

DeepSeek在水文勘测资料处理中的应用与展望

李振鹏, 郝金义

长江水利委员会水文局汉江水文水资源勘测局, 湖北 襄阳

收稿日期: 2026年3月11日; 录用日期: 2026年4月29日; 发布日期: 2026年6月29日

摘要

DeepSeek作为先进的人工智能技术平台, 给水利领域带来了深刻的变革。本文以DeepSeek为基础, 结合VBA技术, 通过用结构化提问、合理拆分任务等技巧, 识别用户需求, 实现代码生成、解释、纠错、调优, 生成水文勘测数据处理软件, 并在生产中得到应用。研究结果表明: DeepSeek生成的软件语言简洁、计算结果准确, 人工干预少, 数据处理效率高。

关键词

DeepSeek, 人工智能, 银行家舍入法

Application and Prospect of DeepSeek in Hydrological Survey Data Processing

Zhenpeng Li, Jinyi Hao

Hanjiang Hydrology and Water Resources Survey Bureau, Bureau of Hydrology, Changjiang River Water Resources Commission, Xiangyang Hubei

Received: March 11, 2026; accepted: April 29, 2026; published: June 29, 2026

Abstract

DeepSeek, as an advanced artificial intelligence technology platform, has brought profound changes to the water conservancy field. Based on DeepSeek and combined with VBA technology, this paper identifies user needs through techniques such as structured questioning and reasonable task decomposition, enabling code generation, interpretation, error correction, and tuning. The resulting hydrological survey data processing software has been applied in production. The results show that the software generated by DeepSeek is concise in language and accurate in calculation, and requires minimal human intervention with high data processing efficiency.

作者简介: 李振鹏(1983-), 男, 湖北丹江口人, 高级工程师, 主要从事河道勘测、地理信息工程方面的研究, Email: 284172856@qq.com

文章引用: 李振鹏, 郝金义. DeepSeek 在水文勘测资料处理中的应用与展望[J]. 水资源研究, 2026, 15(3): 333-340.
DOI: 10.12677/jwrr.2026.153036

Keywords

DeepSeek, Artificial Intelligence, Banker's Rounding

Copyright © 2026 by author(s) and Wuhan University & Bureau of Hydrology, Changjiang Water Resources Commission. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

DeepSeek 问世之前,广大水文工作者只能在 Excel 中通过人机交互操作完成数据表的格式转换,或通过人工编写代码实现简单的功能。单纯依靠 Office 软件处理表格会造成工作效率低下,人工编写代码对水文工作者的编程水平要求高,代码可移植程度差。本文提出一种基于 DeepSeek 的程序代码自动编写方法。由 DeepSeek 生成的程序代码,在 Excel VBA 平台的加持下完成水文勘测资料处理。该研究破解了单纯依靠 Office 软件处理水文勘测数据过程繁琐,效率低下的难题。

2. DeepSeek 与水文勘测的适配性分析

2.1. DeepSeek 简介

DeepSeek 是由杭州深度求索 AI 基础技术研究有限公司开发的系列化 AI 模型,凭借“好用、开源、低成本”三大特点,打破“规模定律”与“生态壁垒”,在大数据、机器学习等新技术的加持下,为各行业的数字化转型提供核心驱动力[1]-[3]。DeepSeek 还优化了中文场景,填补了国际模型的中文短板,成为中文大模型领域的标杆[4] [5]。

2.2. Excel VBA 简介

Microsoft Excel 是一款受多行业用户欢迎且功能强大的办公软件。在数据记录和处理、数据分析、财务管理、项目管理等专业领域发挥着重要作用。VBA (Visual Basic for Applications)是 Excel 的内置编程语言,用户可以创建自定义的函数、宏和用户界面,扩展了 Excel 的功能。由于 Excel VBA 操作简单,容易上手,深受初学者青睐。

2.3. 水文勘测数据处理现状

水文勘测工作不仅需要面对大范围、多尺度、长序列的降水、水位、流量数据以及海量的地形观测和大断面信息,还要对多要素信息整合处理、科学配置,迫使作业人员采用编写能实现数据处理自动化处理的脚本来提高数据处理效率。但多数水文勘测人员编程能力有限,数据处理效率低下,达不到预期目标。

