

T/JAASS

江苏省农学会团体标准

T/JAASS XXXX—2025

基于大数据的二化螟成虫预测规范

Rule for adults population prediction based on Big Data of rice stem borer *Chilo suppressalis* (Walker)

(征求意见稿)

2025 – XX – XX 发布

2025 – XX – XX 实施

江苏省农学会 发布

目 次

前言 II

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 模型开发原则 1

5 模型开发步骤 2

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由江苏省农学会提出并归口。

本文件起草单位：江苏省农业科学院，江苏省植物保护植物检疫站，北京金禾天成科技有限公司。

本文件主要起草人：罗光华，鲍海波，周晨，方继朝，徐玮，徐晴玉，鞠佳菲。

基于大数据的二化螟成虫预测规范

1 范围

本文件规定了利用大数据预测二化螟成虫种群动态的原则、步骤。
本文件适用于不同稻作区水稻二化螟的成虫种群动态预测。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 15792-2009 水稻二化螟测报调查规范

3 术语和定义

下列术语和定义适用于本文件。

3.1

二化螟 *Chilo suppressalis* (Walker)

属鳞翅目，螟蛾总科草螟科，是水稻上的重要害虫之一。以幼虫为害，水稻苗期至穗期均可遭受危害。水稻分蘖期受害，产生枯鞘和枯心苗；孕穗抽穗期受害，产生枯孕穗和白穗；灌浆乳熟期受害，产生枯穗和虫伤株。

3.2

预测靶标 prediction targets

模型预测的对象，本规程所预测的对象是二化螟的每日成虫数量，以“头/天”表示。二化螟各代次的每日成虫数量基于各代次的发生区间得出。

3.3

预测准确率 prediction accuracy

模型预测靶标的预测值与实际观察值的百分比。

3.4

变量 variable quantity

预测模型中，对预测靶标有重要影响的因子，主要包括：每日由诱捕器获得的虫量、田间越冬幼虫量、每日降水量、每日最低温和最高温、每日平均气温、每日平均相对湿度。

4 模型开发原则

4.1 目标

二化螟成虫数量预测准确率在85%以上，见计算公式（1）。

$$1 - |Y - S|/S \times 100\% \dots\dots\dots (1)$$

式中：

Y——表示预测值；

S——表示实际观察值。

4.2 策略

将二化螟每日成虫数量作为预测靶标；将每日由诱捕器获得的虫量、田间越冬幼虫量、每日降水量、每日最低温和最高温、每日平均气温、每日平均相对湿度作为各项变量参数，其中每日由诱捕器获得的虫量作为主变量，其余各项作为协变量。

4.3 算法

利用基于时间序列预测的DeepAR算法。将历史资料数据集划分为训练集和测试集，其中训练集数据量远远多于测试集；训练集用于模型的开发，测试集用于模型预测准确度的测试。

5 模型开发步骤

5.1 数据粒度

所有数据集统一按照年度的日粒度（每日）进行呈现。

5.2 数据排布

所有日粒度数据，按照日期顺序从每年的1月1日起至每年的12月31日止排列。数值有缺失的日期用N替代。

5.3 数据格式

排列整齐的数据，利用R语言工具包转换成DeepAR程序可识别的.CSV文档。

5.4 数据导入

将完整的.CSV数据文档导入DeepAR程序进行模型构建。

5.5 参数调整

为了获得符合要求的模型，进一步调整相关参数，包括：epoch（迭代次数）、batch size（训练批量大小）、学习率、神经网络结构、分布函数类型、协变量的使用等，最终建立较理想的二化螟成虫种群动态预测模型。
