

文章编号:0559-9350(2020)10-1212-11

## 智慧灌区的架构、理论和方法之初探

史良胜, 查元源, 胡小龙, 杨琦

(武汉大学 水资源与水电工程科学国家重点实验室, 湖北 武汉 430072)

**摘要:** 我国灌区面临着多元化、精准化和智能化的管理需求。结合人工智能技术的发展趋势和水利学科本身特点, 尝试提出智慧灌区的定义和基本功能。从灌区的感知智能、认知智能和管理决策三个方面讨论了智慧灌区研究的难点, 初步论述了智慧灌区架构、理论和方法。在灌区感知智能方面, 提出非接触式与接触式观测相结合、移动式和固定式观测相结合的观测体系, 开发适用于灌区不同尺度观测的机器视觉技术、灌区特征解译技术和天地空数据智能交换技术; 建立大数据环境下的信息理论, 支撑灌区多源观测体系的研发和海量数据的解译和分析。在灌区认知智能方面, 提出物理方程和机器学习相结合的灌区建模思路, 建立“数据-物理”混合认知智能的理论基础, 开发非完整先验物理机制下灌区水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化模型, 开发灌区物理机制挖掘方法和动态建模方法。在灌区决策方面, 提出可满足众多决策者需求以及可处理多种目标的智能优化决策系统, 实现灌区尺度上的水量、水质和生态的最优化管理。

**关键词:** 智慧灌区; 人工智能; 大数据; 认知智能; 感知智能

**中图分类号:** S274

**文献标识码:** A

**doi:** 10.13243/j.cnki.slxb.20200169

灌区在我国农业和水利行业中具有举足轻重的地位。目前灌区管理面临新的需求, 包括: (1) 管理目标多元化。灌区是我国农业发展的主体, 如何提高农业水资源利用效率, 对于保障工业和生活用水供应至关重要。在水资源短缺现象日益严重、水土环境污染和生态退化问题加剧的背景下, 灌区管理从传统农业用水管理转变至农业、工业、生活和生态用水管理并重, 有些灌区还服务于旅游和科教产业。灌区管理涉及到多地区、多部门、多行业, 管理目标多元化使得灌区系统的复杂性显著增加; (2) 管理行为精准化。粗放型农业用水方式导致我国农田灌溉水有效利用系数仅为0.554(中国水资源公报, 2019), 且传统的灌区管理以农业用水为主, 缺乏生态和环境方面的监测与管理。为实现多目标优化管理, 需对灌区水资源配置、种植结构、农田水肥药措施等不同时间和空间尺度的行为进行精细化管理, 实时精准地实施抗旱、防洪、排涝、排渍、控污, 以及满足生产、生活、生态的用水需求, 达到水土资源的最优利用以及生态环境保护的目的; (3) 管理手段智能化。随着城镇化的发展, 农业就业人口数量快速下降、老龄化现象日趋突出。灌区管理在面临管理目标多样、管理任务繁重的同时, 也面临管理人员不足的压力; 因此, 需要借助智能化手段, 提高管理服务的效率和水平。总的来说, 我国灌区管理主要沿用传统的理论与技术手段。本文以智慧灌区为论述对象, 试图为未来智慧灌区发展提供架构、理论及相应方法层面的展望并结合人工智能的最新研究进展, 力图梳理出适用于灌区场景的人工智能技术特点以及需解决的难点。

### 1 智慧灌区的定义和基本功能

灌区是指灌溉水源工程所能控制的对农田实施灌溉的地域, 由灌溉水源工程、灌溉排水系统及

收稿日期: 2020-03-17; 网络首发时间: 2020-09-18

网络首发地址: <http://kns.cnki.net/kcms/detail/11.1882.TV.20200917.0944.002.html>

基金项目: 国家自然科学基金项目(51861125202)

作者简介: 史良胜(1982-), 教授, 主要从事农业水利研究。E-mail: liangshs@whu.edu

灌溉农田组成<sup>[1]</sup>。在灌区的管理上，通常以抗旱、防洪、排涝、排渍等作为主要目标。目前，灌区的管理从水量调控转变至水量、水质、生态调控并重，灌区中农业活动也已经从稳产、增产转变至保证产量品质并重。

人工智能(又称机器智能)，是指一个可以感知周围环境并采取行动以最大化概率实现其目标的系统。最近20年来，人工智能研究的核心是在推理、知识表达、规划、学习、自然语言处理、感知、移动与操控工具等方面构建与人类类似甚至超出人类的能力<sup>[2-3]</sup>。人工智能在图像和视频识别、游戏、语言、无人驾驶等方面取得了初步成果。然而，按照Pearl等<sup>[4]</sup>对因果关系的分级，目前的人工智能仅具有初级的关联能力，即可基于被动观测做出预测；更高级的、具有干预能力和反事实能力的人工智能尚待开发。

参照人工智能的定义，本文尝试将智慧灌区定义为：具有智能监测、解译、模拟、预警、决策和调控能力的灌区，全面实时感知灌区水情、墒情、工情、作物长势、生态环境等信息，快速、精准、自主调控水源、输配水及排水系统等工程设施及设备，实现水量、水质和生态等多目标的最优化管理。智慧灌区是现有灌区信息化、自动化和数字化的高级形式；它融合了人工智能技术，以期实现更为智能的灌区监测、信息解译、模拟、预警、决策和调控<sup>[5-6]</sup>，具备自主学习、分析和优化能力。智慧灌区依赖于灌区场景的机器智能，其在感知、认知、管理灌区方面具备超越人类的能力。

智慧灌区包括4个基本功能：(1)能够对不同尺度的灌区要素进行观测，从低信息含量的被动观测过渡到高信息含量的主动观测；(2)能够从多源数据中准确解译出灌区的水情、墒情、作物(植被)长势、生态、环境、工情等定量特征，自动识别出灌区干旱、涝渍、盐碱、水土流失、生态退化、环境污染等表征；(3)针对庞大、复杂的灌区系统，能够准确描述灌区的水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化，具备动态自主建模能力和模型进化能力，具有观测数据之外的推理能力；(4)能够自主、精准制定水资源调度和配置、水旱灾害防治、水环境修复、生物多样性保护等措施，可准确评估各管理行为的效应和效益并具备实时调整的能力。针对某些特定的灌区功能和管理目标，还可定义其它类似的智慧型概念，例如可针对灌溉、供排水等目标，定义智慧灌溉、智慧泵站等概念。智慧灌区建设以开发人工智能技术来提升灌区管理能力为核心，是目前“补短板、强监管”的有效手段，也是现代灌区的全新阶段。图1给出了智慧灌区架构的框图，下文将分别予以论述。