DeepSeek 支持自然语言,具备自动提醒功能(例如在 DeepSeek 服务器中输入“将表 1 的 A 列填充到表 2 的 A 列,将表 1 的 B 列填充到表 2 的 A 列”,这时候服务器会提示要不要覆盖表 1 的 A 列),自主学习功能(例如 DeepSeek 服务器被告知用户的 Excel 不支 Filters.Clear 命令以后,服务器自动生成的代码就不会再出现该项命令),水文工作者可以根据自身需求,在 DeepSeek 服务器上完成结构化提问,然后生成 VBA 代码,加载到 Excel 平台中迭代调试,生成脚本,完成数据处理工作,更加专注于需求分析和成果质量把控,避免在编程上花费大量时间。

3. 开发思路及方案

汉江水文水资源勘测局承担辖区内 14 个工程专用水位站资料收集、数据整编及入库工作。由于数据格式不

匹配，部分测站 8:00 数据缺失等原因，水位站传感器上收集的水位数据不能直接录入水文资料整编库。在项目时间紧急、数据时空跨度大、序列较长的情况下，需要研发一款能满足现有技术要求、流程标准、界面友好、操作便捷、运行高效的数据处理软件来完成数据处理，进而提高资料整编的进度。

本次水文勘测资料处理软件的开发主要针对辖区内专用水位站资料收集和数据整编展开。专用水位站资料收集工作主要包括源数据表格的下载和人工校核数据的录入两项工作。源数据表格从多要素平台下载，源数据表格包括测站编码、测站名称、测站类型、时间、水位、水位 1、水位 2、在线流量、推流流量、人工流量共 10 个字段的内容，如表 1 所示：

表 1. 源数据表格样式

站点编码	站点名称	站点类型	时间	水位	水位 1	水位 2	在线流量	推流流量	人工流量
7557	LH	水文站	2025-10-01 01:20:00	47.49	-	-	-	-	-
7557	LH	水文站	2025-10-01 01:25:00	47.50	-	-	-	-	-
7557	LH	水文站	2025-10-01 01:30:00	47.50	-	-	-	-	-
7557	LH	水文站	2025-10-01 01:35:00	47.50	-	-	-	-	-

资料整编的主要内容包括：

- 1) 水位数据整理。
- 2) 缺失数据的补充与完善。
- 3) 精度验证。
- 4) 数据入库。

水位数据治理主要是对水位报表的处理，首先删除标题行，然后删除表格前 3 列和后 5 列数据，仅保留时间列和水位列的数据，表格另存为 csv 格式，原表格保持不变。处理后的表格格式如表 2 所示。

表 2. 整编数据表格

时间	水位
2025-10-01 1:20	47.49
2025-10-01 1:25	47.50
2025-10-01 1:30	47.50
2025-10-01 1:35	47.50

缺失数据的补充与完善主要是对 A 站和 B 站，两个站由于设备原因没有 8:00 的水位，考虑到这两个站 7:55 和 8:05 分水位记录都是完整的，而且水位变化较为平缓，因此确定用时间内插法求得两站 8:00 的水位完成缺失数据的补充与完善，水位记录报表如表 3 所示。

精度验证主要是完成缺失水位补充表格的精度评定。

数据入库则是将经过数据格式转换，缺失数据补充，粗差别除和精度验证后的正确的水位报表录入到水文资料整编库中的过程。

针对这 4 部分主要内容和环节，采用目前最流行的 DeepSeek，结合 Microsoft Excel 的二次开发工具 Excel VBA，实现批量文件数据格式转换，8 点水位自动完善，水位报表信息自动添加等功能，大大提高了工作效率。

本次水文勘测资料数据处理的流程包括：数据准备，需求分析，任务分解，提示词输入，代码生成，软件调试，成果验证，软件封装 8 个环节，软件开发的技术方案如图 1 所示。

