

## SHUD 研究组实验室手册

舒乐乐和他的学生们、同事们、朋友们

2025-12-22



献给……

呃，爱谁谁吧

# 目录

前言	<b>9</b>
<b>第一章 研究组介绍</b>	<b>11</b>
1.1 基本组织结构	11
1.2 身份角色	12
1.3 硕博士差异	17
1.4 待遇	21
<b>第二章 规则</b>	<b>23</b>
2.1 研究组规则	23
2.2 导师制度	27
2.3 科研经费	29
2.4 组会制度	29
2.5 给学生的建议	30
2.6 年度考核	32
2.7 安全制度	33
2.8 保密制度	35

<b>第三章 研究生培养计划</b>	<b>39</b>
3.1 研究方向 . . . . .	39
3.2 时间表 . . . . .	39
3.3 课程推荐 . . . . .	40
3.4 必读书籍 . . . . .	42
3.5 必读文献 . . . . .	42
<b>第四章 研究的范式</b>	<b>45</b>
4.1 科学问题 . . . . .	45
4.2 研究区 . . . . .	46
4.3 检验、检验、检验 . . . . .	47
<b>第五章 论文写作</b>	<b>49</b>
5.1 IMRaD . . . . .	49
5.2 题目 (Title) . . . . .	50
5.3 摘要 (Abstract) . . . . .	52
5.4 方法、研究区、数据 (Methodology) . . . . .	65
5.5 结果 (Results) . . . . .	65
5.6 讨论 (Discussion) . . . . .	66
5.7 结论 (Conclusion) . . . . .	66
5.8 其他论文附属内容 . . . . .	66
5.9 论文投稿 . . . . .	67
<b>第六章 技能训练</b>	<b>69</b>
6.1 电子邮件 . . . . .	69
6.2 ABT 句式 . . . . .	70

<b>目录</b>	<b>7</b>
6.3 作图与解释 . . . . .	71
6.4 精炼书写原则 . . . . .	71
6.5 格式/形式自查清单 . . . . .	74
6.6 个人简历 (CV) . . . . .	77
<b>第七章 科研工具</b>	<b>79</b>
7.1 空间制图 . . . . .	81
7.2 绘图 . . . . .	82
7.3 文献 . . . . .	82
7.4 AI 大语言模型 . . . . .	87
7.5 远程登录和数据传输 . . . . .	88
7.6 Linux 操作系统 . . . . .	90
7.7 R . . . . .	91
7.8 Python {#python} . . . . .	93
<b>第八章 科研资源</b>	<b>95</b>
8.1 工作环境 . . . . .	95
8.2 图书馆 . . . . .	95
8.3 文献下载 . . . . .	95
8.4 数据 . . . . .	98
<b>第九章 科研设备</b>	<b>99</b>
9.1 NAS 网络存储 . . . . .	101
9.2 Windows 远程桌面 . . . . .	107
9.3 Ubuntu 服务器 . . . . .	108
9.4 GPU 服务器 . . . . .	114
9.5 SHUDHPC 高性能计算集群 . . . . .	115

<b>附录</b>	<b>120</b>
<b>附录 A 文献阅读</b>	<b>123</b>
A.1 任务概述 . . . . .	123
A.2 公众号“阅读笔记”文章安排 . . . . .	124
A.3 忠告 . . . . .	124
A.4 阅读 PPT 指导意见 . . . . .	125
<b>附录 B 入组新手任务</b>	<b>127</b>
B.1 个人材料准备 . . . . .	127
B.2 阅读本研究组重要论文 . . . . .	127
B.3 完成 R 语言入门练习 . . . . .	128
B.4 任意流域 SHUD 模型模拟 . . . . .	128
<b>附录 C 修改日志</b>	<b>131</b>
<b>附录 D 致谢</b>	<b>133</b>

# 前言



## 为什么要阅读这本书？

无论你是本研究组的学生（本硕博），还是客座研究生，还是科研助理，甚至博士后，你都应该以这本书作为了解本研究组以及你未来同学同事的入口。

写这个本是为了公开重要信息，明晰权责，减少误解，确保沟通，保证各方的合理预期并达成预期。

这本书将告诉你：自己应该做什么，知道别人能为你做什么，减少不切实际的要求。

\* 本书中学生一词，同时包括本科生、硕士研究生、博士研究生和客座硕/博研究生。

祝生活如你所愿 ~~~



# 第一章 研究组介绍

## 1.1 基本组织结构

### 1.1.1 团队研究方向

1. **流域水文模型**: 描述和模拟流域内的水文循环过程、确定模型参数的最优值、改进模型结构和方法以提高模型精度和适用性，以及将模型应用于水文预测和水资源管理领域解决实际问题。
2. **洪水预报**: 洪水形成和发展规律的研究，水文观测和数据分析，洪水预报模型的建立和优化，洪水预报技术的应用和推广。
3. **“大气-陆面-水文”耦合模拟**: 将大气模型、陆面模型和水文模型相互耦合，对大气、陆面和水文过程进行综合模拟和分析，以揭示它们之间的相互作用和影响。
4. **数据挖掘与深度学习**: 气象预报和气候模拟、气象图像处理、气候变化分析、土壤水分预测、径流预测、洪水预测。
5. **寒区水文学**: 研究极寒地区水文过程、水文特征、水资源利用及管理等问题的学科，主要研究冰雪过程、融雪径流、冻土水文过程、寒旱区生态水文等内容。

### 1.1.2 单位介绍

中国科学院西北生态环境资源研究院（简称中国科学院西北研究院，[www.nieer.cas.cn](http://www.nieer.cas.cn)）是我国专门从事高寒干旱地区生态环境、自然资源和重大工程研究的国家级研究机构，其主要研究领域（如冰川、冻土、沙漠、高原生态、盐湖、油气地质和资源环境信息等）均处于国内引领地位。本院目

前拥有 2 个国家重点实验室，6 个中国科学院重点实验室，1 个国家数据中心，20 个甘肃、青海省级重点实验室/工程中心，6 个国家级野外观测研究实验站，19 个中国科学院和研究所级野外观测研究实验站。

## 1.2 身份角色

### 1.2.1 团队负责人

科研团队负责人 (Principle Investigator) 是研究团队中的核心成员之一，其角色和任务非常重要。以下是团队负责人的主要角色和任务 (参考 chatGPT 的回答)：

- **领导者**: 团队负责人应该担任领导者的角色，确保团队的目标和愿景得到清晰的阐述，并且指导团队朝着这些目标和愿景前进。他们应该能够鼓舞团队成员，激发他们的潜力，鼓励他们在项目中表现出色。
- **协调者**: 团队负责人应该能够协调团队成员的工作，确保所有成员都明确自己的角色和职责，并且工作合理分配。他们应该与团队成员沟通，了解每个人的进度和工作状态，以便及时调整团队的计划和目标。
- **管理者**: 团队负责人应该能够有效地管理团队，包括资源分配、预算管理、时间管理等。他们应该能够制定计划和目标，并监督团队成员的执行情况。此外，他们还应该能够解决团队成员的问题和冲突，并与上级或其他团队保持联系。
- **战略家**: 团队负责人应该具有战略眼光，能够分析团队所处的地位和行业趋势，并制定相应的策略。他们应该能够推动团队不断创新和进步，并与其他团队进行合作，以实现更高的目标。
- **培训者**: 团队负责人应该是一个有效的培训者，能够培训团队成员并帮助他们提高技能。他们应该能够识别团队成员的优点和弱点，并为他们提供支持和反馈，以便他们能够不断成长和进步。
- **资助者**: 团队负责人要负责从外部获得足够多的经费，以支持团队成员的工资和科研资源需求，并且为团队未来研究做好技术储备。

总之，团队负责人的角色和任务非常重要，需要具备领导能力、协调能力、管理能力、战略眼光和培训能力。只有做好这些方面的工作，才能够推动团队不断创新、进步并实现目标。

### 1.2.2 科研助理

科研助理是科研团队中的重要成员之一，通常是学生或新入职的科研工作者。其主要职责如下：

- **协助科研工作：**科研助理主要的工作是协助科研团队进行研究工作，包括文献检索、实验设计、数据采集、数据处理和分析等。在研究项目中，科研助理通常需要配合团队成员完成具体的工作任务，并且积极主动地向团队成员寻求帮助和指导。
- **组织和管理研究材料：**科研助理需要负责管理研究项目的各种材料，包括文献资料、实验记录、数据结果、研究进展报告等。他们需要确保这些材料有序、准确、完整，并且可以方便地被团队成员查阅和使用。
- **支持团队的学术活动：**科研助理需要支持团队的学术活动，包括组织和参与学术研讨会、讲座、学术会议等。他们需要与团队成员合作，制定学术活动计划，协调各项细节，并确保活动顺利进行。
- **学习和掌握研究技能：**科研助理需要不断学习和掌握研究技能，包括实验技能、数据处理和分析技能、文献检索和阅读技能等。他们需要通过自学和与团队成员交流互动，不断提高自己的研究能力和水平。

总之，科研助理在科研团队中起着非常重要的作用，他们需要协助团队成员完成各项研究任务，并且积极参与团队的学术活动，不断提高自己的研究能力和水平，为科研项目的顺利进行做出贡献。

### 1.2.3 博士生

博士生的核心职责与使命：

- **推进原创性研究课题：**博士生应聚焦创新性科研攻关，系统开展从理论建构到实践验证的全链条研究。具体包括：系统梳理领域前沿文献、构建理论框架与技术路线、严谨实施实验/实证研究、运用先进方法进行数据解析，最终形成具有学术价值的研究成果并通过学位论文答辩，完成从知识消费者向知识生产者的转变。
- **深度参与学术共同体建设：**博士生需主动融入学术生态网络，通过参与国际会议、学术沙龙、跨机构合作等多元形式，在学术对话中实现知识迭代。应注重学术成果的会议报告、墙报展示及期刊发表，在学术争鸣中完善研究设计，同时培育学术网络资源。
- **构建系统性研究能力体系：**博士生应着力打造包含“理论建模-方法创新-技术实现-成果传播”的完整能力链。重点强化文献批判性阅读能力、跨学科技术方法迁移能力、多维数据分析能力（含计算建模与统计分析）、学术成果凝练能力（涵盖 SCI 论文撰写与基金申请），并通过持续性的学术训练实现科研素养的层级跃迁。
- **参与科研团队管理：**作为团队骨干，博士生需承担科研项目管理职能，包括：制定研究进度路线图、统筹实验资源分配、协调跨学科合作接口、编制阶段性研究报告等。应培养项目管理的系统思维，在质量把控、进度协调、成果整合等环节发挥枢纽作用，推动团队形成规范化、高效化的运行机制。
- **学术传承与人才培养：**博士生应扮演导师助教角色，通过构建“传帮带”培养体系，指导本科生/硕士生完成从科研启蒙到独立研究的过渡。重点培养后学者的问题发现能力、实验设计思维、学术写作规范，建立定期组会汇报、文献共读、数据互审等培养机制，同时注重学术价值观塑造与科研诚信引导。

博士生兼具研究者、创新者、管理者、教育者等多重角色，既是学术创新的中坚力量，又是协同创新的关键纽带。通过高质量完成学位论文、产出标志性成果、培养科研后备人才，既实现个人学术能力的跨越式发展，更为学科建设和团队发展注入持续动能。这种多维度的成长轨迹，最终指向独立科研工作者和未来学术领袖的能力塑造。

### 1.2.4 硕士生

硕士生是科研团队中重要的成员之一，其职责包括：

- **构建系统性知识体系：**硕士研究生应在导师指导下完成进阶课程矩阵学习，重点突破专业核心课程（含领域前沿专题、方法论训练模块等），系统掌握学科范式与关键技术。需注重理论工具与实证方法的有机融合，通过课程论文、项目实践等形式实现知识迁移，为后续研究奠定坚实的学术基础。
- **科研项目协同攻关：**作为科研梯队的生力军，硕士生应深度参与团队项目的全流程运作：在预研阶段承担文献图谱构建、技术路线图绘制等基础性工作 在实施阶段介入实验参数校准、数据采集标准化等关键技术环节 在总结阶段参与成果可视化呈现与辅助性论文撰写需主动建立与导师、博士生的跨阶段协作机制，在项目实践中培育科研协作意识。
- **研究实践：**硕士生须经历从研究设计到成果产出的完整训练：在导师指导下完成研究假设提出与技术可行性论证 制定分阶段研究计划并严格执行进度管控 运用专业工具（如 SPSS/Python/R 等）实施数据清洗与基础分析 遵循学术规范撰写学位论文，重点呈现方法论严谨性与结论可靠性该过程需同步提升独立研究能力与质量管控意识。
- **融入学术交流网络：**硕士生应分层次参与学术共同体互动：基础层：定期参加实验室组会汇报、跨校工作坊等，培养学术表达能力 提升层：在领域内区域性学术会议进行墙报展示或分组报告 进阶层：参与学术期刊论文的辅助撰写与修改反馈通过阶梯式参与实现从学术消费者到初级生产者的角色过渡。。

硕士研究生兼具科研学徒与项目执行者的双重属性，其培养过程强调“课程奠基-项目牵引-论文驱动”的发展模型。通过参与团队攻关任务、规范化学术训练、渐进式学术输出，既为课题组注入新鲜动能，又为个人职业发展（学术深造或产业研发）构筑能力基石。这种承上启下的定位使其成为连接本科通识教育与博士创新研究的关键枢纽，在科研生态系统中发挥人才培养的桥梁作用。

### 1.2.5 客座研究生

客座研究生或联培研究生，指在高等教育机构中，不属于该机构但在该机构进行短期或长期学术研究的学生。他们可能是来自其他高等教育机构的交换生，或者是参与合作研究项目的研究生。

本研究组不仅可指导客座研究生顺利发表学术论文、完成学业，同时能够为客座研究室提供丰富的科研资源和生活补贴，为未来科研之路加分。

客座研究生培养方案与本组研究生一致，但管理规则和成果归属方面存在差异。

客座研究生基本管理规则如下：

1. **资格确认：**客座研究生需要获得原培养单位导师的许可方可具有客座研究生的资格。
2. **工作地点及期限：**客座研究生的研究工作需在中国科学院西北生态环境资源研究院（兰州）进行，研究期限为 15 至 36 个月。
3. **指导教师：**客座研究生将由本组导师和原培养单位老师共同指导，以完成硕士毕业所需的期刊论文和毕业论文。
4. **研究方向：**研究方向可以与已开题的硕士论文方向不一致，选题可以根据实际情况适当调整。
5. **期刊论文发表：**发表的期刊论文可以以原单位为第一完成单位，研究生为第一作者，原培养单位导师和本组导师共同担任通讯作者。
6. **研究资源：**研究生可以使用本单位提供的数据、模型和计算资源进行研究。
7. **条件保障：**客座研究生将获得办公空间、电脑、网络和计算资源等必要的支持。
8. **野外工作保障：**参与野外工作会获得相应补贴，并购买野外商业保险。
9. **生活补贴：**研究生的生活补贴将根据其在西北研究院的学习和工作表现发放，薪资标准不低于原培养单位同级别研究生。
10. **住宿条件：**客座研究生将获得与西北研究院研究生相似的住宿条件，住宿费用根据实际情况而定。
11. **医疗保险：**本单位的医疗保险不适用于客座研究生，因此医疗保险由原培养单位负责。

12. 安全规定：在兰州学习期间，必须遵守西北研究院和原培养单位的保密、防疫和安全规定，并及时报告任何异常情况。
13. 问题解决：遇到学习、工作或生活上的困难或疑问，应立即与本组导师或原培养单位老师联系。
14. 行动自由：如需进出兰州，必须提前通知导师并获得明确同意。
15. 报告制度：客座研究生至少每月向原培养单位导师报告一次在兰州的生活和学习情况。
16. 退出机制：如对在兰州的生活、学习或研究工作感到不满意，可选择退出。若决定退出，需在离开前一个月告知原培养单位导师并取得同意。

## 1.3 硕博士差异

### 1.3.1 研究目标的本质差异

1. 硕士：聚焦已知领域的知识应用与技术优化
  - 通过改进现有方法（如算法调参、数据增强）解决具体问题，强调方案的工程落地性
  - 研究路径遵循成熟范式，成果价值体现在技术指标的提升（如检测精度优化 15%）
2. 博士：探索未知领域的原创突破
  - 需构建新理论框架或颠覆性方法（如开发具备数学证明的创新算法）
  - 研究成果需填补学科空白，引用量通常为硕士论文的 8.3 倍以上

### 1.3.2 能力培养的进阶路径

	硕士培养维度	博士培养维度
知识体系	专业核心技术的系统掌握	跨学科知识迁移与理论重构能力

	硕士培养维度	博士培养维度
研究方法	规范执行预设方案（如 3 种算法对比实验）	自主设计研究范式（含实验路径创新）
思维模式	解题思维：在限定条件下优化方案	出题思维：定义科学问题并开发验证工具

### 1.3.3 成果产出的层级要求

- 硕士论文：

- 侧重研究过程的完整性（如文献综述占 30%）
- 创新性要求限于局部改进（如现有模型适配新场景）

- 博士论文：

- 需构建完整的理论体系（如提出新型深度学习框架）
- 原创性成果需经受国际同行评议（如顶刊发表 2 篇以上）

### 1.3.4 团队角色的功能分野

角色	指导对象	核心职能	价值输出
硕士	本科生	技术实施督导（如代码调试指导）	项目执行的中坚力量
博士	硕士生 + 本科生	研究方向把控（如课题可行性论证）	科研创新的引擎驱动

### 1.3.5 学术生态的定位差异

- 硕士：

- 作为科研助理支撑团队运转
- 职业出口偏向产业技术专家

- 博士：

- 担任独立 PI 或课题负责人
- 承担学术传承与学科建设使命

总结表格：

对比维度	硕士研究生	博士研究生
研究目标	聚焦已知领域的知识应用与技术优化 • 改进现有方法解决具体问题 • 遵循成熟研究范式	探索未知领域的原创突破 • 构建新理论框架或颠覆性方法 • 填补学科空白
知识体系	专业核心技术的系统掌握 • 理论工具与实证方法融合 • 课程矩阵学习奠定基础	跨学科知识迁移与理论重构能力 • 创建新型分析框架 • 突破学科边界
研究方法	规范执行预设方案 • 技术路线图实施 • 标准化流程操作 (如算法对比实验)	自主设计研究范式 • 实验路径创新 • 方法论原创性开发
思维模式	解题思维 • 在限定条件下优化方案 • 技术可行性论证	出题思维 • 定义科学问题 • 开发验证工具
成果产出	局部改进型创新 • 侧重研究过程完整性 (文献综述占 30%+) • 技术指标优化 (如精度提升 15%)	体系化理论建构 • 完整理论体系创建 (如新深度学习框架) • 国际同行评议考验 (顶刊论文 2+ 篇)
团队角色	项目执行中坚力量 • 技术实施督导 (代码调试/实验校准) • 数据采集与基础分析	科研创新引擎驱动 • 研究方向把控 (课题可行性论证) • 跨学科接口协调
指导层级	本科生实践指导 • 实验操作规范培训 • 基础数据分析指导	硕士生学术指导 • 研究设计框架构建 • 论文写作范式传授
学术定位	科研助理角色 • 支撑团队日常运作 • 职业出口：产业技术专家	独立 PI 预备阶段 • 课题负责人角色 • 承担学科建设使命

对比维度	硕士研究生	博士研究生
培养周期特征	知识应用系统优化者 • 2-3 年标准化训练 • 技术执行能力培养	知识生产体系创造者 • 4-6 年创新性蜕变 • 理论建构能力塑造
学术贡献度	技术改良型贡献 • 单点问题解决方案 • 团队项目子模块贡献	范式变革型贡献 • 领域新方向开拓 • 团队学术影响力支柱
质量评估标准	过程规范性审查 • 方法复现可靠性 • 技术指标达成度	原创价值评审 • 理论创新显著性 • 学科推动力评估
学术网络参与度	区域性学术活动参与 • 会议墙报展示 • 辅助性论文撰写	国际学术圈层建设 • 大会主题报告 • 期刊审稿/基金评审
能力认证指标	专业技术认证 • 行业资格证书（如 PE）• 技术专利授权	学术影响力指标 • H 指数/高被引论文 • 学术奖项获得
失败容忍度	预设路径偏差容忍度低 • 强调方案可重复性 • 风险规避型探索	探索性试错空间大 • 鼓励高风险创新 • 允许理论重构
职业发展路径	垂直技术通道 • 特定领域专家 • 产业研发工程师 • 科技企业 CTO	多维发展空间 • 学术机构 PI • 智库首席科学家 • 科技企业 CTO

### 系统差异总结

博士与硕士的本质区别在于：前者是知识生产体系的创造者（定义问题 + 构建方法），后者是知识应用系统的优化者（解决问题 + 改进方案）。这种差异映射到培养周期上，博士需要额外 3-5 年完成从技术执行者到理论创新者的蜕变，其成果需经受更严苛的学术检验。二者的角色互补构成科研生态系统的完整闭环。

## 1.4 待遇

### 1.4.1 一般工资或补助

本研究组内每个成员的工资收入由身份决定。

- 硕士生和博士生工资按照研究生处的规定执行；免学费后，博士研究生每人每年不低于 70000 元，硕士研究生每人每年不低于 33000 元（2024 年 9 月起执行）。具体可以与中国科学院西北研究院研究生处沟通。
- 博士后或特别研究助理工资主要按照中国科学院西北研究院的规定，额外津贴与本组 PI 商量决定。
- 一般研究助理工资的变幅较大，依据为中国科学院西北研究院科研助理管理规则；工作属于劳务派遣，长期工作则签署劳动合同，并享有劳动法规定的医疗、养老和失业保险。工资外的收入与具体工作内容和工作性质有关。

保送本研究组的硕士研究生，自动获得 12000 元奖学金（2023 年标准）。

优秀科研助理可获得每年 5 万元自主科研经费，支持个人科研兴趣。

### 1.4.2 奖金

每年过年前 1-2 周，开年底总结会。每位组员对过去一年的工作做一个总结，按照对本组的贡献打分，获得相应的年度奖金。多劳多得，详见2.6。

### 1.4.3 住宿

- 研究生住宿由西北研究院研究生处安排住宿。
- 合作研究者、项目聘用和客座研究生住宿由本组 PI 协助解决。



# 第二章 规则

进组的时候，每个学生本质差异并不大，但是毕业的时候，差异会非常大。其原因就在于是否能够真正把以上这一节“规则”在行动中、实践中坚持下去。

## 2.1 研究组规则

### 1. 对学术不端零容忍

学术诚信是立身之本，如果做不到这一点，请换一个职业。任何抄袭、剽窃、伪造数据等学术不端行为，都会被严肃处理。请务必遵守学术道德规范。

### 2. 事事有回应，件件有着落，凡事有交代

所有组里安排的事情，一定要在规定时间前做出反馈。不要以“我正在做”和“还不完美”作为你没有回音的借口。你可以说“我正在做”，但是一定要给出一个时间节点，说明你预计何时完成。即使你没有完成，也要给出一个合理的解释。例如：我正在做，预计下周一完成；或者我正在做，预计下周一完成，但是因为 XXX 原因，可能会延迟到下周三完成。

