(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2024 年 2 月 29 日 (29.02.2024)

WIPO | PCT

(10) 国际公布号
WO 2024/040845 A1

(51) 国际专利分类号:
G06F 18/2337 (2023.01)

(21) 国际申请号: PCT/CN2022/144067

(22) 国际申请日: 2022 年 12 月 30 日 (30.12.2022)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202211007850.5 2022年8月22日 (22.08.2022) CN

(71) 申请人: 江 苏 大 学 **(JIANGSU UNIVERSITY)** [CN/CN]; 中 国 江 苏 省 镇 江 市 京 口 区 学 府 路 301号, Jiangsu 212013 (CN)。

(72) 发明人:石茂林**(SHI, Maolin)**; 中国江苏省镇江市京口区学府路301号, Jiangsu 212013 (CN)。

(74) 代理人: 南 京 智 造 力 知 识 产 权 代 理 有 限 公 司**(NANJING DOYES INTELLECTUAL PROPERTY CO., LTD.)**; 中国江苏省南京市江宁区芝兰路18号4幢1202室, Jiangsu 211100 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ,

(54) **Title:** COMBINE-HARVESTER BIG DATA RECONSTRUCTION METHOD, AND APPARATUS

(54) 发明名称: 一种联合收获机大数据重构方法与装置

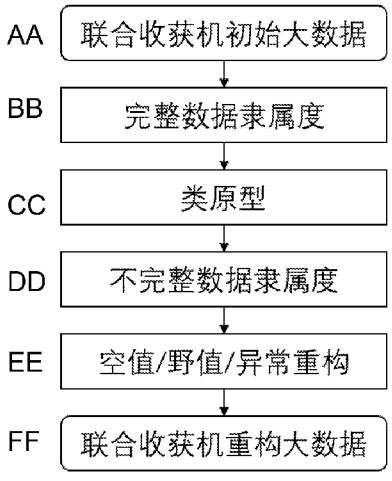


图 1

AA Combine-harvester initial big data
BB Complete data membership
CC Class prototype
DD Incomplete data membership
EE Null value/outlier/abnormal value reconstruction
FF Combine-harvester reconstructed big data

(57) **Abstract:** Provided in the present invention are a combine-harvester big data reconstruction method, and an apparatus. The method comprises: on the basis of a fuzzy clustering framework, describing a parameter association relationship of combine-harvester complete data by using a Kriging method, and inferring the real values of a null value, an outlier and an abnormal value according to an association relationship among data parameters, and information such as a space distribution density and a mean value of the data. In the apparatus, a data input module is used for acquiring field operation data of agricultural machinery; a data reconstruction module is used for receiving combine-harvester big data inputted by the data input module, and reconstructing the data according to a combine-harvester big data reconstruction method; a data output module outputs reconstructed data of the combine-harvester big data; and a data storage module is used for storing original data of the combine-harvester big data, the reconstructed data of the combine-harvester big data and a membership matrix of the combine-harvester big data. The precision of combine-harvester big data reconstruction is improved, a data support is provided for subsequent combine-harvester big data modeling and analysis tasks.

WO 2024/040845 A1

[见续页]

NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:
— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 本发明提供了一种联合收获机大数据重构方法与装置，该方法基于模糊聚类框架，利用克里金法刻画联合收获机完整数据的参数关联关系，通过数据参数之间的关联关系，以及数据空间分布密度、均值等信息共同推断空值、野值、异常值的真实数值。该装置，采用数据输入模块用于获取农机现场作业数据；数据重构模块用于接收数据输入模块输入的联合收获机大数据，并根据联合收获机大数据重构方法重构数据；数据输出模块输出联合收获机大数据的重构数据；数据保存模块用于保存联合收获机大数据的原始数据、重构数据、隶属度矩阵，提高了联合收获机大数据重构的精准度，为后续联合收获机大数据建模与分析任务提供数据支撑。

一种联合收获机大数据重构方法与装置

技术领域

本发明属于联合收获机、大数据、人工智能的交叉技术领域，特别涉及一种联合收获机大数据重构方法与装置。

背景技术

联合收获机是保障水稻、小麦、玉米等主粮及时收获的关键装备，其作业过程的性能表现对于我国粮食生产安全具有重要意义。联合收获机的作业性能主要由割台、输送槽、脱粒、清选等装置组成的工作部件所决定，不同装置之间的相互作用十分复杂，难以通过物理推导、数值仿真等手段准确刻画。在实际作业过程中，联合收获机的作业参数调整多依赖机手经验，难以准确匹配实时收获工况，作业性能时好时坏。

随着联合收获机工作部件运行监测的日益完善，来源于作业现场的实测数据不断累积。这些数据记录下了联合收获机工作部件内部的运行过程，也隐含了联合收获机工作部件与作物的相互作用机理。通过挖掘这些数据的内部信息，可以有效地指导联合收获机的运行作业，提升联合收获机的作业性能表现。

但是联合收获机在田间作业，部分传感器安装于脱粒、清选等装置内部，服役环境温度高、湿度大、粉尘多；田间地域广阔，难以提供稳定、良好的通讯条件。以上因素共同导致实测数据存在大量空值、野值与异常值，数据不完备。因此，就需要对这些空值、野值、异常值进行数据重构，以形成完备的数据集，为后续数据建模与分析任务提供数据支撑。

数据重构是大数据领域的热点话题之一，其是指根据数据内在特征，对空值、野值、异常值的真实数值进行估计，从而将内含空值、野值、异常值的不完备数据转变为完备数据的过程。

目前，对于联合收获机大数据的重构问题，多是从数据空间分布角度重构数据。例如，通过空值、野值、异常值对应参数的空间分布密度、均值等估计其真实数值，并没有考虑数据参数之间的关联关系，难以实现数据的精准重构。

发明内容

针对上述技术问题，本发明的一个方式的目的之一是提供一种联合收获机大数据重构方法，基于模糊聚类框架，利用克里金法刻画联合收获机完整数据的参数关联关系，通过数据参数之间的关联关系，以及数据空间分布密度、均值等信息共同推断空值、野值、异常值的

真实数值，提高联合收获机大数据重构的精准度，为后续联合收获机大数据建模与分析任务提供数据支撑。

本发明的一个方式的目的之一是提供一种实现联合收获机大数据重构方法的装置，采用数据输入模块用于获取农机现场作业数据；数据重构模块用于接收数据输入模块输入的联合收获机大数据，并根据联合收获机大数据重构方法重构数据；数据输出模块输出联合收获机大数据的重构数据；数据保存模块用于保存联合收获机大数据的原始数据、重构数据、隶属度矩阵，提高联合收获机大数据重构的精准度，为后续联合收获机大数据建模与分析任务提供完备的数据支撑。。

