="NODATA_value" kernelType="float" value="-9999" />
    </XDO>
```

Data Node

```
<?xml version="1.0" encoding="utf-8"?>
<UDXDeclaration name="SlopeAnalysis_Input" description="" >
  <UDXNode>
    <UDXNode name="Dem_Data" type="DTKT_STRUCTURE" description="DEM Data">
      <UDXNode name="head" type="DTKT_STRUCTURE" description="Header Info of DEM Data">
        <UDXNode name="ncols" type="DTKT_INT" description="Number of column count" />
        <UDXNode name="nrows" type="DTKT_INT" description="Number of row count" />
        <UDXNode name="xllcorner" type="DTKT_FLOAT" description="Minimun X value" />
        <UDXNode name="yllcorner" type="DTKT_FLOAT" description="Minim
        <UDXNode name="cellsize" type="DTKT_FLOAT" description="Resolu
        <UDXNode name="NODATA_value" type="DTKT_FLOAT" description="Flag of NODATA valu
      </UDXNode>
      <UDXNode name="body" type="DTKT_LIST" description="DEM height value content">
        <UDXNode name="row" type="DTKT_FLOAT|DTKT_LIST" description="Row of value conte
      </UDXNode>
    </UDXNode>
    <UDXNode name="winSize" kernelType="DTKT_VECTOR2D" description=
  </UDXNode>

  <SemanticAttachment>
    <Units>
      <Unit nodeId="cellsize" unitId="Units.Measure.M" />
    </Units>

    <Concepts>
      <Concept nodeId="Dem_Data" conceptId="Concepts.GIS.DIGITAL_ELEVATION_DATA"/>
    </Concepts>

    <SpatialRefs>
      <SpatialRef nodeId="Dem_Data" spatialRefId="ESPG.4326"/>
    </SpatialRefs>

    <GeneralTypes>
      <GeneralType nodeId="Dem_Data" generalTypeId="GeneralData.AsciiGridData" />
    </GeneralTypes>
  </SemanticAttachment>
</UDXDeclaration>
```

Unit

```
<Unit id="Measure.M" type="basic" dimension="L">
  <Localization local="ZH_CN" name="米"
      description="米，国际单位" symbol="米" />
  <Localization local="EN_US" name="Meter"
      description="Meter, International Unit" symbol="m" />
</Unit>
```

Concept

```
