# MODFLOW6的MPI并行化

## MODFLOW6的并行化研究总结

Sutanudjaja, et al. 2018一文主要介绍了全球尺度的水文模型PCR-GLOBWB 2的原理及应用，文中也简要介绍了MODFLOW6在水文模拟系统的单向和双向耦合，系统中可选择性地使用2层的地下水模块。实施了5 arcmin分辨率的水文-地下水耦合模拟。

J. Verkaik, et al. 2021一文主要介绍了并行化的MODFLOW的原理，其中提到其并行化的前后处理依赖于iMOD-Python库。

Jarno Verkaik, et al. 2022一文详细介绍了MPI并行化MODFLOW的前后处理方法，并将全球尺度的地下水与水文耦合模拟尺度提高到30 arcsec。

J. Verkaik, et al. 2021. Distributed memory parallel groundwater modeling for the Netherlands Hydrological Instrument. Environmental Modelling and Software 143: 105092.

<https://gitlab.com/deltares/imod/imod-python>

Edwin H. Sutanudjaja, et al. PCR-GLOBWB 2: a 5 arcmin global hydrological and water resources model. Geosci. Model Dev., 11, 2429-2453, 2018

Jarno Verkaik, et al. 2022.GLOBGM v1.0: a parallel implementation of a 30 arcsec PCRGLOBWB-MODFLOW global-scale groundwater model. GMD. https://doi.org/10.5194/gmd-2022-226

# PCR-GLOBWB 2研究(Sutanudjaja, et al. 2018)

## NHI简介

National Hydrological Instrument in 2013 (NHI; De Lange et al., 2014)

荷兰的大量钻孔资料和地表水网观测，数学模型一般有高的空间分辨率（≤250m）和时间分辨率（≤1day）。

Netherlands Hydrological Model (NHM)

Distributed memory parallel computing

non-uniform memory access architecture (NUMA). In NUMA, the entire computational grid (memory) is first partitioned into multiple subdomains and one (or more) subdomains is assigned (distributed) to a node, each having local main memory (RAM) and one or more multi-core CPUs (processors). Then, the processor cores solve the problem simultaneously while exchanging necessary data between the nodes through a fast interconnection network using the Message Passing Interface (MPI)

本文介绍NHI系统中的5个水文模型代码中的2个模型的分布式内存并行：

（1）饱和地下水MODFLOW模型

（2）不饱和区域(SVAT)的土壤-植被水转移模型(De Lange et al., 2014)。

## MPI并行化的MODFLOW6模型（J. Verkaik, et al. 2021）

NHM (De Lange et al., 2014)由5个耦合的水文模型组件组成，见图1。

NHI系统中的5个水文模型包括(De Lange et al., 2014)：

（1）SWOD模型，输入：河流流量和来自SWSC的水流通量；输出：from/to SWSC分配通量，入海流量，节点处的流量和水位。（可认为是：基于水量平衡的水文模型）。

（2）SWFT模型，输入：河流流量和来自SWOD的节点处的水量通量；输出：入海流量和节点处的流量和水位。（基于水动力模型的水流和温盐场计算）。

（3）SWSC模型，from/to SWOD模型的通量，来自子流域内的GW的水量；输出：到SWOD的子流域，到GW的水位变化（考虑到SVAT的输入）。（基于水量平衡的地表水流和温盐场计算）。

（4）SVAT模型，输入：来自大气的降雨和蒸发、植被类型和来自GW的井水位；输出：到SVAT的井水位和向SWSC的排水量。（地下水与地表水交换的模拟）。

SVAT模型包括：MetaSWAP/TRANSOL，其中，MetaSWAP求解Richards方程，使用恒定态土壤水分剖面数据。TRANSOL求解对流-扩散方程。

（5）GW模型(MODFLOW)

地下水(GW)，由7个承压模型层组成。

SVAT模型组件：土壤-植被大气的定量水转换

the unsaturated zone salt transport (UZST) model component, the surface water for sub-catchments (SWSC) model component and the surface water for optimized distributing (SWOD) model component