笔者认为智慧灌区不仅仅是人工智能技术在灌区内的应用，这主要是因为目前的人工智能技术过度依赖于数据，而对于复杂的灌区系统来说，难以进行空中、地表、土壤、地下全覆盖式的观测，也因此难以获取高维的、高密度的数据集。智慧灌区应是经典农田水利学、水文学、水力学、环境学、生态学等专业学科知识与人工智能的结合。前者为智慧灌区提供先验知识，后者处理数据

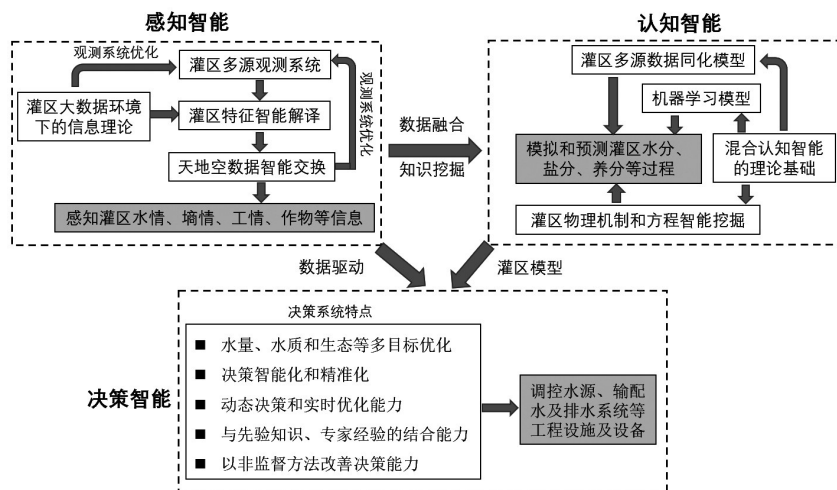


图1 智慧灌区架构

的能力更强,两者结合将更有利于灌区场景的人工智能技术的开发。

## 2 智慧灌区的观测体系

智能观测系统和数据解译系统构成了灌区感知智能,其目的在于快速、准确获取灌区的数字化表示,包括:地形,土地利用类型,土壤水分、盐分和养分状态,气象要素,干旱和洪涝状态,水库、湖泊、河道和沟渠水位、水质、流量、流速,地下水位和水质,作物(植被)生长状态,生物多样性,设备和建筑物(闸门、泵站、沟渠等)运行状态,农田管理信息等主要数据。数字化服务于灌区智能管理,数字化的规划精度一方面取决于灌区调控目标,也取决于技术精度与技术成本之间的均衡;在数据量和精度能够满足灌区特定调控目标的前提下,快速和低成本数字化优先于高精度和高成本数字化。

**2.1 灌区多源观测系统** 灌区传统的地面观测以点尺度观测为主(如墒情站),观测密度和代表性均不足。Li等<sup>[7]</sup>研究表明,点尺度的土壤含水量在农田管理中存在局限性,这主要是由于土壤含水量的强烈空间变异性削弱了点尺度数据的价值。卫星遥感虽然已广泛应用于灌区灾害评估,但仍未成为一种日常观测手段;目前卫星观测的时空分辨率也难以满足精准、精细化的管理要求。低空无人机观测在中尺度问题中已展现出了显著的优势,但在灌区中相应的技术和应用研究仍然滞后。随着卫星观测精度和分辨率的改善、无人机观测成本和地面传感器价格的降低,天地空一体化观测将成为未来灌区的基本观测架构。

目前构建天地空一体化观测体系存在诸多挑战:(1)天地空一体化观测体系的基础设施建设不足,现有体系还不能支持灌区的数字化,而数字化是智慧灌区的前提;(2)大规模布设地面接触式传感器(如TDR、FDR传感器)的成本非常昂贵且不利于田间机械作业,这意味着应该更加重视非接触式传感器;(3)天地空观测还没有形成协同效益,不同尺度和不同来源数据的协同利用缺乏严谨的理论支持和系统的实验研究;(4)灌区缺乏高时空分辨率和高通量的观测手段,缺乏针对海量数据的解译方法。

非接触式观测近年来获得飞速发展,但目前灌区的观测系统设计尚未予以重视。在河流和沟渠流量监测方面,图像测速方法正成为一种具有潜力的非接触式表面流速和断面流量测量方法<sup>[8]</sup>,有可能成为常规接触式测流法(如流速仪法)的替代方法。例如赵浩源等<sup>[9]</sup>在崇阳水文站的比测实验发现,图像测速法估测的表面流速与旋桨式流速仪实测结果之间的相对误差小于15%、流量误差小于5%。在干旱监测方面,土壤含水量和作物水分胁迫指数(CWSI)分别是最为常用的基于土壤的指标(接触式方法)和基于作物的指标(非接触式方法);随着土壤水分遥感监测理论的发展,特别是土壤热惯量和微波方法的兴起,非接触式的干旱监测手段日益丰富。非接触式干旱监测方法通过地表的近中远红外、热红外信号以及穿透冠层的微波信号来反映干旱,应用更为便捷,且兼容无人机等移动平台,然而相关技术在现有灌区中应用很少。在灌区植被表型观测方面(例如病虫害、LAI、作物长势)<sup>[10]</sup>,非接触式观测已能达到很高的精度。在水质监测方面,非接触式传感器已可测量叶绿素a含量、水温、溶解氧、总磷等指标<sup>[11]</sup>。在大坝等建筑物监测方面,全球导航卫星系统(GNSS)方法,在精度上虽略逊于正倒垂法等经典方法,但具有全天候、实时、自动化程度高等特点<sup>[12]</sup>。

非接触式观测大多基于光谱遥感技术;与此同时,激光探测及测距(LiDAR)技术也已经成为成熟应用于地形监测<sup>[13]</sup>、农作物特征(株高、生物量、作物氮素状态)<sup>[14-15]</sup>,以及大坝等建筑物的结构和变形监测等方面。随着低中空遥感技术的发展,这些技术的成本正在快速降低。总的来说,非接触式监测方法已在灌区得以广泛应用,但仍不是灌区观测的主流方法,这主要是因为其技术成熟度、观测精度和可靠性方面仍有待完善。智慧灌区观测系统应结合非接触式和接触式观测,接触式观测系统用于高精度的项目信息获取,以及在现有技术条件下,非接触式所无法替代的观测项目。

