Identification of space-time clusters and hotspots in communicable diseases surveillance in Northern Vietnam

Phnom Penh, December 5th 2022

Presented by: Nguyen Hai Tuan, MD., MPH., PhD.

Department of Communicable Disease Control

National Institute of Hygiene and Epidemiology (NIHE)

Presentation outline

- ❑ Purposes
- ❑ Methods
- ❑ Implementing

Purposes

The need to improve Communicable diseases Surveillance system

- To systematize the monitoring of clusters (occurrence of disease/outbreak signals) based on morbidity, mortality, coordinates, known targeted population
- Support for early detection of clusters of cases to trigger responses through Emergency Management Steps including confirmation, investigation, and activating responses
- Identify signal of clinical burden of disease in population

Why spatio-temporal scan statistics?

- In epidemiology spatio-temporal scan statistics is used to detect spatial or space-time disease clusters, and to determine if they are statistically significant
- Applied for similar problems in other fields such as archeology, criminology, demography, ecology, geography or zoology

Why spatio-temporal scan statistics?

- Support the description of the outbreak situation
- Contribute to Investigation Decision

Methods

References of spatio-temporal scan statistics

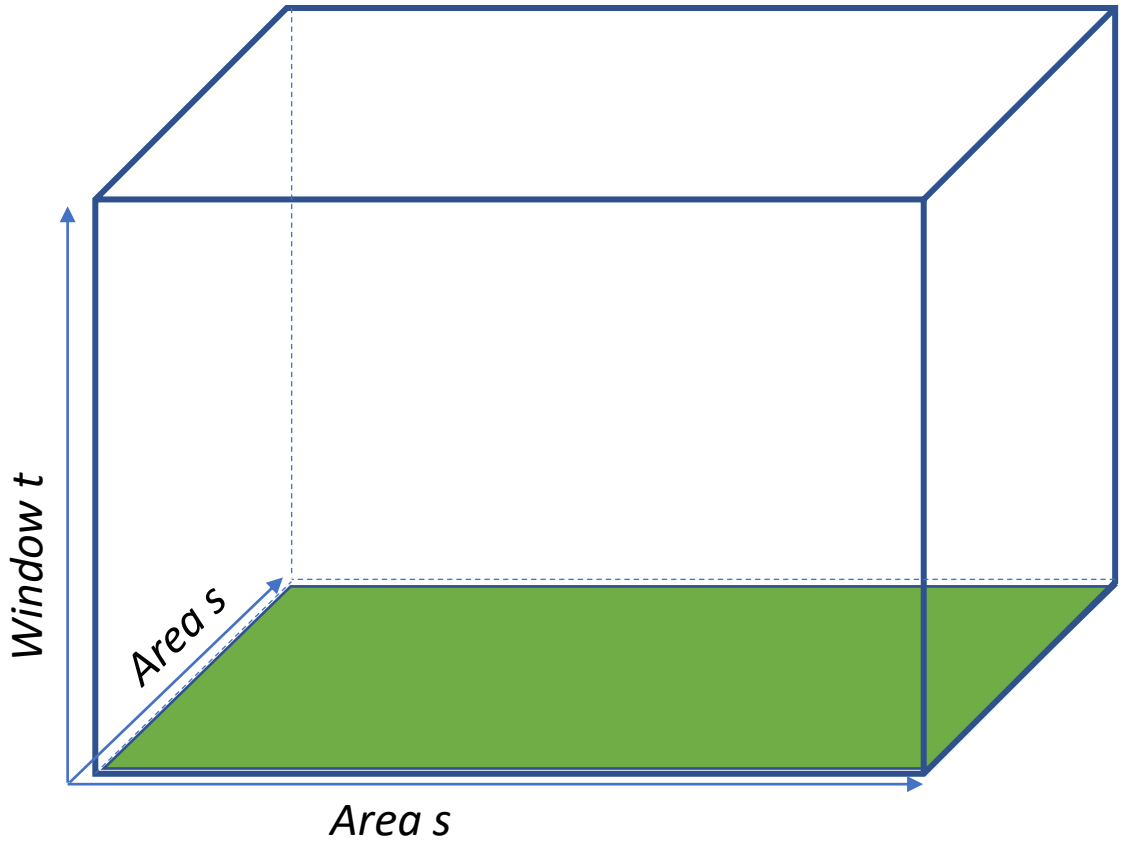
- Martin Kulldorf: www.satscan.org
- Scan Statistics: Theory and Applications, Séminaire de Probabilité et Statistique, Laboratoire de Mathématiques Paul Painlevé, 5 March, 2014, Lille
- Inkyung Jung 2019: The spatial scan statistic is defined as the maximum of likelihood ratio test statistics over a collection of scanning windows
- Wikle, C. K., Zammit-Mangion, A., and Cressie, N. (2019) Spatio-Temporal Statistics with R

