

ENVIFUSION

环境地学数据融合平台线上直播课

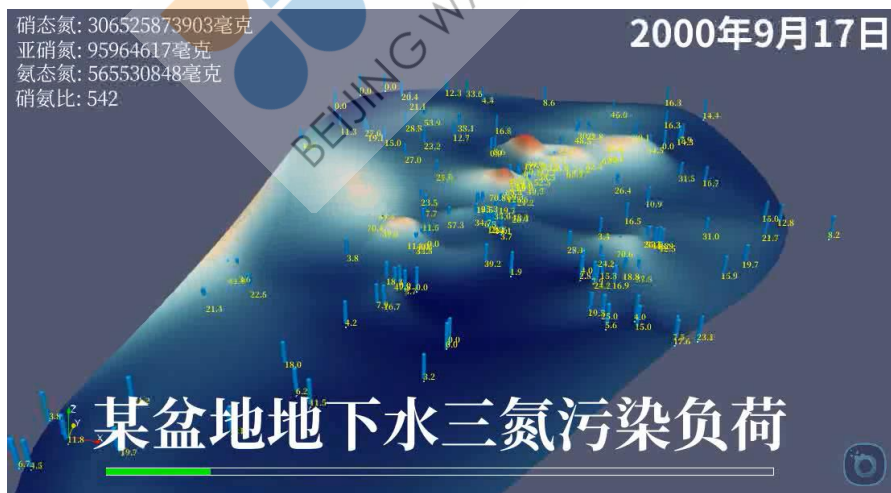
环境水文数据 缺失值补齐

2021年10月

空间缺失

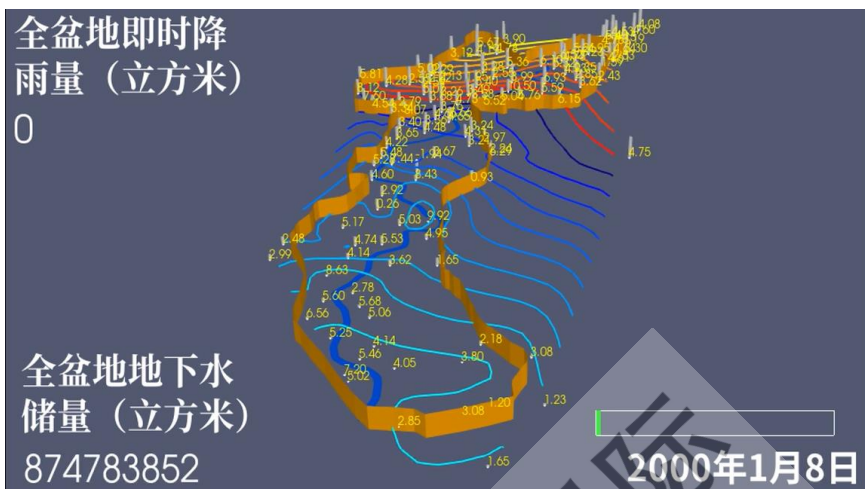
硝态氮: 306525873903毫克
亚硝氮: 95964617毫克
氮态氮: 565530848毫克
硝氮比: 542

2000年9月17日



- 某个关注区域位于监测网内，但未设置监测站点

站点缺失



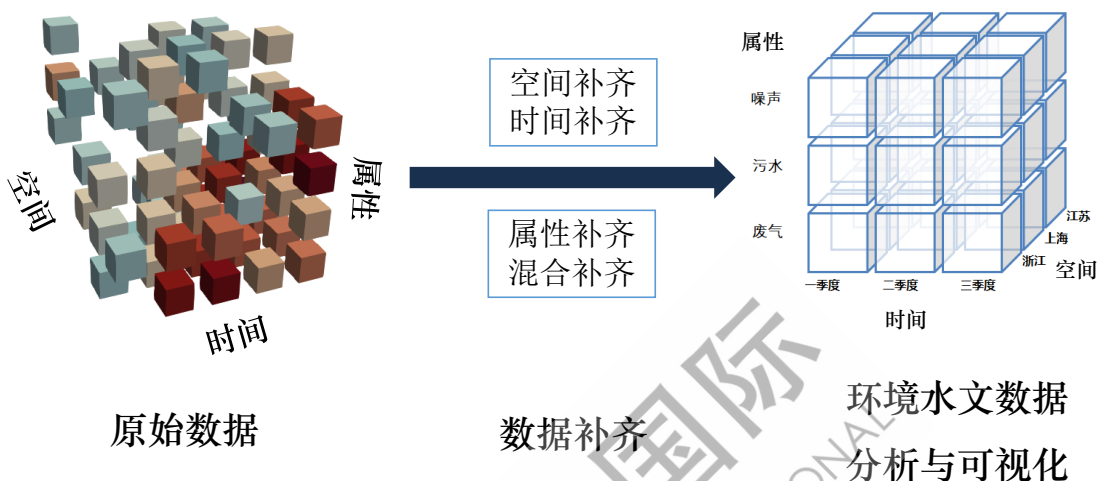
- 各类型监测站 (如水位和水文) 在空间位置往往不同。
- 需要将两种属性的站点进行统一分析。

