

**PRAGMA 25, Beijing, 10.16-18, 2013**

## Qinghai Lake Case

Databases, data analysis and applications

Luo Ze ([luoze@cnic.cn](mailto:luoze@cnic.cn)); Yan Baoping ([ybp@cnci.cn](mailto:ybp@cnci.cn))

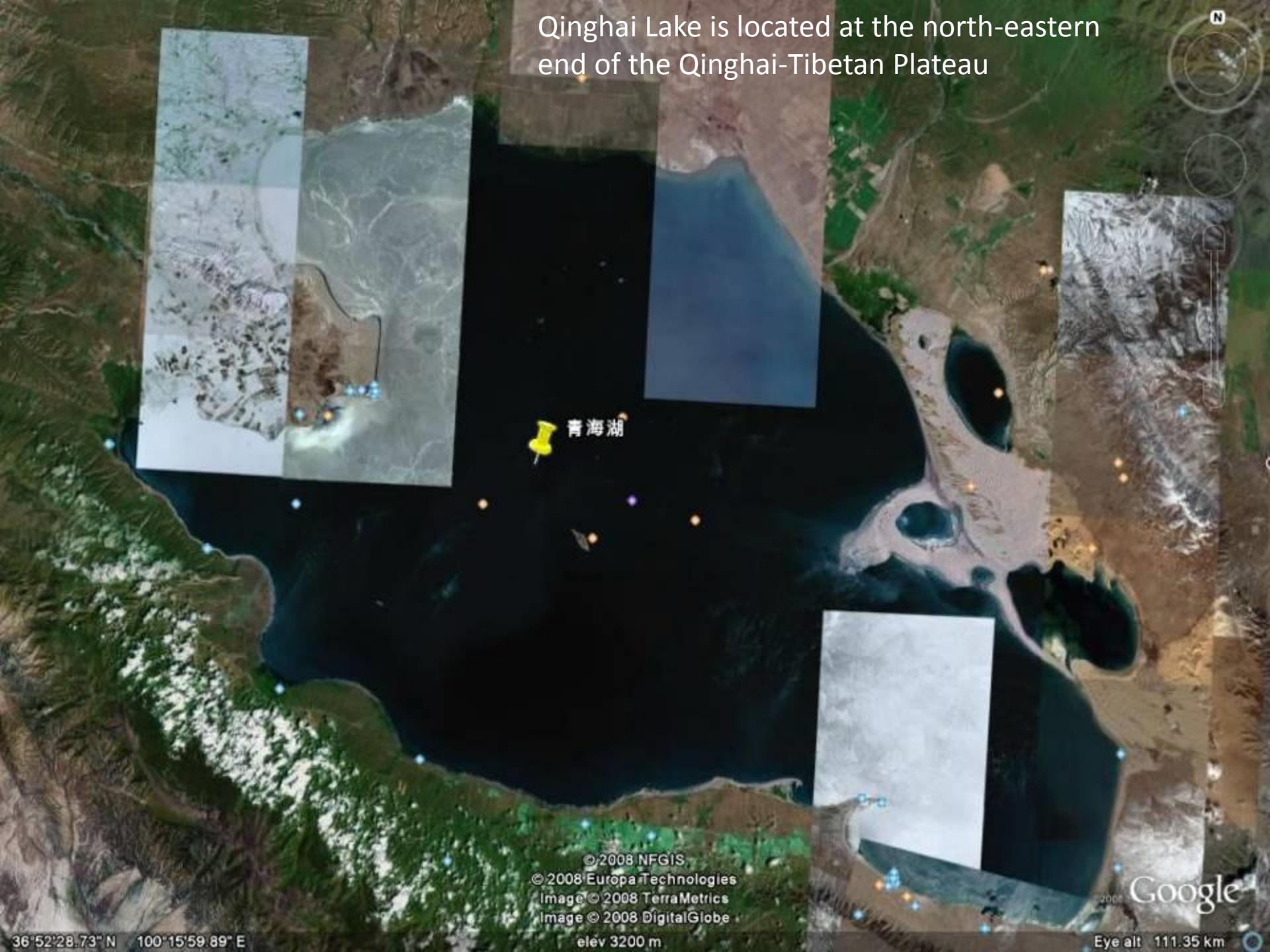
Computer Network Information Center, CAS

October 18, 2013

# Outline

- Introduction
- Database system and consideration
- Data analysis and visualization
- Applications
- Future work

Qinghai Lake is located at the north-eastern end of the Qinghai-Tibetan Plateau



© 2008 NFGIS

© 2008 Europa Technologies

Image © 2008 TerraMetrics

Image © 2008 DigitalGlobe

elev 3200 m

2008 Google

36°52'28.73" N 100°15'59.89" E

Eye alt 111.35 km





## Qinghai Lake

- In Qinghai-Tibet plateau
- the largest inland saline lake in China
- wetland of International Importance
- rich in plants and aquatic biological resources
- a paradise for birds living and breeding

## Qinghai Lake National Natural Reserve

- One of the eight major national reserves for wild birds in China
- Conservation of Migratory Birds and wetlands
- Protection of nature resources and ecological environment
- Protection of Biodiversity

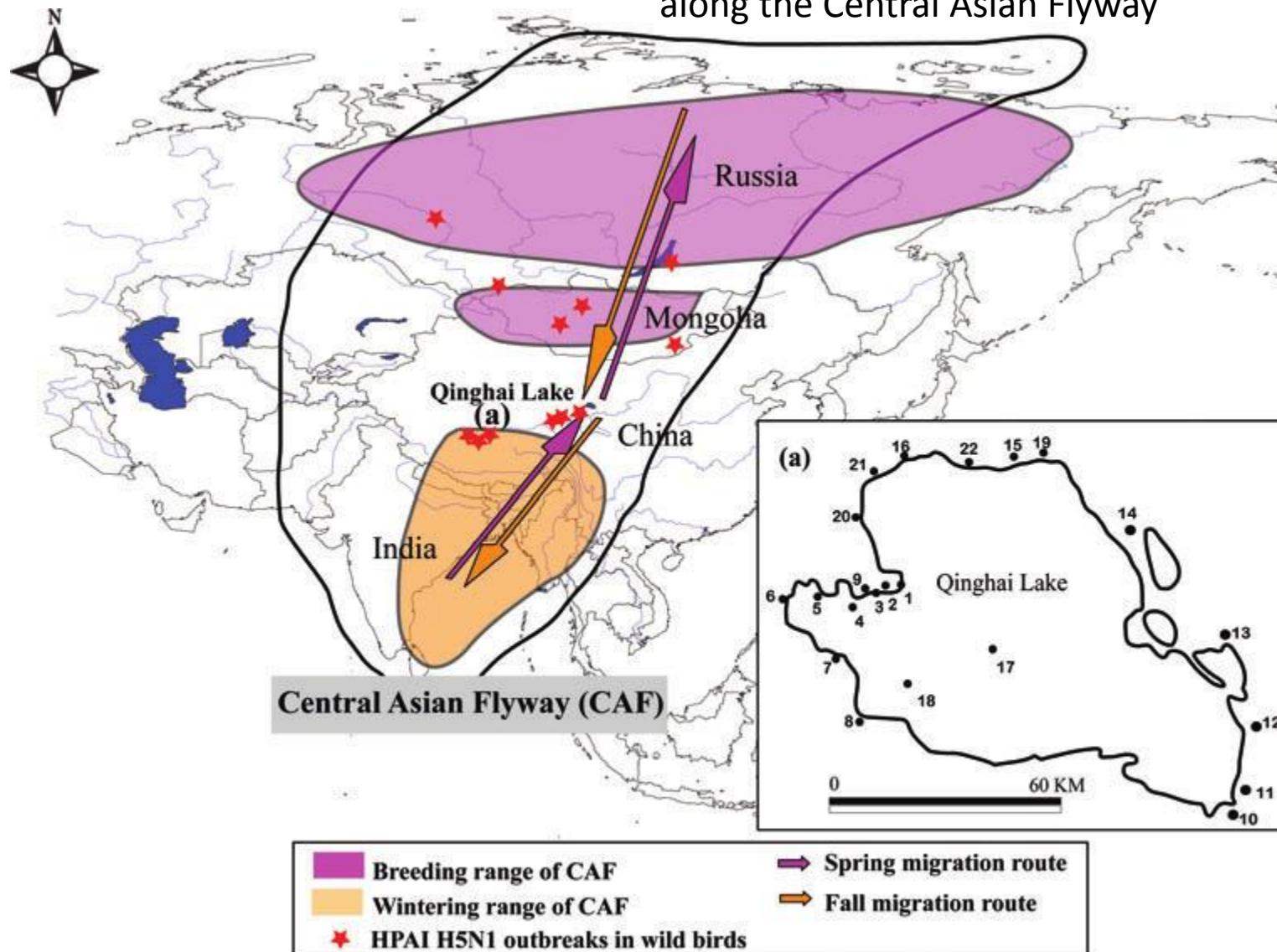




### Przewalski's Gazelle

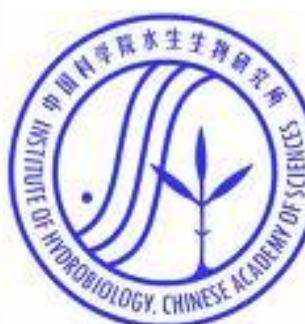
Endangered Species, just live in China,  
total number less than 500 in the world

Qinghai Lake is one of the most important breeding and stopover sites for migratory birds along the Central Asian Flyway



Location map of Qinghai Lake, including the range of the Central Asian Flyway.

# e-Science Application in Qinghai Lake Region



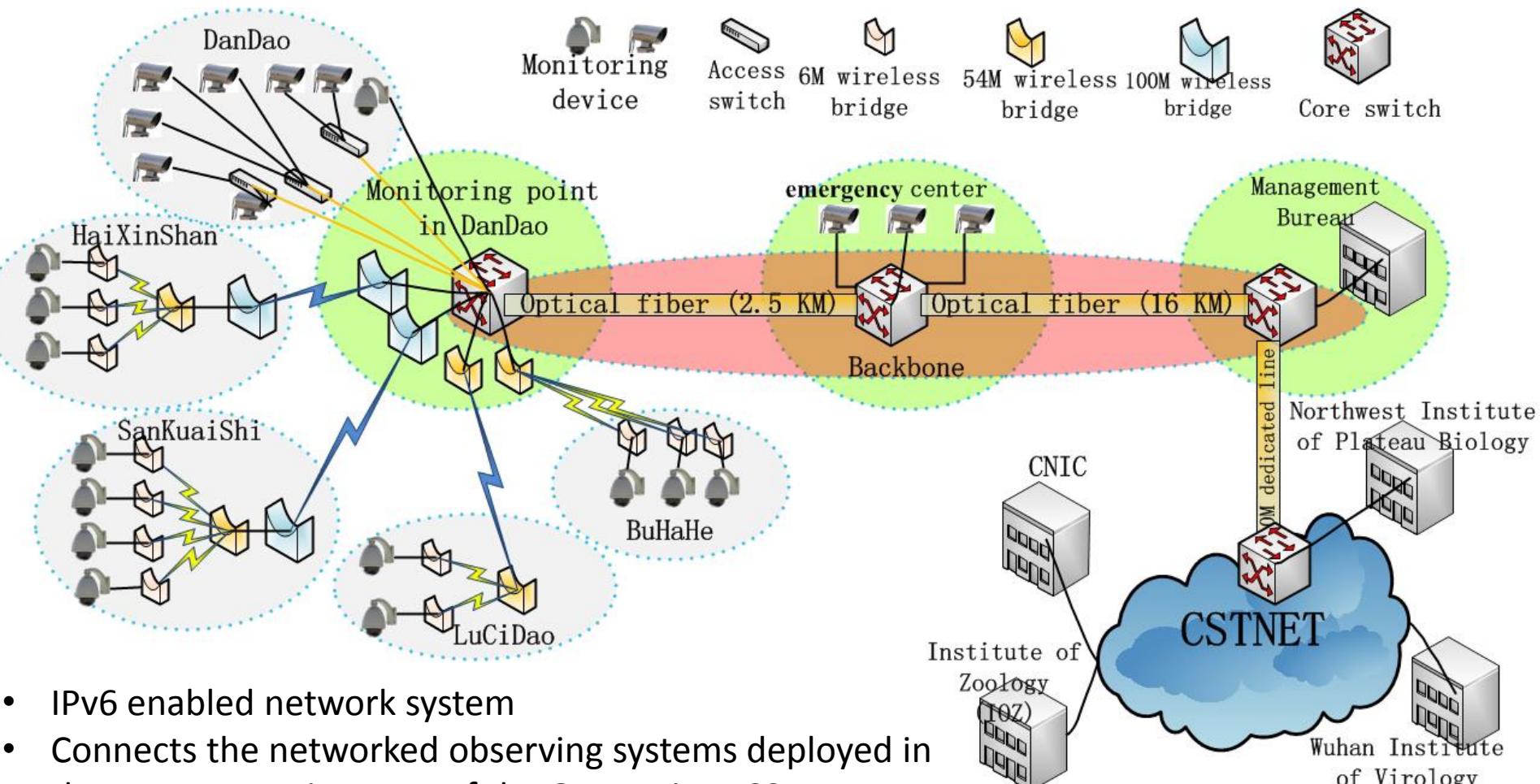
the first e-Science application for supporting ecological protection and research in plateau region, as well as the first application to national nature reserve in China

Joint Research Center of Chinese Academy of Sciences and  
Qinghai Lake National Nature Reserve

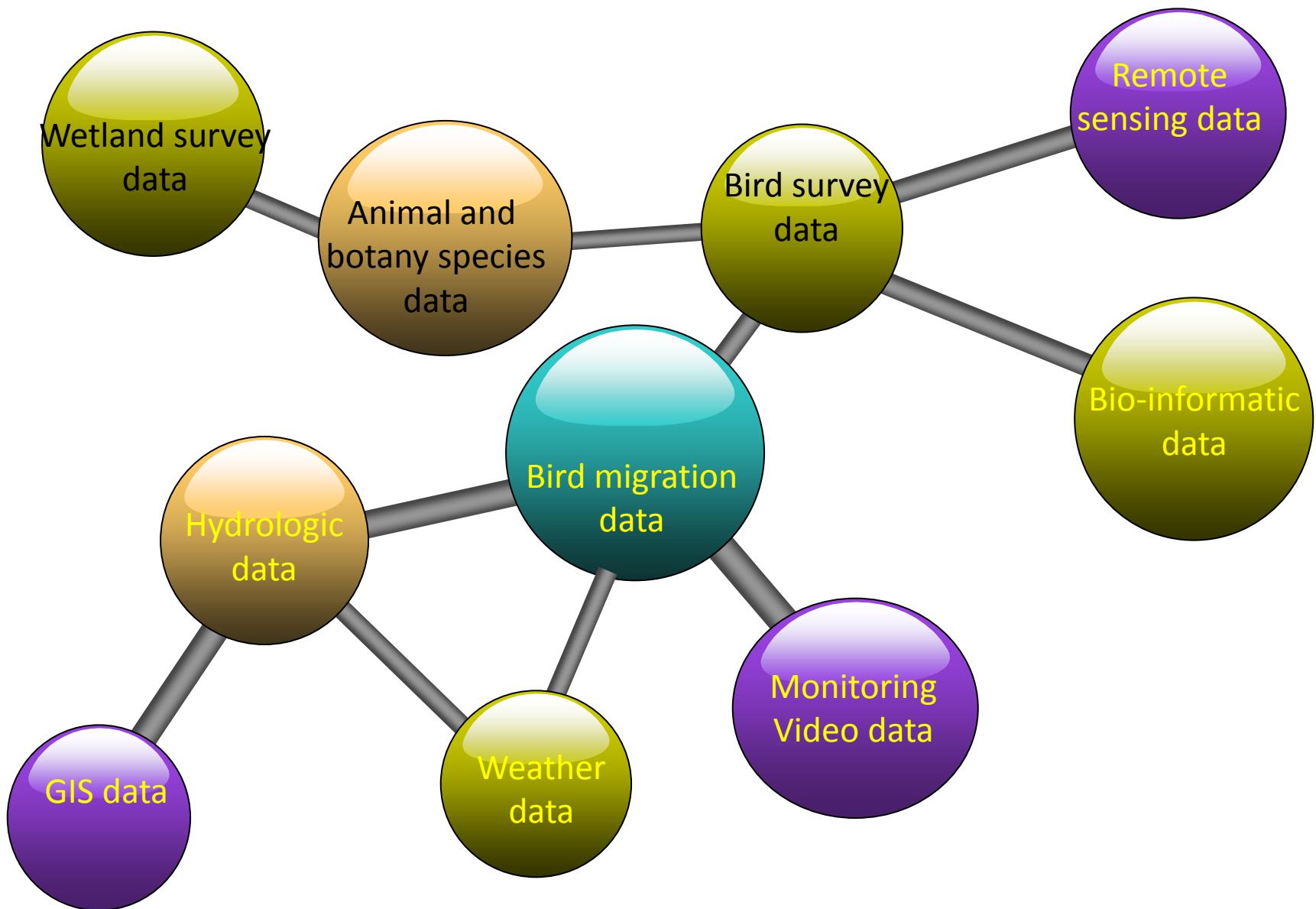
# Outline

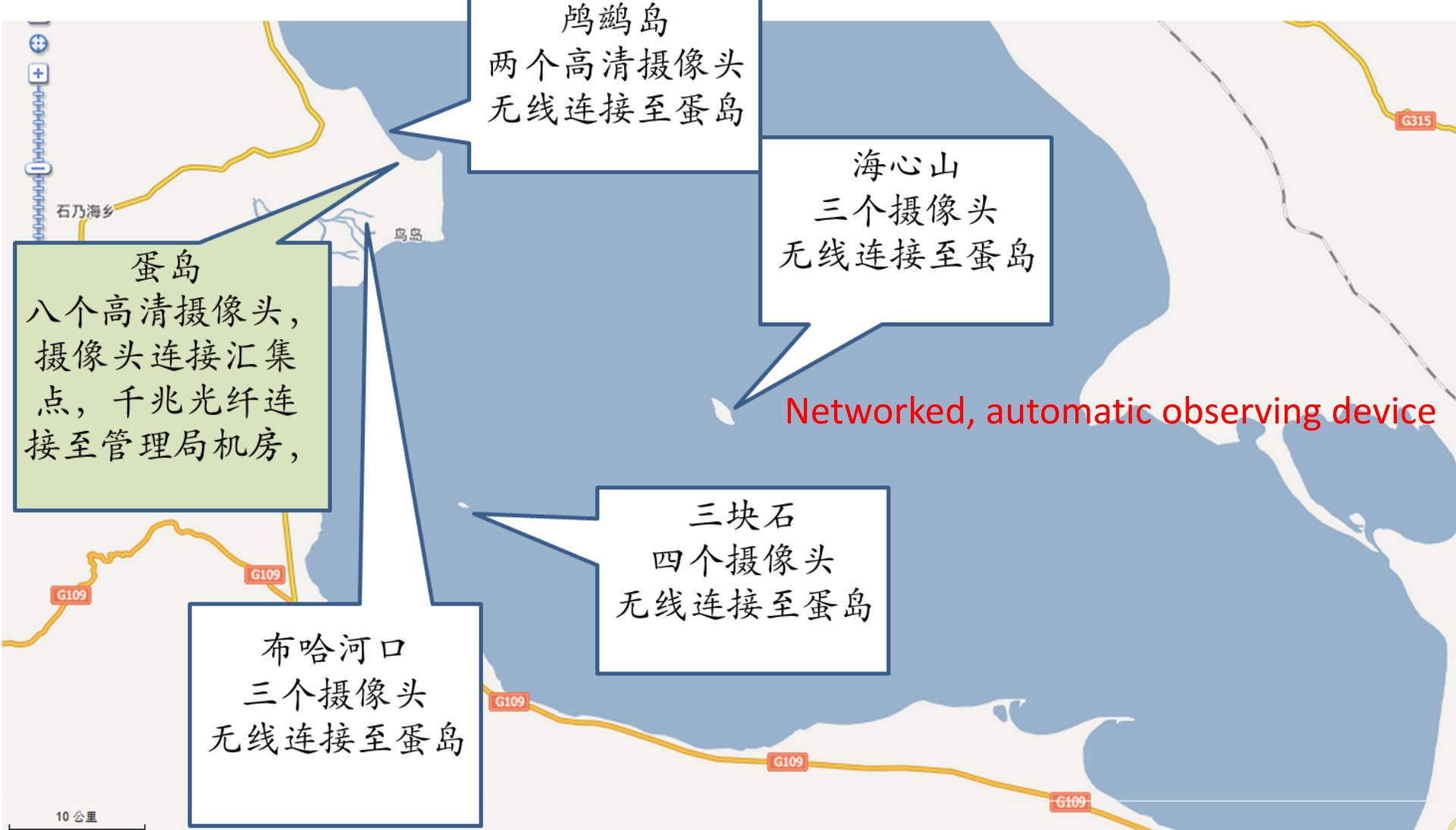
- Introduction
- Database system and consideration
- Data analysis and visualization
- Applications
- Future work