Input data

- Morbidity/mortality: according to surveillance definition
- Coordinates:
  - Coordinates of administrative regions (commune level) of current patients' living address
  - Averaged coordinates (calculated by software)
  - Small surveillance area is preferred (cases in the same surveillance area have same coordinates)
  - Decimal coordinates (Lat/Long), geodetic EPSG 4326 - WGS 84 (European Petroleum Survey Group – World Geodetic System)
- Time:
  - Onset date/detection date of cases (day, week, month, year)
- Population size (averaged population or real-time population)

Two-dimensional Scan Statistics

- Windows statistics:
  + Temporal windows
  + Spatial windows



Setting up scanning parameters

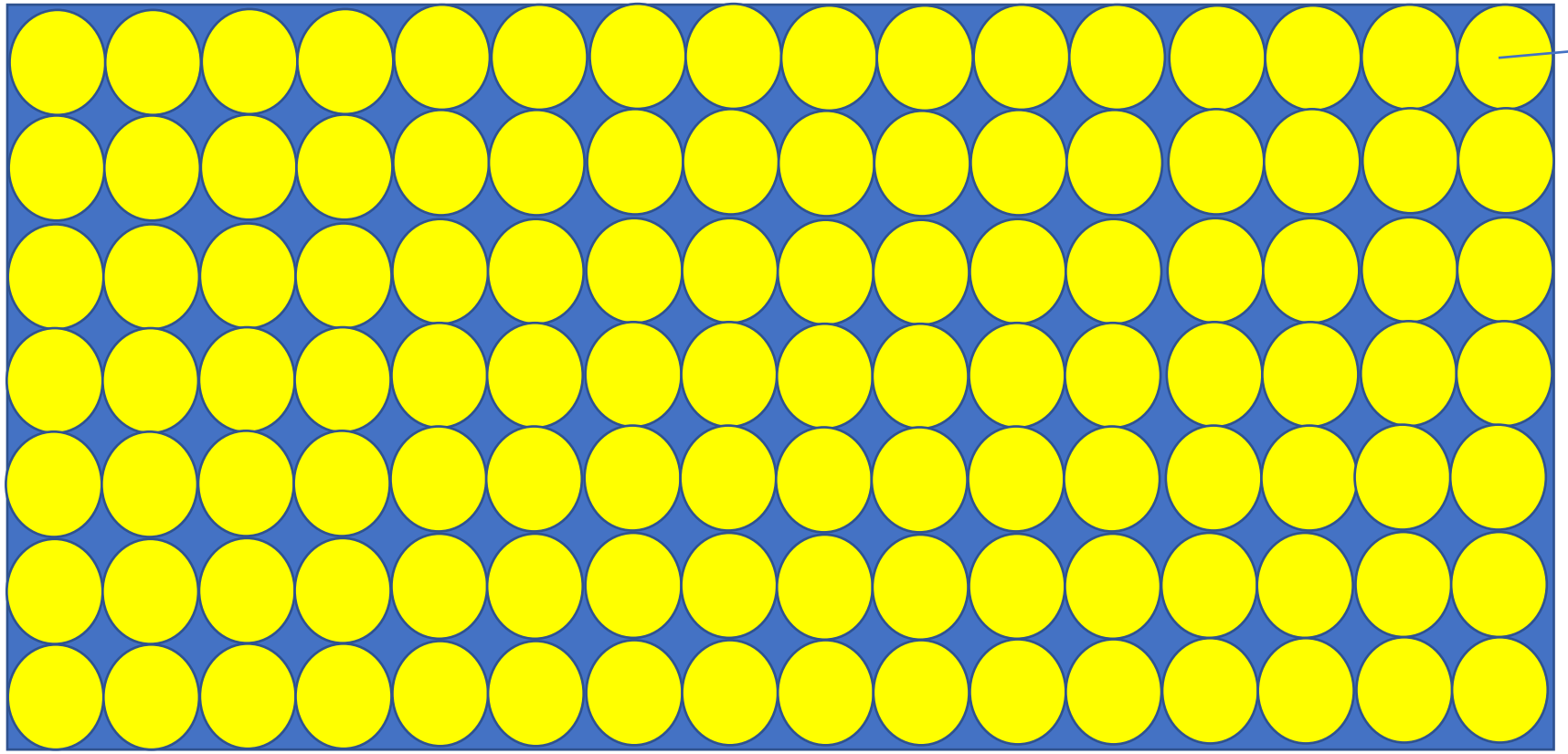
- Cluster threshold
  - More or equal to 2 cases for epidemic diseases
  - Endemic diseases:
    - Average (baseline) + > 2SD: data collected from respective surveillance region (e.g.: district, province, country), counting period of 3-5 years excluding outbreak period
- Time aggregation
  - Time aggregation : usually 7 days, 14 days, 21 days
  - Prospective analysis: averaged time is reporting period (7 days for weekly count)
- Spatial window
  - Percentage of population at risk ≤ 50% (default value 50)
  - Maximum radius (km): meaningful distance for surveillance, investigation (block, hamlet, village) several to dozen of km (20 km).

