Quantitative Land Cover Classification from MODIS - A case study in Northeast China

WANG Zhengxing

Global Change Information and Research Center
Institute of Geographical Sciences and Natural Resources Research (NGSNRR)
Chinese Academy of Sciences

E-mail: wangzx@igsnrr.ac.cn

20th CODATA Conference, Oct., 2006，Beijing

OUTLINE

1、Introduction

2、Study area: NE China

3、Input data: MODIS-NDVI product

4、Output info: Land cover

5、Process: Decision Tree

6、Conclusion

1、Introduction: Repeatability - **Quantitative**

（1）What is Repeatability？

Given same input data and same target classification system, the same methodology should yield same (comparable) results even used by different investigators.

（2） Repeatability is essential in change detection or environmental monitoring. Yet has hardly been achieved in remote sensing application

（3）Three examples

Repeatability – What if not ?

Example[1] – How high is Mt. Everest ?

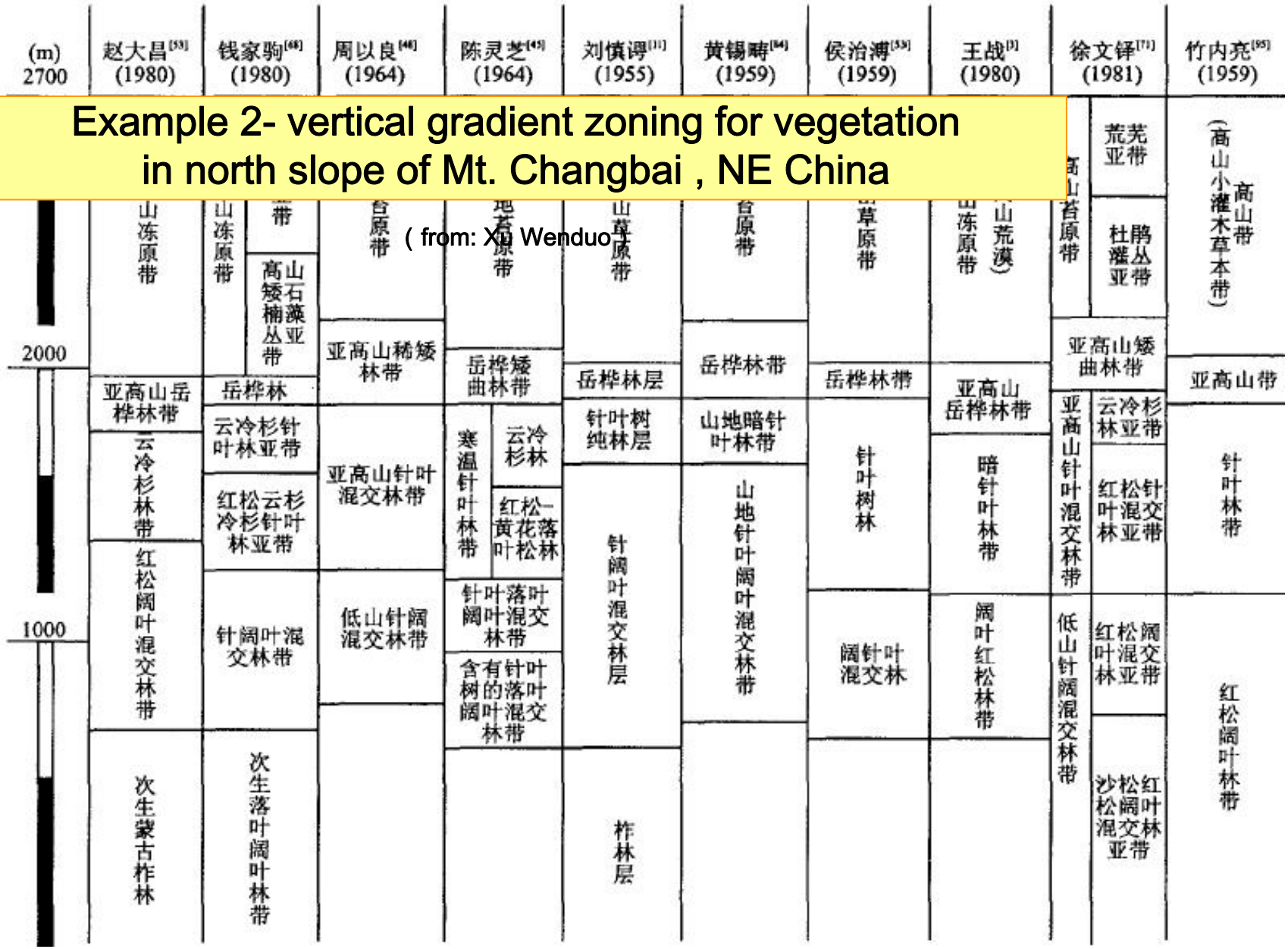
In 1975： 8848.13 meters

In 2005： 8844.43 meters

Has it shrunk 3.7meters in last 30 years ？

Not all people is so luck to know all the stories behind these DATA !

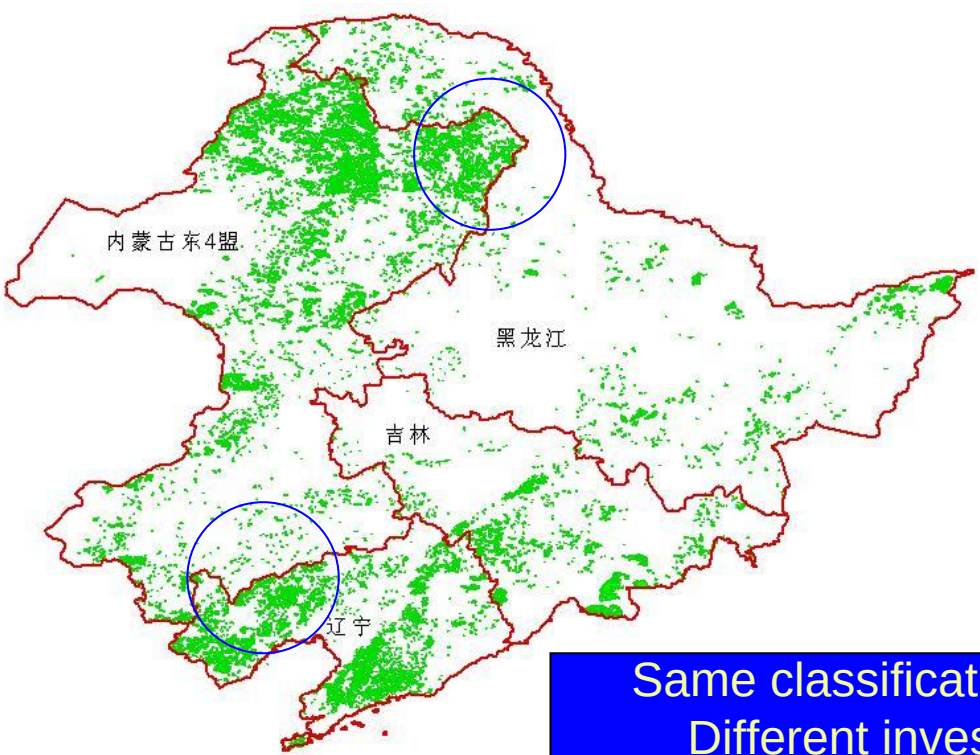
Example 2- vertical gradient zoning for vegetation in north slope of Mt. Changbai , NE China

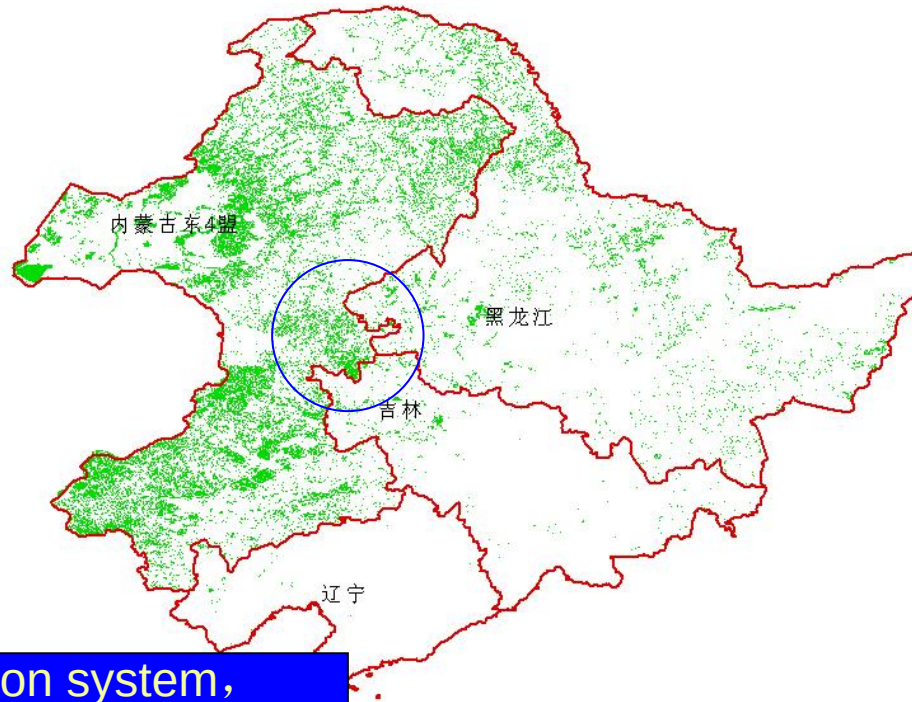


(from: Xu Wenduo)

Example 3. land use map of China, 1 : 100,000

(left : sparse forest , Right : dense grass)