表 3. 缺失 8 点记录的水位站数据表格

日期	时间	水位
2026-01-01	07:55:00	50.12
2026-01-01	08:05:00	50.12
2026-01-02	07:55:00	49.95
2026-01-02	08:05:00	49.94
2026-01-03	07:55:00	48.50
2026-01-03	08:05:00	48.48
2026-01-04	07:55:00	47.16
2026-01-04	08:05:00	47.16

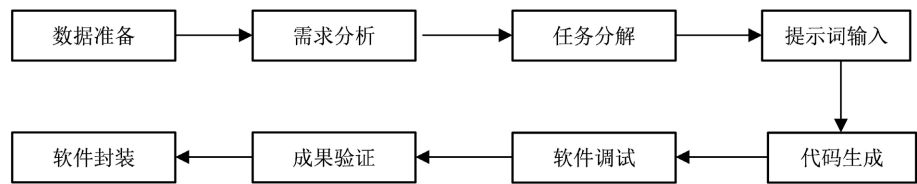


图 1. 水文勘测数据处理流程图

- 1) 数据准备阶段主要是完成 14 个水位站水位资料的收集。具体内容包括完成每个月 14 个水位站水位信息的分类存放，然后新建文件夹为数据处理阶段提供数据存储环境。
- 2) 需求分析阶段主要是分析软件需要完成的功能，具体内容包括：
 - a) 水位站数据治理
 - b) 8 点水位的完善
 - c) 长时段缺失数据的补充
 - d) 水位报表信息完善
- 3) 任务分解则是对每个功能做详细的拆分以便更好完成任务，详见 8:00 水位的自动完善案例。
- 4) 提示词输入则是在 DeepSeek 服务器中输入自然语言完成代码的编写，详见水位站数据治理案例。
- 5) 代码生成则是在 DeepSeek 服务器上生成 VBA 代码，用于 Excel VBA 测试。
- 6) 软件调试则是在 Excel VBA 平台上测试 VBA 代码。
- 7) 成果验证阶段主要是对软件生成的计算成果与第三方软件生成的成果对比分析，完成质量评定。
- 8) 软件封装阶段内容是将软件的源代码通过编译器编译成机器语言包，生成独立的程序。

4. 案例分析

4.1. 水位站数据治理

辖区内有 14 个工程专用水位站，每站 12 个月的水位观测数据单独生成 1 个水位文件，原始文件包括：测站编码、测站名称、站点类型、时间、水位、水位 1、水位 2、在线流量、推流流量、人工流量共 10 列，其中第一行为标题行。目前需要完成每个站 12 个月的水位数据单独入库，入库数据表格格式要求如下：

- 1) 数据无表头；
- 2) 第一列为日期时间，第二列为水位值，保留至小数点后两位。

如果在 Excel 中采用人机交互方式对单个文件操作，要操作 168 个表格文件，耗时费力，而且出错概率高，

故采用编程完成数据文件的批量转换。具体流程如下：利用 DeepSeek 工具，输入“首先用 VBA 选择一个文件夹，打开里面的所有表格，然后对每一个表进行以下操作，先删除第一行，然后删除表格的左边 3 列和右边 5 列，接着所有表格在新文件夹另存为 csv 格式，原文件保持不变，文件夹和新文件夹的选择在对话框中完成”提示词，生成代码，将 DeepSeek 给出的代码复制到 Excel 软件环境中运行后，代码运行成功，读取了 10 个站 12 个月的数据，按照测站方式分月存储。结果表明，DeepSeek 生成的代码能实现水位站数据治理功能，部分代码如图 2 所示：

```
End With
' 检查文件夹路径是否以反斜杠结尾
If Right(sourceFolder, 1) <> "\" Then sourceFolder = sourceFolder & "\"
If Right(targetFolder, 1) <> "\" Then targetFolder = targetFolder & "\"
' 获取第一个Excel文件
filePath = Dir(sourceFolder & "*.xls*")
' 循环处理所有Excel文件
Do While filePath <> ""
    ' 构建完整文件路径
    fileName = sourceFolder & filePath
    ' 打开工作簿
    Set wb = Workbooks.Open(fileName)
    ' 处理每个工作表
    For Each ws In wb.Worksheets
        With ws
            ' 删除第一行
            .Rows(1).Delete
            ' 删除A、B、C列
            .Range("A:C").Delete Shift:=xlToLeft
            ' 查找最后一列
            lastCol = .Cells(1, .Columns.Count).End(xlToLeft).Column
            ' 删除后五列（如果列数足够）
            If lastCol > 5 Then
                .Range(.Cells(1, lastCol - 4), .Cells(.Rows.Count, lastCol)).EntireColumn.Delete
            ElseIf lastCol > 0 Then
                ' 如果列数少于5列，删除所有剩余列
                .Range(.Cells(1, 1), .Cells(.Rows.Count, lastCol)).EntireColumn.Delete
            End If
        End With
        ' 构建新的文件名（CSV格式）
        newFileName = targetFolder & Replace(filePath, ".xls*", ".csv")
        newFileName = Replace(newFileName, ".xls", ".csv")
        ' 另存为CSV格式
        wb.SaveAs fileName:=newFileName, FileFormat:=xlCSV
    Next ws
    ' 关闭工作簿，不保存更改（因为已经另存为CSV）
    wb.Close SaveChanges:=False
    ' 获取下一个文件
    filePath = Dir
Loop
```