注意，这些目的的记载并不妨碍其他目的的存在。本发明的一个方式并不需要实现所有上述目的。可以从说明书、附图、权利要求书的记载中抽取上述目的以外的目的。

本发明是通过以下技术手段实现上述技术目的的。

实施例 1

一种联合收获机大数据重构方法，包括如下步骤：

步骤 S1、在模糊聚类框架下，利用克里金法刻画联合收获机不包括空值、野值或异常值的作业参数的完整数据的参数之间的关联关系，获取联合收获机完整数据的隶属度；

步骤 S2、基于步骤 S1 获得的联合收获机的完整数据及其隶属度计算类原型，继而获得包含空值、野值和或异常值的作业参数的不完整数据的隶属度；

步骤 S3、根据计算的到的类原型、不完整数据的正常数值、不完整数据的隶属度分别估计空值、野值和或异常值的真实数值，实现联合收获机大数据的重构。

上述方案中，所述步骤 S1 中，采用克里金法刻画数据参数之间的关联关系，通过如下公式计算完整数据的隶属度：

$$u_{k,i} = \left[\sum_{t=1}^{c}\left(\frac{\|y_k - \widehat{y_{k,i}}\|_2^2}{\|y_k - \widehat{y_{k,t}}\|_2^2}\right)\right]^{-1}$$

$u_{k,i}$是第k个完整数据对第i个子类的隶属度；

c是聚类的子类个数；

y_k是第k个完整数据的预测目标参数的数值；

$\widehat{y_{k,i}}$是第i个子类的克里金模型给出第k个完整数据的预测目标参数的预测值;

$\widehat{y_{k,t}}$是第t个子类的克里金模型给出第k个完整数据的预测目标参数的预测值。

进一步的，所述步骤 S1 中，聚类的子类个数c为 3，隶属度阈值θ，一般取为为 0.5。

上述方案中，所述步骤 S2 中，通过如下公式计算的类原型：

$$\boldsymbol{v}_i = \frac{\sum_{k=1}^{n} u_{k,i}^2 \boldsymbol{z}_k^w}{\sum_{k=1}^{n} u_{k,i}^2}$$

n是完整数据的样本个数，

$\boldsymbol{v}_i$是第i个子类的类原型，

$\boldsymbol{z}_k^w$是第k个完整数据，$\boldsymbol{z}_k^w = [\boldsymbol{x}_k^w, y_k^w]$，

y_k^w是第k个完整数据的预测目标参数的数值，

$\boldsymbol{x}_k^w$是第k个完整数据的除预测目标参数之外其余参数数值组成的向量。

上述方案中，所述步骤 S2 中，通过如下公式计算不完整数据的隶属度：

$$\mu_{k,i} = \sqrt{\frac{s}{\sum_{j=1}^{s} I_{k,j}} \sum_{j=1}^{s} \left(z_{j,k} - v_{i,j}\right)^2 I_{kj}}$$

$$I_{k,j} = \begin{cases} 0, if\, z_{k,j} \in \boldsymbol{X_M} \\ 1, if\, z_{k,j} \in \boldsymbol{X_P} \end{cases}$$

$\mu_{k,i}$是第k个不完整数据对第i个子类的隶属度，

$z_{k,j}$表示第k个数据的第j个参数的数值，

$v_{i,j}$表示第i个类原型的第j个参数的数值，

$\boldsymbol{X_M}$表示空值、野值、异常值组成的集合，

$\boldsymbol{X_P}$表示正常数值组成的集合，

s是参数的个数。

进一步的，所述步骤 S3 中，通过如下公式估计空值、野值、异常值的真实值，实现联合收获机大数据的重构：

$$x_{k,j} = \frac{\sum_{i=1}^{c} \mu_{k,i}^2 v_{i,j}}{\sum_{i=1}^{c} \mu_{k,i}^2}$$

$x_{k,j} \in \boldsymbol{X_M}$，表示数据中的空值、野值或异常值，

$v_{i,j}$表示第i个类原型的第j个参数的数值，

$\mu_{k,i}$表示第k个不完整数据对第i个子类的隶属度。

一种实现所述联合收获机大数据重构方法的装置，包括数据输入模块、数据重构模块、数据输出模块和数据存储模块；

所述数据输入模块用于获取农机现场作业数据；

所述数据重构模块用于接收数据输入模块输入的联合收获机大数据，并根据联合收获机大数据重构方法重构数据；

所述数据输出和保存模块用于输出联合收获机大数据的重构数据，保存联合收获机大数据的原始数据、重构数据、隶属度矩阵。

上述方案中，所述数据输入模块与联合收获机车载监测系统、远程监控系统或智慧农场平台连接，用于接收标定好空值、野值、异常值的联合收获机大数据。

上述方案中，所述数据重构模块用于在模糊聚类框架下，利用克里金法刻画数据参数之间的关联关系，获取完整数据的隶属度，再基于完整数据及其隶属度计算类原型，继而获得不完整数据的隶属度，最后根据类原型、空值、野值和异常值对应的完整数据的正常数值、不完整隶属度，估计空值、野值和异常值的真实数值，实现联合收获机大数据的重构。

上述方案中，所述数据输出和保存模块将联合收获机的重构大数据与隶属度矩阵输出至云端，将联合收获机原始大数据、重构大数据、隶属度矩阵保存至联合收获机车载监测系统。

与现有技术相比，本发明的有益效果是：

根据本发明的一个方式，提供一种联合收获机大数据重构方法，基于模糊聚类框架，利用克里金法刻画联合收获机不包括空值、野值或异常值的作业参数的完整数据的参数之间的关联关系，通过数据参数之间的关联关系，以及数据空间分布密度、均值等信息共同推断空值、野值、异常值的真实数值，提高联合收获机大数据重构的精准度，为后续联合收获机大数据建模与分析任务提供完备的数据支撑。

根据本发明的一个方式，提供一种实现联合收获机大数据重构方法的装置，采用数据输入模块用于获取农机现场作业数据；数据重构模块用于接收数据输入模块输入的联合收获机大数据，并根据联合收获机大数据重构方法重构数据；数据输出模块输出联合收获机大数据的重构数据；数据保存模块用于保存联合收获机大数据的原始数据、重构数据、隶属度矩阵，最终实现提高联合收获机大数据重构的精准度，为后续联合收获机大数据建模与分析任务提供完备的数据支撑。

注意，这些效果的记载不妨碍其他效果的存在。本发明的一个方式并不一定必须具有所有上述效果。可以从说明书、附图、权利要求书等的记载显而易见地看出并抽出上述以外的效果。

