

多空间分辨率的时变气候集合模拟数据可视化 (Visualization of Time-Varying Weather Ensembles Across Multiple Resolutions)

作者: Fan Hong 日期: 2016年9月23日

在气象相关的集合模拟中，量化模型中的不确定性是领域科学家们十分关心的问题，其对在现实世界中做出决策有十分重要的意义。现如今，随着计算能力的飞速提高，已有的模型模拟已经能够产生出时变的多分辨率的集合模拟数据集。因此而产生了领域科学家十分关心的两个问题：一是输入参数的敏感度信息，二是模型在不同分辨率下的精确程度。根据输入参数的敏感度信息，可以将更多的计算资源倾斜到那些对输出有更高影响上的输入变量上。同时，了解模型在不同分辨率下的精确程度，也能在节省存储和计算代价的前提下，得到足够优秀的模型输出结果。本文就针对WRF (Weather Research and Forecasting) 模型，根据其在不同参数不同分辨率下的输出，并在真实数据的辅助下，研究参数敏感度和空间分辨率精确度的关系。

本文所采用的数据是WRF模型模拟数据。输入参数有5个，各个参数的物理意义和取值范围见表1。输出变量有1个，为降雨量。模拟模拟的时间长度为30天，每天对应一个时间步。共有150个集合模拟成员。同时，本文所用数据特殊的地方在于有3个不同的空间分辨率下的模拟结果，即12千米、25千米和50千米，并有真实数据用以确定模拟结果的精确程度。

Parameter	Description	Range	Default
Pe	Coefficient related to entrainment mass flux rate	[-1,1]	0
Pd	Coefficient related to downdraft mass flux rate	[-1,1]	0
Pt	Maximum turbulent kinetic energy	[3,12]	5
Ph	Starting height of downdraft above updraft source layer	[50,350]	150
Pc	Average consumption time of convective available potential energy	[900,2700]	2700