<Concept id="GIS.DIGITAL_ELEVATION_MODEL" wkName="DEM"
    urn="http://en.wikipedia.org/wiki/Digital_elevation_model">
  <Localization local="ZH_CN" name="数字高程模型"
      description="用网格离散的方式表达地形的高度" />
  <Localization local="EN_US" name="Digital Elevation Model"
      description="3D representation of a terrain's surface" />
</Concept>
```

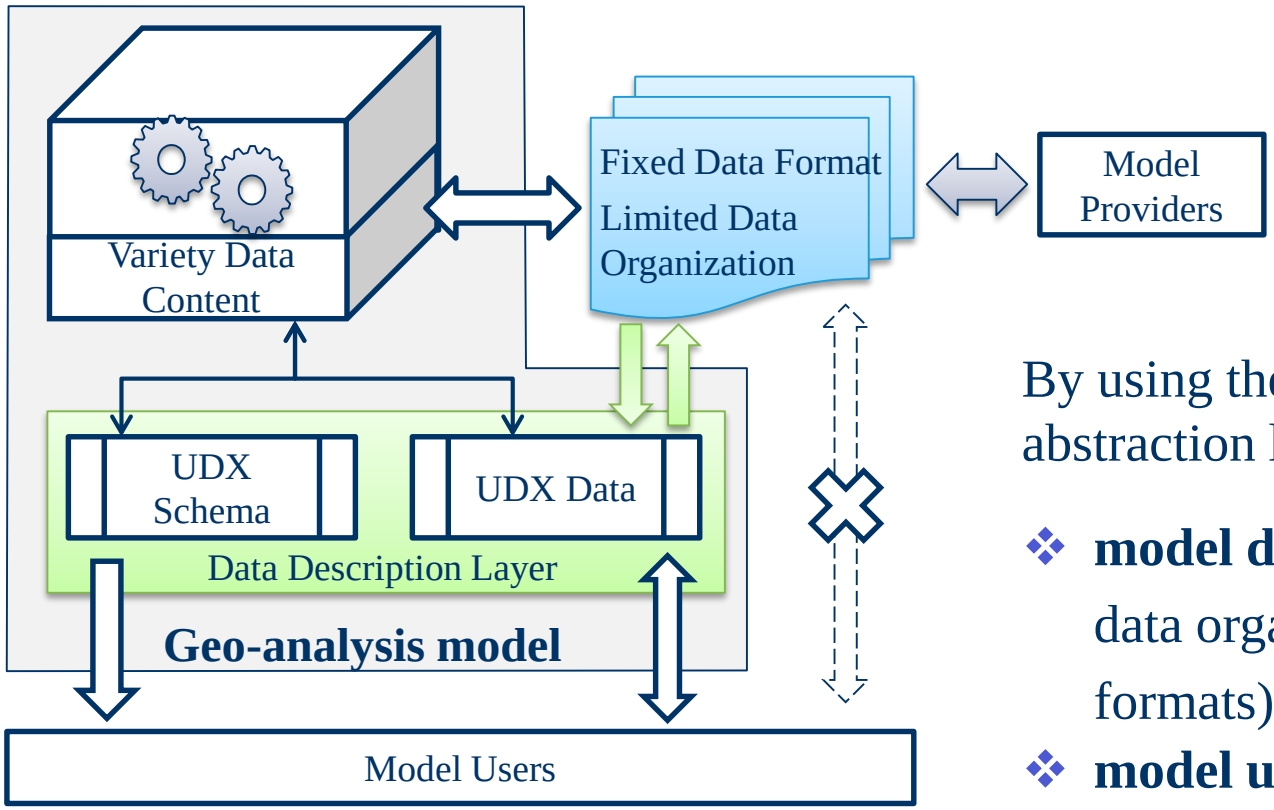
Spatial Reference

```
<SpatialRef id="EPSG.4326" type="Geographic" wkName="WGS_1984">
  <Localization local="ZH_CN" name="WGS_84"
      description="是为GPS全球定位系统使用而建立的坐标系统" />
  <Localization local="EN_US" name="WGS_84"
      description="World Geodetic System 1984" />
</SpatialRef>
```

Data Description Template

```
<GeneralData id="AsciiGridData" wkName="ASCII_Grid_Data"
      description="common used Grid data format (.asc)">
  <Localization local="ZH_CN" name="AsciiGrid"
      description="常用的一种文本文件表示的格网数据格式" />
  <Localization local="EN_US" name="AsciiGrid"
      description="A common used format, such as srtm data" />
</GeneralData>
```

2. Data Description Model



By using the UDX model, a middle abstraction layer is built between:

- **model data** (which require limited data organizations or fixed data formats)
- **model users** (who need to understand model data not only the value content but also its semantic information)

Outline

- Target and Basic Idea
- Data Description Model
- **Architecture Design and Practice**
- Conclusion

3. Architecture Design and Experiment

Ascii GRID

- Name=DEM Data KernelType=*KTStructure*
  - Name=Header KernelType=*KTStructure*
    - Name=NCol KernelType=*KTInt*
    - Name=NRow KernelType=*KTInt*
    - Name=Xllcorner KernelType=*KTFloat*
    - Name=Yllcorner KernelType=*KTFloat*
    - Name=Cellsize KernelType=*KTFloat*
    - Name=Nodata KernelType=*KTFloat*
  - Name=Body KernelType=*KTList*
    - Name=Row KernelType=*KTFloatList*

GeoTiff

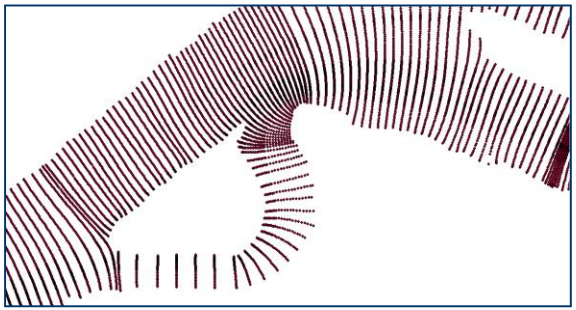