附图说明

图 1 是本发明一实施方式的流程示意图。

图 2 是本发明一实施方式的完整数据隶属度计算流程示意图。

图 3 是本发明一实施方式的联合收获机大数据重构装置的示意图。

图 4 是本发明一实施方式的联合收获机大数据重构方法获取的部分完整数据隶属度矩阵

图。

图 5 是本发明一实施方式的联合收获机大数据重构方法与其他数据重构方法的实验结果对比图。

具体实施方式

下面详细描述本发明的实施例，所述实施例的示例在附图中示出，其中自始至终相同或类似的标号表示相同或类似的元件或具有相同或类似功能的元件。下面通过参考附图描述的实施例是示例性的，旨在用于解释本发明，而不能理解为对本发明的限制。

在本发明的描述中，需要理解的是，术语“中心”、“纵向”、“横向”、“长度”、“宽度”、“厚度”、“前”、“后”、 “左”、“右”、“上”、“下”、“轴向”、“径向”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。此外，术语“第一”、“第二”仅用于描述目的，而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此，限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多个该特征。在本发明的描述中，“多个”的含义是两个或两个以上，除非另有明确具体的限定。

在本发明中，除非另有明确的规定和限定，术语“安装”、“相连”、“连接”、“固定”等术语应做广义理解，例如，可以是固定连接，也可以是可拆卸连接，或一体地连接；可以是机械连接，也可以是电连接；可以是直接相连，也可以通过中间媒介间接相连，可以是两个元件内部的连通。对于本领域的普通技术人员而言，可以根据具体情况理解上述术语在本发明中的具体含义。

图 1 所示一种联合收获机大数据重构方法，包括如下步骤：

步骤 S1、在模糊聚类框架下，利用克里金法刻画联合收获机不包括空值、野值或异常值的作业参数的完整数据的参数之间的关联关系，获取联合收获机完整数据的隶属度；

步骤 S2、基于步骤 S1 获得的联合收获机的完整数据及其隶属度计算类原型，继而获得包含空值、野值和或异常值的作业参数的不完整数据的隶属度；

步骤 S3、根据计算的到的类原型、不完整数据的正常数值、不完整数据的隶属度分别估计空值、野值和或异常值的真实数值，实现联合收获机大数据的重构。

如图 2 所示，所述完整数据的隶属度计算包括如下步骤：

步骤 S1）设定聚类的子类个数c，完整数据中的预测目标参数y，隶属度阈值θ，最大迭代次数g_{max}，初始迭代次数g设定为 1；

步骤 S2）使用模糊c均值算法对完整数据进行初始聚类，获得隶属度矩阵$\mathbf{U}_{\mathrm{FCM}}$，作

为初始隶属度矩阵$\mathbf{U}^{(0)}$；

步骤 S3）针对每个子类，选用隶属度数值大于θ的完整数据作为该子类的训练数据集；

步骤 S4）检查每个子类的训练数据集是否为空集，如果为空集，随机生成一个隶属度矩阵，返回步骤 S3），不为空集，执行下一步骤；

步骤 S5）针对每个子类，利用克里金法建立预测模型；

步骤 S6）利用每个子类的预测模型，获取对所有完整数据的预测值，构建如下预测值矩阵：

$$\mathbf{Y}=\begin{bmatrix} \widehat{y_{1,1}} & \cdots & \widehat{y_{1,i}} & \cdots & \widehat{y_{1,c}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \widehat{y_{k,1}} & \ddots & \widehat{y_{k,i}} & \ddots & \widehat{y_{k,c}} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \widehat{y_{n,1}} & \cdots & \widehat{y_{n,i}} & \vdots & \widehat{y_{n,c}} \end{bmatrix}$$

式中，$\mathbf{Y}$是预测值矩阵，

$\widehat{y_{k,i}}$是第i个子类预测模型对第k个完整数据的预测值；

步骤 S7）利用公式计算完整数据隶属度：

$$u_{k,i}=\left[\sum_{t=1}^{c}\left(\frac{\left\|y_k-\widehat{y_{k,i}}\right\|_2^2}{\left\|y_k-\widehat{y_{k,t}}\right\|_2^2}\right)\right]^{-1}$$

式中，$u_{k,i}$表示第k个完整数据对第i个子类的隶属度；

y_k是第k个完整数据预测目标参数的真实数值；

$\widehat{y_{k,i}}$是第i个子类预测模型对第k个完整数据的预测值；

$y_{k,t}$是第t个子类预测模型对第k个完整数据的预测值；

步骤 S8）对于任意i,j，如果满足 $\max\left|\mathrm{u}_{ik}^{(g)}-\mathrm{u}_{ik}^{(g-1)}\right|<\varepsilon$，或$g\geq g_{max}$，$\varepsilon$为预设阈值，将当前隶属度矩阵视为最终隶属度矩阵，否则，返回步骤 S3），同时，$g=g+1$。

所述利用克里金法建立预测模型包括如下步骤：

步骤 S5.1）假设$\boldsymbol{x}$和y满足如下关系,克里金法建立预测模型：

$$y=\sum_{j=1}^{p}\beta_j f_j(\boldsymbol{x})+Z(\boldsymbol{x})$$

式中，y是完整数据中的预测目标参数；

$\boldsymbol{x}$是除预测目标参数之外其余参数组成的向量；

$f_j(\boldsymbol{x})$为第j个基函数；

β_j为第j个基函数的系数；

p是基函数的个数；

$Z(\boldsymbol{x})$为高斯过程，满足如下条件：

$$E\big(Z(\boldsymbol{x})\big) = 0$$

$$E\left(Z(\boldsymbol{x}_i)Z\big(\boldsymbol{x}_j\big)\right) = \sigma^2\boldsymbol{R}(\boldsymbol{\theta}, \boldsymbol{x}_i, \boldsymbol{x}_j)$$

σ^2为样本方差；

$\boldsymbol{R}(\boldsymbol{\theta}, \boldsymbol{x}_i, \boldsymbol{x}_j)$为相关矩阵；

$\boldsymbol{\theta}$为相关矩阵的参数向量；

$\boldsymbol{x}_i$是第i个数据除预测目标参数之外的其余参数组成的向量；

$\boldsymbol{x}_j$是第j个数据除预测目标参数之外的其余参数组成的向量；

步骤 S5.2）构建基函数矩阵$\boldsymbol{F}$：

$$\boldsymbol{F} = (\boldsymbol{f}(\boldsymbol{x}_1), \boldsymbol{f}(\boldsymbol{x}_2), \dots, \boldsymbol{f}(\boldsymbol{x}_n))^T$$

其中n是训练数据的样本数；

$\boldsymbol{f}(\boldsymbol{x}_i)\ (i = 1,2, \dots, n)$表示第$i$个数据通过基函数得到的数值，如下所示：

$$\boldsymbol{f}(\boldsymbol{x}_i) = (f_1(\boldsymbol{x}_i), f_2(\boldsymbol{x}_i), .., f_p(\boldsymbol{x}_i))$$