SVAT和UZST模型组件是1D离散（垂向），不涉及侧向MPI通信。

## MODFLOW模型的并行化综述

之前关于MODFLOW模型的并行化研究文献见表2。

Schreuder, W.A., 2005. Parallel Numerical Solution of Groundwater Flow Problems, Ph. D. dissertation. University of Colorado. （基于PETSc库）

Ji, X., Li, D., Cheng, T., Wang, X.S., Wang, Q., 2014. Parallelization of MODFLOW using a GPU library. Ground Water 52, 618-623. https://doi.org/10.1111/gwat.12104

## MODFLOW模型的并行化介绍

地下水模型的并行化组件有：

（1）将网格单元分解为子区域（blocks）；

（2）设置子区域之间的通信；

（3）并行化读写模型的文件（I/O）；

（4）线性求解器的并行化

具体细节参考附录A1；求解器的并行化参考附录A2.

子区域的区域分解的两种方法：（1）均匀分解；（2）正交递归二分（ORB）。ORB的ghost cell数目比均匀分解的多一个，但荷载均衡性更好，提高并行效率。

重叠和通信：重叠不需要用户显式指定额外数据（如单元间水力传导）。并且，（物理的）重叠便于实施先进的计算格式。进程之间发送和接收的MPI通信。

图3 四个进程的均匀分区(a)和ORB分区(b)的示例：黑色虚线表和点黑线表示2种分区的边界；红色盒子表示非重叠分区；粉色盒子表示p1进程的重叠分区；绿色单元表示两个进程间的通信界面。

并行化的输入和输出：并行化输入数据支持：栅格数据、点数据和线数据。例如，读取栅格数据，使用二进制地理参考坐标的iMOD数据格式（IDF: IMOD-Python, IMOD），因为这些文件格式支持快速的二进制格式的直接访问读取，容易使用iMOD GUI可视化。IDF文件允许有效并行化读取子区域的数据，同时保证局部内存利用。除了抽水井和地质断层数据，分别以点和线数据格式读取，所有（静态的）模块和（动态的）软件包数据（如河流和排水渠）都以IDF栅格文件格式读取。这意味着对于稀疏栅格文件可能会读取大量的冗余（no-value）数据，例如：在半干旱地区模拟排水渠系统。NHM栅格数据是稠密的，因此使用IDF的冗余较低。

并行化输出是直接的，各进程对不重叠分区，单独输出各进程的IDF文件或标准的MODFLOW ASCII/二进制文件。后处理这些子区域结果需要额外工具，如iMOD，来合并这些数据到一个单独的整体计算网格的数据集。

Vermeulen, P.T.M., Roelofsen, F.J., Minnema, B., Burgering, L.M.T., Verkaik, J., Rakotonirina, A.D., 2019. iMOD User Manual.

## 线性求解器并行化

线性求解器预处理方面，本文实施additive Schward预处理，并行化求解线性系统。但是，本文实施与表2的努力存在一些显著差别。第一，本文方法完全是分布式内存并行，包括输入和输出数据。第二，不规则模型边界的荷载均衡模型，即由于地质边界，提供一个健壮的正交递归二分方法，根据处理器数目，将地下水单元平分荷载块。第三，并行模型仅依赖MPI库，在多种平台上容易编译实施。第四，作为iMOD的部分开源软件代码来维护，iMOD是Windows系统下的GUI软件，与加速版本的MODFLOW-2005集成，被广泛使用。第五，向MODFLOW-2005增加一个新的模块化的非结构并行求解器，称之为Parallel Krylov Solver (PKS) (Verkaik et al., 2016, 2015)，主要基于UPCG线性求解器。

## 结论

图6显示了观测的加速和总的内存使用。使用144个处理器核心。串行运行耗时是4h 48min，并行计算耗时降低至2min 40s，加速比达到108，各单元的水头误差低于指定的*e*=0.001