- Name=GeoTiff Data KernelType=*KTStructure*
  - Name=Header KernelType=*KTStructure*
  - Name=Bands KernelType=*KTList*
    - Name=Band KernelType=*KTFloatList*
      - Name=Nodata KernelType=*KTFloat*
      - Name=Offset KernelType=*KTFloat*
      - Name=Scale KernelType=*KTFloat*
      - Name=Unit KernelType=*KTString*
      - Name=Value KernelType=*KTList*
        - Name=Row KernelType=*KTFloatList*

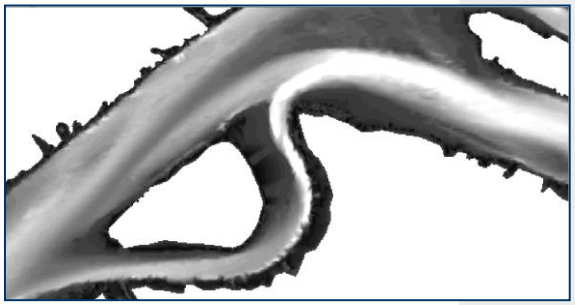
→

- Name=Header KernelType=*KTStructure*
  - Name=Upper left corner X KernelType=*KTFloat*
  - Name=Pixel width KernelType=*KTFloat*
  - Name=Tansform param 1 KernelType=*KTFloat*
  - Name=Upper left corner Y KernelType=*KTFloat*
  - Name=Tansform param 2 KernelType=*KTFloat*
  - Name=Pixel height KernelType=*KTFloat*
  - Name=Width KernelType=*KTInt*
  - Name=Height KernelType=*KTInt*

3. Architecture Design and Experiment

Data Processing With UDX





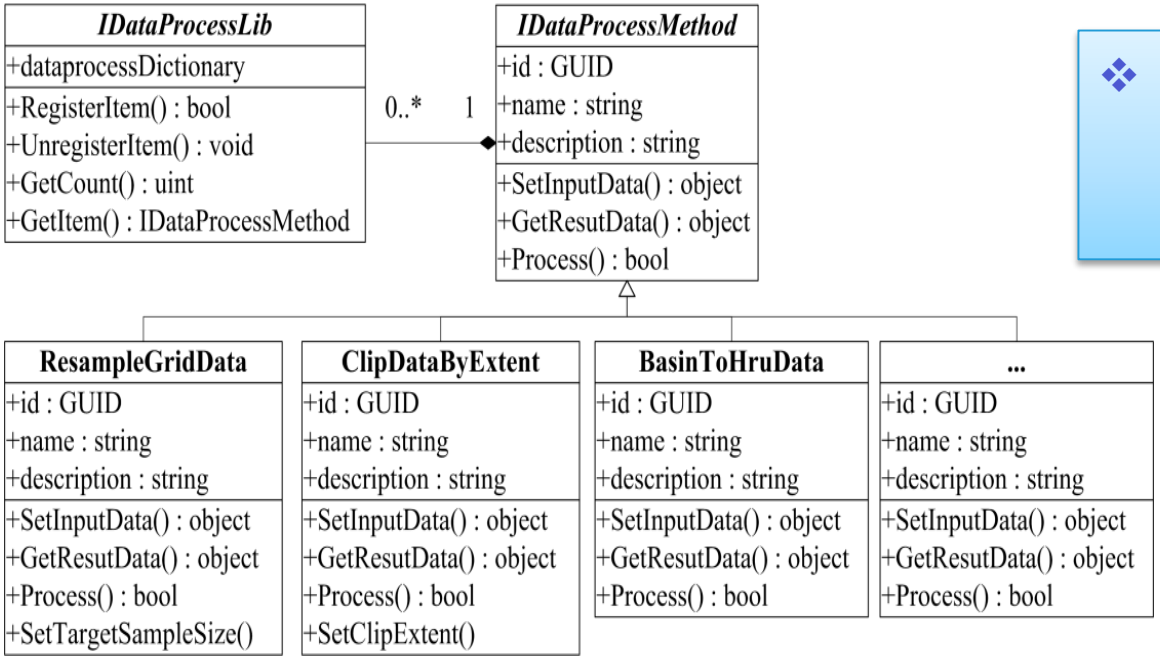
- Name = Curve Grid
  Kernel = *KTStructure*
  - Name = CurveList
    Kernel = *KTList*
    - Name = CurveLine
      Kernel = *KTVector2dList*

- Name = TIN
  Kernel = *KTStructure*
  - Name = Nodes
    Kernel = *KTVector3dList*
  - Name = Triangles
    Kernel = *KTIntList*

- Name = Grid
  Kernel = *KTStructure*
  - Name = Origin
    Kernel = *KTVector2d*
  - Name = XCount
    Kernel = *KTInt*
  - Name = YCount
    Kernel = *KTInt*
  - Name = CellSize
    Kernel = *KTInt*
  - Name = Value
    Kernel = *KTTable*

3. Architecture Design and Experiment

UDX Data Processing Method Library



- An open library, which can be extended by all the geo-analysis modelers

- The interface *IDataProcessMethod* was designed to encapsulate the behavior of actual data process method. And the implemented class of *IDataProcessMethod*, such as *ResampleGridData*, *ClipDataByExtent*, and *BasinToHruData*, etc., can be registered to the *IDataProcessLib*.