综上所述,笔者认为智慧灌区的基本观测架构具有如下特征:(1)接触式和非接触式观测相结合,移动式与固定式观测相结合。非接触式传感器以可见光、近中远红外、热红外、微波等电磁波

段、激光、声纳为探测手段,可搭载在地面固定设备、卫星、移动无人车、无人船和无人机上;相对于接触式观测而言,非接触式设备在单位观测成本以及运行维护上更有优势;(2)适用于灌区不同尺度观测的机器视觉技术。机器视觉是人工智能的基础技术之一,也是目前人工智能领域中成果最为丰富的分支。机器视觉在图像的颜色、形状、纹理和光谱分析上取得了显著进展<sup>[16]</sup>,应是灌区感知智能的基础组成。为了精准捕捉灌区特征的动态变化和三维结构特征<sup>[17-19]</sup>,灌区机器视觉还应具有高通量图像数据(超高分辨率图像、视频数据和超光谱图像)的收集和分析能力;(3)基于多源大数据的灌区特征解译技术。大数据不仅体现在数据容量和类型的增加,更体现在数据分析方法的创新,如卷积神经网络<sup>[20]</sup>和云技术<sup>[21]</sup>等。例如,对干旱的定量诊断,可融合土壤含水量、近中远红外、亮度温度、微波等多源数据中所含的干旱信息,也可融合不同尺度数据来提高估计精度<sup>[22]</sup>。相对于单一数据,多源大数据技术一方面为利用间接数据提供了途径,一方面可降低对单源观测的过度依赖,增加了观测系统的稳定性;(4)天地空数据智能交换技术。鉴于天地空观测手段各自的优劣势,应实现信息流在天地空观测设备之间的交换。数据交换的益处一方面在于可以实现天地空设备之间的互相校准和不同尺度数据的尺度升降,另一方面可根据管理目标,为由被动观测向自主触发观测提供前提。例如,当系统判断卫星数据不足以服务于管理目标时,可自主触发无人机观测,对特定区域自动进行高分辨率的巡测。

**2.2 灌区大数据环境下的信息理论和数据解译方法** 当面临大量的、不同类型的、不同低价值密度的数据时,就需要引入信息度指标(例如熵)以评估各种观测数据在量化灌区特征上的价值大小,并可以通过数据价值反向管理监测系统,提高数据采集效率,降低监测成本。因此,信息理论是灌区感知智能的基础理论之一,应解决如下4个问题:(1)评估某种数据所含的信息量;(2)评估多源数据联合包含的信息量;(3)评估某种数据所含的信息量有多少可通过其它数据推导出来;(4)评估多源数据包含的重复信息量<sup>[23]</sup>。在已有研究中,科学家们提出了边际熵(Marginal entropy)、联合熵(Joint entropy)、相对熵(Relative entropy)、交叉熵(Cross entropy)、转移信息(Transinformation)和全相关性(Total Correlation)等概念,并在水文观测站设计等问题中得以应用<sup>[23-24]</sup>。然而,已有的信息理论在处理多源大数据时是不足的。Calude等<sup>[25]</sup>证实当数据集过大时会出现任意的、无意义的相关性,导致“伪相关泛滥”问题,从而产生数据越多越难提取有效信息现象。不同于推荐系统、广告领域的大数据技术,灌区的多源大数据包含复杂的物理关系,如土壤含水量与热红外、微波反射时间和振幅等数据存在特定的但又具有一定不确定性的物理关系。由于相关性不等同于因果关系<sup>[4]</sup>,笔者认为在多源大数据分析过程,新的信息理论应服从于自然界固有的因果关系,尤其对于包含众多变量和管理目标的复杂系统。在多源大数据分析中考虑这些物理关系的约束,有利于解译出更有价值的信息。总而言之,灌区大数据的信息理论尚不完善,缺乏系统的理论分析和相应的实验研究。下文第3节将从模型的角度进一步论述该点的重要性。

由被动观测转变为主动观测是未来趋势,即根据管理需求,系统自主调整卫星轨道、高度和角度,触发无人机自动巡测,以及自主调整地面观测设备的高度、角度和位置,实现有针对性的观测。实现该功能的最大难点在于如何评估拟采集数据所包含的信息能在多大程度帮助系统实现预定的管理目标。因此,新的信息理论也是主动观测的基础理论,从而可以科学地评估和预判观测数据价值。

在数据解译方法上,传统的统计方法、机器学习方法以及深度学习的方法均有广泛应用。深度学习技术近期获得广泛关注,在灌区相关问题上已有初步的成果,如植被分类<sup>[26-27]</sup>、降雨降雪估计<sup>[28]</sup>、太阳辐射计算<sup>[29]</sup>、物候估计<sup>[30]</sup>、产量估计<sup>[31]</sup>、水位计算<sup>[32]</sup>和溶质参数估计<sup>[33]</sup>等。为同时利用多源数据,学者们提出多种改进的深度学习架构,如Yang等<sup>[31]</sup>提出了一种双支深度神经网络,分别解译可见光和多光谱数据;Yang等<sup>[30]</sup>以热时间作为辅助数据来约束卷积神经网络的计算结果。然而,目前的研究尚不足以支撑灌区不同尺度观测的机器视觉技术,也无法高效解译多源大数据,特别在超高分辨率图像、视频数据、超光谱图像、多源海量数据解译方面仍有待开发。深度学习技术在解译数据方面虽然直接高效,但其缺乏物理可解释性的不足有待弥补。Zhou等<sup>[33]</sup>最近通过深度卷积神经网络

络建立了含水层水力参数与宏观弥散度的映射关系，避免了密集观测水头和求解水流方程，极大地便捷了多孔介质宏观参数的推求；然而，如何在参数估计中满足内在的物理约束，服从连续性方程、能量方程和动量方程等基本方程值得深入研究。