### 3. 完成比完美重要

科研就没有“完美”的时刻。先完成在努力完美，才是正确的科研行为。一味完美实际是逃避和拖延症的外在表现。当别人问你进度的时候，你的回答应该是：先展示现有结果；然后报告已经完成 50%，剩下

的 50% 预截止哪天哪时完成并提交。你要让别人看到你在努力工作，而不是在拖延。

#### 4. 有疑问请尽早问

我们对任何问题都持开放态度，也不会因为你问了问题而挑剔你，但是如果存在问题不问，靠个人猜测，往往会引起很多误解。在我们组里没有“傻问题”。

#### 5. 保持好奇

对学术中出现的每个问题都要有好奇心。虽然不一定深入研究，但是必须有所思考。

#### 6. 尊重和褒奖他人科研成果

使用任何他人的数据、观点、方法、逻辑证明，都要给别人足够的学术肯定（Academic Credit），形式包括但不限于合理作者署名、论文致谢、广而告之等；强烈不建议使用金钱等方式答谢。肯定他人学术贡献，并不会使你的研究成果变得暗淡失色。

#### 7. 持开放心态

对他人的观点持开放态度，学习的态度。若要否定别人的数据、逻辑或观点，要拿出比对方更坚实的证明过程。

#### 8. 工作和生活分开

为了更长久战斗在科研领域，要能够同时享受科研工作和个人生活。不加班、不熬夜的前提是尊重正常的工作时间。

#### 9. 尊重他人的私人空间

若无紧急事态，不要打扰组员的周末、假日和晚上的个人时光。

#### 10. 科研工作需要自由空间

每个人都有不同的工作模式。工作的时间、地点、方式都存在个人倾向，选择个人最合适的方式。但同时，在团队合作与单位管理规定之间做出适当的、合理的妥协。

### 2.1.1 赫尔曼法则 20 条

在行动上、思维上，请参考赫尔曼法则的 20 条 (Herman, I. **Following the law**. Nature 445, 228 (2007). <https://doi.org/10.1038/nj7124-228a>)。

1. **Your vacation begins after you defend your thesis.**
  - 只有论文答辩完，你才会有假期。
2. **In research, what matters is what is right, and not who is right.**
  - 研究中重要的是：什么是正确，而非谁是正确。
3. **In research and other matters, your adviser is always right, most of the time.**
  - 在科研及其他事务中，你的导师大多数时候都对的。
4. **Act as if your adviser is always right, almost all the time.**
  - 行动时，请假设导师几乎总是正确的。
5. **If you think you are right and you are able to convince your adviser, your adviser will be very happy.**
  - 若你确信自己正确并能说服导师，导师会非常欣慰。
6. **Your productivity varies as  $(\text{effective productive time spent per day})^{100}$ .**
  - 你的效率与“每日有效科研时间”的 100 次方成正比。
7. **Your productivity also varies as  $1/(\text{your delay in analysing acquired data})^{100}$ .**
  - 你的效率也与“分析数据的拖延时间”的 100 次方成反比。
8. **Take data today as if you know that your equipment will break tomorrow.**
  - 采集数据时，需假设仪器明天就会损坏。
9. **If you would be unhappy to lose your data, make a permanent back-up copy of them within five minutes of acquiring them.**

- 若数据丢失会让你痛苦，请在获取后五分钟内永久备份。
10. Your adviser expects your productivity to be low initially and then to be above threshold after a year or so.
- 导师默认你的初期效率较低，但一年后应有所进步。
11. You must become a bigger expert in your thesis area than your adviser.
- 你必须成为自己研究领域内比导师更专业的专家。
12. When you cooperate, your adviser's blood pressure will go down a bit.
- 你配合得好，导师的血压能下降一点。
13. When you don't cooperate, your adviser's blood pressure either goes up a bit or it goes down to zero.
- 你不配合，导师血压会上升，或骤降至零。
14. Usually, only when you can publish your results are they good enough to be part of your thesis.
- 通常，只有达到发表水平的成果才足够优秀以纳入论文。
15. The higher the quality, first, and quantity, second, of your publishable work, the better your thesis.
- 可发表成果的质量（首位）与数量（次位）越高，论文越优秀。
16. Remember, it's your thesis. You (I) need to do it.
- 谨记：这是你的毕业论文，需由你自己完成。
17. Your adviser wants you to become famous, so that he/she can finally become famous.
- 导师希望你成名，以便他/她最终因你成名。
18. Your adviser wants to write the best letter of recommendation for you that is possible.
- 导师愿为你而尽量写最有力的推荐信。

**19. Whatever is best for you is best for your adviser.**

- 对你最有利之事，亦对导师最有利。

**20. Whatever is best for your adviser is best for you.**

- 对导师最有利之事，亦对你最有利。

## 2.2 导师制度

本组学生入学都有两个不同的“导师”，一个 Mentor，一个是 Advisor。我们最常说的导师实际是指 Academic Advisor。Advisor 入学已经确定，Mentor 需要学生入学后自己联系确定。

Mentor 和 Advisor 在职责和角色方面存在一些区别，主要表现在以下几个方面：

- **职责不同：**Mentor 的职责是帮助学生在职业和个人发展方面提供支持和指导，通常与学生的整个职业生涯有关，包括个人成长、人际关系、职业规划等。而 Advisor 的职责主要是在学术领域提供指导和建议，如研究课题、实验设计、数据分析、论文写作等。
- **关注的范围不同：**Mentor 的关注范围通常是全面的，包括学术、职业、个人等多个方面，而 Advisor 的关注范围主要是学术研究。
- **指导方式不同：**Mentor 通常通过面对面的交流、指导和建议来帮助学生，建立师生之间的信任和互动，而 Advisor 的指导方式则更加注重实践和技术方面的指导和建议。

综上所述，Mentor 和 Advisor 在职责和角色方面存在一些区别，但在实际工作中，他们的工作内容和职责也会有所重叠和交叉。

### 2.2.1 学术导师 (Academic Advisor)

你的学术导师，不是你的朋友。

学术导师的主要提供给学生的服务：

1. 条件：提供必要的科研环境，包括科研软硬件条件、学习的环境、经费支持等。
2. 学习：指导学生选择适合的课程和学术计划；提供关于学术要求和课程安排的信息和建议；
3. 研究：指导学生的学术研究工作，提供研究思路、数据、方法、分析等研究的建议；
4. 答疑：帮助学生解决学术方面的问题，提供指导和建议；

#### 学术导师的行为准则：

导师与学生天生不平等，任何可能带有不平等而导致学生无法拒绝或不好意思拒绝的请求，导师都应该回避。

1. 不歧视学生——导师不因任何原因（包含但不限于性别、籍贯、疾病、相貌、专业、兴趣、家境等）而对学生有歧视语言或行动。尊重学生人格，尊重学生的个人自由。
2. 不收学生礼物——学生毕业以前，不接受任何学生的礼物，包含但不限于水果、鲜花、特产、贺卡等。不接受学生或其家人请吃饭。
3. 不干涉学生个人生活——导师只负责学习和科研相关内容；学生拥有安排个人生活、保护隐私的权利。

#### 2.2.2 生活导师 (Mentor)

##### 你的生活导师并不是你的保姆

哪些事情你应当寻求生活导师的帮助？

- 帮助学生制定职业规划和目标，并提供相关的建议和指导；
- 指导学生在学术、职业和个人成长方面取得进步；
- 与学生保持密切联系，提供建设性的反馈和指导；
- 为学生提供职业发展机会和推荐信；
- 帮助学生在职业和个人发展方面建立自信和自尊心。
- 科研资源信息（在寻求帮助之前，先看一遍本实验室手册）

- 科研生活的困惑
- 职业规划
- 师生、同学关系
- 重要的且需要与有经验的研究者沟通的疑惑。

## 2.3 科研经费

科研经费是支持本研究组生存的关键，组内成员的学费、工资、生活补助、办公空间、书籍、电脑/服务器/设备、数据购买、差旅费、版面费、甚至电费等等都需要充足科研经费支持，因此每个组员都有义务争取经费，合理支出经费。

基本的规则如下：

1. PI 是经费支配的实际负责人，并对经费支配方式具有解释权。
2. 所有经费支出，都应遵守中国科学院西北生态环境资源研究院经费管理规范，以及各经费预算。
3. 可支出的大类分为：
  1. 固定资产：服务器、台式/笔记本电脑、打印机、显示器、投影仪、办公桌椅等。所有固定资产有规定的使用年限。
  2. 文具/耗材：纸笔、打印耗材、数据线
  3. 书籍：限与学术研究直接相关的中英文书籍。书籍归单位所有。
  4. 差旅：会议/考察等工作差旅中涉及的出租车、火车、飞机、租车费用，住宿费，会议注册费，国际签证费用等。野外工作时同时享受单位规定的野外补助。
  5. 文献版面费：出版论文，专著等。
  6. 技术服务费：如数据采集、程序编写等第三方提供的技术服务。

## 2.4 组会制度

组会是研究团队内部定期开展的集体讨论活动，目的是为了促进团队成员之间的交流和合作，推进研究项目的进展。组会的一些好处：

- **促进交流和合作:** 成员间分享研究成果、问题和解决方案，促进彼此之间的交流和合作，避免孤立地开展研究工作。
- **加深理解和认识:** 组会可以邀请内部或外部专家来讲解前沿的研究进展或关键技术，帮助团队成员加深对研究领域的理解和认识。
- **提高研究能力和学术素养:** 组会可以让团队成员接触不同的研究思路和方法，提高研究能力和学术素养，也可以帮助团队成员更好地理解应用相关研究工具和软件。
- **促进项目进展和成果输出:** 组会可对项目进展进行定期汇报和评估，发现问题并及时解决，也可以帮助团队成员充分讨论项目的研究方向和创新点，促进成果的产生和输出。
- **培养演讲和表达能力:** 成员有机会进行学术报告和演讲，提高其学术交流和表达能力，也为其日后参加学术会议和报告做好准备。

组会组织方式：

- **定期召开:** 每周一次。确保活动的持续性和规律性。
- **轮流主持:** 成员轮流担任组会的主持人，负责安排会议议程、邀请讲者和主持讨论环节等。
- **人员安排:** 团队成员，包括老师、学生和研究助理，都有必要参加。每周一个团队成员主讲，并有下一周的主讲人作为本周的主持人，
- **报告和讨论:** 组会可包括研究成果的报告和讨论，或者邀请内部或外部专家进行报告和讲解，以加深对研究领域的理解和认识，同时也可为团队成员提供一个分享和交流研究思路和方法的平台。
- **问题与解决:** 组会可以安排团队成员分享遇到的问题和解决方案，以促进成员之间的交流和合作，并解决项目中遇到的难点和问题。
- **思维碰撞:** 组会也可安排一些与研究领域相关的主题，引导成员进行讨论和思维碰撞，以激发创新思维和开拓研究思路。

## 2.5 给学生的建议

在前面科研规则的基础上，以下是本组的一些重要规则。

### 2.5.1 重要规则

#### 1. 安全第一

确保个人安全。事关安全的事情，尽早向导师汇报，或者寻求帮助。对导师、同事、同学的可疑或不合理行为保持警惕。

保守国家秘密。

#### 2. 进导师办公室，永远敞开门

为了学生好，也为了导师好。

#### 3. 明白学习与工作的关系

学生在学习上享有充分的自主权。无论是论文的选题，方法，还是实现目标的工具、证明逻辑、书写办法等事关学生个人科学研究的内容，导师只能指导学生，无法替代学生，更不应该替学生做决定；学生需要在其中珍惜和把握自主权。

#### 4. 科研工作（即学生获得工资的缘由）由具体科研项目决定。

学生需要规划合理时间，按照项目设计的工具、方法和思路来完成项目所规定之任务。研究成果可作为学生的论文成果，但也可能与学生自主选择的科研方向不同。学生在论文方向的选择上，可以选择与工作一致的方向，也可以选择不同方向，但必须预料到两种选择所导致时间和精力的冲突。

### 2.5.2 优秀科研人的特点

- **挑战自我的勤奋**——没有什么是学不会的。
- **有组织、有纪律、有规划**——对日常作息、研究工作、文件组织、交流汇报、论文结构、语言沟通都能安排得当。本质就是把“事事有回应，件件有着落，凡事有交代”落在行动上。
- **积极主动**——表现为积极主动地承担责任，积极主动与周边的人交流、积极主动的参与学术活动。
- **自己思考**——不断地思考/完善科学问题，不断地完善方法，不断地追问逻辑关系，不断地挑战解释的可能性。

- **敢于质疑**——敢于质疑导师、质疑权威、质疑自己；能质疑自己才是核心；“吾爱吾师、吾更爱真理”，质疑权威的手段是用客观证据证明自己是对的；
- **会用工具**——不囿于最简单、通用工具，有挑战自我的意愿和能力；能使用、敢使用别人不会的工具；对所有问题有最高效的工具解决方案。
- **团队合作**——能为别人提供价值，也能利用好别人给自己的帮助。
- **会讲话**——能用简单、有逻辑的语言，给别人讲明白自己的想法、概念、成果、经历等。
- **劳逸结合**——能合理安排科研和娱乐。
- **与人为善**——能与他人建立良好的关系，能与他人建立良好的合作关系，能与他人建立良好的信任关系。

## 2.6 年度考核

本组成员需每年完成以下任务：

### 1. 文献阅读 (1 分/篇)

- 有价值英文文献；
- 写成阅读文章发于本组公众号，质量获得审核可得分；
- 使用标准 PPT，在组会汇报（30 分钟）；

### 2. 参加学术会议 (5 分/次)

- 国内正式学术会议，在会议做口头报告；
- 会议汇报内容为在本组的研究内容；
- 汇报 PPT 至少在会议前 14 天以上完成，并通过 PI 审核；
- 全程参会，参会做笔记，会后做总结；
- 材料齐全，会议费用可全额报销。

### 3. 发表学术论文 (10 分/1 篇)

- 核心以上期刊发表学术论文
- 本研究所为第一署名单位，本组 PI 为通讯作者，本组项目为第一资助。

年底考核中，每个积分可获得 100 元奖金，若年底考核超过 20 分，额外奖励 5 个积分。完成任务过程，凡涉嫌“抄袭”等学术不端行为，请直接退学或离开本研究组。

## 2.7 安全制度

安全是科研工作的基石，也是不可逾越的红线。任何科研成果都不能以牺牲安全为代价。本团队坚持“安全第一，预防为主”的原则，所有成员必须时刻保持高度的安全意识。

### 一、人身安全

#### 1. 野外工作安全

- a. 行前报备：野外考察前必须制定详细计划（含路线、日程、应急预案），并向导师报备。
- b. 结伴原则：野外工作期间，严禁单人行动。严禁单人进入野外无人区或高风险区域作业，必须两人以上结伴而行。
- c. 风险防范：作业时需时刻警惕地形、气象、野生动物等潜在风险。如遇极端天气或不可控风险，应立即停止作业，撤离至安全地带。  
数据诚可贵，生命价更高。
- d. 保险与装备：野外人员必须购买足额的人身意外伤害保险。出发前检查车辆、通讯设备（卫星电话/GPS）、急救包等物资是否完备。设备等必须出发前指定专人负责。

#### 2. 实验室与办公安全

- a. 安全区域：所有实验室内、办公室等区域，必须设置严格的安全措施，防止未经授权的人员进入。
- b. 水电安全：最后离开办公室或实验室的人员，必须负责检查并关闭门窗、空调、饮水机及非必要持续运行的实验设备电源。
- c. 消防意识：熟知灭火器位置及使用方法，保持消防通道畅通。严禁在实验室内违规使用大功率电器，严禁私拉乱接电线。

#### 3. 通勤与出行

- a. 遵守交规：无论是市内通勤还是出差，均应严格遵守交通规则，不乘坐黑车。
- b. 禁止驾驶：出差或出野外期间，学生或工作人员不得驾驶车辆。
- c. 出行报备：学生离开常住地（学习、研究所、家、其他合作团队所在地），必须提前通知导师，获得批准后方可离开。安全返回驻地，必须及时向导师报告。

## 二、设备与环境安全

### 4. 设备操作规范

- a. 培训授权：大型贵重仪器或有潜在危险的设备，未经过培训及导师授权，严禁擅自操作。
- b. 使用记录：严格按照操作规程使用设备，使用后按要求填写使用记录。

### 5. 软件环境安全

- a. 正版化：办公电脑应优先使用正版软件或开源替代品。
- b. 软件禁令：严禁在团队电脑上安装盗版软件、游戏软件及存在安全隐患的“流氓软件”（如 360 全家桶、各类非必要的电脑管家、性能监控、网速监控等），以防止系统卡顿或数据后门泄露。

## 三、数据与信息安全

### 6. 账号防护

- a. 高强度密码：电脑及服务器必须设置高强度登录密码（建议字母 + 数字 + 符号组合）。
- b. 禁止共享：严禁将个人账号密码告知他人或多人共用同一账号。
- c. 定期更新：严禁将超算、服务器、工作站用户名和密码与其他人分享。定期（每 12 个月）修改超算登录密码。
- d. 隐私保护：严格保护团队成员隐私，不得在未经允许的情况下向他人提供团队成员的个人信息。

### 7. 物理防护

- a. **自动锁屏**: 必须启用电脑的自动锁屏功能（建议设置为 5-10 分钟无操作自动锁定）。
- b. **手动锁屏**: 离开座位超过 10 分钟，必须手动锁屏（Windows: Win+L; Mac: Ctrl+Cmd+Q）。
- c. **禁止常亮**: 禁止长时间屏幕保持开启状态，避免屏幕上的个人信息泄露，也避免设备无人监管状态可能引起的安全事故。

#### 8. 网络安全

- a. **防范钓鱼**: 办公网络环境下，不随意点击不明链接，不下载来源不明的邮件附件，警惕网络钓鱼。
- b. **禁止 P2P**: 禁止使用电驴、BitTorrent 等下载工具。
- c. **防范恶意软件**: 避免下载包含病毒、木马等恶意软件。
- d. **禁止手机扫描文件**: 禁止使用手机扫描重要文件（合同、信息表等），避免文件被不法软件泄露。
- e. **在线会议安全**: 使用团队内科技云会开启会议时，需向导师报备记录。会议需设置密码。会议纪要或录屏，无许可不得外传。

## 2.8 保密制度

本团队严格遵守国家保密法律法规，高度重视数据安全与隐私保护。团队成员需知悉，我们通常仅作为数据的临时使用者或处理者，并不必然拥有数据的所有权，因此更应严格履行保密义务。

### 一、保密数据范围

#### 1. 保密数据定义

本制度所指“保密数据”包括但不限于：

- a. **科研数据**: 野外观测数据、实验数据、分析过程数据、未发表的研究结果、源代码、专利申请材料、论文草稿等。
- b. **保密文件**: 包括各种项目申请书、研究计划、合同、协议等。
- c. **合作数据**: 合作伙伴提供的商业机密、专有数据或受协议保护的数据。

- d. **水文径流观测数据**: 通过非公共渠道获取的水文径流观测数据。这类数据我们团队签署过或承诺过数据保密责任: 不公开、不分享、不泄密。

## 二、数据处理规范

### 2. 数据存储

- a. **存储介质**: 所有保密数据必须存储于团队指定的服务器或团队配发的工作设备上。
- b. **禁止私存**: 严禁将保密数据保存于任何私人电脑、移动硬盘、U 盘或未经授权的云存储(包括但不限于百度云盘、坚果云、Dropbox、Google Drive 等)中。

### 3. 数据传递

- a. **内部传递**: 保密数据在团队内部传递时, 必须使用团队 NAS 或指定的加密通道。严禁通过微信、QQ 等社交软件传输敏感数据。
- b. **外部传递**: 可通过加密的团队内云存储, 或者加密 U 盘/光盘方式向已经获得授权的外部合作伙伴传递敏感数据。严禁通过微信、QQ 等社交软件传输敏感数据。

### 4. 数据使用

- a. **AI 使用限制**: 不得使用任何公共 AI 工具(如 Claude、ChatGPT、DALL-E、GLM、DOUBAO 等)进行敏感数据的处理或分析。使用 AI 时, 不得向 AI 提供任何保密数据。
- b. **人员限制**: 未获授权或未配备团队专用设备的同学及工作人员, 不得参与保密数据的处理工作。
- c. **离组交接**: 在毕业、离职或退出项目时, 无条件归还所有数据资产, 并彻底删除设备中可能残留的任何相关缓存或副本, 不私自留存。
- d. **五不原则**: 在处理保密数据时, 必须严格遵守五不原则, 即不分享、不公开、不传输、不私存、不扩大范围。

### 5. 设备安全管理

- a. **设备存储**: 所有用于处理保密数据的电脑及服务器, 必须存储在团队指定的安全区域, 禁止未授权人员访问。

- b. **访问控制:** 所有用于处理工作的电脑及服务器必须设置强密码（建议字母 + 数字 + 符号组合），并定期更新密码。
- c. **自动锁屏:** 必须启用自动锁屏功能（建议设置为 5 分钟或更短时间无操作自动锁定），离开座位时应手动锁屏。
- d. **账号独立:** 严禁将 NAS 账号密码、超算服务器账号密码、服务器账号密码、数据库访问权限等借予他人使用；即使是团队内部成员，也必须使用各自的独立账号，禁止共用或转借账号。

### 三、信息披露与责任

#### 6. 信息披露规范

- a. **禁止外泄:** 严禁向任何第三方（包括非本项目的其他人员）泄露保密数据。
- b. **报告与论文:** 对外做报告发论文，必须获得数据提供者和导师许可。不得未经许可，将保密数据用于 PPT、报告或论文。
- c. **公开数据平台:** 禁止向 Zenodo, figshare, ScienceDB 或其他公开数据平台上传保密数据。