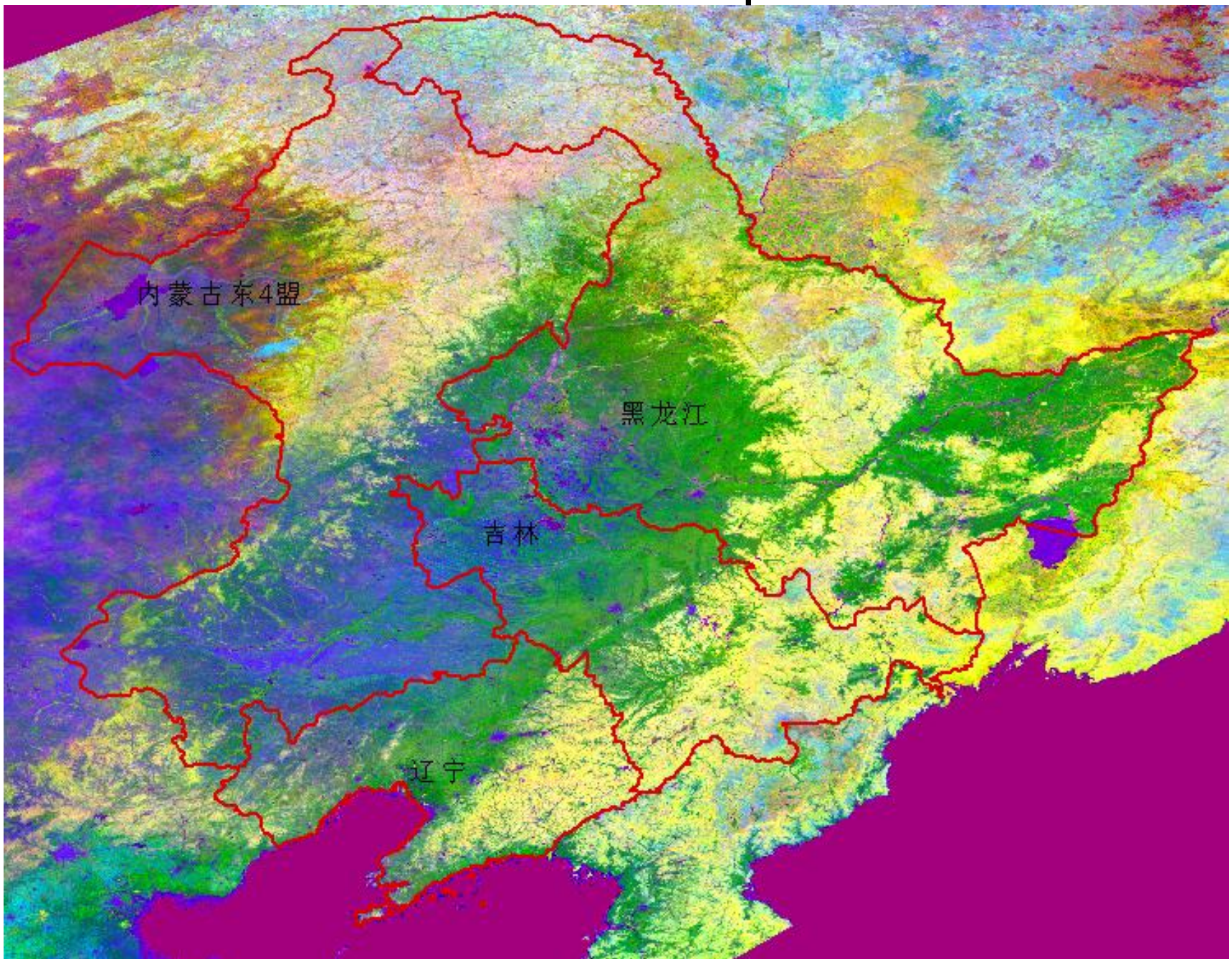




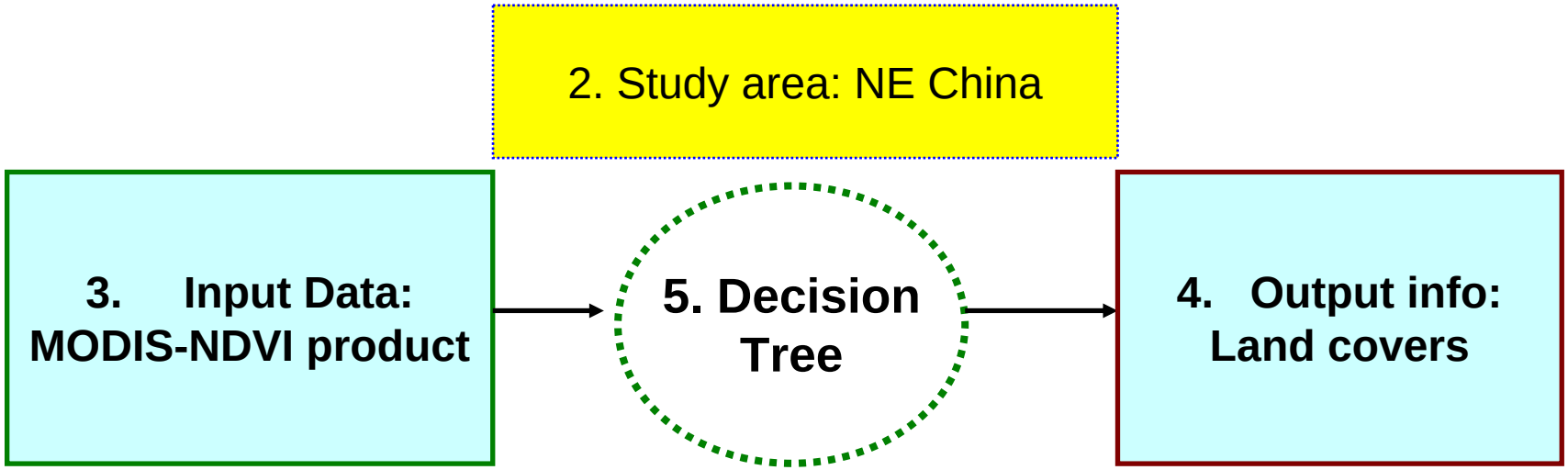
Same classification system，
Different investigators，
Sharp change across working border.

Across Border between
Inner-Mongolia and Hei-long-jiang

MODIS: smooth across borders between provinces

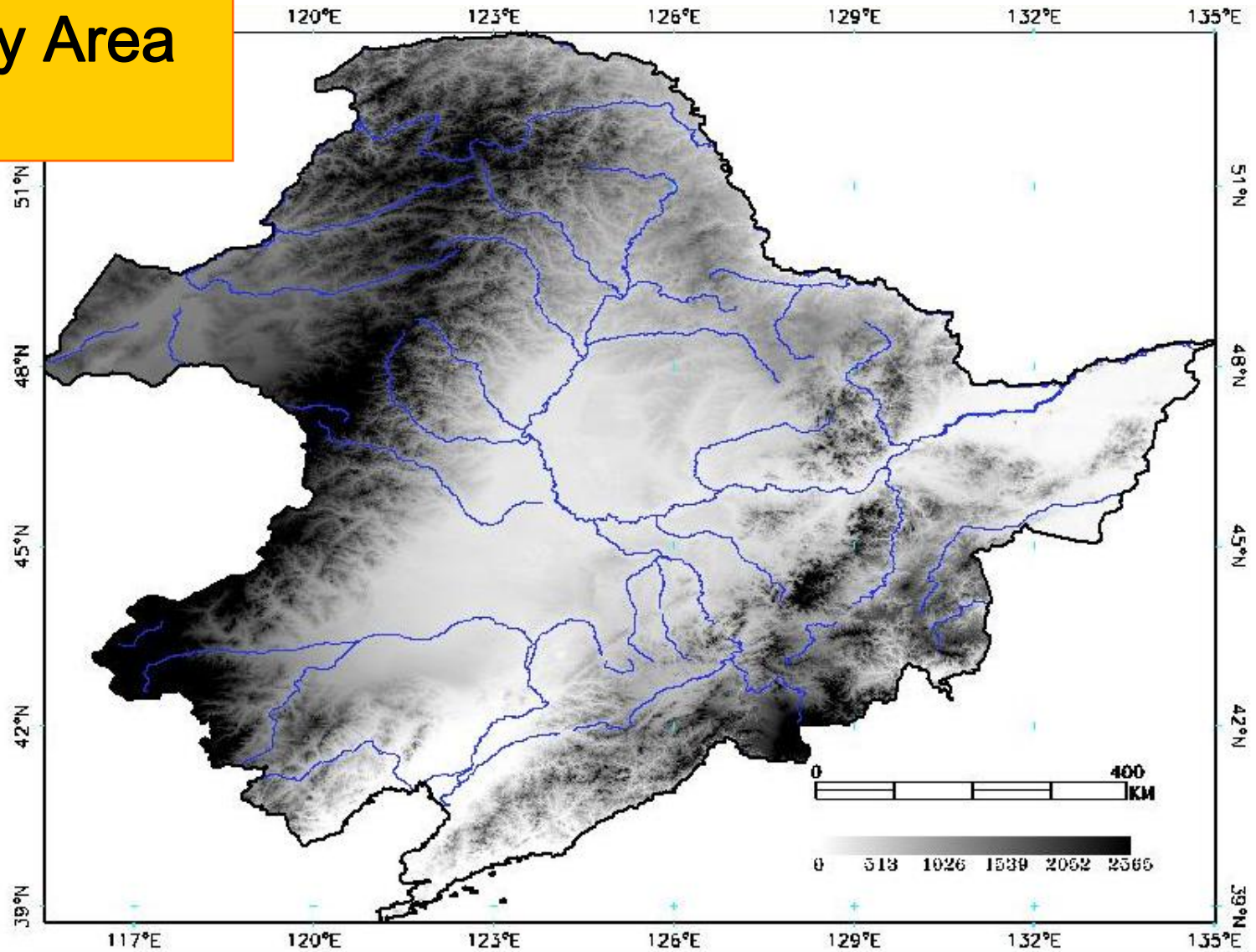


maximize Repeatability in land cover classification



Comparable input;
quantitative target information (output);
transparent process

2、Study Area



NE China:Land cover classification system
Adapted from IGBP 17-class system

3、INPUT DATA: MODIS VI product

- MODIS-NDVI Product: 11 layers
- Noise detection: 4 approaches
- Pre-processing: 3 approaches
- Data Reliability after processing

MODIS-NDVI Product: From NASA LP-DAAC,1km 16day composites ,11 layers,

- **NDVI, EVI, NDVI-QA, EVI-QA,**
- **blue, red, NIR,MIR,**
- **3 geometry parameters,**
- **Derived: NDWI**

$$NDVI = (B2 - B1) / (B2 + B1) \quad (1)$$

$$EVI = 2.5 * (B2 - B1) / (B2 + 6 * B1 - 7.5 * B3 + 1) \quad (2)$$

NDVI : Normalized Difference Vegetation Index

EVI : Enhanced Difference Vegetation Index

B1 : $MODIS_1$ (red: 620-670nm)

B2 : $MODIS_2$ (NIR 841-876nm)

B3 : $MODIS_3$ (blue: 459-479nm)

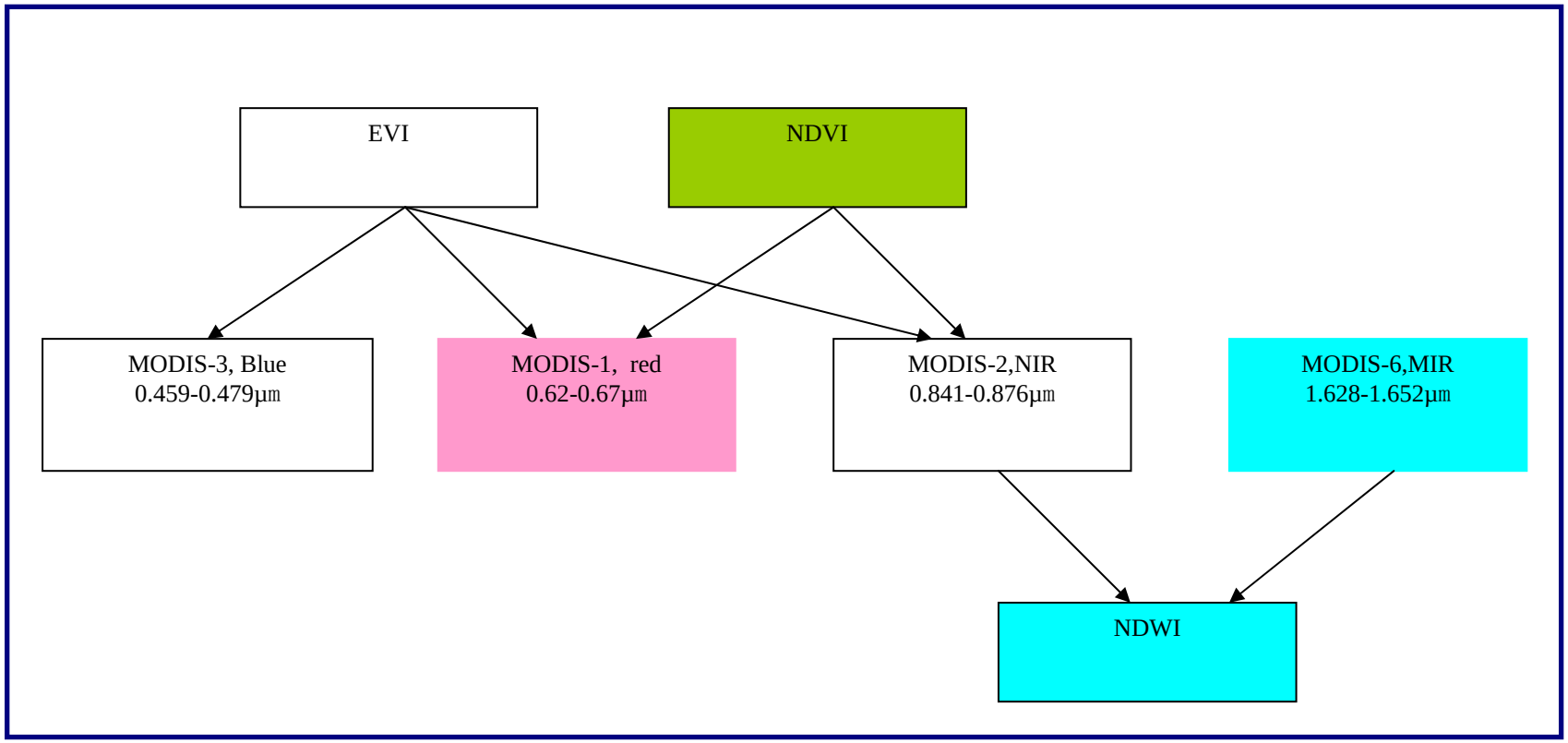
$$NDWI = (B2 - B6) / (B2 + B6) \quad (3)$$

NDWI: Normalized Difference Water Index

B2 : $MODIS_2$ (NIR 841-876nm)

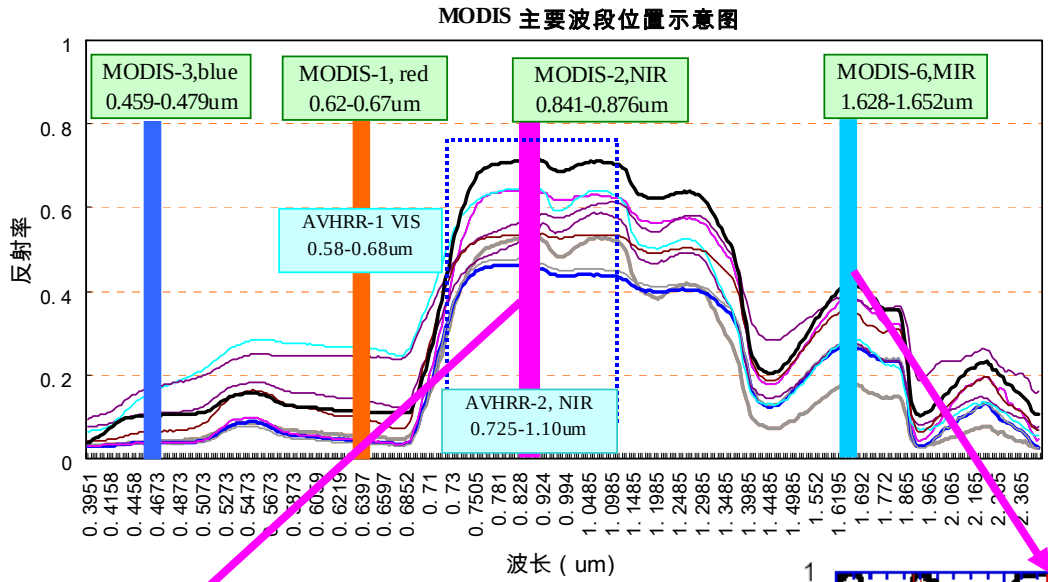
B6 : $MODIS_6$ (MIR 1626-1652nm)

