

基于 GRU 门控循环网络的大气碳浓度预测

摘要

在“双碳”背景下，研究大气中二氧化碳浓度和变化趋势，有助于了解区域内二氧化碳排放情况以及碳汇潜力，并在此基础上制定合理有效的能源政策。本文利用 GRU(Gate Recurrent Unit)作为大气碳浓度的预测模型，以中国中部一区域为研究区域，选取 2007-2019 年该区域的碳排放量、大气碳浓度、陆地温度、降水量以及附近区域的海水表面温度作为训练数据，用 18 个月的数据去预测后 1 个月的数据，并与碳浓度实际值进行比对验证。研究结果表明：GRU 模型对大气碳浓度数据的变化趋势有较好模拟能力；预测模型的训练集和验证集的 MSE (mean squared error) 范围在 0.021~0.064。

关键词：大气碳浓度；神经网络；GRU；预测模型

目录

1 引言	1
2 模型与方法.....	1
2.1 模型假设	1
2.2 GRU 模型	3
2.3 模型评估方法	4
3 实例验证	4
3.1 数据来源与预处理	4
3.1.1 研究区域.....	4
3.1.2 数据来源	5
3.1.3 数据预处理.....	6
3.2 参数设置	7
3.3 实验结果	7
3.3.1 模型评估分析.....	7
3.3.2 模型预测分析.....	8
3.3.3 经验与推广	8
4 结论与政策建议	9
参考文献.....	10
致谢.....	12

表格与插图清单

表 1 变量名称及选取区域

表 2 模型结果评估表

图 1 本文试验的 GRU 模型

图 2 其中一个栅格 5 个变量的现实时间序列数据曲线

图 3 GRU 预测模型各层参数表

图 4 2018 年 9 月到 2019 年 12 月 36 个格子的预测值和真实值的区别

基于 GRU 门控循环网络的大气碳浓度预测

1 引言

《2023 年全球气候状况》报告显示，过去一年，全球温室气体水平、地表温度、海平面上升等主要指标刷新记录的同时，也带来了更多的极端天气事件。专家呼吁重视全球变暖对经济社会和人类生活的风险^[1]。

全球变暖的原因与工业化以来温室气体排放增加脱不开干系，这一点已成为共识。据 IPCC（政府间气候变化专门委员会）评估，人为排放的二氧化碳，86%来自化石燃料^[2]。

如何在推动经济增长的同时，有效应对全球气候变化，是我们当下必须思考的问题。研究大气中二氧化碳浓度和变化趋势，对了解区域内二氧化碳排放情况以及碳汇潜力至关重要，可以为能源政策的制定提供科学依据，继而以绿色低碳发展的新成效激发新质生产力，积极推进“碳达峰”“碳中和”承诺^[3]。

早期大多用时间序列去预估二氧化碳浓度趋势，常用的模型有 ARMA（自回归滑动平均模型）和 ARIMA（自回归差分移动平均模型）^[4]。随着人工智能技术的发展，深度学习作为其中一个重要分支，可以模拟人脑的神经网络系统，自动从大量的数据中提取特征，越来越多人将其应用在二氧化碳浓度预测上，并取得了一定的成就^[5]。例如，邵涛^[6]利用 BP（back propagation）神经网络预测了教室室内空气污染物浓度；谢秋菊等^[7]基于 ISSA-GRU-ARIMA 模型对猪舍的有害气体浓度作了精准预测；付子骏等^[4]通过将二氧化碳浓度变量、污染物排放量和气温等作为输入变量，基于长短期记忆神经网络（long short-term memory neural network, LSTM）构建了一个模型，在二氧化碳浓度预测中取得了良好效果。

大气碳浓度的影响因素多元且难以确定，本文创新性假设了多个合理输入变量，并在此基础上通过比 LSTM 收敛更快的门控循环（gate recurrent unit, GRU）神经网络对区域内二氧化碳浓度进行预测，GRU 模型对大气碳浓度数据的变化趋势有较好模拟能力。

2 模型与方法

2.1 模型假设

碳，无处不在。人类在自然中生产、生活会产生二氧化碳排放到大气中，这

些二氧化碳主要来自于化石燃料的燃烧。据 IPCC 统计，人为排放的二氧化碳约只有一半会残留在大气中，剩下一半会被海洋、陆地这些自然储层所吸收^[2]，这些碳储层相互连接、作用、释放，完成了碳循环的闭环，一定程度上缓解了气候变化及其影响，否则，全球变暖效应可能会是目前观测的 2 倍^[8]。