图 2. 水位站数据治理代码图

4.2. 8:00 水位的自动完善

辖区内有 14 个工程专用水位站，由于仪器设备原因，部分水位站每月都缺少 8:00 的水位，而 7:55 分和 8:05 分的水位记录都是正常的，如果通过人机交互方式在 Excel 软件中逐条追加记录，然后再根据时间计算出 8:00 的水位，并完善记录，不仅费时费力，而且容易出错。通过 DeepSeek 编写程序，完成水位线性内插则事半功倍。可事先将这个任务分解成 3 个子任务，子任务 1：查询没有早 8 点水位记录的日期记作 D1；子任务 2：在该日期内查询距离 8 点最近的两个时刻 T1、T2 的水位值 A1 和 A2；子任务 3：计算出 D1 那天 8:00 的水位 A3 并在 T1 和 T2 之间追加 1 行记录，第一列为“yyyy/mm/dd hh/mm/ss”格式，第二列则根据 T1 和 T2 及 8:00 的时间差线性内插，具体计算公式如下： $A3 = A1 + (A2 - A1) * (8:00 - T1) / (T2 - T1)$ 。

根据《水文资料整编规范》(SL/T 247-2020)¹要求，水位需记至 0.01 m，取用精度位数后一位数字，采用“四舍六入法取舍”，即取用精度位数后一位数字小于五者则舍，大于五者则入，等于五者若其后有非零尾数仍入，

¹<https://www.doc88.com/p-77387164378453.html>

无非零尾数则视取用的末位数字的奇偶取舍，为奇则入，为偶则舍。银行家舍入法，是由 IEEE 754 标准推荐的浮点数取整算法，其规则包含三种情况：当修约位后数字小于 5 时舍去，大于 5 时进位；若等于 5 且后无有效数字，则按修约位前数字奇偶性决定舍入——奇数进位，偶数舍去；若 5 后存在非零数字则直接进位。例如，13.825 保留两位小数得 13.82 (修约位前为偶数)，13.835 则为 13.84 (修约位前为奇数)。

银行家舍入法和四舍六入法其实是一种算法，相对于四舍六入法，银行家舍入法表述更简洁，更容易被 DeepSeek 识别。因此，只需要对 A3 按银行家舍入法保留至小数点后 2 位即可完成水位修约要求。