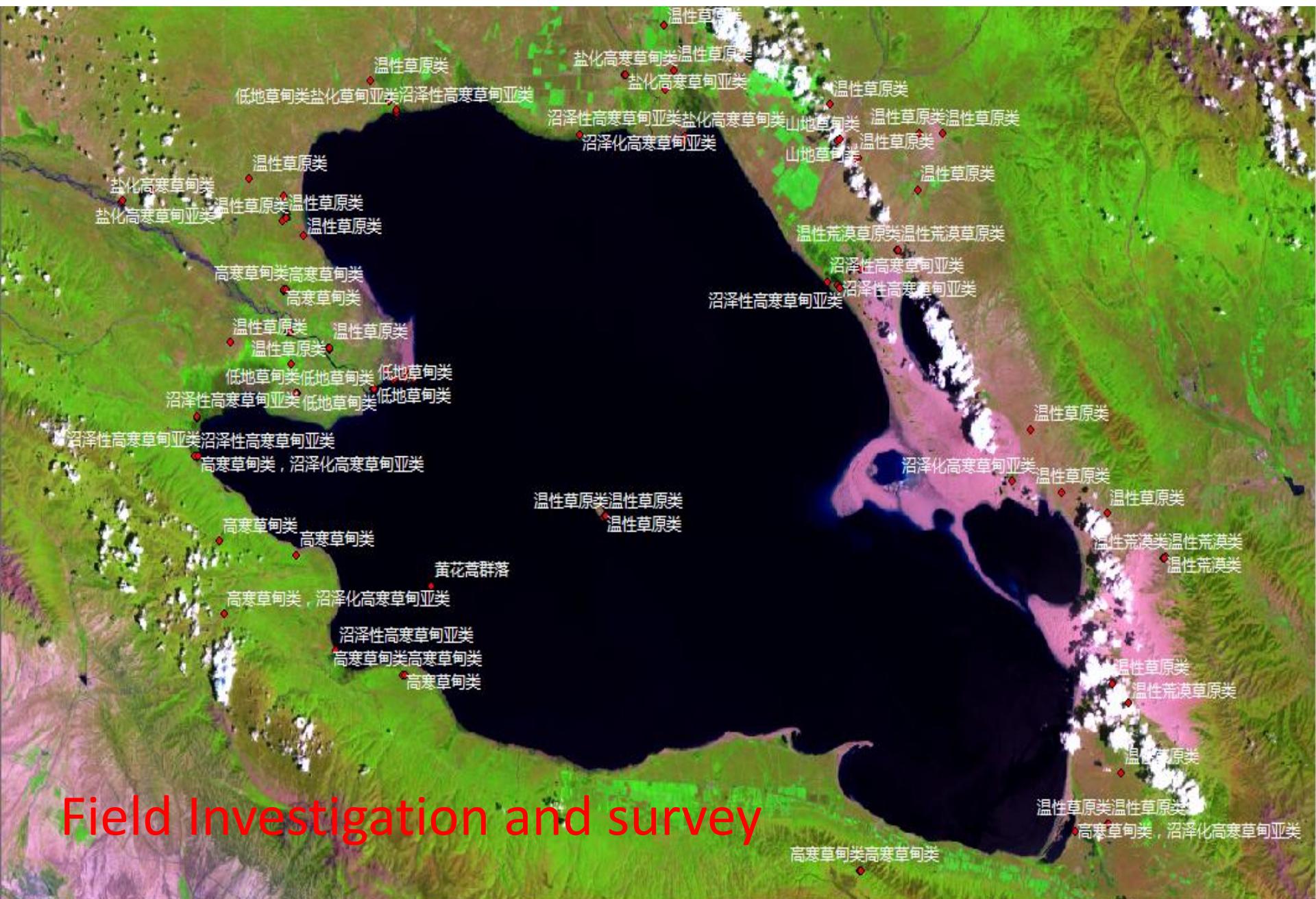
# Network Infrastructure



- IPv6 enabled network system
- Connects the networked observing systems deployed in the core protection area of the QLNNR into CSTNET
- A component of China Next Generation Network (CNGI)



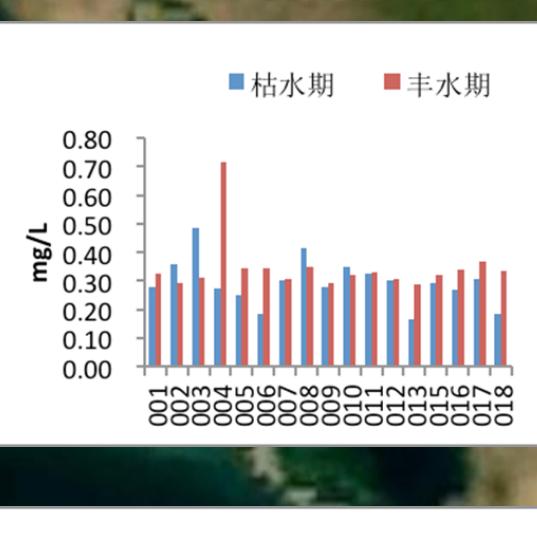
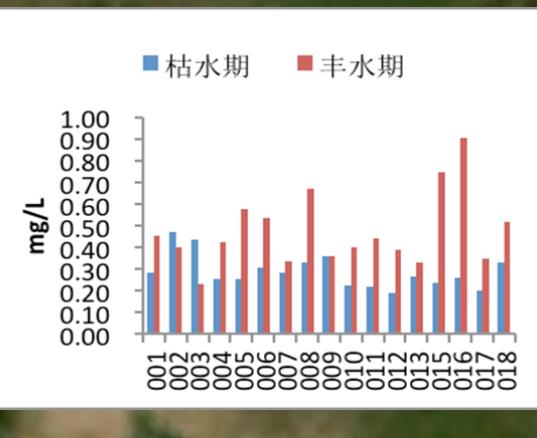




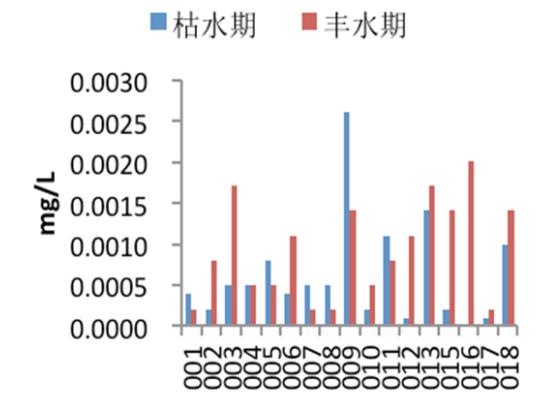
# Field Investigation and survey

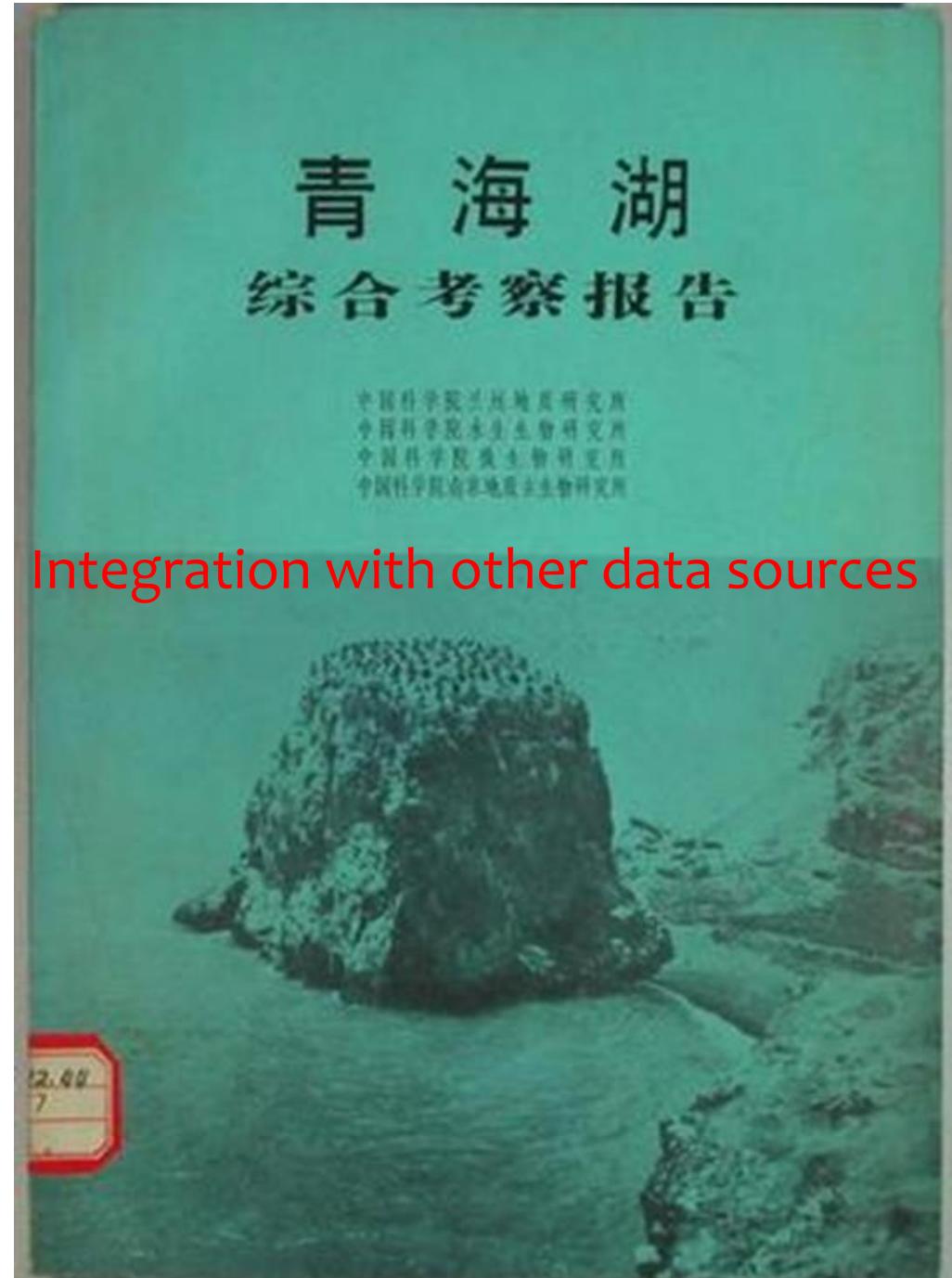
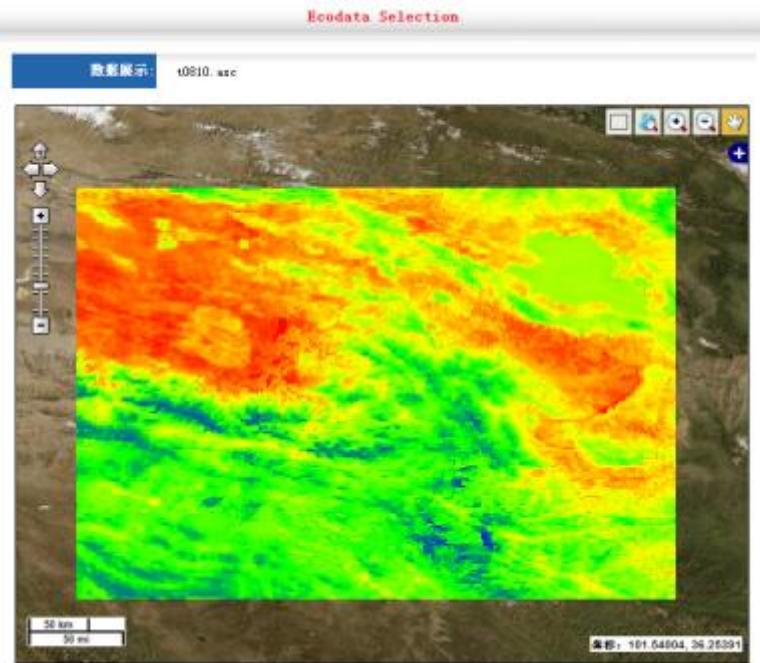
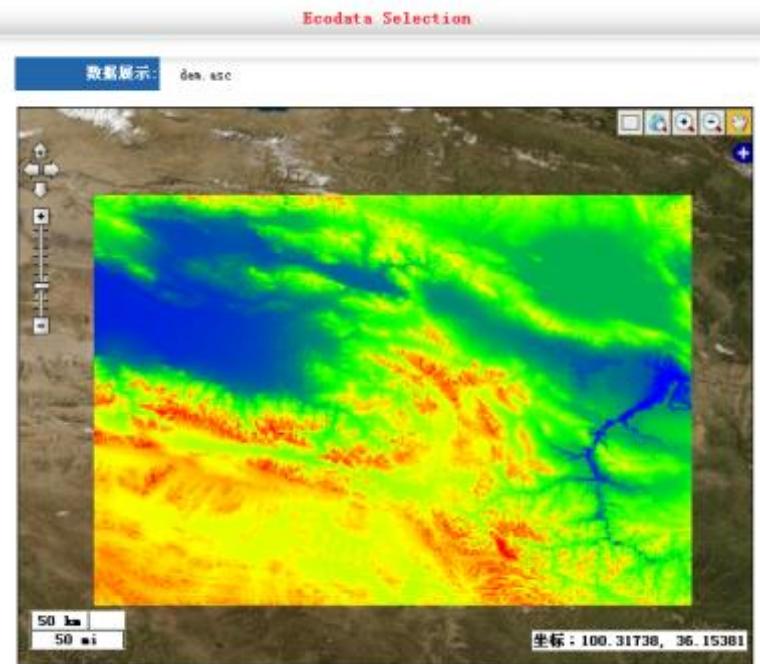
Virus Sampling

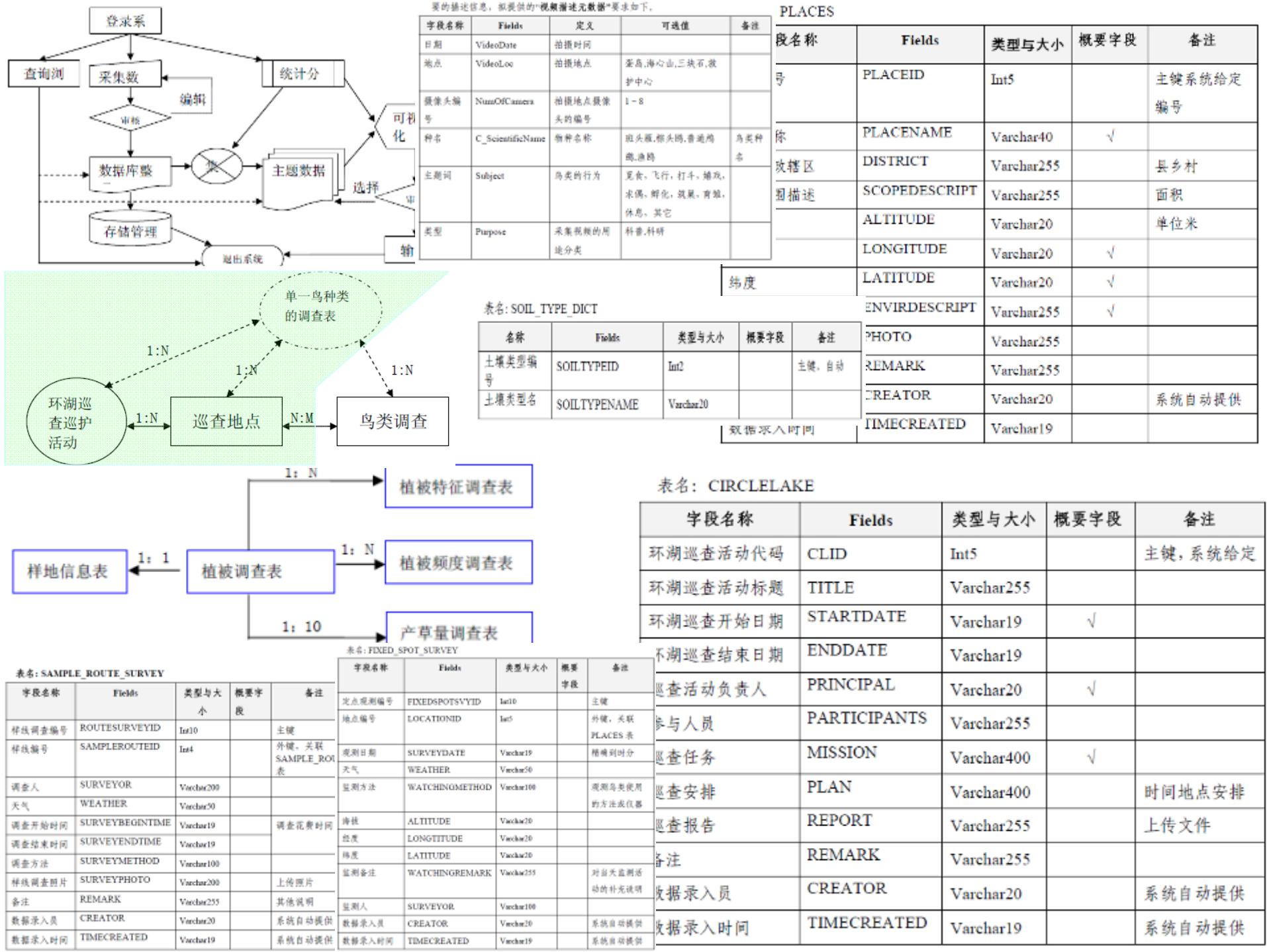




Water sampling





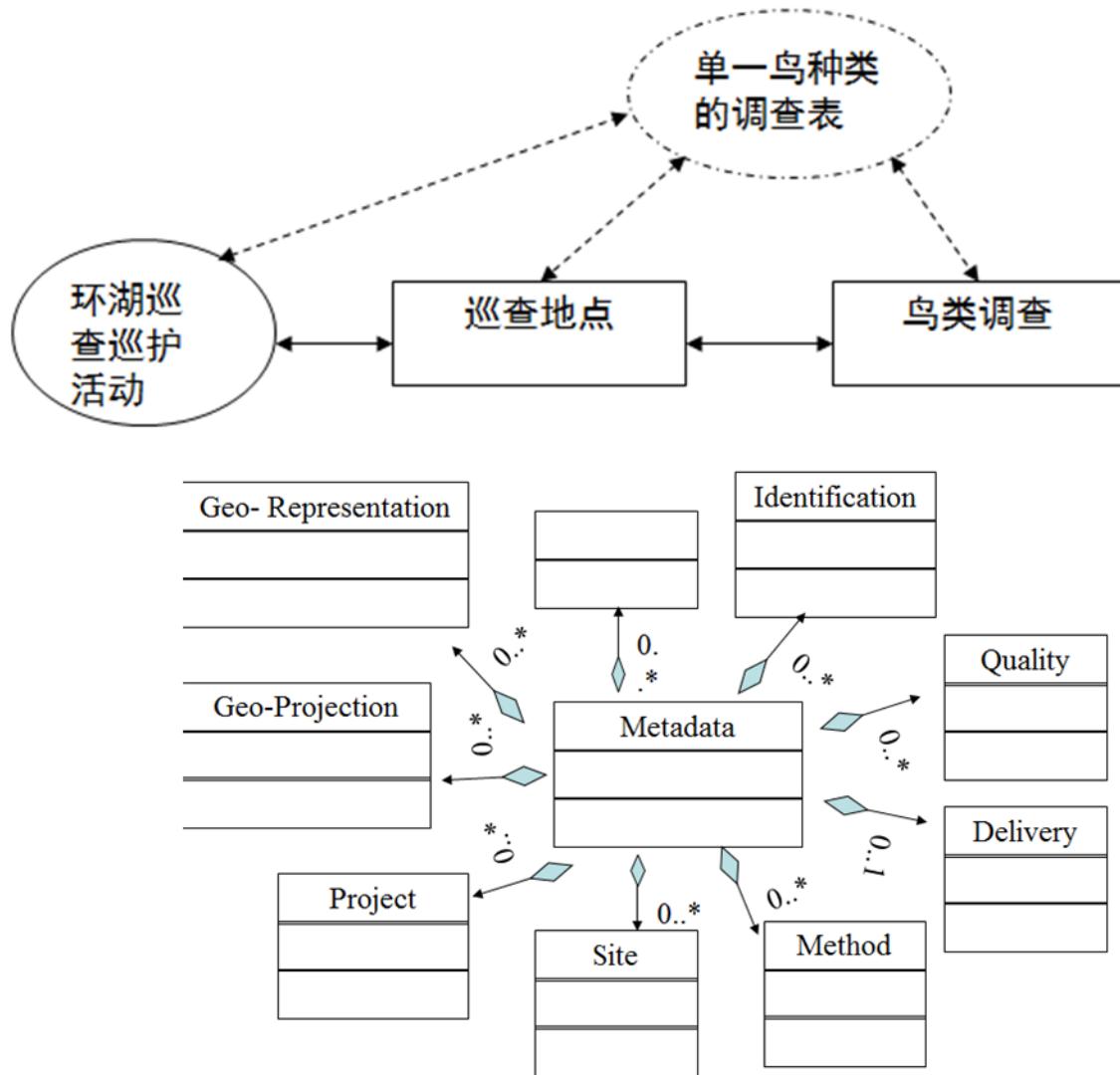


- 原来

- 数据库设计范式
- 降低数据冗余
- 多层级的表设计

- 重构之后

- 尽可能扁平设计
- 通过部分数据冗余实现数据关联和快速检索



1. Flat design as much as possible
2. Data association and quick search by partial data redundancy

# 野外调查登记表 t\_FieldSurvey

t_FieldSurvey		
<b>id</b>	int	<pk>
name	varchar(300)	
start_time	datetime	
end_time	datetime	
member	varchar(200)	
report_path	varchar(500)	
remark	varchar(200)	
ctime	datetime	

t FieldSurvey		
字段名	字段类型	说明
<b>id</b>	int	Id
name	varchar(300)	调查名称
start_time	datetime	开始时间
end_time	datetime	结束时间
member	varchar(200)	参与成员(分隔)
report_path	varchar(500)	考察报告路径(分隔)
remark	varchar(200)	备注
ctime	datetime	入库时间

t_VegeBushwood		
<b>id</b>	int	<pk>
sample_id	varchar(3)	
time	datetime	
site_id	int	<fk1>
vege_id	int	<fk2>
vege_name	varchar(50)	
quadrat_area	float	
length	float	
width	float	
height	float	
coverage	float	
number	int	
biomass_pn	float	
biomass_ps	float	
phenophase	varchar(50)	

t VegeBushwood		
字段名	字段类型	说明
<b>id</b>	int	Id
sample_id	varchar(3)	样地编号
time	datetime	记录时间
site_id	int	站点id
vege_id	int	植物id
vege_name	Varchar(50)	植物名称(冗余)
quadrat_area	float	样方面积m^2
length	float	长度cm
width	float	宽度cm
height	float	高度cm
coverage	float	盖度
number	int	株丛数
biomass_pn	float	生物量(g/标准丛)
biomass_ps	float	生物量(g/m^2)
phenophase	varchar(50)	物候期
life_status	varchar(50)	生活力
remark	varchar(200)	备注
ctime	datetime	入库时间