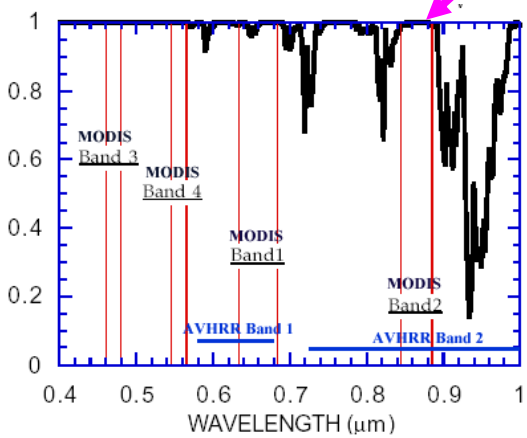
Relationships of NDVI and original bands

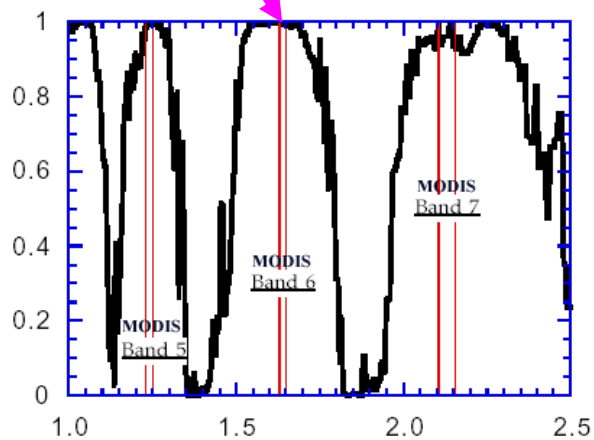


The quality of MODIS has been greatly improved

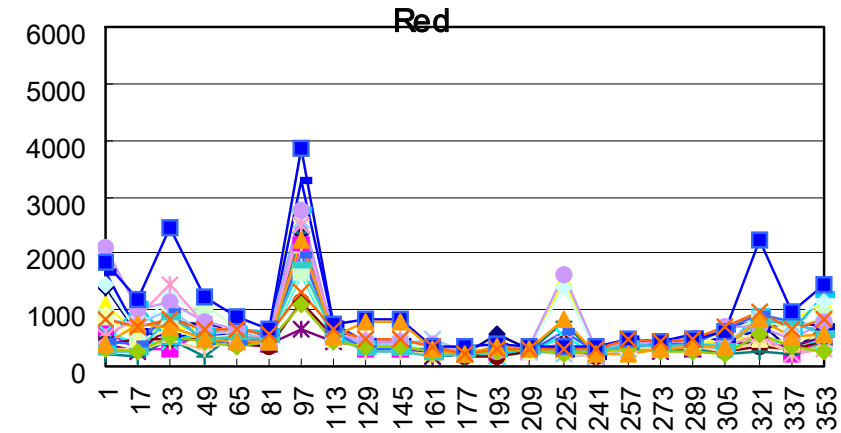
MODIS vs AVHRR: Narrower bands avoid some noises (moisture absorption at 0.95um, 1.4um)

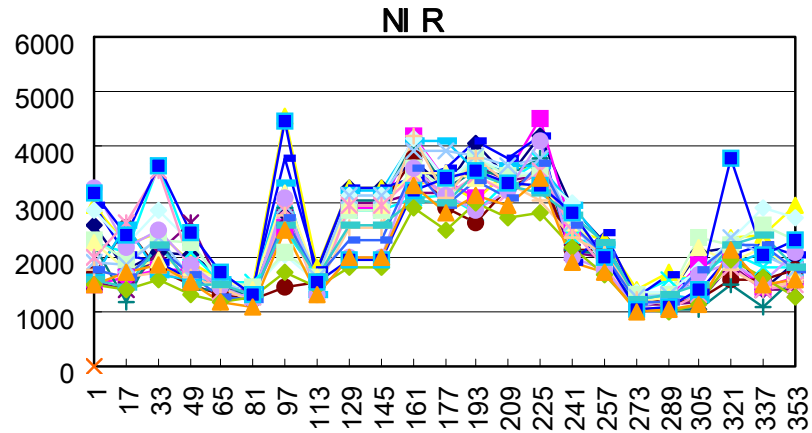


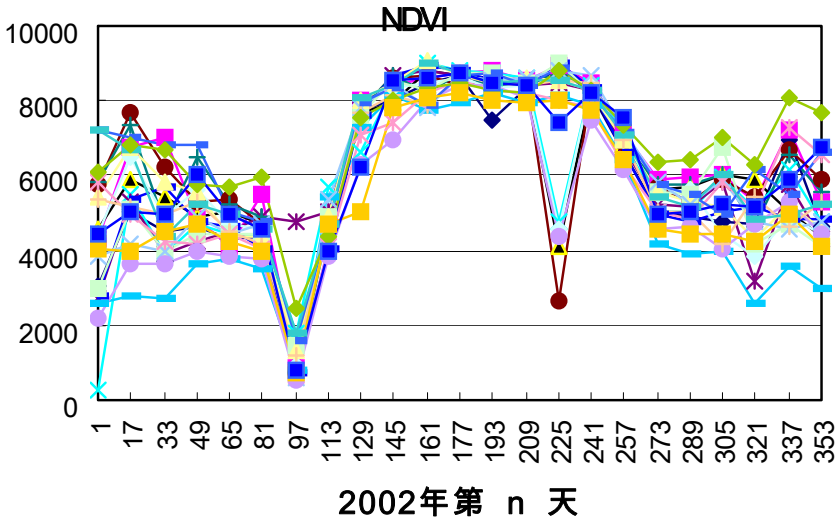


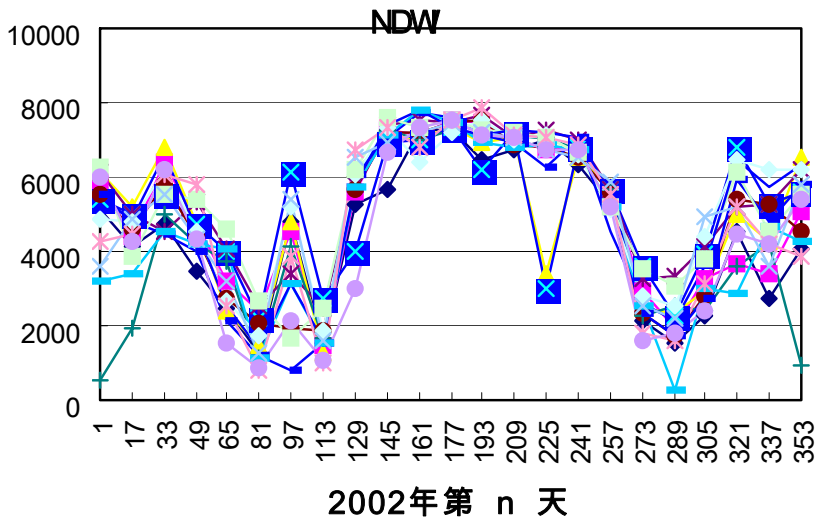


yet still include noises

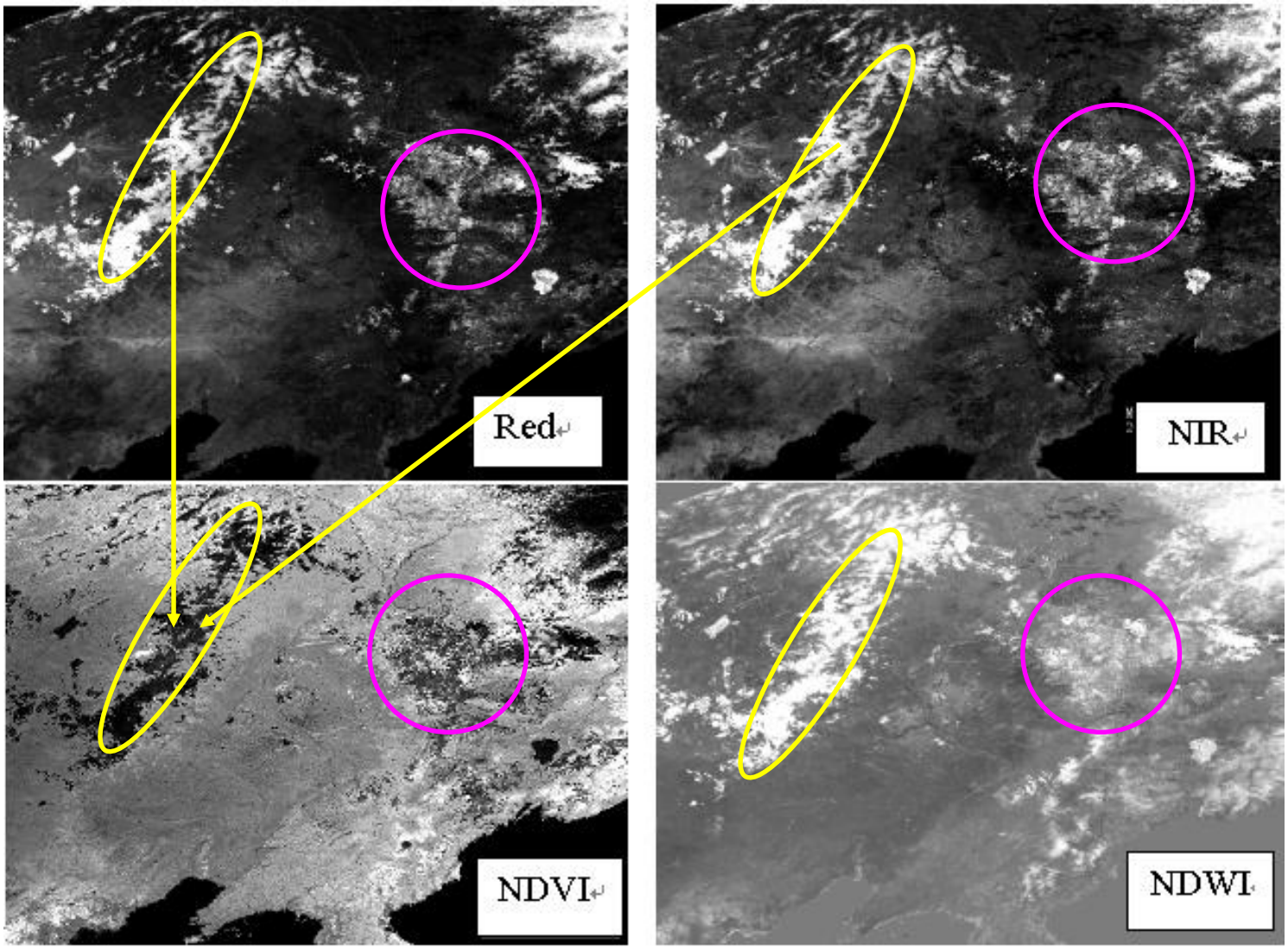








A case for noise detection： 7-22, April, 2002, day097



Major Sources of Noises：Cloud, fog, snow, ice

- **Cloud, fog: anywhere, anytime, yet more frequent in forest in rain season. The signal：**
  - **red、NIR、MIR ↑**
  - **NDVI, EVI ↓**
  - **normally last within 16-32 days, except in Mt. Changbai forest**

- **Snow, ice: spatial and temporal pattern, more occurrence in North and Higher region**
  - **Red, NIR ↑ (> 0.6)**
  - **NDVI, EVI ↓ (<=0)**
  - **MIR ↓ , NDWI ↑ (>=0.6)**

Noise detection: 4 means

（1）Quality Assessment（**QA**）：only detectable (cannot be corrected)

（2）Temporal：as from previous page, with water a exception

（3）Spatial：

- less reliable than temporal, eg，small forest/water;
- Fog and snow may distributed in large region，space filters does not work well

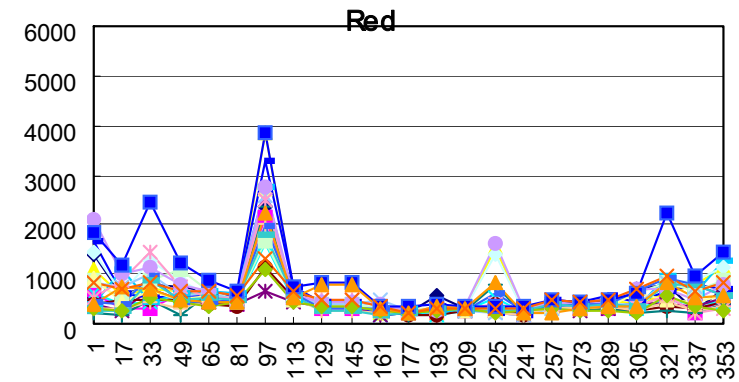
（4）Consistency of indices： Signals may be closely related

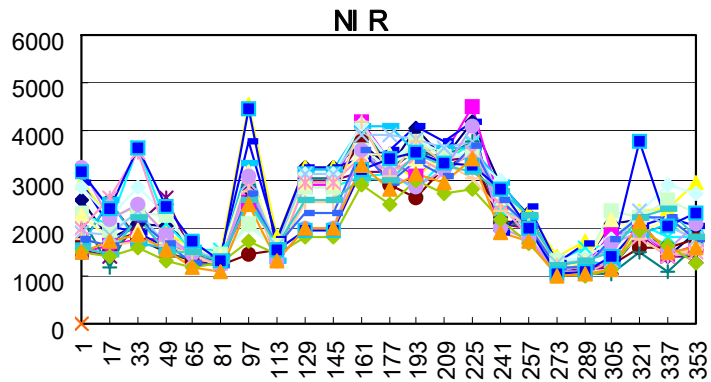
- Snow = red ↑ + MIR ↓
- Cloud = NIR ↑ + NDVI ↓