基于此，我们可以认为大气碳浓度（**atmospheric CO_2** ，记作 C_{atm} ）是关于化石燃烧二氧化碳排放量（ **CO_2 emissions from fossil fuel combustion**，记作 $FFCO_2$ ）与碳储层吸收量即碳汇（**carbonsink**，记作 S ）的函数，得出

$$C_{atm} = f_1(FFCO_2, S) \quad (1)$$

陆地和海洋的碳汇影响因素众多，就海洋碳汇来说，大气碳浓度增加会把更多的二氧化碳送达海洋表面，再通过洋流传输到海洋深处储存。但随着海洋表面温度上升，洋流传输交换的频率会随之减缓，无法将更多的二氧化碳传输到海洋深处。另一边，强劲的风力也可能会提升洋流交换速度，把深水中的碳释放出来，但具体影响目前尚不明朗^[8]。

陆地生态碳循环系统在整个碳循环中占主导地位^[9]，其中，土壤的碳储存量约占陆地生态循环系统的 70%^{[10][11]}，是植物碳汇的 3 倍左右^{[12][13]}。大气碳浓度升高会影响土壤碳的固定^[14]，温度和降雨又会影响森林的碳储存量^[15]，但具体的消长关系目前还并不清楚^[9]。

但可以明确的是，无论碳循环机制具体如何作用，减少人为二氧化碳排放对于抑制全球气候变暖都是必不可少的。

综合以上，我们可以继续认为碳汇 S 又是关于化石燃烧二氧化碳排放量 $FFCO_2$ 、大气碳浓度 C_{atm} 、海洋温度（**ocean temperature**，记作 T_o ）、陆地温度（**land temperature**，记作 T_L ）和降雨量（**precipitation**，记作 P ）的函数，即

$$S = f_2(C_{atm}, T_o, FFCO_2, T_L, P) \quad (2)$$

将 (2) 带入 (1)，我们可以得到

$$C_{atm} = f_3(FFCO_2, C_{atm}, T_o, T_L, P) \quad (3)$$

即大气碳浓度会收到化石燃烧二氧化碳排放量、大气碳浓度、海洋温度、陆地温度和降雨量的综合影响。目前就其中某一两个元素对海洋或森林的碳汇估算较多，但综合以上 5 个元素对大气碳浓度变化趋势的预测鲜见报道，这也本文的

创新之处及研究价值所在。

2.2 GRU 模型

当下用于预测研究的神经网络模型主要有卷积神经网络（convolutional neural networks, CNN）、循环神经网络(recurrent neural network, RNN)和 Transformer 等。其中，CNN 在图像、音频处理等方面更有优势，RNN 的记忆性使它更适合处理时间序列预测，Transformer 更适用于大数据量的模型预测^[16]。

比较而言，RNN 模型更适合本次研究，但是 RNN 的局限性在于难以捕捉长期依赖关系，因此本文选择基于 RNN 和 LSTM 改良的、更流行的 GRU（门控循环神经网络）模型^[17]来完成本次预测。

GRU 模型的门控分别指更新门（一般用 z_t 表示）和重置门（一般用 r_t 表示），更新门用于控制前一时刻的状态信息 h_{t-1} 被带入到当前状态中的程度，值越大说明带入的信息越多；重置门用于控制前一状态有多少信息被写入到当前的候选集 \tilde{h}_t 上，值越小，前一状态的信息被写入的越少。其公式如下：

$$z_t = \sigma(W_{zh} * h_{t-1} + W_{zx} * x_t) \quad (4)$$

$$r_t = \sigma(W_{rh} * h_{t-1} + W_{rx} * x_t) \quad (5)$$

$$\tilde{h}_t = \tanh(W_{hh} * r_t * h_{t-1} + W_{hx} * x_t) \quad (6)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (7)$$

其中， x_t 为 t 时刻的输入， h_t 为 t 时刻的输出， $*$ 表示乘积， σ 和 \tanh 都是激活函数， W_{zh} 、 W_{zx} 、 W_{rh} 、 W_{rx} 、 W_{hh} 、 W_{hx} 是训练过程中需要学习的参数。