基函数矩阵$\boldsymbol{F}$可改写为：

$$\boldsymbol{F} = \begin{pmatrix} f_1(\boldsymbol{x}_1) & \cdots & f_p(\boldsymbol{x}_1) \\ \vdots & \ddots & \vdots \\ f_1(\boldsymbol{x}_n) & \cdots & f_p(\boldsymbol{x}_n) \end{pmatrix}$$

步骤 S5.3）输入向量表示如下：

$$\boldsymbol{Y} = \boldsymbol{F}\boldsymbol{\beta} + \boldsymbol{Z}$$

$\boldsymbol{\beta}$是基函数系数组成的向量；

$\boldsymbol{Z}$是训练数据的高斯过程组成的向量：

$$\boldsymbol{Z} = (Z(\boldsymbol{x}_1), Z(\boldsymbol{x}_2), \dots, Z(\boldsymbol{x}_n))^T$$

步骤 S5.4）待预测样本$\boldsymbol{x}_{new}$的输出估计值为：

$$\hat{y}(\boldsymbol{x}_{new}) = c(\boldsymbol{x})^T\boldsymbol{Y}$$

$\boldsymbol{Y}$是训练数据预测目标参数组成的向量；

$c(\boldsymbol{x})^T$是的系数向量；

步骤 S5.5）样本$\boldsymbol{x}_{new}$的预测误差如下：

$$\hat{y}(\boldsymbol{x}_{new}) - y(\boldsymbol{x}_{new}) = (\boldsymbol{F}^T c(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}))^T \boldsymbol{\beta} + c(\boldsymbol{x})^T \boldsymbol{Z} - z(\boldsymbol{x})$$

$y(\boldsymbol{x}_{new})$是样本$\boldsymbol{x}_{new}$的真实输出值；

由线性无偏条件，得：

$$E(\hat{y}(\boldsymbol{x}_{new}) - y_{new}) = 0$$

$$\boldsymbol{F}^T c(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}) = 0$$

步骤 S5.6）计算回归均方差$\varphi(\boldsymbol{x})$：

$$\varphi(\boldsymbol{x}) = \sigma^2(1 + c(\boldsymbol{x})^T \boldsymbol{R} c(\boldsymbol{x}) - 2c(\boldsymbol{x})^T \boldsymbol{r}_{x_{new}})$$

$$\sigma^2 = D(\boldsymbol{Z}) = E(\boldsymbol{Z}^2) - \big(E(\boldsymbol{Z})\big)^2 = E(\boldsymbol{Z}^2)$$

其中$\boldsymbol{r}_{x_{new}}$是样本$\boldsymbol{x}_{new}$与训练数据输入通过相关矩阵得到的对应向量；

步骤 S5.7）建立如下优化问题求解模型最优参数：

$$\min \varphi(\boldsymbol{x}) = \sigma^2\big(1 + c(\boldsymbol{x})^T \boldsymbol{R} c(\boldsymbol{x}) - 2c(\boldsymbol{x})^T \boldsymbol{r}_{x_{new}}\big)$$

$$\text{s.t.} \boldsymbol{F}^T c(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}) = 0$$

s.t.表示约束条件，

步骤 S5.8）构建朗格朗日函数：

$$L(c(x), \lambda) = \sigma^2\big(1 + c(\boldsymbol{x})^T \boldsymbol{R} c(\boldsymbol{x}) - 2c(x)^T \boldsymbol{r}_{x_{new}}\big) - \lambda^T(\boldsymbol{F}^T c(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}))$$

其中λ是拉格朗日乘子；

求偏导，得：

$$\frac{\partial L(c(\boldsymbol{x}), \lambda)}{\partial c\big(c(\boldsymbol{x})\big)} = 2\sigma^2\big(\boldsymbol{R} c(\boldsymbol{x}) - \boldsymbol{r}_{x_{new}}\big) - \boldsymbol{F}\lambda = 0$$

$$\frac{\eth L(c(\boldsymbol{x}), \lambda)}{\eth \lambda} = \boldsymbol{F}^T c - \boldsymbol{f}(\boldsymbol{x}) = 0$$

得$\begin{cases} \lambda = 2\sigma^2(\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1}(\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r}_{x_{new}}) \\ c(x) = \boldsymbol{R}^{-1}(\boldsymbol{r}_{x_{new}} + \boldsymbol{F}(\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1}(\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r}_{x_{new}})) \end{cases}$

步骤 S5.9）$\boldsymbol{x}$的估计值通过下式获得：

$$\hat{y}(\boldsymbol{x}) = \boldsymbol{r}_{\boldsymbol{x}}^T \boldsymbol{R}^{-1} \boldsymbol{Y} + (\boldsymbol{f} - \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{r}_{\boldsymbol{x}})^T (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{Y}$$

由$\boldsymbol{F}\boldsymbol{\beta} \cong \boldsymbol{Y}$，得$\boldsymbol{\beta}$的最优估计：

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{F})^{-1} \boldsymbol{F}^T \boldsymbol{R}^{-1} \boldsymbol{Y}$$

步骤 S5.10）样本$\boldsymbol{x}_{new}$的输出值如下所示：

$$\hat{y}(\boldsymbol{x}_{new}) = \boldsymbol{f}(\boldsymbol{x}_{new})^T\widehat{\boldsymbol{\beta}} + \boldsymbol{r}_{\boldsymbol{x}_{new}}^T\boldsymbol{R}^{-1}(\boldsymbol{Y} - \boldsymbol{F}\boldsymbol{\beta})$$

所述步骤 S2 利用如下公式计算类原型：

$$\boldsymbol{v}_i = \frac{\sum_{k=1}^{n} u_{k,i}^2 \boldsymbol{z}_k^w}{\sum_{k=1}^{n} u_{k,i}^2}$$

n是完整数据的样本个数；

$\boldsymbol{v}_i$是第i个子类的类原型；

$\boldsymbol{z}_k^w$是第k个完整数据，$\boldsymbol{z}_k^w = [\boldsymbol{x}_k^w, y_k^w]$；

y_k^w是第k个完整数据的预测目标参数的数值；

$\boldsymbol{x}_k^w$是第k个完整数据的除预测目标参数之外其余参数数值组成的向量。

所述步骤 S2 利用如下公式计算不完整数据的隶属度：

$$\mu_{k,i} = \sqrt{\frac{s}{\sum_{j=1}^{s} I_{k,j}} \sum_{j=1}^{s} \left(z_{j,k} - v_{i,j}\right)^2 I_{kj}}$$

$$I_{k,j} = \begin{cases} 0, if\, z_{k,j} \in \boldsymbol{X_M} \\ 1, if\, z_{k,j} \in \boldsymbol{X_P} \end{cases}$$