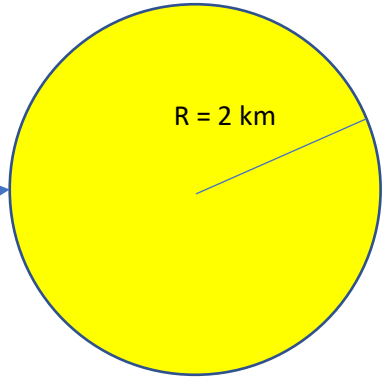
Determine scanning parameters

- Temporal window
  - Retrospective scan: maximal temporal window ≥ Time aggregation
  - Prospective scan: maximal temporal window = reporting period (7 days for weekly count); if choose 14 days, the Time aggregation (mean, SD) should be 14 days.
- Monte Carlo Replication
  - Select value 0, 9, 999, 9999, or any number ending with 999
  - Convenient p-value %, ‰, ...

Simple snapshot of purely spatial dimension

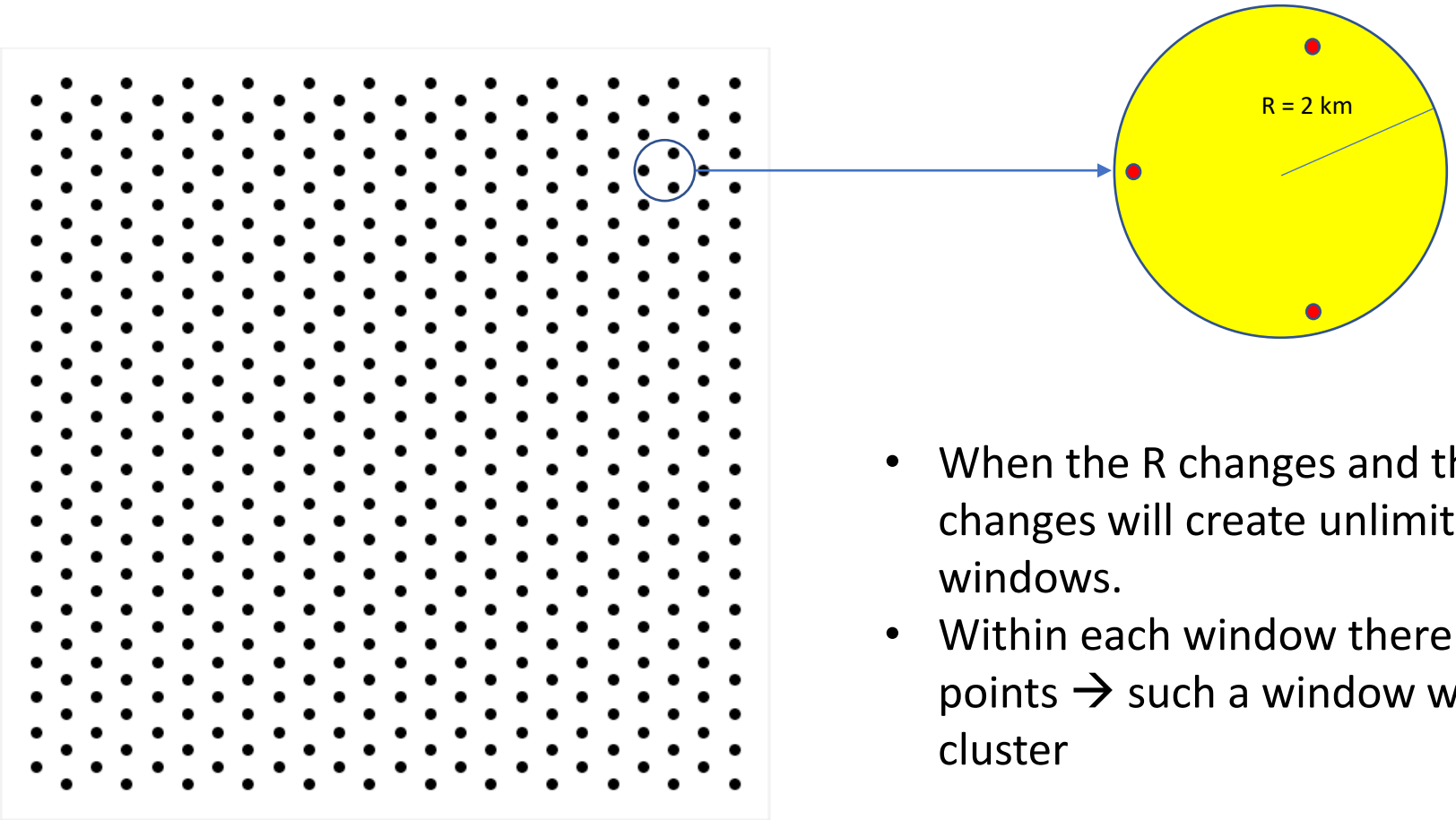
Surface S is divided into m surfaces of equal size.