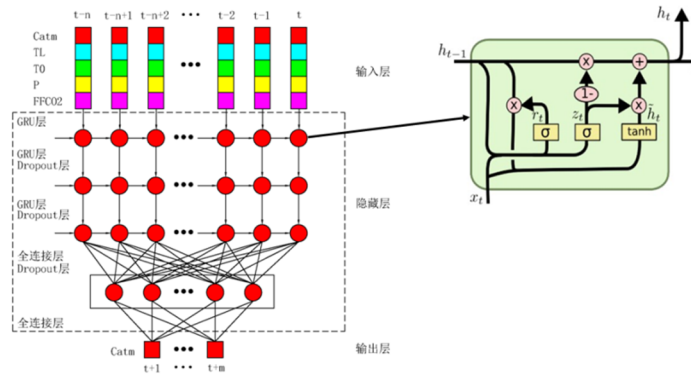


图1 本文试验的 GRU 模型

本文基于 Keras 建立 GRU 模型，具体模型建立如下：

- (1) 将预处理后的样本数据 $Catm$ 、 $FFCO_2$ 、 T_0 、 T_L 和 P 作为输入层输入，

通过 3 层 GRU 层叠加，并在每一层都使用正则化技巧（参数为 0.01），从而获取时序关系，在第二层和第三层后都连接一个 dropout 层（参数为 0.2），以避免过拟合现象。

（2）在 GRU 层后连接两个全连接层，第一层全连接层采用 16 个神经元，并在其后添加一个 dropout 层（参数为 0.2）。第二层全连接层采用 1 个神经元，以得到最终的输出结果。

（3）在建立好模型后，使用预处理后的数据集样本进行训练，采用分批训练，批次大小（batch size）分别取 32, 64, 128。

多层 GRU 的设置可以帮助文辞研究获取大气碳浓度数据与 5 个样本特征数据时间上的非线性关系。正则化技巧和 Dropout 层则可以避免过拟合现象，增加模型处理数据的能力^[18]。

2.3 模型评估方法

在 GRU 模型的样本训练中，我们使用均方误差和平均绝对误差对结果进行评估。计算公式分别如下：

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_{real} - y_{pred})^2 \quad (8)$$

式中： MSE 为均方误差（mean squared error），是大气碳浓度数据标准化处理后训练时真实值和预测值的差值的平方和的平均值， y_{real} 为大气碳浓度数据标准化处理后的真实值， y_{pred} 为大气碳浓度数据标准化处理后的预测值。

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{real} - y_{pred}| \quad (9)$$

式中： MAE 为平均绝对误差（mean absolute error），是大气碳浓度标准化处理后训练时的真实值和预测值之间绝对误差的平均值。

通过 GRU 模型对大气碳浓度进行预测时， MSE 和 MAE 的值越小，表明训练和验证的效果越好。本文训练模型时会通过这两个指标来评估模型训练的效果，以期达到最好的预测效果。

3 实例验证

3.1 数据来源与预处理

3.1.1 研究区域

本文选取中国中部区域（101° E-113.5° E，33.5° N-46° N）为研究区域。因为每个变量数据的空间分辨率不同，为确保每个变量的数据个数都是 6*6=36

个，提取区域稍有差异。另，该区域为内陆地区，缺乏海水表面温度数据， T_0 变量选取的区域往东南方向偏移了一些，特作说明。

表 1 变量名称及选取区域

变量名	选取区域	空间分辨率	单位
$Catm$	101.25° E-113.75° E, 34° N-44° N	$2^{\circ} \times 2.5^{\circ}$	10^{-6}
$FFCO_2$	107.5° E-112.5° E, 35.5° N-40.5° N	$1^{\circ} \times 1^{\circ}$	$gC/m^2/d$
T_0	122° E-127° E, 19.5° N-24.5° N	$1^{\circ} \times 1^{\circ}$	$^{\circ}C$
T_L	101.25° E-113.75° E, 33.75° N-46.25° N	$2.5^{\circ} \times 2.5^{\circ}$	$^{\circ}C$
P	101.25° E-113.75° E, 33.75° N-46.25° N	$2.5^{\circ} \times 2.5^{\circ}$	mm