其中$\mu_{k,i}$是第k个不完整数据对第i个子类的隶属度，

$z_{k,j}$表示第k个数据的第j个参数的数值，

是第个不完整数据对第个子类的隶属度；表示第个数据的第个参数的数值；

$v_{i,j}$表示第i个类原型的第j个参数的数值；

$\boldsymbol{X_M}$表示空值、野值、异常值组成的集合；

$\boldsymbol{X_P}$表示正常数值组成的集合；

s是参数的个数。

所述步骤 S2 利用如下公式估计空值、野值、异常值的真实值：

$$x_{j,k} = \frac{\sum_{i=1}^{c} \mu_{ki}^2 v_{i,j}}{\sum_{i=1}^{c} \mu_{k,i}^2}$$

其中$x_{j,k} \in \boldsymbol{X_M}$，表示数据中的空值、野值或异常值；

$v_{i,j}$表示第i个类原型的第j个参数的数值；

$\mu_{k,i}$表示第k个不完整数据对第i个子类的隶属度。

为了更好地说明本发明所述联合收获机大数据重构方法的优势，引入以下 2 种常用的数据重构方法作为对比。

1. 参数均值重构方法

在参数均值重构方法中，将联合收获机的空值、野值、异常值对应参数正常值的均值视为其正常数值。

2. 最近邻原型重构方法

在最近邻原型重构方法中，利用模糊 c 均值算法对完整数据聚类，并将空值、野值、异常值的最近类原型的对应参数值视为其真实数值，如下所示。

$$x = v_i;$$

式中，x是数据中的空值、野值或异常值；

v_i是通过模糊c均值算法获取的最近邻类原型中，空值、野值或异常值对应的参数值。

利用模糊c均值算法对完整数据的聚类过程如下：

步骤（1）设定聚类的子类数目c，最大迭代次数g_{max}，迭代次数g设定为 1；

步骤（2）生成一个随机矩阵$\mathbf{U}$作为初始隶属度矩阵；

$$\mathbf{U} = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1c} \\ \vdots & \ddots & \vdots \\ \mu_{n1} & \cdots & \mu_{nc} \end{bmatrix}$$

式中，n是完整数据样本个数；

c是子类个数；

$\mu_{i,k}$表示第i个完整数据对第j个子类的隶属程度，取值从 0 到 1，并满足如下条件：

$$\sum_{i=1}^{c} u_{i,k} = 1 \left(k = 1, 2, \ldots, n;\ \forall i, k: u_{i,k} \in [0,1]\right)$$

步骤（3）计算每个子类的类原型；

$$\boldsymbol{v}_i = \frac{\sum_{k=1}^{n} u_{i,k}^2 \boldsymbol{z}_k}{\sum_{k=1}^{n} u_{i,k}^2}$$

式中，$\boldsymbol{v}_i$是第i个子类的类原型；

$\boldsymbol{z}_k$是第k个完整数据；

步骤（4）更新隶属度矩阵；

$$u_{i,k} = \left[\sum_{t=1}^{c} \left(\frac{\|\boldsymbol{z}_k - \boldsymbol{v}_i\|_2^2}{\|\boldsymbol{z}_k - \boldsymbol{v}_t\|_2^2}\right)\right]^{-1}$$

步骤（5）对于任意i, j，如果满足 $\max\left|\mathrm{u}_{ik}^{(g)} - \mathrm{u}_{ik}^{(g-1)}\right| < \varepsilon$，或$g \geq g_{max}$，将当前类原型视

为最终类原型；否则，返回步骤（3），$g = g + 1$。

根据本实施例，优选的，联合收获机收获作物为水稻，数据样本容量 150，包括割台高度、前进速度、风机转速、脱离间隙、清选损失等 11 个参数。

随机选取 10%的数据中作为不完整数据，随机抽取其中的一个数值作为待重构数值，剩余数据作为完整数据。

利用所述联合收获机大数据重构方法、参数均值重构方法、最近邻原型重构方法估计联合收获机大数据中的空值、野值和异常值的真实值。

优选的，本发明提供的一种联合收获机大数据重构方法的参数设定如下：

聚类数设定为 3；

克里金法的预测目标参数设定为清选损失数；

隶属度阈值设定为 0.5；

最大迭代次数设定为 20。

最近邻原型重构方法参数设定如下：

聚类数设定为 3；

最大迭代次数设定为 20。

进一步详细说明所述联合收获机大数据重构方法的数据重构过程：

1. 计算完整数据的隶属度。利用克里金法刻画数据参数之间的关联关系，获取完整数据的隶属度，其中部分数据的隶属度如附图 3 所示。可以看出，数据对于不同子类的隶属程度有所不同。以第一个数据为例，隶属度分别为 0.91，0.08，0.01，其属于第一个子类的可能性最大，为 0.91，因此其归为第一类。

2. 计算完整数据的类原型。基于完整数据及其隶属度，计算完整数据的类原型。

3. 计算不完整数据的隶属度。基于不完整数据和类原型计算其隶属度。

4. 计算空值、野值、异常值的真实数值。基于空值、野值、异常值对应的完整数据的参数值，结合获取的不完整数据隶属度，计算空值、野值、异常值的真实数值，部分不完整数据隶属结果如附图 4 所示。

所述联合收获机大数据重构方法、参数均值重构方法、最近邻原型重构方法获取的数据重构结果如表 1 所示。

表 1 数据重构结果

原始数值	本发明提供方法		参数均值重构方法		最近邻原型重构方法	
	预测值	误差	预测值	误差	预测值	误差
784	795.34	1.45%	664.82	15.20%	901.11	14.94%

951	963.54	1.32%	787.91	17.15%	983.74	3.44%
692	711.25	2.78%	787.91	13.86%	617.00	10.84%
213	214.19	0.56%	172.78	18.88%	217.61	2.16%
平均	-	1.53%	-	16.27%	-	7.85%

本发明所述联合收获机大数据重构方法的平均重构误差为 1.53%，明显低于参数均值重构方法（16.27%）和最近邻原型重构方法（7.85%）。

本发明所述联合收获机大数据重构方法的重构误差均在 3.00%以下，参数均值重构方法误差均高于 10.00%，最近邻原型重构方法部分数据重构误差低于 5.00%，但仍高于本发明所述联合收获机大数据重构方法。由此可见，本发明所述联合收获机大数据重构方法提高了大数据空值、野值、异常值重构的精准度，能够实现更为精准的联合收获机大数据重构。