#### 7. 泄密应对

- a. **泄密报告:** 若发生数据泄露，应立即报告给导师，并采取必要的措施防止再次泄露。
- b. **安全修复:** 一旦发现设备被恶意访问或数据泄露，应立即采取措施修复，包括但不限于删除泄露数据、更新密码、禁用相关账户等。所有泄密或者被利用的情况，都应第一时间报告给导师。

#### 8. 适用范围

- a. **适用人员:** 本制度适用于团队内所有学生、长期工作人员及临时访问人员。
- b. **监督责任:** 每一位成员都有义务相互监督，确保数据保密、信息安全。



# 第三章 研究生培养计划

## 3.1 研究方向

本研究组总体科研方向请参看：1.1.1节。

本组各成员的研究方向都应当是显著不同的。

学生的研究方向主要有两个因素确定：

1. 学生个人兴趣；
2. 本组项目状况。

对于“科研方向选择”的几条建议：

1. 选择和放弃之前，一定要充分了解它。
2. 难易是相对的。例如：有的领域没人做，A 觉得是因为难而无人做，B 觉得无人做才容易出结果。有的领域已经有很成熟的分析/模拟流程，A 觉得好上手容易做，B 觉得已经做烂的事情，再怎么深入也只能是步人后尘，难以独树一帜。
3. 热情是好事，但必须脚踏实地一步一步走出来。
4. 空话、大话很容易说，但是科研要求你知道研究内容中几乎所有细节。  
你要做到每个细节都计算过、了解过、思考过。

## 3.2 时间表

时间点	学业时间	任务	任务目标
-12 个 月	入学前一 年	文献阅读；参与项目；补课； 学软件	熟悉科研环境和 流程
-6 个月	入学前半 年	文献阅读；参与项目；补课； 学软件	掌握软件
0-6 月	入学	文献阅读；完成必修课；参与 项目	充分了解研究方 向
7-12 月	入学 1 年	文献阅读；完成必修课；参与 项目	确认研究方向
13-18 月	入学 1.5 年	研究综述；试验设计	形成综述文章
19-24 月	入学 2 年	开展研究、试验、分析；撰写 论文	完成试验，发表 论文
25-30 月	入学 2.5 年	总结试验，撰写论文；开展改 进试验	毕业论文
31-36 月	临近毕业	完成毕业论文、送审、答辩	学位:-)

### 3.2.1 培养指标

1. 英文文献汇报 100 篇。导师负责推荐其中 50 篇，学生自己自行阅读 50 篇。
2. 公众号论文阅读笔记，15 篇。
3. 一篇一作 SCI 论文
4. 完成毕业论文。

## 3.3 课程推荐

- 地理信息系统导论

- 掌握有关空间栅格/矢量数据的区别、格式和读写方法；
- 熟练掌握空间数据处理、数据分析、数据投影；

- 熟练掌握水文有关的空间数据；
- 能够利用 DEM 生成流域边界和河网，了解 D8 等算法。

- **水文学导论**

- 深入理解流域水循环；
- 掌握降雨、蒸散发、入渗、产流、汇流和地下水为物理过程；
- 了解水文统计和水文模型；
- 掌握水文分析的基本方法和流程。

- **地下水力学**

- 掌握基本的地下水的基本概念，如土壤孔隙度、含水量、水力学传导度、承压/非承压地下水、Duijput Assumption；
- 掌握地下水力学的重要公式，贝努力方程、达西定律、理查兹方程、van Genuchten 方程；
- 能够熟练建立/推导承压和非承压地下水的控制方程。

- **地学中的数值方法（有限差分、有限元）**

- 掌握数值方法的基本原理；
- 掌握从泰勒展开式到一阶、二阶精度有限差分方法的推导/证明过程；
- 明确理解数值方法的稳定性、收敛性、一致性的含义和影响；
- 掌握 CFL 条件的计算和证明方法；
- 能够编程数值方法求解简单的对流扩散方程。

- **气象与气候学**

- 了解基本的气象要素的含义和观测方法；
- 理解水和能量在气候/气象过程中的运动规律；
- 熟练掌握气象和气候学的基础控制方程。

- **R/Python**

- 基本操作 (安装、扩展包安装)
- 能够使用 R/Python 读取各种类型数据 (shapefile, GeoTiff, HDF, NetCDF 等)；
- 熟练掌握数据分析功能 (基本统计、数据挖掘、空间数据处理/分析、简单爬虫)；
- 熟练掌握做图功能 (ggplot, rgl 等)。

### 3.4 必读书籍

- Beven, K., 2012. Rainfall-Runoff Modelling: The Primer: Second Edition, John Wiley & Sons, Ltd, Chichester, UK. <https://doi.org/10.1002/9781119951001>
- Sorooshian, S., Hsu, K.-L., Coppola, E., Tomassetti, B., Verdecchia, M., Visconti, G. (Eds.), 2008. Hydrological Modelling and the Water Cycle, Water Science and Technology Library. Springer Berlin Heidelberg, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-540-77843-1>
- 威尔·布鲁萨著, 王忠静, 梁友, 耿国婷译. 2017. 水文学导论. 气象出版社. 北京
- 徐宗学, 2009. 水文模型. 科学出版社, 北京.
- 贾仰文, 王浩, 倪广恒, 2005. 分布式流域水文模型原理与实践. 中国水利水电出版社, 北京.
- Kang-tsung Chang 著, 陈健飞等译. 2019. 地理信息系统导论. 科学出版社. 北京.
- McDonnell, J.J., 2019. Navigating an Academic Career, Special Publications. Wiley. <https://doi.org/10.1002/9781119642206>
- 

### 3.5 必读文献

- 如何成为合格研究生
  - Herman, I. (2007). **Following the law**. Nature 445, 228. <https://doi.org/10.1038/nj7124-228a>
  - Schwartz, M. A. (2008). **The importance of stupidity in scientific research**. Journal of Cell Science, 121(11), 1771. <https://doi.org/10.1242/jcs.033340>
  - Sabretooth. (2001) **A memo to graduate students and post-docs**. J Cell Sci 114 (14): 2547–2548. <https://doi.org/10.1242/jcs.114.14.2547>
  - Prather, C. Choate, D., Michel, M.(2009). **Putting the “Ph”**

**back into “PhD”: framing graduate research in a theoretical context.** Frontiers in Ecology and The Environment. 389-390. <https://doi.org/10.1890/1540-9295-7.7.389>

- Odom DT. **How to evaluate a graduate studentship, or choosing the right doctoral advisor.** Genome Biol. 2013 Apr 29;14(4):114. <https://doi.org/10.1186/gb-2013-14-4-114>
- McDonnell, J. (2019). **Navigating an Academic Career.** Wiley. <https://doi.org/10.1002/9781119642206>

- **关键科研文献**

- Shu, L., Ullrich, P. A., & Duffy, C. J. (2020). Simulator for Hydrologic Unstructured Domains (SHUD v1.0): numerical modeling of watershed hydrology with the finite volume method. Geoscientific Model Development, 13(6), 2743–2762. <https://doi.org/10.5194/gmd-13-2743-2020>
- 舒乐乐, 陈昊, 孟宪红, 常燕, 胡立堂, 王文科, et al. (2024). 地表-地下过程耦合的数值水文模型综述. 中国科学: 地球科学, 54(5), 1484–1505. <https://doi.org/10.1360/SSTe-2022-0420>
- Duffy, C. J. (2017). The terrestrial hydrologic cycle: an historical sense of balance. WIREs Water, 4(4), e1216. <https://doi.org/10.1002/wat2.1216>



# 第四章 研究的范式

## 4.1 科学问题

具体什么是科学问题，难有准确定义，但是我们可以从一些侧面了解什么是科学问题。

科学问题的一些诠释：

- 一个你能通过科学实验方能解决的问题；
- 一个讨论过、研究过，但尚未形成定论的问题；
- 一个可能被你的科学设计而解决的问题。

### 4.1.1 科学问题的案例

(未完成)

先阅读一下内容：

<https://www.scribbr.com/research-process/research-question-examples/>

## 4.2 研究区

### 4.2.1 资料收集和处理

#### 4.2.1.1 基本资料

- 社会经济发展: GDP, 主要经济支柱
- 人口: 人口分布以及人口时间变化, 重要时间节点
- 行政归属

#### 4.2.1.2 地理数据

- 高程: DEM, 地形图
- 土地利用: 不同分类体系下的土地利用; 不同时期的土地利用
- 土壤质地: 沙粉黏比例, 土壤容重、有机质含量、土壤层厚度
- 地质背景
- 降雨分布
- 温度分布
- 行政边界: 各级行政区划
- 城市/居民点
- 道路: 各级公路铁路

以上数据应当形成标准图件, 可在论文、报告中直接使用。

#### 4.2.1.3 气象再分析资料

- GLDAS
- ERA5-Land
- CMFD
- PRISM

### 4.3 检验、检验、检验

科研工作的每一步，都需要进行检验，而且是双重检验（Double Check），用不同的两个种策略对同一个数据、结果进行检验。

常见的几个检验方法：

- 异常值检验和原因分析
- 值的正确范围检验
- 单位及量级检验
- 手动演算
- Excel 演算
- 对公式进行绘图
- 因果关系、相关关系、逻辑关系检验
- 物质和能量平衡检验
- 文献数据的交叉对比



# 第五章 论文写作

## 5.1 IMRaD

IMRaD 是学术论文和学术讨论的基本框架。无论是写学术论文、毕业论文、毕业答辩、组会报告、给同行讲解，都请不断练习这种表达方式。

**IMRaD 即 Introduction, Method Results and Discussion。** 中文为：引言、方法、结果、讨论。尽管，论文写作中，通常在结尾加入 Conclusion 一节，但 Conclusion 是对前面所述内容的总结，而 IMRaD 则是相互不重复的内容。

IMRaD 的基本结构如下：

### 1. 引言 (Introduction)

1. 研究的背景：具体研究问题之后的大背景。
2. 研究的意义：该研究对于人类认识、生产生活、其他研究的价值。
3. 前人研究：重点关注 5 个要点：
  1. 前人的研究对象；
  2. 前人的研究方法；
  3. 前人研究出来的结论；
  4. 这些研究对于本次研究的借鉴价值或提供的思路/数据/背景；
  5. 前人研究的局限性/不足，或尚未回答的问题。从此引出本次研究的科学问题——即，在前人研究基础之上，我在本次研究中打算回答的问题，或称为零假设 (Null Hypothesis)。
4. 简要地阐明本论文的组织结构。

## 2. 方法 (Method)

1. 研究对象。我们的研究对象通常为研究区。这里需要对研究区的基本地理背景做简要说明。包括研究区的面积、所属行政区、海拔、气候、人类活动情况等等。
2. 方法：模型介绍、计算公式、统计方法、分析工具等等。

## 3. 结果 (Results)

1. 方法验证：通常为模型的校准、方法的初步结果，论证方法的有效性。
2. 直接结果：方法核算出来的结果
3. 间接结果：基于数据间的关系和逻辑
4. 新发现/新结论：对以上结果的分析，此部分内容应该分散在个结果分析之中。

## 4. 讨论 (Discussion)

1. 不确定性讨论：对模型、数据、参数、方法等可能得不确定性讨论和分析。
2. 当前研究可能存在的不足：大方承认研究的尚未全面考虑的方面。这部分往往是受制于客观条件。
3. 后人或者今后研究可能改进和进一步论证的方向和可能性。
4. 研究的局限性：对研究结果的局限性进行讨论。包括空间、时间、方法、数据等方面局限性。

## 5. 结论 (Conclusion) 将以上 IMRaD 的内容总结成三段话：

1. 背景和意义，一两句话。
2. 方法，一句话。
3. 结果，分条目列举文章的新发现。
4. 讨论，总结一句最重要的不足和局限性。
5. 展望和价值，本研究的发现对于科研或生产生活有何指导意义。

## 5.2 题目 (Title)

题目是影响读者第一印象的重要因素。作为地理学家和水文学家，在选择题目时需要注意以下几点：

### 1. 精准定位核心贡献

- 包含：研究区域（如流域/地貌单元）+ 核心方法（如 SWAT 模型/遥感反演）+ 关键结论（如“径流响应机制”）
- 例：青藏高原冰川融水对三江源区径流变化的贡献度研究（对比：“水循环研究”）

### 2. 关键词前置原则

- 将学科关键词置于标题前部，提高检索可见度
- 推荐结构：【研究方法】+【研究对象】+【研究区域】
- 例：基于 InSAR 技术的黄河三角洲地面沉降监测（对比：地面沉降监测新方法）

### 3. 长度控制与信息密度

- 严格控制在 12-20 个汉字/英文 8-15 个实词
- 中文标题建议结构：主标题（核心结论）+ 副标题（方法补充）
- 例：城市化进程中的地表径流响应机制——以深圳河流域多尺度水文模拟为例

### 4. 学科术语规范

- 使用本领域标准术语（如 DEM、NDVI、SWMM 等通用缩写）
- 避免自创缩写（如将“分布式水文模型”简写为 DFHM）

### 5. 问题导向吸引力

- 在严谨性基础上突出科学问题
- 例：气候变化如何影响干旱区内陆河流域的水-生态-经济耦合系统？（对比：某流域水系统研究）

要简明扼要地写一个好题目，首先要明确研究的核心重点。确保题目能够传达以下三个关键元素中的至少一个：**研究主题、研究方法、或者研究结果**。可以使用副标题进一步细化。

### 应该避免的错误

- 冗长和模糊：太长的题目让人难以捉摸关键点，如“关于某些方面的初步研究”这样模糊不清的表达应当尽量避免。

- 专业术语过多：虽然使用行业关键词很重要，但过多复杂术语会吓跑非专业读者。
- 吸引注意但内容无关：题目应与研究内容高度相关，避免吸引眼球的浮夸表达却与实际内容不符。

### 举例

- 主题：城市化对水资源管理的影响
- 好题目：“城市化影响下的区域水资源管理：XX 城市的实证研究”优点：简洁明了，且指出了研究地点和方法，提供了明确的研究背景。
- 坏题目：“我们如何应对城市化导致的水资源问题”缺点：使用问句，范围太宽泛而不具体，没有明确表示研究范围或对象。

## 5.3 摘要 (Abstract)

摘要是论文的缩影，通常是读者首先接触到的部分，

**摘要的作用：**- 吸引读者：好的摘要能够吸引读者深入阅读整篇论文。- 提供背景信息：帮助读者快速了解研究的背景、方法、结果和结论。- 提高可见度：在数据库中，摘要是检索的主要依据之一，好的摘要能够提高论文的可见度。

- 研究背景与目的 (1-2 句)

- 阐明研究的科学问题及其重要性例：“青藏高原冰川消融对亚洲水塔的影响是当前全球变化研究的核心问题”

- 研究方法 (2-3 句)

- 明确技术路线：实验设计/数值模拟/野外观测手段/模拟方法注明关键参数：如“利用 Landsat 8 OLI 数据 (30m 分辨率) 和 MODIS 积雪产品 (500m)”说明时间空间尺度：“基于 2010-2020 年青藏高原 78 个气象站观测数据”

- 核心结果 (3-4 句)

– 定量化呈现主要发现例：“观测到冰川面积年均退缩率  $1.2 \pm 0.3\%$ , 显著高于前十年 ( $0.7 \pm 0.2\%$ )” 突出创新性发现：“首次发现印度季风与西风环流耦合对高原降水相变的调控作用”

- 科学意义 (1-2 句)

– 阐明理论贡献或应用价值例：“本研究为评估第三极地区水循环演变提供了新的观测约束”

### 5.3.1 技术规范

#### 1. 篇幅控制

- 传统期刊：150-400 词（如 Nature 子刊严格限制 150 词）
- 结构式摘要：300 词左右（背景、方法、结果、结论分段）

#### 2. 语体特征

- 使用现在时态陈述结论，过去时态描述方法
- 专业术语首次出现需全称（如 ENSO: El Niño-Southern Oscillation）

#### 3. 信息密度控制

- 关键数据保留 2-3 个核心数值
- 方法学仅保留最具特征性的技术（如“通过 LA-ICP-MS 钻石 U-Pb 定年”）

#### 4. 空间信息标注

- 研究区域应给出地理范围：“羌塘盆地”
- 涉及模式模拟需说明分辨率：“WRF 模式 3km 嵌套网格”

#### 5. 时间尺度表征

- 地质年代使用国际地层年代表（如晚更新世： $126-11.7$  ka BP）
- 现代过程研究需注明观测时段：“2016-2020 年季风期（6-9 月）”

### 5.3.2 自检清单

- 是否包含研究的关键四要素：背景，方法/内容，结果，结论？
- 地理范围/时间尺度是否明确？
- 摘要中无需参考文献
- 术语规范，如避免混淆地质年代单位（如误用“纪/世”）
- 缩略规范，避免非标准缩写（如自创“TPM”代指青藏高原）
- 避免使用缩写和符号，除非是常用的缩写和符号。
- 内容应当与正文一致，避免使用与正文不一致的术语和概念。
- 关键词是否与主题高度匹配？
- 长度是否控制在期刊要求  $\pm 10\%$ ？

### 5.3.3 摘要案例

以下是一些不同领域的研究摘要案例，供参考。

- Kratzert, F., Gauch, M., Klotz, D., and Nearing, G.: HESS Opinions: Never train a Long Short-Term Memory (LSTM) network on a single basin, *Hydrol. Earth Syst. Sci.*, 28, 4187–4201, <https://doi.org/10.5194/hess-28-4187-2024>, 2024.

Machine learning (ML) has played an increasing role in the hydrological sciences. In particular, Long Short-Term Memory (LSTM) networks are popular for rainfall-runoff modeling (研究背景). A large majority of studies that use this type of model do not follow best practices, and there is one mistake in particular that is common: training deep learning models on small, homogeneous data sets, typically data from only a single hydrological basin (研究结论). In this position paper, we show that LSTM rainfall-runoff models are best when trained with data from a large number of basins (研究方法, 结果).

- Sun, H., Yao, T., Su, F., Yang, W., and Chen, D.: Spatiotemporal responses of runoff to climate change in the southern Tibetan Plateau, *Hydrol. Earth Syst. Sci.*, 28, 4361–4381, <https://doi.org/10.5194/hess-28-4361-2024>, 2024.

A comprehensive understanding of spatiotemporal runoff changes in the Yarlung Zangbo (YZ) basin in the southern Tibetan Plateau (TP) at a sub-basin scale (研究区), amidst varying climatic and cryospheric conditions, is imperative for effective water resources management (研究背景). However, spatiotemporal differences of runoff composition and change and their attribution within the YZ basin have not been extensively explored, primarily due to the lack of hydrometeorological observations, especially in the downstream region (研究问题和局限). In this study, we investigated historical and future evolution of annual and seasonal total water availability, as well as glacier runoff and snowmelt contributions across six sub-basins of the YZ, with a particular focus on the comparison between the upstream Nuxia (NX) basin and the downstream Nuxia–Pasighat (NX-BXK) basin, based on a newly generated precipitation dataset and a well-validated model with streamflow, glacier mass, and snow cover observations (研究内容和方法). Our findings revealed that large spatiotemporal differences in changes exist within the YZ basin for 1971–2020. Firstly, runoff generation was dominated by rainfall runoff throughout the YZ basin, with glacier runoff playing a more important role in the annual total runoff (19 %) in the NX-BXK sub-basin compared to other sub-basins. Notably, glacier runoff contributed 52 % of the total runoff at the Pasighat outlet of the YZ basin. Secondly, annual runoff exhibited an increasing trend in the NX basin but a decreasing trend in the NX-BXK, primarily attributed to rainfall runoff changes influenced by atmospheric moisture. Glacier runoff enhanced water supply by offsetting the decreasing contribution from rainfall. Total runoff will consistently increase (27–100 mm (10 yr) $^{-1}$ ) across the sub-basins through the 21st century, resulting from increased rainfall runoff and a minor effect of increased snowmelt and glacier runoff (研究结果与结论).

- Farmani, M. A., Behrangi, A., Gupta, A., Tavakoly, A., Geheran, M., and Niu, G.-Y.: Do land models miss key soil hydrological processes controlling soil moisture memory?,

Hydrol. Earth Syst. Sci., 29, 547–566, <https://doi.org/10.5194/hess-29-547-2025>, 2025.

Soil moisture memory (SMM), which refers to how long a perturbation in soil moisture (SM) can last, is critical for understanding climatic, hydrological, and ecosystem interactions. Most land surface models (LSMs) tend to overestimate surface soil moisture and its persistency (or SMM), sustaining spuriously large soil surface evaporation during dry-down periods (研究背景). We **attempt to answer a question**: do LSMs miss or misrepresent key hydrological processes controlling SMM? (科学问题) We use a version of Noah-MP with advanced hydrology that explicitly represents preferential flow and surface ponding and provides optional schemes of soil hydraulics. We test the effects of these processes, which are generally missed by most LSMs in SMM. We compare SMMs computed from various Noah-MP configurations against that derived from the Soil Moisture Active Passive (SMAP) L3 soil moisture and in situ measurements from the International Soil Moisture Network (ISMN) from the years 2015 to 2019 over the contiguous United States (CONUS) (研究方法). The **results suggest that** (1) soil hydraulics plays a dominant role and the Van Genuchten hydraulic scheme reduces the overestimation of the long-term surface SMM produced by the Brooks–Corey scheme, which is commonly used in LSMs; (2) explicitly representing surface ponding enhances SMM for both the surface layer and the root zone; and (3) representing preferential flow improves the overall representation of soil moisture dynamics. The combination of these missing schemes can significantly improve the long-term memory overestimation and short-term memory underestimation issues in LSMs (研究结果及发现). We suggest that LSMs for use in seasonal-to-subseasonal climate prediction should, at least, adopt the Van Genuchten hydraulic scheme (研究结论).

- Sadri, S., Famiglietti, J. S., Pan, M., Beck, H. E., Berg, A., and Wood, E. F.: FarmCan: a physical, statistical, and machine learning model to forecast crop water deficit for farms, Hydrol. Earth Syst. Sci., 26, 5373–5390, <https://doi.org/10.5194/hess-26-5373-2022>, 2022.