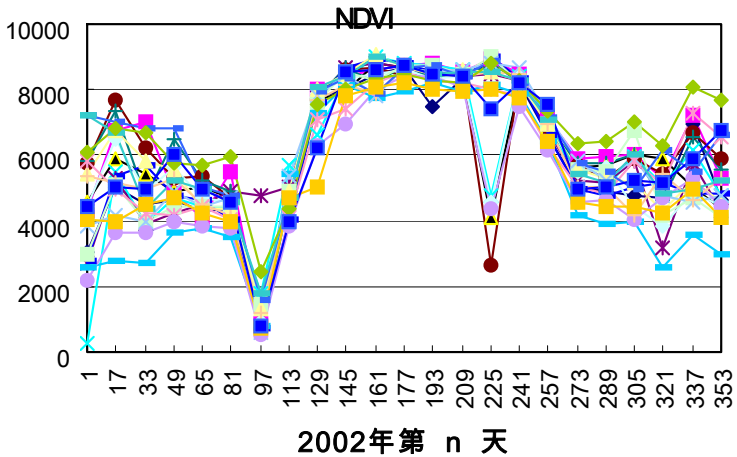
Relationships of indices for various noises：

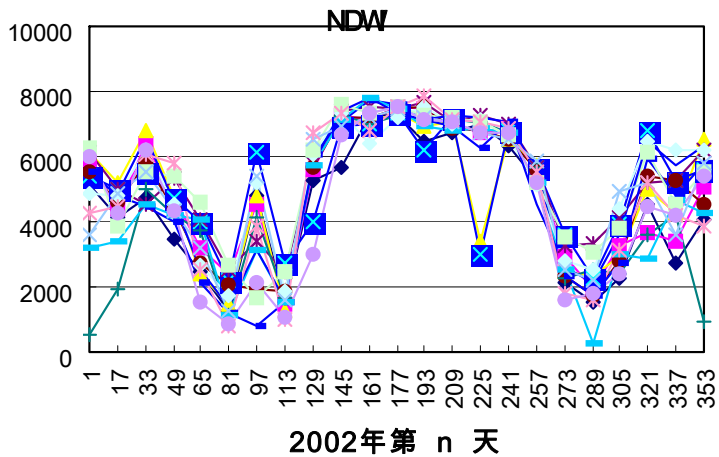
CLOUD and FOG - red ↑, NIR ↑, NDVI ↓ （Day97, 225, 321）；

SNOW and ICE - red ↑, NIR ↑, MIR ↓, NDWI ↑ （Day97）





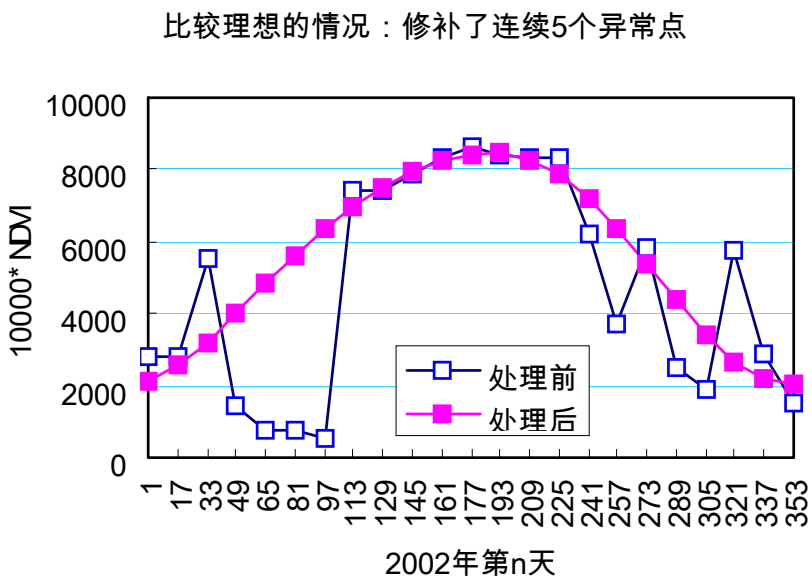


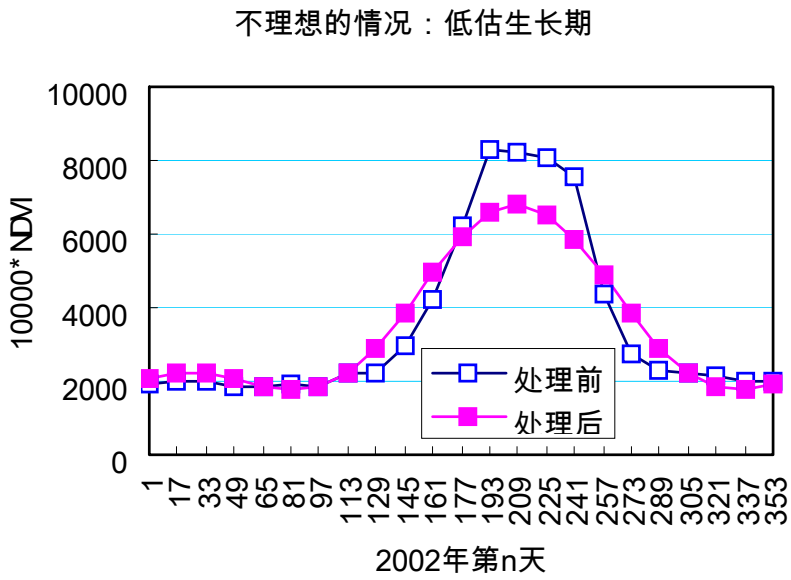


Pre-processing：3 approaches

- Harmonic ANalysis of time Serials, HANTS
- Conditional Smoothing
- some specific means for too noisy regions

Pre-processing (1) - Harmonic ANalysis of time Serils, HANTS





To large extent, the success of HANTs depends on the selection of parameters, which is quite subjective;

Preprocessing (2) - Conditional Temporal Smoothing

[1] NDVI /EVI Time Serials: growing season (Day97-273)

Growing season :

- Dormant season: smooth for natural, expected change for cropland
- With noise: replace corrupt pixel with (average of) neighbor pixels;
- For continuous noises: iterative operations

Dormant season : still exist some problems

- For mixed noise from snow and fog: Max (NDVI)。
- An exception: water

[2] Red Time Serials:

Dormant season: soil, snow, evergreen

Growing season: smooth for all peak noises

[3]NIR Time Serials:

- On condition of lower-NDVI
- replace higher-NIR with (average of) neighbors

[4]MIR Time Serials:

- Dormant season:“凹” = noise
- Growing season:“凸” = noise
- Transitional period (Day097-113) : independent operation
- Only for growing season in this study

[5]NDWI Time Serials:

- Before smooth, NDWI time serials are smoother than NDVI;

Pre-processing(3): tailored means for too noisy regions

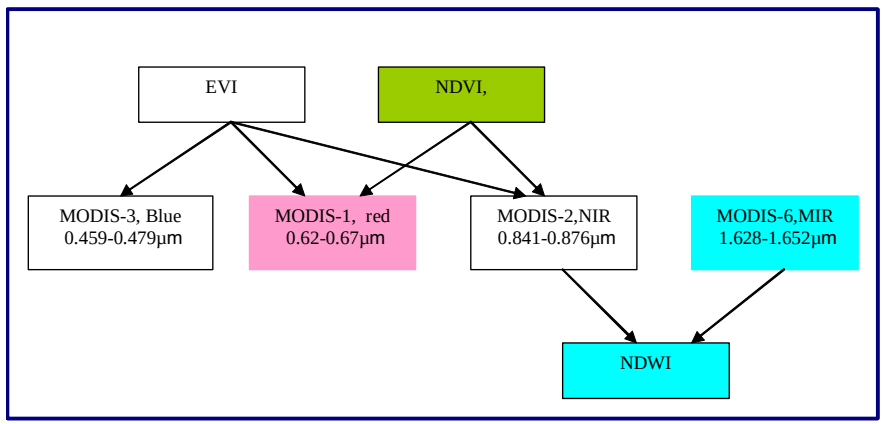
- Spatial: forest of Mt Changbai
- Temporal: growing season (rainy season)
- Operation: subset spatially and temporally, for HANTs or manual operation

Reliability After Pre-processing：NDVI (EVI) > NDWI > Red > NIR > Blue

- **NDVI: most informative for vegetation, reliable**

- **NDWI: from** MIR and NIR.
  - Two advantages of MIR：less sensitive to moisture in atmosphere
  - Sensitive to land surface moisture。
  - NDWI, like NDVI, more comparable than MIR.

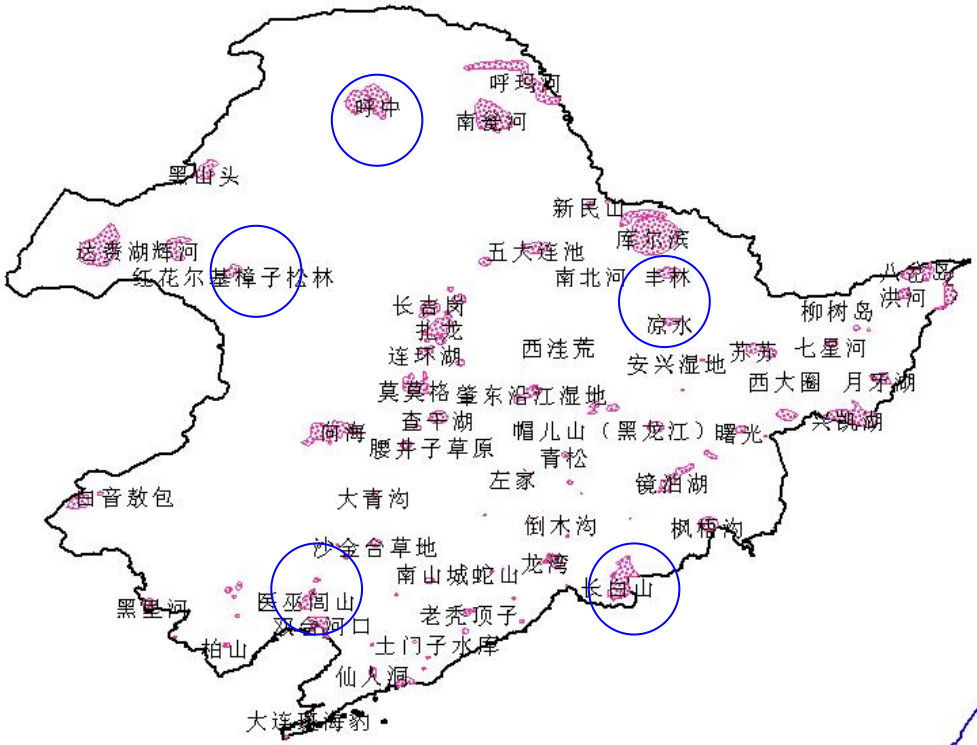
- **Blue,red,NIR: consistent with NDVI, but less reliable than NDVI**



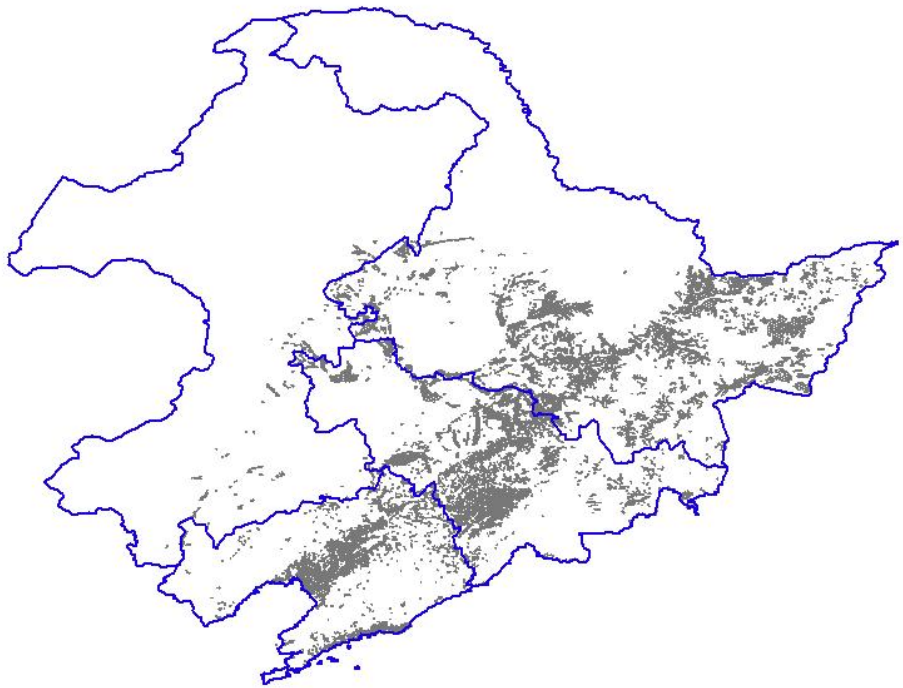
4、Quantitative Description of Land Cover Types

(1)From MODIS-NDVI time serials
(2)From MODIS-NDWI time serials
(3)Other data,
(4)Used as decision rules