# 植被采样点基本信息表 t\_VegeSampleSite

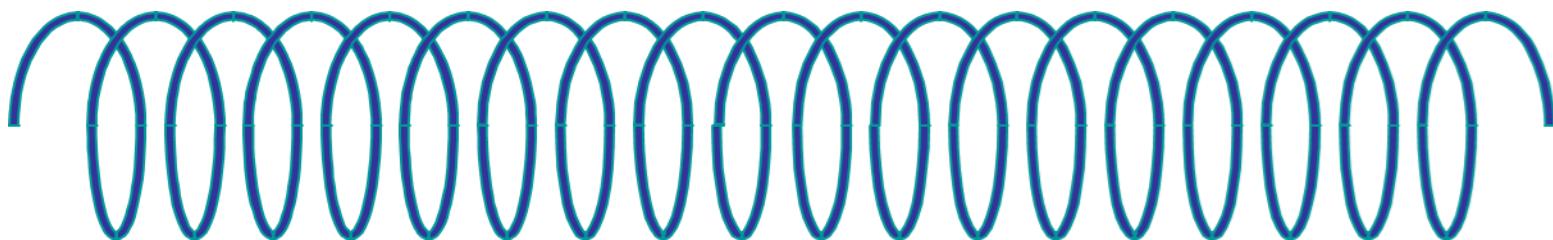
t_VegeSampleSite		
<u><b>id</b></u>	<u><b>int</b></u>	<u><b>&lt;pk&gt;</b></u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	
<u><b>fieldsurvey_id</b></u>	<u><b>int</b></u>	<u><b>&lt;fk1&gt;</b></u>
<u><b>image_close</b></u>	<u><b>int</b></u>	<u><b>&lt;fk2&gt;</b></u>
<u><b>image_long</b></u>	<u><b>int</b></u>	<u><b>&lt;fk3&gt;</b></u>
<u><b>name</b></u>	<u><b>varchar(100)</b></u>	
<u><b>site_code</b></u>	<u><b>varchar(100)</b></u>	
<u><b>site_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>area</b></u>	<u><b>float</b></u>	
<u><b>coverage</b></u>	<u><b>float</b></u>	
<u><b>time</b></u>	<u><b>datetime</b></u>	
<u><b>investigator</b></u>	<u><b>varchar(200)</b></u>	
<u><b>elevation</b></u>	<u><b>float</b></u>	
<u><b>longitude</b></u>	<u><b>float</b></u>	
<u><b>latitude</b></u>	<u><b>float</b></u>	
<u><b>admin_region</b></u>	<u><b>varchar(200)</b></u>	
<u><b>admin_code</b></u>	<u><b>varchar(50)</b></u>	
<u><b>reserve</b></u>	<u><b>varchar(100)</b></u>	
<u><b>proj_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>vege_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>soilero_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>soilero_degree</b></u>	<u><b>varchar(100)</b></u>	
<u><b>community</b></u>	<u><b>varchar(100)</b></u>	
<u><b>survey_index</b></u>	<u><b>varchar(100)</b></u>	
<u><b>survey_method</b></u>	<u><b>varchar(100)</b></u>	
<u><b>geomorphic_type</b></u>	<u><b>varchar(200)</b></u>	
<u><b>geological_matrix</b></u>	<u><b>varchar(100)</b></u>	
<u><b>humaneffect_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>humaneffect_degree</b></u>	<u><b>varchar(100)</b></u>	
<u><b>animaleffect_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>animaleffect_degree</b></u>	<u><b>varchar(100)</b></u>	
<u><b>soil_type</b></u>	<u><b>varchar(100)</b></u>	
<u><b>soil_feature</b></u>	<u><b>varchar(100)</b></u>	
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	
<u><b>ctime</b></u>	<u><b>datetime</b></u>	

字段名	字段类型	说明
<u><b>id</b></u>	<u><b>int</b></u>	<u><b>Id</b></u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	<u><b>样地编号</b></u>
<u><b>fieldsurvey_id</b></u>	<u><b>int</b></u>	<u><b>所属野外调查id</b></u>
<u><b>image_close</b></u>	<u><b>int</b></u>	<u><b>近景照片id</b></u>
<u><b>image_long</b></u>	<u><b>int</b></u>	<u><b>远景照片id</b></u>
<u><b>name</b></u>	<u><b>varchar(100)</b></u>	<u><b>样地名称</b></u>
<u><b>site_code</b></u>	<u><b>varchar(100)</b></u>	<u><b>站点代码 (行政代码-监测类别-序号(区域)-年度)</b></u>
<u><b>site_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>站点属性</b></u>
<u><b>area</b></u>	<u><b>float</b></u>	<u><b>面积</b></u>
<u><b>coverage</b></u>	<u><b>float</b></u>	<u><b>总覆盖度</b></u>
<u><b>time</b></u>	<u><b>timestamp</b></u>	<u><b>调查时间</b></u>
<u><b>investigator</b></u>	<u><b>varchar(200)</b></u>	<u><b>调查员(;分隔)</b></u>
<u><b>elevation</b></u>	<u><b>float</b></u>	<u><b>海拔高度</b></u>
<u><b>longitude</b></u>	<u><b>float</b></u>	<u><b>经度</b></u>
<u><b>latitude</b></u>	<u><b>float</b></u>	<u><b>纬度</b></u>
<u><b>admin_region</b></u>	<u><b>varchar(200)</b></u>	<u><b>所属行政区域</b></u>
<u><b>admin_code</b></u>	<u><b>varchar(50)</b></u>	<u><b>所属行政区域代码</b></u>
<u><b>reserve</b></u>	<u><b>varchar(100)</b></u>	<u><b>所属自然保护分区</b></u>
<u><b>proj_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>工程项目类别</b></u>
<u><b>vege_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>植被类型(;分隔)</b></u>
<u><b>soilero_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>土壤侵蚀类型(;分隔)</b></u>
<u><b>soilero_degree</b></u>	<u><b>varchar(100)</b></u>	<u><b>土壤侵蚀程度(;分隔)</b></u>
<u><b>community</b></u>	<u><b>varchar(100)</b></u>	<u><b>群落名称</b></u>
<u><b>survey_index</b></u>	<u><b>varchar(100)</b></u>	<u><b>监测指标(;分隔)</b></u>
<u><b>survey_method</b></u>	<u><b>varchar(100)</b></u>	<u><b>监测方法(;分隔)</b></u>
<u><b>geomorphic_type</b></u>	<u><b>varchar(200)</b></u>	<u><b>地形地貌(;分隔)</b></u>
<u><b>geological_matrix</b></u>	<u><b>varchar(100)</b></u>	<u><b>基质口(;分隔)</b></u>
<u><b>humaneffect_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>人类影响方式(;分隔)</b></u>
<u><b>humaneffect_degree</b></u>	<u><b>varchar(100)</b></u>	<u><b>人类影响程度(;分隔)</b></u>
<u><b>animaleffect_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>动物活动影响方式(;分隔)</b></u>
<u><b>animaleffect_degree</b></u>	<u><b>varchar(100)</b></u>	<u><b>动物活动影响程度(;分隔)</b></u>
<u><b>soil_type</b></u>	<u><b>varchar(100)</b></u>	<u><b>土壤类型(;分隔)</b></u>
<u><b>soil_feature</b></u>	<u><b>varchar(100)</b></u>	<u><b>土壤特征(;分隔)</b></u>
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	<u><b>备注</b></u>
<u><b>ctime</b></u>	<u><b>datetime</b></u>	<u><b>入库时间</b></u>

# 植被类型登记表 t\_VegeTypeDetail

t_VegeTypeDetail		
<u><b>id</b></u>	<u><b>int</b></u>	<u>&lt;pk&gt;</u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	
<u><b>time</b></u>	<u><b>datetime</b></u>	
<u><b>site_id</b></u>	<u><b>int</b></u>	<u>&lt;fk1&gt;</u>
<u><b>vege_id</b></u>	<u><b>int</b></u>	<u>&lt;fk2&gt;</u>
<u><b>vege_name</b></u>	<u><b>varchar(50)</b></u>	
<u><b>level</b></u>	<u><b>varchar(20)</b></u>	
<u><b>avgheight_yy</b></u>	<u><b>float</b></u>	
<u><b>avgheight_sz</b></u>	<u><b>float</b></u>	
<u><b>number</b></u>	<u><b>int</b></u>	
<u><b>coverage</b></u>	<u><b>float</b></u>	
<u><b>life_form</b></u>	<u><b>varchar(50)</b></u>	
<u><b>branch_form</b></u>	<u><b>varchar(50)</b></u>	
<u><b>life_status</b></u>	<u><b>varchar(50)</b></u>	
<u><b>phenophase</b></u>	<u><b>varchar(50)</b></u>	
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	
<u><b>ctime</b></u>	<u><b>datetime</b></u>	

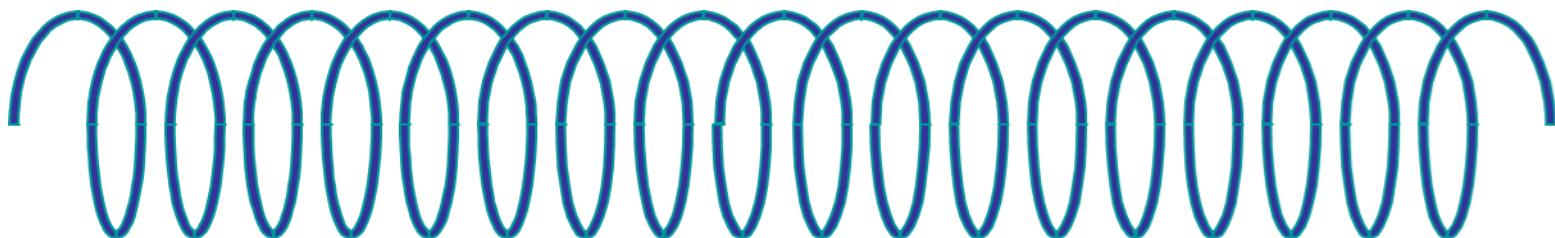
t_VegeTypeDetail		
字段名	字段类型	说明
<u><b>id</b></u>	<u><b>int</b></u>	<u>Id</u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	<u>样地编号</u>
<u><b>time</b></u>	<u><b>datetime</b></u>	<u>记录时间</u>
<u><b>site_id</b></u>	<u><b>int</b></u>	<u>站点id</u>
<u><b>vege_id</b></u>	<u><b>int</b></u>	<u>植物id</u>
<u><b>vege_name</b></u>	<u><b>Varchar(50)</b></u>	<u>植物名称(冗余)</u>
<u><b>level</b></u>	<u><b>varchar(20)</b></u>	<u>层次</u>
<u><b>avgheight_yy</b></u>	<u><b>float</b></u>	<u>营养枝平均高度(cm)</u>
<u><b>avgheight_sz</b></u>	<u><b>float</b></u>	<u>生殖枝平均高度(cm)</u>
<u><b>number</b></u>	<u><b>int</b></u>	<u>株丛数</u>
<u><b>coverage</b></u>	<u><b>float</b></u>	<u>投影盖度</u>
<u><b>life_form</b></u>	<u><b>varchar(50)</b></u>	<u>生活型</u>
<u><b>branch_form</b></u>	<u><b>varchar(50)</b></u>	<u>枝叶型</u>
<u><b>life_status</b></u>	<u><b>varchar(50)</b></u>	<u>生命力</u>
<u><b>phenophase</b></u>	<u><b>varchar(50)</b></u>	<u>物候相</u>
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	<u>备注</u>
<u><b>ctime</b></u>	<u><b>datetime</b></u>	<u>入库时间</u>



# 植物频度样方登记表 t\_VegeFrequency

t_VegeFrequency		
<u><b>id</b></u>	<u><b>int</b></u>	<u>&lt;pk&gt;</u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	
<u><b>time</b></u>	<u><b>datetime</b></u>	
<u><b>site_id</b></u>	<u><b>int</b></u>	<u>&lt;fk1&gt;</u>
<u><b>vege_id</b></u>	<u><b>int</b></u>	<u>&lt;fk2&gt;</u>
<u><b>vege_name</b></u>	<u><b>varchar(50)</b></u>	
<u><b>quadrat_1</b></u>	<u><b>int</b></u>	
<u><b>quadrat_2</b></u>	<u><b>int</b></u>	
<u><b>quadrat_3</b></u>	<u><b>int</b></u>	
<u><b>quadrat_4</b></u>	<u><b>int</b></u>	
<u><b>quadrat_5</b></u>	<u><b>int</b></u>	
<u><b>quadrat_6</b></u>	<u><b>int</b></u>	
<u><b>quadrat_7</b></u>	<u><b>int</b></u>	
<u><b>quadrat_8</b></u>	<u><b>int</b></u>	
<u><b>quadrat_9</b></u>	<u><b>int</b></u>	
<u><b>quadrat_10</b></u>	<u><b>int</b></u>	
<u><b>average</b></u>	<u><b>float</b></u>	
<u><b>frequency</b></u>	<u><b>float</b></u>	
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	
<u><b>ctime</b></u>	<u><b>datetime</b></u>	

t_VegeFrequency		
字段名	字段类型	说明
<u><b>id</b></u>	<u><b>int</b></u>	<u>Id</u>
<u><b>sample_id</b></u>	<u><b>varchar(3)</b></u>	<u>样地编号</u>
<u><b>time</b></u>	<u><b>datetime</b></u>	<u>记录时间</u>
<u><b>site_id</b></u>	<u><b>int</b></u>	<u>站点id</u>
<u><b>vege_id</b></u>	<u><b>int</b></u>	<u>植物id</u>
<u><b>vege_name</b></u>	<u><b>Varchar(50)</b></u>	<u>植物名称(冗余)</u>
<u><b>quadrat_1</b></u>	<u><b>int</b></u>	<u>样方1(1:有, 0无)</u>
<u><b>quadrat_2</b></u>	<u><b>int</b></u>	<u>样方2(1:有, 0无)</u>
<u><b>quadrat_3</b></u>	<u><b>int</b></u>	<u>样方3(1:有, 0无)</u>
<u><b>quadrat_4</b></u>	<u><b>int</b></u>	<u>样方4(1:有, 0无)</u>
<u><b>quadrat_5</b></u>	<u><b>int</b></u>	<u>样方5(1:有, 0无)</u>
<u><b>quadrat_6</b></u>	<u><b>int</b></u>	<u>样方6(1:有, 0无)</u>
<u><b>quadrat_7</b></u>	<u><b>int</b></u>	<u>样方7(1:有, 0无)</u>
<u><b>quadrat_8</b></u>	<u><b>int</b></u>	<u>样方8(1:有, 0无)</u>
<u><b>quadrat_9</b></u>	<u><b>int</b></u>	<u>样方9(1:有, 0无)</u>
<u><b>quadrat_10</b></u>	<u><b>int</b></u>	<u>样方10(1:有, 0无)</u>
<u><b>average</b></u>	<u><b>float</b></u>	<u>平均</u>
<u><b>frequency</b></u>	<u><b>float</b></u>	<u>频度</u>
<u><b>remark</b></u>	<u><b>varchar(200)</b></u>	<u>备注</u>
<u><b>ctime</b></u>	<u><b>datetime</b></u>	<u>入库时间</u>



# 植物登记表 t\_VegetationDic

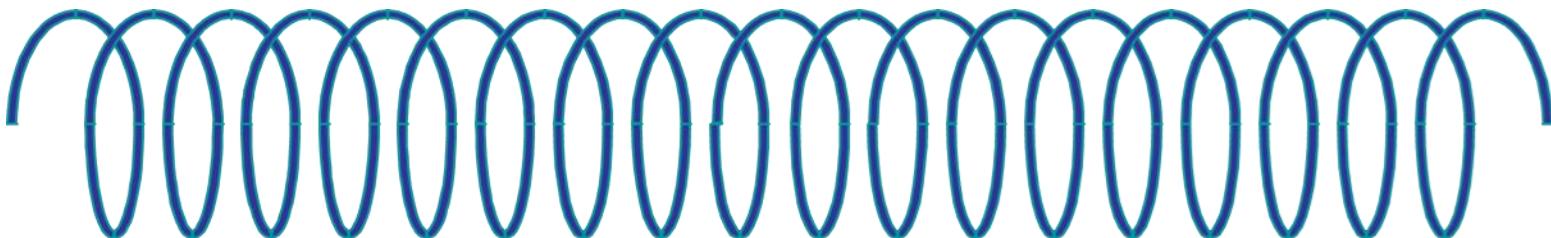
t_VegetationDic		
id	int	<pk>
name	varchar(50)	
sample_id	varchar(50)	
remark	varchar(200)	
image_id	int	<fk>
ctime	datetime	

t_Vegetation		
字段名	字段类型	说明
id	int	id
name	varchar(50)	植物名称
sample_id	varchar(50)	标本编号
remark	varchar(200)	备注
image_id	int	照片id
ctime	datetime	入库时间

# 照片登记表 t\_Image

t_Image		
id	int	<pk>
name	varchar(100)	
type	varchar(100)	
path	varchar(500)	
author	varchar(100)	
remark	varchar(200)	
ctime	datetime	

t_Image		
字段名	字段类型	说明
id	int	Id
name	varchar(100)	照片名称
type	varchar(100)	类型
remark	varchar(200)	备注
author	varchar(100)	拍摄者
path	varchar(500)	路径
ctime	datetime	入库时间



t_BirdSampleSite		
<u><b>id</b></u>	int	<pk>
fieldsurvey_id	int	<fk>
name	varchar(100)	
longitude	float	
latitude	float	
elevation	float	
start_time	timestamp	
end_time	timestamp	
weather	varchar(100)	
observer	varchar(200)	
recorder	varchar(200)	
wetland_type	varchar(200)	
total_species	int	
total_number	int	
remark	varchar(200)	
ctime	datetime	

t_BirdSampleSite		
字段名	字段类型	说明
<u><b>id</b></u>	int	Id
fieldsurvey_id	int	调查id
name	varchar(100)	样点名称
longitude	float	经度
latitude	float	纬度
elevation	float	海拔高度
start_time	timestamp	开始时间
end_time	timestamp	结束时间
weather	varchar(100)	天气状况
observer	varchar(200)	观察者
recorder	varchar(200)	记录者
wetland_type	varchar(200)	湿地类型
total_species	int	种类数
total_number	int	总数
remark	varchar(200)	备注
ctime	datetime	入库时间

t_BirdSurvey		
<u><b>id</b></u>	int	<pk>
time	datetime	
site_id	int	<fk1>
bird_id	int	<fk2>
bird_name	varchar(100)	
number	int	
remark	varchar(200)	
ctime	datetime	

t_BirdSurvey		
字段名	字段类型	说明
<u><b>id</b></u>	int	Id
time	datetime	记录时间
site_id	int	站点id
bird_id	int	鸟类id
bird_name	varchar(100)	鸟类名称
number	int	数量
remark	varchar(200)	备注
ctime	datetime	入库时间

# Outline

- Introduction
- Database system and consideration
- Data analysis and visualization
- Applications
- Future work

# Satellite Telemetry

- Since March and April 2007, an international team from U. S. Geological Survey, United Nations-FAO, Chinese Academy of Sciences, and Qinghai Lake National Nature Reserve developed a project of satellite-marking migrant birds at the Qinghai Lake region to track their movements and migration routes.