### 3 智慧灌区的认知体系

土壤-植物-大气连续体(SPAC)、彭曼公式、圣维南方程、谢才公式、达西定律、对流弥散方程等基本概念和方程是灌区水分、盐分、养分、污染物迁移转化和作物生长模拟的基础。目前应用于灌区管理的软件也众多，包括SWAT、TOPMODEL和SHE等概念性流域水文模型，WOFOST、DS-SAT、ORYZA、EPIC和CropSyst等作物模型，Fluent、MODFLOW和HYDRUS等水动力学模型，SWAP、DRAINMOD、AquaCrop和AHC等田间水管理模型及其它等等。然而，仅对少数灌区问题(如短距离渠道输水)，模型才具有良好的精度；对于不同尺度耗水量估测、湖泊水库河道水质模拟、农田水土环境模拟、地下水运动及污染物运移、多渠段复杂明满流输配水系统自动控制等众多问题，这些概念、方法和模型仍不足以支撑灌区精准化和智能化的管理。Gupta等<sup>[34]</sup>指出，如何在复杂系统中调和过程模型与观测数据已成为现代环境科学的关键问题。Savenije<sup>[35]</sup>认为模型是基于特定理论的工具，概念模型不等同于现实。为对灌区水量、水质、生态实施精准管理，增强模型的现实性至关重要。模型与现实的脱钩，也是现有灌区模型难以胜任精准化和智能化智能管理的主要原因。这种脱钩体现在3个地方：(1)灌区数字化程度不足，模型运行缺乏足够的支撑。让模型在基于现实的数字环境中运行是灌区感知智能解决的主要问题；(2)对灌区的机理认识不足，尚未透彻地理解水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化过程。灌区水转化与伴生的多种过程在不同时空尺度上的机理和规律也有待深入揭示；(3)建模方法的不足。合理的模型架构应具有利用已知观测来描述仍含有大量未知信息(如土壤质地、水力性质等)的系统的功能<sup>[35]</sup>。后两点是灌区认知智能需解决的核心问题。

**3.1 灌区模拟的纯数据驱动方法和数据同化方法** 许多与灌区管理相关的模型采用了大量的假设和简化，导致难以准确模拟灌区内的各种过程。为了减弱或摆脱模型对复杂物理机制的依赖，多种数据驱动的模拟技术得以发展和应用，包括机器学习为代表的纯数据驱动方法和数据同化方法。

纯数据驱动方法是指具有特定架构的机器学习模型。这类模型通过历史数据来构建输入和输出之间的关系，从而达到直接预测的目的。在径流模拟方面，Hsu等<sup>[36]</sup>利用非线性神经网络模拟降雨-径流过程，取得了甚至优于物理模型的精度。Jain等<sup>[37]</sup>检验了在神经网络方法在降雨-径流模拟中保留物理机理的能力。后续研究者还结合了多种不同机器学习方法来进一步改进径流模拟<sup>[38-39]</sup>。在灌区耗水计算方面，由于蒸散发物理过程复杂，数学模型形式众多，不同模型均存在较大的局限性。例如能量平衡模型虽物理机制明确，但能量不闭合问题、气孔阻抗及地表粗糙长度参数化难题大大降低了模型的精度。考虑到蒸散发过程本身是一个复杂非线性的现象，其依赖于众多相互影响的气象要素，Kumar等<sup>[40]</sup>尝试摆脱特定的模型，利用神经网络直接建立气象要素与蒸散发关系。在灌溉排水系统控制方面，也有学者采用神经网络来预测沟渠水位<sup>[41]</sup>，但相关研究较少，目前主流模型还是依赖于水力学原理。可以看出，在涉及到复杂机理、多变量、多过程的问题中，纯数据驱动模型受到了更多的关注。

为弥补经典机理模型的不足，有学者尝试将机理模型与观测相结合，用于提高模型的性能<sup>[34]</sup>，该过程被称之为数据融合或数据同化。与纯数据驱动方法不同的是，数据同化方法以物理过程的数学模型为基础，通过融合观测数据不断调整模型的参数和状态，从而获得参数和状态的最优估计。数据同化方法虽已成功应用于水文预报<sup>[42]</sup>、地下水资源管理<sup>[43]</sup>、农田水分管理<sup>[44]</sup>等方面，但现有的数据同化技术依赖于物理模型，只能加入与物理模型变量对应的观测数据，尚无法融合多源大数据。

总而言之，随着灌区数据的累积，特别是灌区感知智能的建立，机器学习方法在建模中会起到越来越重要的作用。尽管纯数据驱动方法目前取得了一定的成功，但它属于黑箱模型，忽略了物理

过程的数学描述。对于强烈非线性问题,如果缺乏或没有足够多的输出变量的历史观测(或观测噪声大),机器学习方法则表现不佳<sup>[45]</sup>。由于缺乏极端样本,机器学习方法通常难以捕捉小概率事件;因此,应用纯数据驱动方法管理大坝等重要建筑物具有一定风险。相比而言,数据同化方法的机理性更强,外插效果更好。近年来,两种不同的数据利用方法有混合的趋势。例如,Wang等<sup>[24]</sup>最近在数据同化的框架内,通过机器学习(高斯过程回归)的方法将间接观测(土壤温度)融合到土壤含水量的估计。这种混合方法一定程度上保留了物理机制,同时避免了构建更为复杂的物理模型,为利用多源数据提供了一种思路。相关的研究仍处于初步阶段,如何避免多源大数据“伪相关泛滥”的干扰仍有待研究。第3.3节将深入论述混合方法。

**3.2 灌区的物理机制挖掘方法** 目前绝大多数灌区相关模型都具有固定的模型结构,即无论研究区域数据的丰富程度和物理过程的复杂程度如何,均采用具有固定结构的模型。同时,已有研究主要通过数学方程和参数来表达物理机制。

Fenicia等<sup>[46-47]</sup>提出了一种逐步建模法,随着观测数据的累积不断调整模型结构,让模型结构适应性地进化,在模型复杂度和数据量之间实现平衡。Clark等<sup>[48]</sup>提出了一种多假设建模框架,在建模过程中考虑物理过程的不同表达形式。Bui等<sup>[49]</sup>提出用神经模糊推理系统创造初始洪水风险模型,并通过元启发算法优化模型。目前已有多种机器学习方法(包括稀疏回归、高斯过程和人工神经网络方法等)从数据反向推理常微分和偏微分方程。但是,在实际问题中挖掘物理机理面临着众多挑战<sup>[50]</sup>,例如:如何处理大噪声和异常数据,如何处理来自不同实验环境的数据,以及如何在没有观测数据的区间内进行外插等。在复杂的灌区系统中,数据大多来源于非控制性的环境,数据的噪声通常也是未知的,物理机制的反向推理尤其困难。