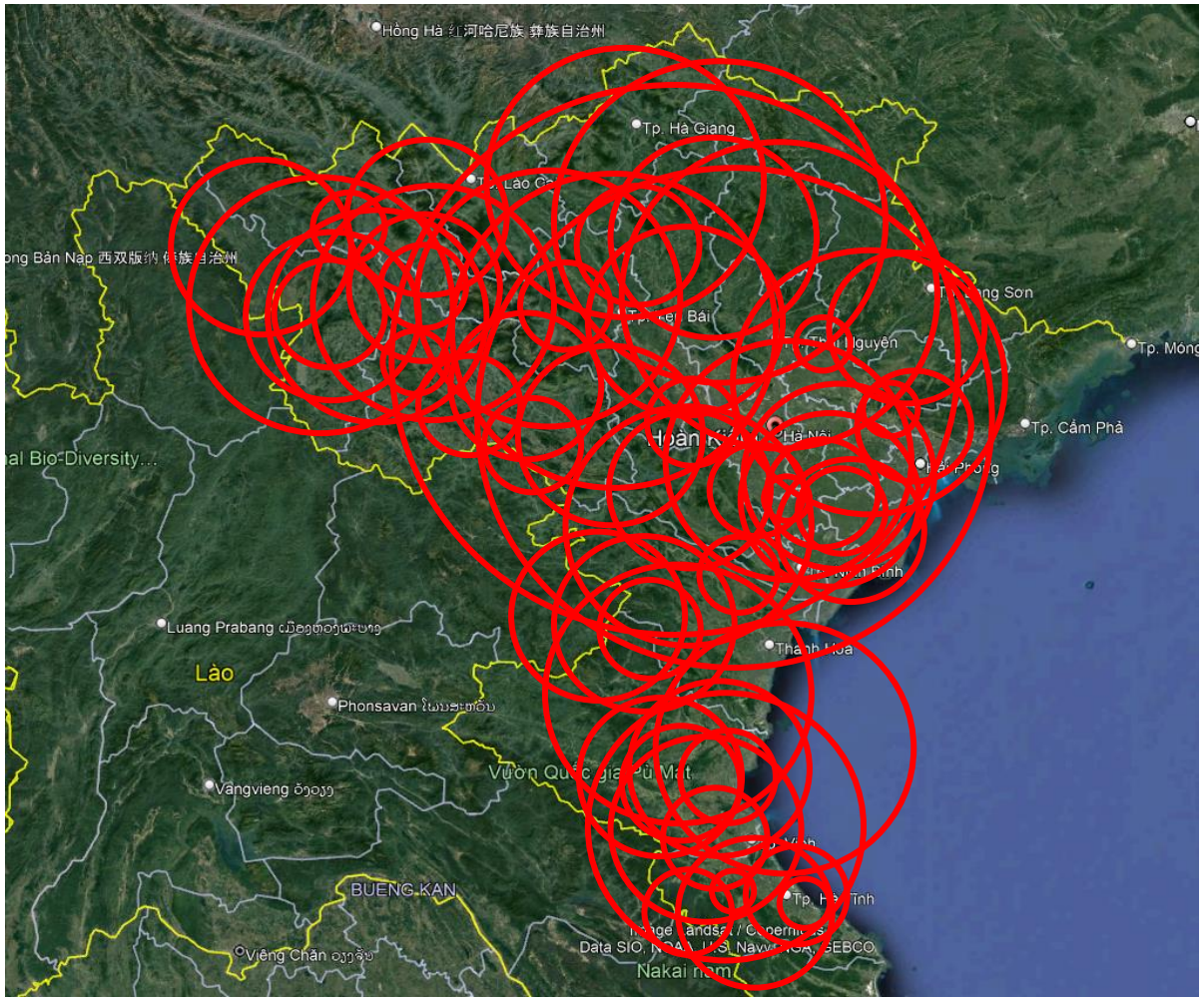
- E.g. with R = 2 km, the S is divided into 112 circles of 2 km
- The m changes according to R varying 0 to designed size.

How potential purely spatial clusters are identified?



- When the R changes and the center of circle changes will create unlimited number of windows.
- Within each window there are neighboring points → such a window will be a potential cluster