实施例 2

如图 5 所示，本实施例提供一种实现实施例 1 所述联合收获机大数据重构方法的装置，因此具有实施例 1 的有益效果，此处不再赘述。

所述联合收获机大数据重构方法的装置包括数据输入模块、数据重构模块、数据输出和数据存储模块。所述装置可集成于但不限于联合收获机车载监测、远程监控、智慧农场等平台。

所述数据输入模块对接联合收获机车载监测、远程监控、智慧农场等平台，接收已经标定好空值、野值、异常值的联合收获机大数据；

所述数据重构模块根据本发明所述联合收获机大数据重构方法，重构、联合收获机大数据；

所述数据输出和数据存模块输出联合收获机大数据的重构数据，保存联合收获机大数据的原始数据、重构数据、隶属度矩阵。

在本实施例中，优选的，将本发明实现所述联合收获机大数据重构方法的装置集成于联合收获机车载监测系统，主要用于在数据传输至云端前，将获取的联合收获机大数据进行重构。

联合收获机车载监测系统利用各子系统数据采集装置获取联合收获机大数据，通过 CAN 总线汇总。

联合收获机车载监测系统通过集成的空值判断、野值识别、异常识别等算法标定数据中的空值、野值与异常值。

将标定的联合收获机大数据输入到本发明实现所述联合收获机大数据重构方法的装置中的数据输入模块；

利用本发明提供的一种联合收获机大数据重构方法，重构输入的联合收获机大数据。

通过本发明提供的一种实现所述联合收获机大数据重构方法的装置的数据输出模块和数据存储模块，将联合收获机的重构大数据与隶属度矩阵输出至云端，将联合收获机原始大数据、重构大数据、隶属度矩阵保存至联合收获机车载监测系统。

以上公开的本发明优选实施例只是用于帮助阐述本发明。优选实施例并没有详尽叙述所有的细节，也不限制该发明的具体实施方式。本说明书选取并具体描述实施例，是为了更好地解释本发明的原理和实际应用，从而使所属技术领域技术人员能很好地理解和利用本发明。本发明仅受权利要求书及其全部范围和等效物的限制。

权 利 要 求 书

1.一种联合收获机大数据重构方法，其特征在于，包括如下步骤：

步骤 S1、在模糊聚类框架下，利用克里金法刻画联合收获机不包括空值、野值或异常值的作业参数的完整数据的参数之间的关联关系，获取联合收获机完整数据的隶属度；

步骤 S2、基于步骤 S1 获得的联合收获机的完整数据及其隶属度计算类原型，继而获得包括空值、野值和或异常值的作业参数的不完整数据的隶属度；

步骤 S3、 根据计算的到的类原型、不完整数据的正常数值、不完整数据的隶属度分别估计空值、野值和或异常值的真实数值，实现联合收获机大数据的重构。

2.根据权利要求 1 所述的联合收获机大数据重构方法，其特征在于，所述步骤 S1 中，采用克里金法刻画数据参数之间的关联关系，通过如下公式计算完整数据的隶属度：

$$u_{k,i}=\left[\sum_{t=1}^{c}\left(\frac{\|y_k-\widehat{y_{k,i}}\|_2^2}{\|y_k-\widehat{y_{k,t}}\|_2^2}\right)\right]^{-1}$$

$u_{k,i}$是第k个完整数据对第i个子类的隶属度；

c是聚类的子类个数；

y_k是第k个完整数据的预测目标参数的数值；

$\widehat{y_{k,i}}$是第i个子类的克里金模型给出第k个完整数据的预测目标参数的预测值；

$\widehat{y_{k,t}}$是第t个子类的克里金模型给出第k个完整数据的预测目标参数的预测值。

3.根据权利要求 2 所述的联合收获机大数据重构方法，其特征在于，所述步骤 S1 中，聚类的子类个数c为 3，隶属度阈值θ，一般取为为 0.5。

4.根据权利要求 1 所述的联合收获机大数据重构方法，其特征在于，所述步骤 S2 中，通过如下公式计算的类原型：

$$\boldsymbol{v}_i=\frac{\sum_{k=1}^{n}u_{k,i}^2\boldsymbol{z}_k^w}{\sum_{k=1}^{n}u_{k,i}^2}$$

n是完整数据的样本个数，

$\boldsymbol{v}_i$是第i个子类的类原型，

$\boldsymbol{z}_k^w$是第k个完整数据，$\boldsymbol{z}_k^w=[\boldsymbol{x}_k^w,y_k^w]$，

y_k^w是第k个完整数据的预测目标参数的数值，

$\boldsymbol{x}_k^w$是第k个完整数据的除预测目标参数之外其余参数数值组成的向量。

5.根据权利要求 1 所述的联合收获机大数据重构方法，其特征在于，所述步骤 S2 中，通过如下公式计算不完整数据的隶属度：

$$\mu_{k,i}=\sqrt{\frac{s}{\sum_{j=1}^{s}I_{k,j}}\sum_{j=1}^{s}\left(z_{j,k}-v_{i,j}\right)^2 I_{kj}}$$

$$I_{k,j}=\begin{cases}0, if\, z_{k,j}\in \boldsymbol{X_M}\\ 1, if\, z_{k,j}\in \boldsymbol{X_P}\end{cases}$$

$\mu_{k,i}$是第k个不完整数据对第i个子类的隶属度，

$z_{k,j}$表示第k个数据的第j个参数的数值，

$v_{i,j}$表示第i个类原型的第j个参数的数值，

$\boldsymbol{X_M}$表示空值、野值、异常值组成的集合，

$\boldsymbol{X_P}$表示正常数值组成的集合，

s是参数的个数。

6. 根据权利要求 5 所述的联合收获机大数据重构方法，其特征在于，所述步骤 S3 中，通过如下公式估计空值、野值、异常值的真实值，实现联合收获机大数据的重构：

$$x_{k,j}=\frac{\sum_{i=1}^{c}\mu_{k,i}^2 v_{i,j}}{\sum_{i=1}^{c}\mu_{k,i}^2}$$

$x_{k,j}\in \boldsymbol{X_M}$，表示数据中的空值、野值或异常值，

$v_{i,j}$表示第i个类原型的第j个参数的数值，

$\mu_{k,i}$表示第k个不完整数据对第i个子类的隶属度。

7.一种实现权利要求 1-6 任意一项所述联合收获机大数据重构方法的装置，其特征在于，包括数据输入模块、数据重构模块、数据输出模块和数据存储模块；

所述数据输入模块用于获取农机现场作业数据；

所述数据重构模块用于接收数据输入模块输入的联合收获机大数据，并根据联合收获机大数据重构方法重构数据；

所述数据输出和保存模块用于输出联合收获机大数据的重构数据，保存联合收获机大数据的原始数据、重构数据、隶属度矩阵。

8.根据权利要求 7 所述的联合收获机大数据重构方法的装置，其特征在于，所述数据输入模块与联合收获机车载监测系统、远程监控系统或智慧农场平台连接，用于接收标定好空值、野值、异常值的联合收获机大数据。