借助机器学习技术和新的建模思路推理物理机制、自动构建灌区某些过程描述方法或针对调控目标自主制定决策模型应是灌区认知智能的重要组成部分。在灌区感知智能成熟的前提下,监测系统可以在不同尺度上收集灌区的各类数据,能否利用这些数据,结合物理方程挖掘技术和建模方法,来重新理解灌区的水分、盐分、养分、污染物迁移转化规律;能否在灌区各类错综复杂的因果关系未知或部分已知的条件下,建立具有更加完整机制的模型或根据数据自动调整模型结构,提高模拟和预测精度,特别是改善观测数据区间之外的预测能力。相关的研究非常欠缺。

**3.3 灌区的混合认知智能** 针对真实的灌区系统,现有的各种模型均不具备完整的物理机制,也难以精准预测灌区的水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化规律。以SPAC系统为例,从概念化物理过程的角度来说,目前SPAC系统使用到的经典理论本身(如彭曼公式、Richards方程、对流弥散方程等)还难以描述自然界复杂的、完整的物理机制,原因包括三个方面:(1)在建模过程中,需要对土壤、水分、溶质、冠层等研究对象做理想性的概化。孔隙率、多孔介质理想流体、叶面积指数(LAI)等统计性质的概念虽便于方程和模型的建立,但也可能削弱模型中物理问题的真实性。(2)SPAC系统侧重于刻画土壤水运动、蒸散发、光合作用、干物质累积和分配等过程,而忽略或简化了很多其它过程。例如,现有模型大多采用简化的函数来描述作物根系分布,而很难考虑根系趋水、趋肥、趋氧和趋温的生长特征。(3)农田中各个过程受到众多因素的影响,而在构建模型的过程中无法考虑所有因素与模型输出之间的因果关系。例如,微气象对农田耗水具有重要影响,但由于监测难度大,现有模型大多忽略了田间的微气象条件。鉴于灌区环境的复杂性和时空变异性,现有灌区模型在还存在诸多不足,经常出现模型愈复杂、预测能力愈差的现象。

建立包含所有物理机制的过程模型是很困难的。在灌区各类错综复杂的因果关系未知或部分已知的条件下,应建立非完整先验物理机制下的灌区模型,降低灌区系统模拟精度对模型结构完整性的依赖,在提高模拟精度的同时放宽建模要求。Reichstein等<sup>[51]</sup>认为物理过程建模(理论驱动)与机器学习建模(数据驱动)过去往往被看成是两个领域;前者外推能力强,后者更灵活可发现新规律,两者可以互为补充。鉴于物理机制的重要性以及直接从数据发掘复杂物理机制的巨大难度,数据驱动的机器学习可能不会替代物理模型,但会起到补充和丰富的作用,最终构成认知系统的混合智能。

近年来,越来越多学者将物理过程模型与数据驱动模型结合到一起,以提高模型的可解释性、

改善预测精度。目前已有多种不同的混合策略。例如, Young 等<sup>[52]</sup>结合了过程模型(HEC-HMS)与支持向量机,用于预测降雨-径流过程; Zhang 等<sup>[53]</sup>在数据同化的框架内,结合了集合卡尔曼滤波与高斯过程,提高了土壤含水量的预测精度。有些学者仅对物理模型的参数实施机器学习,保留物理过程模型的数学描述。例如,蒸散模型中的许多参数有明确的物理意义,但缺乏精准的量化方案(如地表粗糙长度); Chaney 等<sup>[54]</sup>利用极端随机树方法对地表粗糙长度模型的参数进行估计,从而在原始物理框架下改进蒸散精度。Hu 等<sup>[55]</sup>利用随机森林、神经网络等机器学习技术,直接建立地表粗糙长度与相关变量之间的关系,从而克服了地表粗糙长度物理模型自身的不确定性。

总而言之,观测数据是若干因素在现实世界的综合表现。在控制实验中(固定某些因素),通过数据分析来发现物理机理是较易的;但在灌区这样的复杂系统中,特别是在多因素共同作用下,如何发掘新的物理机理仍有待研究。笔者认为灌区的混合认知智能是指构建出新的模型架构,一方面自主发掘物理机理,另一方面利用已知观测(灌区的数字实现)准确模拟仍存在固有未知信息的灌区内的水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化,准确评估管理行为的效应和效益,实时调整管理行为,提高完成调控目标的概率。

针对灌区复杂系统来说,混合认知智能需解决如下问题:(1)当灌区错综复杂的因果关系未知或仅仅部分已知时,以不牺牲模拟精度为前提放松对建模的要求,建立非完整先验物理机制的灌区水分、盐分、养分、污染物迁移转化以及作物生长和生态系统演化模型;(2)结合基于机器学习的物理方程挖掘技术和数据同化技术,提出基于“数据-物理”混合策略的灌区建模方法;(3)现有模型大多具有固定的模型结构;鉴于灌区的很多环境要素是未知的、难以观测的,固定的模型结构会导致模型无法进化,无法适应于特定的使用环境<sup>[35]</sup>。需根据数据、环境和目标为模型自动设定适宜复杂度,让模型结构具有数据、环境和目标适应性,提高模型的适用性;(4)为混合认知智能建立相应的基础理论支撑。由于机器学习本身的理论限制,能否为“数据-物理”混合方法建立严谨的理论基础,决定了混合认知智能是一种描述自然界因果关系的新方式还是说仅仅是一种实用性技术。

大数据为科学研究和工程应用提供了新的机遇,但在大数据时代,纯数据驱动的方法能否完全取代具备物理意义的理论,数据的相关性能否取代因果关系,在学术界存在巨大的争议<sup>[25]</sup>。笔者认为因果关系超越相关性,水利学科的基本理论模型是该学科主要因果关系的数学表达,基于大数据的模型应首先服从已有的基本理论、方程和因果关系,并通过发掘和描述新的因果关系来形成高级智能(干预能力和反事实能力)。物理与机器学习的混合是灌区感知和认知智能的合理实现途径。但从更长远的角度来看,物理和数据之间的能力界限可能会逐渐模糊。

需要特别强调的是,当数据集不够大时(即所谓的小数据问题),物理机制的作用会尤其重要。鉴于小数据问题在灌区系统中的普遍性(例如在地下水资源管理中,参数、水量、水质的观测数据通常都有限),如何利用物理机制并基于小数据来构建认知智能也应予以关注。