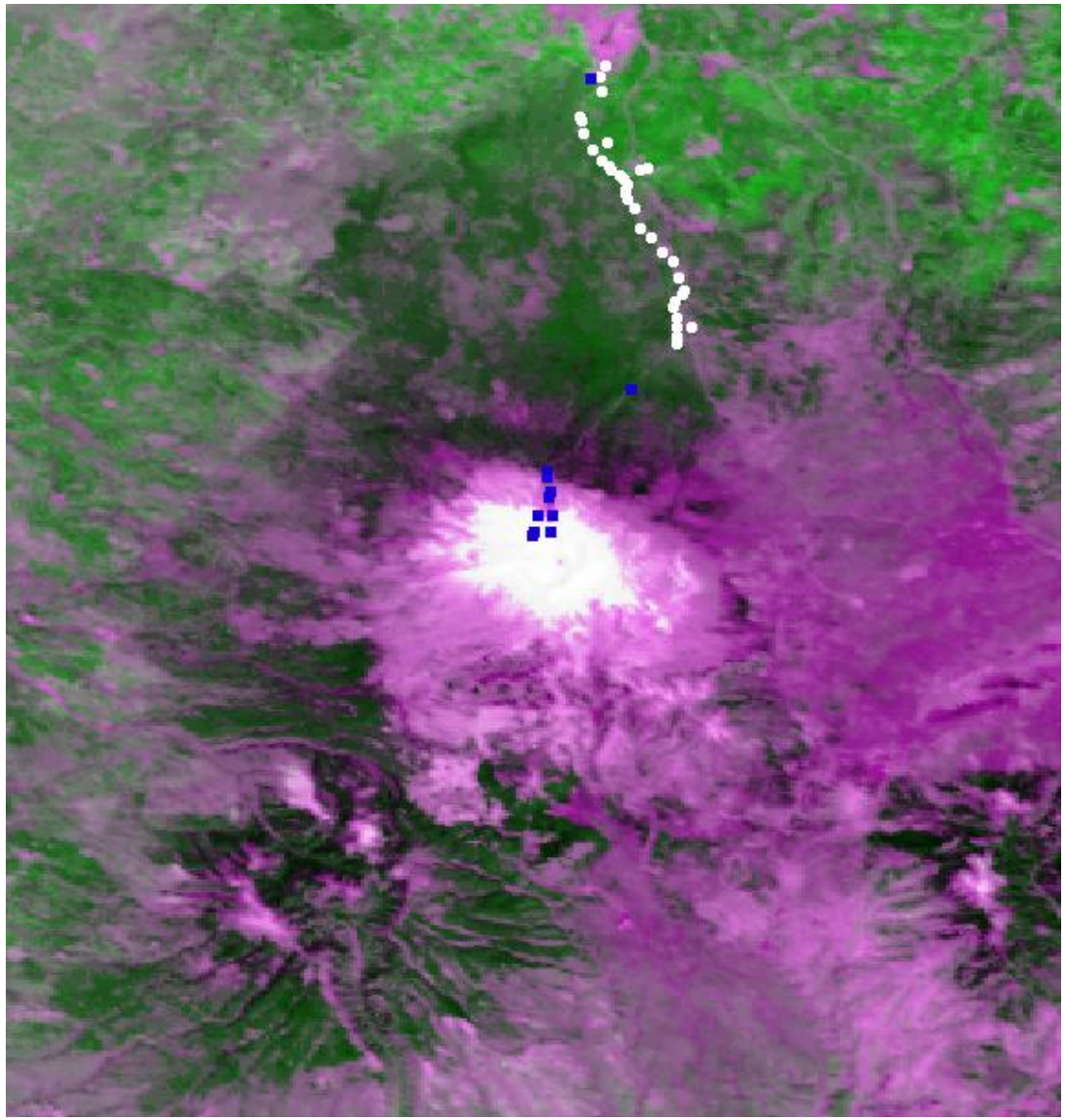
Auxiliary: Natural reserves



Landuse Map, 1 : 100000
IGSNRR, CAS



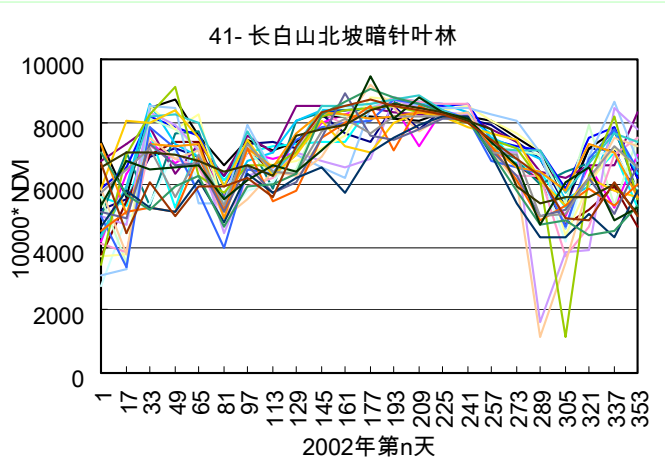
Field sample：Mt. Changbai, National Forest Reserve

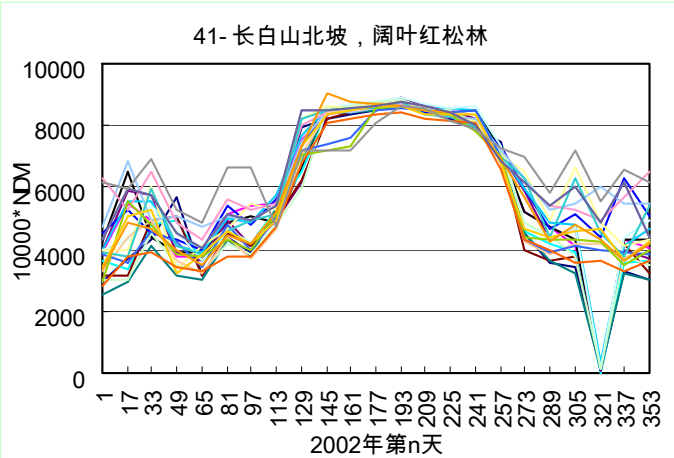


(1) From MODIS-NDVI, Site = Mt. Changbai

Needle leaf – evergreen (left)

Mixture of evergreen and deciduous(right)





Needle ever-green forest: (average)

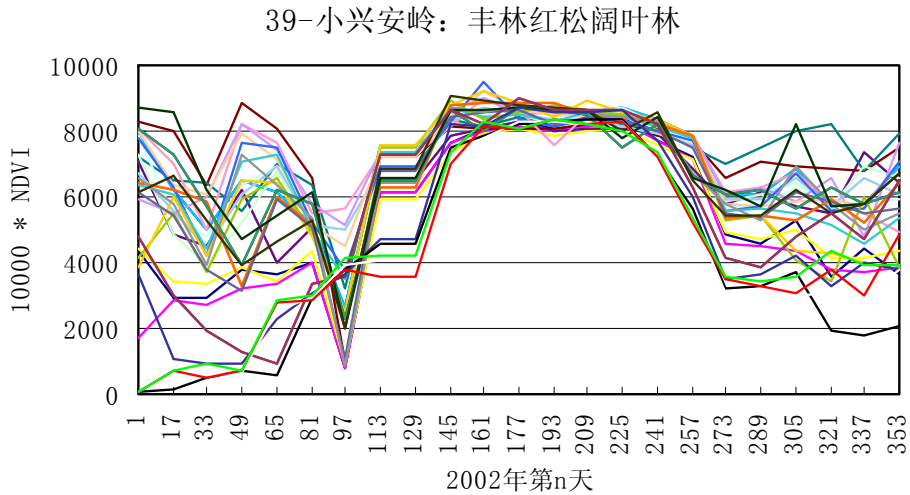
- Full year: NDVI > 0.75
- Growing season: NDVI > 0.80
- Dormancy: NDVI > 0.60

Needle leaf -Evergreen ： Mt. Changbai, evergreen/all = 28/30

表 4-1 苔藓岳桦云冷杉林的立木因子（郁闭度: 0.93）（王战）

树种	林层	组成		胸径(cm)		树高(m)		年龄	每公顷	
		按株数	按材积	最大	平均	最大	平均		株数	材积
鱼鳞云杉	I	10	10	84	50	30	28	160	640	389.0
冷杉	I	+	+		16		25		16	8.0
落叶松	I	+	+		20		25		16	11.0
云杉	II	6	6		25		17	120	96	46.0
冷杉	II	3	3		20		16	80	48	31.0
岳桦	II	1	1		20		12		16	4.0
杂木	II	+	+		20		12		8	2.0
云杉	III	5	6		12		11	60	260	24.0
冷杉	III	4	3		12		11	60	196	11.0
岳桦	III	+	+		12		10		32	1.0
杂木	III	1	1		12		10		72	4.0
合计					30.8		19.6		1400	531.0

Mixture of Needle leaf–evergreen and deciduous ：
Site 2 = Mt. Small Xing'an (Fenglin reserve)



15 forest populations, components of the largest five:

云冷杉红松林： 6 红松-2冷杉- 1云杉-1枫- 1 其它。郁闭度0.5-0.8。

枫桦红松林： 6 红松-3枫树-1冷杉- 1云杉、椴、白桦。

针阔混交林： 2 红松-2冷杉- 1云杉-1落叶松 -3枫- 1白桦。

云冷杉林： 4 云杉-4冷杉- 1 红松-1白桦。

冷杉林： 8 冷杉-2云杉

针叶混交林： 2 红松-3冷杉- 2云杉-3 其它。

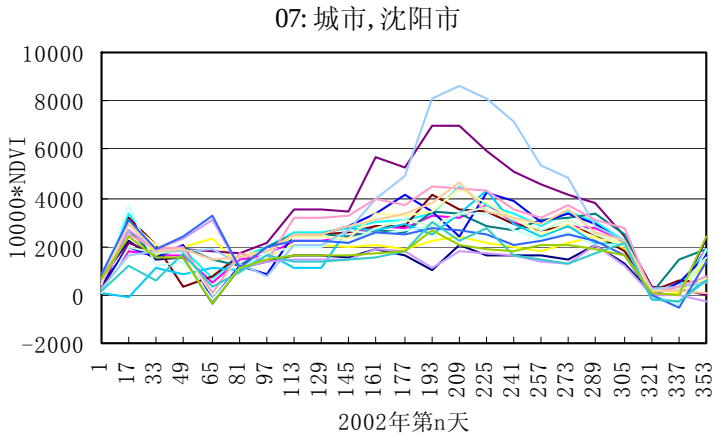
Mixed forest: evergreen and deciduous

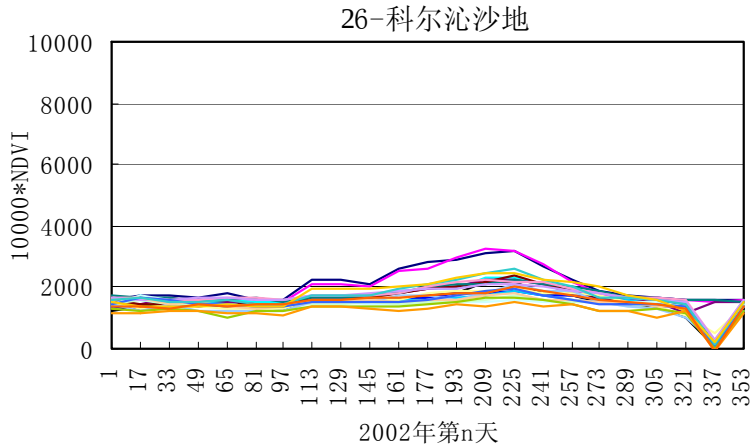
- In study area, there is a continuous distribution of various forest. How to classifying them is a matter of definition. There is the place where subjectivity occurs.

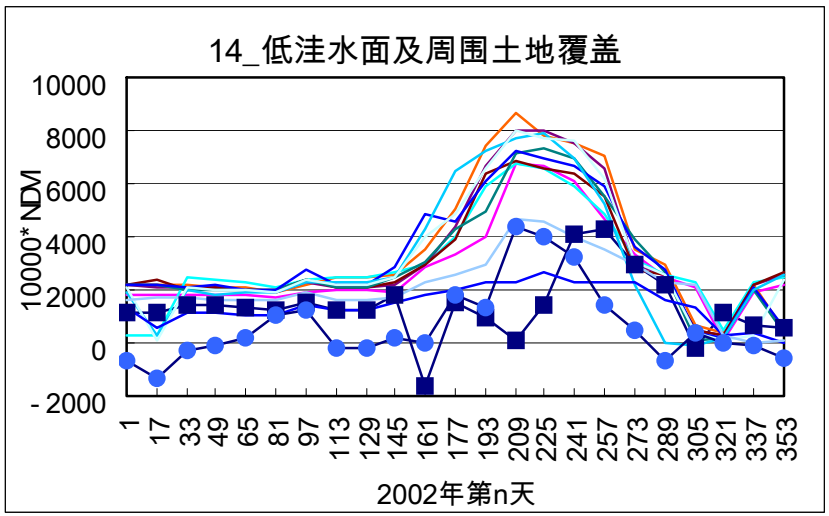
- At 740m, the fraction of evergreen trees in mixed forest is 40%，this has been correctly described by NDVI time serials. During growing season, NDVI could be about 0.85; during dormancy，NDVI=0.4-0.5. Given the fact of baseline during dormancy (NDVI=0.25)，$NDVI_{winter}$ = 0.4-0.5 is a good indicator of fraction of evergreen in winter.