3.1.2 数据来源

大气碳浓度 ($Catm$) 数据来源于国家科技基础条件平台—国家地球系统科学数据中心—地理资源分中心 (<http://gre.geodata.cn>)，其 1992-2020 年大气碳浓度月均值数据集是根据正弦估算模型逐像元参数率定与模拟得到的，并通过了站点观测数据的验证^[19]；

化石燃烧二氧化碳排放量 ($FFCO_2$) 数据来自全球气候日本国立环境研究所 (NIES) 开发的全球化石燃料二氧化碳排放数据集。数据是结合夜间光照数据和单个发电厂的排放位置概况，通过创新的排放模型估算得出^[20]，当前版本为 ODIAC2022，是时间跨度从 2000 年 1 月至 2021 年 12 月的逐月数据。下载地址为 ODIAC Fossil fuel emission dataset | Center for Global Environmental Research (nies.go.jp)。

海水表面温度 (T_0) 数据来源于中国 Argo 实时资料中心的全球海洋 Argo 网格数据集，其 2007-2022 年逐月数据为用 Barnes 逐步订正法构建的^[4]，具体下载地址为 <http://www.argo.org.cn>。

陆地温度 (T_L) 和降雨量 (P) 数据均来自于欧洲联盟气候监测机构哥白尼气候变化服务局 (C3S)，包含 2000-2019 年来自于世界各地 80000 个气象观测点的逐月数据，具体可以通过 cds.climate.copernicus.eu 下载。

3.1.3 数据预处理

首先，对数据做一些基本处理。具体包括：通过 matlab 将下载的 nc 文件数据信息转成 xlsx 格式方便后续处理；按变量根据前文标注的经纬度范围分别提取 36 个横截面数据信息；

然后，确定 5 个变量时间跨度的交集 2007 年 1 月至 2019 年 12 月共计 13 年数据为研究对象，得到 5（变量数） \times 156（逐月数据的时间序列数） \times 36（横截面数据）=28080 个数据；

接着，为了消除量纲的差异，利用 python 的 sklearn 中的 standardScaler 函数对所有数据执行标准化处理，将各变量中的数据分别转换为均值为 0、方差为 1 的标准正态分布。

标准化处理后的 $Catm$ 数据用作生成器的标签数据，同时，将标准化处理后的 $Catm$ 、 $FFCO_2$ 、 T_0 、 T_L 和 P 数据作为样本特征集数据。将 28080 个数据按 36 个横截面栅格、分变量各自建立长度为 156 的时间序列。