时间缺失

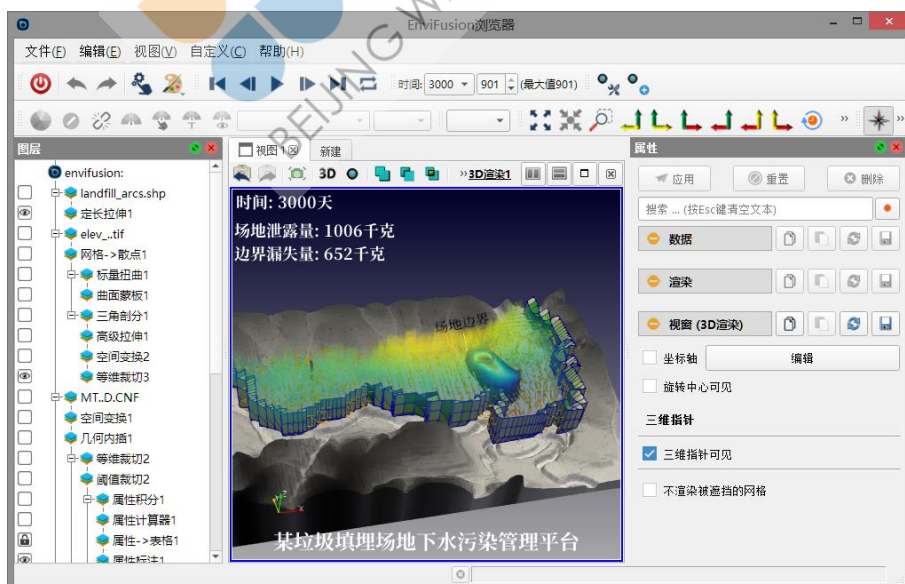


- 监测站点的采样频率常有不同
- 需要分析相同时间不同站点的监测结果。

环境水文数据分析



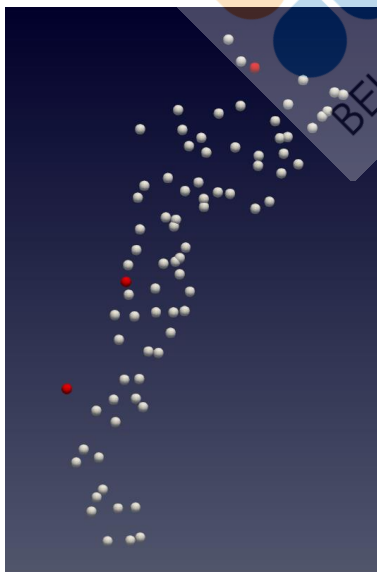
环境地学计算平台EnviFusion



数值模拟引擎

- 梯度场引擎：势场->流场->追踪相关工具；
- MODFLOW引擎：边界+参数->势场->流场；
- MT3DMS引擎：流场->浓度场；
- SEAWAT引擎：变密度流场->浓度场；
- SUB模块引擎：地面沉降模拟；
- TFPR引擎：势场变化->含水层参数；
- 地球化学引擎：高维度地球化学数据->降维分析；
- 文本分析引擎：热词提取与语义关联；
- **数据补齐引擎。**