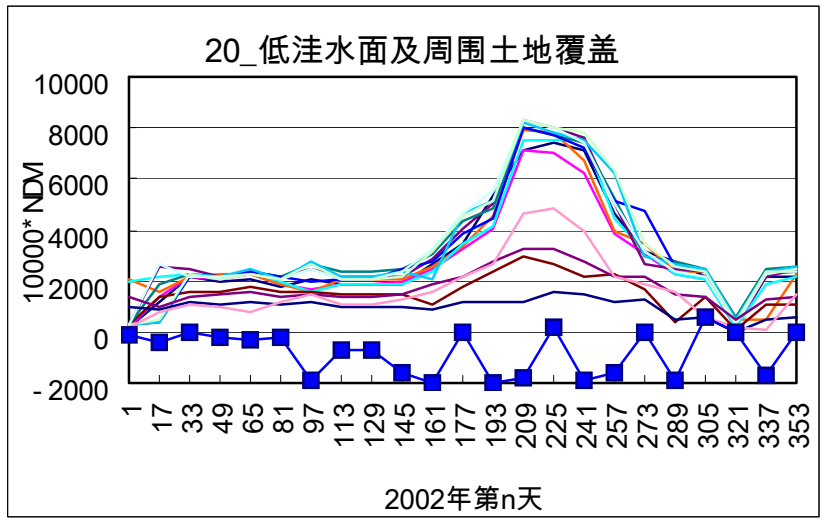
- At 1800m, the fraction of evergreen in the purest forest is as high as 90%，and its NDVI time serial has a high level of 0.60-0.85 in whole year.

Problem of NDVI： City；Sand；Water Body

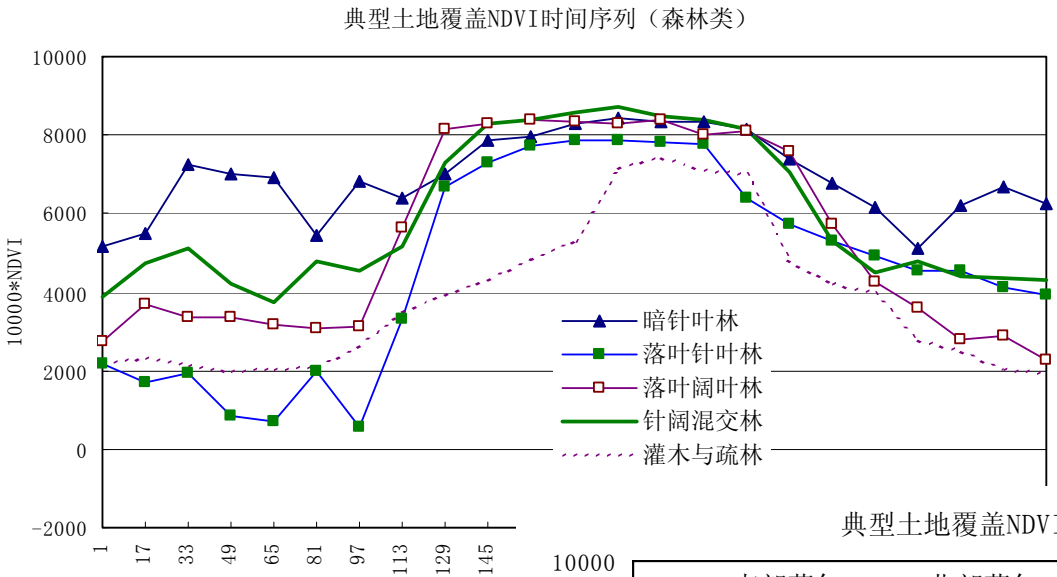


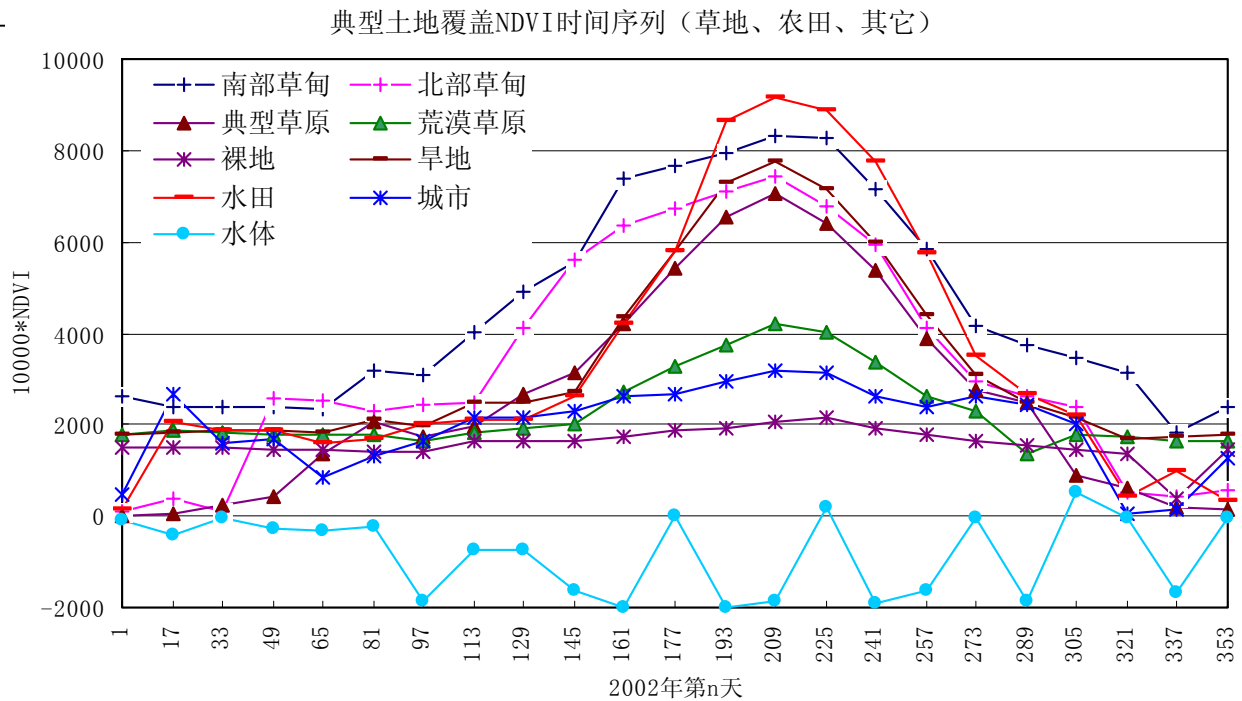






NDVI wrap-up


典型土地覆盖NDVI时间序列（森林类）
10000*NDVI
暗针叶林
落叶针叶林
落叶阔叶林
针阔混交林
灌木与疏林


典型土地覆盖NDVI时间序列（草地、农田、其它）
10000*NDVI
南部草甸
北部草甸
典型草原
荒漠草原
裸地
旱地
水田
城市
水体
2002年第n天

Average NDVI of Major Land Covers （10000*NDVI）

Landcover	Full Year	Growing Season	Dormancy Season	Green-up Season
	Day001-353	**Day113-273**	**Day001-097, 289-353**	Day113-145
Needle leaf-Evergreen	6929	7716	6208	7088
Needle leaf-Deciduous	4600	6704	2671	5756
Broad leaf-Deciduous	5354	7717	3187	7353
Nixed forest	5965	7615	4453	6918
Bush and sparse forest	3836	5422	2383	3924
Meadow grassland (South)	4531	6478	2747	4827
Meadow grassland (North)	3326	5423	1404	4066
Typical grassland	2597	4494	858	2587
Sparse grassland	2291	2909	1725	1913
Bare land / sand	1586	1825	1367	1657
Rain-fed cropland	3331	4864	1926	2557
Irrigated cropland	3413	5504	1496	2280
City and town	1955	2623	1343	2200
Water body	-807	-1117	-524	-1032

(2) Additional information From MODIS-NDWI: Irrigated vs. Non-irrigated

(1) Before Day145, all croplands have the similar signature as soil, NDVI < 0.25-0.3； during Day145-209, NDVI can be 0.7-0.99;

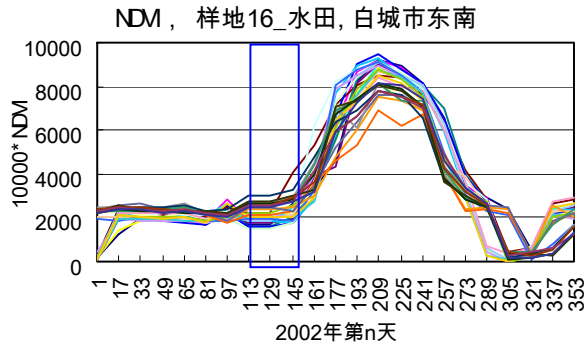
(2) The NDVI difference between Irrigated and non-irrigated crop land exists, but not reliable for classification

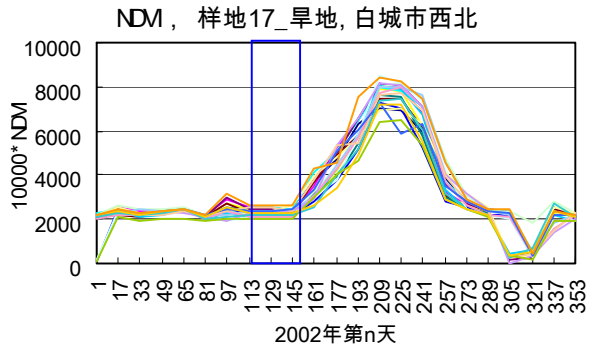
(3) The signature from WATER can be best described with NDWI.

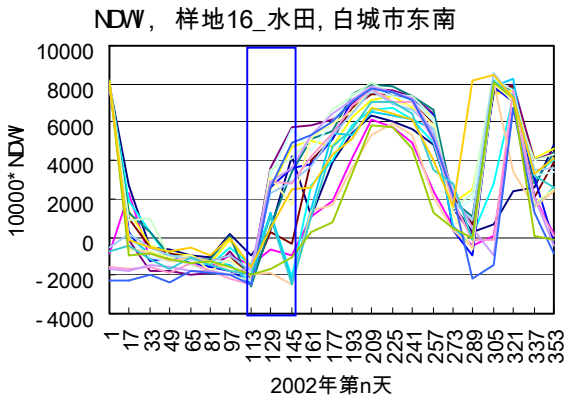
Irrigated (left) vs Rain-fed (right)