表1：WRF模型所考虑的输入参数的物理意义及其取值范围

作者针对所提出的这两个问题，提出了两部分较为独立的分析方法，并将其综合放入一个分析系统中。图1是所提出分析方法的工作流程图。

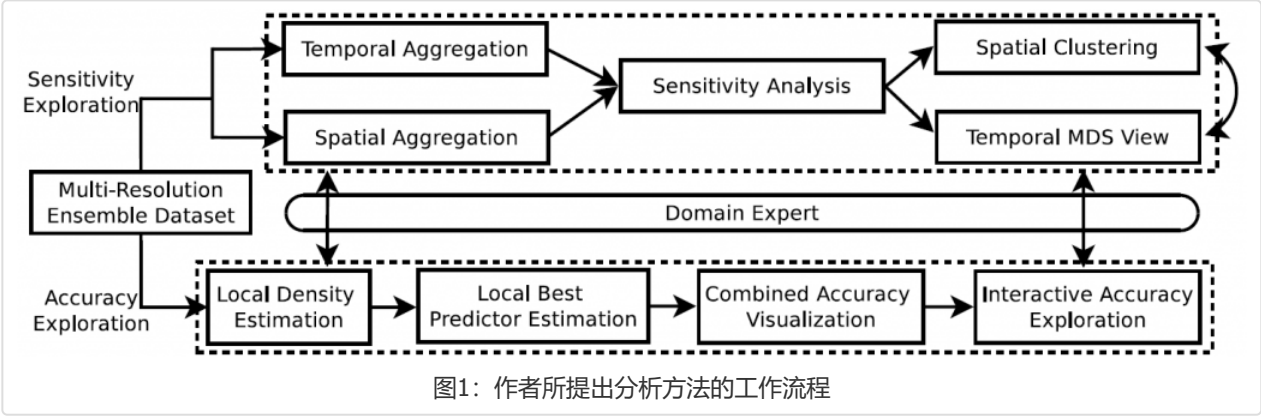


图1：作者所提出分析方法的工作流程

我们先来看输入参数敏感度分析部分。

作者首先定义如何计算一个参数对于一个模型的敏感度，这已经有比较成熟的方法。对于一个有n个输入参数和1个输出变量的模型，当输入参数取各种不同值的组合时，输出变量也会随着变化，此时输出变量的值会形成一个分布。当考虑将某一个输入参数固定为某个特定值，而其他参数仍然可以取多个值时，输出变量的值的分布会产生变化。此时，把两个分布的差值就称作这个输入参数取当前值是的“偏移” (shift)。而考虑这个参数的所有可能取值，其偏移的期望就是这个偏移的敏感度。作者之后的分析就是基于这个定义。