**5194/hess-26-5373-2022, 2022.**

In the coming decades, a changing climate, the loss of high-quality land, the slowing in the annual yield of cereals, and increasing fertilizer use indicate that better agricultural water management strategies are needed (研究背景) . In this study, we designed FarmCan, a novel, robust remote sensing and machine learning (ML) framework to forecast farms' needed daily crop water quantity or needed irrigation (NI) (研究内容) . We used a diverse set of simulated and observed near-real-time (NRT) remote sensing data coupled with a random forest (RF) algorithm and inputs about farm-specific situations to predict the amount and timing of evapotranspiration (ET), potential ET (PET), soil moisture (SM), and root zone soil moisture (RZSM) (研究方法) . Our case study of four farms in the Canadian Prairies Ecozone (CPE) shows that 8 d composite precipitation ( $P$ ) has the highest correlation with changes ( $\Delta$ ) of RZSM and SM. In contrast, 8 d PET and 8 d ET do not offer a strong correlation with 8 d  $P$ . Using  $R^2$ , root mean square error (RMSE), and Kling-Gupta efficiency (KGE) indicators, our algorithm could reasonably calculate daily NI up to 14 d in advance. From 2015 to 2020, the  $R^2$  values between predicted and observed 8 d ET and 8 d PET were the highest (80 % and 54 %, respectively). The 8 d NI also had an average  $R^2$  of 68%. The KGE of the 8 d ET and 8 d PET in four study farms showed an average of 0.71 and 0.50, respectively, with an average KGE of 0.62 (研究结果) . FarmCan can be used in any region of the world to help stakeholders make decisions during prolonged periods of drought or waterlogged conditions, schedule cropping and fertilization, and address local government policy concerns (研究结论) .

- **Shen, H., Tolson, B. A., & Mai, J. (2022). Time to update the split-sample approach in hydrological model calibration. Water Resources Research, 58, e2021WR031523. <https://doi.org/10.1029/2021WR031523>**

Model calibration and validation are critical in hydrological model robustness assessment. Unfortunately, the commonly used split-sample

test (SST) framework for data splitting requires modelers to make subjective decisions without clear guidelines (研究背景及问题). This large-sample SST assessment study empirically assesses how different data splitting methods influence post-validation model testing period performance, thereby identifying optimal data splitting methods under different conditions (研究内容). This study investigates the performance of two lumped conceptual hydrological models calibrated and tested in 463 catchments across the United States using 50 different data splitting schemes. These schemes are established regarding the data availability, length and data recentness of continuous calibration sub-periods (CSPs). A full-period CSP is also included in the experiment, which skips model validation. The assessment approach is novel in multiple ways including how model building decisions are framed as a decision tree problem and viewing the model building process as a formal testing period classification problem, aiming to accurately predict model success/failure in the testing period (研究方法). Results span different climate and catchment conditions across a 35-year period with available data, making conclusions quite generalizable. Calibrating to older data and then validating models on newer data produces inferior model testing period performance in every single analysis conducted and should be avoided. Calibrating to the full available data and skipping model validation entirely is the most robust split-sample decision. Experimental findings remain consistent no matter how model building factors (i.e., catchments, model types, data availability, and testing periods) are varied (研究结果). Results strongly support revising the traditional split-sample approach in hydrological modeling (研究结论).

- Kunadi, A. S., Lardner, T., Silberstein, R. P., Leopold, M., Callow, N., Veneklaas, E., et al. (2024). Introducing pour points: Characteristics and hydrological significance of a rainfall-concentrating mechanism in a water-limited woodland ecosystem. Water Resources Research, 60, e2023WR035458. <https://doi.org/10.1029/2023WR035458>

The interception of rainfall by plant canopies alters the depth and spatial distribution of water arriving at the soil surface, and thus the location, volume, and depth of infiltration (研究背景). Mechanisms like stemflow are known to concentrate rainfall and route it deep into the soil, yet other mechanisms of flow concentration are poorly understood (科学问题). This study characterizes pour points, formed by the detachment of water flowing under a branch, using a combination of field observations in Western Australian banksia woodlands and rainfall simulation experiments on Banksia menziesii branches. We aim to establish the hydrological significance of pour points in a water-limited woodland ecosystem, along with the features of the canopy structure and rainfall that influence pour point formation and fluxes. Pour points were common in the woodland and could be identified by visually inspecting trees. Throughfall depths at pour points were up to 15 times greater than rainfall and generally comparable to or greater than stemflow. Soil water content beneath pour points was greater than in adjacent controls, with 20%–30% of the seasonal rainfall volume infiltrated into the top 1 m of soil beneath pour points, compared to 5% in controls (研究方法). Rainfall simulations showed that pour points amplified the spatial heterogeneity of throughfall, violating assumptions used to close the water balance (研究结果). The simulation experiments demonstrated that pour point fluxes depend on the interaction of branch angle and foliation for a given branch architecture. Pour points can play a significant part in the water balance, depending on their density and rainfall concentration ability (研究结论).

- Wi, S., & Steinschneider, S. (2022). Assessing the physical realism of deep learning hydrologic model projections under climate change. *Water Resources Research*, 58, e2022WR032123. <https://doi.org/10.1029/2022WR032123>

This study examines whether deep learning models can produce reliable future projections of streamflow under warming (科学问题). We train a regional long short-term memory network (LSTM) to daily

streamflow in 15 watersheds in California and develop three process models (HYMOD, SAC-SMA, and VIC) as benchmarks. We force all models with scenarios of warming and assess their hydrologic response, including shifts in the hydrograph and total runoff ratio (研究方法). All process models show a shift to more winter runoff, reduced summer runoff, and a decline in the runoff ratio due to increased evapotranspiration. The LSTM predicts similar hydrograph shifts but in some watersheds predicts an unrealistic increase in the runoff ratio (研究结果). We then test two alternative versions of the LSTM in which process model outputs are used as either additional training targets (i.e., multi-output LSTM) or input features (研究方法). Results indicate that the multi-output LSTM does not correct the unrealistic streamflow projections under warming. The hybrid LSTM using estimates of evapotranspiration from SAC-SMA as an additional input feature produces more realistic streamflow projections, but this does not hold for VIC or HYMOD (研究结果). This suggests that the hybrid method depends on the fidelity of the process model (研究结论). Finally, we test climate change responses under an LSTM trained to over 500 watersheds across the United States and find more realistic streamflow projections under warming (研究方法). Ultimately, this work suggests that hybrid modeling may support the use of LSTMs for hydrologic projections under climate change, but so may training LSTMs to a large, diverse set of watersheds (研究结论).

- Lee, R. M., McGuire, K. J., Strahm, B. D., Knoepp, J. D., Jackson, C. R., & Stewart, R. D. (2020). Revisiting the Hewlett and Hibbert (1963) Hillslope Drainage Experiment and Modeling Effects of Decadal Pedogenic Processes and Leaky Soil Boundary Conditions. *Water Resources Research*, 56, e2019WR025090. <https://doi.org/10.1029/2019WR025090>

Subsurface flow dominates water movement from hillslopes to streams in most forested headwater catchments. Hewlett and Hibbert (1963, <https://doi.org/10.1029/JZ068i004p01081>) constructed an idealized

hillslope model ( $0.91 \times 0.91 \times 15.0$  m;  $21.8^\circ$ ) using reconstituted C horizon soil to investigate importance of interflow, a type of subsurface flow. They saturated the model, covered it to prevent evaporation, and allowed free drainage for 145 days. The resulting recession drainage curve suggested two phases: fast drainage of saturated soil in the first 1.5 days and then slow drainage of unsaturated soil. Hydrologists interpreted the latter as evidence interflow could sustain baseflow, even during extended drought. Since that experiment, typical forest vegetation grew in the model, providing root and litter inputs for 55 years (研究背景). We removed all aboveground live biomass and repeated the experiment physically and numerically (HYDRUS-2D) (研究方法), hypothesizing that pedogenesis would change the drainage curve and further elucidate the role of unsaturated flow from hillslopes (科学零假设). Contrary to this hypothesis, drainage curves in our twice-repeated physical experiments and numerical simulation were unchanged for the first  $\sim 10$  days (结果), indicating pedogenesis and biological processes had not largely altered bulk hydraulic conductivities or soil moisture release characteristics (研究结论). However, drainage unexpectedly ceased after about 2 weeks ( $14.3 \pm 2.5$  days), an order of magnitude sooner than in the original experiment, due to an apparent leak in the hillslope analogous to commonly observed bedrock fractures in natural systems (结果). Thus, our results are a more natural recession behavior that highlight how incorporation of alternative hydrologic outputs can reduce drainage duration and volume from soils to baseflow (研究结论).

- **Illuminating water cycle modifications and Earth system resilience in the Anthropocene**

Fresh water—the bloodstream of the biosphere—is at the center of the planetary drama of the Anthropocene. Water fluxes and stores regulate the Earth's climate and are essential for thriving aquatic and terrestrial ecosystems, as well as water, food, and energy security. But the water cycle is also being modified by humans at an unprecedented

scale and rate. A holistic understanding of freshwater's role for Earth system resilience and the detection and monitoring of anthropogenic water cycle modifications across scales is urgent, yet existing methods and frameworks are not well suited for this (研究背景) . In this paper we highlight four core Earth system functions of water (hydroclimatic regulation, hydroecological regulation, storage, and transport) and key related processes. Building on systems and resilience theory, we review the evidence of regional-scale regime shifts and disruptions of the Earth system functions of water. We then propose a framework for detecting, monitoring, and establishing safe limits to water cycle modifications and identify four possible spatially explicit methods for their quantification (研究方法) . In sum, this paper presents an ambitious scientific and policy grand challenge that could substantially improve our understanding of the role of water in the Earth system and cross-scale management of water cycle modifications that would be a complementary approach to existing water management tools (研究结论) .

- **MERIT Hydro: A High-Resolution Global Hydrography Map Based on Latest Topography Dataset Dai Yamazaki, Daiki Ikeshima, Jeison Sosa, Paul D. Bates, George H. Allen, Tamlin M. Pavelsky**

High-resolution raster hydrography maps are a fundamental data source for many geoscience applications (研究背景) . Here we introduce MERIT Hydro, a new global flow direction map at 3-arc sec resolution (~90 m at the equator) derived from the latest elevation data (MERIT DEM) and water body data sets (G1WBM, Global Surface Water Occurrence, and OpenStreetMap) (研究内容) . We developed a new algorithm to extract river networks near automatically by separating actual inland basins from dummy depressions caused by the errors in input elevation data. After a minimum amount of hand editing, the constructed hydrography map shows good agreement with existing quality-controlled river network data sets in terms of flow accumulation area and river basin

shape. The location of river streamlines was realistically aligned with existing satellite-based global river channel data. Relative error in the drainage area was <0.05 for 90% of Global Runoff Data Center (GRDC) gauges, confirming the accuracy of the delineated global river networks. Discrepancies in flow accumulation area were found mostly in arid river basins containing depressions that are occasionally connected at high water levels and thus resulting in uncertain watershed boundaries (研究结果). MERIT Hydro improves on existing global hydrography data sets in terms of spatial coverage (between N90 and S60) and representation of small streams, mainly due to increased availability of high-quality baseline geospatial data sets. The new flow direction and flow accumulation maps, along with accompanying supplementary layers on hydrologically adjusted elevation and channel width, will advance geoscience studies related to river hydrology at both global and local scales (研究意义).

- Herrera-Estrada, J. E., Diffenbaugh, N. S. (2020). Land-falling droughts: Global tracking of moisture deficits from the oceans onto land. *Water Resources Research*, 56, e2019WR026877. <https://doi.org/10.1029/2019WR026877>

Droughts threaten food, energy, and water security, causing death and displacement of millions of people and billions of dollars in damages. However, there are still important gaps in the understanding of drought mechanisms and behaviors, inhibiting the accuracy of early-warning systems designed to protect communities worldwide (研究背景). We use an object-tracking algorithm to track clusters of precipitation-minus-evaporation moisture deficits across land and ocean areas of the globe from 1981–2018 (研究方法). This analysis reveals a new type of “landfalling drought” that originates over the ocean and “migrates” onto land. We find that 16% of droughts that affected the continents worldwide from 1981–2018 were landfalling droughts. These droughts were significantly larger (220–425%) and more intense (4–30%)—and grew (253–285%) and intensified (9–28%) faster—than droughts that developed solely over the land or ocean (研究结果). To

identify potential underlying mechanisms, we analyze moisture transport associated with landfalling droughts over western North America. We find that landfalling droughts in this region are associated with anomalously anticyclonic atmospheric pressure patterns that reduce moisture fluxes over the Pacific Ocean toward the continent (研究结论). By advancing understanding of the spatiotemporal evolution of droughts, our findings offer the potential to improve seasonal-scale prediction and long-term projection of global drought risks (研究价值).

- **Global Investigation of Soil Moisture and Latent Heat Flux Coupling Strength**

As a key variable in the climate system, soil moisture (SM) plays a central role in the Earth's terrestrial water, energy, and biogeochemical cycles through its coupling with surface latent heat flux (LH). Despite the need to accurately represent SM/LH coupling in Earth system models, we currently lack quantitative, observation-based, and unbiased estimates of its strength (研究背景). Here we utilize the triple-collocation (TC) approach introduced in Crow et al. (2015) to SM and LH products obtained from multiple satellite remote sensing platforms and land surface models (LSMs) to obtain unbiased global maps of SM/LH coupling strength (研究方法). Results demonstrate that relative to coupling strength estimates acquired directly from remote sensing-based data sets, the application of TC generally enhances estimates of warm-season SM/LH coupling, especially in the western United States, the Sahel, central Asia, and Australia. However, relative to triple-collocation estimates, LSMs (still) overpredict SM/LH coupling strength along transitional climate regimes between wet and dry climates, such as the central Great Plains of North America, India, and coastal Australia. Specific climate zones with biased relations in LSMs are identified to geographically focus the reexamination of LSM parameterizations (研究结果). TC-based coupling strength estimates are robust to our choice of LSM contributing SM and LH products to the TC analysis. Given their robustness, TC-based coupling strength

estimates can serve as an objective benchmark for investigating model-predicted SM/LH coupling (研究结论) .

- **If Precipitation Extremes Are Increasing, Why Aren't Floods?**

Despite evidence of increasing precipitation extremes, corresponding evidence for increases in flooding remains elusive. If anything, flood magnitudes are decreasing despite widespread claims by the climate community that if precipitation extremes increase, floods must also (研究背景与问题) . In this commentary we suggest reasons why increases in extreme rainfall are not resulting in corresponding increases in flooding. Among the possible mechanisms responsible, we identify decreases in antecedent soil moisture, decreasing storm extent, and decreases in snowmelt. We argue that understanding the link between changes in precipitation and changes in flooding is a grand challenge for the hydrologic community and is deserving of increased attention (研究结论) .

## 5.4 方法、研究区、数据 (Methodology)

### 5.4.1 方法 (Methodology)

### 5.4.2 研究区 (Study Area)

### 5.4.3 数据 (Data)

## 5.5 结果 (Results)

结果与讨论是研究的核心部分，应当是对研究的结果和发现的总结和分析，以及对研究的意义和价值的阐述。

## 5.6 讨论 (Discussion)

讨论与展望应当是对研究的不足和局限性的总结, 对研究的未来发展方向和可能性的展望。

## 5.7 结论 (Conclusion)

## 5.8 其他论文附属内容

### 5.8.1 关键词

### 5.8.2 附录

### 5.8.3 补充材料

### 5.8.4 致谢

### 5.8.5 利益冲突声明

### 5.8.6 数据/软件可用性声明

### 5.8.7 参考文献

- 参考文献选取
  - 参考文献应当是与研究相关的、最新的、权威的、具有代表性的。
- 参考文献格式
  - 参考文献应当按照期刊的要求进行格式化。
  - 参考文献应当按照作者姓氏的字母顺序进行排列。
- 参考文献的引用
  - 参考文献应当在正文中进行引用。
  - 参考文献的引用应当按照期刊的要求进行格式化。
  - 参考文献的引用应当在正文中进行标注。

## 5.9 论文投稿

### 5.9.1 期刊选择

### 5.9.2 投稿材料

- Cover Letter
- Plain Abstract

### 5.9.3 投稿流程

### 5.9.4 投稿后续

### 5.9.5 论文修改

### 5.9.6 论文撤回



# 第六章 技能训练

## 6.1 电子邮件

电子邮件是科研工作中最重要的沟通工具之一，电子邮件写作是基本技能之一。

题目和收发人：

- 主题应该简洁明了，能够清楚地传达邮件的目的。
- 发信人名称为人名，禁止使用网名作为邮件的发信人姓名。例如：张晓明 <Zhangxm@qq.com>，或者 Zhang Xiaomin<Zhangxm@qq.com>，而不能用胖死在减肥的道路上 <Zhangxm@qq.com> 一类的。
- 收件人应该是邮件的主要接收者，避免使用过多的收件人。
- 抄送人应该是需要知晓邮件内容的人，避免使用不必要的抄送人。
- 邮件的抄送和密送应该谨慎使用，避免不必要的抄送和密送。
- 邮件的签名应该包含发件人的姓名、职务、单位和联系方式，避免使用过长的签名。注意中英文的格式上的差异。举例：

– 英文

Xiaoming Zhang  
Phd Candidate  
College of Resources and Environment, Lanzhou University  
Lanzhou, Gansu, China

### - 中文

张晓明  
博士研究生  
兰州大学，资源环境学院

#### 正文：

- 邮件的写作风格和语气应该是正式的。
- 正文应该简洁明了，避免使用复杂的句子和术语。
- 结尾应该礼貌地结束，感谢收件人的时间和关注。
- 邮件的格式应该规范，避免使用过多的颜色和字体，保持简洁明了。
- 邮件文字应该经过仔细检查，避免出现拼写和语法错误。
- 邮件的发送时间应该选择在工作日的工作时间内，避免在节假日或周末发送邮件。
- 邮件回复应该及时，避免长时间不回复。

#### 附件要求：

- 当带有附件时，应该在邮件中清楚地提及。
- 附件应该使用常见的文件格式，如 PDF、Word、Excel 等，避免使用不常见的文件格式。
- 附件应该清晰地标明文件名和内容，避免使用模糊的文件名。
- 附件的大小应该控制在合理范围内，避免发送过大的附件。
- 附件的内容应该清晰明了，避免使用复杂的格式和排版。

## 6.2 ABT 句式

ABT 句子结构是：..... and ...., but ...., therefore .....。

中文句式为：....., 并且....., 但是....., 因此.....。

#### 举例

- 山洪研究

- 气候变化导致山洪等灾害越来越严重,
- 同时山洪带来的生命财产损失也越来越受关注,
- 但是当前山洪灾害预警的时间和空间准确性远远不能满足人民和社会的需求,
- 因此我们的目标是提高山洪的预见期、预警时间精度和空间位置准确性。

- 模型开发:

- 水文模型能够解析气候变化对流域水文和湖泊的贡献,
- 同时古气候研究需要模型来解释湖泊水位的历史变化,
- 但是古气候数据无法满足现有水文模型对日/小时尺度的气象驱动数据的需求,
- 因此开发万年尺度、体现冰川、积雪、植被影响下的流域水文和湖泊动态的水文模型非常必要。

## 6.3 作图与解释

### 6.3.1 DTM 作图方法

Discover-Transform-Method 作图方法

## 6.4 精炼书写原则

\* 以下内容参考源: Utah Valley University Writing Center ([www.uvu.edu/writingcenter](http://www.uvu.edu/writingcenter))

Writing with concision means intentionally using language and sentence structure to improve clarity and cut excess or unnecessary elements. Listed below are strategies you can use to make your writing more concise. Remember to always work with your audience and assignment in mind.

- **Limit Modifying Phrases** (减少修饰短语)

Eliminate phrases that modify the noun if they do not clarify the meaning of the sentence or if they include unnecessary information.

- Wordy: The dishwasher **in the kitchen** is broken.
- **Concise:** The dishwasher is broken.

- **Minimize Prepositional Phrases** (减少介词从句)

Avoid unnecessary prepositional phrases (preposition plus a noun).

- Wordy: The **assignment on writing about discourse communities** is due on Friday.
- **Concise:** The **discourse community writing assignment** is due by Friday.

- **Avoid Unnecessary Relative Clauses** (避免不必要的从句)

Relative clauses beginning with that, who, and which should be changed to short phrases, or deleted.

- Wordy: The candidate **who ran for president** was likeable.
- **Concise:** The **presidential candidate** was likeable.

- **Reword Infinitive Phrases** (少用不及物动词)

Eliminate infinitive phrases (to plus a verb) by using another form of the verb.

- Wordy: The shortage of essential items has **caused customers to become dissatisfied**.
- **Concise:** The shortage of essential items has **dissatisfied customers**.

- **Modify Other Phrases** (能用一个字, 决不用两个字)

Change wordy phrases into single words or direct expressions when possible.

- Wordy: At **this point in time**, add the eggs to the mixture.
- **Concise:** Now, add the eggs to the mixture.

- **Use Active Voice** (使用主动语态)

Avoid passive-voice phrases by changing them to active voice.

- Wordy: An account was created by Emma Garcia last week.
- Concise: Emma Garcia created an account last week.

- **Eliminate Extra or Empty Words** (删除空洞词, 用精确词)

Use one word that captures the main idea instead of multiple words that approximate it.

- Wordy: The view was pretty, nice, and interesting.
- Concise: The view was breathtaking.

- **Avoid Inflated Writing** (避免浮夸)

Use common and precise language whenever possible. Using overly academic language to impress readers often makes writing sound inflated.

- Wordy: My sanguineous professor disesteemed my election of “grandiloquent.”
- Concise: My wise professor disapproved of my choice of big words.

- **Avoid Nominalization** (避免“动词名词化”)

Nominalization occurs when a verb functions as a noun. Avoid this by using one active verb instead.

- Wordy: The employee made the decision to quit their job.
- Concise: The employee decided to quit their job.