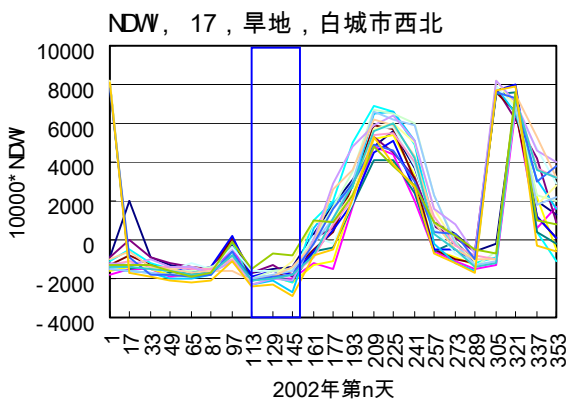
NDVI

NDWI

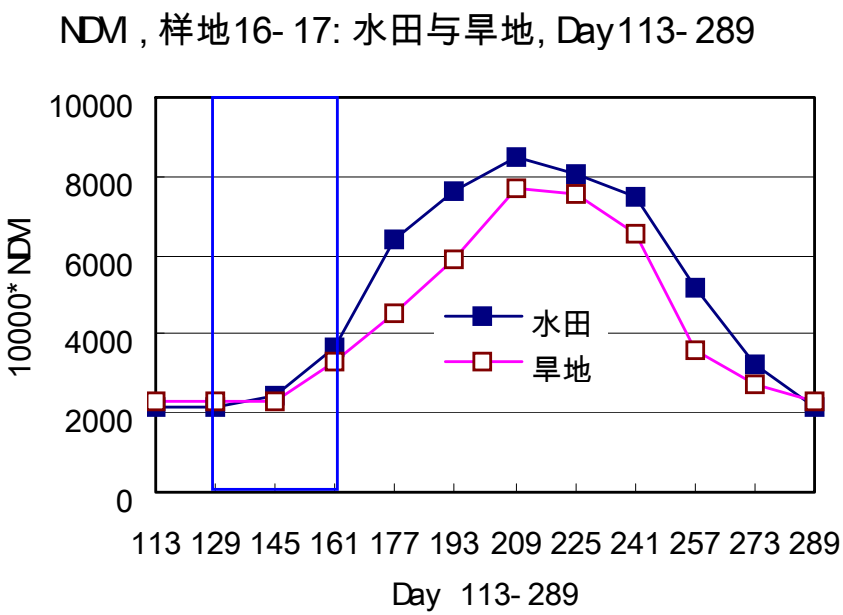


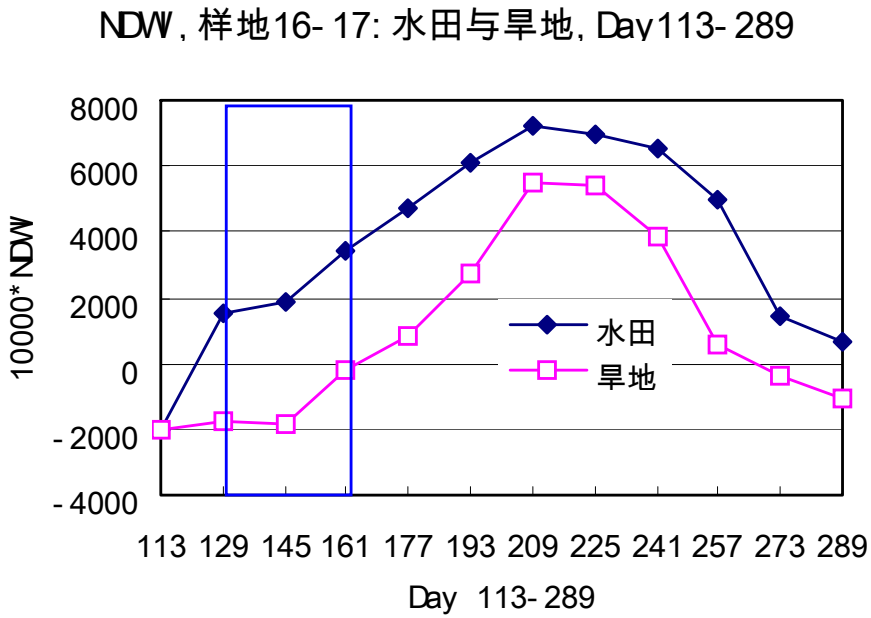






Most sensitive period of NDWI: $Day_{129\text{-}161}$

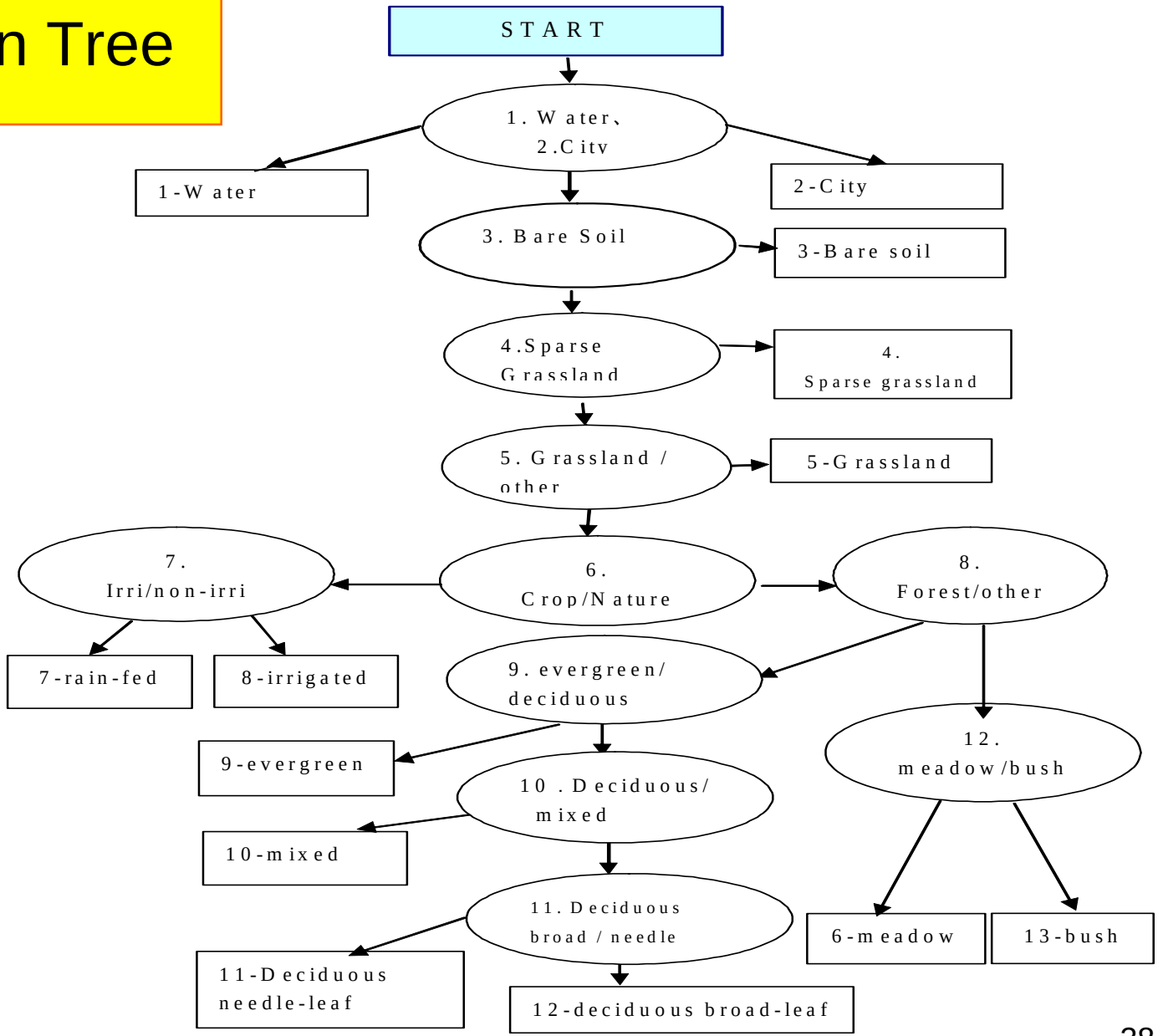




More information from other indices, DEM

…

5.Decision Tree



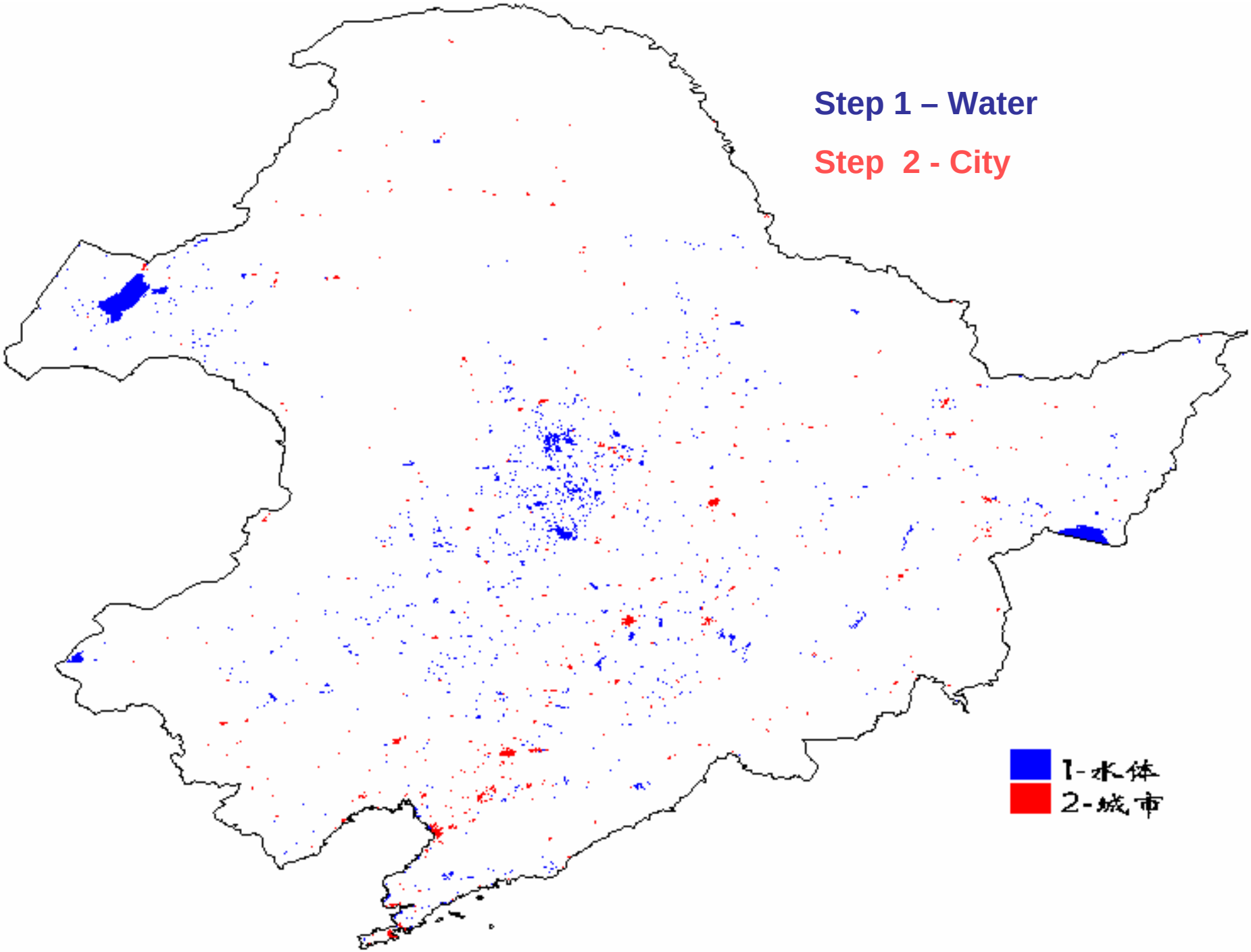
(2)、Node and Decision Rules (Indices)

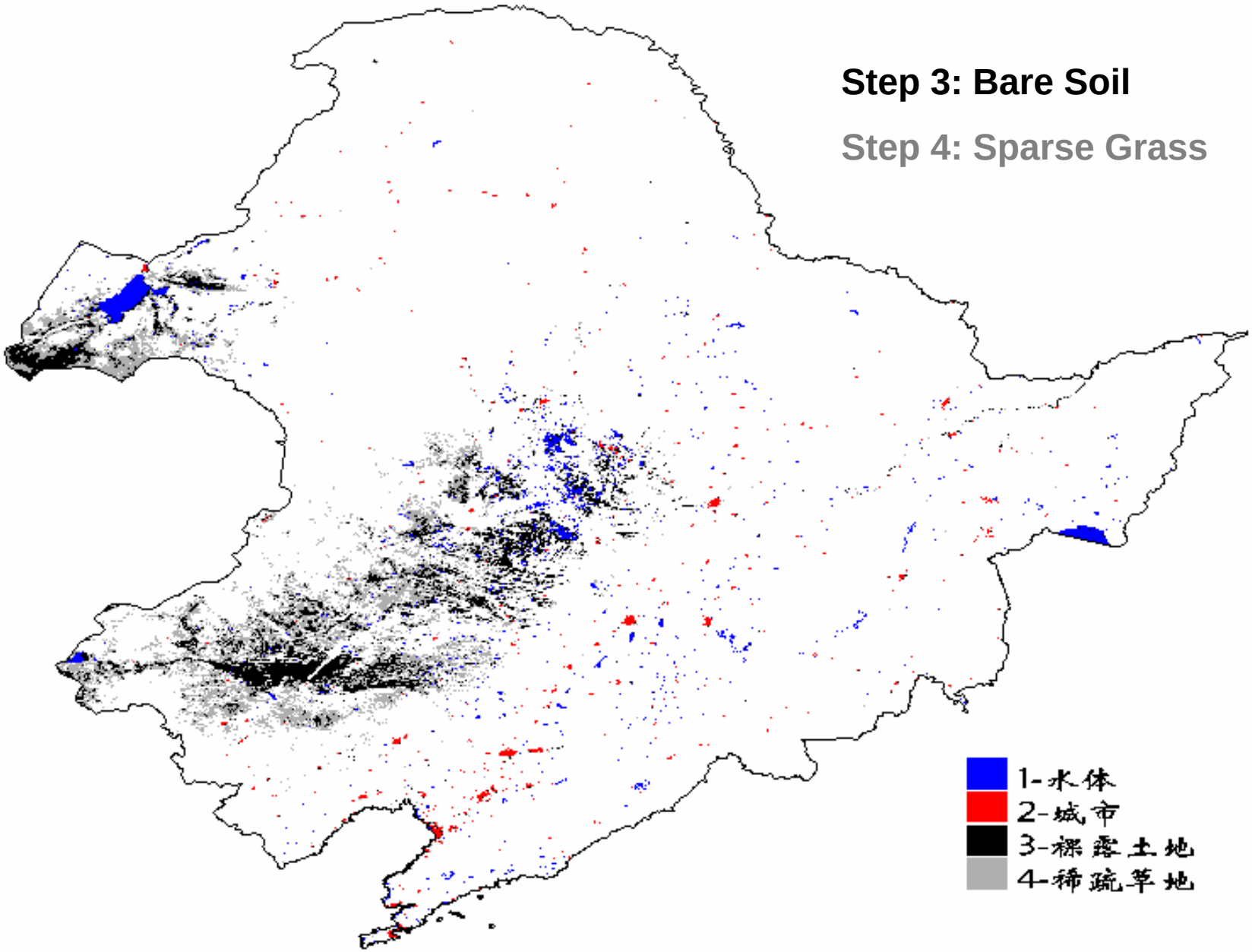
Step 1-2：Water and City, from 1:100,000 landuse map, IGSNRR,CAS.
Step 3-5：growing season ($_{day113\text{-}273}$)
If NDVI $_{day113\text{-}273}$ < 0.30 then ‘bare soil’
If NDVI $_{day113\text{-}273}$ = 0.30-0.35 then ‘Sparce grass’
If NDVI $_{day113\text{-}273}$ = 0.35-0.45 then ‘Typical grass’

Step 6：If $NDVI_{day145}$ ≤ 0.4 then ‘crop land’
Step 7：If $NDWI_{day129\text{-}145sum}$ > 0.05 and DEM<500 then ‘irrigated’