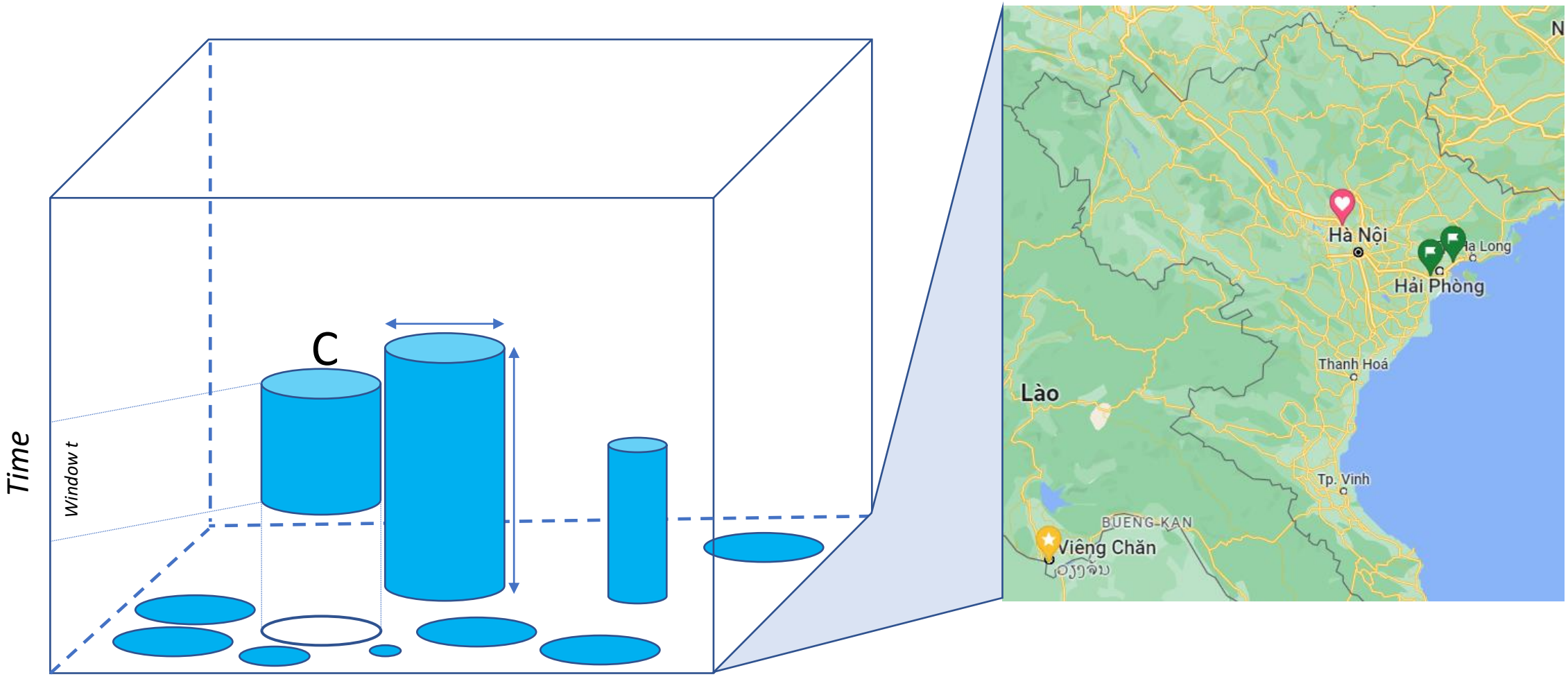
Potential purely spatial clusters



- Example: a small sample of windows in purely spatial scan statistics.
- At each location, the R of the circle varies from 0 to a designed size (km).

Potential spatio-temporal clusters

Bottom surface is surveillance area

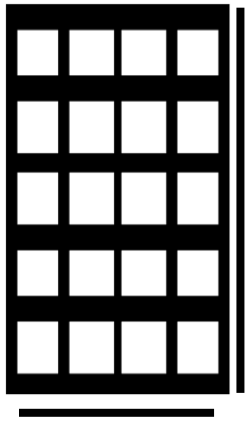


- Number and size of cylinder vary according to spatial window (t) and temporal window (R)
- Cases observed within a cylinder is potential cases of a cluster