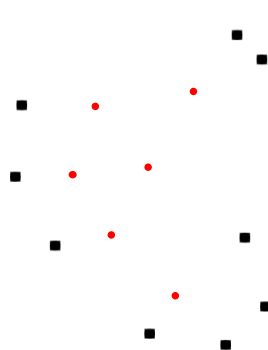
空间补齐-站点互采



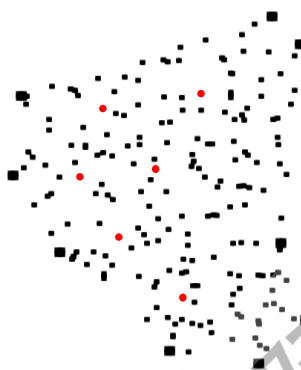
	ef散点ID	Day	WL	WellID	X	Y	ef散点		ef散点_矢量长度	
0	0	384	20.91	15	251950	4.057e+06	251950	4.057e+06	0	4.06482e+06
1	1	384	23.96	17	258059	4.05749e+06	258059	4.05749e+06	0	4.06569e+06
2	2	384	22.18	18	256743	4.05683e+06	256743	4.05683e+06	0	4.06495e+06
3	3	384	28.29	19	259384	4.06198e+06	259384	4.06198e+06	0	4.07025e+06
4	4	384	28.05	20	260743	4.06099e+06	260743	4.06099e+06	0	4.06936e+06
5	5	384	29.03	21	259791	4.06658e+06	259791	4.06658e+06	0	4.07487e+06
6	6	384	30.91	22	262970	4.06546e+06	262970	4.06546e+06	0	4.07396e+06
7	7	384	23.97	26	257044	4.06178e+06	257044	4.06178e+06	0	4.0699e+06
8	8	384	24.65	27	257000	4.06084e+06	257000	4.06084e+06	0	4.06897e+06

	ef散点ID	Day	Precip	WL	WellID	X	Y	ef散点		ef散点_矢量长度	
0	0	384	0.0155	20.91	15	251950	4.057e+06	251950	4.057e+06	0	4.06482e+06
1	1	384	0.0105	23.96	17	258059	4.05749e+06	258059	4.05749e+06	0	4.06569e+06
2	2	384	0.0105	22.18	18	256743	4.05683e+06	256743	4.05683e+06	0	4.06495e+06
3	3	384	0.0105	28.29	19	259384	4.06198e+06	259384	4.06198e+06	0	4.07025e+06
4	4	384	0.0105	28.05	20	260743	4.06099e+06	260743	4.06099e+06	0	4.06936e+06
5	5	384	0.0105	29.03	21	259791	4.06658e+06	259791	4.06658e+06	0	4.07487e+06
6	6	384	0.0105	30.91	22	262970	4.06546e+06	262970	4.06546e+06	0	4.07396e+06
7	7	384	0.0105	23.97	26	257044	4.06178e+06	257044	4.06178e+06	0	4.0699e+06
8	8	384	0.0105	24.65	27	257000	4.06084e+06	257000	4.06084e+06	0	4.06897e+06
9	9	384	0.0105	27.75	28	254866	4.06257e+06	254866	4.06257e+06	0	4.07056e+06