中国科学院  
Chinese academy of sciences  
青海湖国家级自然保护区联合科研基地





Equipment was programmed to record GPS locations every 2 hours, and the record data were uploaded to the Argos satellite tracking system every 2 days



GPS solar-powered Platform Terminal Transmitters (PTT)



TABLE I. THE STATISTICS OF BIRDS' SATELLITE TELEMETRY DATA

Species	Statistical Information			
	<i>Tracked Birds Number</i>	<i>Record Start Time</i>	<i>Record End Time</i>	<i>Record Number</i>
BHG	29	2007-03-21	2010-01-06	973,649
RSD	20	2007-03-21	2010-05-24	393,072
GBG	10	2007-06-21	2008-06-07	37,242
Total	59	2007-03-21	2010-05-24	1,403,963

TABLE II. RELATION REPRESENTATION OF BIRD MIGRATION DATA

<b>Id</b>	<b>Animal</b>	<b>PTT</b>	<b>Date</b>	<b>Time</b>	<b>Latitude</b>	<b>Longitude</b>	<b>LC94</b>	<b>Speed (km/h)</b>
85	BH07_67695	67695	2008-03-02	3:27:10	29.275	88.731	LZ	32
86	BH08_67688	67688	2008-03-02	4:27:10	30.275	89.253	LG	43

# Bar-headed Geese *Anser indicus* home range and local movement analysis in Qinghai Lake and Lhasa River Watershed, China

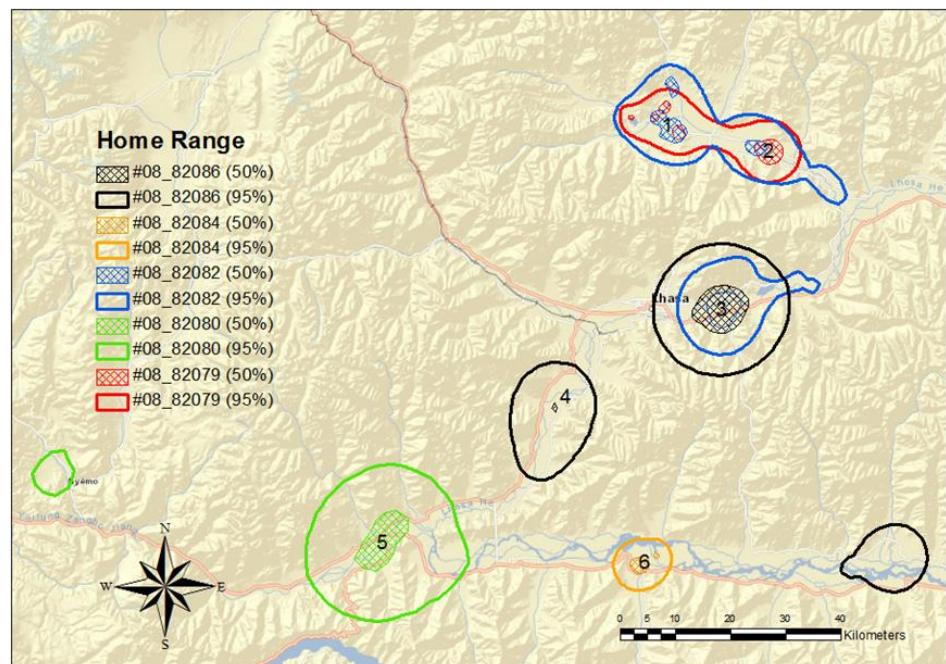
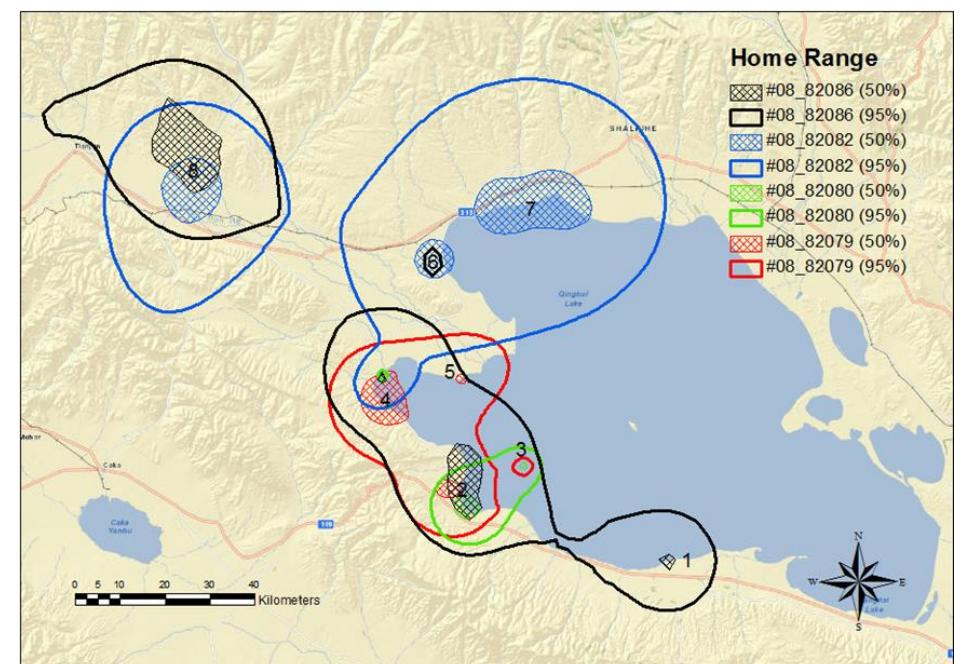
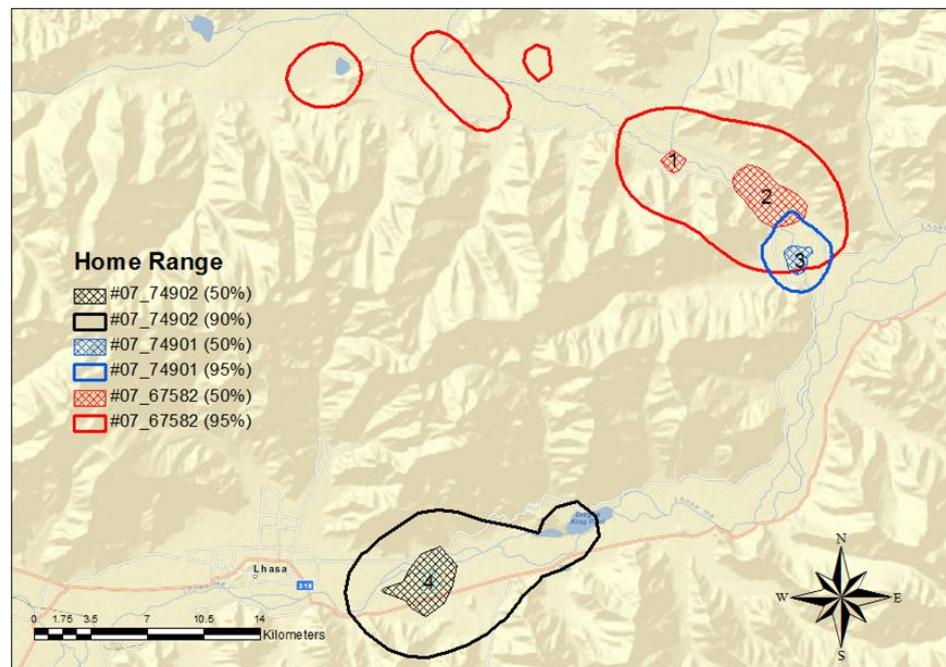
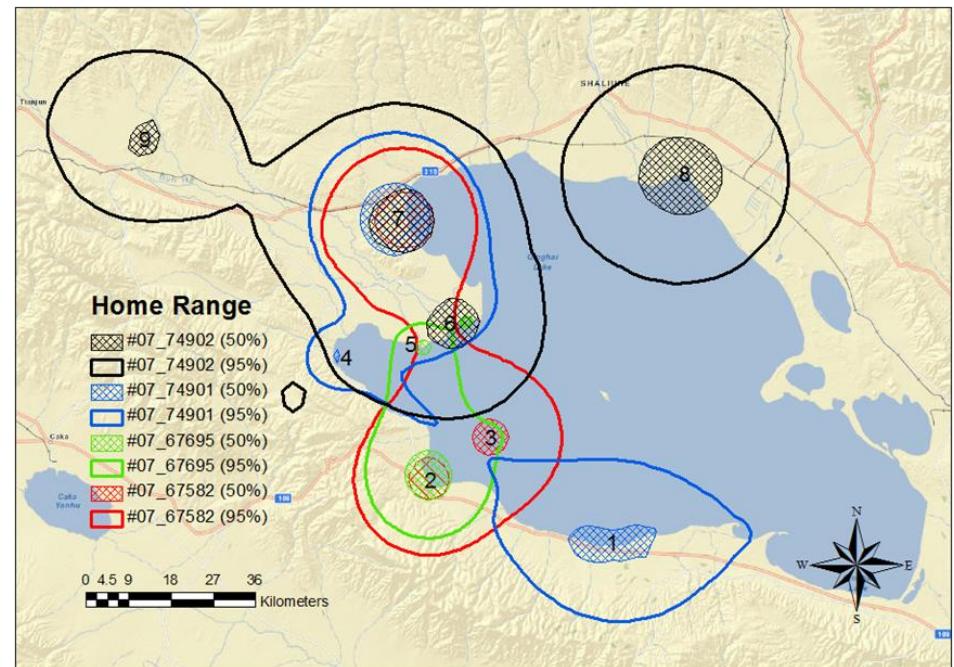
- The detailed information about home range, utilization distribution, and local daily movement pattern in these two key sites
- Home ranges of the marked birds in Qinghai Lake and Lhasa River Watershed
- Daily home range
- Local movement pattern
- Daily movement distance and daily movement rate

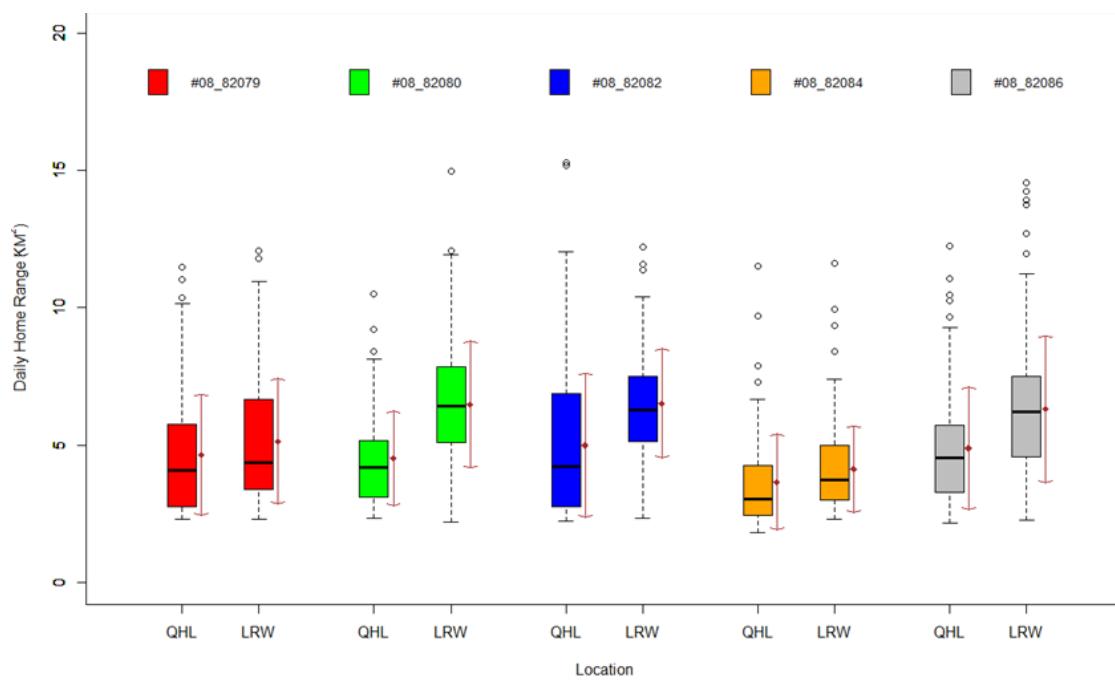
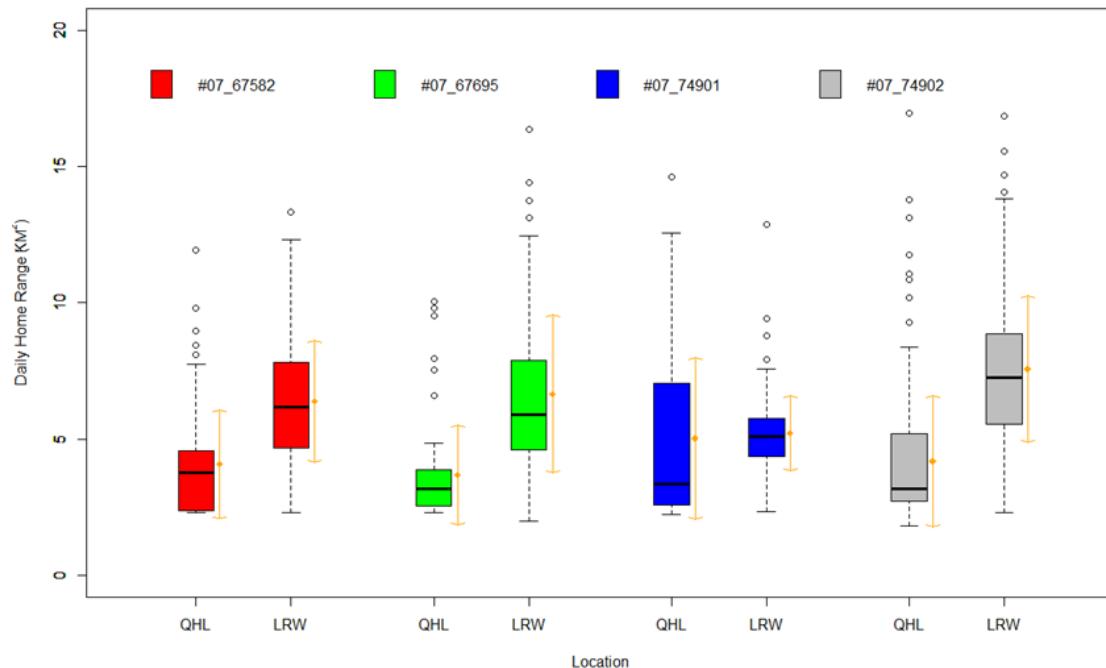
# Daily movement pattern analysis

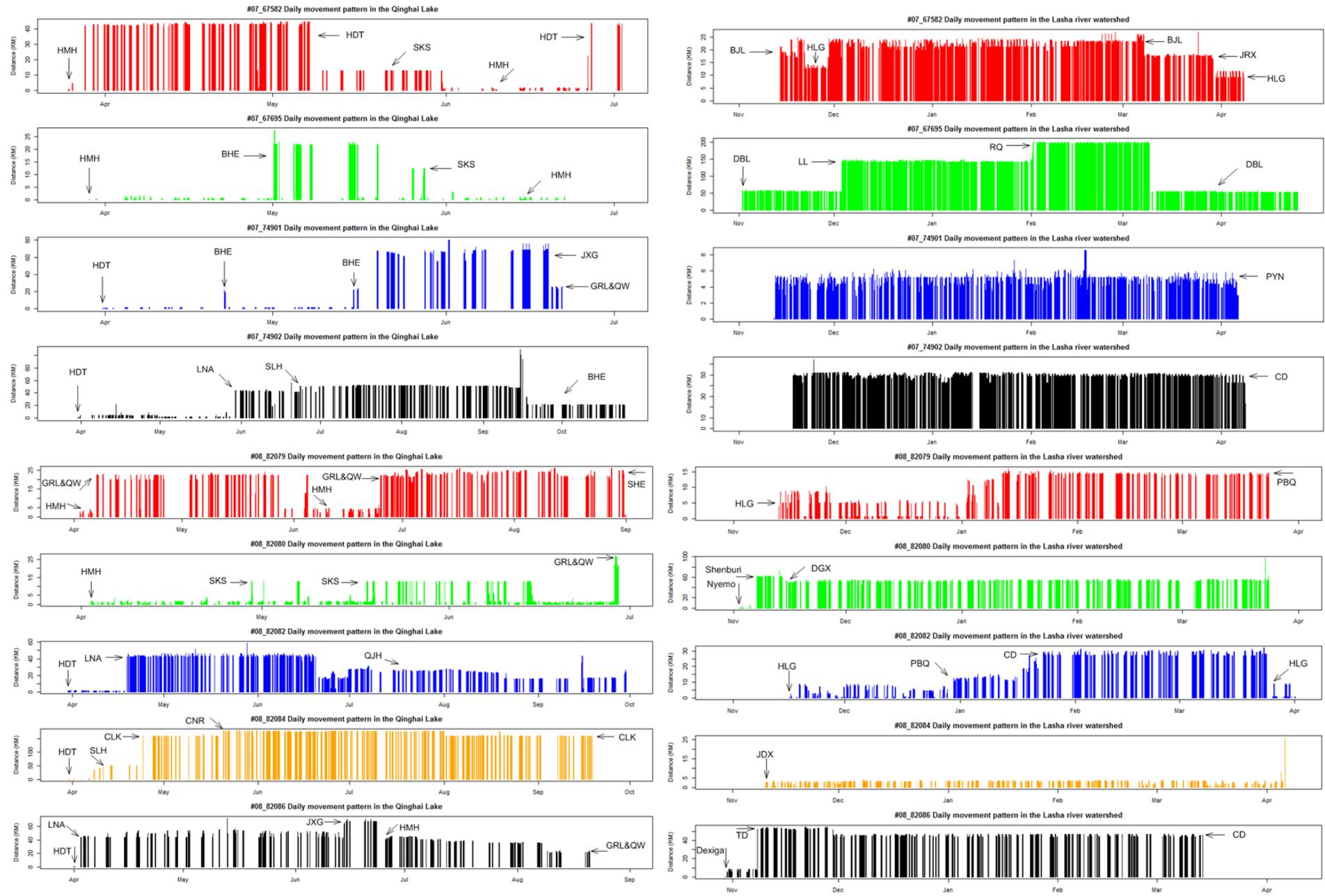
- Daily movement pattern analysis and home range estimation of Bar-headed Goose in breeding site (Qinghai Lake) and Wintering site (Tibet)
  - Daily movement pattern related with behavior (breeding, molting?)
  - Home range in breeding site and Wintering site (comparation?)
  - Speed is slow in breeding site but fast in wintering site (?)

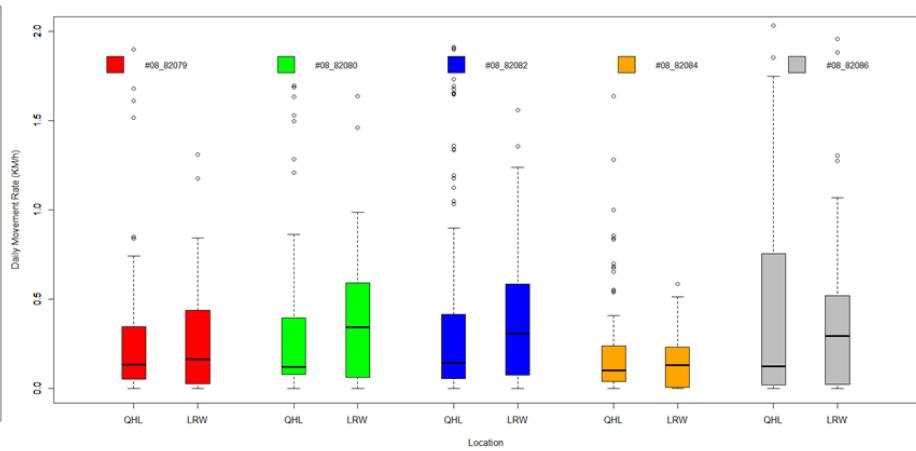
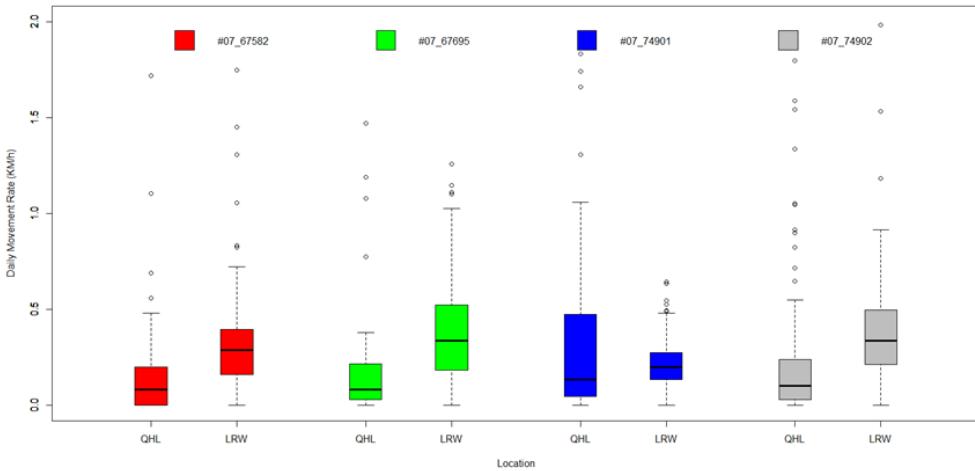
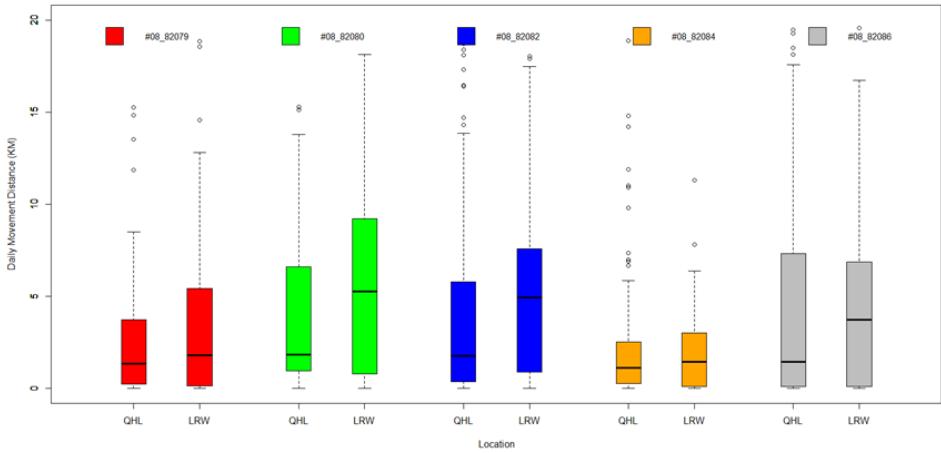
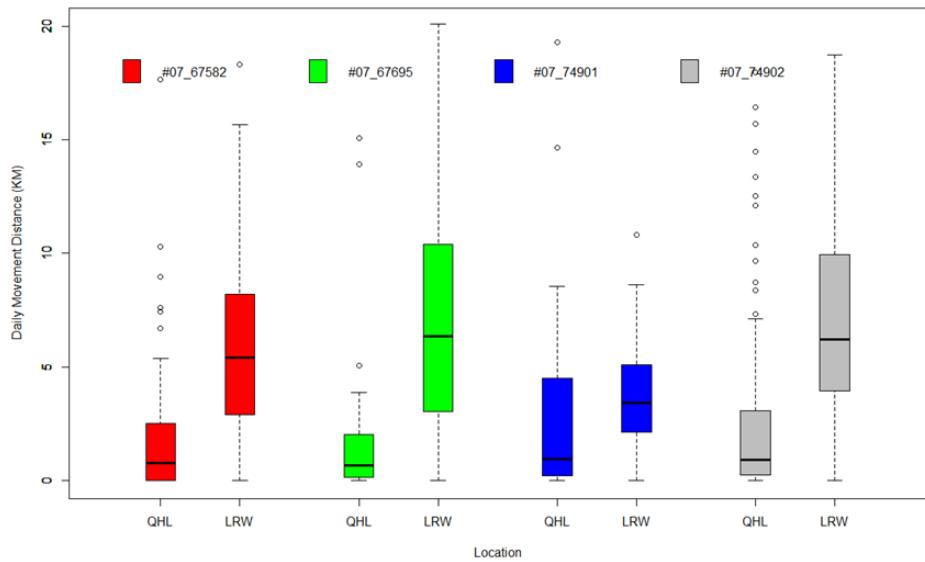
# Daily movement pattern analysis

