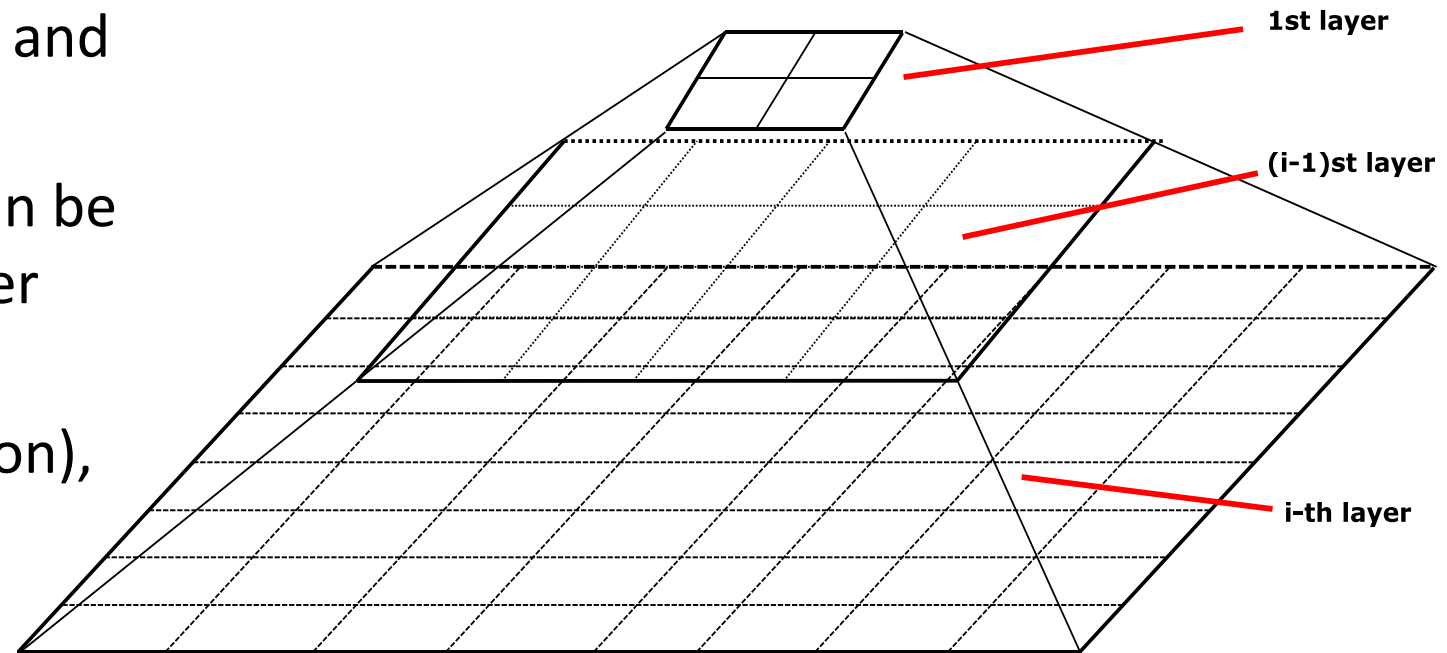




# **STING: A Statistical Information Grid Approach**

# STING: A Statistical Information Grid Approach

- ❑ STING (Statistical Information Grid) (Wang, Yang and Muntz, VLDB'97)
- ❑ The spatial area is divided into rectangular cells at different levels of resolution, and these cells form a tree structure
- ❑ A cell at a high level contains a number of smaller cells of the next lower level
- ❑ Statistical information of each cell is calculated and stored beforehand and is used to answer queries
- ❑ Parameters of higher level cells can be easily calculated from that of lower level cell, including
  - ❑ *count, mean, s*(standard deviation), *min, max*
  - ❑ type of distribution—*normal, uniform, etc.*



# Query Processing in STING and Its Analysis

---

- ❑ To process a region query
  - ❑ Start at the root and proceed to the next lower level, using the STING index
  - ❑ Calculate the likelihood that a cell is relevant to the query at some confidence level using the statistical information of the cell
  - ❑ Only children of likely relevant cells are recursively explored
  - ❑ Repeat this process until the bottom layer is reached
- ❑ Advantages
  - ❑ Query-independent, easy to parallelize, incremental update
  - ❑ Efficiency: Complexity is  $O(K)$ 
    - ❑  $K$ : # of grid cells at the lowest level, and  $K \ll N$  (i.e., # of data points)
- ❑ Disadvantages
  - ❑ Its probabilistic nature may imply a loss of accuracy in query processing