然后作者进行基于空间的分析。作者首先对一个空间分辨率一个集合成员下的一个空间位置在时间上的取值做一个加权平均。然后可以计算一个空间分辨率下，各个空间位置的各个参数的敏感度，它们实际上可以看作是一个5维的向量（因为有5个输入参数）。接着作者基于这个敏感度向量和位置向量，对各个空间分辨率下的所有空间位置一起做聚类分析。这里，在敏感度向量和位置向量之间，使用者可以调节两者的权重。图2展示了聚类结果，其中有4个聚类。我们可以很容易地看到12千米和25千米分辨率的敏感度表现与50千米的有很大的不同。

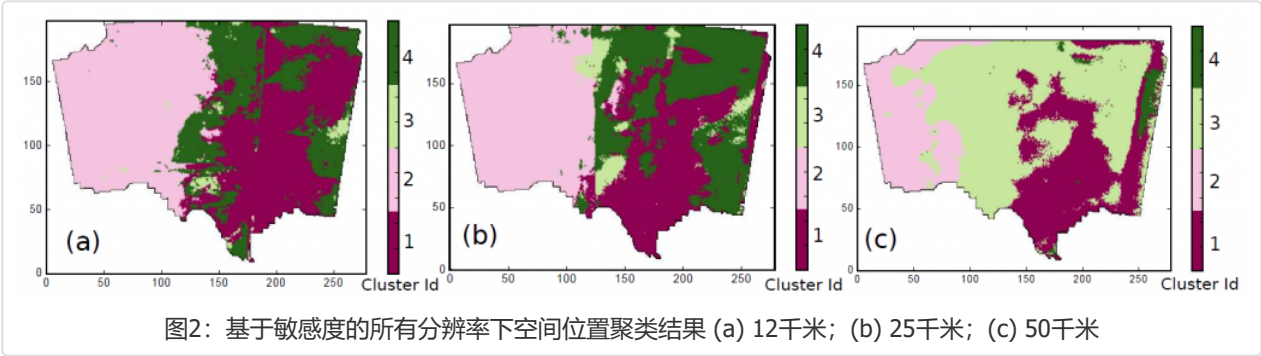
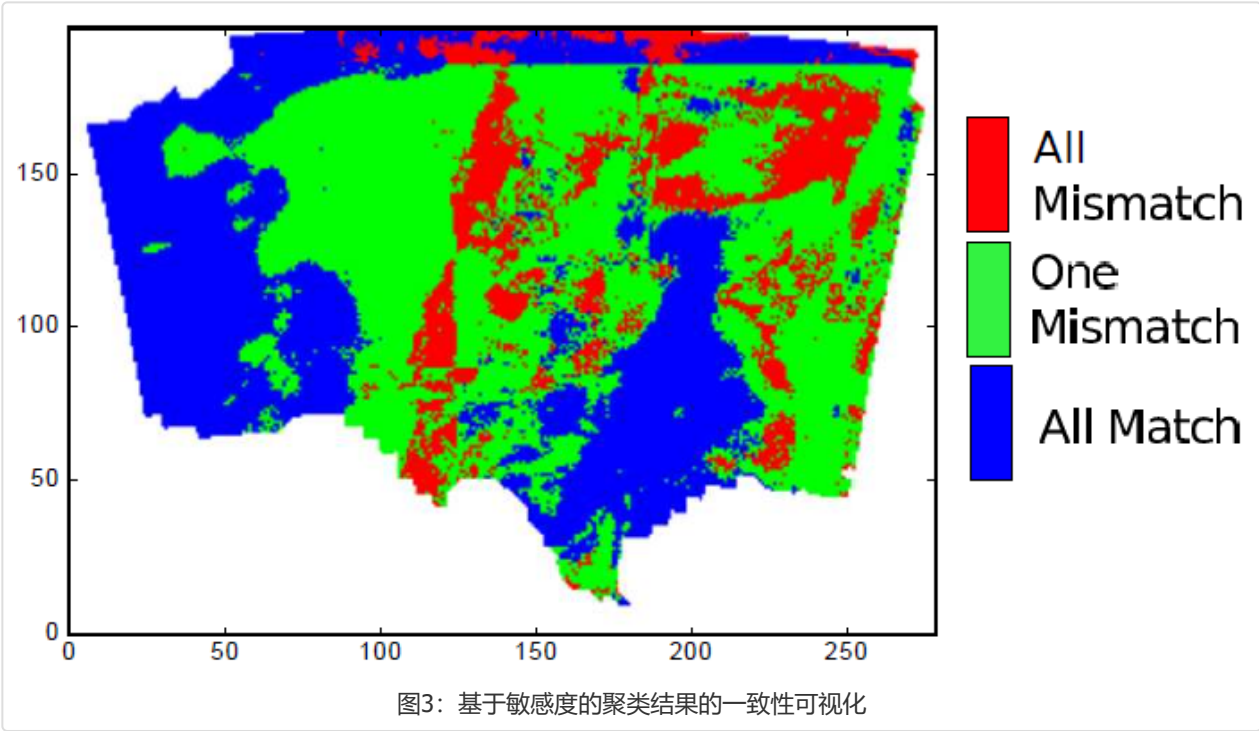