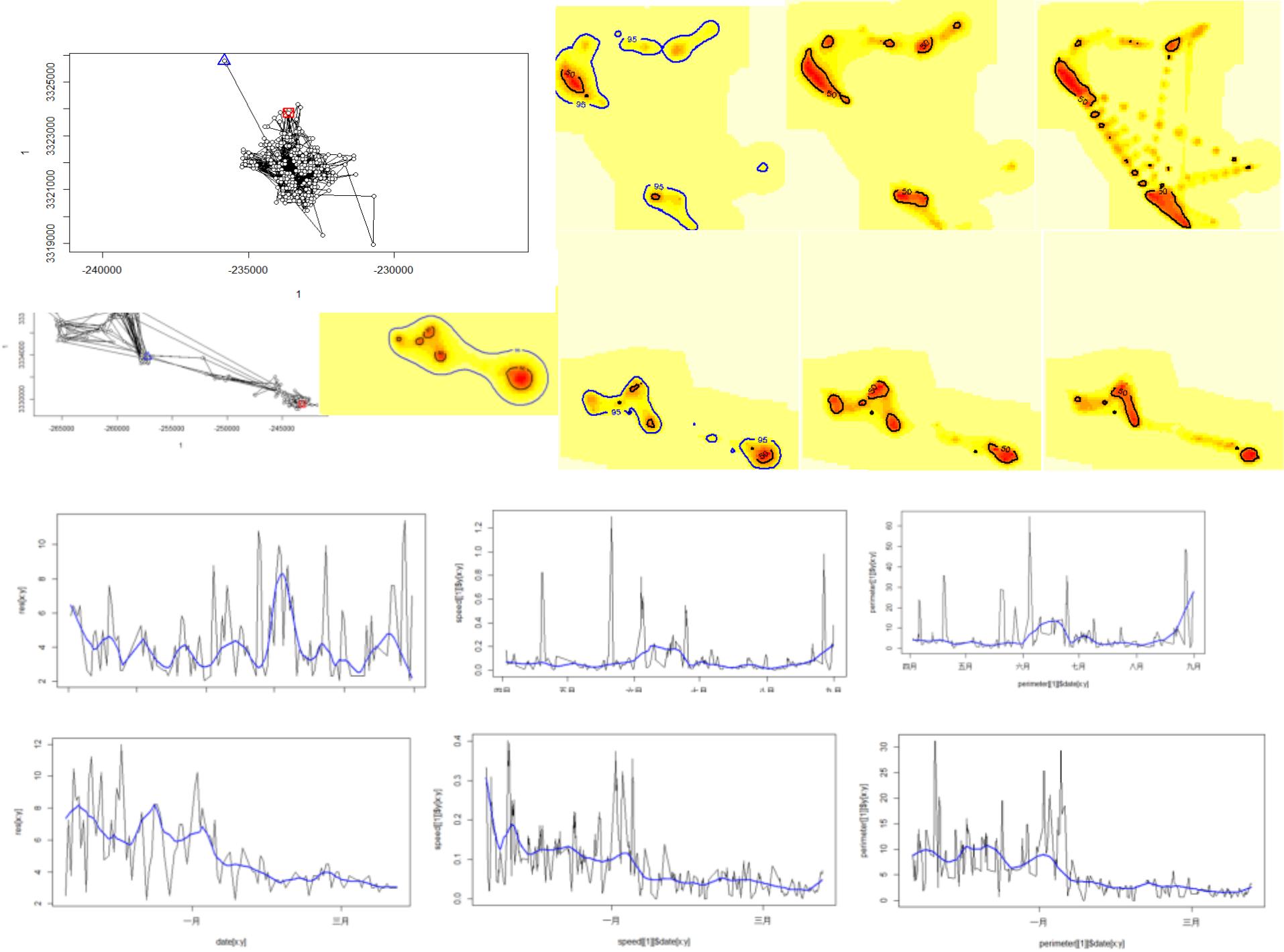
- Methods:
  - Distribution in breeding site and wintering site, arrive time and leave time
  - Home range estimate in breeding site and wintering site by BBMM and BRB method (produce 50-95 area for each bird) and compare home range area in breeding site and wintering site for each bird
  - Intensive home range estimation in breeding site and wintering site by BRB method (produce 50-95 area for each bird)
  - Daily home range estimation and daily speed, perimeter, forage distance by the product kernel algorithm, and slide windows statistic for daily movements.





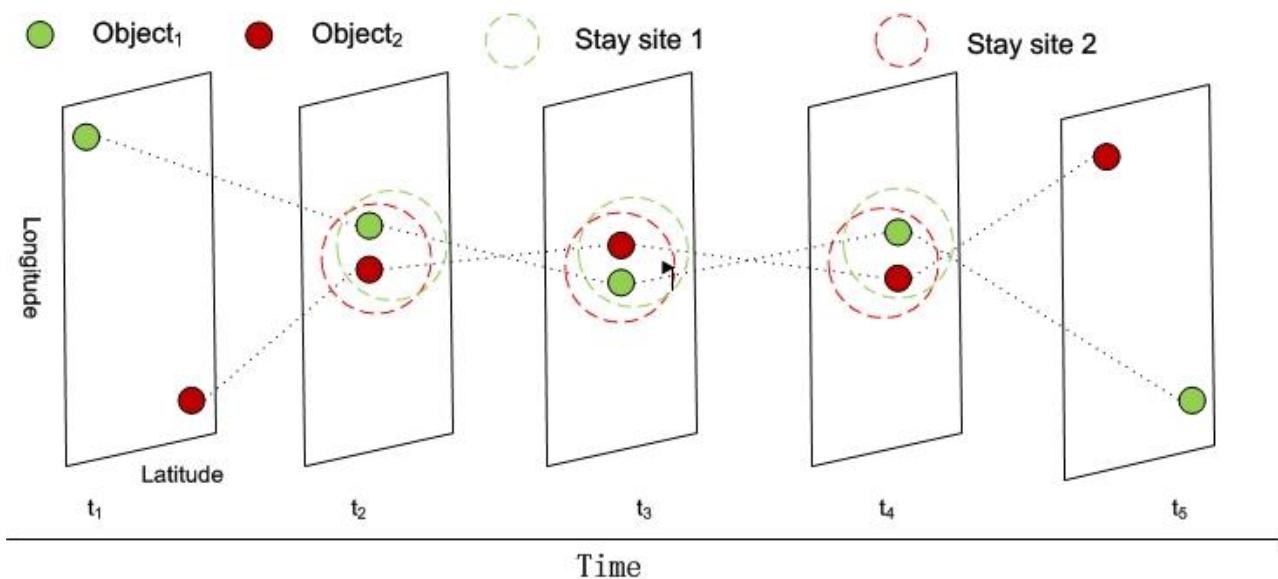




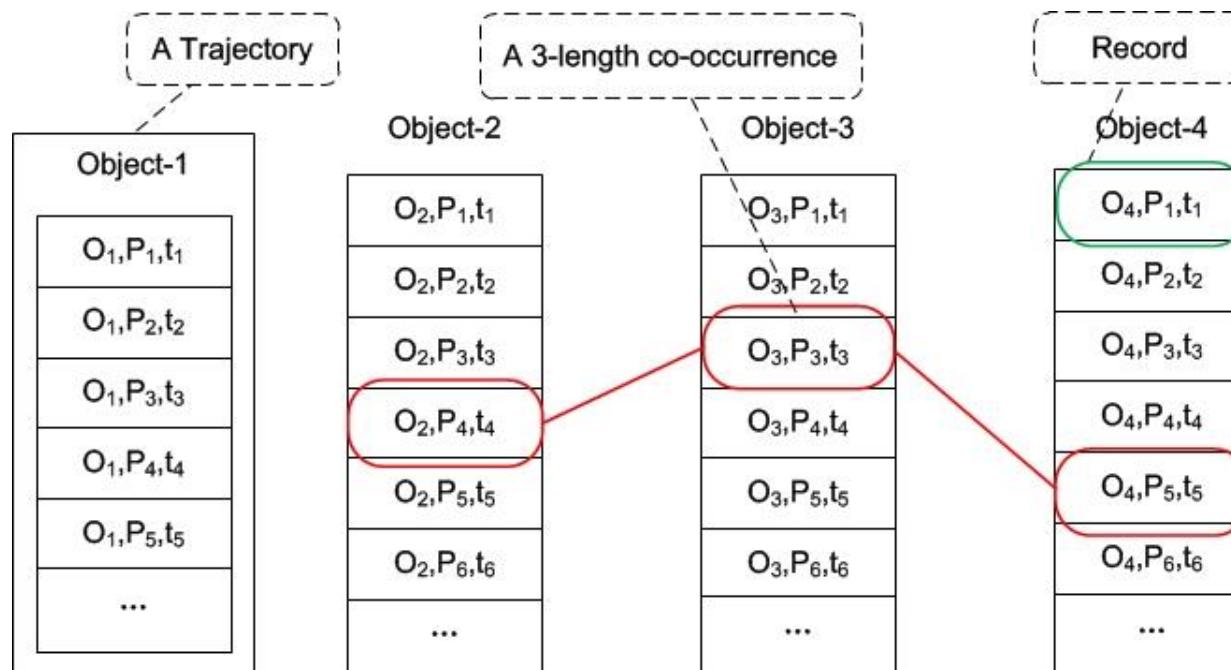


# Fast Mining of Important Co-occurrence Cases on GPS Location History Database

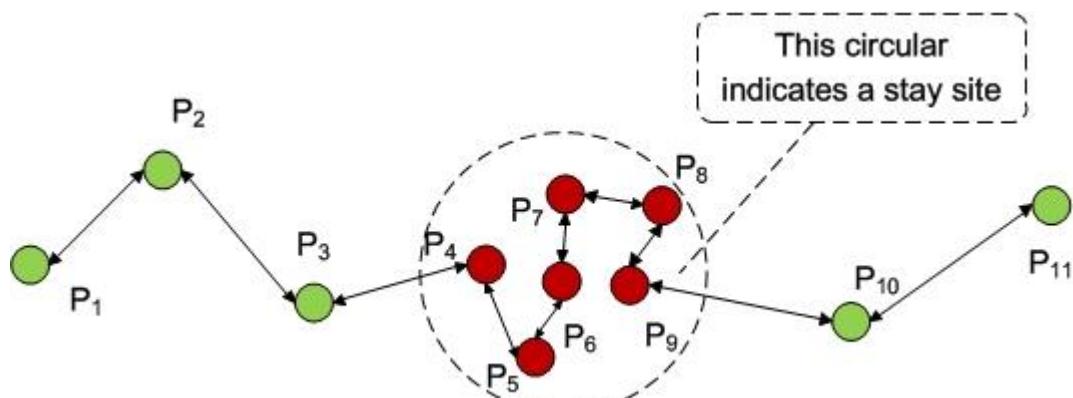
- we propose a concept of important co-occurrence case
- We propose a naïve approach and a stay-site-based approach to discover the important co-occurrence cases on GPS location history
- We conduct experiments on two real world data to verify the correctness and effectiveness of our algorithms.
- Experiment results show the stay-site-based approach outperforms the naïve approach significant.



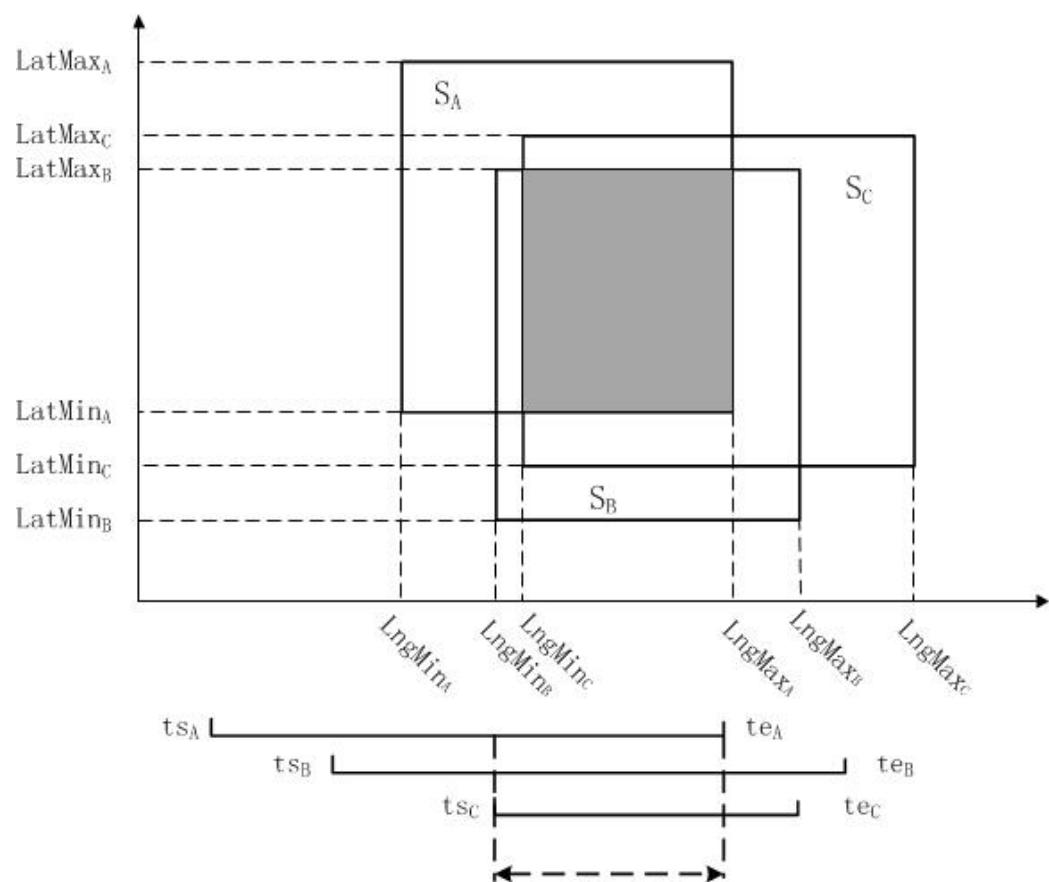
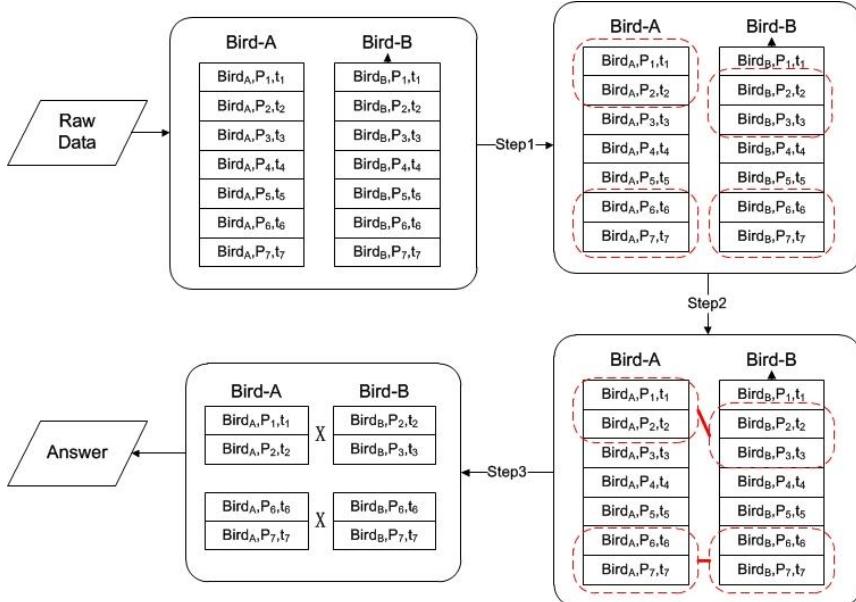
An conceptual demo of an important co-occurrence case.



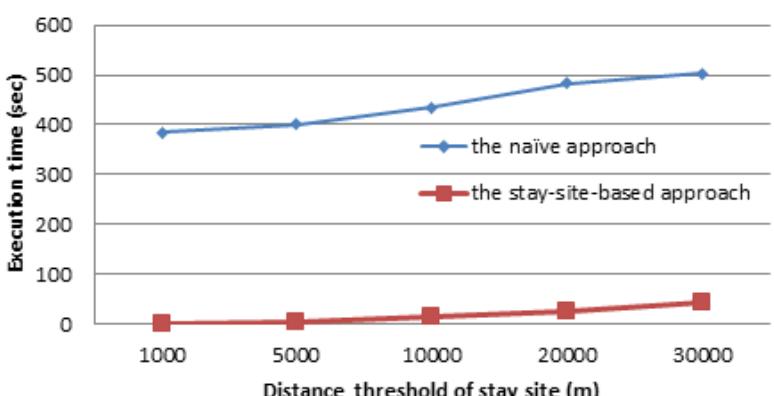
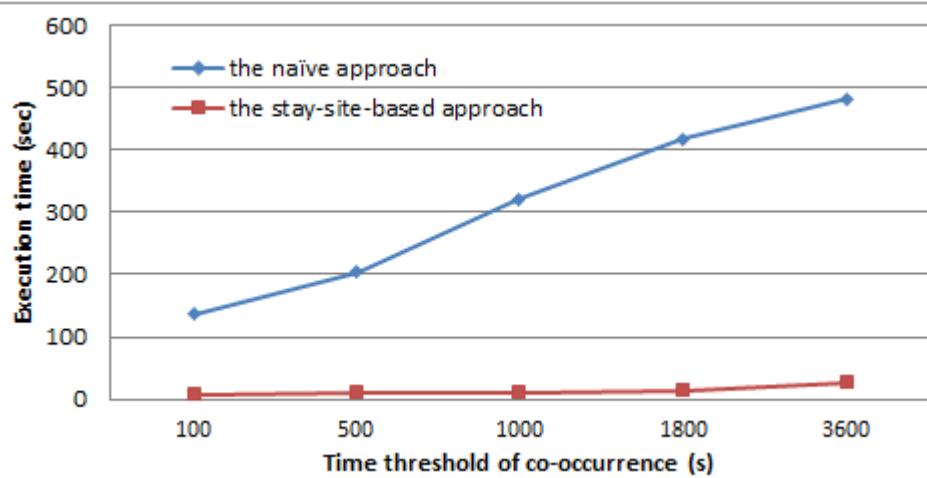
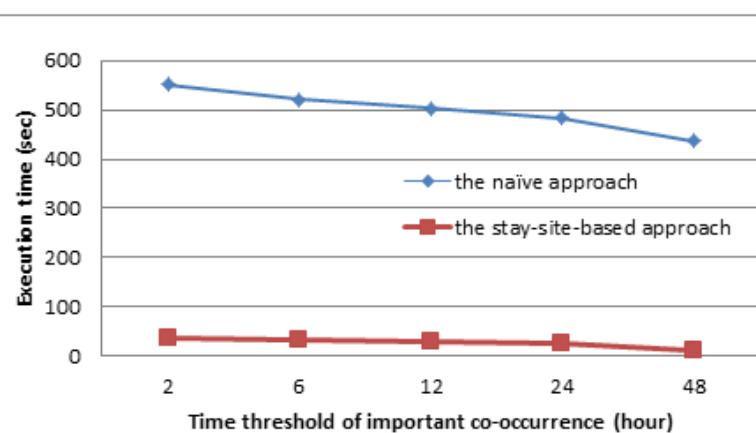
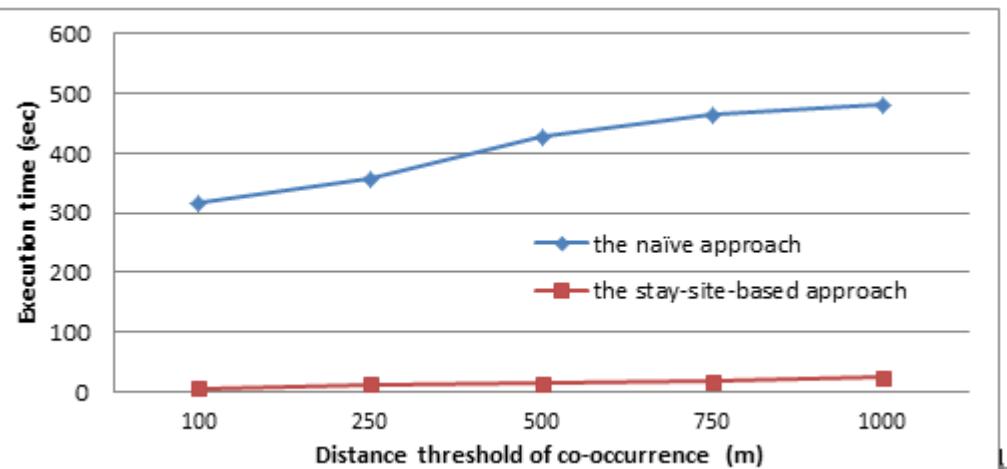
The concept of record, trajectory and co-occurrence case