**3.4 智慧灌区管理决策与认知和感知智能的关系** 现有灌区决策系统能够在相对简单的任务上实现优化决策(例如闸门开度的优化控制、单一水库灌区用水调度、以增产为目标的灌溉决策等),但对复杂问题(例如多水源灌区的用水调度、多目标决策等)进行智能化决策面临诸多困难。由于灌区系统的复杂性和不确定性,面向抗旱、防洪、排涝、排渍、控污等多个目标的智能优化决策极具挑战性。这主要是因为难以准确量化管理决策与灌区状态变量(即水分、盐分、养分、污染物、作物和生态系统的状态)之间的复杂关系,经济、环境与生态效益难以均衡,以及问题规模庞大。优化管理决策的主要优化方法包括线性规划、非线性规划、动态规划等经典方法,遗传算法、粒子群算法等进化算法,以及强化学习方法等。就处理复杂优化问题的能力来看,进化算法不依赖于目标函数的梯度信息,适用于处理灌区这类庞大的非线性问题<sup>[56]</sup>。深度强化学习是传统强化学习与神经网络的结合,近年来受到广泛关注,但在灌区管理决策上的研究仍为空白。

灌区观测系统和认知体系的完善分别从数据和基础模型方面为灌区决策系统提供了更好的支撑。灌区的环境感知系统可为智慧决策提供当前的状态,而认知智能可更好评价决策效果。然而,为对水源、沟渠、建筑物、田间等实体进行调控,实现水量、水质和生态的最优化管理,并逐渐形

成系统自主制定管理决策的能力,可能面临以下挑战:(1)灌区管理优化方案的确立相当复杂,涉及到工业、农业、生活、生态、环境多目标协同,通常由多地区、多部门、多行业参与,优化规则的确立相当复杂。例如,灌区用水决策可能涉及到成千上万个农田管理者、渠道管理者、水库管理者,以及农业和环境部门人员等,不同决策者想达到目标也不尽相同,这样导致多目标 Pareto 解集难以生成,即使生成后也难以决策。如何对决策参与者众多、目标多样的管理行为进行优化仍有待研究;(2)灌区的认知和感知智能建立并不能消除灌区决策过程中固有的随机性和不确定性。即使灌区认知体系为决策过程提供了更为准确的约束条件,决策优化过程仍面临着如何获取全局最优解的挑战,特别是实时动态决策。与此同时,对于庞大的灌区系统,评价决策对环境的影响以及相应管理行为为不同决策者带来的效益可能需要消耗巨大的计算成本;(3)根据管理目标的不同,灌区管理将是多种决策优化方法并存的局面;但核心难题仍在于建立一套决策系统,可以胜任整个灌区尺度上的水量、水质和生态的最优化管理;(4)具备从历史数据中逆向提取优化灌溉、排水等方案的能力,从而提高在复杂系统中的决策优化能力;(5)具备专家知识和经验与计算机决策相结合的能力;(6)基于强化学习算法,决策系统在与真实环境和虚拟模型的不断交互中获取经验,而不需要假设经典马尔科夫决策过程中状态转移概率方程的先验知识,以降低其对历史数据的依赖,以非监督方法改善决策系统的决策能力。

## 4 结论

智慧灌区的发展一方面依赖于新型基础设施(如高通量观测设备、数据中心和各种智能设备等)的建设;另一方面依赖于计算能力的突破和先进算法和模型的研发,以及算法、模型与数据中心、灌区设备的连接;再一方面仍依赖于基础学科研究,特别是以农田水利学、水文学、力学、作物学、遥感学等学科为基础,使人工智能在解读灌区因果关系上具有更高精度。在灌区感知智能方面,以多源大数据技术为手段,精准、实时识别灌区特征,并逐渐开发主动观测功能;需要建立大数据环境下的信息理论,用以支撑多源海量数据的解译和分析。在灌区认知智能方面,摆脱对复杂物理机理描述的过度依赖,开发“数据-物理”混合型建模技术,提高模型在非完整先验物理机制下的模拟和预测能力,让模型具有适宜复杂度和相应的模型结构,同时为混合认知智能建立“数据-物理”混合方法的理论基础。在灌区决策方面,需研发针对众多决策参与者、目标函数多样的智能决策系统,实现整个灌区尺度上的水量、水质和生态的最优化管理。

与目前主流的人工智能技术不同,智慧灌区的底层是一种混合智能,是多源大数据技术的成熟形式,是观测系统与管理设施的协作,是物理机制和数据的融合,也是管理决策与灌区感知和认知的结合。本文的观点虽主要围绕智慧灌区来论述,但对其它学科的智能研究也有一定的参考作用。

致谢:感谢武汉大学杨金忠、王修贵、胡铁松、黄介生、崔远来、曾玉红、刘德地、管光华、常剑波教授,中国农业大学任理教授给与的指正和建议。

## 参 考 文 献:

- [ 1 ] 高占义.我国灌区建设及管理技术发展成就与展望[J].水利学报,2019,50(1):88-96
- [ 2 ] RUSSELL S J, NORVI P. Artificial Intelligence: A Modern Approach[M]. Malaysia: Pearson Education Limited, 2016.
- [ 3 ] LUGER G F. Artificial Intelligence: Structures and Strategies for Complex Problem Solving[M]. England: Addison Wesley, 2005.
- [ 4 ] PEARL J, MACKENZIE D. The Book of Why: the New Science of Cause and Effect[M]. New York: Basic Books, 2018.
- [ 5 ] 戴玮,李益农,章少辉,等.智慧灌区建设发展思考[J].中国水利,2018(7):48-49.
- [ 6 ] 周亚平,陈金水,高军.智慧灌区建设要素及关键技术[J].水利信息化,2019(2):11-18.