- **Omit Redundant Word Pairs** (避免重复含义)

When using word pairs, look to see if both words are necessary. (如果关键词已经包含修饰词的含义, 则直接删除修饰词)

- Wordy: This gift is an unexpected surprise.
- Concise: This gift is a surprise. (surprise 已经包含了 unexpected——意想不到的含义。)

## 6.5 格式/形式自查清单

### 6.5.1 PPT

- 图件规范检查：可视化表达的核心要素
  1. 每张图是否有图题 (Caption)?
  2. 是否标注变量单位 (单位范围合理)?
  3. 非常见简写是否首次出现时解释? (如: NGS= 非生长季)
  4. 组合图是否可合并为一张图? (减少页面碎片化)
  5. 色盲测试是否通过? <https://www.color-blindness.com/color-blindness-simulator/>
  6. 颜色映射是否符合直觉? (例: 红 → 高温/干旱, 蓝 → 低温/湿润)
  7. X/Y 轴名称与单位是否清晰?
  8. 坐标比例是否合理? (避免数据失真)
  9. 多 Y 轴时是否明确区分左右对应关系? 读者能否一眼读懂数据对应左右 Y 轴的对应关系?
- 引用规范检查：学术诚信与格式统一
  1. 他人图表是否标注来源?
  2. 引用是否规范? 建议格式为: 人名 + 年份 + 期刊 (例: Duffy et al., 2016, WRR)
- 视觉设计规范：专业性与美观性平衡
  1. 定制属于自己的模板 (避免 PowerPoint 默认模板), 针对报告场合和听众使用不同的模版。
  2. 主色调 3 种 (优先使用单位 LOGO 色系)

3. 禁用大面积背景色（推荐纯白/浅灰底）
  4. 避免大红色字体
  5. 强调字体加粗或高亮
  6. 非必要，不使用 PPT 动画。
  7. 字体是否统一?
    - 标题：微软雅黑/Arial/Helvetica
    - 正文：微软雅黑/Arial/Helvetica/Times New Roman；宋体，可用但少用。
    - 特殊字体：仿宋/楷体，可用但少用。
  8. 文字断行是否合理？避免单字成行。
  9. 页面是否保留足够留白？（上下左右边距 1cm）
  10. 封面是否包含：
    - 单位 LOGO（完整版）
    - 主旨图（强化主题记忆）
    - 作者姓名/单位/邮箱
    - 会议名称/时间/地点
  11. 目录页是否清晰展示逻辑框架？
  12. 章节过渡页是否有视觉标识？（例：色块/图标分隔）
  13. 是否添加页码？
- 内容逻辑检查：学术表达的严谨性
    1. 章节顺序是否符合研究叙事？
    2. 是否存在“前繁后简”的页面失衡？“前一页堆满图文，后一页空空如也”？
    3. 关键内容或者关键词加粗突出显示
  - 提交前终检：细节决定专业度

1. 听众致谢语句
2. 资助机构 LOGO 及编号（如：NSFC-XXXX）
3. 每页角落是否有单位小 LOGO

- 讲解及解说词

1. 是否有解说词？（避免“读 PPT”）
2. 是否有时间限制？（避免“超时”）
3. 汇报之前，必须反复练习，熟悉每一页的内容。
4. 禁止拿文稿朗读。需脱稿讲解。
5. 讲的时候，眼睛要看着听众，而不是看着 PPT。
6. 语速适中，语调有起伏。
7. 发音清晰，声音自信而洪亮。

- 问答环节

1. 注意听清楚问题，回答时要简洁明了。
2. 遇到没听清的问题，可以请对方重复或换个问法，不能胡诌。
3. 回答一开始，感谢对方的提问，肯定对方的提了一个“重要/好/关键/专业”问题。
4. 注意礼仪，保持微笑，眼神交流。
5. 遇到不懂的问题，可以说“我不太了解这个问题”或者“我没有研究过这个问题”。
6. 如果对方“咄咄逼人”，保持礼貌，尽量做到就事论事，针对科学的研究/问题本身谈看法。
7. 不同观点时，不要试图说服对方，而是让对方理解你的观点，同时对别人的观点表示尊重，表示未来会继续探索该问题。如果对方的提问是“无理取闹”，可以说“谢谢你的提问，我会考虑这个问题”或者“谢谢你的提问，我会继续研究这个问题”。

### 6.5.2 Word 文档

1. 图题（caption）是否合理？图题要能够清楚解释图件里面的含义——读者即使不阅读正文，可以借助图本身、图例和图题读懂图件表

达的含义。

2. 章节标题的层级关系是否得当? 是否可以通过 Word/WPS 的自动提纲功能提取文章的目录结构?
3. 是否在合适的地方加入引用? 一般引用规则: 重要的数据、重要论述、重要的因果关系、引自他人的观点/数据/结果等。
4. 引用格式是否全文一致?
5. 图件中的信息与正文阐述是否一致? 例如: 图件中包含研究区的经纬度、高程范围、降水量、气温等, 该数值是否与正文中对应的数值一致?
6. 前后文逻辑、观点是否一致?
7. 标点符号是否中英文混用?
8. 文字行距是否一致?
9. 文字格式推荐使用模版。

## 6.6 个人简历 (CV)

个人简历是求职的第一步, 简历的质量直接影响到面试的机会。

- 内容简洁明了, 突出个人的优势和特点。
- 内容包含: 个人信息、教育背景、工作经历、学术成果、技能特长等。
- 个人成果中论文按照论文引用格式。
- 成果使用数字编号。
- 避免使用过多的颜色和字体, 保持简洁明了。
- 排版整齐, 避免错位和不对齐。
- 易读的字体, 避免花哨字体。
- 行距适中、段落适当, 字号适中
- 拼写/语法检查
- 学生简历不超过两页, 避免过长。





# 第七章 科研工具

推荐使用的科研软件请先查看下面这张图。绝大部分为免费软件，个别是收



费软件。

Presented with xmind

## 7.1 空间制图

### 7.1.1 QGIS

QGIS 是一款免费、开源的地理信息系统（GIS）软件，相对于其他商业 GIS 软件，它有以下几个优势：

- **免费：**QGIS 是一款完全免费的 GIS 软件，用户可以自由地下载、使用和修改它。
- **开源：**QGIS 是一个开源软件，这意味着用户可以随意查看其源代码、修改和扩展其功能，同时也可以通过用户社区获得技术支持。
- **跨平台：**QGIS 可以在 Windows、MacOS、Linux 等多个操作系统上运行，并提供相应的安装程序和二进制文件。
- **多功能：**QGIS 具有很多功能，包括地图制作、数据处理、空间分析、地理编码和地理数据编辑等。它支持许多不同的文件格式，包括 ESRI shapefile、MapInfo 文件、PostGIS 和 Oracle 空间数据库、GeoTIFF 和其他栅格数据格式等。
- **易于使用：**QGIS 提供了易于使用的用户界面，包括地图绘制、数据导入和空间分析等功能。此外，用户可以通过插件来扩展软件的功能，满足不同的需求。
- **社区支持：**QGIS 有一个活跃的用户社区，用户可以通过该社区获得技术支持、交流使用经验和分享资源等。

综上所述，QGIS 是一款免费、开源、跨平台、功能丰富、易于使用且有着强大的社区支持的 GIS 软件。它是许多研究人员、学生和行业专业人士的首选工具之一。

下载地址：<https://www.qgis.org/en/site/forusers/download.html>

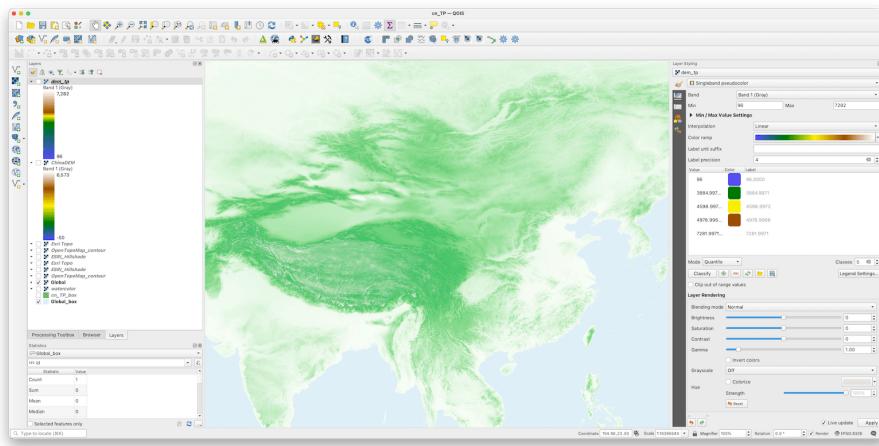


图 7.1: QGIS 界面

## 7.2 绘图

### 7.2.1 XMind

XMind 是一款流行的思维导图软件，它帮助用户通过可视化的方式进行思维整理和知识管理。支持多种导图类型，界面友好，操作简单，且具有丰富的样式和主题，用户可以根据需要选择适合自己的导图方式和风格。此外，XMind 还支持多种文件格式，如 PDF、Excel、Word 等，方便用户进行跨平台工作和分享。

下载地址：<https://xmind.cn/download/>

## 7.3 文献

### 7.3.1 Mendeley Desktop

Mendeley Desktop 是一款强大的文献管理软件，它帮助用户组织、整理和分享学术文献。其优势包括：支持文献的快速检索、全文搜索，以及跨平台同步；提供丰富的文献管理功能，如文献分类、标签管理、引用格式自动生成等；支持文献协作和共享，可以与同行合作整理文献，共同撰写文章。

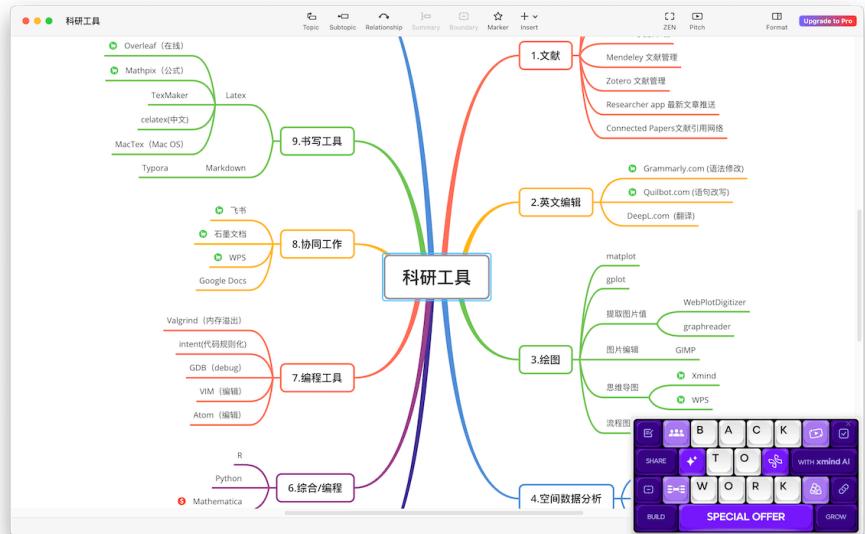


图 7.2: xmind

下载地址: <https://www.mendeley.com/download-reference-manager>

### 7.3.2 Consensus.app

Consensus.app 是一个利用人工智能技术驱动的科研搜索引擎，它的核心价值在于帮助研究人员高效地查找、总结和理解大量的科学研究论文。Consensus.app 能够针对用户提出的问题，综合多篇研究论文的内容，提供一个基于研究证据的共识性回答，这对于科研工作者在进行文献综述、寻找支持自己研究假设的证据时非常有用。此外，Consensus.app 还提供了一个便捷的方式来引用相关论文，确保了答案的真实性和可靠性，从而在学术研究中发挥了重要作用。

访问地址: <https://Consensus.app>

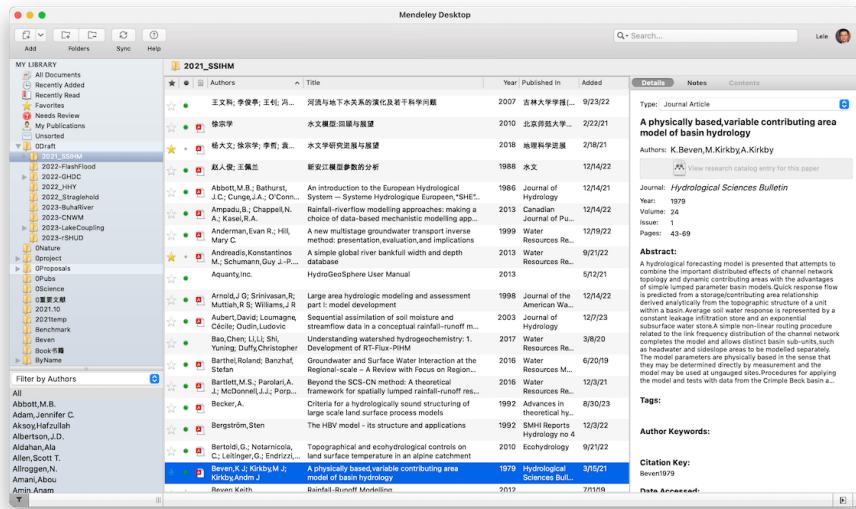


图 7.3: Mendeley desktop

### 7.3.3 Researcher App

Researcher App 是一个移动应用程序，旨在为研究人员提供一个便捷的方式来访问和管理科学文献。它允许用户快速搜索学术文章，阅读摘要，管理自己的引用和阅读列表，以及同步图书馆资料。Researcher App 的优势在于其用户友好的界面和实用的功能，如离线阅读、文献分享和跨平台同步，这些功能都有助于提高研究人员的工作效率。

访问地址：<https://www.researcher-app.com>

### 7.3.4 Publish or Perish (PoP)

Publish or Perish (PoP) 是一款广受学术界欢迎的科研评价工具，由 Google 工程师 Timothy original 所开发。该软件的宗旨是帮助科研工作者更有效地管理和评估他们的学术成果。Publish or Perish 的主要功能包括：

1. 学术成果分析：它可以分析学者在 Google Scholar 上的引用情况，提供关于其研究成果的即时反馈。

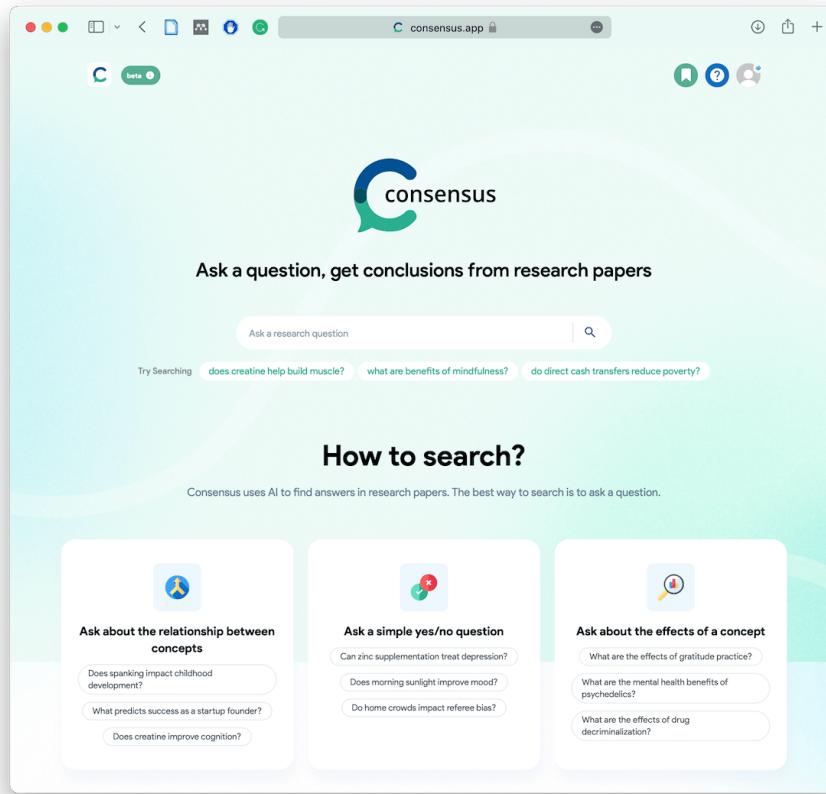


图 7.4: Consensus.app

2. 引用排名: PoP 能够显示个人或机构的引用排名, 以及他们在特定学科领域的表现。
3. 趋势分析: 通过分析引用数据, PoP 可以帮助用户了解特定研究主题或领域的趋势和动向。
4. 成果比较: 用户可以比较不同研究者或论文的引用情况, 从而对学术成就进行量化评估。尽管 Publish or Perish 是一个免费软件, 但用户需要意识到, 它所依赖的数据源——Google Scholar——可能会受到搜索引擎算法更新的影响, 从而影响评价结果的准确性和时效性。因此, 虽然 PoP 是一个有用的工具, 但它应该与其他评价方法结合使用, 以获得更全面的学术评价。

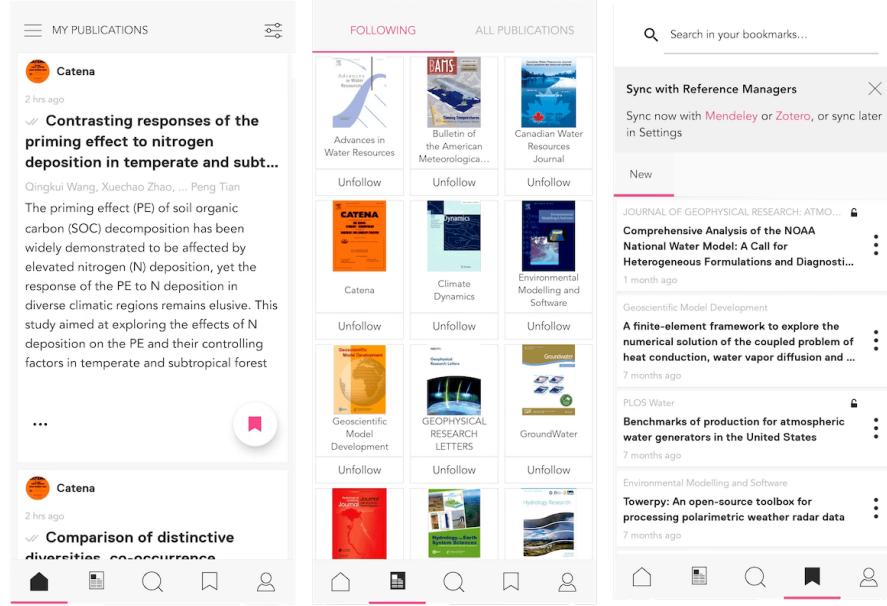


图 7.5: researcherapp 手机截屏

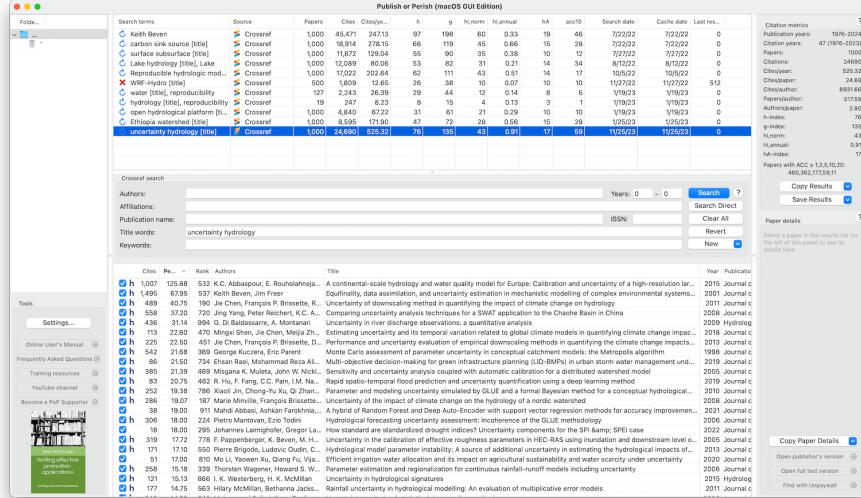


图 7.6: pop

## 7.4 AI 大语言模型

### 7.4.1 AI 及 ChatGPT

当从事科学研究时，使用 AI 或大语言模型（LLM）可以作为辅助工具提供帮助和灵感。但以下是一些建议：

- **准备好问题：**确保你有清晰的问题或概念，以便从 LLM 中获得有用的回答。明确你想要什么样的帮助，这样 LLM 才能更好地回应你的需求。
- **\*\* 验证信息：**LLM 提供的答案可能不总是准确或可靠的，所以记得验证这些信息。进行进一步的研究和检查，确保你得到的结果是可信的。
- **适度使用：**LLM 是一个很有用的工具，但也有一些局限。明确 LLM 的作用，将其作为辅助手段，而不是替代你自己的研究工作。
- **避免偏见：**LLM 模型的训练数据是从互联网上收集的，这可能包含了偏见或不准确的信息。所以在使用 LLM 的结果时，要保持谨慎，注意避免偏见。
- **不断训练改进：**利用 LLM 与其他研究人员进行交流，共享反馈和经验，这样可以帮助改进和训练 LLM，提高它在科研任务中的效果。

需要注意的是，尽管 LLM 可以为科学研究提供一些有价值的帮助，但它仍然是一个语言模型，并不能替代严格的实验设计和科学方法。将 LLM 作为辅助工具，与自己的专业知识和判断相结合，才能更好地实现科研任务的目标。

### 7.4.2 ChatGPT

<http://chat.openai.com>

### 7.4.3 基于 API 的私有部署

学术 GPT 部署：[https://github.com/binary-husky/gpt\\_academic](https://github.com/binary-husky/gpt_academic)

#### 7.4.4 私有部署

- ollama: <https://ollama.ai>。Ollama 是一个开源的大型语言模型服务工具，它提供了类似 OpenAI 的 API 接口和聊天界面，使得用户能够方便地部署和使用最新版本的 GPT 模型
- llama2 <https://github.com/facebookresearch/llama> Llama2 是由 Meta 公司开发的开源大语言模型，它在多个基准测试上取得了超越现有开源模型的成绩，具有优秀的多轮对话能力和安全性。
- ChatGLM3 <https://github.com/THUDM/ChatGLM3>。ChatGLM3 是由智谱 AI 公司和清华大学 KEG 实验室联合发布的新一代对话预训练模型。它是基于 GLM-3 模型开发的，具有出色的多轮对话能力和良好的上下文理解能力。ChatGLM3 模型在多种场景下都表现出了优秀的性能，例如客服、教育、娱乐等。它还支持多种语言的交互，能够满足不同用户的需求。
- 通义千问 (Qwen) <https://github.com/QwenLM/Qwen>。通义千问 (Qwen) 是阿里巴巴开源的大型语言模型，旨在实现通用人工智能 (AGI)，具备强大的语言处理能力和多模态交互能力，适用于多种下游任务和领域。

## 7.5 远程登录和数据传输

### 7.5.1 ssh 远程登录

ssh 是默认的远程服务器访问软件，小巧、安全、快速。命令一般执行方式如下：

```
ssh [username@remotehost]
```

第一次登录要求在本地保存登陆指纹，输入 y 确认。然后屏幕提示输入远程的登录密码。

其他登录参数：

- **-p 22** - 采用 22 端口登录。ssh 的默认端口是 22。
- **-Y** - 启用远程 GUI
- **-i**  
*file*  
- 使用指定的登录密钥

Windows 平台登录时可使用 PowerShell；但是 Windows 平台推荐使用MobaXTerm软件。

在 Linux 或者 mac 平台可以直接使用 terminal 命令行；Mac 平台也推荐使用iTerm2，可以使用多开方式。

ssh 同时支持文件传输，例如 sshfs 和 sftp。快速的文件传输和查看，在 Mac 和 Linux 平台，可以使用CyberDuck；Windows 平台推荐使用WinSCP。