A stay and a stay site

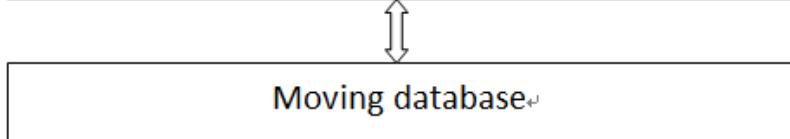
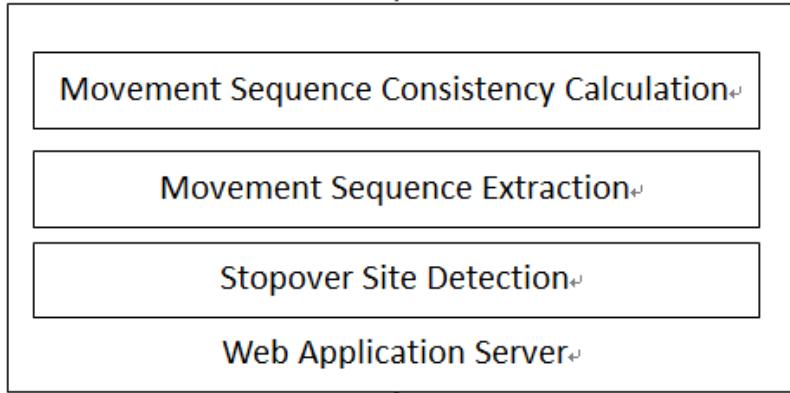
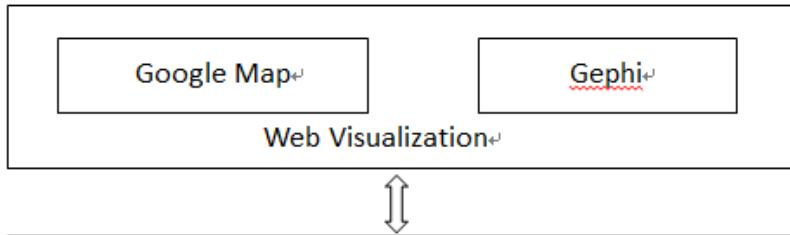


A 3-length stay-site-based co-occurrence case

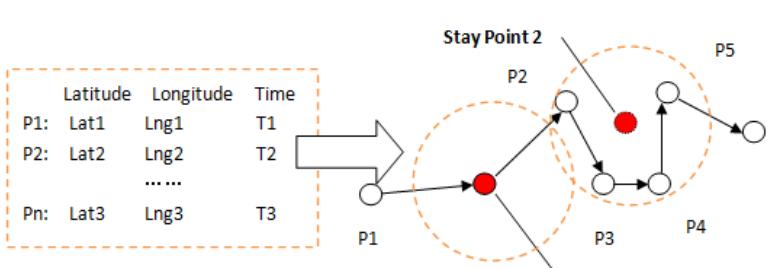


# GPS Location History Data Mining and Anomalous Detection: the Scenario of Bar-headed Geese Migration

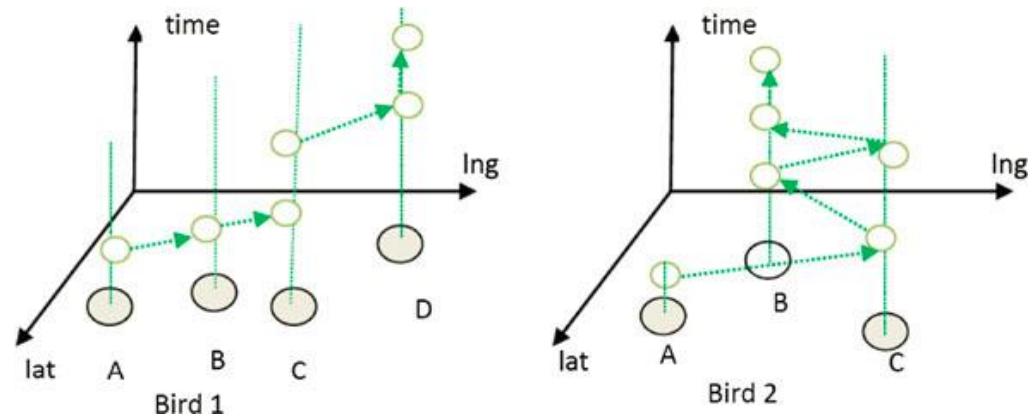
- We propose a new approach to analyze the GPS location history data of migratory birds.
- The stopover sites are first extracted from the location history data of birds, and their movement sequences are generated automatically.
- Then, a consistency calculation method is introduced for calculating the movement sequence consistency degrees among the birds.
- The common movement sequences and uncommon behaviors can be recognized on the basis of consistency.
- Published on Lecture Notes in Electrical Engineering, Springer



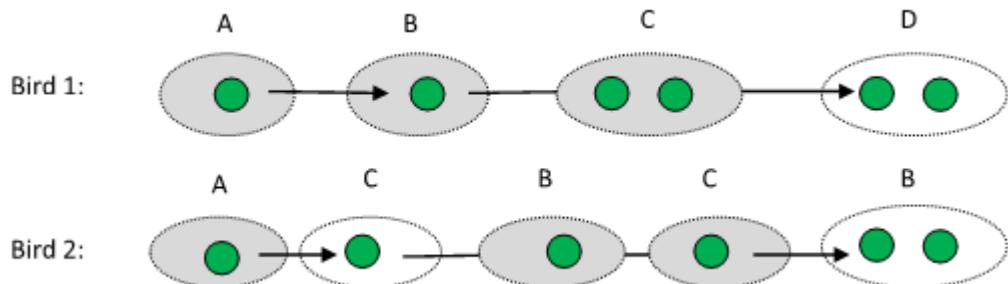
## LHMAN System Architecture



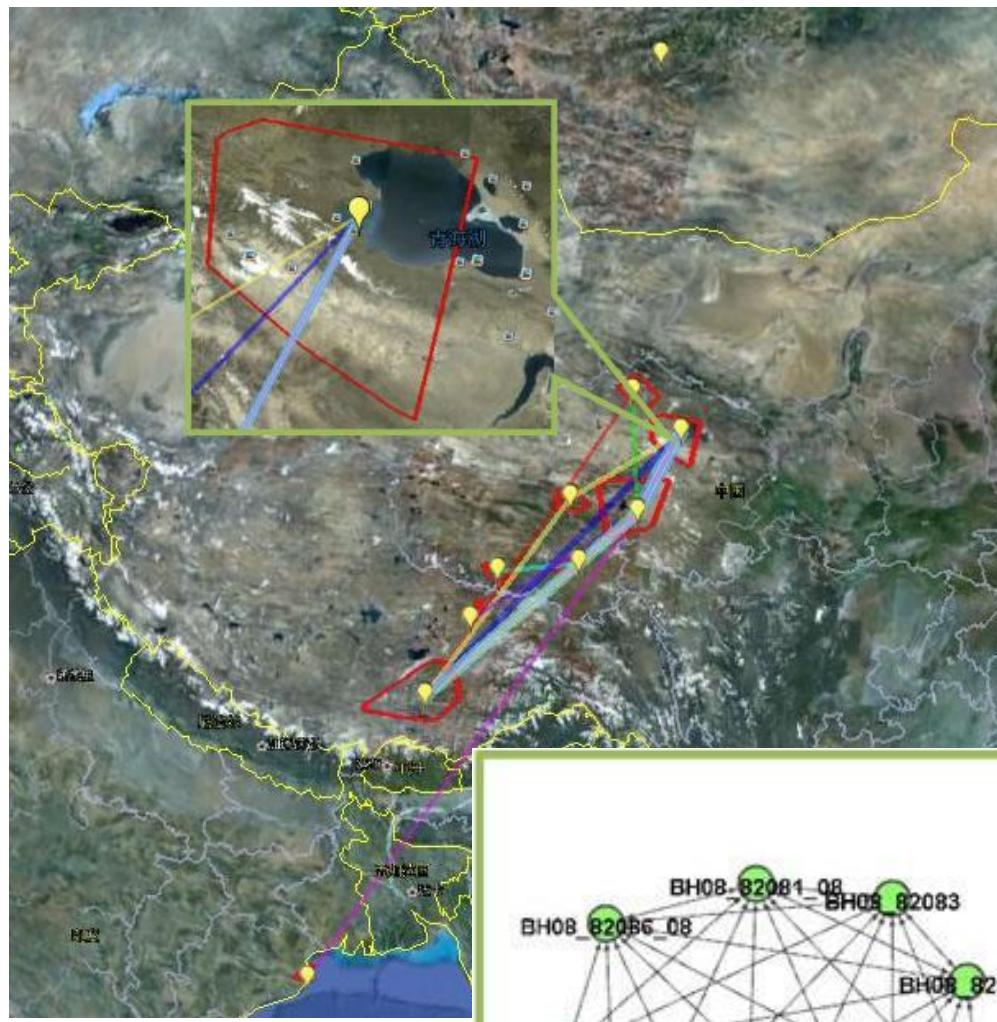
Stopover site



Movement sequence representations

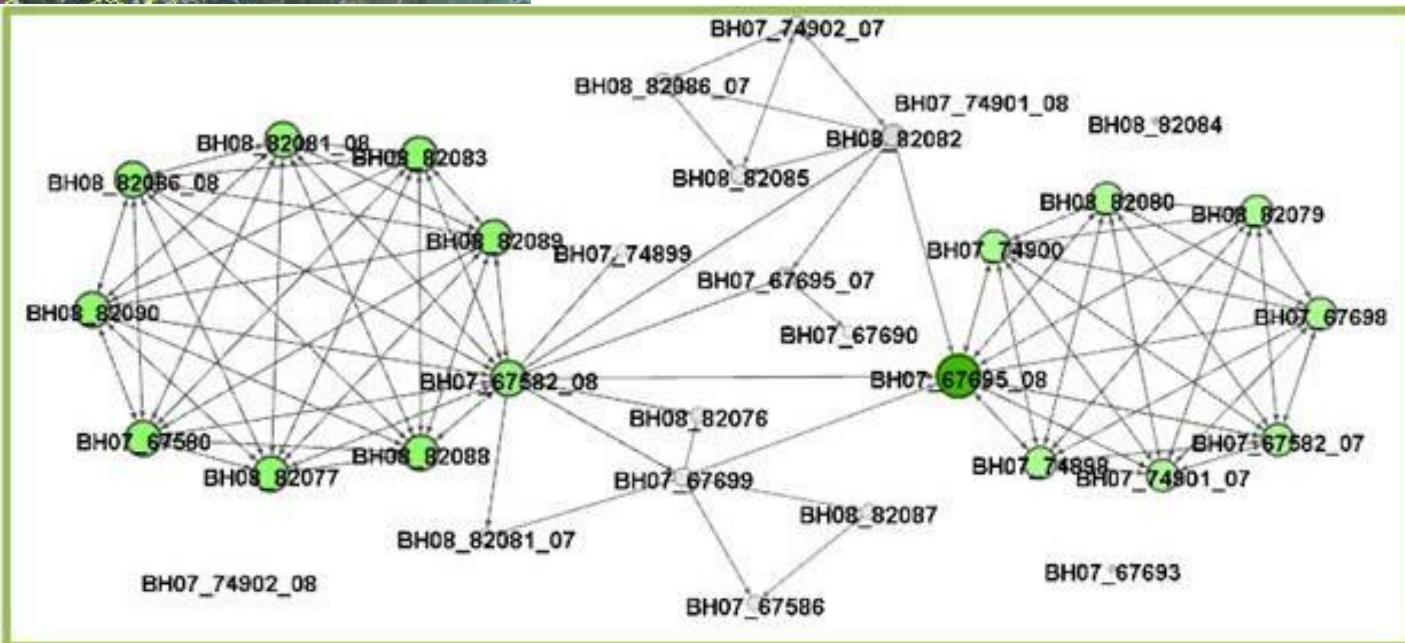


Maximum shared movement sequence



## Stopover sites and migration pathways

## Consistency results of 33 movement sequences

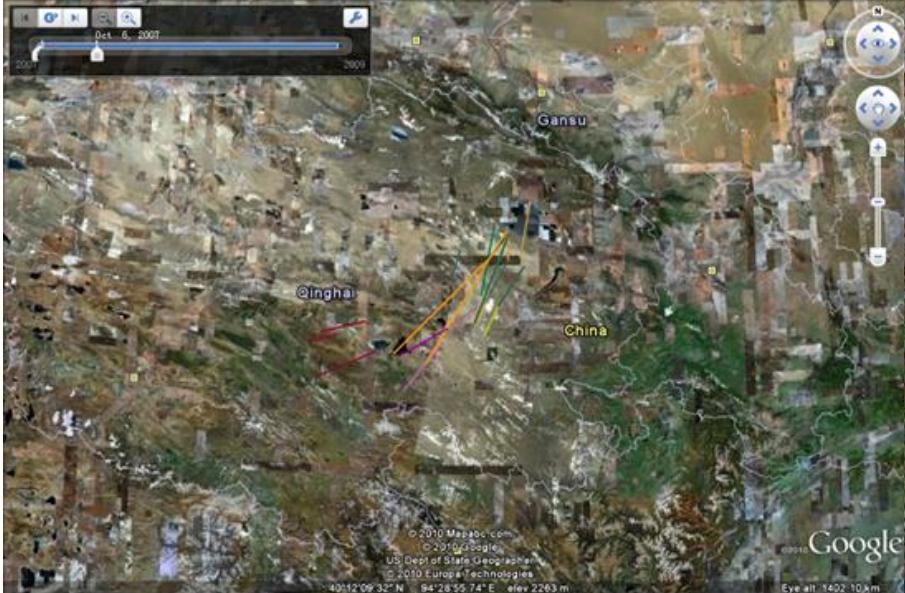


## Discovering Shared Segments of Bar-headed Goose's migratory pathway by Time-based Plane-sweeping Trajectory Clustering

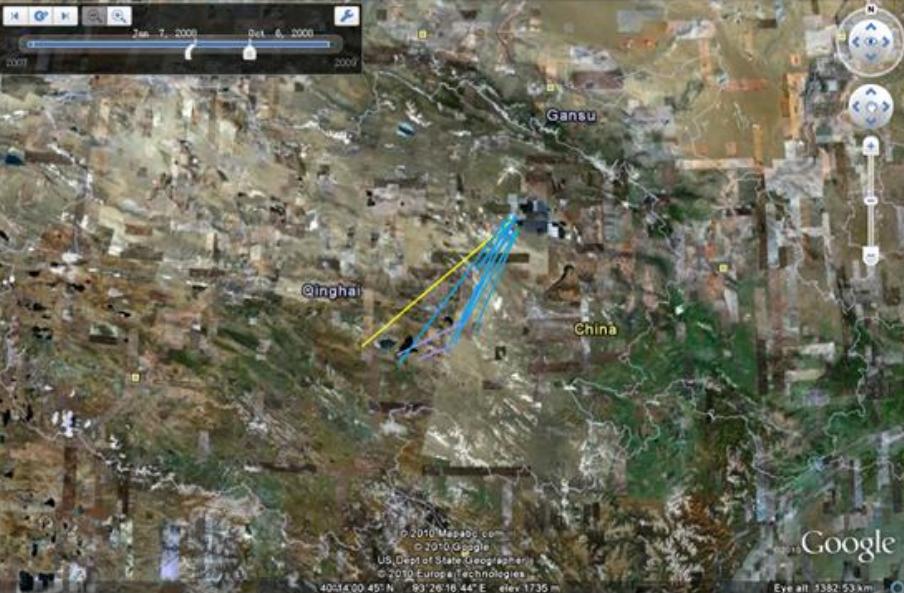
- We propose a new method to discover shared segments on the migratory pathway of the bar-headed geese by time-based plane-sweeping trajectory clustering
- We present findings on the migratory behavior of bar-headed geese determined from this new analytical approach.
- Published on Journal of Information & Computational Science 9: 16 (2012) 5093–510



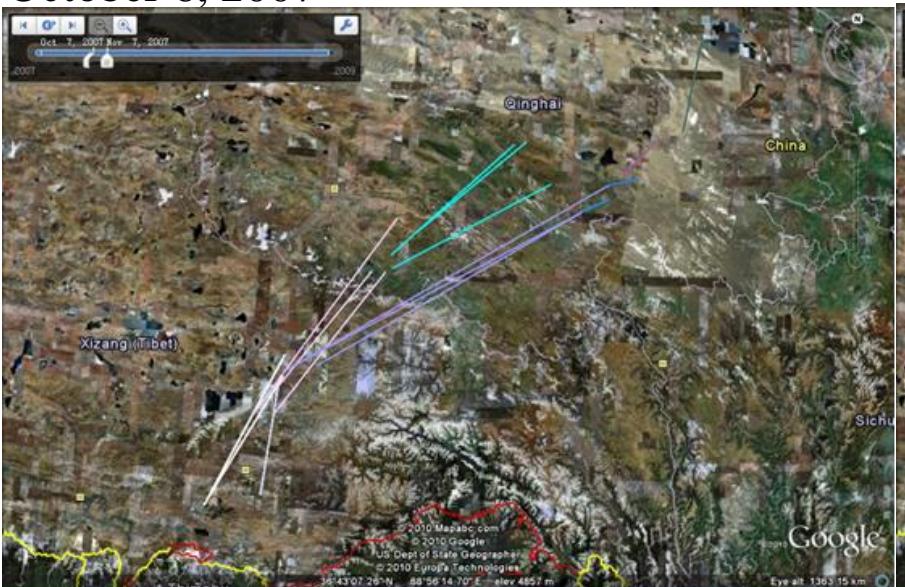
Clustering result of shared segments on migration route



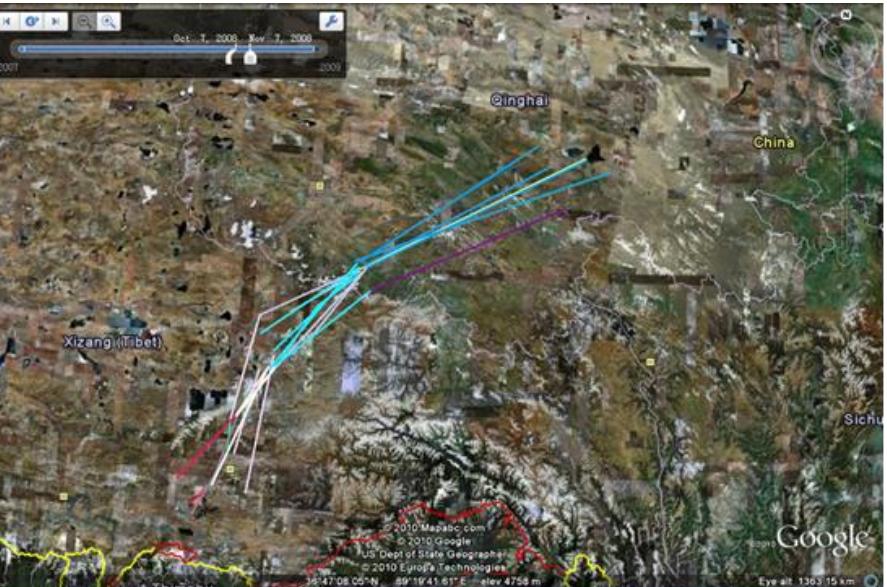
Shared segments from June 7, 2007 to October 6, 2007



Shared segments from June 7, 2008 to October 6, 2008



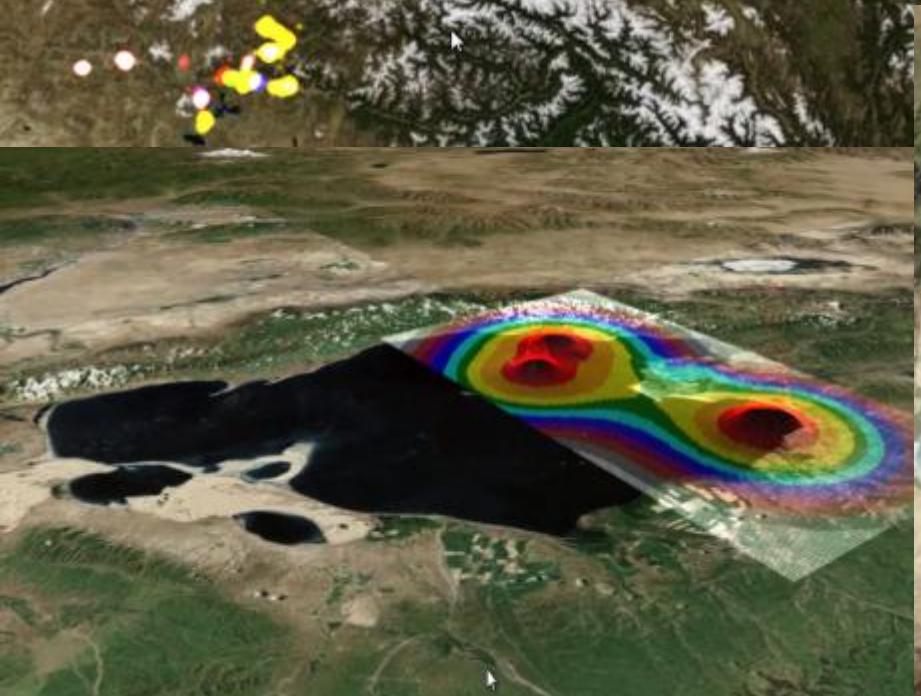
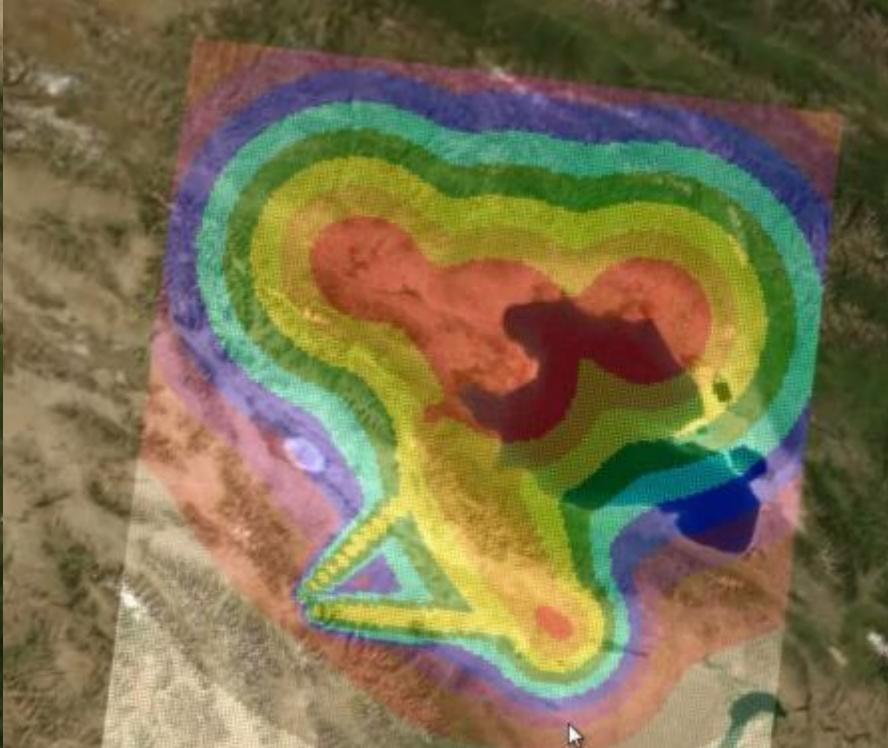
Shared segments from October 7, 2007 to November 7, 2007