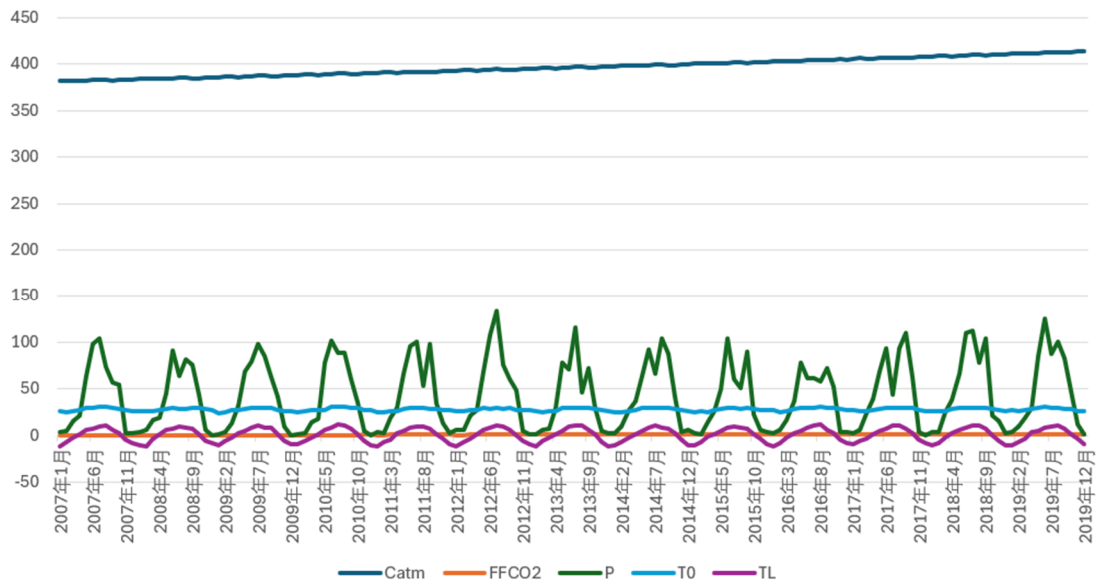


图 2 其中一个栅格 5 个变量的现实时间序列数据曲线

通过滑动窗口构建数据集，使用前 18 个月的 $Catm$ 、 $FFCO_2$ 、 T_0 、 T_L 和 P 数据对后 1 个月的 $Catm$ 进行预测，滑动步长为 2。按照此参数，每个栅格可得到 69 组数据，36 个栅格即可得到 2484 组数据。这些数据按照 72%、18% 和 10% 的占比划分为训练集、验证集和测试集。训练集用于模型训练，验证集作为校验数据用于验证训练出的模型，测试集用于和预测值作比对。

3.2 参数设置

GRU 预测模型各层参数见图 3, 本文选择了 ReLU 函数作为激活函数, 指定 MSE 作为损失函数, 使用 Adam 优化器来优化参数, Dropout 参数设置为 0.2, L2 正则化参数为 0.01, 初始学习率为 0.001。

Layer (type)	Output Shape	Param #
gru_145 (GRU)	(None, 18, 8)	360
gru_146 (GRU)	(None, 18, 16)	1248
dropout_144 (Dropout)	(None, 18, 16)	0
gru_147 (GRU)	(None, 32)	4800
dropout_145 (Dropout)	(None, 32)	0
dense_95 (Dense)	(None, 16)	528
dropout_146 (Dropout)	(None, 16)	0
dense_96 (Dense)	(None, 1)	17
Total params: 6953 (27.16 KB)		
Trainable params: 6953 (27.16 KB)		
Non-trainable params: 0 (0.00 Byte)		

图 3 GRU 预测模型各层参数表

3.3 实验结果

3.3.1 模型评估分析

本文对 GRU 预测模型分别作了 400 轮训练, 模型训练结果见表 3:

表 2 模型结果评估表

批次	参数 batch_size	参数 epochs	训练集 MSE	验证集 MSE	测试集 MSE	测试集 MAE
1	32	400	0.02586	0.01066	0.06375	0.21454
2	64	400	0.02661	0.00801	0.02785	0.13804
3	128	400	0.02185	0.03238	0.03238	0.14828

通过对模型训练结果分析发现: 训练集的均方误差最佳值为 0.02185, 验证集的均方误差最佳值为 0.00801, 总体上可以看出, GRU 模型的训练结果较好; 模型训练完后, 再对其训练结果模型进行评估, 发现其测试集的均方误差最佳值为 0.02785, 平均绝对误差最佳值为 0.13804, 预测效果一般。

3.3.2 模型预测分析

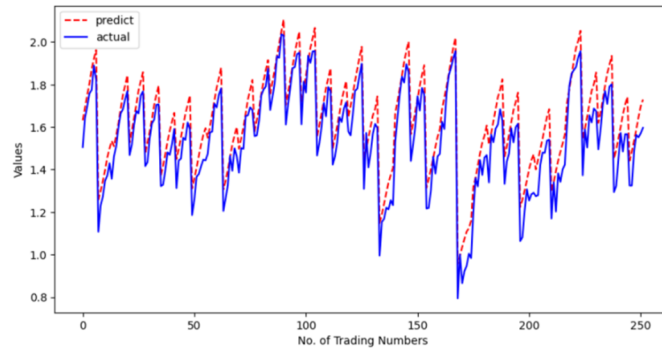


图 4 2018 年 9 月到 2019 年 12 月 36 个格子的预测值和真实值的区别

为了更好地呈现 GRU 模型在大气碳浓度趋势的预测效果，本文将 2018 年 9 月到 2019 年 36 个栅格的真实值和预测值进行了比较，结果如图 4 显示，实测值曲线和预测值拟合曲线的趋势一致，但实际数值仍有一定差距。

3.3.3 经验与推广

在本研究中，我们通过 GRU 神经网络模型对中国中部区域的未来大气碳浓度进行了预测，验证了 GRU 模型在时间序列数据预测中的有效性和准确性。以下是本文在实验过程中总结的经验，并将对其在更大范围内的推广应用进行探讨。

