

气象领域人工智能的现在与未来

——张平文院士及团队骨干访谈录

采访人：韩佳芮（本刊特约，中国气象局气象干部培训学院）

受访专家：张平文（北京大学）

李昊辰（北京邮电大学）

夏江江（中科院大气物理研究所）

采访方式：座谈

采访时间：2021年3月13日

DOI: 10.3969/j.issn.2095-1973.2021.03.0002

韩佳芮：您认为，目前人工智能应用在气象预报业务的哪个领域最有优势和潜力？人工智能可以解决这些领域中的什么关键问题？

张平文：随着多年的发展，气象预报已经有了巨大进展。综合气象观测体系更加完善，积累了海量、长时间序列、质量较好的多源气象观测数据。全球数值天气预报能力也不断提高，目前，数值预报模式基本可以解决大部分区域的大尺度预报问题，如欧洲中心数值预报模式已经可以实现对全球13 km、逐小时较为准确的预报。

气象业务中所获取、产生的海量气象数据是当之无愧的“大数据”，然而，这些海量数据之间还存在着难以融合应用、价值挖掘不够的问题。在实际预报中，实况数据是不是只能用来反映当时的局地天气状况，而模式数据仅能反映一定大尺度上的预报情况？这两大类数据如何融合应用？人工智能的优势之一是解决数据融合问题、挖掘数据价值，不同类型气象数据融合应用中存在的问题恰好为人工智能在气象预报领域发挥作用留下了空间。通过人工智能算法将小微尺度、时空高频的实况观测信息与具有一定准确率、中大尺度、低频的模式预报信息相结合，进而实现对局地的高时空精度的客观预报产品，达到在局地提高预报准确率的目的。

另外，在预报业务中，针对小尺度、精细化的



【编者语】

北京大学副校长张平文院士，带领团队与气象部门合作，将大数据等人工智能(AI)技术引入天气预报技术，以全新视角开启新的预报方法的尝试。2021年3月13日，本刊特约编辑韩佳芮博士对张院士及其团队进行了采访。

预报需求，预报员需要综合对天气形势、实况信息及多种数值预报产品的融合分析，结合自身的预报经验和局地特征，通过会商给出预报结论。会商过程中花费了预报员大量的时间和精力整合分析多源信息。这个过程能不能简化？这实际也是人工智能在预报领域应用中的契机。由于人工智能算法在融合、处理信息中的先天优势，它在一定程度上可以代替预报员在会商中整合、分析信息的过程，通过数据挖掘、学习，将预报员的经验内化在算法中，实现智能、高效的预报。这也是人工智能在气象预报领域未来发展的必然方向，我将它称之为“智能天气预报系统”。

韩佳芮：人工智能应用在气象领域中所面临的最大挑战是什么？传统气象很重视对物理过程和机理的解释，而人工智能包括机器学习的算法虽然可以提高预报准确率，但在一定程度上难以解释，您怎么看待这个问题？

张平文：人工智能应用在气象领域中所面临的最大挑战之一是对预报的机理解释。在对科学认识上，有“强机理”和“弱机理”之分。如牛顿第二定律适用性强，可以认为是一种“强机理”；实际上，科学规律还存在着适用性较弱、非线性关系的“弱机理”问题，如社会科学属于典型的“弱机理”。

对预报来说，构建数值预报模式底层理论的天文学原理、理论假设的物理过程和机制比较清晰，可认为是“强机理”。但模式物理参数化过程的改进对预报准确率的提高并不是一一对应的关系，在这一程度上属于“弱机理”的范畴，如调参之后的模式结果或许在A地的准确率提高了，但在B地可能是下降

的；夏季的预报准确率提高了，但可能春季的预报准确率下降了等。由于大气运动本身是非线性、湍流运动，具有显著的不确定性，对于实际的预报来说，预报的影响因素比较多，并不是直接的线性关系，因此可以认为是一种“弱机理”。

实际上，人工智能应用在对相关信息的优化整合中也是属于“弱机理”，它的特点是可解释性差、局地性强、泛化效果较差，但准确率高。如果与传统预报相比，对相对符合“强机理”的大尺度天气形势来说，人工智能的优势可能并不明显。但对受多种条件影响、具有显著局地特征的地方来说，人工智能则可能比数值预报模式体现出较强的优势，但从机理上很可能无法解释应用人工智能提高预报准确率内在的物理机理、机制。

应用模式和人工智能的“初心”是提高局地天气的预报准确率，因此，需要不忘初心，以提高局地高精度预报准确率为目标，让人工智能可以发挥出应有的作用。天气预报的本身就是一种湍流模式，希望有新的非线性关系的研究发现，不能把机理解释的作用放大，更不能神话，最重要的目标是提高预报准确率。在这个过程中，也需要行业内专家认识的统一，接受预报理念、方式和方法的变革。

除了机理解释之外，人工智能应用在气象领域还面临着其他挑战：

一是开放的心态。气象业务中应用人工智能的目的是帮助整个气象行业提高效率、使业务流程更加科学化。人工智能可以解决观测数据和模式产品融合、应用等问题，同时，在应用中还可以进一步反馈精细化预报中应获取和使用什么数据等信息，解决数据获取和应用的盲目性，提高多种数据的使用效率和价值。因此，业内科学家和业务人员需要有更开放的心态来迎接人工智能技术对气象业务的推动。

二是多部门合作共享。气象预报的进步需要行业内、外多部门各取所长、共同努力。如，气象部门有海量的数据、成功的预报经验，而大学及科研机构具有强大的科研团队以支撑底层算法的研发。因此，各部门可以以开放的心态，通过共享、合作打破行业壁垒，共同推进气象预报的进步。