Shared segments from October 7, 2008 to November 7, 2008



Clustering result of shared segments in Qinghai Lake region





Pattern charts (Habitat 0)



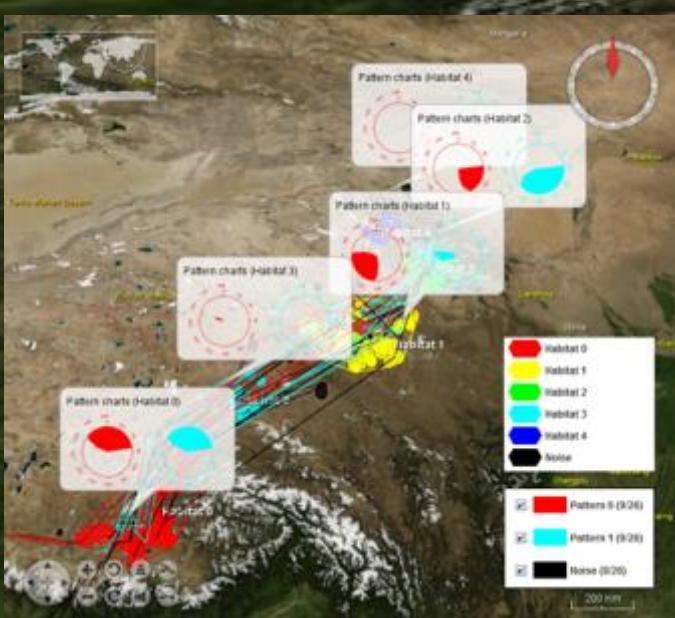
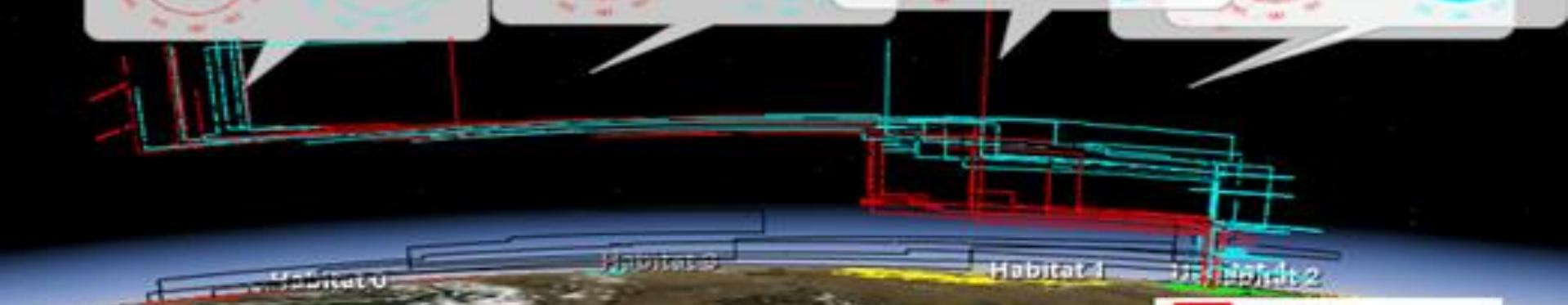
Pattern charts (Habitat 3)



Pattern charts (Habitat 1)



Pattern charts (Habitat 2)



200 Km

青海湖水鸟

159.226.15.225

排序: 目

搜索... P

目  
排序: 数量

鹤形目  
雁形目  
鹳形目  
鹭形目  
鹬形目  
鸥形目  
鸭形目  
鹤形目

39  
24  
12  
5  
3  
3  
2  
1

科  
属  
英文名  
拉丁学名  
上传者

## 鸬鹚岛 2013年

温度

电压

空气湿度

风速

降雨量

盐渍度

电导率

含水量体积比

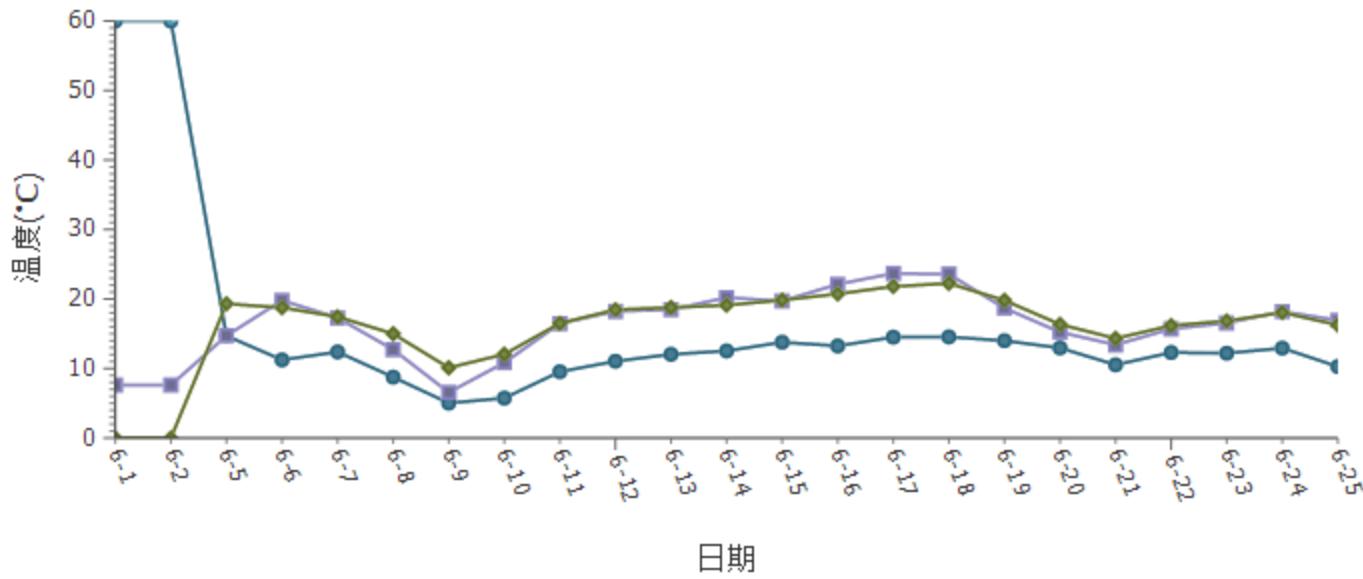


土壤温度



表面温度

空气温度



# Outline

- Introduction
- Database system and consideration
- Data analysis and visualization
- Applications
- Future work

VisualDB DataForge : 数据管理 - Microsoft Internet Explorer

文件(F) 编辑(E) 查看(V) 收藏(A) 工具(T) 帮助(H)

后退(B) 前进(F) 搜索 收藏夹 地图 打印 复制 粘贴 撤消 恢复 转到 链接

地址(1) http://159.226.13.184:8080/console/editor/template1.vpage

**VisualDB DataForge**

**表格编辑**

浏览模式 编辑模式 新增 弹出新增 保存 X 删除 复制 粘贴 撤消 恢复 高级查询 SQL查询

转到 1 第1页 共157页 共1561条记录

	采集者	采集日期	所属大洲	国家	省自治区	地区	区县	具体地点
1		2005-04-29	Asia	Argentina	Rio Negro			San Antonio Oes
2		2001-01-01	Asia	Canada	Ontario			
3		2001-01-01	Asia	Canada	Ontario			
4		2005-01-24	Asia	Argentina	Rio Negro			
5		2000-11-17	Asia	Sweden	Västernorrland			
6	Jan T. Lifjeld	2005-06-25	Asia	Norway	Romsdalsfjord			
7	Jan T. Lifjeld	2005-06-25	Asia	Norway	Romsdalsfjord			
8	A. J. Baker	1981-11-22	Asia	Argentina	Buenos Aires			
9	A. J. Baker	1981-11-22	Asia	Argentina	Buenos Aires			
10	Louheed, S C	1988-01-01	Asia	Argentina	Córdoba			

**从表信息**

转到 1 第1页 共1页 共1条记录

	采集号	鸟类名	基因名/Mark名	拼接好的序列	记录修改时间
1	7	家麻雀	COI-5P	-----ACAAAGACATTGGCACCTGTACCTAATCTCGCGCATGAGCCGGG	2007-05-22

**基因数据表**

查询

基因数据表 关键字... 全表检索

采集号	鸟类名	基因名/Mark名	记录修改时间
1560	暗绿柳莺	COI-5P	2008年07月31日
1559	暗绿柳莺	COI-5P	2008年07月31日
1558	暗绿柳莺	COI-5P	2008年07月31日
1557	暗绿柳莺	COI-5P	2008年07月31日
1556	暗绿柳莺	COI-5P	2008年05月22日
1555	暗绿柳莺	COI-5P	2008年06月07日
1554	暗绿柳莺	COI-5P	2008年07月31日
1553	暗绿柳莺	COI-5P	2008年07月31日

共 1660 条记录, 当前第 1 条到第 8 条 12345 GO> 1

We collect gene data, sample data and other relevant data of wild birds around Qinghai Lake and build a database as the standard data for follow-up species identification.

## 基因序列鉴别

## 基因序列鉴别

请输入基因序列  
(FASTA格式, >80) :

```
ATAGTAGGTACCGCCCTAACGTCTCCTCATTGAGCAGAGCTGGGACAACCCGGAGCCCTCAGGAGACATCAAGTATAACGTAGTCGTCACAGGCCATGCCTCGTAATAATTTCCTTAGTTTACCAATTATAATTGGGGATTGGTAACTGACTAGTCTCTAATAATTGGAGCCGGACATGGCATTCCCACGAATAAACATAAGCTCTGACTACTACCCCCATCCTCCCTACTAGCATCCTCTACAGTAGAAGCAGGAGTAGGTACAGGCTAACAGTATAACCAACACTAGCTGGCACCTGGCCACGCCGGAGCCTCAGTAGACCTAGCAATCTCTCCCTACCTGGCCGGCATCTCCTCAATTAGGTGCAATCAACTTATCACAAACAGCAATCAACATAACCTCTGCCTTACAGTACCAAAACCCCCCTGTTGTTGATCCGTAATCACCGC/TGCTCTGCTCTGTCACTACCAGTCTAGCTGAGGTATTACAATGCTCTGACAGACAAACCTCAATACTACATTCTTGACCCCTGCAGGAGGAGACCCAGTCCTATACCAACATCTTGATTCTCGGACACCCAGAAGTATATA
```

从文件中输入：

浏览...

显示记录：

前10条

搜索

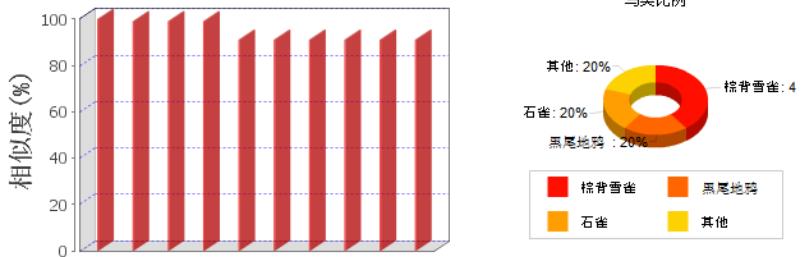
清空

## 基因序列鉴别

## 鉴别结果为:

基因序号	相似度	拉丁文名	中文名	英文名
1590	100%	Montifringilla blanfordi	棕背雪雀	Blanford's snow finch

## 物种鉴别分析结果统计图



## 前10条记录

基因序号	拉丁文名	中文名	英文名	相似度	本条记录详情
1590	Montifringilla blanfordi	棕背雪雀	Blanford's snow finch	100%	<a href="#">详情</a>
1591	Montifringilla blanfordi	棕背雪雀	Blanford's snow finch	99%	<a href="#">详情</a>
1592	Montifringilla blanfordi	棕背雪雀	Blanford's snow finch	99%	<a href="#">详情</a>
1593	Montifringilla blanfordi	棕背雪雀	Blanford's snow finch	99%	<a href="#">详情</a>
1452	Podoces hendersoni	黑尾地鸦	Mongolian Ground Jay	91%	<a href="#">详情</a>
1451	Petronia petronia	石雀	Rock Petronia	91%	<a href="#">详情</a>
1452	Podoces hendersoni	黑尾地鸦	Mongolian Ground Jay	91%	<a href="#">详情</a>
1451	Petronia petronia	石雀	Rock Petronia	91%	<a href="#">详情</a>
1565	Montifringilla taczanowskii	白腰雪雀	White-rumped Snowfinch	91%	<a href="#">详情</a>
1561	Montifringilla taczanowskii	白腰雪雀	White-rumped Snowfinch	91%	<a href="#">详情</a>

拉丁文名:

Montifringilla blanfordi

中文名字:

棕背雪雀

英文名字:

Blanford's snow finch

查询**BOLD**上该物种的详细信息基因记录ID: 1590 1591 1592 1593 [全部详情](#)

&gt;&gt;&gt;&gt;&gt;&gt; 基因ID为1590的详细数据记录 &lt;&lt;&lt;&lt;&lt;&lt;

## 基因数据记录详细信息

基因记录ID: 1590

GenBank Accession: FJ624121

PCR引物1: null

PCR引物1序列: null

PCR引物2: null

PCR引物2序列: null

测序引物1: null

测序引物1序列: null

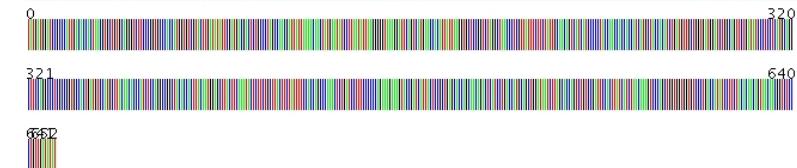
测序引物2: null

测序引物2序列: null

## 核酸序列

总长度: 852 碱基A: 177 碱基C: 198 碱基G: 114 碱基T: 163 未明确: 0  
 ATAGTAGGTACCGCCCTAACGTCTCCTCATTGAGCAGAGCTGGGACAACCCGGAGCCCTCTAGGAGACGATCAAGTATAAACGTAGTCACGCCATGCCTCGIAATAATTCTCTCATAGTATAACCAATTATAATTGGGGATTGGTAACTGACTAGTCTCTCTAATAATTGGAGCACCGGACATGGCATTCCCACGATAAACACATAAGCTCTGACTACTACCCCATCTTCTCTCTACTAGATCCTCTACAGTAGAGCAGGAGTAGGACAGGGCTGACAGTACCTCTCTCCCTACACTGGCCGCATCTCTCAATCTTAGGTGCAATCAACTTATCACAAACAGCAATCAACATAAACCTCTGCCCTTCACAGTACCCAACCCCCCTGTTGTTGATTCGTAATCACCGCAAGTGTCTCTGTCACTACCCAGTCTAGCTGAGGTATTACAATGCTCTGACAGACCGTAACCTCAACTACATTCITGATTCTCGGACACCCAGAAGTATATA

## 基因条形码样例



## 标本数据详细信息

本条数据记录结束

拉丁文名: Montifringilla blanfordi 中文名字: 棕背雪雀 英文名字: Blanford's snow finch

查询**BOLD**上该物种的详细信息

地图 卫星 地形



# 基因序列鉴别

该物种名称：

拉丁文名:	中文名字:	英文名字:
Montifringilla blanfordi	棕背雪雀	Blanfordi' snow finch

从BOLD共查询到30条相关记录: [下载](#)

记录ID	processid	taxID	样本id	catalognum	机构	详情
1113378	GBIR1502-09	119780	EU382120	EU382120	Mined from GenBank, NCBI	<a href="#">详情</a>
1113379	GBIR1503-09	119780	EU382119	EU382119	Mined from GenBank, NCBI	<a href="#">详情</a>
1113380	GBIR1504-09	119780	EU382118	EU382118	Mined from GenBank, NCBI	<a href="#">详情</a>
1113381	GBIR1505-09	119780	EU382117	EU382117	Mined from GenBank, NCBI	<a href="#">详情</a>
1113382	GBIR1506-09	119780	EU382116	EU382116	Mined from GenBank, NCBI	<a href="#">详情</a>
1113383	GBIR1507-09	119780	EU382115	EU382115	Mined from GenBank, NCBI	<a href="#">详情</a>
1113384	GBIR1508-09	119780	EU382114	EU382114	Mined from GenBank, NCBI	<a href="#">详情</a>
1113385	GBIR1509-09	119780	EU382113	EU382113	Mined from GenBank, NCBI	<a href="#">详情</a>
1113386	GBIR1510-09	119780	EU382112	EU382112	Mined from GenBank, NCBI	<a href="#">详情</a>
1113387	GBIR1511-09	119780	EU382111	EU382111	Mined from GenBank, NCBI	<a href="#">详情</a>
1113388	GBIR1512-09	119780	EU382110	EU382110	Mined from GenBank, NCBI	<a href="#">详情</a>
1113389	GBIR1513-09	119780	EU382109	EU382109	Mined from GenBank, NCBI	<a href="#">详情</a>
1113390	GBIR1514-09	119780	EU382108	EU382108	Mined from GenBank, NCBI	<a href="#">详情</a>
1113391	GBIR1515-09	119780	EU382107	EU382107	Mined from GenBank, NCBI	<a href="#">详情</a>
1113392	GBIR1516-09	119780	EU382106	EU382106	Mined from GenBank, NCBI	<a href="#">详情</a>
1113393	GBIR1517-09	119780	EU382105	EU382105	Mined from GenBank, NCBI	<a href="#">详情</a>
1113394	GBIR1518-09	119780	EU382104	EU382104	Mined from GenBank, NCBI	<a href="#">详情</a>

Rollover  
 Show Internal Data  
 Taxonomy Colorize  
 Annotation Colorize  
 Colorize Branches  
**Display Data:**  
 Node Name  
 Taxonomy Code  
 Taxonomy Name  
 Prot/Gene Symbol  
 Prot/Gene Name  
 Annotation  
 Confidence Value

**Click on Node to:**  
 Display Node Data  
 Zoom:  
V+ F + V-

Order Subtrees  
 Uncollapse All  
 Search:

[VisualDB DataForge - 数据管理 - Microsoft Internet Explorer](#)  
 地址栏: http://158.226.13.194:8080/console/editor/template1.page

**表编辑**  
 表模式: 表格模式 | 列模式 | 导出新表 | 保存 | X | 复制 | 打印 | 撤销 | 恢复 | 高级查询 | SQL查询  
 表名: 第1页共157页 共1561条记录

采集者	采集日期	所属大洲	国家	省自治区	地区	区县	具体地址
1	2005-04-29	Asia	Argentina	Rio Negro			San Antonio Oeste
2	2001-01-01	Asia	Canada	Ontario			
3	2001-01-01	Asia	Canada	Ontario			
4	2005-03-24	Asia	Argentina	Rio Negro			
5	2000-11-17	Asia	Sweden	Vasterbotten	Vasterbotten		Hosjö, Umeå
6	Jan T. Lifeld	2005-06-25	Asia	Norway	Iriegland		Pierneeskad
7	Jan T. Lifeld	2005-06-25	Asia	Norway	Iriegland		Pierneeskad
8	A. J. Baker	1981-11-22	Asia	Argentina	Buenos Aires		La Plata, near Bs
9	A. J. Baker	1981-11-22	Asia	Argentina	Buenos Aires		La Plata, near Bs
10	Loughheed, S.C.	1989-01-01	Asia	Argentina	Catamarca		Belen, Catamarca

**从表组**  
 表模式: 表格模式 | 列模式 | 导出新表 | 保存 | X | 复制 | 打印 | 撤销 | 恢复 | 高级查询 | SQL查询  
**基因信息表 (基础)**

Barcode 采集号	高商品名	基因名	测定序号
7 7	棕背雪雀	COI-5P	ACAAAGACATTGGCACCCTGTAATCTTGGCGCATGAGCGGGG