首先，数据预处理的重要性。经过预处理方法的尝试，我们发现数据的标准化处理对于模型的性能至关重要。通过将范围分布较大的数据转化为均值为 0、方差为 1 的标准正态分布，能够显著提高模型的训练效率和预测精度。这就为我们提供了一种思路，如在其他区域进行类似研究时，同样进行标准化处理，以保证数据在不同量纲下的可比性以及模型自身的鲁棒性。

从中展开，以 Min-Max 归一化以及标准化两种归一化方式为例。对于神经网络和深度学习，Min-Max 归一化通常表现会更好，因为输入数据在统一的范围内可以加快模型收敛；而对于一些线性模型，如线性回归、逻辑回归等，标准化则可能更合适，因为这些模型对数据服从正态分布假设更敏感。

其次，变量选择与模型假设。本案例中创新性地将化石燃料二氧化碳排放量、海洋温度、陆地温度和降雨量作为输入变量，与大气碳浓度的变化建立联系，对于大气科学这个领域，这种多因素综合考虑的方法，是科学且合理的。

再次，GRU 模型的优越性。相较于传统的时间序列预测方法，GRU 模型在处理非线性关系和长期依赖性方面表现更为优异。同时经过分析可以发现，在其他

区域或不同时间尺度上进行大气碳浓度预测，尤其是在处理复杂气候数据时，GRU模型同样具有较高的适应性和推广价值。

最后，模型的训练与评估。参数可以是激活函数、损失函数、优化器等等。在实际的应用中，可针对不同的参数组合进行多次实验，以确定最终最佳参数的设置。在案例中，以损失函数为例，通过对各种参数值，如 MSE、MAE 等进行效果上的比较，确定了最终的配置。除此之外，采用交叉验证等方法对模型进行评估，也能确保其稳定性和可靠性。

4 结论与政策建议

基于本研究的结论和中国当前面临的碳排放挑战，提出如下政策建议：

1. 加强数据收集与管理。建立全国性的大气碳浓度检测网络，定期收集和更新相关数据，确保数据的全面性和准确性，势在必行。与此同时，推动数据共享机制，促进科研机构之间的数据交流与合作，提高数据的利用效率。

2. 推动绿色低碳技术发展。加大对绿色低碳技术的研发投入，鼓励企业和科研机构创新，并推广清洁能源和节能减排技术。同样地，支持碳捕集与封存技术的发展，减少化石燃料燃烧带来的二氧化碳排放。

3. 强化政策引导与落实。制定更加严格的碳排放标准和法规，督促企业减少碳排放，提升环保意识。除此之外，深入推行碳排放交易制度，通过市场机制引导企业自觉减少碳排放，实现经济效益与环境保护的双赢。

4. 提升公众环保意识。加强环保教育，提升公众对碳排放和气候变化问题的认识，鼓励绿色消费和低碳生活方式。可以考虑的方案有，开展宣传活动，提高社会各界对“双碳”目标的重视与参与度，共同推动绿色低碳发展。