### 7.5.2 免密码登录

执行 ssh-keygen，可在客户端生成用户密码，生成的用户密码可用于免密码的 SSH 登录。

```
ssh-keygen
```

然后执行查看命令，查看密码申请状况。

```
ls ~/.ssh/  
cat ~/.ssh/id_rsa
```

将本地生成的密钥拷贝到远程服务器上，此处需要输入登陆远程服务器的密码。

```
ssh-copy-id [username@remotehost]
```

如果远程服务器端口为 32099，则命令改为：

```
ssh-copy-id -p 32099 [username@remotehost]
```

成功之后，可以 SSH 免密方式登陆远程服务器。

```
ssh -i ~/.ssh/id_rsa [username@remotehost]
```

如端口变为 32099，则命令为：

```
ssh -p 32099 -i ~/.ssh/id_rsa [username@remotehost]
```

### 7.5.3 wget 批量下载数据

关于批量下载数据，请参考博客文章：<https://www.shulele.net/zh/eosdata/>。

## 7.6 Linux 操作系统

### 7.6.1 用户名管理

新建用户 userx

```
sudo useradd -s /bin/bash -d /home/userx/ -m -G sudo userx
```

/bin/bash 是用户默认的登录 shell 界面。/home/userx/ 使用用户的 Home 目录位置。-G sudo 是指定用户的所属的组。sudo 组用户则在输入 sudo 命令时的具有 root 权限。

登录指定用户 userx

```
su userx
```

### 7.6.2 Linux 常用的命令。

部分命令的默认 Ubuntu 系统中不存在时，需要使用 apt 安装。

- 安装软件 apt: sudo apt install tree

- man: 使用 man ls 则可以查看 ls 命令的使用说明; man 命令组合可以看所有命令的说明。
- cd: 切换不同的目录
- ls: 查看目录和文件
- cat: 在屏幕打印出文本文件
- more, head, tail, 查看文本文件
- tree 查看目录树结构
- ip a: 查看服务器 IP 信息
- lsblk: 查看硬件设备
- df: 查看磁盘设备挂载情况
- du: 查看磁盘使用率
- jobs : 查看后台运行的用户程序
- ps: 查看进程
- kill: 杀死/关闭某一个进程
- scp: 通过 ssh 通道拷贝数据
- rsync: 使用数据更新方式拷贝数据, 支持本地数据拷贝或者远程 ssh 数据拷贝。
- wget/curl: 数据下载
- mount/umount: 挂载和卸载磁盘。
- grep 正则表达式: 用于查找文件里符合条件的字符串。
- find: 用于查找目录中的文件

## 7.7 R

R 是高效且灵活的编程语言, 可以高效的完成数据读写、统计分析、空间数据处理处理、并行计算等任务。

R 语言是一种免费的、开源的数据分析和统计建模语言, 相对于其他统计分析软件, R 语言有以下显著优势:

- **免费且开源:** R 语言是一款完全免费、开源的软件, 用户可以自由地下载、使用和修改它。这使得 R 语言成为许多研究人员、学生和行业专业人士的首选工具之一。

- **强大的统计分析能力:** R 语言提供了许多统计分析和建模的方法，包括线性回归、非线性回归、时间序列分析、聚类分析、因子分析和机器学习等。R 语言还提供了许多常用的数据处理和可视化工具，如数据清洗、数据可视化和报告生成等。
- **社区支持:** R 语言拥有一个庞大的用户社区，用户可以通过该社区获得技术支持、交流使用经验和分享资源等。R 语言社区提供了大量的免费学习资源，包括教程、示例代码和数据集等。
- **可扩展性:** R 语言可以通过许多扩展包（packages）来扩展其功能。这些扩展包提供了各种各样的功能，从数据可视化到高级统计分析和机器学习。
- **易于学习和使用:** R 语言拥有易于学习和使用的语法和语言结构，许多人认为 R 语言比其他统计分析软件更加容易学习和使用。
- **跨平台支持:** R 语言可以在 Windows、MacOS、Linux 等多个操作系统上运行，并提供相应的安装程序和二进制文件。

本研究组的 rSHUD(<https://github.com/shud-System/rshud>)，AutoSHUD(<https://github.com/shud-System/autoshud>) 和全球数据云平台 (<https://ghdc.ac.cn>) 都由 R 语言实现。

具体的 R 语言教程请参考《R 在地球科学中的应用》<https://www.shud.xyz/bookr/>。

### 7.7.1 Rstudio Server

Rstudio Server 服务入口，请使用超算的账户登录。

- SHUDHPC: <https://rstudio.shud.vip> 或者 <http://210.77.77.22:8787>

Rstudio Server 的使用手册：<https://s3.amazonaws.com/rstudio-server/rstudio-server-pro-0.98.507-admin-guide.pdf>

## 7.8 Python {#python}

SHUD-HPC 的 Python 入口: <https://py.shud.vip>



# 第八章 科研资源

## 8.1 工作环境

单位：中国科学院西北生态环境资源研究院

地址：甘肃兰州城关区东岗西路 320 号

办公室：一号科研楼

## 8.2 图书馆

中国科学院兰州分院的图书馆位于园区西大门旁边的中国科学院情报中心一楼，可以自习、看书、借阅图书。具体时间表和借阅政策请到图书馆咨询。

甘肃省省图书馆距离研究所只需 10 分钟走路距离，请查阅地图。

## 8.3 文献下载

- CNKI <https://www.cnki.net>。只要使用单位的 IP，自动具有在 CNKI 下载文献的权限。
- 中国科学院兰州文献情报中心 <http://www.llas.cas.cn>。多个中文和英文文献下载入口。

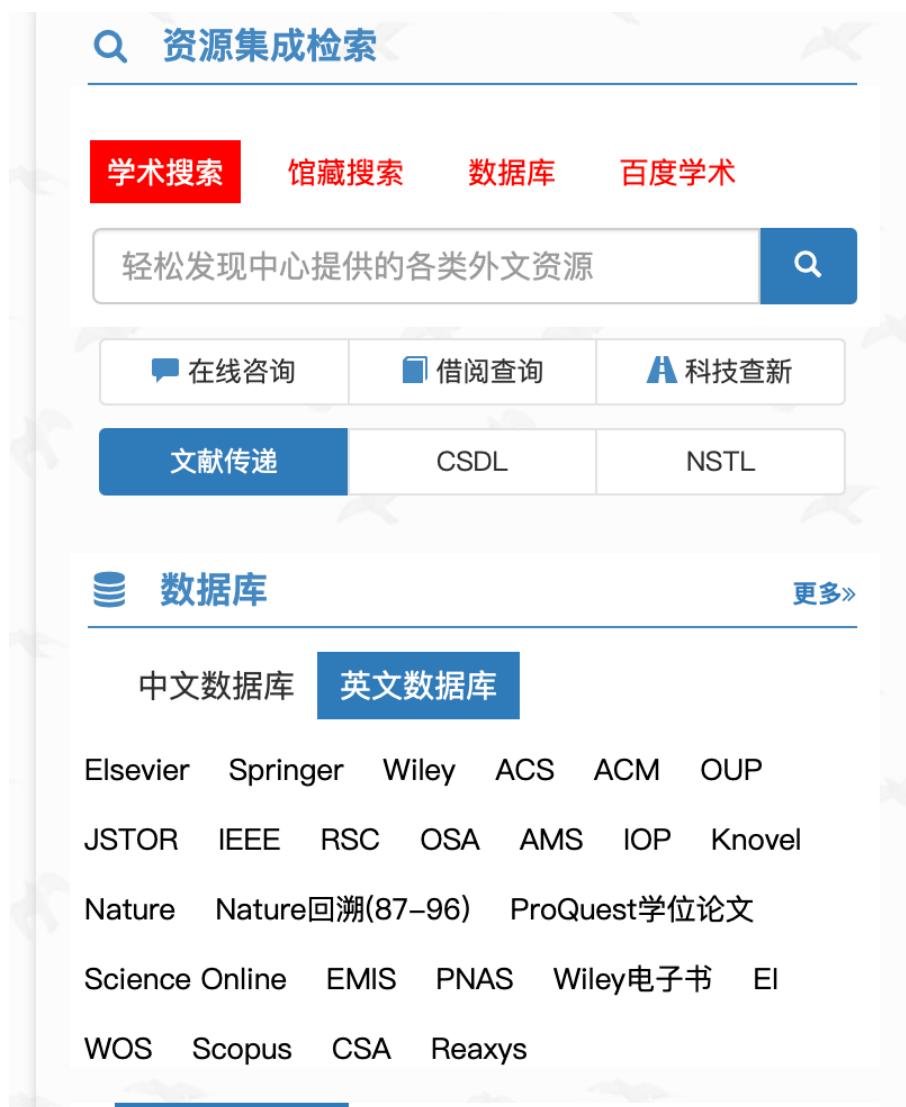


图 8.1: 外文数据库



图 8.2: 中文数据库

## 8.4 数据

绝大部分我们使用的数据已经可以从网上公开渠道获取，利用搜索引擎即可找到下载入口。很多数据存放于各大数据中心，我们列举一部分常用的数 据中心。

- 国家冰川冻土沙漠数据中心 <http://www.ncdc.ac.cn>
- 国家青藏高原科学数据中心 <http://www.tpdc.ac.cn>
- 资源环境科学与数据中心 <https://www.resdc.cn>
- NASA Earth Observation Data <https://www.earthdata.nasa.gov>
- USGS EarthExplorer <https://earthexplorer.usgs.gov>
- World Climate Research Programme, CMIP6 data <https://esgf-node.llnl.gov/projects/cmip6/>
- World Soil Information <https://www.isric.org>
- USDA Web Soil Survey <https://websoilsurvey.sc.egov.usda.gov/App/HomePage.htm>
- FAO Soil Portal <https://www.fao.org/soils-portal/data-hub/soil-maps-and-databases/en/>
- MERIT Hydro: global hydrography datasets [http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT\\_Hydro/](http://hydro.iis.u-tokyo.ac.jp/~yamadai/MERIT_Hydro/)
- MERIT-Basins <https://www.reachhydro.org/home/params/merit-basins>
- National Hydrography Dataset (NHD) <https://www.usgs.gov/national-hydrography/national-hydrography-dataset>
- HydroSHEDS <https://www.hydrosheds.org/>

# 第九章 科研设备

本研究组分别有四台供组员使用的计算服务器和网络存储。四台服务器分别为 Windows Server 远程桌面、Ubuntu 远程桌面、多节点高性能计算机群 (SHUDHPC)、GPU 服务器。

各服务器访问 IP 分别为：

平台	主机名	公网 IP 地址	内网 IP 地址	端口转发地址
计算集群	xnode	shud.vip (210.77.77.22)	10.0.1.100	210.77.77.25:318xx
GPU 计算服务器	gpu	无	10.0.1.98	无
Ubuntu 远程桌面	uDesk	210.77.77.23	10.0.1.23	210.77.77.25:319xx
Windows 远程桌面	Win-Desk	210.77.77.24	10.0.1.24	210.77.77.25:320xx
路由器	H3C	210.77.77.25	10.0.1.1	无
NAS 共享网盘	shudxyz	nas.shud.vip (210.77.77.36)	10.0.1.26	无

注意：ssh 登录端口统一为 32099

SHUDHPC 与 uDesk 通过 NIS 共享同一套账户系统，因此，在 SHUDHPC 和 uDesk 上面的登录账户和用户主目录完全相同；uDesk 上支持 Linux 图形桌面，方便用户进行可视化操作。

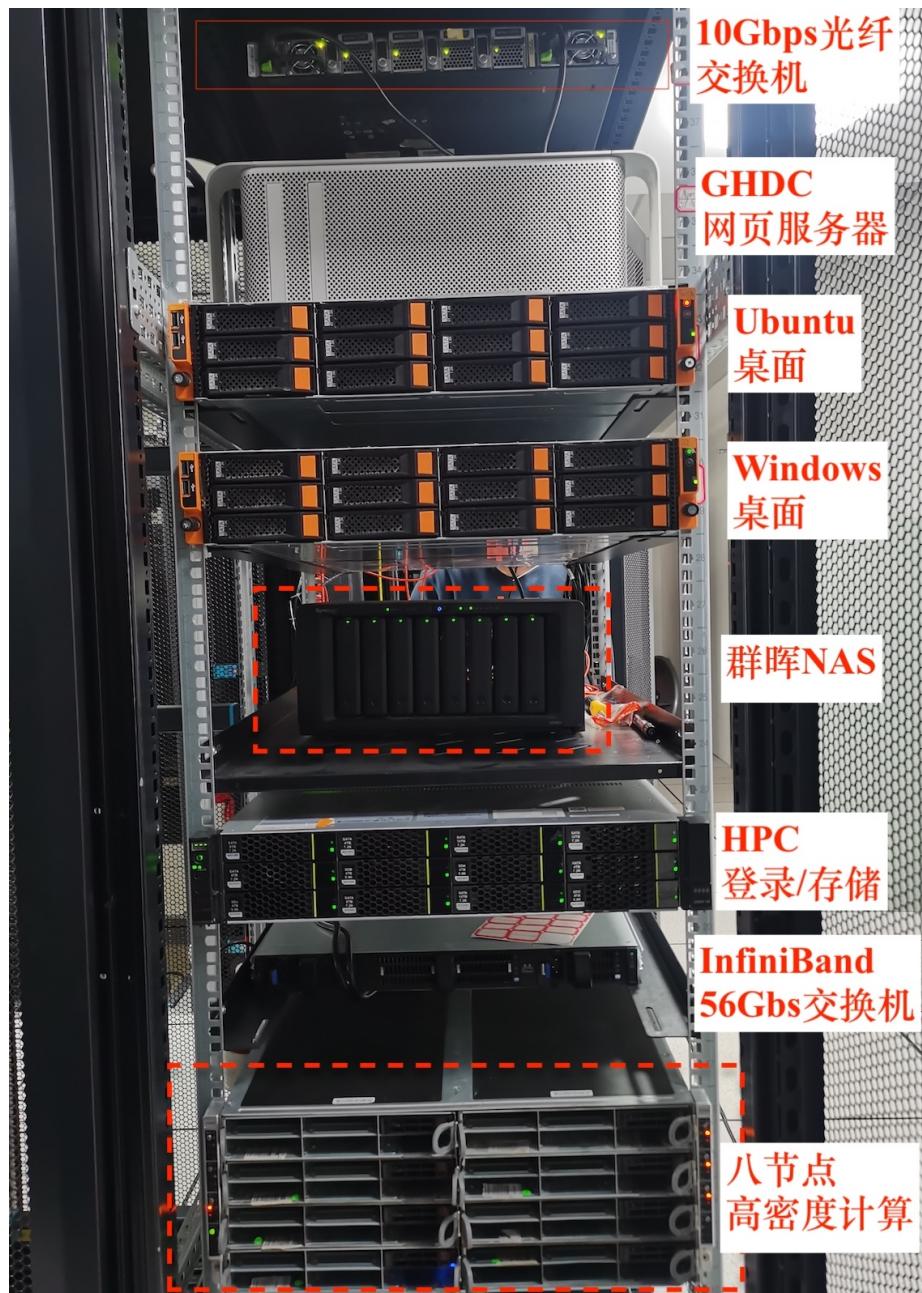


图 9.1: 所有主机照片

## 9.1 NAS 网络存储

本组使用群辉 DS1821+ 建设了 100TB 的网络存储空间，供组内成员和合作者使用。

Synology Drive 是依托本组的 NAS 网络存储建设的，可以远程以网络磁盘的方式访问，也可以使用 sftp,、Synology Drive、rsync、SMB、 AFP、NFS 等方式访问。

### 9.1.1 账号开通/注销

- 账号开通

- 在加入本组后，向 PI 报告你的邮箱；由 PI 给你开通账号，并告知用户名和密码。
- 账号的权限等级由个人身份决定，不同权限对 NAS 的读写权限略有不同。
- 账号开通后，你会收到一封邮件，邮件包含了 NAS 服务的访问地址。请在收到邮件后登录 NAS 页面，并修改秘密。

- 账号注销

- 成员离开本组后，账号会继续保留至少一年。
- 你可以在一年内，备份个人网盘内的数据。一年后，你的账号将被删除，无法再访问 NAS 系统。

### 9.1.2 NAS 文件结构

当前的 NAS 文件结构与读写权限见下图：

更具体的信息见下表：

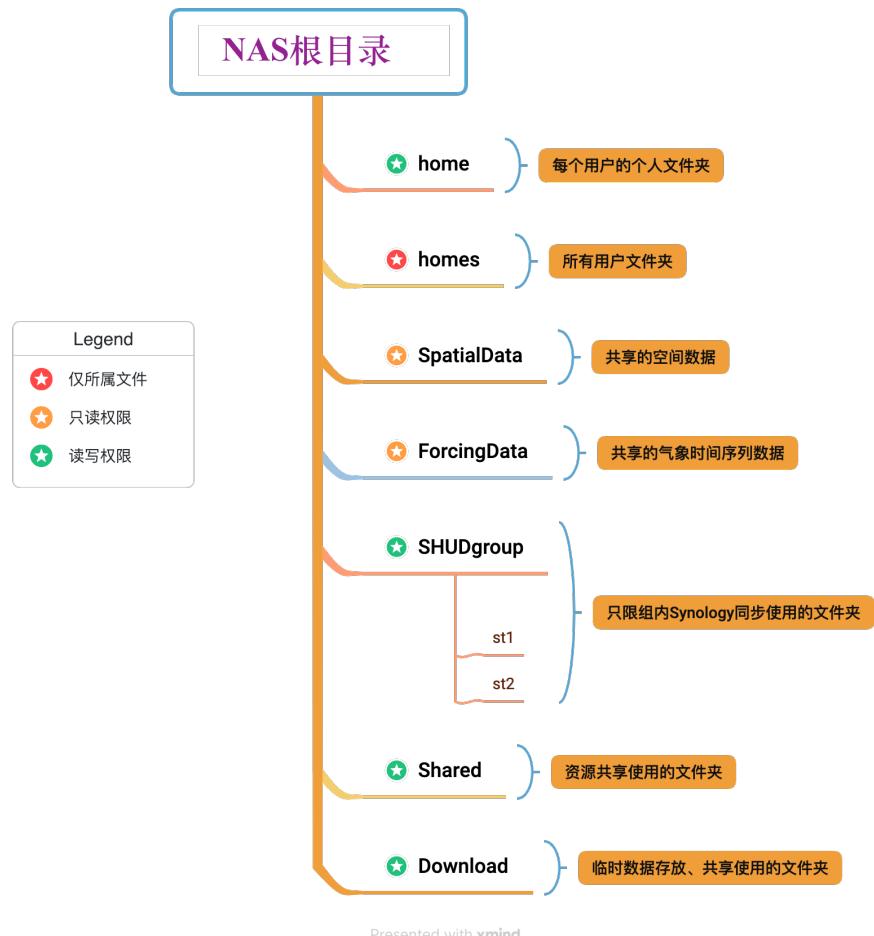


图 9.2: NAS 上的文件结构与权限

文件夹	学生权限	合作者权限	容量	注意事项
homes	无	无	无限	所有用户文件都存放于此，普通用户无法访问

文件夹	学生权限	合作者权限	容量	注意事项
home	读写	读写	10TB	用户个人主目录，其他用户无法访问
SpatialData	只读	只读	10TB	仅特定用户有权更新
ForcingData	只读	只读	10TB	仅特定用户有权更新
SHUDgroup	读写	读写	10TB	组内成员间自动同步。 勿随意存放文件
Shared	读写	读写	10TB	共享资源
Download	读写	读写	无限	文件临时存放。所有人可写，勿长期存放重要文件
Baomi	无	无	无限	保密数据，仅在特殊需要时共享给指定用户

### 9.1.3 访问 NAS

- NAS 名称: shudxyz
- NAS 访问地址: <https://nas.shud.vip> 或者 <https://210.77.77.36>
- 用户名: username
- 密码:

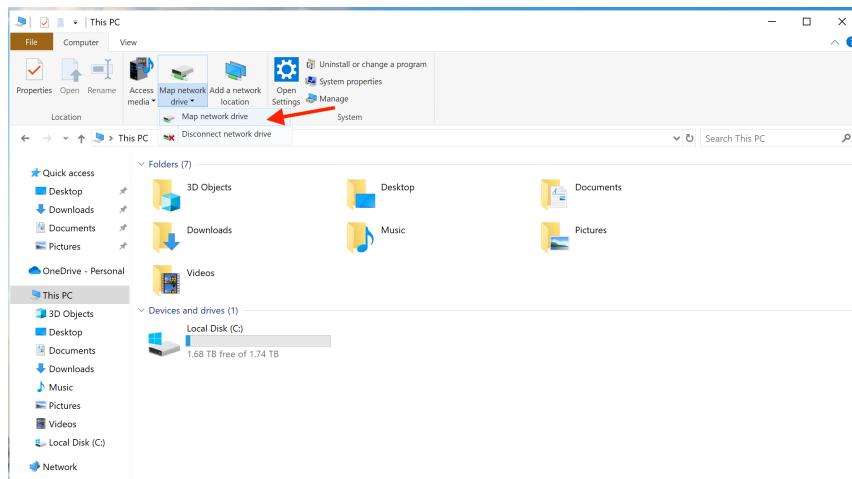
### 9.1.4 NAS 作为网络磁盘

- Mac OS

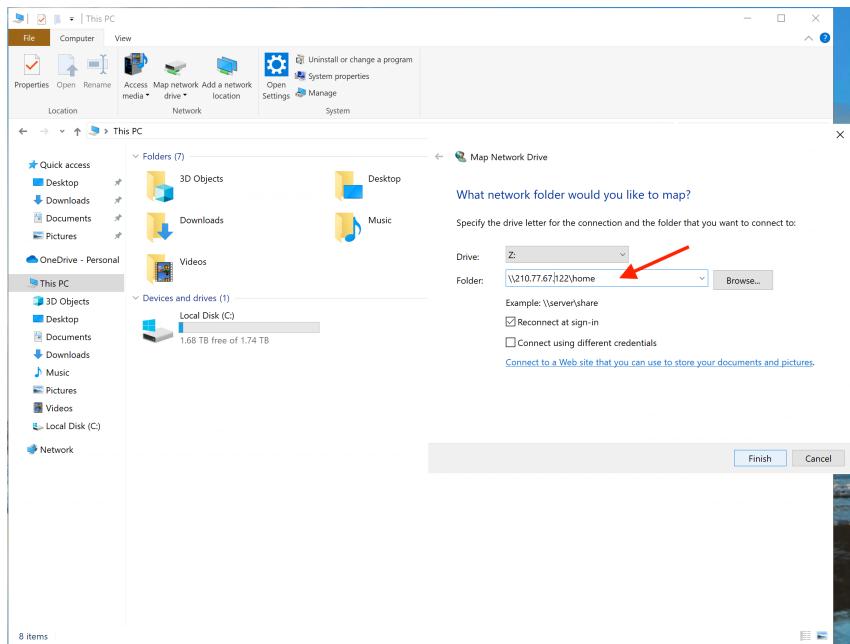
- 打开 Finder (访达), 然后使用键盘 CMD+k, 访问网络地址。
- 在打开的窗口中输入:afp://nas.shud.vip 或者 afp://210.77.77.36。然后点击连接 (Connect)
- 然后提示框要求输入用户名和密码。
- 选择你需要加载的磁盘。