图2：基于敏感度的所有分辨率下空间位置聚类结果 (a) 12千米; (b) 25千米; (c) 50千米

接着，作者又用颜色来编码三中空间分辨率下聚类结果的一致性，如图3所示。如果某位置在三种分辨率下，各参数的敏感度表现较为一致（即聚类结果一致），则用蓝色表示；如果三者各不相同，则用红色表示；否则用绿色表示。通过与实际降雨量的比较发现，在降雨量高或者极低的区域，三种分辨率下的参数敏感度表现一致，即图中的蓝色区域。而在其他区域就不那么一致了。



然后作者进行基于时间的分析。与前面基于空间的分析类似，作者首先对一个空间分辨率一个集合成员下的一个时间步的所有空间位置的取值做一个加权平均。需要注意的是，为了提高之后分析结果在时间上的连续性，作者还采用了滑动窗口技术，即在相邻的时间步上做加权平均。之后，计算一个空间分辨率下，各个时间步的参数敏感度向量。接着，作者利用这个向量对各个空间分辨率的所有时间步一起做了一个多维尺度投影，结果如图4所示。在这个投影图中，红、绿、蓝色边所连接的节点分别表示三种分辨率对应的时间步下的敏感度向量。使用者选择某个输入参数之后，节点的大小用来表示这个参数的敏感度的值，而节点颜色表示其值在5个变量中的排名。从图中可以看到，50千米分辨率的敏感度与另外两种分辨率距离较远，差别较大；12千米和25千米分辨率的敏感度更加接近。此外，作者也提供了一个折线图，来展示各个输入参数敏感度随时间的变化情况。