**采集日期:** 1998-03-04  
**采集人:** John  
**生境:** 沼泽

**标本存放信息**

所属单位:	博物馆
保藏位置:	X-117
保藏方式:	假剥制
鉴定人:	Mary
鉴定日期:	2009-02-03

**图片数据信息:**

图片ID:	PIC0128
摄影者:	Dan
拍摄地点:	青海湖北岸
拍摄时间:	2009-01-01
备注:	无

鸟类图片鉴别

从文件中输入： 浏览

显示记录： 搜索 清空

---







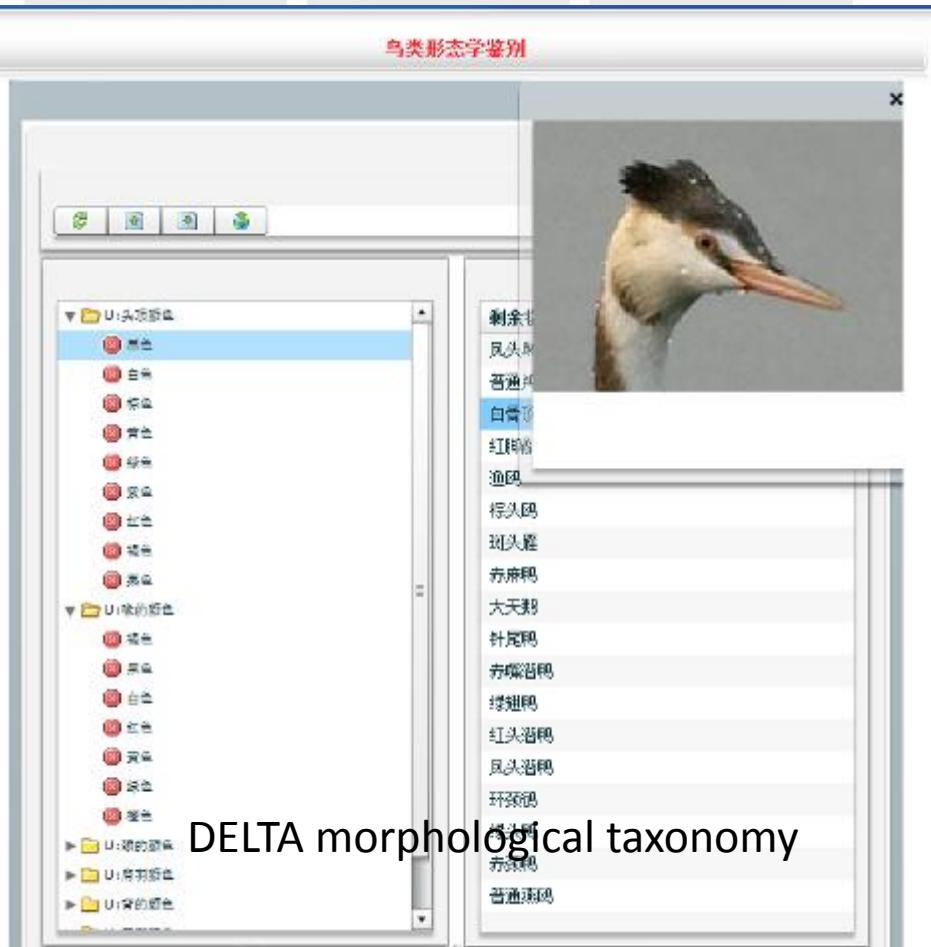
正圖錄

白额雁.jpg  
拉丁学名字: Anser albifrons  
相似:99.999977%  
红沙嘴.jpg  
拉丁学名字: Gallinago solitaria  
相似:56.672573%  
牛背鹭.jpg  
拉丁学名字: Bubulcus ibis  
相似:54.771469%

Content-based image retrieval



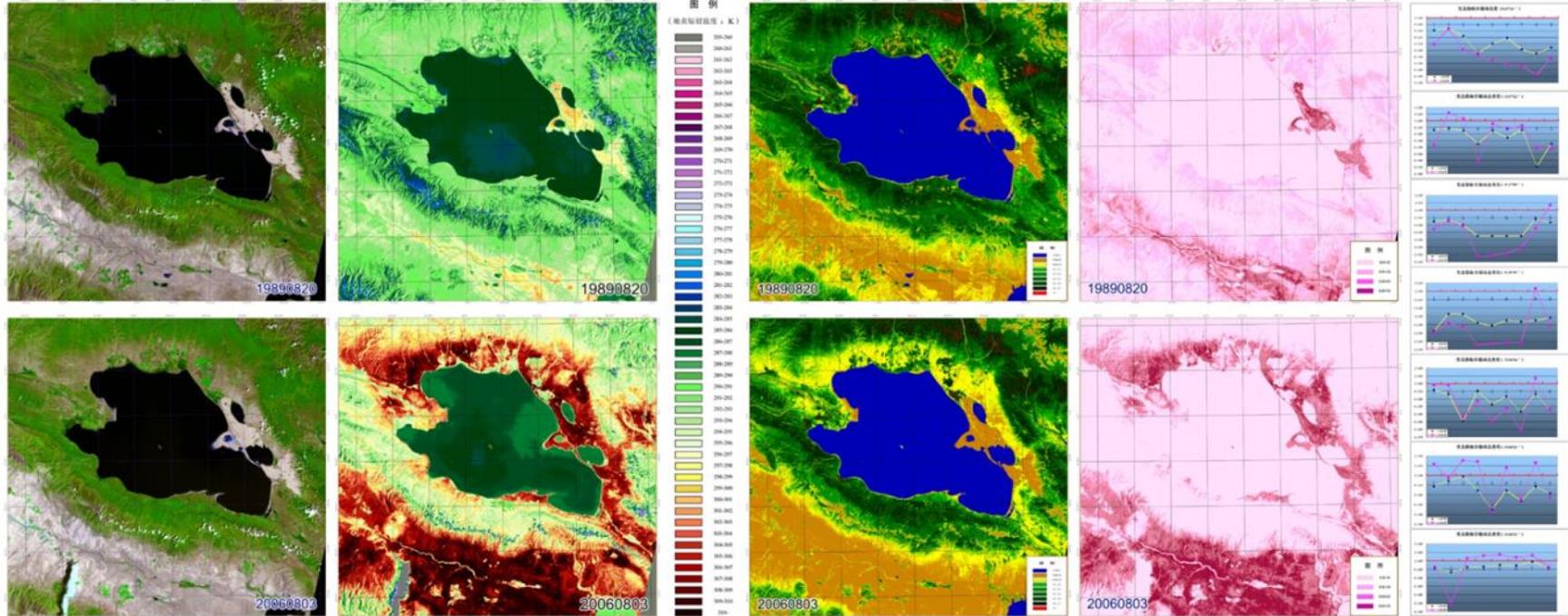
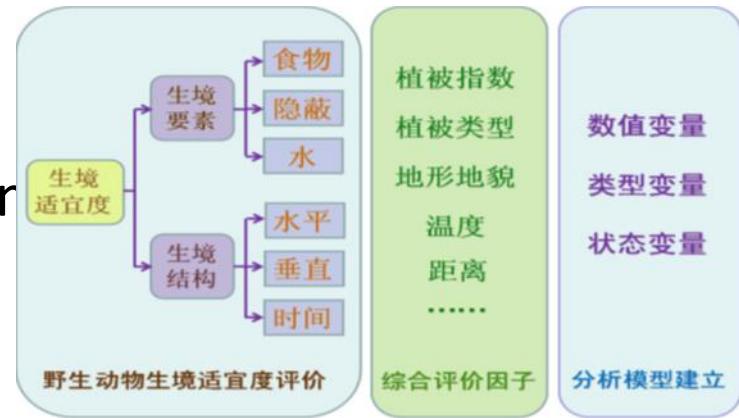
鸟类形态学鉴别



# DELTA morphological taxonomy

# Model-based analysis and prediction of spatial distribution patterns

- GPS-based Telemetry Data
- Ecological Feature extracted from Remote Sensing
- Maxent Model
- Visualization on GIS

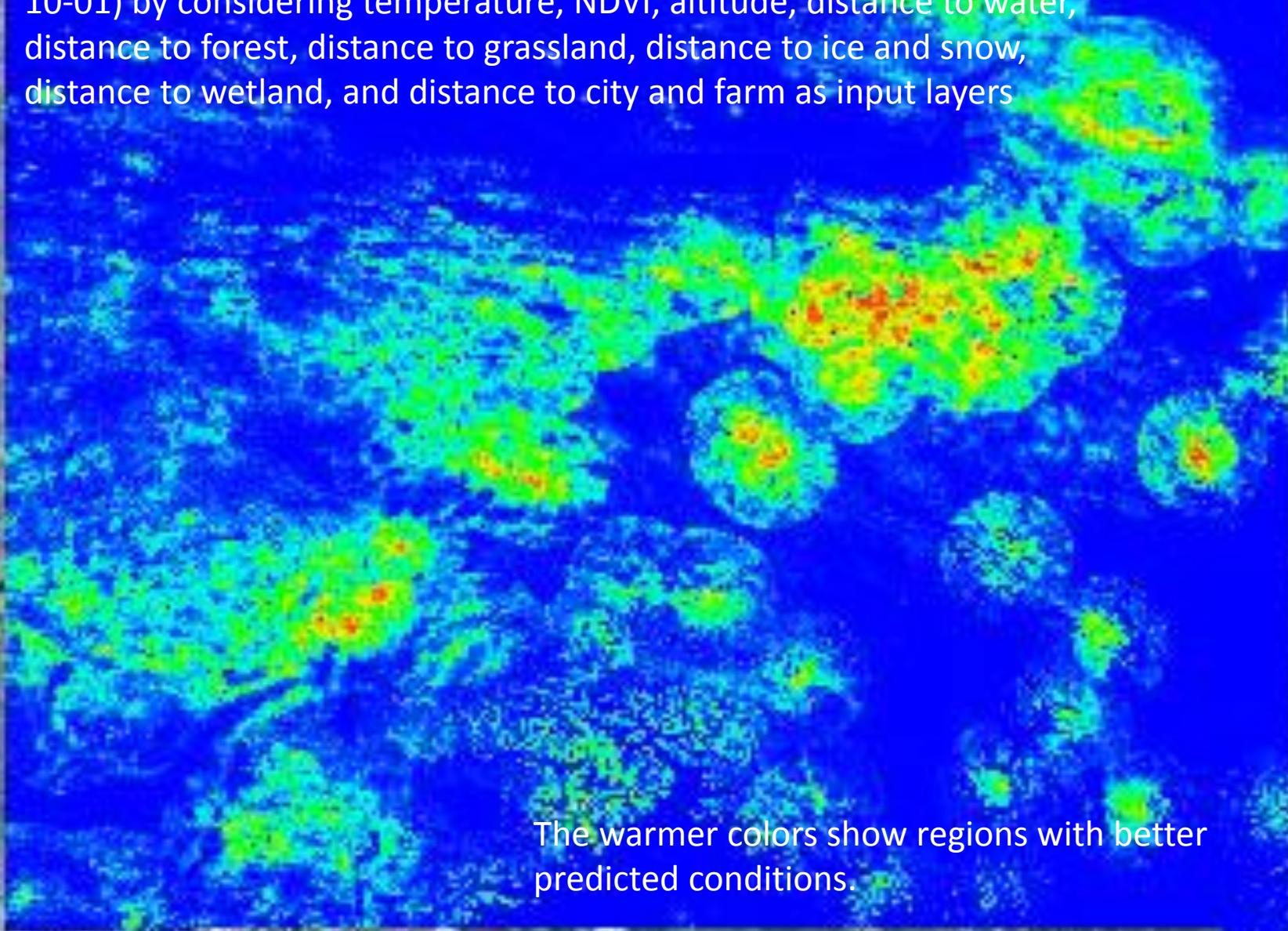


# Ecological Feature extracted from Remote Sensing Images

TABLE IV. LIST OF ECOLOGICAL FEATURE

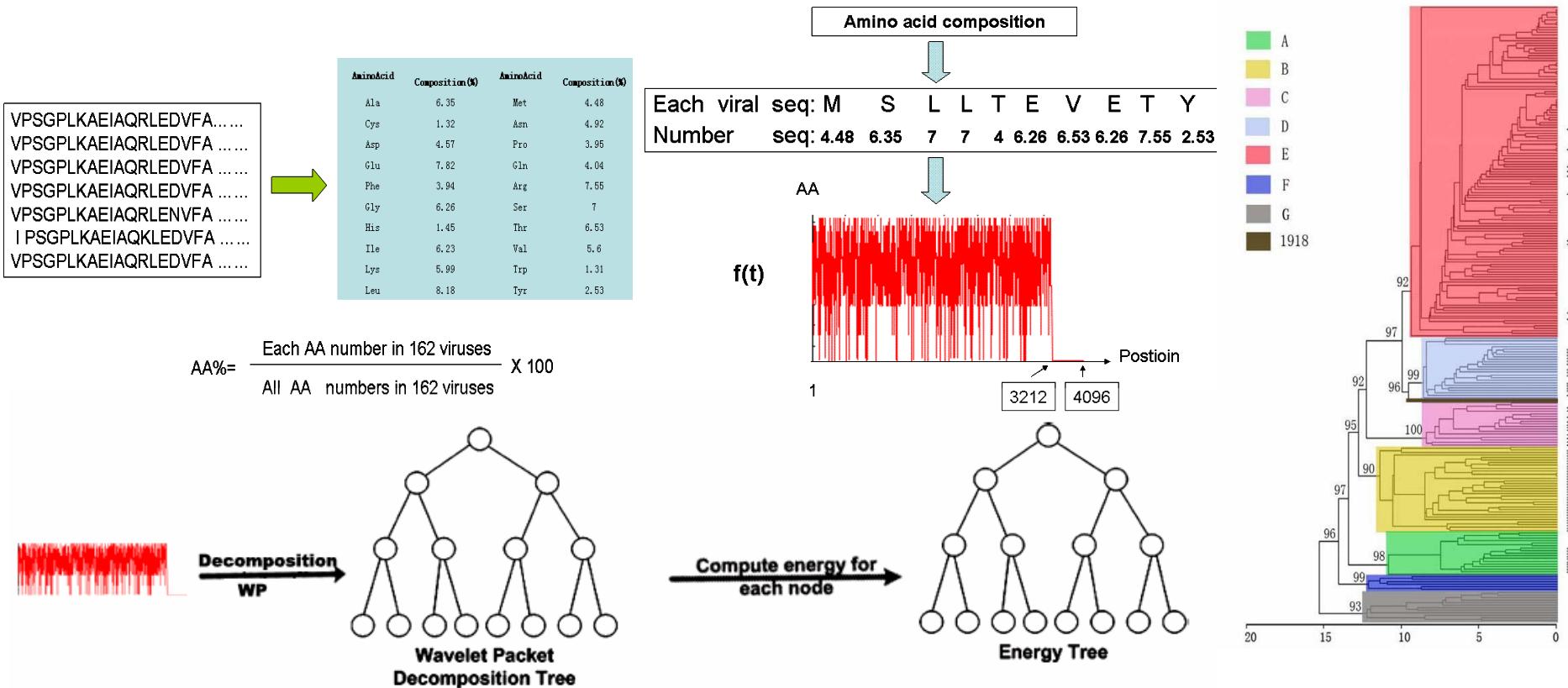
<b>Feature</b>	<b>Sensor/Model</b>	<b>Units</b>	<b>Resolution</b>
Temperature	MODIS (MOD11A2)	°C	1km
NDVI	MODIS (MOD13A2)		1km
Elevation	SRTM-DEM	meter	0.09km
Distance to Water	MODIS (MOD12Q1)	meter	1km
Distance to Meadow	MODIS (MOD12Q1)	meter	1km
Distance to Forest	MODIS (MOD12Q1)	meter	1km
Distance to Wetland	MODIS (MOD12Q1)	meter	1km

prediction result of bar-headed goose (date from 2008-08-01 to 2008-10-01) by considering temperature, NDVI, altitude, distance to water, distance to forest, distance to grassland, distance to ice and snow, distance to wetland, and distance to city and farm as input layers



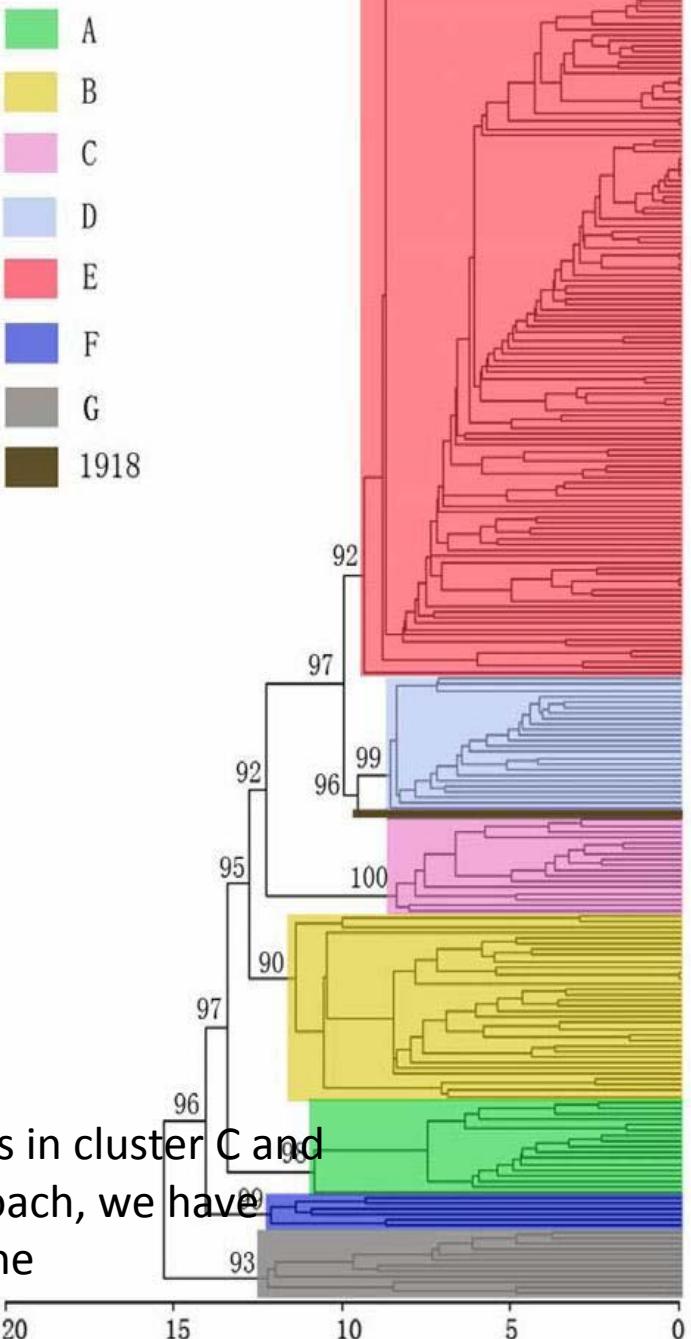
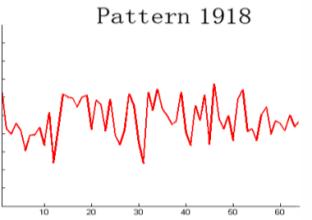
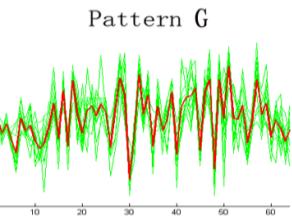
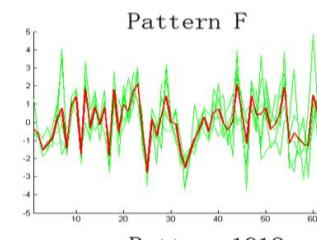
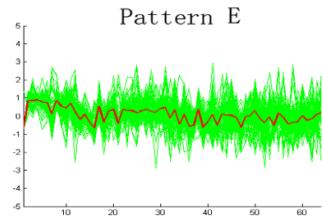
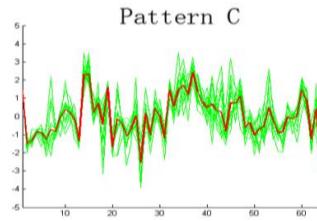
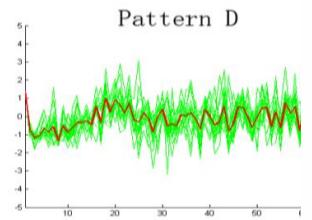
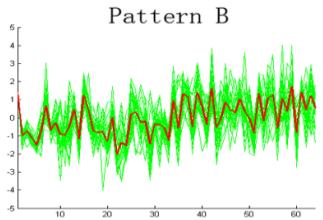
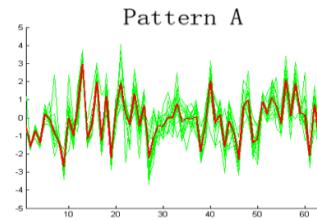
The warmer colors show regions with better predicted conditions.

# cross-host transmission ability of avian influenza virus



One paper describe this part of work was accepted by the 7th international conference on e-Science: Application of Data Mining in Research of Avian Influenza Virus Cross-Species Infection

# molecular patterns



Experiments results suggested that avian influenza A viruses in cluster C and E could directly infect human beings. Based on above approach, we have constructed a web-based early warning system to predict the transmissibility of avian influenza A virus to humans.

# Outline

- Introduction
- Database system and consideration
- Data analysis and visualization
- Applications
- Future work

# Future Work

- Data collection - more data accumulated
- Data integration for multi-sources data from different domain
- Data application
  - Application of biodiversity and spatial-temporal distribution
  - Application for the evolution of ecological structure and function
- Web-based interactive analysis platform

# Thanks!

Questions and Comments