并行化线性求解器迭代随着子区域数目的增加而增加，导致低频率的特征模式会妨碍线性求解器的收敛，这需要额外的多层预处理来改善收敛。但是，本测试中，观测的最大线性迭代增加为~15%，低频的特征模式影响效果相对较低，因此无需实施预处理。图8显示：线性求解迭代（串行迭代3269步），RNHM为18738步，FNHM为30390步。

大量时间耗在线性求解器，见图9，线性扩展性，数据输入耗时。

荷载均衡很重要，见图10，MPI全局通信在p36和p42占主要耗时，这两个进程不与SVAT连接，导致荷载不均衡。

## 附录A1 :区域分解

如图3，不规则区域由nr行，nl列组成，分解为P=4个分区。

均匀分区和ORB，荷载不均衡度分别为1.48和1.0

## 附录A2 :并行化线性求解器

地下水流方程有限体积离散后，经过Picard线性化和施加Dirichlet边界（恒定条件），导出线性方程组系统：

**Ah=b**

**h**为未知向量，**b**为RHS，**A**为方形、正定的系数矩阵。**A**为7点带状矩阵。为求解方程组，使用Krylov子空间加速，在PCG法中实施该预处理方法。不是直接求解**Ah=b**，而是对称化预处理系统：



**M**称为preconditioner。

矩阵**A**可写为块矩阵的形式：



用**A**的块对角矩阵作为**M**则导出非重叠的additive Schwarz预处理方法，使用**MAS**表示：



PCG法的additive Schwarz预处理方法的伪代码为：



上述方法的并行化涉及：（1）在执行稀疏矩阵向量乘积之前的子分区间向量的局部MPI点对点通信；（2）确定内积全局求和和停止迭代准则的全局最大值的全局集合MPI通信。

## 参考文献

NHI（荷兰水文模拟系统）：

De Lange, W.J. Prinsen, G.F. Hoogewoud, et al. 2014. An operational, multi-scale, multi-model system for consensus-based, integrated water management and policy analysis: The Netherlands Hydrological Instrument. Environ. Model. Software 59, 98-108.

IMOD

Imod-Python Development Team, 2017. iMOD-Python: make massive MODFLOW models. https://gitlab.com/deltares/imod/imod-python.

Vermeulen, P.T.M., Roelofsen, F.J., Minnema, B., Burgering, L.M.T., Verkaik, J., Rakotonirina, A.D., 2019. iMOD User Manual

# GLOBGM v1.0 (Jarno Verkaik, et al. 2022)

本文执行了30 arcsec分辨率的全球尺度的水文-地下水（双层）耦合模拟，是5 arcmin的后续。

使用非结构化网格，共278 million激活计算单元（排除了冗余的海洋和陆地单元）。

4个地下水模型都分解为多个不重叠的子模型，在MODFLOW线性求解器内紧密耦合，其中各子模型唯一地分配到一个处理器核心，相关的子模型数据使用data tiles形式前处理写作并行化数据文件。

为平衡计算荷载，使用METIS库，以2种方式做区域分解：（1）直接应用于所有模型网格单元；（2）基于面积的用于HydroBASINS流域，分配到子模型，做pre-sorting，用于将来与地表水模型耦合。

模拟1958~2015，daily时间步长，monthly输入，包括20-year spinup。

串行模拟需要~4.5个月的运行耗时；使用12个计算节点（每个节点32核心，共384个核心），可实现138x加速比，将运行时间降低到16小时。

模型输出使用NWIS水头观测（contiguous United States实施的观测）。

将来得继续细化水文地质参数化方案，提供位置、深度和抽水井区域的更多信息。

## 1引言

使用transient模拟

应用de Graaf et al. (20176)的简单方法获得GLOBGM的初始条件。

使用恒定态结果（自然条件下，没有抽水）spin-up模型