- Windows

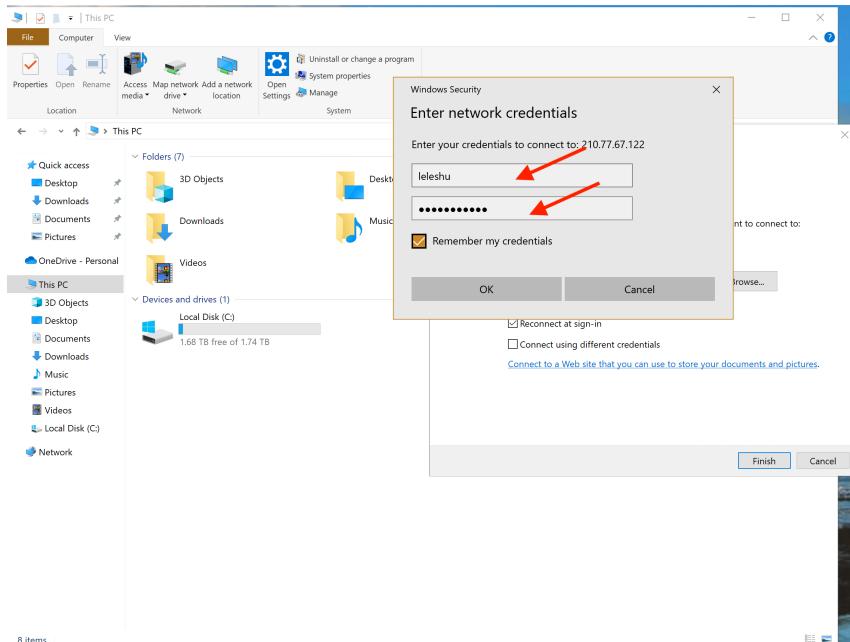
1. 打开文件管理器, 找到映射网络磁盘 (Map network driver),



1. 在打开的窗口中输入需要加载的磁盘 IP 和路径, 例如\\nas.shud.vip\home 或者\\nas.shud.vip\ForcingData。点击完成

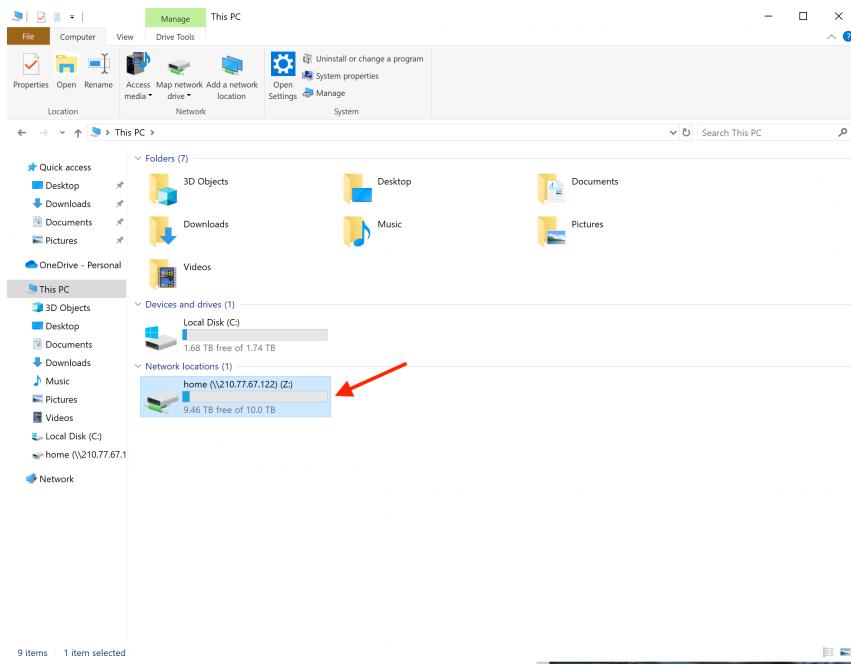


1. 根据提示框输入登录 NAS 使用的用户名与密码。点击 OK



1. 然后系统会加载指定的网络磁盘，映射其为本地磁盘。此后该网

络磁盘将可以像一般本地磁盘一样操作。



注：如果在 Windows 远程桌面（Win-Desk）上加载 NAS，使用本地 IP(10.0.1.x 段)，可以显著提高网络磁盘的读写速度。

### 9.1.5 云盘 Synology Drive

这应该是进入本组第一个学会的软件。软件的安装使用都非常简单，但是你需要学会的是重新思考：

- 如何组织自己的文件？
- 如何高效的定位自己的文件？

加入本组的工作全部在 Synology Drive 上进行同步。同步文件夹的内容仅自己可以看到，其他用户看不到。文件备份在 NAS 云端，任何地点任何电脑上都可以查看/编辑你云盘里面的文件，不会出现“电脑坏了，文件丢失”，“文件染病毒打不开”，“电脑忘记带了，没法交作业”的情况。

注意：Synology Drive 使用中有同步（sync）和备份（backup）两种模式，建议一定要使用同步模式，保证自己每一次保存的文件都实时同步在云端。是否启用备份模式，可以自行选择。

## 9.2 Windows 远程桌面

### 服务器硬件配置

类别	配件	参数	备注
平台	浪潮雷神 SA5212 H5	12 盘位 2U 机架式服务 器	
CPU	至强 Xeon 金牌 6133	每 CPU 20 核 40 线程， 2.5GHz	双 CPU
内存	192GB	DDR4 2666Mhz, 12x 16GB	6 通道内存， 12 条
硬盘	U.2 NVME	2TB	
网卡		10Gbps 光口	

### 9.2.1 操作系统与软件

- 操作系统：Windows Server 2019，支持远程 GUI 登录。
- Office：Microsoft Office 365
- 文本编辑器：Notepad++, vim.
- PS 工具：GIMP
- GIS 软件：ArcGIS, QGIS
- 编程：Visual Studio Code
- R：R, Rstudio
- 计算：Octave

### 9.2.2 登录方式

- 软件：Microsoft Remote Desktop

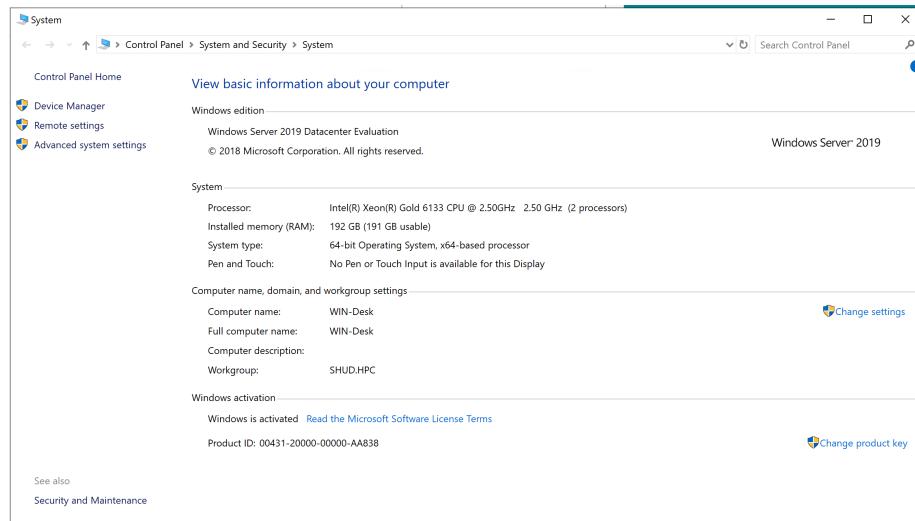


图 9.3: Windows 系统基本信息

- 登录 IP: 210.77.77.24
- 端口: 默认
- 转发 IP: 210.77.77.25
- 转发端口: 31810

### 9.3 Ubuntu 服务器

#### 服务器硬件配置

类别	配件	参数	备注
平台	浪潮雷神 SA5212 H5	12 盘位 2U 机架式服务器	
CPU	至强 Xeon 金牌 6133	每 CPU 20 核 40 线程, 2.5GHz	双 CPU
内存	16GB x12	DDR4 2666Mhz, 16 GBx12	6 通道内存, 12 条
硬盘	U.2 NVME	2TB	

类别	配件	参数	备注
网卡		10Gbps 光口	两路光纤

### 9.3.1 操作系统与软件

- 操作系统: Ubuntu 20.04, 支持远程 GUI 登录。
- 远程桌面系统: xfce4
- GIS 软件: QGIS 3.10
- R: R, Rstudio
- Octave

### 9.3.2 SSH 登录

#### 公网 IP 登录

- 登录 IP: 210.77.77.23
- 软件: ssh (命令行);
- 端口: 32099

示例:

```
ssh zhangsan@210.77.77.23
```

#### 路由器转发登录

- 登录 IP: 210.77.77.25
- 软件: ssh;
- 端口: 32022

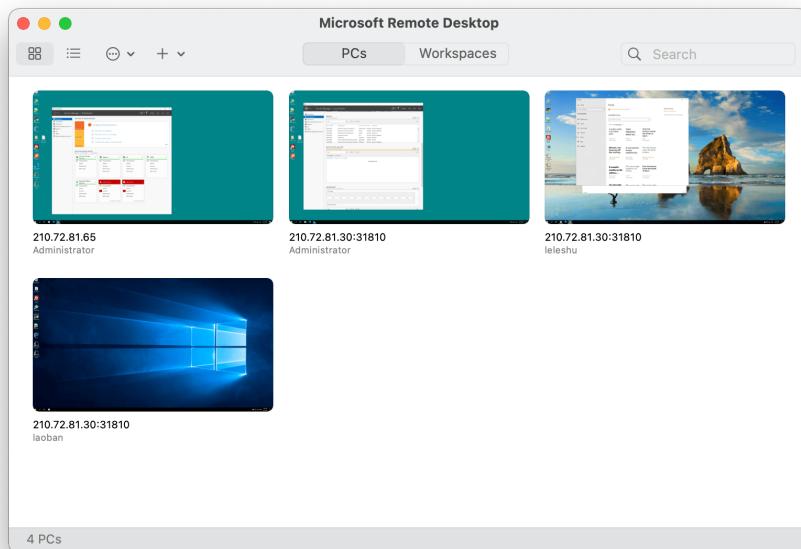


图 9.4: Microsoft Remote Desktop 界面

访问端口	服务器 SSH 端口
32022	32099

示例：

```
ssh -p 32022 zhangsan@210.77.77.23
```

### 9.3.3 VNC 远程桌面

- 软件：VNC（远程桌面）
- 端口：共 20 个端口，最多支持 20 个 VNC 远程桌面，端口对应关系如下表：

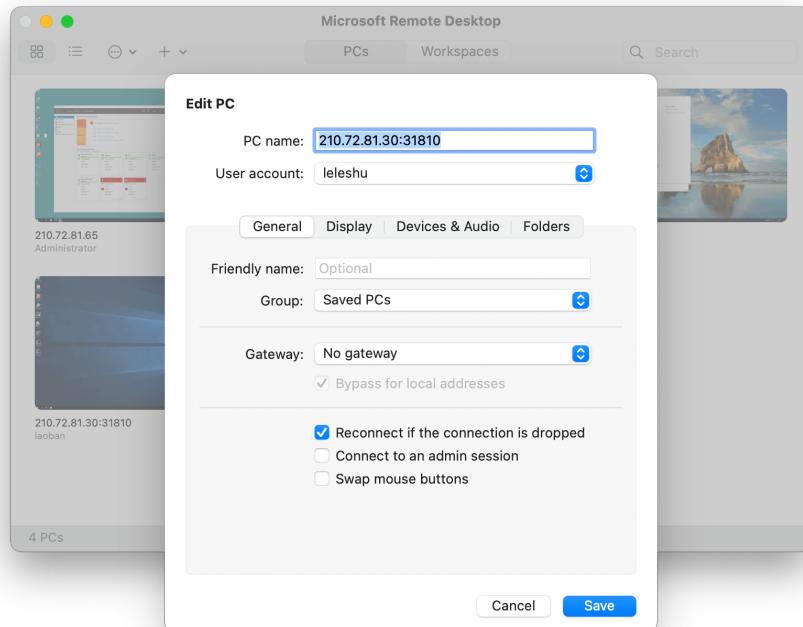


图 9.5: Microsoft Remote Desktop 登录配置

访问端口	服务器 VNC 端口
32001	5901
32002	5902
32003	5903
320xx	59xx
32020	5920

Ubuntu 服务器端配置：建立 vnc 密码，并修改 `~/.vnc/xstartup` 文件。

```
vncpasswd  
touch ~/.vnc/xstartup  
chmod +x ~/.vnc/xstartup  
vim ~/.vnc/xstartup
```

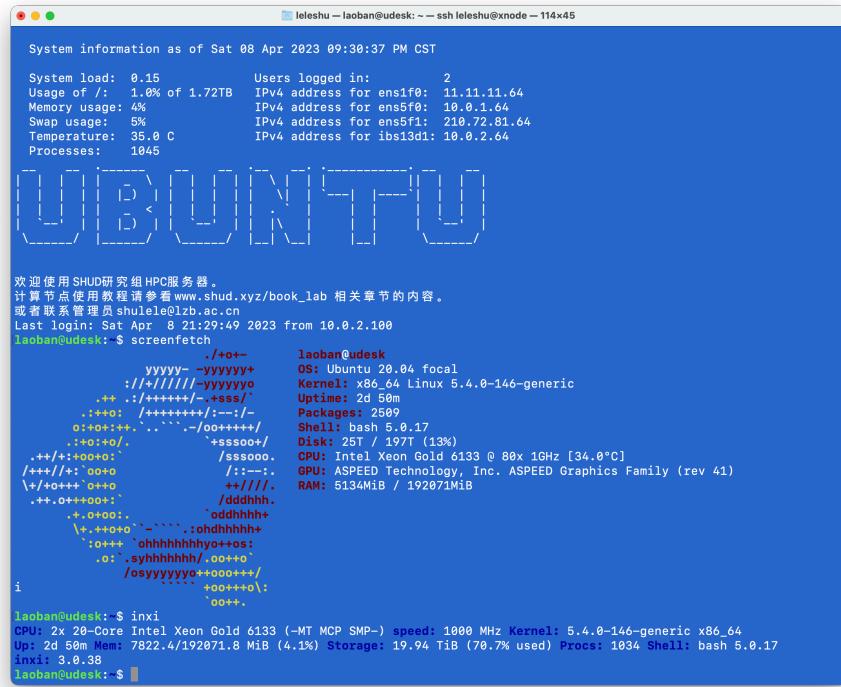


图 9.6: uDesk 基本配置

在文件中输入文件内容：

```
#!/bin/sh
unset SESSION_MANAGER
unset DBUS_SESSION_BUS_ADDRESS
startxfce4 &
```

登录远程桌面：

```

leleshu -- laoban@udesk: ~ -- ssh leleshu@node -- 114x45
laoban@udesk: $ vncserver -geometry 1920x1080 ← step 1
New 'X' desktop is udesk:4

Starting applications specified in /home/laoban/.vnc/xstartup
Log file is /home/laoban/.vnc/udesk:4.log

laoban@udesk: $ ss -ltn ← step 2
State Recv-Q Send-Q Local Address:Port Peer Address:Port Process
LISTEN 0 4096 0.0.0.0:59075 0.0.0.0:* 0.0.0.0:*
LISTEN 0 64 0.0.0.0:2049 0.0.0.0:*
LISTEN 0 4096 0.0.0.0:994 0.0.0.0:*
LISTEN 0 64 0.0.0.0:43557 0.0.0.0:*
LISTEN 0 4096 0.0.0.0:47373 0.0.0.0:*
LISTEN 0 5 0.0.0.0:5901 0.0.0.0:*
LISTEN 0 4096 0.0.0.0:37613 0.0.0.0:*
LISTEN 0 5 0.0.0.0:5902 0.0.0.0:*
LISTEN 0 5 0.0.0.0:5903 0.0.0.0:*
LISTEN 0 4096 0.0.0.0:111 0.0.0.0:*
LISTEN 0 5 0.0.0.0:5904 0.0.0.0:*
```

... (many more ports listed)

```

laoban@udesk: $ vncserver -kill :4 ← step 4
Killing Xtightvnc process ID 316108
laoban@udesk: $
```

1. 在服务器端启动 vncserver。-geometry 参数可以设置远程桌面的分辨率。以下命令三选一

```

vncserver -geometry 1440x900
vncserver -geometry 1920x1080
vncserver -geometry 2560x1440
```

2. 然后使用 ss -ltn 当前 vnc 桌面的端口号，默认第一个 vnc 桌面端口号为 5901，第二个为 5902，以此类推。或者通过命令 server -list 查看。使用路由器端口转发，需要将 590x 的端口转换为 320xx 的端口，转换规则见前表。如果直接使用服务器公网 IP，则直接使用 590x 端口。

3. 在客户端启动 vnc 工具，访问  
*serverIP*

:

。如使用公网 IP，且 vnc 服务端口为 5904:

```
vnc://210.77.77.23:5904
```

如使用路由器 IP，且 vnc 服务端口为 5904，则转发端口为 32004:

```
vnc://210.77.77.25:32004
```

- 完成远程桌面工作后，退出 vncserver 时，需要关闭响应的 vnc 服务  
器，请在远程服务器端输入：

```
vncserver -kill :4
```

使用完远程桌面，请主动关闭该 vncserver

## 9.4 GPU 服务器

GPU 服务器配置：

类别	配件	参数	备注
平台	浪潮雷神 SA5212 H5	12 盘位 2U 机架式服务器	
CPU	至强 Xeon 金牌 6133	每 CPU 20 核 40 线程，2.5GHz	双 CPU
内存	32GB x12	DDR4 2666 Mhz, 384GB	6 通道内存，12 条
硬盘	U.2 NVME	2TB	
GPU	NVIDIA V100	24GB HBM2	4 卡
网卡		10Gbps 光口	单光纤
网卡		100Gbps IB 光口	IB

### 9.4.1 操作系统与软件

- 操作系统：Ubuntu 22.04，支持远程 GUI 登录。
- LLM 模型：Ollama

### 9.4.2 登录方式

- 先登录 SSH 计算集群
- 然后使用 ssh 登录 GPU 服务器: `ssh gpu`

## 9.5 SHUDHPC 高性能计算集群

高性能计算机 (High-performance computer, HPC) 是由多台计算机构成的服务集群 (cluster)。HPC 主要由计算服务器、管理服务器和存储服务器组成, 但只设有一个登录入口。用户登录到 HPC 后, 只需将需要计算的相关任务提交至超算平台, 平台计算任务调度系统将根据任务的需求和实际可用资源对任务进行排队和资源分配。

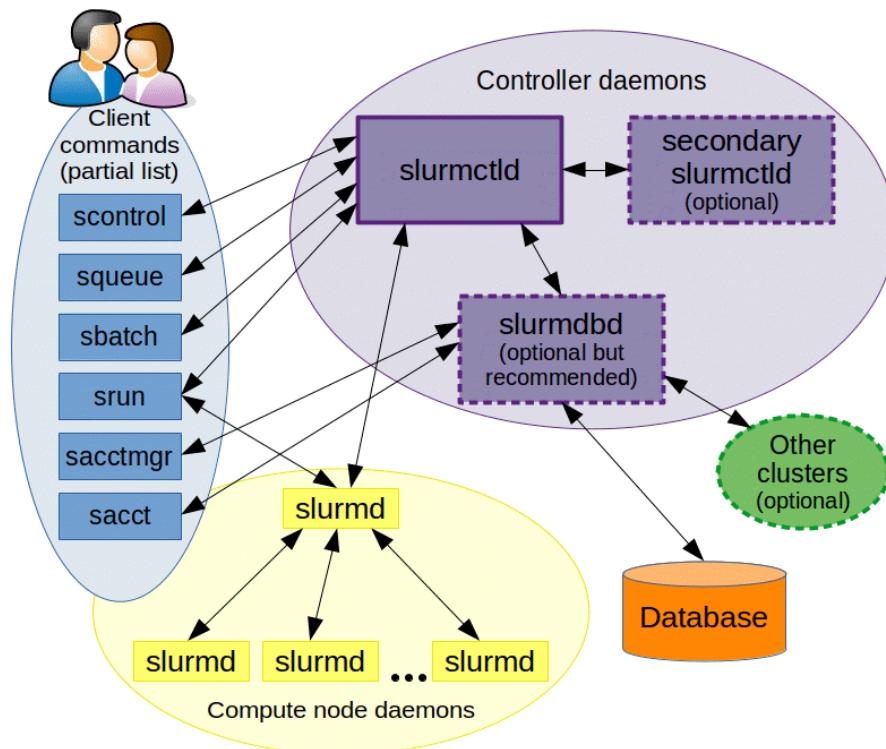


图 9.7: slurm 集群基本结构 (来自: <https://slurm.schedmd.com/overview.html>)

### 9.5.1 登录

- 访问地址: **shud.vip**
- ssh 端口: 32099
- 网页 SSH 访问: **ssh.shud.vip**
- RStudio Server 访问: **rstudio.shud.vip**

### 9.5.2 服务器硬件配置

大类	平台	参数	备注
平台	高性能计算集群	24 计算节点, 1 登录节点, 1 高 速存储	56G IB 网络, 10G 以太网
登录节点	超聚变 2288 V6	2x 至强 6318Y (56 核), 128G, 2TB U.2 NVME	
存储	超聚变 2288 V6	14x 16TB, RAID 5	173TB 可用
全闪存储	泰安 3036	12x 3.76TB, ZFS RAIDZ2	34TB 可用
计算节点	超微 24 节点高密 计算平台	每节点: 至强 6133 40 核 (2x 20), 192GB 内 存	
IB 交换机	SX6025	36x 56Gbps	InfiniBand
万兆以太交换机	IBM	48x 10Gbps, 4x 40Gbps	全光口
千兆以太交换机	H3C	48x 1Gbps	全光口