9. 根据权利要求 8 所述的联合收获机大数据重构方法的装置，其特征在于，所述数据重构模块用于在模糊聚类框架下，利用克里金法刻画数据参数之间的关联关系，获取完整数据

的隶属度，再基于完整数据及其隶属度计算类原型，继而获得不完整数据的隶属度，最后根据类原型、空值、野值和异常值对应的完整数据的正常数值、不完整隶属度估计空值、野值和异常值的真实数值，实现联合收获机大数据的重构。

10. 根据权利要求 8 所述的联合收获机大数据重构方法的装置，其特征在于，所述数据输出和保存模块将联合收获机的重构大数据与隶属度矩阵输出至云端，将联合收获机原始大数据、重构大数据、隶属度矩阵保存至联合收获机车载监测系统。

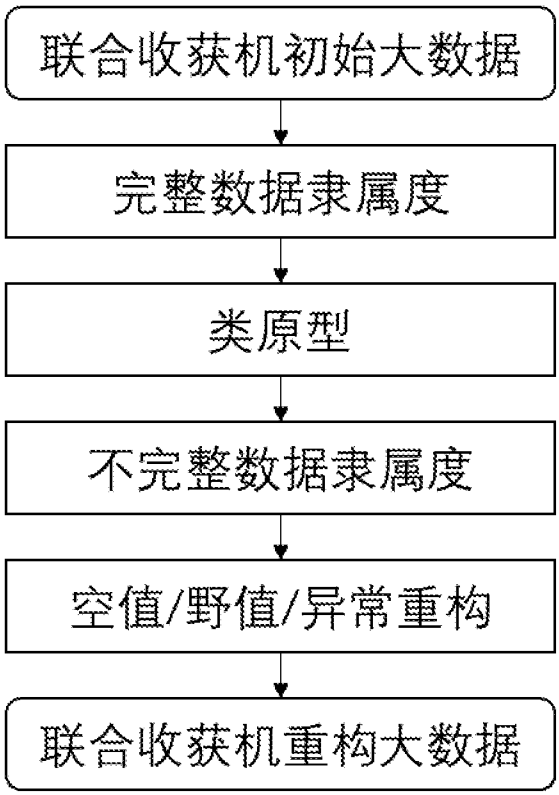


图 1

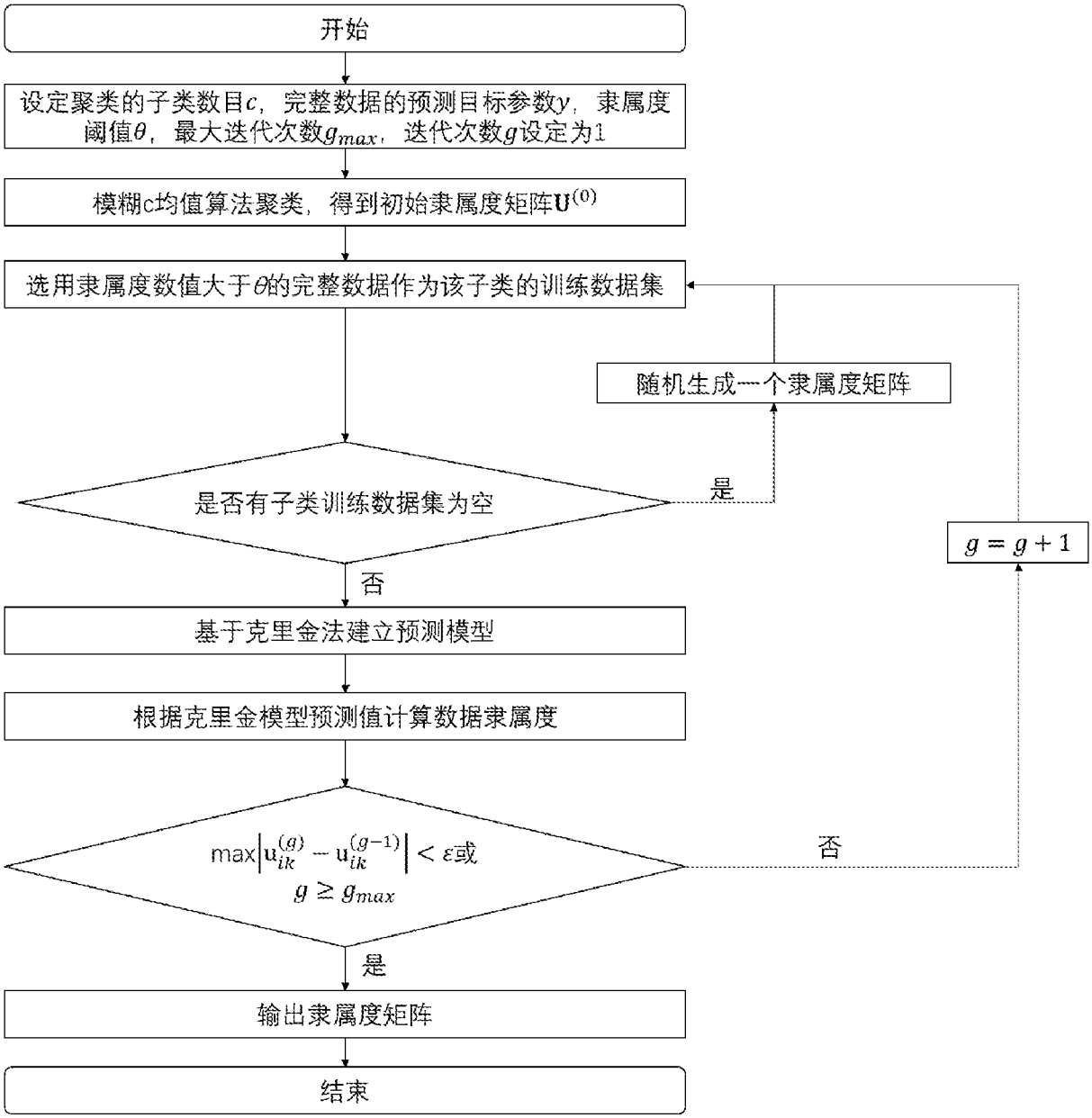

开始
设定聚类的子类数目c，完整数据的预测目标参数y，隶属度阈值θ，最大迭代次数g_{max}，迭代次数g设定为1
模糊c均值算法聚类，得到初始隶属度矩阵$\mathbf{U}^{(0)}$
选用隶属度数值大于θ的完整数据作为该子类的训练数据集
随机生成一个隶属度矩阵
是否有子类训练数据集为空
是
否
$g = g + 1$
基于克里金法建立预测模型
根据克里金模型预测值计算数据隶属度
$\max\left|\mathbf{u}_{ik}^{(g)} - \mathbf{u}_{ik}^{(g-1)}\right| < \varepsilon$或 $g \geq g_{max}$
否
是
输出隶属度矩阵
结束