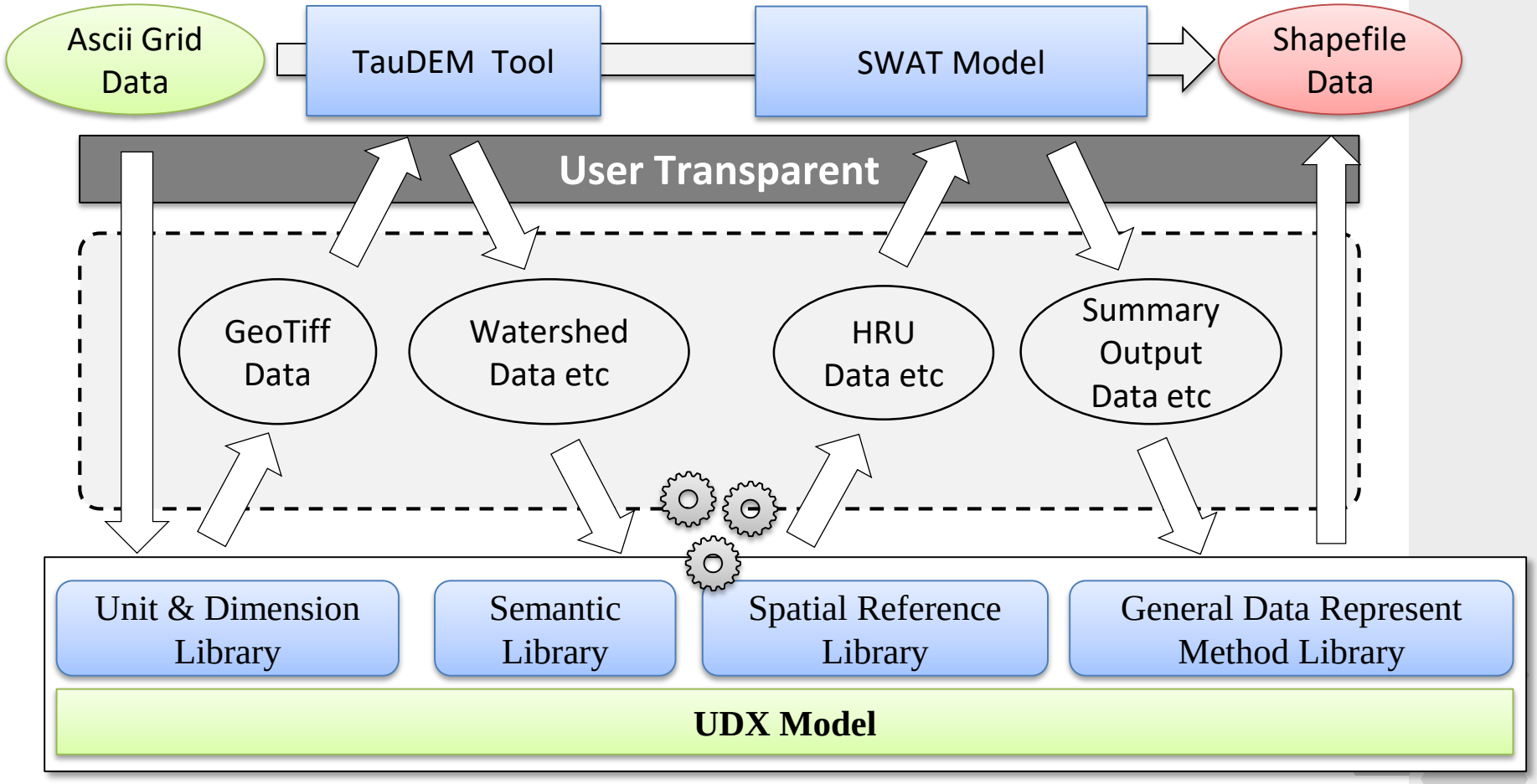
3. Architecture Design and Experiment

Integration of TauDEM and SWAT

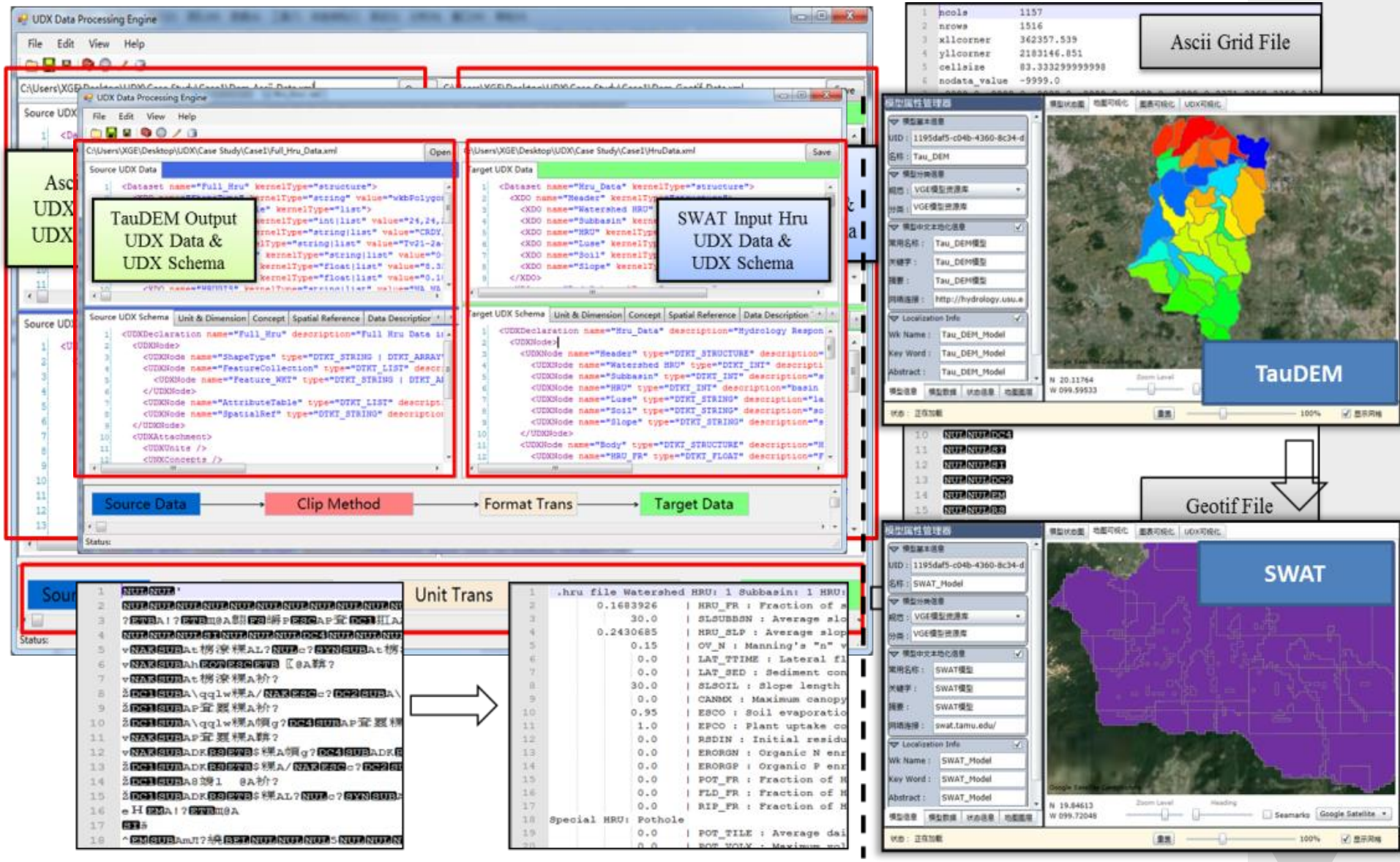
Differences	TauDEM	SWAT
Relevant discipline	Terrain Analysis	Hydrology
Program language	C++	Fortran
Release mode	Source code / executable	Source code / executable
Input data type	Binary File: Geotiff & Shapefile	Plain text File: Hydro-responded unit data (.hru) Basin data (.sub) Ground water data (.gw) Etc.