- [ 7 ] LI X, SHI L, ZHA Y, et al . Data assimilation of soil water flow by considering multiple uncertainty sources and spatial-temporal features: a field-scale real case study[J] . *Stochastic Environmental Research and Risk Assessment*, 2018, 32: 2477–2493 .
- [ 8 ] FUJITA I, WATANBE H, TSUBAKI R . Development of a non-intrusive and efficient flow monitoring technique: The space-time image velocimetry (STIV)[J] . *International Journal of River Basin Management*, 2007, 5(2): 105–114 .
- [ 9 ] 赵浩源, 陈华, 刘维高, 等 . 基于河流表面时空图像识别的测流方法[J] . *水资源研究*, 2020, 9(1): 1–11 .
- [ 10 ] 杨琦, 叶豪, 黄凯, 等 . 利用无人机影像构建作物表面模型估测甘蔗 LAI[J] . *农业工程学报*, 2017, 33(8): 104–111 .
- [ 11 ] GHOLIZADEH M H, MELESSE A M, REDDI L . A comprehensive review on water quality parameters estimation using remote sensing techniques[J] . *Sensors*, 2016, 16(8) . doi: org/10.3390/s16081298.
- [ 12 ] BARZAGHI R, CAZZANIGA N E, DE GAETANI C I, et al . Estimating and comparing dam deformation using classical and GNSS techniques[J] . *Sensors*, 2018, 18(3) . doi: org/10.3390/s18030756.
- [ 13 ] HODGSON M E, BRESNAHAN P . Accuracy of airborne lidar-derived elevation[J] . *Photogrammetric Engineering and Remote Sensing*, 2004, 70(3): 331–339 .
- [ 14 ] ZHANG L, GRIFT T E . A LIDAR-based crop height measurement system for *Miscanthus giganteus*[J] . *Computers and Electronics in Agriculture*, 2012, 85: 70–76 .
- [ 15 ] EITEL J U, MAGNEY T S, VIERLING L A, et al . LiDAR based biomass and crop nitrogen estimates for rapid, non-destructive assessment of wheat nitrogen status[J] . *Field Crops Research*, 2014, 159: 21–32 .
- [ 16 ] REHMAN T U, MAHMUD M S, CHANG Y K, et al . Current and future applications of statistical machine learning algorithms for agricultural machine vision systems[J] . *Computers and Electronics in Agriculture*, 2015, 156: 585–605 .
- [ 17 ] SINGH A, GANAPATHYSUBRAMANIAN B, SINGH A K, et al . Machine learning for high-throughput stress phenotyping in plants[J] . *Trends in Plant Science*, 2016, 21(2): 110–124 .
- [ 18 ] GE Y, BAI G, STOERGER V, et al . Temporal dynamics of maize plant growth, water use, and leaf water content using automated high throughput RGB and hyperspectral imaging[J] . *Computers and Electronics in Agriculture*, 2016, 127: 625–632 .
- [ 19 ] MADEC S, BARET F, DE SOLAN B, et al . High-throughput phenotyping of plant height: comparing unmanned aerial vehicles and ground LiDAR estimates[J] . *Frontiers in Plant Science*, 2017, 8 . doi: 10.3389/fpls.2017.02002.
- [ 20 ] SILVER D, HUANG A, MADDISON C J, et al . Mastering the game of Go with deep neural networks and tree search[J] . *Nature*, 2016, 529(7587): 484–489 .
- [ 21 ] DELIPETREV B, JONOSKI A, SOLOMATINE D P . Development of a web application for water resources based on open source software[J] . *Computers and Geosciences*, 2014, 62: 35–42 .
- [ 22 ] ZHU P, SHI L, ZHU Y, et al . Data assimilation of soil water flow via ensemble Kalman filter: Infusing soil moisture data at different scales[J] . *Journal of Hydrology*, 2017, 555: 912–925 .
- [ 23 ] LI C, SINGH V P, MISHRA A K . Entropy theory-based criterion for hydrometric network evaluation and design: Maximum information minimum redundancy[J] . *Water Resources Research*, 2012, 48(5) . doi: org/10.1029/2011WR011251. .
- [ 24 ] WANG Y, SHI L . A robust data worth analysis framework for soil moisture flow by hybridizing sequential data assimilation and machine learning[J] . *Vadose Zone Journal*, 2020, 19(1) . doi: org/10.1002/vzj2.20026.
- [ 25 ] CALUDE C S, LONGO G . The deluge of spurious correlations in big data[J] . *Foundations of Science*, 2017, 22: 595–612 .
- [ 26 ] KUSSUL N, LAVRENIUK M, SKAKUN S, et al . Deep learning classification of land cover and crop types using remote sensing data[J] . *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(5): 778–782 .
- [ 27 ] ZHONG L, HU L, ZHOU H . Deep learning based multi-temporal crop classification[J] . *Remote Sensing of Environment*, 2019, 221: 430–443 .
- [ 28 ] TANG G, LONG D, BEHRANGI A, et al . Exploring deep neural networks to retrieve rain and snow in high lati-