任务分解完毕之后，利用 DeepSeek 工具，输入“首先用 VBA 打开一个文件夹，选择里面的 csv 表格，然后从表格第一行开始，查找每天 8:00:00 的记录，如果没查找到 8:00:00 的记录，则线性内插出 8:00:00 的记录。具体算法如下：第一步：定位表格文件中 8:00:00 之前 5 分钟的记录，查询 A 列和 B 列对应字段数值，记作 T1，H1；第二步：定位表格文件中 8:00:00 后 5 分钟的记录，查询 A 列和 B 列对应字段数值，记作 T2，H2，如果没有 5 分钟前的记录，就查找 10 分钟的，以此类推，直至 30 分钟；第三步：在这两行中间插入一行，A 列对应位置的数值中日期与 T1 一致，时间统一为：8:00:00，数据格式为“yy-m-d h:m:m”格式，B 列对应位置的数值 H3 则是按 H1，H2，T1，T2 线性内插，B 列对应位置的数值按银行家舍入法保留至小数点后两位，C 列对应位置则统一填充为“CB”，然后对表格文件依次操作，直至表格最后一行，最后保存表格文件。选择文件夹，打开表格，保存表格文件在对话框中完成提示词。生成代码后，将 DeepSeek 给出的代码复制到 Excel 软件环境中运行后，代码运行成功。每个站缺失的 8:00 水位都得到了补充，而且在表格的 C 列都有内插标记，便于人工交互检查。试验结果表明：DeepSeek 生成的代码能实现 8:00 水位的自动完善。部分代码如图 3 所示。

```

' 从后往前插入新行（避免行号变化问题）
For k = newRows.count To 1 Step -1
    ' 计算需要插入的行号
    Dim insertRow As Long
    insertRow = newRows(k)

    ' 检查插入行是否有效
    If insertRow > 1 And insertRow <= ws.Cells(ws.Rows.count, "A").End(xlUp).Row Then
        ' 获取前后记录的数据（重新获取，因为行号可能已变化）
        T1 = ws.Cells(insertRow - 1, 1).value
        H1 = ws.Cells(insertRow - 1, 2).value
        T2 = ws.Cells(insertRow, 1).value
        H2 = ws.Cells(insertRow, 2).value

        ' 检查数据是否有效
        If IsValidDate(T1) And IsValidDate(T2) And IsNumeric(H1) And IsNumeric(H2) Then
            ' 计算线性插值
            Dim timeDiffTotal As Double, timeDiffPart As Double
            timeDiffTotal = (T2 - T1) * 24 * 60 ' 转换为分钟
            timeDiffPart = (DateValue(T1) + TimeSerial(8, 0, 0) - T1) * 24 * 60

            ' 避免除以零
            If timeDiffTotal <> 0 Then
                H3 = H1 + (H2 - H1) * (timeDiffPart / timeDiffTotal)
            Else
                H3 = H1
            End If

            ' 应用银行家舍入法保留两位小数
            H3 = BankersRound(H3, 2)

            ' 插入新行
            ws.Rows(insertRow).Insert Shift:=xlDown

            ' 设置新行的值
            ws.Cells(insertRow, 1).value = DateValue(T1) + TimeSerial(8, 0, 0)
            ws.Cells(insertRow, 1).NumberFormat = "yy-m-d h:mm:ss"
            ws.Cells(insertRow, 2).value = H3
            ws.Cells(insertRow, 3).value = "CB"
        End If
    End If
Next k

```

图 3. 8:00 水位自动完善

4.3. 成果验证

采用 DeepSeek 生成的水文勘测数据处理软件对辖区内 14 个水位站完成数据治理和早 8 点水位自动完善，并对计算成果和人机交互方式在 Excel 中生成的成果进行了对比。结果如下：

1) 在实现水位站数据治理方面：采用在 Excel 中采用人机交互法完成 14 个水位站 12 个月的数据格式转换需要 1 整天时间，而且容易出错，而采用 DeepSeek 生成的水文勘测数据处理软件完成同样的任务只需 10 min，转换的成果能满足资料整编软件的格式要求。

2) 打开某站 1 月份水位报表发现，表格中缺少 1 月 2 日至 1 月 31 日共 30 条 8:00 水位记录，经过水文勘测数据处理软件处理后，补充了 1 月 3 日至 31 日共 29 条 8:00 水位记录。通过比对 29 条软件自动生成水位记录和人机交互法生成的水位记录，用软件内插的水位和 Excel 中采用人机交互法生成的 8:00 水位较差均为 0，因此可以得出如下结论：DeepSeek 生成的水文勘测数据处理软件能完成缺失的 8 点水位的批量完善，计算成果满足《水文资料整编规范》(SL/T 247-2020)对水位取用位数的要求。两种方法生成的水位报表及较差统计如表 4 所示。