3. Architecture Design and Experiment

- Integration of TauDEM and SWAT



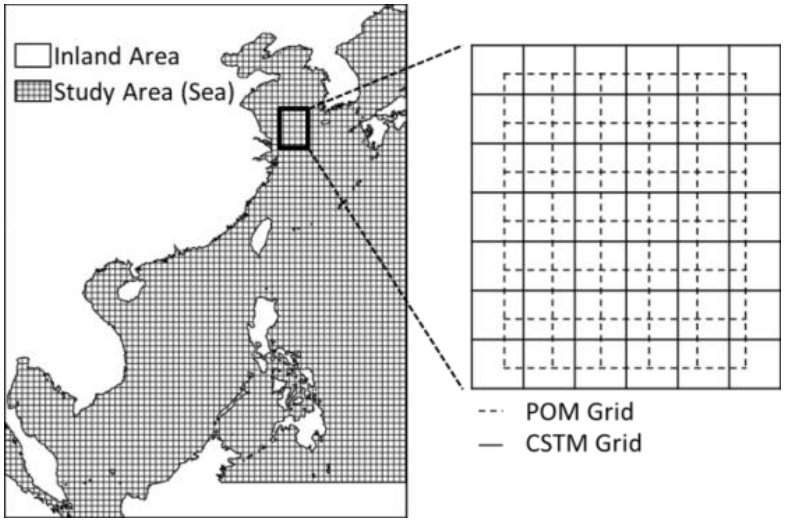
3. Architecture Design and Experiment



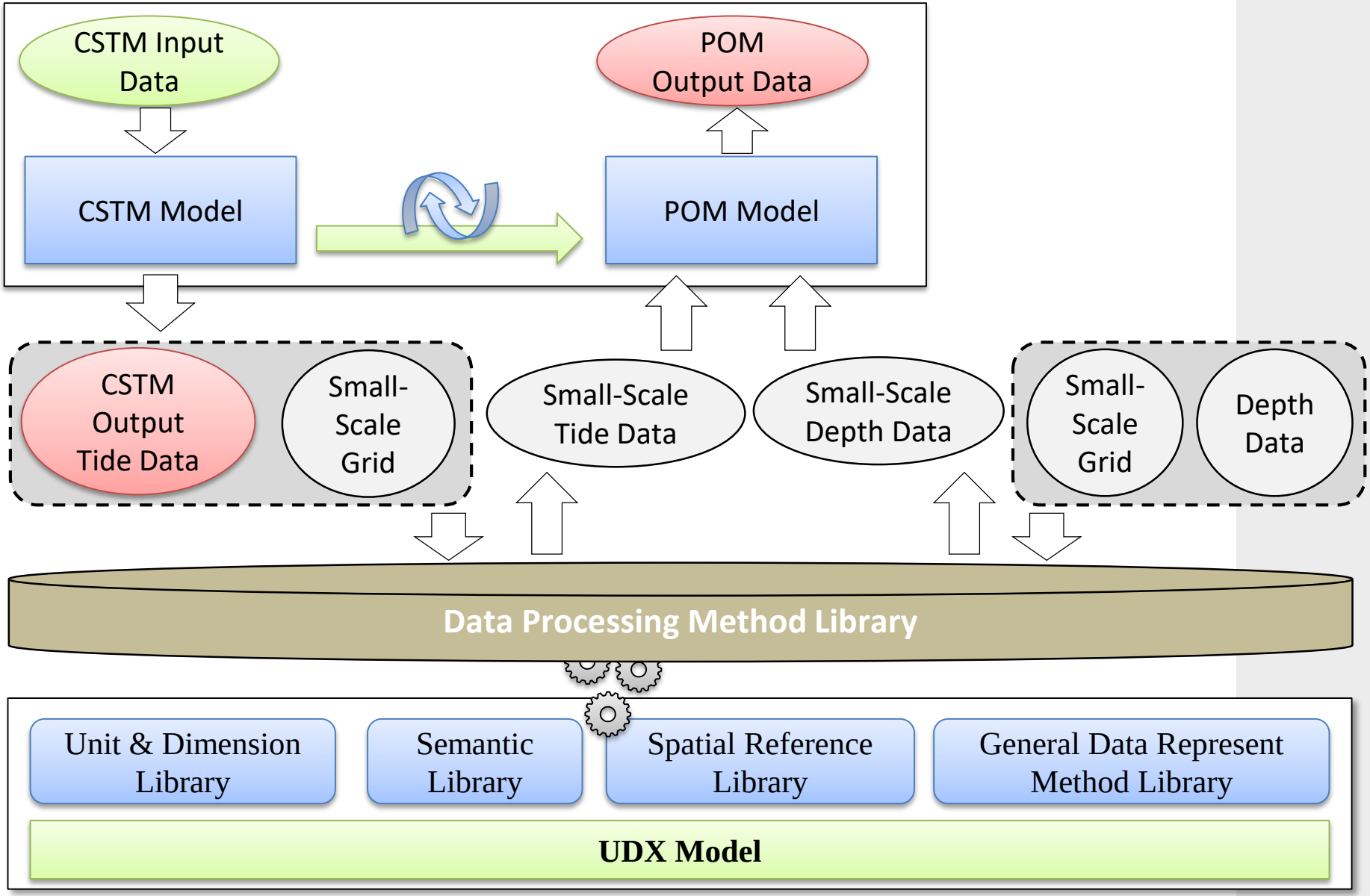
3. Architecture Design and Experiment

❖ Integration of CSTM and POM

Differences	CSTM	POM
Relevant discipline	Ocean tide simulation	Ocean circulation
Program language	Fortran	Fortran
Release mode	Executable	Source code / executable
Spatial dimensions	2D	3D
Input data	Plan text file (split by *tab* char)	Plan text file (split by *space* char)

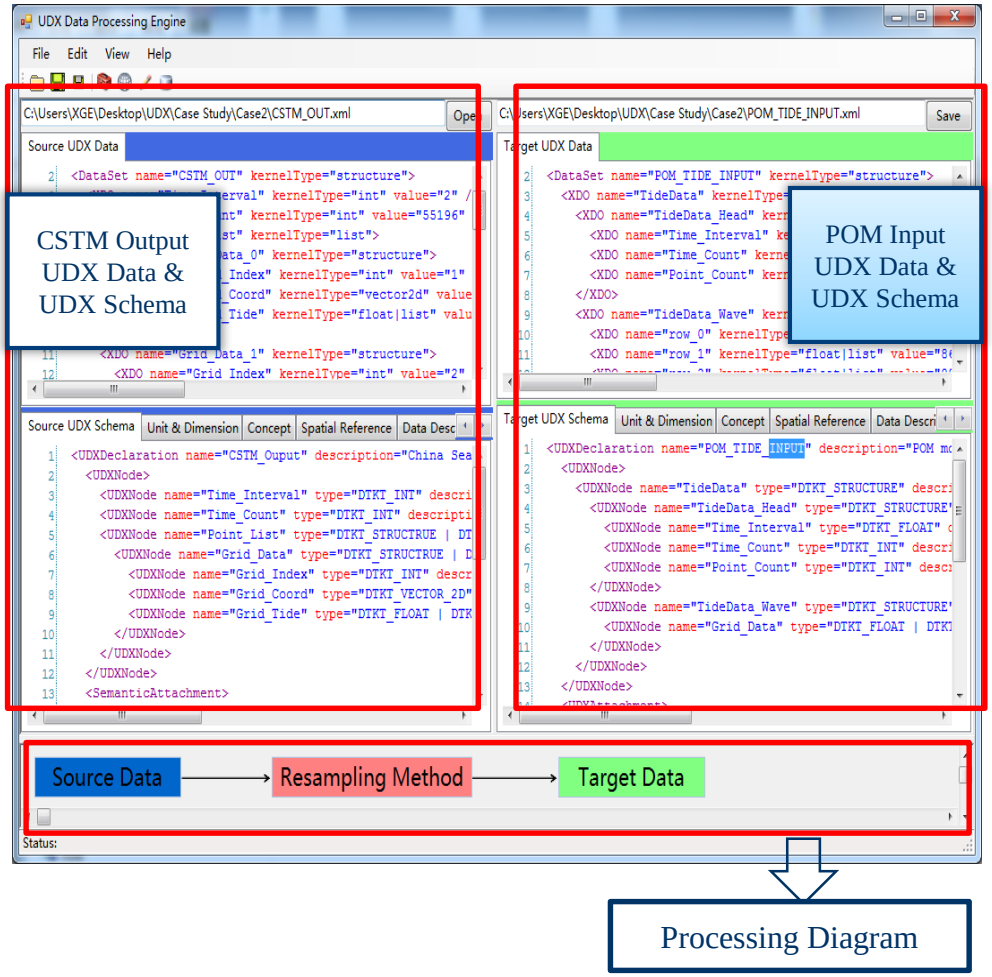


3. Architecture Design and Experiment