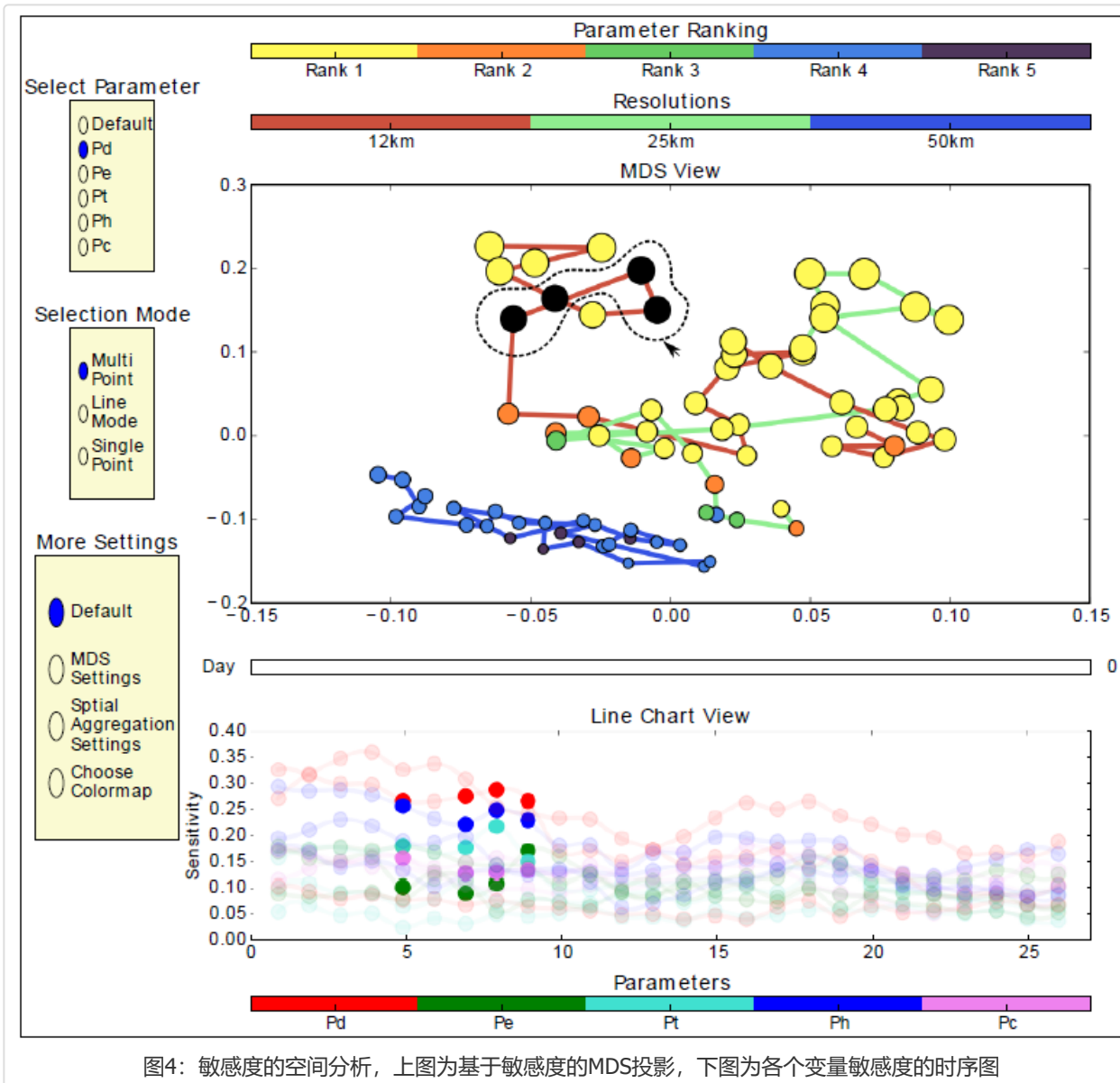


图4：敏感度的空间分析，上图为基于敏感度的MDS投影，下图为各个变量敏感度的时序图

我们再来看不同分辨率下模型精确度分析的部分。

同样地，作者首先定义如何量化这个精确度，作者采用了贝叶斯模型。考虑一个时空位置，在每种分辨率下，集合模拟已经产生了一系列预测值，并且有观测值，现在需要计算哪种分辨率最有可能。可以将其形式化为图5的公式（其中 $D=d_i$ 表示取不同的分辨率， $R=r$ 表示观测值为 r ），并对其按贝叶斯公式变换为两项的乘积。其中前一项表示各个分辨率的先验概率，可以设置让三种分辨率相等，也可以设置为与计算存储花费成反比。而后一项则表示，在有一系列预测值的情况下，观测值的概率。作者根据预测值用核密度估计（KDE）的方法求得分布，然后用观测值在这个分布中对应的密度值来表示概率。