实际上，正是为了推动人工智能在气象领域的应用，从2017年开始，北京大学数学科学学院、中国气象局北京城市气象研究院、中国科学院大气物理研究所、北京大数据研究院等四家科研业务单位就开始提前谋划，于2018年6月签署合作协议共建气象大数据实验室，充分发挥各科研单位优势。

韩佳芮：目前，在人工智能的具体技术上有什么新进展？未来，您团队的研究工作将着重针对哪些问题？您计划围绕人工智能在气象领域应用的哪些方面继续深耕？

张平文：针对局地的精细化预报是人工智能在气象领域应用的很好的个例。目前，任何数值预报模式都解决不好小尺度、时空精细的预报问题。根据赛场对气象服务的要求，气象部门已经在赛场布设了大量的自动气象观测站，有大量新的观测数据加入。但是，仍有很多关键问题处在研究阶段，如这些观测数据是否会对预报有直接的改善？如何利用这些高频的观测数据，并结合模式结果进行精细化预报，提高预报准确率？类似于这样的高精细化预报就是应用人工智能的方法将海量的、高频的、小尺度观测数据与大量的、低频的、大中尺度模式客观预报信息相融合，提高精细化预报的准确率。在技术方面我们团队做了大量的工作，研发了MOML算法。

李昊辰：简单来说，MOML算法是基于多元线性回归（LR）、随机森林（RF）、支持向量回归（SVR）、梯度提升树（GBDT）和极端梯度提升（XGBoost）等多种机器学习算法对模式数据处理，得到站点预报结果，并选取最优的机器学习模型进行最终预报的算法模型。

夏江江：实现MOML算法的核心任务之一是对数据的处理和特征的选取。进入模型的数据质量直接关系到预报结果的好坏，由于观测数据不可避免地存在一些质量问题，我们首先对各类数据中缺失数据进行插补插值，对异常值进行质量控制等，以保证进入MOML算法的数据质量。同时，我们还将海量的模式数据、观测数据和地理信息数据经过基于气象领域专业知识和机器学习算法的筛选和重建，降低了数据特征的维度，可以在保证准确率的情况下提高算法模型运算速度，节省运算时间，为模型的业务化运行提供支撑。实验结果表明，与常规模型输出统计（MOS）和ECMWF模型结果相比，MOML方法将预测准确性提高了10%以上。

李昊辰：目前，MOML算法已经部署到气象台业务服务器实现实时的业务化天气预报。根据目前测试的结果统计，MOML算法对温度、相对湿度、平均风速、风向、阵风等要素能在常规模型输出统计（MOS）和ECMWF模型结果的基础上平均提高10%以上（图1），后续还将通过大量数据的加入和算法的优化，进一步改进预报结果。

张平文：实际上，人工智能的方法解决了观测数

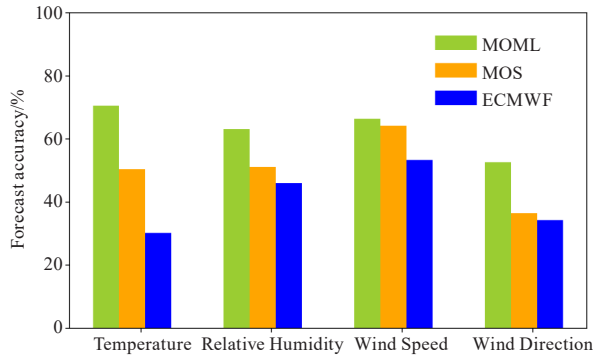


图1 不同模式对温度、湿度和风向预报准确性比较

据与模式预报数据“割裂”的问题，弥补了观测数据与模式预报数据之间不同空间尺度，不同频次信息的转换问题，使多源信息可以真正的融合应用，最终提高局地预报的精确度。有些学者认为说，这实际上是“模式后处理”问题，个人认为并不尽然。在这个项目中主要使用的是自动气象站的观测资料结合EC模式的预报，目前看已经取得了不错的效果，可以理解为是对EC模式结果的订正。但未来，除了应用大量自动气象站的观测数据之外，我们还会进一步应用分析场的格点数据，雷达、卫星等资料以及包含EC在内的，GRAPES、Rmaps-ST等不同尺度的多种数值预报模式产品，甚至包含局地的特征数据等，这不仅仅是对客观预报产品的订正，实际上实现了预报员结合当地预报经验进行预报的过程，相当于通过对多源数据融合应用把预报员经验内化在算法中，实现了“客观预报”超过“人工预报”。从这个角度出发，“模式后

处理”的说法显然已经不合适。

我团队主要是想通过人工智的方法开展“智能天气预报”，解放预报员看大量实况、预报产品进行信息分析、处理的预报过程，替代会商中较为繁琐的会商过程，同时通过应用人工智能提高全国的预报准确率。

韩佳芮：请您畅想下人工智能应用在气象领域的最终场景？

张平文：未来，预报员不需要花大量的时间、精力分析实况数据、环流形势等，也不需要比对多个数值预报模式产品，可以通过基于人工智能算法的“智能预报系统”直接拿到已经综合多种信息的预报产品，预报员只需要做最后的把关或是制作一些特殊的服务产品等工作，这样可以大幅提高预报员的工作效率。同时，“智能预报系统”还大大简化了会商的流程，最终甚至将取代会商，可以不再依赖于预报员们主观经验的好坏，直接给出有较高准确率的、客观精细化的预报结论。这样不仅提高了全国气象预报的准确率，还解决东、中、西部预报水平不一致的问题，针对无人区也可以实现精细化的预报，解决了预报中的难点，也解放了预报员的劳动力。人工智能在气象领域的应用最终不仅体现在提供更精细的预报，提升全国的预报业务的效率，还将带来气象业务或者说是整个气象行业的变革。

韩佳芮：感谢您和团队接受采访，愿您的团队取得更好的成果！