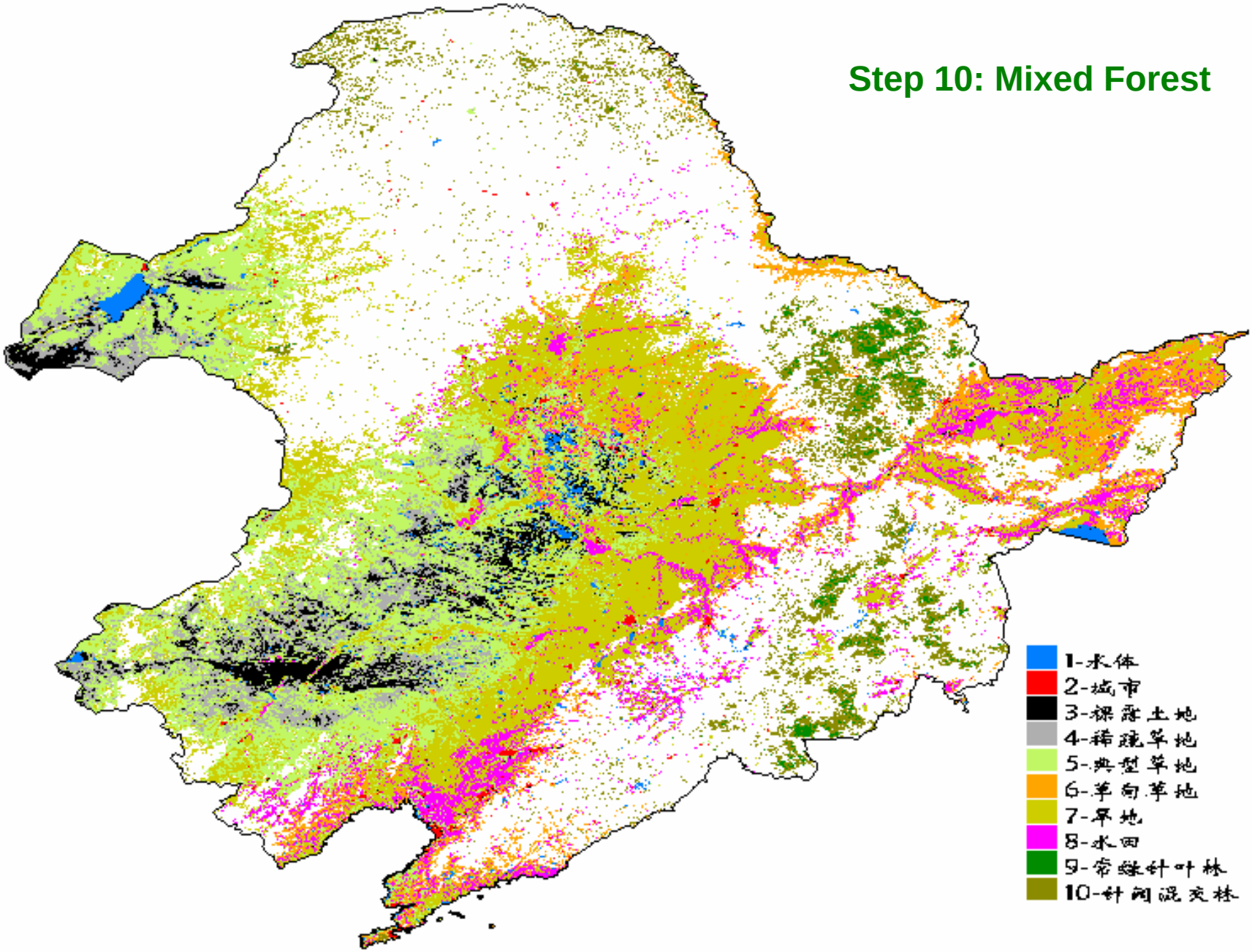
Step 8： If $NDVI_{day113\text{-}273}$>0.65 then ‘forest’
Step 9： If $NDVI_{dormancy}$ >0.53 then ‘forest-evergreen’
Step 10： If $NDVI_{dormancy}$ > =0.38-0.53 then ‘forest-mixed’ else ‘forest-deciduous’

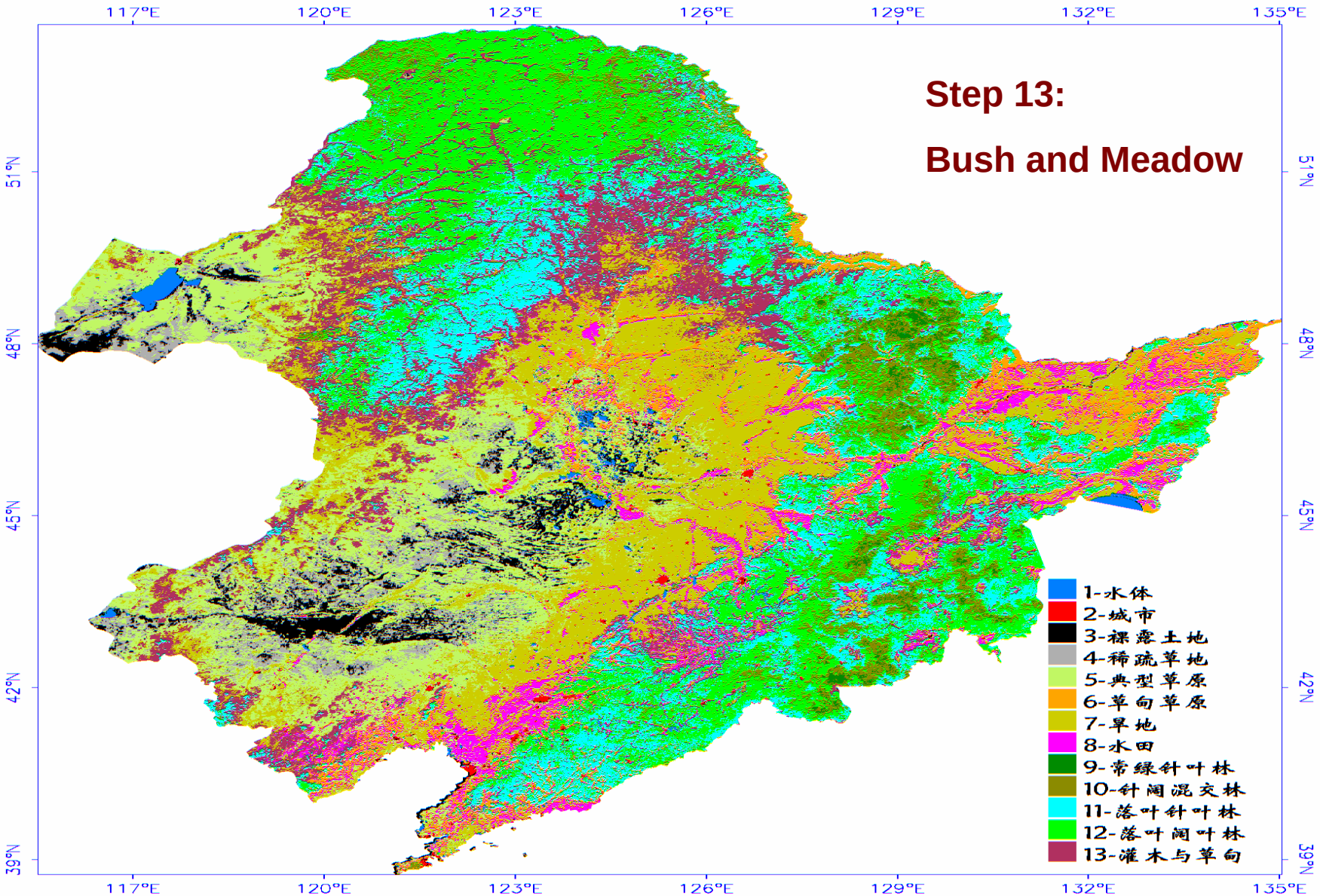
Step 11：If $NDWI_{day97\text{-}129}$<0 then ‘forest-deciduous – larix and birch ’
Step 12：If DEM<200 then ‘meadow grass’ else ‘bush’


Step 1 – Water
Step 2 - City
1-水体
2-城市


Step 3: Bare Soil
Step 4: Sparse Grass
1-水体
2-城市
3-裸露土地
4-稀疏草地

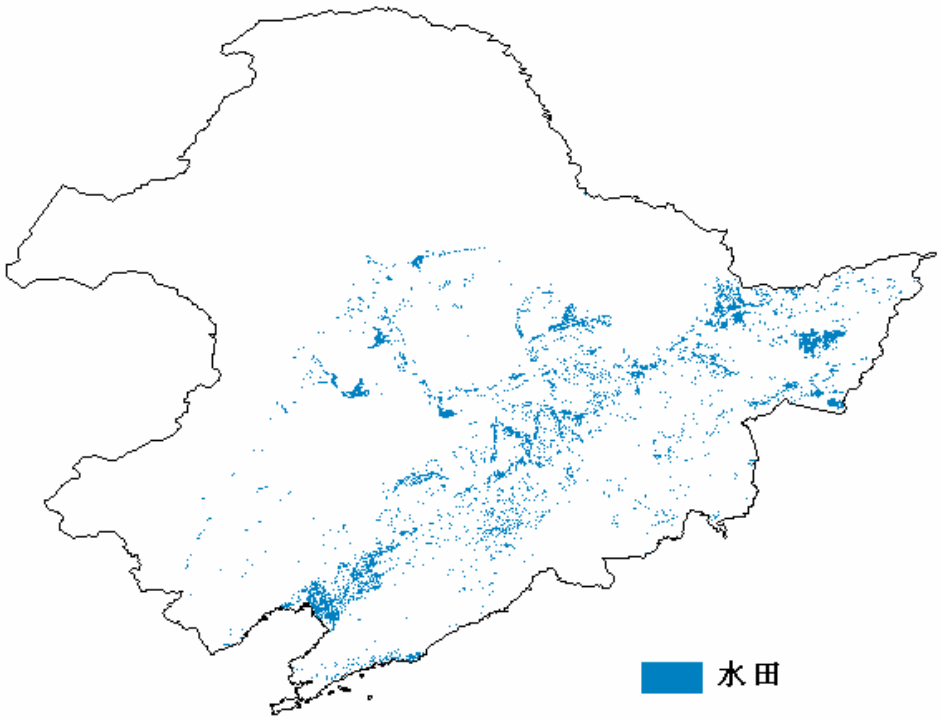
Step 10: Mixed Forest

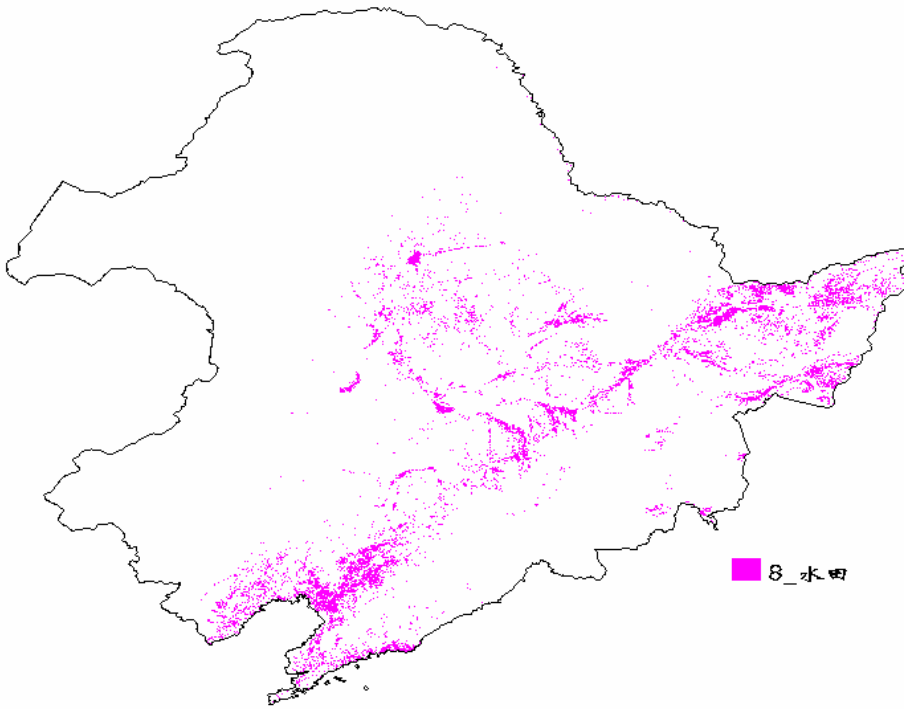



117°E
120°E
123°E
126°E
129°E
132°E
135°E
51°N
48°N
45°N
42°N
39°N
Step 13:
Bush and Meadow
1-水体
2-城市
3-裸露土地
4-稀疏草地
5-典型草原
6-草甸草原
7-旱地
8-水田
9-常绿针叶林
10-针阔混交林
11-落叶针叶林
12-落叶阔叶林
13-灌木与草甸

Validation of Step 8

Left：manually interpretated from TM (1：100000), about 2000
Right：from MODIS 1Km NDVI-NDWI, in 2002





6、Conclution

(1).Repeatability should be taken into consideration during long-term environmental monitoring.

(2).Quantitative description of land cover types could reduce subjectivity; especially in transitional regions.

(3).Input data quality should be carefully examined and processed.

(4).Both MODIS-NDVI and MODIS-NDWI are useful for quantitative description of land covers, other indices may have further information.

(5).More data and knowledge are needed to enhance the seperability of some land cover types: less productive cropland / grass…