通过以上政策的推行，中国可以更有效地应对碳排放问题，推动经济与环境的协调发展，为实现“碳达峰”、“碳中和”目标奠定坚实基础。

参考文献

- [1] 韩雪萌. 全球气候主要指标亮红灯[N]. 金融时报, 2024-04-25(008)
- [2] 周天军, 陈晓龙, 张文霞, 张丽霞. 气候变化与碳中和[J]. 自然杂志, 2024, 46(01): 1-11.
- [3] 习近平在中共中央政治局第十一次集体学习时强调 加快发展新质生产力 扎实推进高质量发展[N]. 新华社, 2024-02-01.
- [4] 付子骏, 吴永明, 徐计. TD-LSTM-S 模型在二氧化碳浓度预测中的应用[J]. 重庆理工大学学报(自然科学), 2023, 37(04): 192-199.
- [5] 蒋玮. 深度学习技术对碳循环经济的影响研究[J]. 科技创新与应用, 2024, 14(13): 26-29+34.
- [6] 邵涛. 基于神经网络的教室室内 PM2.5 浓度预测研究. 昆明理工大学. 2023(03)
- [7] 谢秋菊, 马超凡, 王圣超, 包军, 刘洪贵, 于海明. 猪舍氨气与二氧化碳浓度变化时序预测模型优化[J]. 农业机械学报, 2023, 54(07): 381-391.
- [8] Dame Julia Slingo DBE FRS, etc. Next generation climate models: building strong foundations for climate action. The Royal Society. 2021. 06
- [9] 马红亮, 朱建国, 谢祖彬. 大气 CO₂ 浓度升高对陆地生态系统土壤固碳的可能影响[J]. 土壤通报, 2008, (05): 1184-1191.
- [10] PARTON W J, SCHIMEL D S, et al. Aanalysis of factors controlling soil organic matter levels in great plains grasslands [J]. Soil Science Society of American Journal, 1986, 51: 1173- 1179.
- [11] PERRUCHOUD D, FISCHLIN A. The response of the carbon cycle in undisturbed forest ecosystems to climate change: a review of plant-soil models[J]. Journal of Biogeography, 1995, 22: 759- 774.
- [12] POST W M, PENG T H, EMANUEL W R, et al. The global carbon cycle[J]. American Scientist, 1990, 78: 310- 326.
- [13] HOUGHTON J T, JENKINS G J, EPHRAUMS J J. Climate change the IPCC scientific assesment [R]. New York: Cambreidge University Press, 1990
- [14] 刘树伟, 纪程, 邹建文. 陆地生态系统碳氮过程对大气 CO₂ 浓度升高的响应与反馈[J]. 南京农业大学学报, 2019, 42(05): 781-786.
- [15] 令狐大智, 罗溪, 朱帮助. 森林碳汇测算及固碳影响因素研究进展[J]. 广西大学学报(哲学社会科学版), 2022, 44(03): 142-155.
- [16] 钟耿钦, 王学梅, 康龙云. 基于 GRU-BP 的锂电池部分充电数据 SOH 估计[J]. 电源学报: 1-15.
- [17] 何清, 李丽琳, 林子安. 基于 FEEMD-GRU-FC 模型的滑坡位移预测[J]. 人民长江: 1-8.
- [18] 张雪薇, 韩震. Argo 温度数据的 ConvGRU 模型预测分析[J]. 海洋环境科学, 2022, 41(04): 628-635.
- [19] 侯炜烨, 金佳鑫, 严涛, 刘颖. 全球大气二氧化碳浓度 2° × 2.5° 栅格模拟数据集 (1992-2020) [J]. 全球变化数据学报(中英文), 2022, 6(02): 191-199+363-371.
- [20] Tomohiro Oda, Shamil Maksyutov (2015), ODIAC Fossil Fuel CO₂ Emissions Dataset (Version name1 : ODIACYYYY or ODIACYYYYa), Center

for Global Environmental Research, National Institute for
Environmental Studies, doi:10.17595/20170411.001.

致谢

Deadline 迫在眉睫，三个职场人白天工作晚上加班加点建模型、跑程序、写论文，能一路走到这里，首先最想感谢的是我们自己。

非全日制研究生，本来就工作学习“两头烧”；跨行跨专业的经验不足，也没有成为我们退却的理由。经历一个项目一定会多学到一点东西，是我们仨在统计建模大赛报名前的一致初心。多少个上完课的周六晚上，我们在研习室研究选题、讨论模型；五一假期难得有空闲休息，图书馆里是我们查阅资料的身影；工作上脱不开身投入比赛时，队友间互相托底，接过对方手上的代码直接干活。因为时间有限、经验不足，我们踩了很多坑，本次的预测模型也有很多需要继续完善的地方，但每一次失败都为我们下一次的努力指引了方向。统计建模比赛会成为我们应用统计学习中很重要的一站，这次比赛的经验也会带着我们在统计应用中走向更开阔的远方。

当然，我们更要感谢本次建模大赛的指导老师，因为大家都有工作的原因，我们只能在休息时间打扰老师。老师牺牲午休时间指导选题、晚上开远程会指导我们建模方向，出差前也不忘指出我们模型设计中的问题，截稿前更是事无巨细地提醒我们提交作品的细节。

其次，我们还要感谢本次建模大赛为小组提供帮助的多位同学和朋友！正是你们的无私帮助，我们才能顺利完成本次建模及论文写作。

最后，我们要向百忙之中抽出时间对本文进行审阅、评议以及答辩的各位老师表示诚挚的感谢！