图 2

$$\begin{bmatrix} 0.91 & 0.08 & 0.01 \\ 0.86 & 0.09 & 0.05 \\ 0.74 & 0.05 & 0.21 \\ 0.62 & 0.21 & 0.17 \\ 0.19 & 0.09 & 0.72 \\ 0.04 & 0.27 & 0.69 \\ 0.12 & 0.77 & 0.11 \\ 0.20 & 0.78 & 0.02 \\ 0.04 & 0.93 & 0.03 \end{bmatrix}$$

图 3

$$\begin{bmatrix} 0.87 & 0.11 & 0.02 \\ 0.81 & 0.14 & 0.05 \\ 0.82 & 0.09 & 0.09 \\ 0.10 & 0.54 & 0.36 \end{bmatrix}$$

图 4

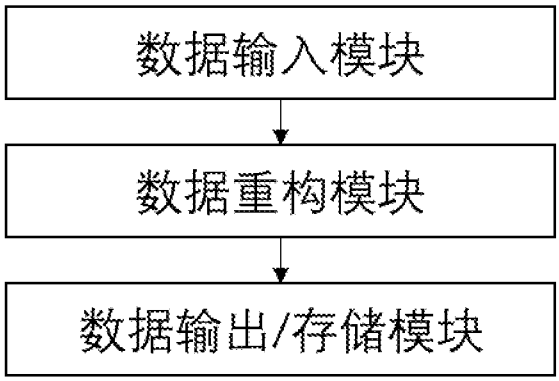


图 5

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2022/144067

A. CLASSIFICATION OF SUBJECT MATTER

G06F18/2337(2023.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC:G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS; CNTXT; EPTXT; VEN; WOTXT; USTXT; CNKI; IEEE: 模糊, 聚类, 克里金, 异常, 关联, 隶属度, 真实, 估计, 重构, 联合收获机, fuzzy, cluster, Kriging, abnormal, related, membership, real, estimate, reconstruct, combine harvester

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 115423005 A (JIANGSU UNIVERSITY) 02 December 2022 (2022-12-02) claims 1-10	1-10
PX	SHI, MaoLin et al. "A robust prediction method based on Kriging method and fuzzy c-means algorithm with application to a combine harvester" *STRUCTURAL AND MULTIDISCIPLINARY OPTIMIZATION,* Vol. 65, No. 9, 30 September 2022 (2022-09-30), 1-18 text, pages 1-15	1-10
X	张立勇 (ZHANG, Liyong). "基于信息粒的模糊聚类方法研究 (Non-official translation: Research on Fuzzy Clustering Method Based on Information Granules)" *中国博士学位论文全文数据库信息科技辑 (月刊) (Information Science & Technology (Monthly), C鰀na Doctoral Dissertations Full-Text Database),* No. 12, 15 December 2018 (2018-12-15), I138-50 ISSN: 1674-022X, text, pages 3 and 41-101	1-10

☑ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
04 May 2023	**09 May 2023**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **China No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088**	Telephone No.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	王鹭 (WANG, Lu). "改进蚁群优化的区间监督模糊聚类算法研究 (Non-official translation: Research on Interval Supervised Fuzzy Clustering Algorithm for Improved Ant Colony Optimization)" *中国优秀硕士学位论文全文数据库信息科技辑 (月刊) (Information Science & Technology (Monthly), C*[illegible]*na Master's Theses Full-text Database),* No. 02, 15 December 2017 (2017-12-15), I138-2185 ISSN: 1674-0246, text, pages 9-47	1-10
A	CN 107729943 A (LIAONING UNIVERSITY) 23 February 2018 (2018-02-23) entire document	1-10
A	JP 2001032679 A (OHBAYASHI CORP.) 06 February 2001 (2001-02-06) entire document	1-10

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2022/144067

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 115423005 A	02 December 2022	None	
CN 107729943 A	23 February 2018	CN 107729943 B	30 November 2021
JP 2001032679 A	06 February 2001	JP 3697957 B2	21 September 2005

国际检索报告

国际申请号
PCT/CN2022/144067

A. 主题的分类

G06F18/2337(2023.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

IPC:G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNABS;CNTXT;EPTXT;VEN;WOTXT;USTXT;CNKI;IEEE: 模糊，聚类，克里金，异常，关联，隶属度，真实，估计，重构，联合收获机，fuzzy，cluster，Kriging，abnormal，related，membership，real，estimate，reconstruct，combine harvester

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
PX	CN 115423005 A (江苏大学) 2022年12月2日 (2022-12-02) 权利要求1-10	1-10
PX	SHI, MaoLin 等. "A robust prediction method based on Kriging method and fuzzy c-means algorithm with application to a combine harvester" STRUCTURAL AND MULTIDISCIPLINARY OPTIMIZATION, 第65卷，第9期，2022年9月30日 (2022-09-30)，1-18 正文第1-15页	1-10
X	张立勇. "基于信息粒的模糊聚类方法研究" 中国博士学位论文全文数据库信息科技辑（月刊）， 第12期，2018年12月15日 (2018-12-15)，I138-50 ISSN：1674-022X, 正文第3、41-101页	1-10

☑ 其余文件在C栏的续页中列出。 ☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“D” 申请人在国际申请中引证的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2023年5月4日	2023年5月9日
ISA/CN的名称和邮寄地址 中国国家知识产权局 中国北京市海淀区蓟门桥西土城路6号 100088	受权官员 赵静 电话号码 (+86) 020-28958139

国际检索报告

国际申请号
PCT/CN2022/144067

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
X	王鹭. "改进蚁群优化的区间监督模糊聚类算法研究" 中国优秀硕士学位论文全文数据库信息科技辑（月刊）， 第02期，2017年12月15日 (2017 - 12 - 15)，I138-2185 ISSN：1674-0246, 正文第9-47页	1-10
A	CN 107729943 A (辽宁大学) 2018年2月23日 (2018 - 02 - 23) 全文	1-10
A	JP 2001032679 A (OHBAYASHI CORP) 2001年2月6日 (2001 - 02 - 06) 全文	1-10

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2022/144067

检索报告引用的专利文件	公布日 (年/月/日)	同族专利	公布日 (年/月/日)
CN 115423005 A	2022年12月2日	无	
CN 107729943 A	2018年2月23日	CN 107729943 B	2021年11月30日
JP 2001032679 A	2001年2月6日	JP 3697957 B2	2005年9月21日