- tudes using multisensor and reanalysis data[J]. *Water Resources Research*, 2018, 54: 8253–8278.
- [ 29 ] KABA K, SARIG Ü L M, AVCI M, et al . Estimation of daily global solar radiation using deep learning model[J]. *Energy*, 2018, 162: 126–135 .
- [ 30 ] YANG Q, SHI L, HAN J, et al . A near real-time deep learning approach for detecting rice phenology based on UAV images[J]. *Agricultural and Forest Meteorology*, 2020, 287 . doi: org/10.1016/j.agrformet.2020.107938.
- [ 31 ] YANG Q, SHI L, HAN J, et al . Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images[J]. *Field Crops Research*, 2019, 235: 142–153 .
- [ 32 ] PAN J, YIN Y, XIONG J, et al . Deep learning-based unmanned surveillance systems for observing water levels [J]. *IEEE Access .*, 2018, 6: 73561–73571 .
- [ 33 ] ZHOU Z, SHI L, ZHA Y . Seeing macro-dispersivity from hydraulic conductivity field with convolutional neural network[J]. *Advances in Water Resources*, 2020, 138 . doi: org/10.1016/j.advwatres.2020.103545.
- [ 34 ] GUPTA H V, WAGENER T, LIU Y . Reconciling theory with observations: elements of a diagnostic approach to model evaluation[J]. *Hydrological Processes: An International Journal*, 2008, 22(18): 3802–3813 .
- [ 35 ] SAVENIJE H H G . HESS Opinions" The art of hydrology"[J]. *Hydrology & Earth System Sciences Discussions*, 2008, 5(6): 3157–3168 .
- [ 36 ] HSU K L, GUPTA H V, SOROOSHIAN S . Artificial neural network modeling of the rainfall-runoff process[J]. *Water Resources Research*, 1995, 31(10): 2517–2530 .
- [ 37 ] JAIN A, SUDHEER K P, SRINIVASULU S, et al . Identification of physical processes inherent in artificial neural network rainfall runoff models[J]. *Hydrological Processes*, 2004, 18(3): 571–581 .
- [ 38 ] HOSSEINI S M, MAHJOURI N . Integrating support vector regression and a geomorphologic artificial neural network for daily rainfall-runoff modeling[J]. *Applied Soft Computing*, 2016, 38: 329–345 .
- [ 39 ] KIM J, HAN H, JOHNSON L E, et al . Hybrid machine learning framework for hydrological assessment[J]. *Journal of Hydrology*, 2019, 577 . doi: org/10.1016/j.jhydrol.2019.123913.
- [ 40 ] KUMAR M, RAGHUWANSHI N S, SINGH R, et al . Estimating evapotranspiration using Artificial Neural Network[J]. *Journal of Irrigation and Drainage Engineering*, 2002, 128(4): 33–224 .
- [ 41 ] CHINH L V, HIRAMATSU K, HARADA M, et al . Estimation of water levels in a main drainage canal in a flat low-lying agricultural area using artificial neural network models[J]. *Agricultural Water Management*, 2009, 96 (9): 1332–1338 .
- [ 42 ] FENG M, LIU P, GUO S, et al . Deriving adaptive operating rules of hydropower reservoirs using time-varying parameters generated by the EnKF[J]. *Water Resources Research*, 2017, 53(8): 6885–6907 .
- [ 43 ] BAUSER G, FRANSSSEN H J H, KAISER H P, et al . Real-time management of an urban groundwater well field threatened by pollution[J]. *Environmental Science and Technology*, 2010, 44(17): 6802–6807 .
- [ 44 ] HU S, SHI L, HUANG K, et al . Improvement of sugarcane crop simulation by SWAP-WOFOST model via data assimilation[J]. *Field Crops . Research*, 2019, 232: 49–61 .
- [ 45 ] TONGAL H, BOOIJ M J . Simulation and forecasting of streamflows using machine learning models coupled with base flow separation[J]. *Journal of Hydrology*, 2018, 564: 266–282 .
- [ 46 ] FENICIA F, SAVENIJE H H, MATGEN P, et al . Understanding catchment behavior through stepwise model concept improvement[J]. *Water Resources Research*, 2008, 44(1) . doi: org/10.1029/2006WR005563.
- [ 47 ] FENICIA F, MCDONNELL J J, SAVENIJE H H . Learning from model improvement: On the contribution of complementary data to process understanding[J]. *Water Resources Research*, 2008, 44(6) . doi: org/10.1029/2007WR006386.
- [ 48 ] CLARK M P, KAVETSKI D, FENICIA F . Pursuing the method of multiple working hypotheses for hydrological modeling[J]. *Water Resources Research*, 2011, 47(9) . doi: org/10.1029/2010WR009827.
- [ 49 ] BUI D T, PRADHAN B, NAMPAK H, et al . Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibility modeling in a high-frequency tropical cyclone area using GIS[J]. *Journal of Hydrology*, 2016, 540: 317–330 .
- [ 50 ] ZHANG S, LIN G . Robust data-driven discovery of governing physical laws using a new subsampling-based sparse Bayesian method to tackle four challenges (large noise, outliers, data integration, and extrapolation) in

- the discovery of physical laws from data[Z]. arXiv preprint, 2019,
- [ 51 ] REICHSTEIN M, CAMPS-VALLS G, STEVENS B, et al . Deep learning and process understanding for data-driven Earth system science[J] . Nature, 2019, 566(7743): 195–204 .
  - [ 52 ] YOUNG C C, LIU W C, WU M C . A physically based and machine learning hybrid approach for accurate rainfall-runoff modeling during extreme typhoon events[J] . Applied Soft Computing, 2017, 53: 205–216 .
  - [ 53 ] ZHANG Q, SHI L, HOLZMAN M, et al . A dynamic data-driven method for dealing with model structural error in soil moisture data assimilation [J] . Advances in Water Resources, 2019, 132 . doi: org/10.1016/j.advwatres.2019.103407.
  - [ 54 ] CHANEY N W, HERMAN J D, EK M B, et al . Deriving global parameter estimates for the Noah land surface model using FLUXNET and machine learning[J] . Journal of Geophysical Research: Atmospheres, 2016, 121 (22): 13–218 .
  - [ 55 ] HU X, SHI L, LIN L, et al . Improving surface roughness lengths estimation using machine learning algorithms [J] . Agricultural and Forest Meteorology, 2020, 287 . doi: org/10.1016/j.agrformet.2020.107956.
  - [ 56 ] 郭生练, 陈炯宏, 刘攀, 等 . 水库群联合优化调度研究进展与展望[J] . 水科学进展, 2010, 21(4): 496–503 .

## A preliminary exploration of framework, theory and method for intelligent irrigation district

SHI Liangsheng, ZHA Yuanyuan, HU Xiaolong, YANG Qi

( State Key Laboratory of Water Resources and Hydropower Engineering Sciences, Wuhan University, Wuhan 430072, China)

**Abstract:** Irrigation district in China is confronted with the requirement of diverse, precise and intelligent management. Integrating the development trend of artificial intelligence and water resources engineering, the definition and basic functions of intelligent irrigation district (IID) is tentatively proposed in this paper, identifying the major challenges of IID study from the aspects of perception intelligence, cognition intelligence and management decision making. The framework, theories and approaches for IID are preliminarily explored. With respect to perception intelligence, it should be a combination of non-contact and contact measurement, and a mixture of mobile and fixed measurement. Machine vision, data interpretation and “space-sky-ground” data exchange technology need development for multi-scale observation of irrigation district. Information theory under big-data environment is required to support the multi-source observation system and support the interpretation and analysis of massive data. With respect to cognition intelligence, a modelling framework is proposed by integrating physical equation and machine learning. The theoretical basis of hybridizing data and physical mechanisms should be built. Competent models for water, salt, nutrient, pollutant transport and transformation, crop growth, and ecosystem evolution should be developed even with incomplete priori physical mechanisms. Machine learning physical mechanism from data and adaptive modelling technique should be explored. The intelligent decision-making system is expected to satisfy the requirements from a large number of decision makers and handle multiple objective functions, so that optimal strategies can be made to manage water quantity, water quality and ecosystem.

**Keywords:** intelligent irrigation district; artificial intelligence; big data; perception intelligence; cognition intelligence

(责任编辑: 王冰伟)