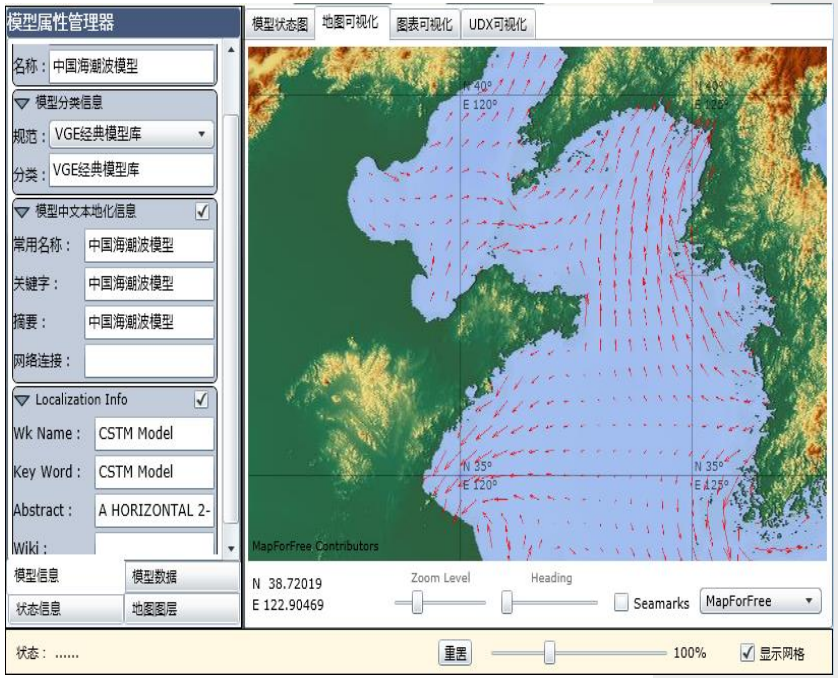


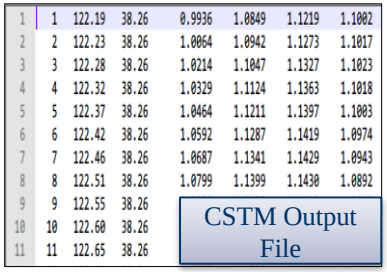
3. Architecture Design and Experiment

❖ Integration of CSTM and POM

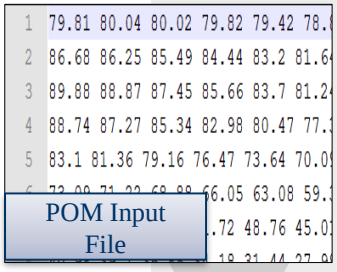


Processing Diagram





CSTM Output File



POM Input File

3. Architecture Design and Experiment

Outline

- Target and Basic Idea
- Data Description Model
- Architecture Design and Practice
- Conclusions

5. Conclusions

- The purpose of this study is to reduce the difficulty of model data preparation and pre-processing work, which is important in reusing, sharing and integrating of geo-analysis models. By taking advantage of the proposed UDX model, model data can be described in a uniform and unambiguous way.

- With the UDX model, the model data processing method library was designed to accumulate and organize various data processing methods, thus making it more convenient for modelers to reuse, share and integrate heterogeneous geo-analysis models.

- However, as geo-modeling and geo-model integration research is synthetic work, future research is needed: Transmission control of UDX Data in a network; Import the semantic reasoning method into the UDX model; The design and development of the UDX data operation script language; etc.

This is an ongoing work

http://geomodeling.njnu.edu.cn

Thank You !