表 4. 软件自动完善水位与人机交互水位对照表

软件自动完善的水位			人机交互完成的水位			较差
时间	水位	备注	时间	水位	备注	
1/1/2026 7:55	47.63		1/1/2026 7:55	47.63		
1/1/2026 8:00	47.63	CB	1/1/2026 8:00	47.63	人工	0.000
1/1/2026 8:05	47.63		1/1/2026 8:05	47.63		
1/2/2026 2:25	47.62		1/2/2026 2:25	47.62		
1/2/2026 14:05	47.61		1/2/2026 14:05	47.61		
1/3/2026 7:55	47.60		1/3/2026 7:55	47.60		
1/3/2026 8:00	47.60	CB	1/3/2026 8:00	47.60	人工	0.000
1/3/2026 8:05	47.60		1/3/2026 8:05	47.60		
1/4/2026 7:55	47.61		1/4/2026 7:55	47.61		
1/4/2026 8:00	47.61	CB	1/4/2026 8:00	47.61	人工	0.000
1/4/2026 8:05	47.61		1/4/2026 8:05	47.61		
1/5/2026 7:55	47.59		1/5/2026 7:55	47.59		
1/5/2026 8:00	47.59	CB	1/5/2026 8:00	47.59	人工	0.000
1/5/2026 8:05	47.59		1/5/2026 8:05	47.59		
1/6/2026 7:55	47.59		1/6/2026 7:55	47.59		
1/6/2026 8:00	47.59	CB	1/6/2026 8:00	47.59	人工	0.000
1/6/2026 8:05	47.59		1/6/2026 8:05	47.59		
1/7/2026 7:55	47.60		1/7/2026 7:55	47.60		
1/7/2026 8:00	47.60	CB	1/7/2026 8:00	47.60	人工	0.000
1/7/2026 8:05	47.60		1/7/2026 8:00	47.60	人工	0.000

5. 结论与展望

5.1. 结论

上述 2 个案例利用了 DeepSeek 自然语言编译能力以及代码自动生成功能实现了水位站数据治理和缺失的

8:00 水位完善, DeepSeek 生成的代码简洁、计算结果准确, 人工干预少, 计算结果满足《水文资料整编规范》(SL/T 247-2020)要求。

5.2. 展望

借用 DeepSeek 的自然语言功能结合 Excel VBA 技术开发的水文勘测数据处理软件, 能完成水位站和雨量站数据处理工作, 这只是 DeepSeek 最为直接简便的使用方式。今后可紧密围绕水文勘测核心业务, 重点研究以下 2 个方面:

- 1) 根据上下游水位站水位变化规律, 采用过程线插补法, 实现测站长时段缺失水位的自动完善, 例如表 4 中 1 月 2 日 8:00 水位的补充。
- 2) 成果校验自动化。结合水位站洪水特性, 设置洪水期和枯水期不同的异常水位配置规则, 自动剔除异常水位值, 提高水位摘录效率。

参考文献

- [1] 令小雄. DeepSeek 开启后 ChatGPT 时代——基于数字范式革新及其运演哲思[J]. 西北工业大学学报(社会科学版), 2025, 45(2): 59-67.
- [2] 闵斌, 林立芳, 吴健, 等. DeepSeek 在航天科研生产中的应用研究[J]. 上海航天(中英文), 2025, 42(2): 1-8, 18.
- [3] 高宇辰, 李蔚林, 陈翔, 等. DeepSeek 在储能研究中的应用前景展望[J]. 储能科学与技术, 2025, 14(2): 467-478.
- [4] 高子轩, 宋昕熠. DeepSeek 在工程水文领域的应用探索与展望[J]. 长江科学院院报, 2025, 42(8): 179-187.
- [5] 靖翔. 大数据背景下的智慧水务系统开发分析[J]. 工程技术研究, 2024, 9(3): 228-230.