### 9.5.3 SHUDHPC IP 地址配置

节点	主机名	万兆 IP	InfiniBand IP 地址
登录节点	xnode	10.0.1.100	10.0.2.100
全闪存储	flash	10.0.1.99	10.0.2.99
存储	stor	10.0.1.100	10.0.2.100
计算节点	cn01 - 24	10.0.1.101~124	10.0.2.101~124

#### 9.5.4 操作系统与软件

- 操作系统: Ubuntu 20.04, 仅支持命令行登录。
- 作业调度: slurm
- R: R
- Octave

#### 9.5.5 存储信息

假设用户名为 **zhangsan**

目录	所属	权限	使用规则
/volume/repo/zhangsan	用户冷数据空间	700	用于存放用户程序、文件等, 限 10TB 容量
/users/zhangsan	用户主目录	700	仅用于编译、文件库等文件存放; 禁止大文件, 禁止计算。
/scratch/zhangsan	高速计算空间	700	可进行计算任务, 任何超过 60 天的文件会自动删除
/volume/data	冷数据盘	777	公共空间。请勿存储个人数据

目录	所属	权限	使用规则
/tmp*	临时数据	777	公共临时目录；如需高速运算，输出数据可存在/tmp里面
/opt	软件安装/编译 盘	755	公共程序安装目录

注：每一个计算节点上面都有/tmp 目录，这个磁盘属于计算节点的 U.2 NVME 高速磁盘，写入速度约 2000MB/s，计算结果存储在/tmp 速度最快。但是，当计算完成后，用户无法访问计算节点的/tmp。例如：用户在主节点 xnode 上提交任务（./shud -o /tmp/ccw.out ccw）给 cn01 节点，计算过程中，cn01 上的任务将结果保存在 cn01:/tmp/ccw.out 里面。但是，当任务完成后，用户无法直接访问 cn01:/tmp 下的文件，用户只能访问到 xnode:/tmp 下的文件。因此，如果任务提交时选择/tmp 写出，那么需要在任务脚本中加入一个拷贝/tmp/ccw.out 到用户个人目录的命令，以此保证用户可以获得计算结果。

### 9.5.6 用户管理

- zhangsan：一般用户

研究组的每个成员将获得一个登录账号，账号类型为超算一般用户，具有提交任务的权限。如果任务需要特殊要求，请与 PI 进行沟通。

### 9.5.7 作业管理 Slurm

SLURM (Simple Linux Utility for Resource Management) 是一个流行的开源的作业调度和集群管理系统，主要用于高性能计算和科学计算领域。SLURM 的主要特点包括：灵活的资源管理、可扩展性、高可用性、高可靠性、多种作业调度算法和管理工具等。

在 SLURM 中，用户提交的作业会被分配到可用的计算节点上进行计算。SLURM 会根据可用资源的情况，按照用户指定的优先级和作业调度算法来安排作业的执行顺序，以达到最大化集群的利用率和性能。同时，SLURM 还提供了一系列管理工具，如节点管理、队列管理、用户管理、资源限制等，方便管理员对集群进行管理和监控。

SLURM 的使用十分广泛，被许多知名的超算中心采用。

当前 24 个计算节点全部归属同一个计算分区（partition）。`partition = suan`。

命令代码

```
srun -N 5 -n 5 hostname #5个节点，5个CPU  
srun -N 5 -n 50 hostname #5个节点，50个CPU
```

代码示例 1：

```
#!/bin/bash  
#SBATCH --job-name=hostname  
#SBATCH --partition=suan # 计算分区名称。  
#SBATCH -N 1 # 计算节点数量  
#SBATCH --mail-type=end  
#SBATCH --mail-user=YOU@EMAIL.COM  
#SBATCH --output=%j.out # 屏幕输出文件  
#SBATCH --error=%j.err # 屏幕错误信息输出文件。  
  
/bin/hostname
```

以上代码保存为 `submit1.sh`。然后在命令行中执行以下命令：

```
sbatch submit1.sh
```

代码示例 2：

```
#!/bin/bash
```

```

#
#SBATCH --job-name=echo_number # 任务名称
#SBATCH --output=slurm_%j.out # 屏幕输出及错误信息输出文件。
#SBATCH --ntasks=30 # CPU数量。

for i in {2000..2030}
do
    srun -n1 --exact echo $i &
done
wait

```

开启交互模式

```
srun --nodes=1 --ntasks-per-node=1 --time=01:00:00 --pty bash -i
```

**最简单的 SHUD 模型 slurm 任务案例：**

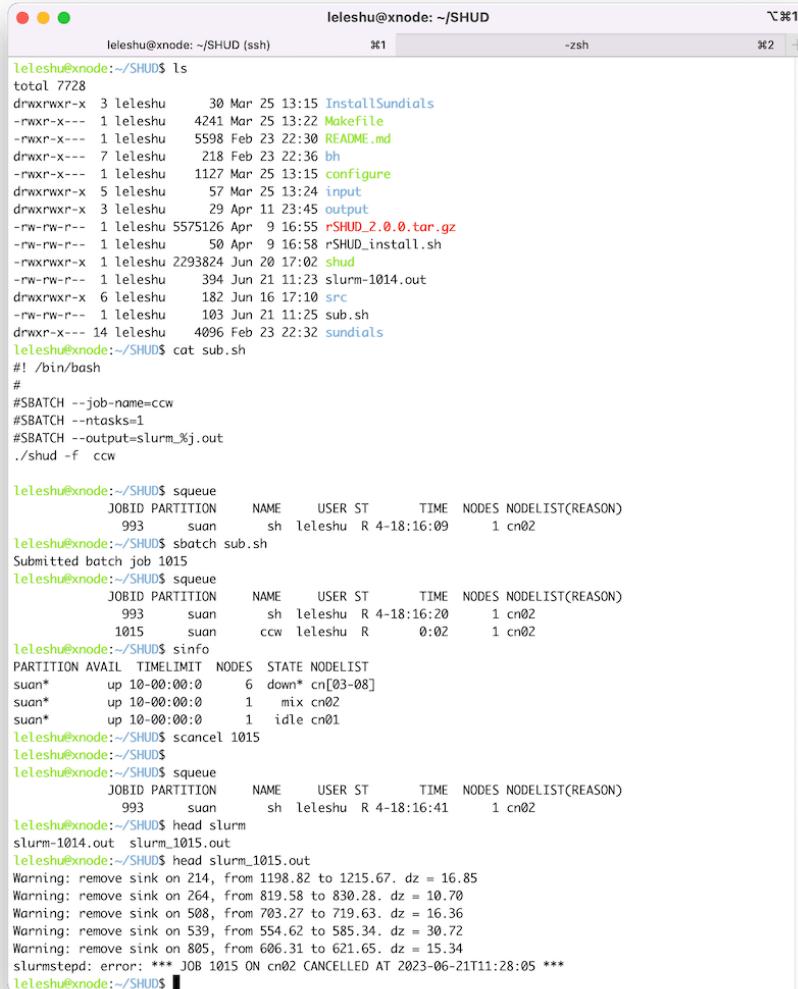
slurm 教程参考

- 上海交大：<https://docs.hpc.sjtu.edu.cn/job/slurm.html>
- 北京大学：[https://hpc.pku.edu.cn/\\_book/](https://hpc.pku.edu.cn/_book/)

### 9.5.8 Slurm 使用注意事项

1. 不要使用主节点进行计算工作，任务全部通过 slurm 提交到计算节点。
2. 不要使用 `exclusive` 选项。如果确实有必要，请提前告知 PI 和其他组员。
3. 如果计算任务的数据写入量巨大，临时数据（药渣数据）请存放 to /tmp 目录下——该文件夹是数据计算节点的 /tmp；然后修改程序，在程序结束之后从 /tmp 目录下移动/复制数据到 /user/zhangsan, /volume/repo/zhangsan 个人文件夹，再做数据后处理。

注意，移动 /tmp 下文件的操作，比如在提交 slurm 的脚本或者程序中体现——因为，如果 slurm 计算任务结束之后，在主节点上访问的 /tmp 是主节点 /tmp，无法再访问计算节点的 /tmp 文件夹。



The screenshot shows a terminal window titled "leleshu@xnode: ~/SHUD". The window contains the following command history:

```

leleshu@xnode:~/SHUD$ ls
total 7728
drwxrwxr-x 3 leleshu 30 Mar 25 13:15 InstallSundials
-rw-r-x---
```

... (long list of files)

```

Leleshu@xnode:~/SHUD$ cat sub.sh
#!/bin/bash
#
#SBATCH --job-name=ccw
#SBATCH --ntasks=1
#SBATCH --output=slurm_%j.out
./shud -f ccw

Leleshu@xnode:~/SHUD$ squeue
JOBID PARTITION      NAME      USER ST      TIME NODES NODELIST(REASON)
993     suan          sh leleshu R 4-18:16:09      1 cn02
Leleshu@xnode:~/SHUD$ sbatch sub.sh
Submitted batch job 1015
Leleshu@xnode:~/SHUD$ squeue
JOBID PARTITION      NAME      USER ST      TIME NODES NODELIST(REASON)
993     suan          sh leleshu R 4-18:16:20      1 cn02
1015    suan          ccw leleshu R 0:02      1 cn02
Leleshu@xnode:~/SHUD$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
suan*      up 10-00:00:0 6 down* cn[03-08]
suan*      up 10-00:00:0 1 mix cn02
suan*      up 10-00:00:0 1 idle cn01
Leleshu@xnode:~/SHUD$ scancel 1015
Leleshu@xnode:~/SHUD$ squeue
JOBID PARTITION      NAME      USER ST      TIME NODES NODELIST(REASON)
993     suan          sh leleshu R 4-18:16:41      1 cn02
Leleshu@xnode:~/SHUD$ head slurm
slurm-1014.out slurm_1015.out
Leleshu@xnode:~/SHUD$ head slurm_1015.out
Warning: remove sink on 214, from 1198.82 to 1215.67. dz = 16.85
Warning: remove sink on 264, from 819.58 to 830.28. dz = 10.70
Warning: remove sink on 508, from 793.27 to 719.63. dz = 16.36
Warning: remove sink on 539, from 554.62 to 585.34. dz = 30.72
Warning: remove sink on 805, from 606.31 to 621.65. dz = 15.34
slurmstepd: error: *** JOB 1015 ON cn02 CANCELLED AT 2023-06-21T11:28:05 ***
Leleshu@xnode:~/SHUD$ 
```

图 9.8: slurm 任务案例



# 附录 A 文献阅读

- 文献阅读是学术研究的认知基石。通过系统梳理领域内经典与前沿成果，研究者得以构建学科知识谱系，掌握核心理论范式与研究范式演化路径，避免低水平重复劳动。批判性阅读可训练学术敏感度，帮助博士生在庞杂信息中识别关键科学问题，形成具有创新潜力的研究切口。
- 文献综述具有学术对话的必要性。研究不是孤立的知识生产，而是与既有成果持续互动的过程。深度解析前人研究的逻辑框架与方法论局限，既能验证研究问题的学术价值，又可规避论证偏差。地理学强调空间异质性与尺度嵌套，唯有通过跨学科文献整合，方能建立多维视角的理论解释模型。
- 文献批判承载着学术传承与突破的双重价值。经典理论为研究提供锚点，而方法论争鸣则孕育创新机遇。博士生需在文献精读中锤炼理论迁移能力，例如将遥感技术嵌入传统地貌分析，或通过社会网络模型重构人文地理研究框架。这种知识再生产过程既是对学术伦理的恪守，更是推动学科边界拓展的核心动力。

## A.1 任务概述

1. 每位学生每年需完成 **100** 篇英文文献阅读，分为两部分：**50** 篇由导师推荐，**50** 篇由学生自行选择。
2. 每篇论文阅读后，向导师或组会进行汇报。
3. 论文选择要求，**必须英文**，研究内容与本组研究领域高度相关，具有明显的借鉴和学习价值。论文可自行选择，但需经过导师同意。

4. 要求进行全文精读，不借助任何全文翻译工具或 AI 进行全文或大面积翻译。
5. 其中 1/4~1/5 的论文需要总结为“阅读笔记”文章，至“流域水文模型”公众号。

## A.2 公众号“阅读笔记”文章安排

1. 每周周一发送一篇。负责的同学需提前两周确定将阅读的论文题目；提前三周完成“阅读笔记”的初稿。
2. “阅读笔记”初稿统一放在 NAS 上 SHUDgroup/公众号 2025 共享文件中，命名格式为文章发布日期-姓名（例如：2025-04-28 张三）。
3. “阅读笔记”的内容符合公众号的定位和阅读笔记基本要素。
4. 文章格式符合公众号文章格式，具体格式可参见公众号往期好文分享。
5. 严守学术道德。任何抄袭/拷贝/参考他人的阅读笔记、公众号文章、解读、翻译等，同样视为学术不端。

## A.3 忠告

论文阅读计划的目的：

1. 增加你们的学术知识储备
2. 提高英文阅读能力
3. 提高快速理解他人工作、提取重要信息的能力
4. 学习学术表达逻辑
5. 学习作图、读图、看图论述的能力
6. 增加对专业术语的敏感度，避免学术词汇的低级翻译错误。
7. 精准表达科学问题
8. 针对研究缺陷表达和讨论的能力
9. 学习论文格式

虽然这只是论文阅读 PPT 和阅读笔记，对学生而言，无法发论文、也不会影响毕业，但我们认为这是学生从学术前辈那里学习论述的好机会，是学习

如何阐述、总结、思考的好机会。所以老师会认真给 PPT 提意见、阅读并修改你的阅读笔记，希望学生从中学习。但如果哪个学生拿 AI 和翻译工具来糊弄交差，那这个训练就不合适你，我们的培养模式也就不合适你。

可以做得差，做得差我们有义务教；但如果想糊弄，就别走科研路。世上的宽敞的路很多，也不是一定要走科研这一条路。每条人生路都有不同的要求，但是科研要求认真、较真。

## A.4 阅读 PPT 指导意见

### PPT 版式

1. 封面是否包含：
  - 文献名称、关键图件或者首页截屏
  - 单位 LOGO（完整版）
  - 论文完整引用信息
  - 论文基本信息介绍，例如：
    - 论文作者/单位
    - 论文发表时间/期刊
2. 添加页码
3. 全程使用模板版式

### 一般研究论文的阅读

1. 摘要及关键词。摘要需认真翻译，并且标记其，研究背景、研究目的、研究方法、研究结果、研究结论等部分。
2. 目录页，展示论文基本结构
3. 生词表，配翻译。
4. 论文关键术语，要展开解释，网上搜图（标记出处）进行含义解释。
5. 关键图件，展开解释。关键术语翻译成中文。
6. 关键数据的列表
7. 论文方法，方法的评价方法和评价指标

- 模型：模型简单介绍、模型概念图
  - 统计方法：方法简单介绍
  - 评价指标：简单介绍、公式
8. 论文结果，结果的评价方法和评价指标
  9. 论文讨论，作者对问题的解释、研究缺陷、研究不足、未来研究方向等
  10. 结论，详细总结其结论。
  11. 关键参考文献，从中推荐 5 个以内的关键文献。
  12. 阅读者评语/总结。

### 综述文章阅读

综述论文有别于一般研究论文，主要是针对某一领域的研究现状进行总结和分析。

阅读笔记中除了摘要、关键词、目录页、术语、生词表、关键图件等要求之外，应当重点应该阅读和总结这些方面：

1. 研究背景：综述文章通常会介绍该领域的研究背景和重要性，阐明综述的目的和意义。
2. 研究历史背景：综述文章会回顾该领域的研究历史，介绍重要的里程碑和关键的研究成果。
3. 研究出发点/思路/角度的差异
4. 研究方法的差异
5. 不同研究见的对比
6. 综述文章强调系统性
7. 局限性、不足、存在的挑战
8. 未来研究方向

# 附录 B 入组新手任务

## B.1 个人材料准备

- 个人简历：PDF 版本。
- 个人形象照。
  - 形象照要求：略侧身、免冠、清晰可见。
  - 不要穿着过于正式的衣服。
  - 风格：智慧、阳光、成熟。不要拍成房产中介、保险推销员、模特等风格，比如手扶领带、双臂抱胸、拳头托下巴、食指点颧骨等——不要拍成这样！不要拍成这样！不要拍成这样！
  - 风格参考：<https://www.cee.psu.edu/department/faculty-list.aspx>
- 填写个人信息表格：包括个人基本信息、Office365/WPS 账号、紧急联系人、个人身份证正反面等。
- 

## B.2 阅读本研究组重要论文

本研究组的论文列表（建议按照以下顺序阅读）：

1. 舒乐乐, 陈昊, 孟宪红, 常燕, 胡立堂, 王文科, 束龙仓, et al. 2024. “地表-地下过程耦合的数值水文模型综述.” 中国科学：地球科学 54 (5): 1484–1505. <https://doi.org/10.1360/SSTe-2022-0420>.

2. Shu, Lele, Hao Chen, Xianhong Meng, Yan Chang, Litang Hu, Wenke Wang, Longcang Shu, et al. 2024. “A review of integrated surface-subsurface numerical hydrological models.” *Science China Earth Sciences* 67 (5): 1459–79. <https://doi.org/10.1007/s11430-022-1312-7>.
3. Shu, Lele, Paul A Ullrich, and Christopher J Duffy. 2020. “Simulator for Hydrologic Unstructured Domains (SHUD v1.0): numerical modeling of watershed hydrology with the finite volume method.” *Geoscientific Model Development* 13 (6): 2743–62. <https://doi.org/10.5194/gmd-13-2743-2020>.
4. Shu, Lele, Paul Ullrich, Xianhong Meng, Christopher Duffy, Hao Chen, and Zhaoguo Li. 2024. “rSHUD v2.0: advancing the Simulator for Hydrologic Unstructured Domains and unstructured hydrological modeling in the R environment.” *Geoscientific Model Development* 17 (2): 497–527. <https://doi.org/10.5194/gmd-17-497-2024>.
5. Shu, Lele, Xiaodong Li, Yan Chang, Xianhong Meng, Hao Chen, Yuan Qi, Hongwei Wang, Zhaoguo Li, and Shihua Lyu. 2024. “Advancing understanding of lake–watershed hydrology: a fully coupled numerical model illustrated by Qinghai Lake.” *Hydrology and Earth System Sciences* 28 (7): 1477–91. <https://doi.org/10.5194/hess-28-1477-2024>.

### B.3 完成 R 语言入门练习

练习资源：R 语言基本教程：<https://www.shud.xyz/bookR/>

### B.4 任意流域 SHUD 模型模拟

目标：在目标流域部署 SHUD 模型

要求：

1. 选择任意一个流域，建议面积小于 5000 平方公里。
2. 可利用任何我们组公开的建模工具、代码或平台。
3. 编译并运行 SHUD 模型，开展该流域 2 年以上的径流模拟。

4. 利用 rSHUD 提取流域出口的径流量，绘制径流过程线 (Hydrograph)。
5. 完成参数优化/校准
6. 日尺度径流结果高于  $NSE > 0.36$ ,  $PIBAS < 10\%$ 。
7. 提交一份 PDF 格式的报告，包含以下内容：
  - 流域位置图
  - 流域基本信息分析
  - SHUD 建模结果（输入文件）的基本信息。
  - SHUD 模型径流输出结果
  - 流域出口径流过程线图
  - SHUD 模型参数优化/校准的详细流程描述和结果分析
  - 代码和数据的 GitHub 链接

本研究组已有的日尺度（及以上）径流数据支撑的流域：

1. GRDC 网站上存在的中国流域。
2. 黑河，甘肃
3. 黄河源（唐乃亥以上）及其支流
4. 塔里木河，新疆
5. 汾河，陕西
6. 河北省 7 大河流
7. 嘉陵江，四川/重庆
8. 武江，广东韶关段
9. 博斯腾湖，新疆
10. 青海湖，青海
11. 美国 CAMELS 数据集中的流域
12. CRAVAN 数据集中的流域

参考资料：

- SHUD 模型源代码: <https://github.com/SHUD-System/shud>
- SHUD 模型中文说明书: [https://www.shud.xyz/book\\_cn](https://www.shud.xyz/book_cn)
- rSHUD 源代码: <https://github.com/SHUD-System/rshud>
- AutoSHUD 源代码: <https://github.com/shud-system/autoshud>
- GHDC 网站: <https://ghdc.ac.cn>

- GHDC 使用教程: [https://www.shud.xyz/ghdc\\_cn/](https://www.shud.xyz/ghdc_cn/)
- GHDC 案例视频: <https://www.bilibili.com/video/BV1Km4y147Wz/#reply699200607>
- R 语言基本教程: <https://www.shud.xyz/bookr/>

## 附录 C 修改日志

这个日志只记录关键规则的变化，不记录一般文字修改、图片增减、版式调整等。

- 2025-12-22

1. 增加 2.7 节，安全制度。
2. 增加 2.8 节，保密制度。
3. 删除 2.2.1 节。“学术导师的行为准则”中，“唯一可接受礼物为书籍.....”。留了这一条，就有学生向钻空子。
4. 从今天开始，增加修改日志。

- 2025.01 - 2025.12.22

1. 增加附件 A，文献阅读。
2. 增加附件 B，新手入组任务。



## 附录 D 致谢

本书一边摸索，一边完善，都是我和组内成员一起努力的成果。

感谢我的导师，给了我探索的自由。

感谢我的学生，给了我继续探索的动力。

现有成员

姓名	角色	研究方向	入组时间	背景
赵晨	硕士生	洪水预报	2022.10	兰州理工大学保送；土木工程
赖吉	硕士生 (联培)	古气候与湖泊水文	2023.10	兰州大学；导师：李国强
张海昕	博士生	陆面-水文耦合模型	2025.05	兰州大学硕博连读
李文忠	博士生	深度学习支持的小流域山洪预报	2025.05	郑州大学毕业
程亮华	硕士生	未定	2025.05	西南石油大学；测绘工程
温鑫	硕士生	流域旱涝急转	2025.05	四川大学；水利工程
王洁均	硕士生	水文预报的不确定性	2025.05	安徽师范大学；地理科学
李韵涵	硕士生	未定	2025.05	西北师范大学；地理科学

**曾经成员**

姓名	角色	研究方向	入组时间	离开时间
郭家乐	研究助理	全球流域分级算法	2021.10	2022.06
甘亚斌	硕士生 (联培)	未定	2021.01	2021.03
邱心叶	硕士生 (联培)	晚第四纪干旱区生态水文过程	2022.10	2023.01