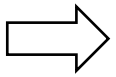
Monte Carlo simulation

- Number of replication : 0, 9, 999, 9999, …

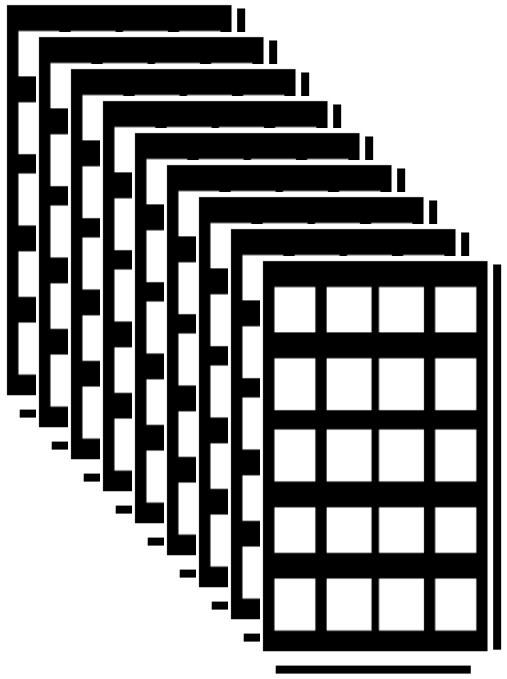
Real dataset



9

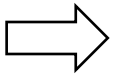


9 simulated datasets



or

999



//////////// …. ////////////

Computing and comparing the maximum likelihoods

The assumption that the case is equally distributed in time and in space:

+ Sliding windows with varying size from 0 to a designed sized.
+ Accumulated number of cases inside and outside the windows.
+ Compute and compare likelihoods. The H0 inside = outside, H1 inside != outside.
+ Monte Carlo replication (Monte Carlo Rank Test for likelihoods).

Steps of space-time analysis for surveillance

- Step 1. Preparing data table
- Step 2. Setting up input data for software/tools
- Step 3. Setting up statistical methods, parameters, output display format
- Step 4. Run program and verify the output
- Step 5. Presentation, visualization and interpretation of the results

Implementing

Implementing activities

- Training provincial CDC staffs on application of space-time analysis in communicable diseases surveillance
  - Provide participants the key concepts and knowledge to conduct scan statistic applied in routine communicable diseases surveillance and responses
  - Hand-on practices to identify significant clusters and communicable diseases outbreaks in community.
- Integrate approach for communicable disease surveillance in 4 selected provinces in the Northern Vietnam

Implementing activities

- Organize and support data collation at provincial CDCs
  - + Number of cases/Number of death
  - + Date of onset/date of event
  - + Coordinates at commune/ward level
  - + Population size of commune/ward level from District Statistic Office
  - + Determinant factors

Implementing activities

- Public health surveillance
  - Highly pathogenic avian influenza (HPAI) (A/H5N1, A/H5N6, A/H7N9, …)
  - Dengue Haemorrhagic Fever
  - Hand-Foot-Mouth Disease
  - Coronavirus disease 2019 (COVID-19)
  - Seasonal Influenza Diseases (A, B)
  - Measles, rabies
- One year pilot
- Implementing provinces: 4 provinces

Implementing activities

- Preparing input data from multiple data sources:
  - + Event-based surveillance
  - + Indicator based surveillance (Hospital based surveillance, CHS surveillance, International Quarantine Network: HPAI, Dengue, Measle, Rabies, …)
  - + Sentinel surveillance (COVID-19/ILI, HFM, Plague Surveillance and Control)
  - + Laboratory surveillance
- Assist weekly analysis to identify potential clusters
  - + Prospective analysis to identify active clusters
  - + Report significant clusters/event
- Assist CDC staff for cluster investigation and outbreak confirmation
  - + Call to request more information (what, when, where, how)
  - + Conduct confirmation process (SOPs for Public Health Event/outbreak confirmation)

Implementing activities

- Reporting clusters to stakeholders
  - + Reporting to NIHE
  - + Reporting to Provincial Department of Health
  - + Reporting to General Department of Preventive Medicine
  - + Reporting to Regional Animal Health Office
- Activating RRT for responses
  - + Laboratory investigation: collection of sample and laboratory diagnosis
  - + Activate response functions and enhance monitoring & reporting