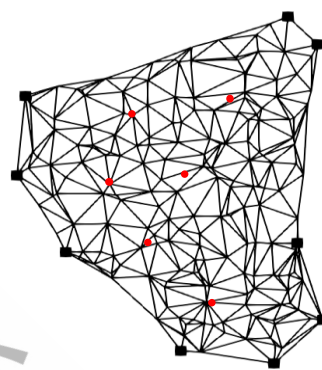
空间补齐-插值



建立凸多边形



随机网格顶点

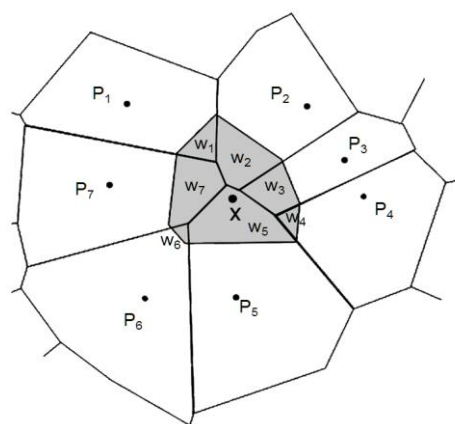


三角剖分

水头空间插值（自然邻点法）



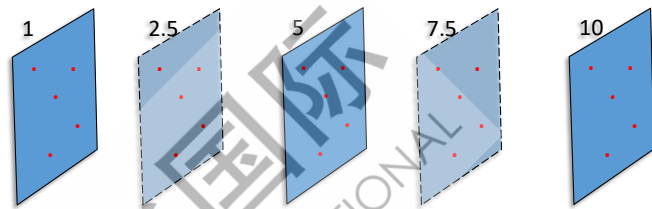
三角网格的对偶Voronoi多边形



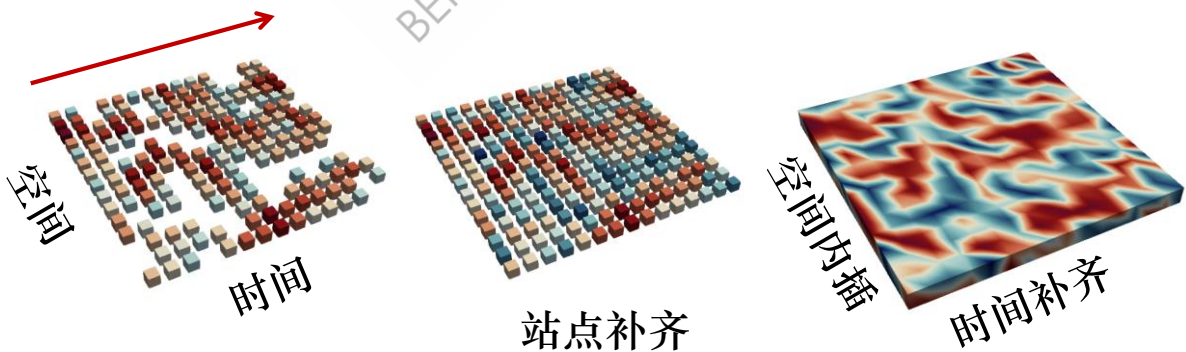
使用Voronoi多边形在X点处插值

时间补齐-插值

- 在离散时间步间生成虚拟时间步，使用上一步和下一步进行线性插值；
- 要求时间步之间站点数量一致；
- 分定序补齐和站点补齐两类；
- 定序补齐按顺序依次插值；
- 站点补齐针对数量一致但顺序混乱的情形，根据站点的名称对应进行补齐。



时间补齐-站点补齐

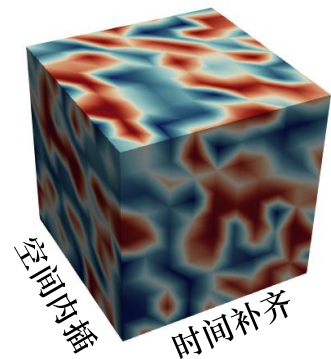
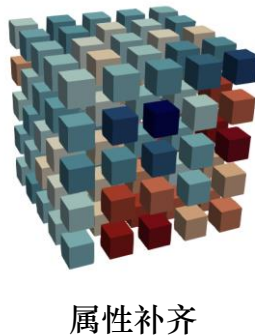
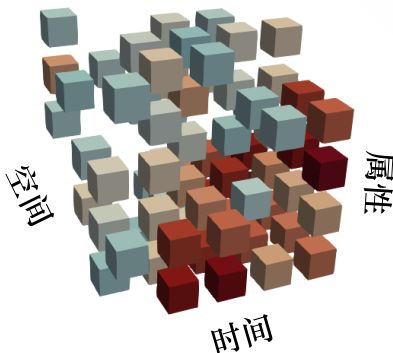


时间补齐-站点补齐

- 针对时间步之间站点数量不齐的情形；
- 遍历所有时间步，找出所有曾经存在过的站点；
- 为每一个时间步填满缺失站点；
- 对于某时间步中的缺失站点，提供两种补齐缺失值方法。
 - 线性补齐：向前和向后查找上一值和下一值，使用线性内插方法。
 - 空缺值补齐：对所有缺失值进行“-999”补充。

属性补齐

- 适用于高维关联属性数据体（如水质）

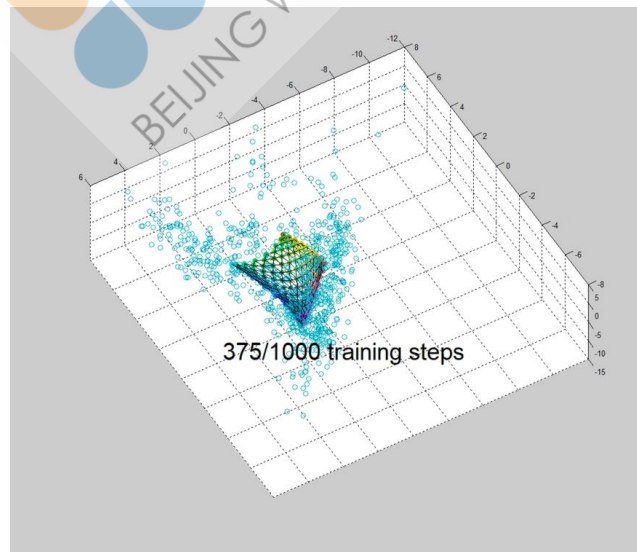


属性补齐

- 属性补齐与时空类补齐不同，需要使用相邻属性推断空缺属性。
- 例如在水化学中通过K、Ca、Na浓度推断 NO_3^- 浓度。
- 首先使用神经网络对全局数据进行学习。
- 再将含有空缺数据的向量映射到神经网络，由神经网络进行属性补齐。



神经网络训练过程



混合补齐

- 混合补齐是从数据体的两个或三个方向共同补齐，如**时间序列补齐**；
- 时间序列补齐是时间补齐和属性补齐的混合体；
- 使用属性的时间差分值进行神经网络训练，而非属性值本身；
- 补齐结果由上一时间步的**属性值**和训练之后的**差分值共同决定**。