$$\operatorname{argmax}_{i \in k} P(D = d_i | R = r)$$

$$P(D = d_i | R = r) \propto P(R = r | D = d_i) P(D = d_i)$$

图5：最优预测的定义，以及根据贝叶斯公式变换后的两项

这样，我们就可以计算在当前观测值下，各个空间分辨率的概率了。作者直接选取最大概率对应的空间分辨率来代表这个时空位置的最优预测。然后，作者对各个空间位置的最优预测分辨率作一个时间上的聚集，即统计某个空间位置在所有时间步上各个分辨率为最优预测的频率。然后对其进行可视化，如图6左上图所示。作者首先用红、绿、蓝三种颜色表示各个位置上最优预测最多的分辨率，然后用颜色的深浅来表现其作为最

优预测的频率。例如，图中中间偏右大部分为橙色，说明高分辨率（12千米）的预测更优。而对于大部分蓝色区域，低分辨率（12千米）的预测更优。并且，对于一些深蓝色区域，低分辨率的优势十分明显。这个发现之前没有被发现过，它也让领域科学家感到十分惊奇。

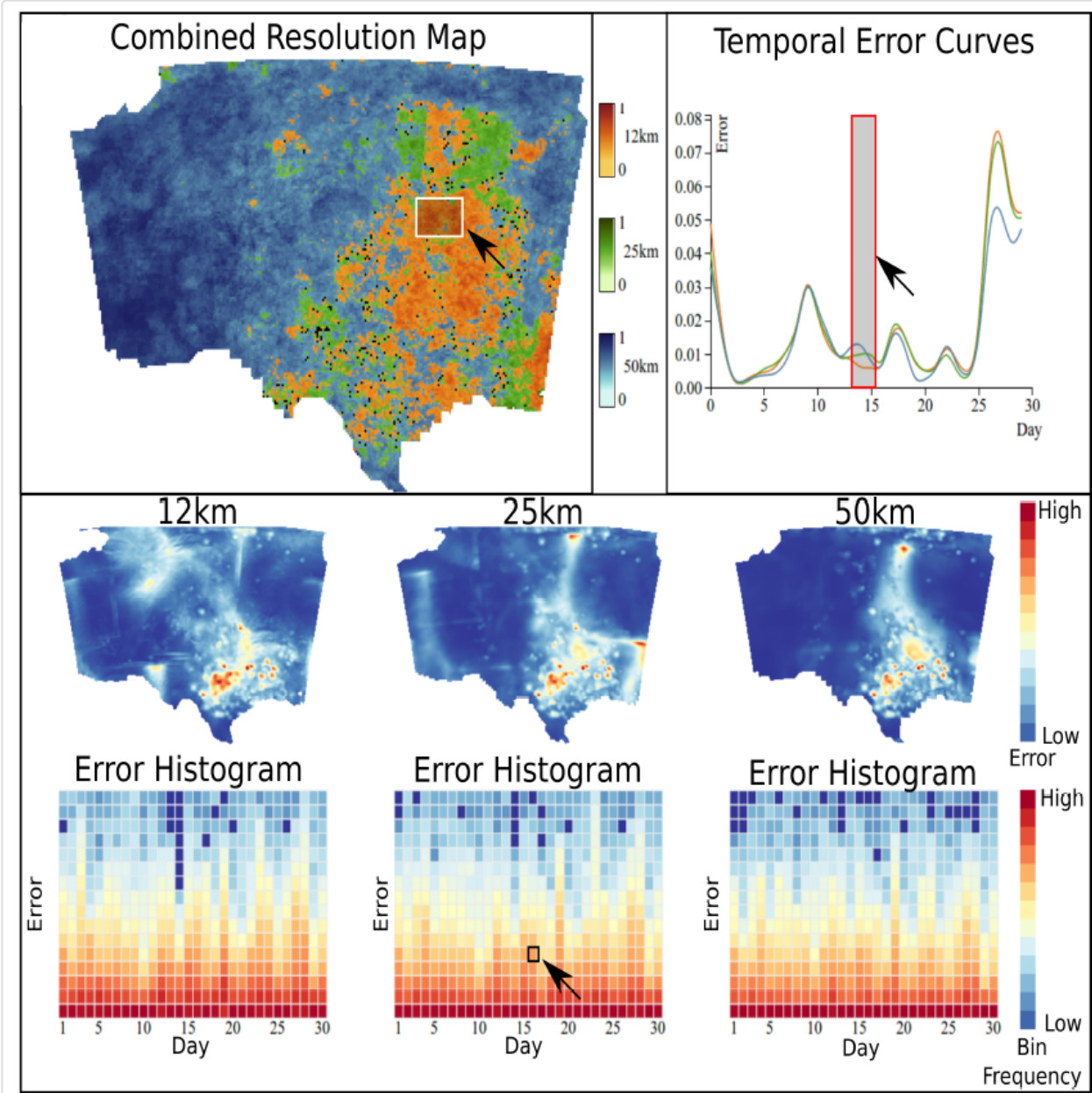


图6：各空间位置精确度的可视化结果

此外，对于每个时空位置，作者还计算其预测值与观测值的均方根误差，并分别做时间或空间的聚集，用时序图、空间视图和直方图来展示误差的分布，如图6右上图和下图所示。

最后，作者基于这两大分析模块构建了一个可视分析系统。领域专家在使用后，有比较好的反馈：这个系统使得他们能够在不同层次的数据聚集操作下分析敏感度信息，而之前只能进行时间或者空间的聚集操作；精确度分析所产生的可视化能够直观地描述各个分辨率的好坏。同时，在系统的使用过程中，也发现了许多有用的信息，甚至包括之前工具所不能发现的。

总起来看，本文提出了针对多分辨率时变集合模拟数据的两类分析方法：输入参数敏感度分析，和不同分辨率下的精确度分析。分析本身主要都是基于由数据所衍生出来的敏感度数据和精确度数据，方法上不外乎是

数据聚集、投影、聚类 and 颜色编码。但由于数据新颖，并且解决了对应的问题，得到了领域专家的认可，所以有了突出的贡献。

[1] Ayan Biswas, Guang Lin, Xiaotong Liu, Han-Wei Shen. Visualization of Time-Varying Weather Ensembles Across Multiple Resolutions. IEEE Transactions on Visualization and Computer Graphics (TVCG), 2016.

论文报告

不确定性, 时变数据可视化, 科学可视化, 集合模拟数据

← 从文章关键字透析VIS领域 (Visualization as Seen Through its Research Paper Keywords)

AxiSketcher: 基于用户绘制的非线性数轴映射 →

评论关闭。

RSS 订阅

功能

- 登录
- 文章RSS
- 评论RSS
- WordPress.org

链接

- 北京大学可视化与可视分析研究小组主页 – PKU Vis Home Page
- 北京大学可视化研究维基 – PKU Vis WIKI

分类目录

- 应用
- 新闻
- 未分类
- 活动
- 研究
- 论文报告

标签

ChinaVis graph interaction PacificVis
pacificvis2019 pvis2016 不确定性 主题模型 交互
交互设计 人机交互 会议 体可视化 体绘
制 信息可视化 动态图可视化
动态网络 北京大学 可视分析
可视化 可视化前沿讲习班 可视化应用
图可视化 城市数据 多层次 大规模数据可
视化 微博可视化 散点图 数据立方体 数据降维